Horizon 2020 Programme
SwafS-14-2018-2019-2020
Supporting the development of territorial Responsible Research and Innovation

Project[1] Number: 872550
Project Acronym: TetRRIS
Project title: Territorial Responsible Research and Innovation and Smart Specialization

D8.2 Data Management Plan

Due date: 31.1.2021
Dissemination Level: Public
Author(s): Ilona Koski and Juha Oksanen

[1] The term 'project' used in this template equates to an 'action' in certain other Horizon 2020 documentation

Peer-reviewed by: Emad Yaghmaei and other consortium partners

TetRRIS Data Management Plan

Document History

Name	Date	Version	Description
Ilona Koski and Juha Oksanen (VTT)	19.1.2021	1.0	First draft up for peer review
Emad Yaghmaei	25.1.2021	2.0	Comments
Ilona Koski and Juha Oksanen (VTT)	29.1.2021	3.0	Finalizing
Mika Nieminen (coordinator)	1.2.2021	3.0.	Final review

Disclaimer: This document is a living document, thus it will be updated in the course of the Project when required. Some issues which will be addressed in this data management plan need further discussion with TetRRIS Steering Group (SG) and approval by them. A complete list of these issues will be provided in the table 2. Any changes to the Data Management Plan will be communicated and introduced in the next iteration of the document.

LEGAL DISCLAIMER

1. Data Summary

The TetRRIS consortium will collect and produce different types of data in the course of implementation of the project. Data from external sources (e.g. material sourced through desk research) is collected for instance to map innovation ecosystems in which the regional pilot cases will be embedded in later stage, to prepare ground for the launch of the pilots and in later stage for purposes of mutual learning and dissemination.

Most of the data in the project is internally generated and used only for research purposes or project's management and administrative functions. Internal data will be generated as a by-product of the project execution: this type of data includes for example stakeholder communications, email correspondence, minutes of meetings as well as other communications among consortium members. Internally generated data may also include indicators derived from external data and results from surveys or outputs and analyses of the interviews and workshops.

During interviews, surveys, workshops or interactive co-creation events personal data will be collected. This may include name, age, gender, occupation, professional background, education and personal opinions. Furthermore, sessions will be recorded for research purposes and screen captures or videos might be taken. Personal data will not be made publicly available if it is not justified by research purposes (for example necessary transparency or validation of the results). In addition, publication permission will always be asked. Personal data, collected as part of the project, will only be shared publicly if agreed with individuals included in the dataset. Otherwise it will be either anonymized or pseudonymized. TetRRIS consortium members are required to use informed consent form to inform about these issues and confirm each human participant's consent regarding them.

The consortium uses protected web-based Microsoft Teams platform hosted by VTT for internal communication and collaboration as well as data exchange and storage for the project documents. All research data collected and generated is recommended to be shared with all consortium members as widely as possible. However, data containing personal information should be anonymized our pseudonymized when shared with the whole consortium. Additionally, all consortium partners have their own web-based platform for each's internal communication, secondary data storage and project management activities. Regarding data storage, these platforms are required to be protected and accessible only for research team.

All data generated in TetRRIS will be of low or moderate risk. High risk data refers to either strictly confidential information, such as patient records, passwords, and credit card numbers, or high-risk research such as biomedical research, high risk chemical experiments. Any high-risk data won't be collected or generated in TetRRIS.

What is the purpose of the data collection/generation and its relation to the objectives of the project? What types and formats of data will the project generate/collect?

In the table below can be found types of the data collected/generated and the objectives of the data collection/generation for each Work Package. Table is based on preliminary plans of each WP according to the TetRRIS Grant Agreement.

However, changes might occur during the execution and implementation of the project and each pilot partner will apply project plan to their own specific needs. Therefore, also more specific table will be provided in the course of the study. This table can be found in Annexes of this document (Table 3). Each WP leader is asked to routinely update their part of the table during the project according to their executed actions. This table will reflect realized inputs and

outputs regarding project's metadata and it will be used for e.g. management and administration purposes.

Table 1. TetRRIS metadata *not all data types can be foreseen at this stage of the project thus this table only presents preliminary plans according to the TetRRIS Grant Agreement

WP Number and name	Type(s) of Data Collection/Generation	Purpose of Data Collection/Generation and it's relation to project's objectives	Formats of Data*
WP1 Ethics Requirements	The 'ethics requirements' that the project must comply with are based on agreed procedures, TetRRIS Grant Agreement and communication between consortium partners (e.g. verbal conversations and written correspondence on ethics deliverables).	Generating ethics requirements and guidelines for administrative use.	Written documents
WP2 Mapping and Analysis of Actors	General framework and analytical tool is built on external sources of data (e.g. existing concepts from the literature and findings from previous research and innovation projects) Interviews and desk research in order to execute the mapping and analysis of pilot regions innovation ecosystems	Mapping and analyzing each pilot's territorial research and innovation system (TRIS) focusing on the degree of integration of RRI into the TRIS. The insights generated by WP 2 directly feed into WP 3. Describing the main insights gained and identifying possible pilot actions. Each report documents the relevant stakeholders, their networks, hard and soft framework conditions, and RRI-relevant activities in the regions.	Written documents (e.g. project deliverables, notes for internal use and meeting minutes) Recorded and transcribed interviews Other written research material (e.g. literature reviews)
WP3 Scoping and Co-Design of actions	Workshops to bring together consortium members and the TRIS stakeholders; e.g. universities, civil society organizations, business, local authorities and intermediaries.	Validating and refining the findings from WP 2 through dialogue with the stakeholders and justify actions and interventions to further integrate RRI into the TRIS and regional development policy. Creating a work plan (e.g. actions and interventions are carried out in WP 4) with clearly stated responsibilities and an at least general timeline with commitments to a defined range of activities is each pilot territory. Workshops between scientific partners, regional	Recorded workshops and co-creation events, possibly recorded and transcribed interviews as well. Written documents (e.g. project deliverables, notes for internal use and meeting minutes) based on existing literature and analysis of the

		partners and regional stakeholders of the different pilots will introduce them to each other's and begin the process of structured exchange that continues in WP 5.	executed events or interviews.
WP4 Pilots	A two-day kick–off workshop for stakeholders in charge of the ongoing or planned measure into which the RRI-related actions are to be embedded. Workshops in the middle and at the end of the project and other events if considered relevant A short survey to collect impressions on the work progress, its perceived value and remaining shortcomings. Dissemination activities to gain insight (for that purpose, different types of communication materials will be developed) and reach broad groups of stakeholders identified in WP2.	Scientific partners will help regional partners further improve the pilot actions. Regular discussions between consortium members and the advisory board for analyzing gained insights Dissemination and efficient communication of project targets, actions, interventions and outcomes. Summarizing the lessons drawn from the pilots and already by that time realized outcomes.	Recorded workshops, possibly recorded and transcribed interviews as well. Quantitative and/or qualitative survey data Written documents (e.g. project deliverables, notes for internal use and meeting minutes, promotional materials) based on existing literature and analysis of the executed events or interviews.
WP 5 Experimenting, Learning & Sharing	Guidelines for the design, development and evaluation of the TetRRIS Policy Lab Four Policy Lab sessions to bring together the pilots and other regions. Short survey for evaluating each Policy Lab session	Mutual learning, networking, and the creation of enriching synergies between the regions. Each session will be evaluated in order to improve contents, designs and dynamics between sessions,	Written documents (e.g. project deliverables, notes for internal use and meeting minutes, promotional materials) based on existing literature and analysis of the executed events or interviews. Recordings of the policy lab sessions Quantitative and/or qualitative survey data
WP 6 Analyses, Reflections and	Policy briefs for practitioners in each pilot territory, outlining tools, good practices and policy	Turn findings and the lessons from previous findings into actionable tools, good practices and	Written documents (e.g. project deliverables,

Recommendations	recommendations tailored to their specific circumstances General handbook for practitioners elsewhere in Europe, summarizing the experiences, tools, good practices and illustrative examples generated by TetRRIS. Material generated in the previous WP's will be deepened with interviews and participant observation during the policy lab sessions of WP 5.	recommendations for pilot regions and other similar projects in the future elsewhere in the Europe. Support embedding RRI in regions' innovation systems and development policy approaches	notes for internal use and meeting minutes) Recorded and transcribed interviews Other written research material (e.g. literature reviews)
WP 7 Dissemination and communication activities	Newsletter, Leaflets, brochures, Project branding materials, Social media, Website of the project Plan for dissemination and exploitation of project results and various internal and external dissemination events	Dissemination reporting outcomes and successes of the project and supporting the overall implementation of the project through a set of communication activities. Raising awareness about the TetRRIS project and allow for the mobilization of as many relevant stakeholders as possible	Written documents (e.g. project deliverables, notes for internal use and meeting minutes) Promotional materials and dissemination material: written blog posts, social media posts, pictures, videos, etc.
WP 8 Management	Internal communication SG as well as consortium meetings	Administrative and financial management Quality assurance and risk management	Written documents (e.g. project deliverables, notes for internal use and meeting minutes)

Will you re-use any existing data and how?

Generally, the data of the project will be used only for research purposes and handled ethically and lawfully. There are not any commercial aims. Regarding the re-use of data externally of this project, the project aims to make public data broadly available and improve the open access of collected and generated data. This manner will be further addressed later in this document.

Regarding internal re-use of the data, in many occasions the data collected or generated will be re-used in the upcoming WP's. For example, the insight gained in the WP2, 'Mapping and Analysis of Actors', will be used when deciding e.g. participants for workshops or policy lab sessions or a target group for dissemination activities. In addition, WP 6, 'Analyses, Reflections and Recommendations', will be built on all previous WP's and participant observation during the policy lab sessions of WP 5. Also, dissemination materials of WP 7 will contain references to previous writings, research findings, project results etc.

Additionally, existing external data will be re-used in this project. For example, project deliverables often contain information from existing literature which will be sourced through a desk research. Publicly available data will also be used for analyses. Any re-used data will be clearly cited and acknowledged.

What is the origin of the data?

Data used in the project will originate and be collected through relevant primary and secondary literature, interviews, workshops and policy labs to be organised in each of the pilot regions of TetRRIS project as well as observation. In addition, as a part of the internal quality assurance there will be carried out an evaluation of outcomes of the project in terms of impacts. For that purpose, data will be collected by conducting key person interviews and pilot territory specific surveys.

Regarding recruitment of participants to the TetRRIS activities producing data, a recruitment plan/procedure has been prepared and described in context of Deliverable 1.3 on identification and enrolment of people in the project activities.

Recruitment of human participants consists of contacting identified actors in each pilot region's territorial research and innovation system and providing them with sufficient information on the TetRRIS project in order to make them interested to engage in the project activities. Initial mapping of actors in pilot territories' research and innovation systems is part of the implementation of Work Package (WP) 2. In context of scoping and co-design of pilot actions (WP3), the core partners and the regional partners jointly identify and recruit local stakeholders presenting territorial research and innovation system and beyond to participate in workshops. In start of the actual pilot actions (WP4), there will be reached out to relevant stakeholders who have been identified in preceding stages of the project (WP2 and WP3), but who have not yet been directly involved. In following phase (WP5), the project partners organize four policy labs to which also interested external stakeholders will be invited. Project Information Sheets and GDPR compliant Informed Consent Forms will be used for each interview and event when dealing with human participants.

What is the expected size of the data?

The size of the data will be rather small since mainly qualitative nature of data collection. In case more extensive quantitative analysis and therefore increasing the expected size of data, the data management plan will be updated accordingly.

To whom might it be useful ('data utility')?

Data might be useful e.g. for

- researchers in the fields of Responsible Research & Innovation (RRI) or
- organizations who are aiming to implement RRI methods in their research activities or
- regions and policy makers who are aiming to promote RRI aspects and implementation in their own and their stakeholders actions and operations.

2. FAIR data

FAIR Data Principles sets guidelines for data management to make research data Findable, Accessible, Interoperable, and Re-usable (FAIR) to the greatest possible extent. Especially when needed to validate data and results presented in scientific publications.[2]

2. 1. Making data findable, including provisions for metadata

The first FAIR principle is to make data findable which means that the datasets collected and used in the project can be easily discovered by other researchers.

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

To ensure that data produced is findable, quality control measures will be taken to maintain the accuracy of data during the project. Labelling the produced data material in a consistent and transparent manner with metadata elements and identifiers is a key element in ensuring that the data is findable afterwards. List of metadata elements and metadata standards used will be compiled and provided in a separate spreadsheet in the end of the project. Metadata of opened data will be made available for research and re-use after project closure.

The Steering Group will discuss and make a decision if the project should obtain a Digital Object Identifier (DOI) entry for all project outputs labelled public. In addition to scholarly publications, DOI entries can be created for data sets and associated outputs. DOI entries contribute to accessibility of research data and its discovery and reuse. The DOI system is conceived as a generic framework for managing identification of content over digital networks, recognising the trend towards digital convergence and multimedia availability.

What naming conventions do you follow?

Deliverables produced during the project will be named in the following format:

DX.X_NameOfTheDeliverable_VX.Dataformat

DX.X stands for the deliverable identifier, NameOfTheDeliverable for the name and VX for the version and Dataformat for the document type/extension such as pdf, docx and so on. As an example: *D8.2_DataManagementPlan_V1.0.pdf*

Possible changes regarding naming conventions can be decided by the project steering group. In case this happens, the changes will be updated in the Data Management Plan accordingly.

Will search keywords be provided that optimize possibilities for re-use?

Keywords will be used whenever it fits to the purpose and features of the publication or document. The list of key words will be updated continuously throughout the project.

[2] http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf

Following is an exemplary list of optional keywords used for TetRRIS documents and publications: *responsible; research; innovation; science; policy; Responsible Research and Innovation; RRI; RRI Tools; Public engagement; Participatory/Participation; Regional innovation systems; responsibility; ethics; governance*

Do you provide clear version numbers?

The documents produced will be provided with clear version numbers and dates of release. Information of the document history is recorded in a control table, which includes name(s) of contributors, date, version number as well as a short document status description/description of change(s) made. Example of the version control table used is illustrated below.
Version numbering system applied will be discussed and agreed on in the project Steering Group. It is suggested that in version numbering is followed an approach reflecting the status of the document. For example, the whole number (e.g. 1.x) is used to refer to a document involving major changes, e.g. the initial draft version or subsequent revision(s) requiring screening and (re-)approval, while the decimal number (e.g. x.5) refers to a commented version.

An example of a version control table:

Name	Date	Version	Description
Name of the person	**xx.xx.xxxx**	**1.0**	**First draft up for peer review**

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Anticipations of created metadata throughout the project are described in the table 1. Anticipations are based on the TetRRIS Grant Agreement and realized creation of metadata might differ from the planned. Additionally, the table of reflect realized inputs and outputs regarding project's metadata is provided in the Annexes of this document (Table 3).

2.2. Making data openly accessible

The second FAIR principle is making data openly accessible which means providing open access to the data used in a research project, enabling other researchers to verify the published findings of the project and conduct further research using the same data.

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Deliverables classified as public (Deliverables in WP's from 2 to 8) in the grant agreement will be shared with a wider audience. Deliverables classified as confidential (Deliverables in WP 1) will only be shared with members of the consortium, including the Commission Services. The principle is that the deliverables not restricted to the consortium should be disseminated and made available for free.

Final decisions concerning the sharing of (selected) datasets will be taken by project steering group. The Project Coordinator in collaboration with project partners will take the

appropriate measures to make relevant data openly available and usable for third parties for study, teaching and research purposes.

Personal data collected as part of the project will only be shared publicly if agreed with individuals included in the dataset and if it can be justified by research purposes. Other than that, regarding the publicity of research's metadata collected by each consortium partner will be discussed and decided in the course of the TetRRIS project.

Open issues regarding this are:

- What kind metadata will be public and openly accessible (e.g. anonymized or pseudonymized outputs and outcomes from interviews or surveys)
- When metadata will be published (real-time or at the end of the project)

The Data Management Plan will be updated accordingly.

Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement and are in line with the reasons for opting out.

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

Following issues will be taken into account regarding accessibility of the data and accessibility of the data will be constantly improved taking into consideration more issues when the project goes forward.

- Data used in the project will be shared internally in the consortium and therefore consortium members will have access to the data collected in TetRRIS. For this, web-based Microsoft Teams platform will be used which is hosted by VTT as a project coordinator.
- As previously pointed, all deliverables classified as "public" will be made freely available on the TetRRIS website after the deliverables are submitted and approved by the Commission.
- Datasets will be made available in common file formats such as Microsoft Excel or Word. No advanced or prohibitively expensive software will be required to access the data.
- In releasing and sharing results/publications/documents there will be taken into account that it will be made by openly accessible way, e.g. by aiming consistently to publish in Open Access journals which are free of charge to readers.

Open access data issues will be discussed further in the Steering Group meeting and the Data Management Plan will be updated accordingly.

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

Required Software to access the data will be depended on the type of the data format. Main software tools for accessing and sharing data in this project will be Microsoft Office, such as Ms. Word, Ms, Excel and Ms. PowerPoint. Other software, such as SPSS, might be required for accessing additional data. The priority aim is to provide data formats that can be accessed with open source software.

Is documentation about the software needed to access the data included?

The data made openly accessible will be complemented with a documentation of recommended software for further analysis.

Is it possible to include the relevant software (e.g. in open source code)?

There are no plans to include the relevant software as such, but in the data documentation references and links can be provided to download software where appropriate.

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.

The data shared between the partners and produced by the consortia will be saved on the VTT Teams TetRRIS platform. The platform can be used to deposit also confidential data. In sharing confidential data, particular attention is paid to the data minimisation principle to avoid unnecessary storage of personal data in the shared workspace. Personal data, such as interviewee's name and contact information should be deleted or in minimum pseudonymised by the partner(s) before uploading the data on the TetRRIS Teams platform. The personal data has to be kept in a safe way at the local repository of the partner(s) that collected the data.

Regarding open data, the Steering group considers available options via which to make project metadata, documentation and code available. Aim is to archive published and FAIR - compatible data in a common and open data repository. Recommended generic and certified repository services include CSC's IDA or CERN's Zenodo, either of them can be used to enhance long-term accessibility and re-usability of the data.

Have you explored appropriate arrangements with the identified repository?

VTT has an extensive understanding and experience of the use of secure storage opportunities for data on dedicated Microsoft Teams project platforms.

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

Access to data will be managed by the Project Coordinator. Data on consortium Teams channel can only be accessed on invitation. Consortium partners will be instructed about the appropriate conduct with sensitive data.

If, after project closure, permission to re-use the data is required, all requests for further use of data will be considered carefully and whenever possible approved by Project Coordinator or the person mandated with the task. Permission for data use will be granted providing there are no IPR or confidentiality issues involved or any direct overlap of research questions with the primary research. Permission will be provided by contacting Project Coordinator. Contact information and appropriate procedure will be provided in connection with other metadata.

Is there a need for a data access committee?

Currently there is not envisaged need to set up a data access committee. In case situation changes, the Data Management Plan will be revised accordingly.

Are there well described conditions for access (i.e. a machine readable license)?

The aim is to provide each dataset made openly accessible with a clear description and instruction how to access and process the data. Opportunity to use open source software in data processing and analysing will be taken into account when selecting data saving format for each dataset planned to be made openly accessible.

How will the identity of the person accessing the data be ascertained?

Access to the dataset(s) will only be granted to trustworthy and known mail addresses. The list of e-mail addresses of consortium members' personnel involved in the project have been collected at the beginning of the project. If there are changes in the personnel participating in the project, the list is updated accordingly.

2.3. Making data interoperable

The third FAIR data principle is making data interoperable which means enabling and promoting data exchange between researchers, institutions, organizations and countries and ensuring the usability and utility opportunities of the data.

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

The data collected and generated will be shared in an interoperable format for further use if it is decided to be appropriate for a given data set. When data is made interoperable, data anonymisation and pseudonymisation standards will be used.

In pseudonymization, numerical codes will be used. Numerical codes won't have any connection with the personal identifying information. Personal identifying information and its corresponding numerical code will be stored separately from the document containing the anonymized data.

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

These issues will be decided in the course of the TetRRIS project and the Data Management Plan will be updated accordingly.

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

This issue will be decided in the course of the TetRRIS project and the Data Management Plan will be updated accordingly.

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

It is likely that only standard vocabulary will and ontology will be used in the TetRRIS project. In case of changes in this matter, the Data Management Plan will be updated accordingly and mappings for specific ontologies and vocabularies will be provided.

2.4. Increase data re-use (through clarifying licences)

The final FAIR data principle concerns increasing data re-use which can be done for example by promoting opportunities to data access, mining, exploitation, reproducing and dissemination.

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

The project aims to make public data broadly available. Ownership of datasets will belong to project consortium after the project completion. Creative Commons licence CC-BY-SA or CC-BY will be used for any opened datasets, unless there are compelling reasons to select more restricted type of CC-licence. Creative commons licences will by default include also a disclaimer of liability for the re-use of opened data.

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Data assessed appropriate for open access will be made available after decision of the project Steering Group/Consortia and if needed confirmation of the European Commission. Prior to this and depending on the nature of data in question, it may need to go through the anonymisation or pseudonymisation process. In addition, if necessary, the project team acquires permission(s) of the data subject(s) described in the data.

According to the Grant agreement, the data will be stored in minimum five years. However, no definite period or time limit is planned for access or re-use of the data. Justification for possible case-spesific embargo for published data will be decided by project consortium. Embargo will be sought primarily in connection with any potential patent application based on project results.

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

It is foreseen that the project's outputs and data can be useable for instance to actors and stakeholders involved in planning and implementation of regional innovation activities and policies. As a rule, aim is to make data usable for third parties if it is classified as public data by the project Steering Group. Any changes to the release policy will be updated in the Data Management Plan and communicated publicly, through e.g. the project website.

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

Article 18.1 in the Grant Agreement sets an obligation for the beneficiaries to keep records and other supporting documentation for a period of five years after the payment of the balance. Following this requirement, data access will be preserved and maintained for a minimum of 5 years after the completion of the project and after the moment of publication of its results. In case there is made changes to the duration of data access, it will be revised and indicated in the DMP.

Are data quality assurance processes described?

Qualitative data, produced through interviews, workshops, conferences etc. will be documented through notes and minutes. Especially interviews will be recorded and transcribed.

All deliverables of the project will be peer-reviewed internally by consortium partners. Improvements and adjustments to the quality assurance processes will be reviewed throughout the project and added to the Data Management Plan.

Further to the FAIR principles, DMPs should also address:

No identified issues.

3. Allocation of resources

What are the costs for making data FAIR in your project?

The costs to make data FAIR have not been discussed yet in the consortium. MIKA. Costs related to research data management and opening are eligible as part of the project grant; costs needed to make research data quality-controlled, FAIR-compatible and as open as possible.

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

These costs are allocated under “Other costs” defined in the Grant Agreement. Each partner has budgeted funds for e.g. covering open access publication costs.

Who will be responsible for data management in your project?

The Steering Group (SG) together with the coordinator of TetRRIS will be responsible for the general data management in the project. General data management in the project including preparation and necessary updates of the Data Management Plan is part of Work Package 8 lead by VTT. During the project, consortium partners will be responsible for managing and curating datasets at their possession. At the project ending, consortium steering group will mandate Project Coordinator to take care of long-term preservation and sharing of jointly created datasets.

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

This has not been under the discussion yet in the consortium meetings, but it will be addressed later on when it is topical. Information considering these issues will be added in the Data Management Plan on that time. In arrangement of long-term and secure preservation of published research data, it will be ensured that only certified and OpenAIRE guidelines compatible repositories are used.

4. Data security

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

Following data access guidelines are covered in TetRRIS Grant Agreement:

- Confidentiality: Every partner will treat information from the other partners as confidential and not disclose unless the information is already public or disclosure is required by law.
- Ownership of Knowledge: the knowledge is co-owned by the partners in the project.

- The access to the rights: Partners privilege each other with free access rights to knowledge generated in the project. Access rights to knowledge generated in the project and to pre-existing knowledge for use outside the project are, when needed to make use of the project result, given between partners in different WPs on preferential conditions. Access rights to knowledge generated in a WP, when needed to make use of the project result, is given royalty free to the other partners participating in the same WP. Access rights to a partner of pre-existing knowledge for use outside the project is, when needed and only to the extent necessary to make use of the project result, given on preferential conditions to the other partners in the same WP.

During the project, research datasets will be available only to those project partners or project consortium members, who have been accredited by and their data usage has been approved by Project Coordinator or authorized project consortium member. Project partners will be responsible for curating, preserving, disseminating and deleting in appropriate manner the datasets in their possession. Retention time for curated datasets will be the same as for other project results at the project consortium partners.

Data collected or acquired within the project will be stored in secure cloud environment provided by VTT′s IT service provider Fujitsu Ltd. The actual solution selected by the project consortium is the web-based Microsoft Teams platform which is used for internal communication and collaboration between the partners as well as data exchange and storage for the project documents. Microsoft Teams platform uses industry standard technologies such as user and client authentication through Azure Active directory and OAuth, as well as TLS (Transport Layer Security) and SRTP (Secure Real-Time Transport Protocol) protocols to encrypt all data in transit between users' devices and Microsoft datacenters, and between Microsoft datacenters. Enterprise data is also encrypted at rest in Microsoft datacenters.

Access to the Teams project workspace needs registration and authentication. Project Coordinator will check applications for the use of data. The workspace can be accessed only by on invitation to prevent unauthorised access to project related material. If person discontinues being involved in the project s(he) will be denied access to data and other project material on the Teams platform and any other repository as soon as notice has been given to the Coordinator.

All tools used in the project for the collection, transfer and storage of personal data (e.g. registration tools for workshops and events such as policy labs, survey tools) are selected so that they are GDPR 2016/679 compliant. In the same way, the compliance with GDPR regulation will be followed in data handling, storage and possible transfer after the project is completed.

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

The project consortium will discuss and make decision on long term preservation of project data and deliverables in due time. Decision will cover questions regarding which type of project data, on which platform and for how long will be preserved.

5. Ethical aspects

The TetRRIs project pertains to the following ethics issues as described in the Ethics Issue table of the proposal: *involvement of human participants and processing of personal data*. Ethical standards and guidelines of Horizon 2020 will be rigorously applied in the project. In addition, further consideration shall be given to the relevant regulation, decisions and

guidelines including the Charter of Fundamental rights of the European Union. The TetRRIs project is based on the idea of co-creation and stakeholder participation and will include interviews, workshops and interactive co-creation events. The project will respect and fully apply ethical principles and current international, EU and national law and regulation for any activities that involve human participants and processing of personal data. The participants in interviews, workshops and co-creation activities will be adults with voluntary consent to their involvement in research, on the basis of full, accurate, and clear information.

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

All activities involving human participants and processing of personal data will be organised by project partners with appropriate data protection policies and information security provisions in place. The project follows Horizon 2020 ethical standards and guidelines, and GDPR regulation in all activities involving participants and processing of personal data.

Data subjects shall be informed about intended use of their personal data and granted right to erasure of their personal data under provisions of GDPR 2016/679. The TetRRIS project follows data minimization principle – only strictly necessary personal data required in implementation of the project will to be collected. No sensitive personal data (e.g. health, ethnicity, religious convictions) will be collected in the project. Previously collected data will only be used if the subject has given explicit consent for data usage. Consent is given if the person is participating in interviews, workshops, surveys or other activities of the project on a voluntary basis after invitation by project partner(s) and by signing the consent form (see D1.1). Further processing or profiling of personal data that exceeds the scope of the project, or processing of such data regarding as genetic, biometric and/or health information will not be performed within the TetRRIS project.

As defined in the Consortium Agreement (4.4), in the event personal data is processed in the project, the partners undertake to respect their obligations in application of regulations in force and, especially, the GDPR regulation on the protection of natural persons with regard to the Processing of personal data and on the free movement of such data. The partners will not disclose to each other personal data without first entering into a separate written agreement for such purpose, except for the necessary personal data of persons participating in the Project or conclusion of the Consortium Agreement, which the Party is legally entitled to disclose.

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

Yes. Participants are informed of the data sharing and long term preservation regarding personal information in the consent form.

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

At this stage, no other procedures are applied. Any changes to this will be stated in the next iteration of the Data Management Plan.

7. Further support in developing your DMP

The Research Data Alliance provides a Metadata Standards Directory that can be searched for discipline-specific standards and associated tools.
The EUDAT B2SHARE tool includes a built-in license wizard that facilitates the selection of an adequate license for research data.
Useful listings of repositories include:

- ***Registry of Research Data Repositories***
- ***Some repositories like Zenodo, an OpenAIRE and CERN collaboration), allow researchers to deposit both publications and data, while providing tools to link them.***
- ***Other useful tools include DMP online and platforms for making individual scientific observations available such as ScienceMatters.***

Open issues in TetRRIS Data Management Plan

As mentioned in the beginning of Data Management plan, this document is a living document, thus it will be updated in the course of the Project when required. Some issues which will be addressed in this data management plan need further discussion with TetRRIS Steering Group (SG) and approval by them. Any changes to the Data Management Plan will be communicated and introduced in the next iteration of the document.

A list of these issues can be found in the table below.

Table 2. Issues to be addressed with the TetRRIS Steering Group

Topic	Purpose/Goal	When this will be addressed?	Decision/Output
All research data collected and generated is recommended to be shared with all consortium members as widely as possible. Data containing personal information should be anonymized our pseudonymized when shared with the whole consortium. All consortium partners have their own (web-based) platform/system for each's internal communication, secondary data storage and project management activities. Regarding data storage, these platforms are required to be protected and accessible only for research team.	These issues should be further discussed and agreed		
Pseudonymization principles	To confirm awareness of these principles		

What data and metadata vocabularies, standards or methodologies TetRRIS will follow to make your data interoperable? Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability? In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?	-Discussion and decisions regarding these questions		
How to improve and take into consideration open data access – in TetRRIS in general and in each pilot.	-To get more official tools and activities to improve this -Awareness about this issue among the partners		
What kind of metadata will be provided public and openly accessible? (e.g. will anonymized or pseudonymized outputs and outcomes from interviews or surveys be published in certified repository services?) What repository services will be used in this purpose? Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers) When metadata will be published? (during the project or at the end of the project?) This question refers to the previous one since first it must be discussed what kind of metadata will be published in the first place.	-Discussion and decisions regarding these questions		

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?	-Discussion and decisions regarding this question		
Naming and numbering conventions	Informative purposes		
Updateable metadata table (Table 3)	Informative purpose		

SUMMARY TABLE 1

FAIR Data Management at a glance: issues to cover in your Horizon 2020 DMP

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

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 mechanism. Do you make use of persistent and unique identifiers such as Digital Object Identifiers? • Outline naming conventions used • Outline the approach towards search keyword • Outline the approach for clear versioning • Specify standards for metadata creation (if any). If there are no standards in your discipline describe what 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 re-use. If applicable, specify why and for what period a data embargo is needed • Specify whether the data produced and/or used in the project is useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why • Describe data quality assurance processes • Specify the length of time for which the data will remain re-usable
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)

ANNEX

Table 3. Metadata by WP reported by WP leaders

WP Number and name, Deliverables and WP Leader		Type(s) of Data Collection/Generation	Purpose of Data Collection/Generation and it's relation to project's objectives	Formats of Data
WP1 Ethics Requirements				
D1.1 Description and Template for Informed Consent Procedure	VTT			
D1.2	VTT			
D1.3	VTT			
WP2 Mapping and Analysis of Actors				
D2.1 Validated Mapping and Analysis Framework	Fraunhofer			
D2.2 Mapping Report for each Pilot Territory	Fraunhofer			
WP3 Scoping and Co-Design of actions				
WP4 Pilots				
WP5 Experimenting, Learning & Sharing				

WP6 Analyses, Reflections and Recommendations				
WP7 Dissemination and communication activities				
WP8 Management				
D8.1 Work Plan for the project				
D8.2 Data Management Plant	VTT			
D8.3 Interrim report of the project	VTT			
D8.4 Final report of the project	VTT			
D8.5 Policy Brief on challenges and interventions	VTT			
D8.5 Policy Brief on challenges and interventions	VTT			