DATA MANAGEMENT PLAN

Grant Agreement number	23IND11
Project short name	ThermoSI
Project full title	Thermometry with embedded SI traceability for industrial applications
Data management plan	1st ☒ 2nd ☐
Confidentiality status	PU - Public, fully open

TABLE OF CONTENTS

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EURAMET. Neither the European Union nor the granting authority can be held responsible for them.

The project has received funding from the European Partnership on Metrology, co-financed from the European Union's Horizon Europe Research and Innovation Programme and by the Participating States.

1 Data management plan

1.1 Data summary

Questions	Answers
1 Will you re-use any existing data and what will you re-use them for? State the reasons if re-use of any existing data has been considered but discarded.	This project will re-use: • Internal pre-existing data from the participants • Publicly available data (e.g. from scientific literature) These data will be used for the selection and assessment of uncertainty of the instrumentation and methods developed in the project, as input for calculations or for comparison with new measurements carried out in the project, and for validation of the project's results.
2 What types and formats of data will the project generate or re-use?	The actual produced data formats will be in accordance with the requirements, in particular: • Microsoft Excel ® or CSV format for measurement data and their reduction • Pdf files for reports and notes; occasionally these might be in Microsoft Word ® or PowerPoint ® • JPG, TIFF, SVG, etc… for images e.g., photographs of experiments, AVI, MP4 for video recording of trials • Data from models may have formats such as Matlab ® and R ®, python code
3 What is the purpose of the data generation or re-use and its relation to the objectives of the project?	The data will be used in meeting the project's objectives, at conference, and publications. Data will be generated by the consortium in order to meet objectives 1 – 4. Measurement, calibration and comparison data will result from objectives 1 – 4, Data from questionnaires and market surveys may be used to support end-user uptake (objective 5).
4 What is the expected size of the data that you intend to generate or re-use?	The overall size of the data is expected to be in the range 1 – 50 GB.
5 What is the origin/provenance of the data, either generated or re-used?	<u>Data generated in the project</u> The data generated will be from questionnaires, measurements, calibrations, models, comparisons and validations. <u>Re-used data</u> The existing data will originate from several sources, which will include: participant's pre-existing data, data from the scientific literature, real-world measurement data and data from simulation experiments.
6 To whom might your data be useful ('data utility'), outside your project?	The data generated during the project realization will be and should be useful to industry, partners, and stakeholders of the project. The project will also contribute to standardisation committees like IEC TC 65, CCT, EURAMET TC-IM and TC-T, IEC/TC 65 UkrNDNC TK 65, IEC/TC 65 Ukrainian National Mirror Committee TK 85, VDI, IMEKO TC12, and Nadcap/SAE.

1.2 Findable, Accessible, Interoperable and Re-usable (FAIR) Data

1.2.1 Making data findable, including provisions for metadata

Questions	Answers
7 Will data be identified by a persistent identifier?	Yes, each of the project's deposited datasets will be identified by the Digital Object Identifier (DOI).
8 Will rich metadata be provided to allow discovery? What	The metadata created for all of the project's datasets will fulfil the repository's (Zenodo) requirement for a minimum set of metadata

metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.	(i.e. 1) description, 2) creator / ownership, 3) access, 4) lifecycle, 5) persistent identifiers).
9 Will search keywords be provided in the metadata to optimise the possibility for discovery and then potential re-use?	Yes, discipline-specific search keywords will be provided in the metadata to optimise the discovery and potential re-use of the deposited datasets.
10 Will metadata be offered in such a way that it can be harvested and indexed?	Yes. The project's datasets and associated metadata will be deposited only in repositories ensuring the FAIR principles (https://www.go-fair.org/fair-principles) of data publishing. Zenodo complies with FAIR principles (https://about.zenodo.org/principles/).

1.2.2 Making data accessible

Questions	Answers
Repository:	
11 Will the data be deposited in a trusted repository?	Yes. The data and associated metadata, documentation and code will be deposited in the open access repository Zenodo (https://zenodo.org).
12 Have you explored appropriate arrangements with the identified repository where your data will be deposited?	No, the data will be uploaded via a standard procedure and require no special arrangements.
13 Does the repository ensure that the data are assigned an identifier? Will the repository resolve the identifier to a digital object?	Yes (see question 7, 11), it will be made sure that the project's datasets are published in a trusted repository where each dataset is assigned a persistent identifier. For example, Zenodo will assign an identifier (DOI) to each of the project's deposited datasets. The repository will resolve the identifier to a digital object.
Data:	
14 Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.	All the data associated with scientific publications will be made openly available as default unless there is a specific reason not to publish the data or the full dataset is included in the publication. Datasets which cannot be shared – voluntary restrictions Other data may be made available on a case-by-case basis if it is relevant for third parties. The following data will not be made publicly available: • Data obtained with the permission of third parties, but the third parties have not agreed to make the data publicly available. • Data that discloses the identity of a manufacturer. • Data that compromises the protection of a partner(s) intellectual property.

Questions	Answers
	The level of data made available will also be considered, for example, pre-processed data will not be provided unless there is a clear reason for doing so. Datasets which cannot be shared - legal and contractual reasons Market or customer survey data, which are commercially sensitive, will not be shared. Data with GDPR restrictions will not be made available, for example raw data with identifiable personal information.
15 If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.	The data used in scientific publications, posters and oral communications will be made available for re-use as soon as is reasonably possible. It is expected that the final data will be available for re-use after the end of the project. The first data used during project in scientific publications, posters and oral presentations will be made available for re-use earlier. The possibility of applying an embargo on some of the datasets will be decided by the Project Management Board on case-by-case basis. Should such a case arise, the embargo period will be discussed, set, and documented.
16 Will the data be accessible through a free and standardised access protocol?	Yes, in case of Zenodo, well described conditions for free and standardised access are provided (http://about.zenodo.org/policies).
17 If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?	There are no restrictions on the use of the published data, but users will be required to acknowledge the European Partnership on Metrology program, the project, the consortium and the source of the data in any resulting publications and in any public use of the data (e.g., in presentations).
18 How will the identity of the person accessing the data be ascertained?	There is no need to ascertain the identity of persons accessing the data.
19 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?	This consortium will not have a Data Access Committee, because the project is not going to produce sensitive data. All results will be publicly available without restrictions, with the exception of data specified in Question 14.
Metadata:	
20 Will metadata be made openly available and licensed under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?	The metadata created for all of the project's datasets will fulfil the repository's requirement for a minimum set of metadata. In case of Zenodo, metadata are licensed under CC0, except for email addresses. All metadata are exported via OAI-PMH and can be harvested. The metadata will contain information to enable the user to access the data.
21 How long will the data remain available and findable? Will metadata be guaranteed to remain available after data are no longer available?	The data will remain available, findable and reusable for the lifetime of the repositories used (see Question 11), which is expected to be at least 10 years. If data are withdrawn from the repository (e.g. Zenodo), the DOI and the URL of the original object are retained. In case of closure of the given repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.
22 Will documentation or reference about any software be	It is not expected that documentation will be required – the data are in a common format and can be read using widely available software (open

Questions	Answers
needed to access or read the data and will this be included? Will it be possible to include the relevant software (e.g. in open source code)?	source or commercial). In addition, the data can be read using specialised scientific software (open source or commercial).

1.2.3 Making data interoperable

Questions	Answers
23 What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?	The project's datasets and associated metadata will use the trusted repository's basic metadata schema for administrative data, which is compliant with the recommended standards used by DataCite (https://search.datacite.org), OpenAIRE (https://www.openaire.eu) and BASE search (https://www.base-search.net). To guarantee the interoperability of project's data/research outputs, individual datasets will be described using affirmed discipline-specific vocabularies, standards, formats, and methodologies, including but not limited e.g., GUM, VIM, OBO foundry, DICOM, NetCDF, HDF5, CityGML, INSPEC, ISO 9001, ISO 80000.
24 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? Will you openly publish the generated ontologies or vocabularies to allow their re-use, refinement or extension?	No mappings will be necessary, as the datasets will be described using standard terminologies.
25 Will your data include qualified references[1] to other data (e.g. other data from your project, or datasets from previous research)?	Yes, the project's datasets that will be deposited in the chosen repository (Zenodo) will include qualified references to other datasets from the same project and/or to other datasets from previous research.

1.2.4 Increase data re-use

Questions	Answers
26 How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?	A short README file (e.g. in Markdown format) will be provided together with the data, in order to enable data analysis and to facilitate data re-use.
27 Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard re-use licenses, in line with the	The data will either be licensed under the latest available version of the Creative Commons Attribution International Public License (CC BY) or a license with equivalent rights as set out in the Grant Agreement. Users will be required to acknowledge the European Partnership on

[1] A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source: https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/)

Questions	Answers
obligations set out in the Grant Agreement?	Metrology program, the project, the consortium and the source of the data in any resulting publications.
28 Will the data produced in the project be useable by third parties, in particular after the end of the project?	Any data published in open-access journals will be usable by third parties after the datasets have been deposited in a trusted repository (see Question 11). The data that do not relate to peer-reviewed publications will be made available for re-use on a case-by-case basis.
29 Will the provenance of the data be thoroughly documented using the appropriate standards?	Yes, the provenance and context of the data will be thoroughly documented to meet relevant standards using the Provenance and Context Content Standard (PCCS) Matrix (cf. https://eos.org/opinions/the-importance-of-data-set-provenance-for-science). Data will be accompanied by information on how they were captured, processed, analysed, and validated. Other information that enables interpretation and use will also be provided.
30 Describe all relevant data quality assurance processes.	Each participant will assure the data quality by working according to each internal certified quality management system and certified technical procedures. Generally, data quality will be assured through repeated measurements, adherence to standards for data recording, the use of controlled vocabularies and standard terminology, through the metrological characterisation of the measurement set-ups, provision of test results along with the data and through the validation of the data collected.
31 Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.	Allocation of resources The estimated costs for making the (data and) other research outputs FAIR are 3,000 € (personnel costs), see Question 34. The costs for making other research outputs FAIR are included in the project's budget and will be claimed if compliant with the Grant Agreement's conditions. The Project Management Board (PMB) will also have overall responsibility for managing other research outputs (see Question 36). Where feasible, long-term preservation will be ensured by depositing the other research outputs in repositories. The PMB will decide on a case-by-case basis on which other research outputs will be deposited and for how long. Security of other research outputs The participants will store other research outputs on their organisations' networks, which are protected by firewall, backups etc. Other research outputs will also be stored in the project's SharePoint environment, with password-protected login. Deposition in public repositories will provide additional security as they have multiple replicas in a distributed file system which is backed up on a nightly basis. This project will not generate sensitive other research outputs. The other research outputs will be safely stored in open access repositories. Ethical aspects Ethical issues will be discussed and addressed by the PMB on case-by-case basis. The project will not share other research outputs with identifiable personal information. Sensitive information relating to the other research outputs will be collected, separated as soon as possible, and kept secure.

1.3 Other research outputs

Questions	Answers
32 In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).	The software developed in the project will be released under a GNU-GPL license. The new calibration methods and protocols produced by the project will be published in form of peer-reviewed scientific publications and/or project deliverables. The most suitable form of publication will be decided by the Project Management Board on case-by-case basis.
33 Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.	As far as possible, the FAIR data approaches specified in Questions 7 – 30 above will be applied to the management of this project's other research outputs. This commitment will be met by releasing the new software that will be developed in the project under license, by placing the new calibration methods, and protocols, in a repository and by patenting the new materials that will be developed in the project in line with the requirements of the project's consortium agreement.

1.4 Allocation of resources

Questions	Answers
34 What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?	The estimated costs for making the data and other research outputs Findable, Accessible, Interoperable and Re-usable (FAIR) are 3,000 € (personnel costs). The costs will be kept to a minimum by using a free trusted repository (see Question 11) and by making only relevant data and other outputs FAIR.
35 How will these be covered? Note that costs related to research data/output management are eligible as part of the European partnership on metrology grant (if compliant with the Grant Agreement conditions).	The costs for making the data FAIR are included in the project's budget and will be claimed if compliant with the Grant Agreement's conditions.
36 Who will be responsible for data management in your project?	Each participant will assure sound data quality (see Question 30) as well as appropriate research data archiving. The initial responsibility for the proper data management of a given dataset (or record, software, etc.) lies with the partner who created the dataset. Furthermore, the following project participants have been appointed as data coordinators: • DTI as project coordinator and WP6 produced data • WP leaders: • PTB for WP1 produced data • PTB for WP2 produced data • NPL for WP3 produced data • DTU for WP4 produced data • STRATH for WP5 produced data • STRATH as project webpage content administrator

37 How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?	Long-term preservation will be ensured by depositing the data within trusted repositories. There are no costs associated with the long-term preservation of the data in these repositories. The potential value of the data will likely increase over time because of its fundamental impact in a wide range of applications (there is the possibility that some data may become superseded by more up-to-date data). The data will also be of value as it underpins the results of published datasets. The Project Management Board will decide on a case-by-case basis on what data will be kept and for how long.

1.5 Data security

Questions	Answers
38 What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?	Data recovery and secure storage All participants will store data on their organisations' networks, which are protected by firewall, backups, etc. Data will also be stored in the project's SharePoint environment, with password protected login. Deposition in the public repositories will provide additional security as it has multiple replicas in a distributed file system which is backed up on a nightly basis. Transfer of sensitive data This project will not generate sensitive data.
39 Will the data be safely stored in trusted repositories for long term preservation and curation?	Yes, the data will be safely stored in open access repository. The data will be safely stored in the Zenodo open access repository. Zenodo and the underlying Invenio Framework for digital repositories were designed according to the Open Archival Information Systems (OAIS) reference model. Zenodo is working towards ISO 16363 certification. The repositories are certified for long-term data preservation and curation.

1.6 Ethics

Questions	Answers
40 Are there, or could there be, any ethics 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 report(s) and the ethics section in the Annex 1.	It is not expected that there will be any ethical or legal issues that will affect the project data sharing. All data coming from third parties, e.g. manufacturers, will be shared only after obtaining explicit consent by the owners. Data from the market surveys/questionnaires will be made anonymous to comply with the General Data Protection Regulation (GDPR).
41 Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?	At present, this item has not been included in any questionnaires, but the project has no plans to share data with identifiable personal information. If any sensitive data will be collected, they will be separated and kept secure.

1.7 Other issues

Questions	Answers
42 Do you, or will you, make use of other national / funder /	Data management will be compliant with:

sectorial / departmental procedures for data management? If yes, which ones (please list and briefly describe them)?	• The research data policy of the European Partnership on Metrology; • European laws about data security and the protection of privacy (e.g. GDPR); • Institutional guidelines; • Scientific community guidelines.

2 Open science: research data management

Statement	**Put an X in the box to confirm**	**Or, list any exceptions to this**
All participants have adhered to the requirements of the project's GA and CA with respect to open science: research data management (GA Article 17 and its Annex 5) for this reporting period	☒	