DATA MANAGEMENT PLAN

Grant Agreement number	23IND05
Project short name	H2FlowTrace
Project full title	Flow measurement traceability for hydrogen in gas networks
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: • Confidential data • Internal data from the participants • Publicly available data These data will be used for the following purpose: • CFVN design specifications (from MetHyInfra) in objective 1 • Good practice guide for CFVN dimensional characterisation (from MetHyInfra) in objective 1
2 What types and formats of data will the project generate or re-use?	The project will generate the following types and formats: • For graphics: jpeg, pdf, png, pptx, tiff • For tables: xlsx, csv, .ods • For texts: docx, pdf, txt, ods • For numerical data: csv, txt • For proprietary software: Labview (VI, LVPROJ, LVLIB…), SolidWorks (SLDPRT), .exe, Ansys Fluent or OpenFoam
3 What is the purpose of the data generation or re-use and its relation to the objectives of the project?	<u>Purpose of the data generation or re-use</u> The data generated and re-used will be from measurements, calibrations, comparisons and validations. They will be used in meeting the project's objectives and in conference and peer-reviewed publications. <u>Data generated in relation to the objectives of the project</u> Data will be generated in relation to project objectives 1, 2, 3, and 4: Development of reference methods/systems, calibrations of the nozzles and the skids, method validation, comparison of methods, development of guidelines, risk assessment and training and contribution to the development of standards. No datasets have been deposited yet. <u>Data re-used in relation to the objectives of the project</u> Measurement, calibration, comparison and validation data will be re-used by the consortium in order to meet objective 4. The design of the CFVN is based on the report "Design of CFVNs for high pressure applications with hydrogen" from MetHyInfra to meet the objective 1. The Good Practice Guide on dimensional characterisation (D3 from MetHyInfra) will be used to meet objective 1. Internal data from the participants is being used for making design choices for the SSTS and LSTS.
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: • 50 GB – 200 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 measurements, calibrations, comparisons and validations. The data collected from domestic properties will be anonymized and will be included in the repository.

	Re-used data 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. Reports or good practice guides from previous EURAMET's JRPs (e.g. MetHyInfra).
6 To whom might your data be useful ('data utility'), outside your project?	The data will be suitable for use by other research groups working on the following topics: hydrogen, flow metering, Distribution System Operators (DSOs) and Transmission System Operators (TSOs). It will also be useful for standards committees including OIML TC8/SC7, ISO/TC 30, CEN/TC 237, CEN/TC 234 and WG11. The data might be useful to: • Stakeholders from industry: TSOs (including participants GRTgaz and EnagasTSO), DSOs and energy companies, Manufacturers (including participants Emerson M and SICK) • Standardisation bodies: in particular CEN/TC 237 • Other scientists working in the field: WGFF (BIPM), EURAMET TC Flow, EMN for Energy Gases, EuReGa, IRENA • NMIs/DIs involved in the project or not and the EMN for Energy Gases

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 is identified by: • Handle • DOI
8 Will rich metadata be provided to allow discovery? What 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.	The metadata created for all of the project's deposited datasets will be open under a Creative Commons Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles (in particular machine-actionable) and provide information at least about the following: datasets (description, date of deposit, author(s), venue and embargo); the European Partnership on Metrology funding; grant project name, acronym and number; licensing terms; persistent identifiers for the dataset, the authors (ORCID) involved in the action, and, if possible, for their organisations (ROR, ISNI) and the grant. Where applicable, the metadata will include persistent identifiers for related publications (DOI) and other research outputs.
9 Will search keywords be provided in the metadata to optimise the possibility for discovery and then potential re-use?	Yes, the following search keywords are provided in the metadata to optimise the discovery and potential re-use of the deposited datasets: hydrogen, flow metering, budget uncertainties, Transportation and Domestic Gas Operators, gas networks, metrology.
10 Will metadata be offered in such a way that it can be harvested and indexed?	Zenodo complies with FAIR principles (https://about.zenodo.org/principles/). The metadata are indexed in a searchable resource. Metadata are licensed under CC0, except for email addresses. All metadata are exported via Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and can be harvested.

1.2.2 Making data accessible

Questions	Answers
Repository:	
11 Will the data be deposited in a trusted repository?	The data and associated metadata, documentation and code will be made accessible by deposition in the following trusted open access repository: https://zenodo.org/ We created a community in Zenodo.
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, 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 of the data that are needed to validate the results presented in scientific publications will be made openly available as the default unless there is a specific reason not to publish the data. 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 while no approval was obtained for the disclosure. • Data that compromises the protection of one or more participant(s) intellectual property.
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. No datasets deposited yet.
16 Will the data be accessible through a free and standardised access protocol?	The data will be accessible through the interface of the chosen open access repository. Zenodo offers well-described conditions for free and standardised access (see 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?	If some data have restricted access, access will only be provided after personal contact to the authors via the repository interface.
18 How will the identity of the person accessing the data be ascertained?	Users are required to register to use the repository.

Questions	Answers
19 Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?	This consortium does not have a Data Access Committee. While the project could produce sensitive data, it will be anonymized. All results will be publicly available without restrictions. Raw data from measurements will remain confidential and stored in a restricted repository after the end of the project.
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?	In Zenodo, metadata are licensed under CC0, except for email addresses. All metadata are exported via OAI-PMH and can be harvested.
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 and findable for the lifetime of the Zenodo repository or, at least, for 20 years. If data are withdrawn from the repository (Zenodo), the DOI and the URL of the original object are retained. In case of closure of Zenodo, best efforts will be made by Zenodo to integrate all content into suitable alternative institutional and/or subject based repositories.
22 Will documentation or reference about any software be 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)?	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 datasets use the trusted basic metadata scheme of for bibliographic metadata, which is compliant with the recommended standards used by: • DataCite (https://search.datacite.org/), • BASE search (https://www.basesearch.net/) • OpenAIRE (https://www.openaire.eu/search) For individual datasets, the following discipline-specific vocabularies, standards, formats, and methodologies will be used: 1. GUM (procedure; subject-independent). 2. VIM (JCGM 200:2012) 3. INSPEC (vocabulary + classification; physics).
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	The generated ontologies and vocabularies for the published data are provided, together with appropriate mappings in an informal, lightweight form (glossaries, alignment tables, …) to more commonly used ontologies.

allow their re-use, refinement or extension?	
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 (e.g. Zenodo) will include qualified references to other datasets from previous research (ex: MetHyInfra).

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. Markdown) 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 obligations set out in the Grant Agreement?	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 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 Zenodo. 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?	Data are accompanied by information on how they were captured, processed, analysed, and validated.
30 Describe all relevant data quality assurance processes.	Data quality will be assured through several quality assurance procedures: • Repeated and comparison measurements. • Adherence to standards for data recording. • Use of controlled vocabularies and standard terminology. • Metrological characterisation of the measurement set-ups. • Validation of the data collected. • Provision of test results along with the data. • Peer-review of publications based on the data.
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 costs for making the (data and) other research outputs FAIR is expected to be 1 000 € (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 consortium's management board had overall

[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
	responsibility for managing other research outputs (see question 36). Long term preservation will be ensured by depositing the other research outputs in repositories. The project management board will decide on a case-by-case basis which other research outputs is deposited and for how long. Security of other research outputs: For data protection, the same infrastructure is used to ensure that all drafts and final outputs are kept secure and protected from loss. The participants are storing other research outputs on their organisations' networks, which are protected by firewall, backups etc. Other research outputs are also stored in the project's SharePoint environment, accessible only by invitation. Deposition in public repositories provides additional security as these have backup systems. This project does not plan to generate sensitive other research outputs. The other research outputs are safely stored in open access repositories. Ethical aspects: At present, there are no ethical issues in our view. But we will keep the matter under review.

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 new calibration methods, and protocols produced by the project will be stored in the Protocol Exchange repository. The management of the IP issues surrounding the new materials that will be developed in the project have been planned for in the project's consortium agreement.
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 placing the new calibration methods, and protocols, in a trusted repository.

1.4 Allocation of resources

Questions	Answers
34 What will the costs be for making data or other research	The estimated costs to make the data and other research outputs FAIR (findable, accessible, interoperable, and reusable) are 0.5 PM

outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.)?	(personnel costs) and 1000 € (non-personnel costs). These costs have been kept to a minimum by using a free repository 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 as they were compliant with the Grant Agreement's conditions.
36 Who will be responsible for data management in your project?	The project management board (unit and WP leaders) had overall responsibility for research data management. The coordinator leads the committee and is responsible for coordinating updates to the data management plan. The project management board will be responsible for organising data backup and storage, data archiving and for depositing the data within Zenodo.
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 storage is ensured by depositing the data in trusted repositories. There are no costs associated with the long-term storage of the data.

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)?	<u>Data recovery and secure storage</u> All participants are either accredited to, or work in compliance with, the ISO 17025 standard on the "General requirements for the competence of testing and calibration laboratories". The 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 Zenodo public repository will provide additional security as it has multiple replicas in a distributed file system which is backed up on a nightly basis. Participants will store the data on their organisations' networks, secured by file backups, Virtual Private Network (VPN) access, firewalls, etc. The data will also be stored in the SharePoint environment of the project whereby access and limitations are granted by invitation. <u>Transfer of sensitive data</u> Sensitive data will only be transferred in a depersonalised / deidentified or pseudonymised form.
39 Will the data be safely stored in trusted repositories for long term preservation and curation?	Yes, 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.

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.	There are issues that could impact on data sharing: • Data acquired from third parties, e.g. manufacturers, will not be shared without their explicit consent. • Data collected by the consortium at commercial sites will not be shared without the site owner's explicit consent. • The data from the market surveys will be made anonymous to comply with the General Data Protection Regulation (GDPR). Ethical issues will be addressed as the project will prepare and submit a report on the Dual Use of the project's results.
41 Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?	Only for analysis / use of the data within the project informed consent for data sharing and long-term preservation will be included in the market and customer surveys, but the project has no plans to share data with identifiable personal information. If any sensitive data are collected, they will be separated as soon as possible and kept secure.

1.7 Other issues

Questions	Answers
42 Do you, or will you, make use of other national / funder / sectorial / departmental procedures for data management? If yes, which ones (please list and briefly describe them)?	Research data management is compliant with: • The research data policy of the funding program "European partnerships on metrology" • European laws about data security and the protection of privacy (e.g., GDPR)

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	☒	