

Climateurope2

Preliminary consistent guidelines of good practices for vocabularies, formats, metadata and technical standards, including FAIR principles for data and software

Deliverable 2.1

Authors: Martina Stockhause, Lars Barring, Pierre-Antoine Bretonnière, Douglas Cripe, Amir Delju, Tjerk Krijger, Aleksandra Krzic, Dragan Mihic, Ángel G. Muñoz, Christian Pagé, Alison Pamment, Charlotte Pascoe, Alessandro Spinuso, Peter Thijsse



Funded by
the European Union

Climateurope2

This project has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No 101056933.

Document Information

GRANT AGREEMENT	101056933
PROJECT TITLE	Supporting and standardising climate services in Europe and beyond
PROJECT ACRONYM	Climateurope2
PROJECT START DATE	01/09/2022
RELATED WORK PACKAGE	WP2
RELATED TASK(S)	T2.1
LEAD ORGANISATION	DKRZ
AUTHORS	Martina Stockhause, Lars Barring, Pierre-Antoine Bretonnière, Douglas Cripe, Amir Delju, Tjerk Krijger, Aleksandra Krzic, Dragan Mihic, Ángel G. Muñoz, Christian Pagé, Alison Pamment, Charlotte L. Pascoe, Alessandro Spinuso, Peter Thijssse
SUBMISSION DATE	04/2024
DISSEMINATION LEVEL	public

History

DATE	SUBMITTED BY	REVIEWED BY	VISION (NOTES)
03 Apr 24	Martina Stockhause / Ángel Muñoz		Draft for CE2 internal review
26 Apr 24	Martina Stockhause	Eelco Kruizinga	Submitted version

Please cite this report as: Stockhause, Martina; Barring, Lars; Bretonnière, Pierre-Antoine; Cripe, Douglas; Delju, Amir; Krijger, Tjerk; Krzic, Aleksandra; Mihic, Dragan; Muñoz, Ángel G.; Pagé, Christian; Pamment, Alison; Pascoe, Charlotte L.; Spinuso, Alessandro; Thijsse, Peter (2024). Preliminary consistent guidelines of good practices for vocabularies, formats, metadata and technical standards, including FAIR principles for data and software - Deliverable D2.1 of the Climateurope2 project. Zenodo. <https://doi.org/10.5281/zenodo.10912719>.

Disclaimer: *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 the European Climate, Infrastructure and Environment Executive Agency (CINEA). Neither the European Union nor the granting authority can be held responsible for them.*

Table of Contents

1	Introduction	4
2	State-of-the-art	5
	2.1 Data.....	5
	2.1.1 Metadata	6
	2.1.2 Vocabularies	11
	2.1.3 Formats	15
	2.1.4 Technical Standards supporting data use	17
	2.1.5 FAIR Data Guidelines and Measures	18
	2.1.6 Data Licensing and Citation	24
	2.2 Software.....	25
	2.2.1 Metadata	27
	2.2.2 Technical Standards for Software and Services	27
	2.2.3 FAIR Guidelines and Measures	28
	2.2.4 Software and Services Licensing and Citation	33
	2.3. Organisational Standards.....	35
	2.3.1 Quality Management System (QMS).....	35
	2.3.2 Avio - The International Civil Aviation Organization (ICAO) standards.....	39
	2.3.3 Laboratory Standards.....	39
	2.3.4 The role of Education.....	41
	2.3.5 Trustworthy Digital Repository Framework.....	42
3	User Engagement in Developing Standards	44
	3.1 CF Community Engagement.....	44
	3.2 CMIP Community Engagement.....	44
4	Gaps, Challenges and Opportunities	46
	4.1 Sectorial Standardisation Approaches.....	46
	4.2 Identified Gaps and Challenges.....	46
5	Conclusion	47
	References	48

About Climateurope2

Timely delivery and effective use of climate information is fundamental for a green recovery and a resilient, climate neutral Europe, in response to climate change and variability. Climate services address this through the provision of climate information for use in decision-making to manage risks and realise opportunities.

The market and needs for climate information has seen impressive progress in recent years and is expected to grow in the foreseeable future. However, the communities involved in the development and provision of climate services are often unaware of each other and lack interdisciplinary and transdisciplinary knowledge. In addition, quality assurance, relevant standards, and other forms of assurance (such as guidelines, and good practices) for climate services are lagging behind. These are needed to ensure the saliency, credibility, legitimacy, and authoritativeness of climate services, and build two-way trust between supply and demand.

Climateurope2 aims to develop future equitable and quality-assured climate services to all sectors of society by:

- Developing standardisation procedures for climate services
- Supporting an equitable European climate services community
- Enhancing the uptake of quality-assured climate services to support adaptation and mitigation to climate change and variability

The project will identify the support and standardisation needs of climate services, including criteria for certification and labelling, as well as the user-driven criteria needed to support climate action. This information will be used to propose a taxonomy of climate services, suggest community-based good practices and guidelines, and propose standards where possible. A large variety of activities to support the communities involved in European climate services will also be organised.

Executive Summary

This deliverable collates and discusses the state-of-the-art data management practices regarding vocabularies, formats, metadata and technical standards –including Findable, Accessible, Interoperable and Reusable (FAIR) principles– for processes related to climate services data and software. The analysis largely draws from existing literature and reports, but also from the experience of the co-authors.

The document is organised in three main parts, discussing the current state-of-the-art and common practices regarding data, software and organisational standards (Section 2); the engagement of users in developing standards, guidance and protocols (Section 3); and missing elements, gaps, challenges and opportunities (Section 4).

In summary, this report has found that:

1. User engagement is essential for the development of climate services that meet the evolving needs (demand) of different sectors. Data and software processes in the climate services realm that are supply driven, do not tend to be fit for purpose, and hence tend to lack key desired attributes by stakeholders and decision makers, such as the actual required quality and reliability needed (although the services might be reliable and have quality in other senses).
2. FAIR principles for data and software have been and are widely used by the community involved in the production of data and software for climate services, but these principles have

become guiding ones also outside said community, being complemented by implementation guidelines emphasising the roles of Persistent IDentifiers (PIDs) and metadata standards (see Section 2 for details).

3. WMO Standards and Guiding Documents are usually the *de facto* protocols used by the climate service community, providing guidelines on many different dimensions of the services. These protocols are complemented by developments of good practices within the climate research community, such as those proposed by the World Climate Research Program (WCRP; in particular the Coupled Model Intercomparison Project, or CMIP) and by EU regulations such as the INSPIRE directive (infrastructure for Spatial Information In Europe).
4. Quality management standards and institutional commitments are important for reliable and quality-assured climate service data and software co-development, production, delivery and use.

Further analysis and input from climate service providers is required to get an overview how these different data management practice components are combined by individual (domain-specific) providers. With this prerequisite common sets of consistent guidelines for a variety of climate service providers can be derived in the final deliverable of this task 2.1.

Keywords

Climate service, guideline, data, metadata, standard, best practice, interoperability, sustainability, climate change, vocabulary

1 Introduction

The current number and diversity of climate services in Europe - and worldwide- require a consistent set of guidelines on good practices to support and ensure the quality, functionality, compatibility and interoperability of processes and tools. Therefore, the aim of this work is to compile guidelines that align various sectoral standards and common practices to enhance consistency, interoperability and sustainability of climate services. As part of the Climateurope2 project, Work Package 2 has been gathering and reviewing best practices for vocabularies, formats, metadata and technical standards for data and software processing as well as more formal organisational standards. As applied practices of climate service providers are diverse, this document focuses on those data management practices with a fairly high uptake across communities and discusses user engagement towards developing standards in a dedicated section. These preliminary findings will be updated with feedback and additional information gathered from the climate services community for the Deliverable D2.6. This report is divided in three main sections, regarding (1) the state of the art on data, software and organisational standards, (2) the user engagement in developing standards, and (3) the preliminary identification of gaps, challenges and opportunities in the sectors.

2 State-of-the-art

2.1 Data

With the increasing demand for high-quality data underpinning climate services, to aid for example disaster risk reduction and climate-change adaptation and mitigation, it is important to ensure that a robust regulatory framework that defines standards and recommended practices and procedures for data management is established.

The World Meteorological Organization (WMO) Manual on the High-quality Global Data Management Framework for Climate (WMO-1238, 2019) sets out the high-level standards and recommended practices for data stewardship that will ensure that data for climate purposes are reliable, accessible, sustainable, and as far as possible complete. Note that these practices are mainly concerned with historic and real-time data collected, from a variety of observational platforms, such as remotely-sensed, in-situ, marine, and aircraft. Guidelines with respect to climate modelling and projections of future climate, while similar in many aspects, merit a separate discussion which follows in the **Metadata for climate model data** in Section 2.1.1 below.

The Manual provides definitions of commonly encountered terms with regard to climate data, as well as general requirements for the management of data, along with more specific considerations of data from different sources. It also provides details on the required standards and recommended practices for managing climate data, and also introduces the concept of dataset maturity assessment as an objective tool for evaluating the level of compliance with these standards and recommended practices. In addition, although interoperability is not addressed specifically, the structure provided by the Manual below enhances interoperability as a by-product of encouraging a common approach to metadata provision and data management.

The provisions in the Manual are based on universally agreed data management goals and relate essentially to the following fundamental aspects:

1. **Accessibility:** In the context of data management, “accessibility” refers to ensuring that the dataset is easily and conveniently downloadable by users;
2. **Data integrity:** Ensuring that data are recorded, preserved and are free from corruption or loss when transferred between systems or in storage throughout the data life cycle;
3. **Data quality control:** Using community best practices, including ensuring that the results of quality control are also documented;
4. **Discoverability:** Ensuring that data and relevant information about the dataset may readily be found, including visibility in online catalogues;
5. **Documentation:** Including all elements necessary to access, provide guidance for users, understand, and process the data, and including documentation on replication, reprocessing and updates;
6. **Governance:** Ensuring that accountability, responsibility and compliance mechanisms are well-defined and transparent;
7. **Metadata:** Ensuring that information about data is publicly available, including full details of the origin and processing history of raw observations and derived products (“provenance metadata”), to ensure full traceability of the processing chain;
8. **Portability:** Ensuring that data are easily incorporated into users' working environment based on community standards;
9. **Preservation:** Ensuring that data and metadata are protected from loss and preserved for future use in line with well-specified retention policies;
10. **Quality assessment of datasets:** Ensuring that datasets and data products are routinely reviewed, and the results of the review are transparent;

11. **Uncertainty:** Ensuring that uncertainty estimates are documented and made available (the quantification and communication of climate uncertainty is discussed in Climateurope2 deliverable 2.3 (Pascoe et al., 2024).;
12. **Usability and usage:** “Usability” describes how easily the data product may be understood and used by users and incorporated into a user’s own working environment. It incorporates aspects of compatibility of the publication medium with community standards and supporting documentation. “Usage” refers to the degree of scientific credibility of the data among users, including through citations in peer-reviewed literature.

Several attempts have been made to define the application of such high-level goals, including those associated with data management principles of the Group on Earth Observation System of Systems (GEOSS). The following subsections offer additional guidelines for ways in which these goals may be met in practice.

2.1.1 Metadata

According to ISO/IEC 2382:2015(en) “Information technology — Vocabulary” the term **metadata** is defined as

data about data or data elements, possibly including their data descriptions, and data about data ownership, access paths, access rights and data volatility.

That is, metadata is crucial for achieving the mentioned high-level data management goals, being information processing and dissemination related to climate services to ensure data integrity, usability, and accessibility.

Based on a survey of documentation from the World Meteorological Organization (WMO), the Global Framework for Climate Services (GFCS), the Open Geospatial Consortium (OGC), the Group on Earth Observations (GEO), and other actors in the metadata sphere for climate services, best practices for climate services metadata in connection with data can be summarised as follows:

1. **Standardised Metadata Formats:** Use established metadata standards such as WMO Core Metadata Profile (WCMP) 2019 for weather and climate data, or ISO 19115-1 or DataCite Metadata Schema 4.4 (2021) for datasets more generally. These standards provide a structured way to document information.
2. **Clear and Comprehensive Descriptions:** Provide detailed information about the dataset. Include titles, abstracts, keywords, and descriptions that explain the purpose, content, and context of the data. This helps users understand what the data is about.
3. **Data Provenance:** Document the data source, including the instruments used, data collection methods, and organisations responsible for data collection. This helps establish data credibility and traceability (see Deliverable 2.2 on provenance for more details).
4. **Temporal and Spatial Information:** Clearly specify the temporal and spatial coverage of the data. This should include start and end times, as well as geographical coordinates, resolution, and projection information.
5. **Data Quality and Uncertainty:** Include information on data quality, accuracy, precision, and any known uncertainties. This helps users assess the reliability of the data.
6. **File Naming Conventions:** Use consistent and informative file naming conventions for data files. Include relevant timestamps, variables, and other metadata in the file names.

7. **Version Control:** Maintain version control for datasets. Document changes, updates, and revisions to ensure users are aware of the dataset's history.
8. **Units and Parameters:** Clearly define the units of measurement for variables and parameters in the dataset. Include standard abbreviations and conversion factors where applicable.
9. **Data Access and Distribution:** Specify how users can access the data, including download links, APIs, or other methods. Ensure that data are available in open and machine-readable formats.
10. **License and Usage Terms:** Clearly state the terms of use, licensing, and any restrictions on data usage. Make it explicit if the data are open access or require permissions (see [section 2.1.6 on licences and citation](#)).
11. **Metadata Cross-Referencing:** Link related datasets, publications, and documentation. This helps users find additional information and resources.
12. **Data Citation:** Encourage data citation by providing a standardised citation format. This helps acknowledge the data source and promotes proper attribution (see [section 2.1.6 on licences and citation](#)).
13. **Data Visualization:** Include visual representations of data coverage, quality, and relevant graphs, maps, or images that help users quickly understand the dataset's context.
14. **Metadata Updates:** Ensure that metadata is regularly updated to reflect changes in the dataset, project, or data source.
15. **Metadata Accessibility:** Make metadata discoverable and accessible through metadata catalogues, data repositories, or data portals. Use standardised metadata search and discovery mechanisms.
16. **User Support:** Provide contact information for users who have questions or need assistance with the data. Offer user support and FAQs when possible.
17. **Interoperability:** Ensure that metadata conform to international standards and can be easily integrated with other climate data sources and tools, including machine interpretable or machine-to-machine exchange.
18. **Training and Documentation:** Offer documentation, tutorials, and training materials to help users understand and use the data effectively (see [section 2.3 on organisational standards](#) for details on training).
19. **Engage Stakeholders:** Involve stakeholders, including the scientific community and end-users, in the development of metadata standards and requirements to ensure they meet user needs (see [section 3 for on user engagement](#)).
20. **Review and Quality Assurance:** Implement a review and quality assurance process for metadata to ensure consistency and accuracy.

Metadata standards for observational as well as model climate data have been developed by the dedicated authorities, which are introduced in the following two subsections.

Metadata for observational climate data

A comprehensive metadata standard covering all types of weather and climate observations is by nature complex to define. A user should be able to use metadata to identify the conditions under which the observation, or measurement, was made, and any aspects that may affect its use or understanding, that is, to determine whether the observations are fit for the purpose. As an example, National Meteorological and Hydrological Services (NMHSs) affiliated with the WMO establish, maintain and update explanatory metadata and ensure data homogeneity. Data homogeneity is very important for the calculation of climate normals and averages. An adequate set of metadata must be available to inform future users about the nature of the data in the system, how the various datasets were collected and any inherent problems.

The WMO Integrated Global Observing System (WIGOS) observations consist of an exceedingly wide range of data, from manual observations to complex combinations of satellite hyper-spectral frequency bands, measured in situ or remotely, from single dimension to multiple dimensions, and those involving processing. The WIGOS metadata describe the observed variable, the conditions under which it was observed, how it was measured or classified, and how the data have been processed, in order to provide users with confidence that the data are appropriate for their application. In the Manual on the WMO Integrated Global Observing System (WMO-No. 1160), Appendix 2.1, the observing network design principle 10 refers to the need of “Providing information so that the observations can be interpreted” (metadata), while in Appendix 2.2, the Global Climate Observing System (GCOS) Climate Monitoring Principle 2.2.1(c) describes the relevance of metadata as follows: “The details and history of local conditions, instruments, operating procedures, data-processing algorithms and other factors pertinent to interpreting data (that is, metadata) should be documented and treated with the same care as the data themselves.” The WIGOS weather and climate metadata standards are thus presented here as an example of best practices for observational data.

Ten categories of WIGOS metadata have been identified. These are listed in Table 1 below. They define the WIGOS Metadata Standard, each category consisting of one or more metadata elements. All of the categories listed are considered to be important for the documentation and interpretation of observations made, and even for their use in the distant future. Hence, the standard currently declares many elements that are clearly not needed for applications focusing on more immediate use of observations. For these applications, such as numerical weather prediction, aeronautical, transport sector applications and advisories, or other climate services, profiles of the standard may be developed. The categories are in no particular order but reflect the need to specify the observed variable; to answer why, where and how the observation was made; how the raw data were processed; and what the quality of the observation is.

Table 1. WIGOS metadata categories

Category	Description
1. Observed variable	Specifies the basic characteristics of the observed variable and the resulting datasets. It includes an element describing the spatial representativeness of the observation as well as the biogeophysical compartment the observation describes.
2. Purpose of observation	Specifies the main application area(s) of the observation and the observing programme(s) and networks the observation is affiliated to.
3. Station/platform	Specifies the observing facility, including fixed station, moving equipment or remote-sensing platform, at which the observation is made.
4. Environment	Describes the geographical environment within which the observation is made. It also provides an unstructured element for additional meta-information that is considered relevant for adequate use of the observations and that is not captured anywhere else in this standard.
5. Instruments and methods of observation	Specifies the method of observation and describes characteristics of the instrument(s) used to make the observation. If multiple instruments are used to generate the observation, then this category should be repeated.
6. Sampling	Specifies how sampling and/or analysis are used to derive the reported observation or how a specimen is collected.
7. Data processing and reporting	Specifies how raw data are transferred into the observed variables and reported to the users.
8. Data quality	Specifies the data quality and traceability of the observation.
9. Ownership and data policy	Specifies who is responsible for the observation and owns it.
10. Contact	Specifies where information about the observation or dataset can be obtained.

The quality of this information is essential for ensuring sound and consistent management in any technical programme or process that leads to producing outputs and delivering them to the user community. The WMO Guidelines on Quality Management in Climate Services (No. 1221, 2018) elaborates the ISO 9001 framework in the context of the climate services value chain. Accurate and curated metadata plays an integral role in a Quality Management System (QMS), an end-to-end system covering all activities from raw measurements and observations to services delivered to end users. It seeks to improve quality and performance so that customer expectations can be met or exceeded, taking into account the NMHS context as well as interested party expectations and requirements. It is an important part of the climatological practices of NMHSs and plays a key role in driving quality through the whole value chain, from the selection and installation of instrumentation, storage and quality control of data to the production of climatological products and services. QMS and ISO 9001 are discussed in detail in section 2.3.

Metadata for climate model data

Climate models and earth system models are the essential tools for producing projections of future climate pathways, which is an essential source for information for climate services. To a large extent the overarching requirements on metadata for climate model data follow those of climate data from observational and remote sensing sources. The twelve general points outlined in *Section 2.1* all apply to climate model data, but there are significant differences in relation to what is described in the preceding sections on WIGOS data. Data homogeneity is not an issue for climate model data, at least not in the same sense as for observational data. Descriptions of instrument and measurement technique and environment are replaced by description of the modelling system, its configuration and performance, which is a complex and far-reaching requirement. Basically, it falls into two categories: (i) technical description of the modelling system and its output, and (ii) the model's performance and related uncertainties.

For the first part several well established community standards exist (Table 2) that in combination cover applicable parts of the 20 points presented in the *Section 2.1.1*. Together these form a comprehensive set of community standards for documenting climate model data and in particular global and regional climate projections. While the different standards are managed independently there is close collaboration because of substantial overlap and interaction between the communities. Moreover, as these standards are essential vehicles for implementing both CMIP and CORDEX, which are core projects of the World Climate Research Programme (WCRP) these standards fall under the auspices of WCRP.

For the second part, model performance and related uncertainties, the situation is more complex as there is no direct way to assess performance and uncertainties in future projections. In general this is an area of intense research where definitive standards yet have to emerge. However, as there is a pressing need from users and stakeholders to provide such information various best practices have emerged. This is dealt with in a separate ClimateEurope2 deliverable (Pascoe et al., 2024).

Table 2. Key community standards for documenting climate model data and experiments.

Standard name	Focus area	Key Reference
Climate and Forecasting (CF) Conventions	NetCDF metadata conventions and vocabularies for describing climate data, including climate model data.	https://cfconventions.org Eaton et al. (2023)
Attribute Convention for Data Discovery (ACDD)	Discoverability of climate datasets.	https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery_1-3
CMIP Specification	Specification of global climate and earth system model experiments, model output data formats, dissemination and archiving.	https://pcmdi.llnl.gov/CMIP6/ Taylor et al. (2018) Petrie et al. (2021)
CORDEX Specification	Specification of downscaling experiments by regional climate models, model output format, dissemination and archiving.	https://cordex.org/ https://github.com/WCRP-CORDEX
ES-DOC	System and database for documentation of CMIP experiment and participating modelling systems and configurations.	https://es-doc.org/

2.1.2 Vocabularies

Controlled vocabularies can be used to provide semantic information about many aspects of data:

- model, observing platform or instrument used to create or collect the data
- experiment protocols and methodologies used in data collection and processing
- scientific parameter naming
- units of measure

In practice, vocabularies may take a number of forms. A controlled vocabulary is a standardised list of terms used to attach semantic information to data. It does not impose any structure or express any relationships between data entities that are labelled with vocabulary terms. Thesauri are similar to controlled vocabularies but also list synonymous terms so that comparable data can be identified even when labelled differently. According to <https://op.europa.eu/en/web/eu-vocabularies/taxonomies>, ‘A taxonomy is a controlled vocabulary in which all the terms belong to a single hierarchical structure and have parent/child or broader/narrower relationships to other terms. The structure is sometimes referred to as a tree.’ Taxonomies are widely used in the biological and ecological sciences, for example, to classify species and subspecies of organisms. An ontology is a kind of taxonomy, but the types of relationships are greater in number and more specific in their function.

In this section we provide examples of widely used vocabularies relevant to climate data and describe associated technical standards and good practices.

Semantic content with examples for relevant controlled vocabularies

The European GEMET (General Multilingual Environmental Thesaurus) provides a high level vocabulary for many topics, including climate. In addition to providing terms for some specific climate related variables such as albedo and cloud, it lists broader concepts relevant to climate services, for example, [adaptation strategy](#) and [climate experiment](#). Importantly, each of the terms is explained in clear language. When referring to such concepts within the Climateurope2 project, and when communicating the purpose of climate services to a broader audience, it is recommended to use terms from GEMET where possible so as to maintain consistency and aid understanding.

i. Observing platforms and instruments

The World Meteorological Organisation (WMO) maintains a series of manuals, the [Guide to Instruments and Methods of Observation](#), in particular [Volume III – Observing Systems](#), but currently does not publish controlled vocabularies listing the instruments and methods as individual semantic artefacts.

The NASA GCMD (National Aeronautics and Space Administration Global Change Master Directory) is a taxonomy covering many categories related to earth science, including platforms and instruments. Platforms are divided into categories such as “air-based” (e.g. radiosondes, Uncrewed Aerial Vehicles (UAV)), “water-based” (e.g., buoys, ships), “space-based” (a list of satellite missions), “ground-based” (includes meteorological stations) and “other” (e.g., computers, climate models). Instruments are divided into categories for “Earth remote sensing instruments”, “in-situ/laboratory instruments” and “solar/space observing instruments”. The Earth remote sensing instruments are further subdivided into categories which ultimately resolve to the names of individual satellite-borne instruments. The in-situ category contains further subcategories such as temperature/humidity sensors, which again subdivides into further categories and ultimately resolves to instrument names such as “wet bulb thermometer” and “psychrometer” which are relevant to meteorological observations. The NASA GCMD keywords can be browsed using the [GCMD Keyword Viewer](#), downloaded in a variety of formats (RDF, JSON, XML, CSV) or accessed via a [RESTful API service](#).

A comprehensive list of Earth Observation satellites, the instruments on board and other details of their mission such as start and end dates is available at <https://www.eoportal.org/>. The SeaVoX Device Catalogue of marine observing instruments is available from the NERC Vocabulary Server (NVS): <https://vocab.nerc.ac.uk/collection/L22/current/>.

ii. Experiment protocols and methodologies used in data collection and processing

The WMO manuals [Guide to Instruments and Methods of Observation](#) contain very detailed descriptions of observing instruments and the methods of using them. These manuals should be regarded as the reference text when making meteorological observations of any kind, but the various methods of data collection are not currently available as controlled vocabularies.

Metadata to describe the constituent Model Intercomparison Projects (MIPs) that formed phase 6 of the Coupled Model Intercomparison Project (CMIP6) and the design of individual experiments is provided by [ES-DOC](#) (Pascoe et al., 2020). The structure of ES-DOC metadata is based on the Common Information Model (CIM) which describes climate data, the models and software from which they derive, the geographic grids used to calculate and project them, and the experimental processes (typically simulations) that produced them (Lawrence et al., 2012). ES-DOC metadata makes use of controlled vocabularies which provide standardised terms for concepts such as MIP name, experiment name, the name of the institution producing the data, and model component (e.g. land-ice scheme, aerosol scheme) (Moine et al., 2014). The CMIP6 vocabularies are published at https://github.com/WCRP-CMIP/CMIP6_CVs/.

iii. Scientific parameter naming

A commonly used vocabulary for parameter naming in the climate research community is the CF Standard Names: <https://vocab.nerc.ac.uk/collection/P07/current/>. The CF (Climate-Forecast) Conventions, <http://cfconventions.org/conventions.html>, are designed for use with the NetCDF format, although many of the principles could be applied to other data formats. The Standard Names are part of the Conventions and governed by the CF community. New Standard Names can be proposed as needed and undergo a process of community discussion prior to acceptance and publication (see also section 3 on community involvement in standards development). Updates of the standard name table are published periodically, the version number being incremented with each release. Version 84 is the latest at the time of writing and contains terms for 4710 distinct variables. The increasing sophistication of climate models and earth system models has driven much of the vocabulary development, for example, many ocean biogeochemistry and sea ice dynamics terms were introduced in 2018 in response to the data requirements of the CMIP6 project. Over time, CF Standard Names have also been adopted beyond the climate modelling community and now include terms applicable to satellite products, radar remote sensing, ocean gliders and other in situ observations.

Another widely used vocabulary in environmental science is the BODC (British Oceanographic Data Centre) Parameter Usage Vocabulary: <http://vocab.nerc.ac.uk/collection/P01/current/>. The CF Standard Names (P07) and P01 vocabularies have some overlapping content and are sometimes used alongside one another, for example, both are required in the metadata for data shared through SeaDataNet (Pan-European infrastructure for ocean and marine data management) (<https://www.seadatanet.org/Publications/Technical-documentation/Metadata>).

NASA GCMD keywords are parameter names for many areas of earth science, including the atmosphere. The full list of available terms can be viewed in a browser at: https://gcmd.earthdata.nasa.gov/KeywordViewer/scheme/all?gtm_scheme=all. The keywords are arranged in a hierarchy, for example, EARTH SCIENCE > ATMOSPHERE > CLOUDS > CLOUD PROPERTIES > CLOUD BASE HEIGHT would be the full path to the cloud base height parameter name. The same term can also be accessed in [RDF](#) or [JSON](#) formats.

The WMO GRIB message format (2.1.3) identifies scientific parameters not by name, but using a code number. GRIB messages are used to share operational weather forecasts globally. Modelling centres may define their own parameters using GRIB local tables, but the contents of these are not standardised. The table of globally standardised GRIB codes is available at <https://codes.wmo.int/grib2/codeflag/4.2>. Individual parameter codes may be accessed by a unique URI, for example <http://codes.wmo.int/grib2/codeflag/4.2/0-1-18> is the URI for “absolute humidity”.

Specialised vocabularies can be used in combination with the more general environmental ones, for example, to ensure standard nomenclature is used when referring to chemical and biological species. As an example of this, CF standard names refer to two biological taxonomies, WoRMS (World Register of Marine Species) (<https://www.marinespecies.org/>) for ocean species and ITIS (Integrated Taxonomic Information System) (<https://www.usgs.gov/tools/integrated-taxonomic-information-system-itis>) for terrestrial and freshwater species. IUPAC chemical species names are used either in the CF standard name itself or referred to in the standard name description.

The RDA (Research Data Alliance) I-ADOPT (InterOperability of Observable Property Terminology) Working Group conducted a survey in 2020 for the purpose of compiling a list of controlled vocabularies, taxonomies, thesauri and ontologies, collectively referred to as terminologies, used in the environmental sciences to provide parameter descriptions. The results of the survey (approximately a hundred terminologies) are published as a searchable catalogue on GitHub <https://i-adopt.github.io/terminologies/list/all/>.

The I-ADOPT Interoperability Framework facilitates mappings and cross-walks between similar terms in different terminologies (Magagna et al., 2022). The framework is being implemented by a number of groups, and has been implemented in the NERC Vocabulary Server (NVS) where it has been used to create mappings between synonymous terms, specifically those relating to ocean chemistry, in the CF Standard Name/P07 and P01 vocabularies. The work of mapping the Standard Names to the I-ADOPT Interoperability Framework is continuing at present.

Metadata ontologies such as the Simple Standard for Sharing Ontological Mappings (SSSOM) (Matentzoglou et al. 2022) provide a potential framework for maintaining a shared understanding between different communities of practice throughout the value chain. SSSOM is analogous to the I-ADOPT interoperability framework, but is also able to explicitly capture the imprecision, inaccuracy and incompleteness of mapped concepts. The effective passing of information between different communities of practice has been identified as a requirement by the Climateurope2 deliverable 2.3 (Pascoe et al., 2024) on the quantification and communication of climate uncertainty.

iv. Units of measure

Units vocabularies are used alongside parameter names and are essential for correct interpretation of the quantities in the data. For example, a quantity that is labelled as a “concentration” could be a mass concentration, molar concentration or volume concentration. Some parameter vocabularies, such as CF Standard Names, distinguish between such quantities while others use the more generic label of “concentration” and the scientific variable can only be understood by also referring to the units.

Within the CF Conventions units are required to be compatible with the UDUNITS software package <https://doi.org/10.5065/D6KD1WNO> which contains its own units database, mostly consisting of SI units and derived units. The NVS serves a unit collection, the BODC-approved (British Oceanographic Data Centre) data storage units <http://vocab.nerc.ac.uk/collection/P06/current/>, to describe the measurement units for data held in its repositories. The NVS representation of the CF Standard Names uses mappings between individual terms in the P07 and P06 collections to express the association between each standard name and its appropriate unit.

UCUM (Unified Code for Units of Measure) <https://ucum.org/> is ‘a code system intended to include all units of measures being contemporarily used in international science, engineering, and business.’ It imposes a strict syntax on all units, combinations of units and symbols. It is primarily intended for unambiguous electronic communication of quantities and units.

QUDT (Quantities, Quantity kinds, Units, Dimensions, and data Types) <https://doi.org/10.25504/FAIRsharing.d3pgw7> provides a collection of ontologies and vocabularies based on SI units of measure and also includes CGS units for mechanics, CGS EMU (electromagnetic) units, CGS-ESU (electrostatic) units, Gaussian units for electrodynamics, and the Planck system of natural units. Many US Customary and British Imperial units for length, weight, and heat are also included. The aim of the QUDT ontologies is to improve interoperability of data by providing unambiguous and machine readable units specifications and supporting conversion between equivalent units and quantities. The QUDT units vocabulary is available at <https://doi.org/10.25504/FAIRsharing.d3pgw7>. It provides multilingual support for unit names and is available in a variety of formats (CSV, RDF/XML, TURTLE, JSON, JSON-LD).

Technical standards for semantic services and the use of controlled vocabularies

Semantic services are technologies that serve semantic content (such as vocabularies and ontologies) in a variety of human and machine readable formats to support semantic interoperability.

Examples of semantic service providers include the NERC Vocabulary Server (NVS), the ICES (International Council for the Exploration of the Sea) vocabulary server <https://vocab.ices.dk/> and the ARDC (Australian Research Data Commons) vocabulary server <https://vocabs.ardc.edu.au/>. These services are useful for discovering existing vocabularies thus helping to avoid unnecessary duplication of terminology and the barrier to interoperability that this can cause. Tools can be built to utilise the vocabulary services to aid in data discovery, for example, by returning search results for data labelled with synonymous terms drawn from different vocabularies.

There are multiple advantages to publishing via vocabulary servers, including the implementation of versioning and tracking vocabulary changes such as the addition or deprecation of individual terms. One of the most powerful features is the ability to assign a PID (Persistent IDentifier) to each vocabulary list and to individual terms, allowing them to be referenced unambiguously by both humans and machines. Use of a SPARQL endpoint allows vocabularies to be presented in multiple ways, conforming to different metadata profiles and machinable formats. Recommendations for FAIR semantics are detailed in Le Franc et al. (2022).

A widely used data model for encoding semantic information is SKOS (Simple Knowledge Organization System) <https://www.w3.org/2001/sw/wiki/SKOS>. A complete vocabulary is called a SKOS *collection*, while an individual term within a vocabulary is a SKOS *concept*. The schema is relatively simple - the *skos:preflabel* is used to specify the term itself, while an abbreviation or other alternative can be specified in the *skos:altlabel*. Relationships between pairs of terms within or across vocabularies are represented by mappings which can express that the terms are synonymous, broader/narrower, or associated (related in some other way). For example, in the NVS a CF standard name in the P07 collection is associated with the appropriate unit in the P06 collection using a 'related' mapping.

Schema.org (<https://schema.org/>) was developed by Google, Microsoft, Yahoo and Yandex to create structured data models for a broad range of concepts or *Things*, for example, there are schemas to describe an organisation, a person or a product. Each schema lists *properties* appropriate to describing the *Thing*, for example, an organisation may have a postal address, an email address and a number of employees. An employee can in turn be described using the person schema, creating a series of linked data. Schema.org is maintained by a [community](#) process. Biochemical entity and taxon schemas have been added relatively recently.

Available evaluation tool/method for compliance

The CF metadata conventions are designed to work with the NetCDF format. CF compliant NetCDF files can be checked for standards compliance using the CF-checker software <https://cfchecker.ncas.ac.uk/>. The CF-checker can also be installed as a command line utility. Checks include determining whether CF standard names used in data files are consistent with the latest published version of the table.

2.1.3 Formats

A comprehensive collection of data and file formats used for meteorological and climatological data are maintained by WMO in <https://codes.wmo.int/wmdr/DataFormat>. Once the data has been processed and enriched to produce more user-oriented climate data and information other data formats may become more relevant either to interface with geographical information systems (GIS) and/or graphical data formats.

“Raw” climate data, i.e. data close to the source are often of huge volumes and/or are frequently multi-dimensional, which either are impossible or impractical to handle using standard generic desktop

software tools like Office suites or text editors. Hence a range of different data (file) formats have been developed over time. The state-of-the-art is represented by

- **BUFR** (Binary Universal Form for the Representation of meteorological data) is a highly compressed binary format defined and maintained by WMO for a wide range of meteorological data. It was first created in 1988, and the current GRIB-4 version is documented in WMO (2022). BUFR is a table-driven format, which means that the information in the files are coded as compact references to a set of tables that are maintained separately, i.e. in the documentation.
- **GRIB** (GRIdded Binary, or General Regularly-distributed Information in Binary form) is a WMO standard format that share its overall conceptual design with BUFR, but is targets gridded (i.e. raster) data from numerical weather prediction models are similar sources. It is the internal format used by ECMWF for all its reanalysis products. The current version is GRIB-2 documented in WMO (2022). The previous version GRIB-1 was discontinued by WMO in 2016, but the documentation is still available as reference for archived data (WMO, 2016). Similar to BUFR the data format relies on a series of external tables, of which some are defined in the documents. For GRIB-1 local table extensions were commonly defined by individual organisations and for specific projects, which due to insufficient and/or only locally available documentation lead to complications and was a source for mistakes and errors.

As WMO standards both the BUFR and GRIB formats were specifically developed for operational exchange of weather data. As both formats essentially depend on external tables which recently has become available through the WMO Codes Registry (<https://ci.codes.wmo.int/>) with some documentation available from <https://ci.codes.wmo.int/ui/help>. As binary formats they are intended for machine-to-machine communication. Applications outside the sphere of national meteorological services include for example data dissemination of weather forecast data to news and TV media, as well as to web services.

- **NetCDF** (Network Common Data Form) is a format that essentially is defined by the reference software library developed and maintained by University Corporation for Atmospheric Research (UCAR), Boulder, Colorado. It is a collection of several closely related machine-independent binary data formats. The formats target multi-dimensional array-oriented datasets, which from a user perspective mainly differ in maximum allowed file size and data compression methods they support. The recent version builds on the HDF5 data format (see below). The netCDF formats are self-describing, which means that the data files contain essential metadata rather than relying on external sources. In combination with the Climate and Forecasting CF conventions the netCDF format has become the *de facto* standard format for exchange and dissemination of climate model data. The documentation is hosted by the UCAR Unidata program at <https://www.unidata.ucar.edu/software/netcdf/>. NetCDF is adopted by OGC through a suite of standards (cf. <https://www.ogc.org/standard/netcdf/>).
- **HDF** (Hierarchical Data Format) is a self-describing machine-independent general purpose binary data format for storing numeric data. There are two main versions: HDF4 and HDF5, which are closely related but not fully compatible. Both versions are actively maintained by the producer <https://www.hdfgroup.org/>. As a binary data format it includes a range of data compression mechanisms that allows for efficient storage of data, which allow for handling huge data volumes of e.g. multi-dimensional arrays, single and stacks of raster images, as well as tables. Except for what is needed to describe the internal data structures as such there are no specific requirements on metadata, which instead have to be organised by user communities. Such communities exist for various sectors, for example the climate data community make use of HDF via the netCDF format, and the earth observation and remote sensing community make use of HDF-EOS. HDF5 is included as OGC Core standard OGC 18-043r3 (cf. <https://www.ogc.org/standard/hdf5/>).

- **Zarr** is a recent data format specifically targeted at storing huge multidimensional numeric datasets on cloud platforms. It is a community effort, cf. <https://zarr.dev/>, where the format specification is now at version 3, which is in development. Similar to HDF it is a generic format where metadata standards have to be added on top of the basic Zarr specification. Zarr version 2 is adopted as OGC standard 21-050r1 (<http://www.opengis.net/doc/CS/zarr/2.0> [N.B. file download]), which specifically mentions links to netCDF and CF Conventions for metadata standards. Specific functionality and conventions to handle geolocated data is under development through the recently formed OGC GeoZarr Standards Working Group.

These data formats are mainly intended for handling large datasets produced by climate models, reanalyses, satellites and other remote sensing sources, as well as data intensive field campaigns. They are all included into the OGC palette of geoinformatics standards, which also include a wide range of other standards focusing on data formats or metadata specification, cf. <https://www.ogc.org/standards/> for an overview of available standards.

2.1.4 Technical Standards supporting data use

The metadata and data standards described in the previous subsections are complemented by technical standards supporting data access, sharing, interoperability and processing. Several organisations govern such standards for different technical aspects, relevant for climate services. Important standardisation organisations are the Open Geospatial Consortium (OGC), RDA, ISO and WMO. Additionally, best practices are developed by research infrastructures and projects like the Earth System Grid Federation (ESGF) and the Coupled Model Intercomparison Project (CMIP). Technical standards are discussed separately for metadata and data.

Metadata access and sharing

Common general metadata access interfaces are REST or RESTful APIs (REpresentational State Transfer Application PRogramming Interfaces) and the Open Archives Initiative Protocol for Metadata Harvesting OAI-PMH. GraphQL APIs (<https://graphql.org/>) that support the analysis of relationships in knowledge graphs are becoming increasingly available.

Apart from these general metadata access interfaces, specific interfaces built on existing metadata standards for geographical data are in use. The OGC Web Catalogue Service (WCAS), partly also called Catalogue Service for the Web (CSW), allows the provision and search of metadata about geodata sets and spatial processes.

The INSPIRE (infrastructure for Spatial Information In Europe; <http://inspire.ec.europa.eu/>) Directive of the European Union has defined several standards for metadata and metadata access. The GeoDCAT-AP (<https://semiceu.github.io/GeoDCAT-AP/>) profile of the W3C DCAT application for data portals is used to make spatial datasets and services discoverable on general data portals. OGC services WMS, WFS, WCS and SPARQL endpoints can be specified.

ESGF/CMIP developed best practice project standards with specific information stored in the NetCDF file header and a Data Reference Syntax (DRS), which are indexed and thus made accessible through a RESTful API (https://esgf.github.io/esg-search/ESGF_Search_RESTful_API.html) for metadata access and data discovery or for use in compute services.

Data access, subsetting and processing

Depending on the metadata standards, aligned data access interfaces are in use.

OGC services used in connection with CSW include the Web Map Service (WMS) for access of image files, Web Feature Service (WFS) for spatial vector data and Web Coverage Service (WCS) for spatial

raster data. INSPIRE uses these OGC standards. The OGC Web Processing Service (WPS) allows for subsetting (selecting parts) and processing of the accessed data.

The OGC Climate Resilience Domain Working Group (DWG) functions as a climate domain group within OGC and is therefore of relevance for climate services. For example, the Climate Resilience DWG coordinates a climate resilience pilot (<https://www.ogc.org/initiatives/crp/>) targeting on improving the interoperability for climate resilience information systems.

ESGF offers as direct download methods HTTP Server data access and Globus and GridFTP file transfers. Additionally, community tools OpenDAP, which enables a server-side data subsetting, and the Live Access Server (LAS) for a graphic-based data subsetting can be offered. Multiple file download is supported by a wget script creation option. More sophisticated data processing options based on the OGC WPS standard are being developed in the ESGF Compute Working Team. The WPS Birdhouse (<https://bird-house.github.io/>) framework includes examples from different organisations and its templates are used in related WPS services like rook (<https://rook-wps.readthedocs.io>).

2.1.5 FAIR Data Guidelines and Measures

2.1.5.1 Introduction to FAIR data

The impact of climate services depends on the trustworthiness of its output datasets including, for example, being able to trace back the full data flow, as well as processing software and its settings, by the end-user. In this Section, we focus first on the data aspect. In order to provide optimised access to, and (re)use of, data it is important that the data are as FAIR as possible. FAIR data means that data is Findable, Accessible, Interoperable and Reusable for humans **and machines**, which can only be achieved if the data is readily available through machine accessible services, supported by complete machine interpretable metadata, and follows international metadata and data standards. This means that best practices for metadata and data will need to be adopted as early as possible in the data life cycle.

A well established data and metadata management system at the source is key to providing access to FAIR data throughout the data life cycle, which implies not only for first use, often national/regional, but also later in re-use via international aggregators. Increased FAIRness can be achieved by developing new data management best practices for climate services data, combining the WMO metadata requirements, with more detailed practices on how to enable machine-to-machine accessibility regarding each of these aspects. Best practices include metadata formats, creation of new vocabularies (increases the I and R), references to services and software and their settings (increases the R, if referred to in the metadata).

In this section we start by explaining what the FAIR metrics are, and then relate that to the current practices for climate data and metadata, to determine how optimised FAIRness could be achieved. This analysis is supported by existing data flow practices in climate services from CE2 partners. In chapter 4 the main challenges and gaps for future work will be summarised.

2.1.5.2 GO-FAIR metrics

A set of 14 principles to determine the level of FAIRness of data has been developed by GO-FAIR (<https://www.go-fair.org/fair-principles/>). Each of the principles contains a reference to recommendations of how they could be met in practice.

The path to FAIR data can be seen as a “movement”, a route that parties involved in a certain domain agree on how to achieve it. Especially for aspects such as Interoperability and Reusability, the path depends a lot on agreed solutions (e.g. vocabularies) in the domain. But of course it helps to have a common target and means to validate how far a data provider is in the process. The FAIR principles themselves can be used as the first generic metrics to evaluate the FAIRness of data from the climate services.

Tools are also available to support the convergence process in a domain, because for each FAIR metric, developers need to make implementation choices (called FAIR Enabling Resources - FERs) and the more agreement there is within the domain on the solution, the more convergence and the more FAIR the data will become. The principles can be used to determine a FAIRness Implementation Plan (<https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/>) via which a community can determine the current state of FAIRness and suggest and agree on implementation choices for improvement. There are also more automatic assessment tools available:

- F-UJI validator <https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>
- FAIR Checker <https://fair-checker.france-bioinformatique.fr/>

For more information on this see <https://www.thehyve.nl/articles/evaluation-fair-data-assessment-tools>.

These tools are still in development, and because their feedback is quantitative (black/white), they are more helpful later in the process once a community has agreed on solutions, implemented them and wishes to revisit certain conditions adopted.

2.1.5.3 Mapping best practices for climate services metadata to GO-FAIR metrics

To support increasing FAIRness of data in the Climateurope2 community, guidelines and measures can be formulated on top of existing best practices in the field of climate services. Based on a review of documentation from the World Meteorological Organization (WMO), the Global Framework for Climate Services (GFCS), the Open Geospatial Consortium (OGC), the Group on Earth Observations (GEO), and other actors in the metadata sphere for climate services, best practices for climate services metadata in connection with data have been summarised in chapter 1.

Using the abovementioned GOFAIR metrics (<https://www.go-fair.org/fair-principles/>) these existing best practices are now grouped per metric, as seen below, in order to find the matches, challenges and gaps. The metrics that could not be mapped to the best practices are listed in the next section on identified gaps.

Findable

F1. (Meta)data are assigned a globally unique and persistent identifier

- File Naming Conventions: Use consistent and informative file naming conventions for data files. Include relevant timestamps, variables, and other metadata in the file names.
 - **The convention is a good starting point, but it is not a globally unique identifier (Digital Object Identifier - DOI) or Persistent Identifier (PID) as such, which might lead to inconsistencies.**
- Data Citation: Encourage data citation by providing a standardised citation format. This helps acknowledge the data source and promotes proper attribution.
 - **Especially for citation, a persistent identifier is much needed.**

F2. Data are described with rich metadata (defined by R1 below)

- Clear and Comprehensive Descriptions: Provide detailed information about the dataset. Include titles, abstracts, keywords, and descriptions that explain the purpose, content, and context of the data. This helps users understand what the data is about.
 - **Important to keep in mind the machine user, keywords should be annotated from common vocabularies as much as possible.**
- Temporal and Spatial Information: Clearly specify the temporal and spatial coverage of the data. This should include start and end times, as well as geographical coordinates, resolution, and projection information.
- Data Quality and Uncertainty: Include information on data quality, accuracy, precision, and any known uncertainties. This helps users assess the reliability of the data.
 - **Also important to keep in mind the machine user, information on quality, accuracy and precision should be annotated via common vocabularies as much as possible.**
- Units and Parameters: Clearly define the units of measurement for variables and parameters in the dataset. Include standard abbreviations and conversion factors where applicable.
 - **Using agreed common code lists/vocabularies as much as possible, e.g. Global Change Master Directory (GCMD), or Climate and Forecast (CF) conventions**
- Review and Quality Assurance: Implement a review and quality assurance process for metadata to ensure consistency and accuracy.

F4. (Meta)data are registered or indexed in a searchable resource

- Metadata Accessibility: Make metadata discoverable and accessible through metadata catalogues, data repositories, or data portals. Use standardised metadata search and discovery mechanisms.

Accessible

A1. (Meta)data are retrievable by their identifier using a standardised communications protocol

- Data Access and Distribution: Specify how users can access the data, including download links, APIs, or other methods.
 - **Important that the metadata contains the PID's of the datafiles it describes, and vice versa.**
- User Support: Provide contact information for users who have questions or need assistance with the data. Offer user support and FAQs when possible.
- Training and Documentation: Offer documentation, tutorials, and training materials to help users understand and use the data effectively.
 - **In case of data access API's document these for developers (e.g. OpenAPI standard/Swagger).**

A1.1. The protocol is open, free, and universally implementable

- Ensure that data are available in open and machine-readable formats.
 - **For FAIRness “free” does not mean that the data should be free of charge and open, but rather the protocol to access it should be open and free and available to all. This means that anyone with a computer and an internet**

connection can access at least the metadata and clearly know how to access the data. It could then also be possible to include restricted data while still remaining 'FAIR'.

Interoperable

I2. (Meta)data use vocabularies that follow FAIR principles

- Interoperability: Ensure that metadata conforms to international standards and can be easily integrated with other climate data sources and tools.
 - **This FAIR metric goes one step further than the syntax (metadata format), but also puts emphasis on agreed international vocabularies to support the semantics in the metadata.**

I3. (Meta)data include qualified references to other (meta)data

- Metadata Cross-Referencing: Link related datasets, publications, and documentation. This helps users find additional information and resources.
 - **This works best if persistent unique identifiers are used for each metadata resource.**

Reusable

R1. (Meta)data are richly described with a plurality of accurate and relevant attributes

- Repeated from F2:
 - Clear and Comprehensive Descriptions: Provide detailed information about the dataset. Include titles, abstracts, keywords, and descriptions that explain the purpose, content, and context of the data, e.g. the experimental protocols. This helps users understand what the data is about.
 - **Important to keep in mind the machine user, keywords should be annotated from common vocabularies as much as possible.**
 - Temporal and Spatial Information: Clearly specify the temporal and spatial coverage of the data. This should include start and end times, as well as geographical coordinates, resolution, and projection information.
 - Data Quality and Uncertainty: Include information on data quality, accuracy, precision, and any known uncertainties. This helps users assess the reliability of the data.
 - **Important to keep in mind the machine user, quality, accuracy, precision and uncertainties, should be described via agreed common vocabularies as much as possible.**
- Plus:
 - Version Control: Maintain version control for datasets. Document changes, updates, and revisions to ensure users are aware of the dataset's history.
 - Metadata Updates: Ensure that metadata is regularly updated to reflect changes in the dataset, project, or data source.
 - **CE2 recommendation on this point is to be stricter: When updating a dataset (reprocessing, or e.g. changing the collection of data), it is recommended to update the related DOI as well to make sure that end-**

users are not citing the wrong version. If there are only changes made to the metadata, for the sake of simplicity the DOI can be maintained.

R1.1. (Meta)data are released with a clear and accessible data usage licence

- License and Usage Terms: Clearly state the terms of use, licensing, and any restrictions on data usage. Make it explicit if the data are open access or require permissions.
 - **Important to keep in mind the machine user, licensing and usage (and e.g. authentication systems) should be specified via common vocabularies as much as possible.**

R1.2. (Meta)data are associated with detailed provenance

- See also the elements under R1.
- Data Provenance: Document the data source, including the instruments used, data collection methods, and organisations responsible for data collection. This helps establish data credibility and traceability.
 - **Important to keep in mind the machine user, instruments, methods (including, where applicable, how the data were combined to achieve a data product or result), organisations, persons, all should be annotated from common vocabularies or agreed directories (e.g. ORCID) as much as possible. Here it is also important to include machine accessible/interpretable quality assessment information, such as the accuracy and/or resolution of the measurement devices and datasets in order to assess the Reusability.**

R1.3. (Meta)data meet domain-relevant community standards

- Standardised Metadata Formats: Use established metadata standards such as WMO Core Metadata Profile (WCMP) 2019 for weather and climate data, or ISO 19115-1 or DataCite Metadata Schema 4.4 (2021) for datasets more generally. These standards provide a structured way to document information.
 - **Community standards listed are a good starting point, especially when including community vocabularies. Metadata and data formats could be reviewed for machine actionability, and extended where needed (see gaps/challenges).**
- Engage Stakeholders: Involve stakeholders, including the scientific community and end-users, in the development of metadata standards and requirements to ensure they meet user needs.
 - **The Climateurope2 community could play an important role here.**

2.1.5.4 Identified Gaps

Several of the FAIR metrics could not be mapped (yet) to WMO requirements. They are listed below.

F3. Metadata clearly and explicitly include the identifier of the data they describe

In the case that the dataset and the accompanying metadata are separate files, the association between the two should be made explicit by mentioning the dataset's globally unique and persistent identifier in the metadata.

A1.2. The protocol allows for an authentication and authorisation procedure, where necessary
The 'A' in FAIR does not necessarily mean 'open' or 'free', but rather that the exact conditions should be clear under which the data is accessible, meaning that even restricted data can be FAIR. Preferably, the accessibility is specified such that a machine can understand the requirements, and then either automatically execute the requirements or alert the user to the requirements. By having repository users create an account, the owner (or contributor) of each dataset can be authenticated and user-specific rights can be set-up.

A2. Metadata are accessible, even when the data are no longer available

Principle A2 states that metadata should persist even when the data are no longer sustained. Datasets tend to degrade or disappear over time because there is a cost to maintaining an online presence for data resources. When this happens, links become invalid and users waste time hunting for data (e.g. which is used in a climate service product) that might no longer be there. Storing the metadata generally is much easier and cheaper. Even if the original data are missing, tracking down people, institutions or publications associated with the original research can be extremely useful.

I1. (Meta)data uses a formal, accessible, shared, and broadly applicable language for knowledge representation

Humans and **machines** should be able to exchange and interpret each other's data. The main goal of this principle is to provide a "common understanding" of digital objects by means of a language for knowledge representation to be used to represent these objects. The chosen language (RDF, OWL, JSON LD, etc.) should have a formal specification and the specifications should be shared and accessible so others can learn the language.

→ **NetCDF for the dataset is a good fit. For the metadata XML is not ideal for machines, RDF is a better alternative. Similarly access to vocabularies and other metadata resources could benefit from increased machine accessibility via RDF, OWL or JSON LD.**

2.1.5.5 Outlook on next steps for FAIR Guidelines and Measures

In order to ensure the right level of trust in climate services output datasets it is important that the data is published in accordance with FAIR principles. This means that not only the output itself is FAIR, but also that the underlying dataflow is FAIR. As we have seen in previous sections this puts additional emphasis on the data management of climate data. The WMO guidelines and existing metadata and data standards are an excellent starting point, but when it comes to FAIRness the guidelines could be "tightened", and we can conclude additional guidance in the form of Best Practices are needed.

Some first directions that need further attention when it comes to Best Practices for climate service data management:

- FAIR is focused on machine-2-machine interaction to the data and metadata. This puts additional focus on data management, e.g. using as little as possible free text in metadata and data, but instead using vocabularies for these attributes. Not only for parameters, units, but also for data access restrictions, usage policy, quality information, organisations, persons, etc.
- To achieve optimum interoperability the vocabularies used should be community supported and FAIR by itself in order to be widely used and accessible by machines (e.g. to make mappings, request ontology).

- The best practices for File Naming Conventions mentions the use of consistent and informative file naming conventions for data files, including the relevant timestamps, variables and other metadata in the file names. Looking at the GO-FAIR metrics, this relates and differs to what Metric F1 states, which is that (Meta)data should be assigned a globally unique and persistent identifier. → **to be further investigated and discussed**
- Processing services involved should be clearly indicated in the metadata with a link to the service and version, as well as to documentation (OpenAPI or software) where possible. See chapter 3.3.

A WP2 working group could be started to discuss further steps in this process and document the additional recommendations and Best Practices for the community.

2.1.6 Data Licensing and Citation

Data licensing establishes the terms and conditions under which data can be accessed, utilised and shared by granting legal permission. Licences are assigned by the data owner or copyright holder. An overview over existing licences with their legal conditions is available at <https://spdx.org/licenses/>. Open licences allow the reuse of the data by all stakeholders. Open data licensing is encouraged by the Open Data and FAIR data principles aims. Imposed restrictions limit the reuse of the data and the creation of derivative products. As climate services often use data from multiple copyright holders, the handling of these imposed restrictions, which might be incompatible, might curtail the publication of derivative products.

Examples for data use restrictions are non-commercial, not allowing any commercial use, and non-derivative, not allowing to share data derivatives. However even the share share-alike attribution of data disseminated under an open licence might restrict commercial reuse, as data derivatives have to be distributed under the same open licence.

Climate service providers have to comply with the data licences and all the restrictions imposed by the copyright holder of the data it uses when distributing its products. IPCC TG-Data recommends to licence data derivatives under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence, where this does not infringe the interests of relevant licence holders (Huard et al., 2022) in order to ensure maximal reusability of the data. This open licence CC BY 4.0 permits the reuse of data under the condition that appropriate credit is given, which includes credit to the data creators (see <https://creativecommons.org/licenses/by/4.0>).

In science credit is commonly given by citation. Data citation principles were formulated by FORCE11 in the Joint Declaration of Data Citation Principles (Data Citation Synthesis Group, 2014). The international membership organisation DataCite provides the infrastructure and services for the citation of research data and further scholarly records and the registration of standardised metadata. DataCite mints DOIs for their members as registration agency of the International DOI Foundation. DataCite was founded in 2009 as an international organisation based on the experience and pilot implementation of a project at the German National Library of Science and Technology (TIB) that was funded by the German Research Foundation (DFG) in 2003. The first DOI in DataCite's catalogue was minted in 2004 (see Brase, 2019).

Most data providers have implemented DOIs on data and have included a cite as statement on their data pages, which climate service provider should comply with when citing the used data for credit assignment. In addition the modification of the data should be indicated, which also supports the transparency and traceability of the climate service product. Products of climate services should be assigned a DOI to make them referenceable and citable for climate service customers. Organisations providing best practices for data citation include apart from DataCite and FORCE11, the [Data](#)

[Curation Centre \(DCC\)](#), the [Earth Science Information Partners \(ESIP\)](#) and the [Coalition for Publishing Data in the Earth and Space Sciences \(COPDESS\)](#).

2.2 Software

The pathway from data to the results of an analysis requires the use of software tools and code (figure 2.1.1). Therefore for science to be truly open and transparent, the code used to process and analyse data should also be open and transparent. Fortunately, good software development practices that make code easier to maintain, extend and collaborate also facilitate this transparency. At the most basic level software should be accessible¹ and well documented² for reuse and reproducibility of results. Software plays a crucial role as the vehicle for provision of climate services, enabling the collection, analysis, and dissemination of climate-related data and information. Best practices for climate services metadata in connection with software can be summarised as follows:

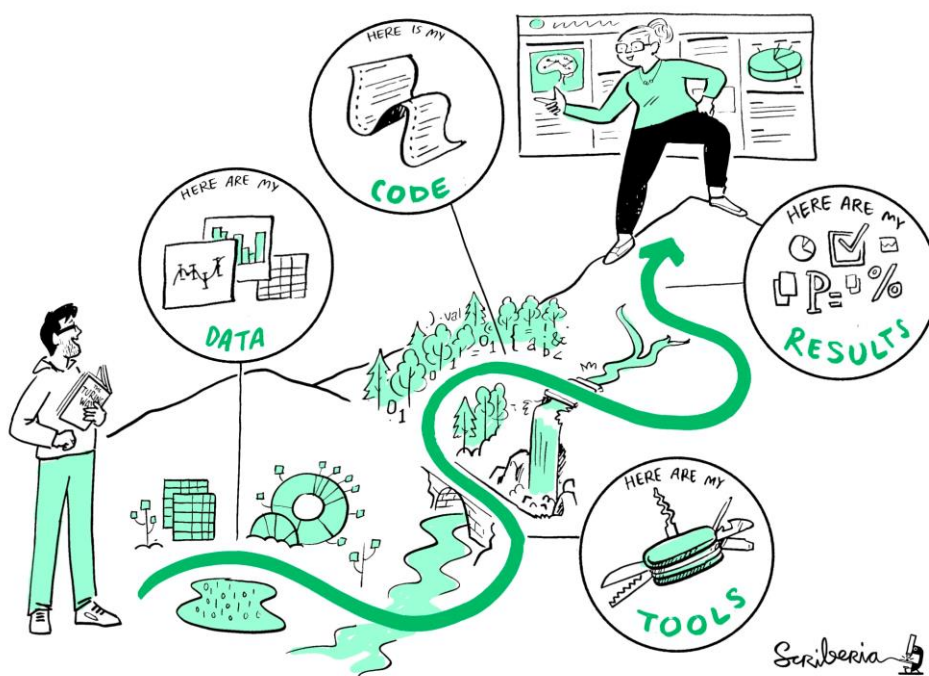


Figure 2.2.1 The pathway from data to the results of an analysis requires the use of software tools and code (Source: *The Turing Way Community, & Scriberia. 2024. Illustrations from The Turing Way: Shared under CC-BY 4.0 for reuse. Zenodo. <https://doi.org/10.5281/zenodo.10556824>*).

1. User-Centric Design:

- Understand User Needs: Conduct user research to understand the specific needs and requirements of climate service users, including scientists, policymakers, and the general public.
- Intuitive Interface: Design software with a user-friendly and intuitive interface to enhance accessibility for users with varying levels of technical expertise.

¹ Accessible software: Software is shared with an appropriate licence via findable and accessible platforms.

² Software Documentation: in general, it must be easy to understand the software (how it works, what it does, and why it does it the way it does), easy to find what needs to be changed, easy to make changes and easy to check that the changes have not introduced any bugs (Crouch, 2024).

2. Data Management and Integration:
 - Standardised Data Formats: Use standardised data formats and protocols to facilitate interoperability and seamless integration with other climate data sources.
 - Data Quality Assurance: Implement mechanisms for data quality assurance, validation, and error handling to ensure the accuracy and reliability of the information provided.
3. Scalability and Performance:
 - Scalable Architecture: Design software with a scalable architecture to handle increasing amounts of data and users over time.
 - Performance Optimization: Optimise code and algorithms to ensure efficient data processing and analysis, especially for large datasets.
4. Open Data and APIs:
 - Open Data Access: Provide open access to climate data through APIs (Application Programming Interfaces) to encourage collaboration, innovation, and the development of third-party applications.
 - Documentation: Document APIs thoroughly to guide developers in using and integrating climate data into their applications.
5. Visualisation and Analysis Tools:
 - Graphical Representation: Include visualisation tools that allow users to easily interpret and analyse climate data through charts, graphs, maps, and other graphical representations.
 - Customizable Dashboards: Enable users to create customizable dashboards to tailor the display of information based on their specific interests and needs.
6. Machine Learning and Predictive Modelling:
 - Integration of Models: Incorporate machine learning algorithms and predictive modelling techniques to enhance the accuracy of climate predictions and analysis.
 - Real-Time Forecasting: Implement real-time forecasting capabilities to provide timely and relevant climate information.
7. Security and Privacy:
 - Data Security: Implement robust security measures to protect sensitive climate data from unauthorised access or manipulation.
 - User Privacy: Adhere to privacy regulations and best practices to safeguard user data and ensure compliance with privacy standards.
8. Accessibility and Inclusivity:
 - Multi-language Support: Provide multi-language support to cater to a diverse user base.
 - Accessibility Features: Design software with accessibility features to ensure that individuals with disabilities can use and benefit from the climate services.
9. Community Engagement:
 - Feedback Mechanisms: Establish feedback mechanisms to gather input from users and stakeholders, allowing continuous improvement of the software based on user experiences and needs.
 - Community Collaboration: Foster collaboration within the climate science and software development communities to leverage collective expertise and resources.
10. Documentation and Training:
 - Comprehensive Documentation: Provide comprehensive documentation, including user manuals, API documentation, and tutorials to assist users in understanding and effectively using the software.
 - Training Programs: Offer training programs to educate users on the features and capabilities of the software.
11. Interdisciplinary Collaboration:

- Collaboration Platforms: Facilitate interdisciplinary collaboration by integrating collaboration platforms or tools that allow researchers, policymakers, and other stakeholders to work together seamlessly.

2.2.1 Metadata

The international discussion on metadata standards for software and services is still ongoing. Consensus exists on software citation metadata standards (see section 2.2.4) and the general need to describe software in order to make it reusable. Github provides recommendations for this in the form of questions to be answered in the readme of a repo³ :

- What the project does
- Why the project is useful
- How users can get started with the project
- Where users can get help with your project
- Who maintains and contributes to the project

There are few initiatives to standardise metadata provision. The CodeMeta Project⁴, supported by amongst others Github and DataCite, has extended the software citation metadata standard by code-specific information like programming language, memory and compute or operating system requirements (see version 3.0 at <https://w3id.org/codemeta/v3.0>). The Software Preservation Network has recently published its Software Metadata Recommendations (Christophersen et al., 2023), in which it provides crosswalks to several schemas in use incl. CodeMeta. Specific adjustments to the description of climate software or climate services still need to be made. According to a recent investigation on the application of FAIR principles to software (Lamprecht, A. L., et al, 2020), it is also recommendable that software and associated metadata should be registered in a suitable, searchable software registry or repository that is dedicated to a particular community, or general ones which are known and widely used (eg. Zenodo, Research Software Directory⁵). This in order to improve findability and gain visibility and impact. For instance as already put into practice by the ESMValTool community.⁶

2.2.2 Technical Standards for Software and Services

Technical Standards in Software covers the technical implementation of standardised input and output data as well as standards in how to write code properly. The input and output formats themselves were described in detail in the previous section 2.1.3 on Formats. For a specific data format there can be several types of implementation in software, this is why those standards are important in the context of software development. This short section lists the most important standards that impact software development.

- Most important standard file formats that are used in climate data distribution and that must be implemented in software
 - File format: NetCDF (OGC): <https://www.ogc.org/standard/netcdf>
 - Metadata Conventions for NetCDF: CF-Convention (Latest 1.11) <https://cfconventions.org/Data/cf-conventions/cf-conventions-1.11/cf-conventions.pdf>

³ <https://docs.github.com/en/repositories/managing-your-repositorys-settings-and-features/customizing-your-repository/about-readmes>

⁴ <https://codemeta.github.io/>

⁵ <https://research-software-directory.org/>

⁶ <https://research-software-directory.org/software/esmvaltool>

- Cloud-optimised file format based on NetCDF data model: zarr <https://zarr-specs.readthedocs.io/en/latest/specs.html>
- GRIB2 standard from WMO: <https://library.wmo.int/records/item/35625-manual-on-codes-volume-i-2-international-codes#.X18yfpMza3l>
- Code writing standards for Python (very popular in current software and tools) and also other languages (ISO standard)
 - PEP8 (python): <https://peps.python.org/pep-0008/> (flake8, black, isort, ruff)
 - ISO/IEC 5055:2021 (SOFTWARE QUALITY STANDARDS)
- Standardised representation of geographical projections and the most complete and used libraries that implement functions to deal with those projections
 - PROJ library (<https://proj.org/en/9.3/index.html>) that implements (or draws on) various standards, e.g. [WKT2:2019 / ISO-19162:2019: Geographic information - Well-known text representation of coordinate reference systems](#), and the open GeoTIFF and Cloud Optimised GeoTIFF (COG) file formats.
 - EPSG Geodetic Parameter Dataset (also EPSG registry) is a public registry of geodetic datums, spatial reference systems, Earth ellipsoids, coordinate transformations and related units of measurement. <https://epsg.org/home.html>
- WMO has developed a framework for WMO data sharing for the WMO Information System 2.0 (WIS 2.0), covering all domains and disciplines. It is defined as the WMO Core Metadata Profile (WCMP) of the ISO 19115 Metadata Standard – Version 2, (Draft, 2024) <https://community.wmo.int/en/activity-areas/wis/wcmp>

2.2.3 FAIR Guidelines and Measures

2.2.3.1 FAIRsoft indicators

The FAIR for Research Software (FAIR4RS) Working Group has adapted the FAIR Guiding Principles for data to create the FAIR Principles for Research Software (FAIR4RS Principles). The contents and context of the FAIR4RS Principles are summarised in (Barker et al., 2022). Del Pico et al. (2022) reasoned that “This interpretation and indicators, albeit fully consistent with FAIRsoft approach, are still too abstract to be applied in a quantitative manner”. To overcome this, Del Pico et al. (2022) generated a set of measurable indicators (see Tables below).

2.2.3.2 Mapping best practices for climate services software to FAIRsoft indicators

It is an interesting exercise to: i) compare the outcomes of the earlier mentioned best practices for climate services metadata in connection with software from the “*survey of documentation from the World Meteorological Organization (WMO), the Global Framework for Climate Services (GFCS), the Open Geospatial Consortium (OGC), the Group on Earth Observations (GEO), and other actors in the metadata sphere for climate services*” to the FAIR metrics; and ii) check how well these guidelines map, and in which parts CE2 could work on further Best Practices in this area. This mapping is listed below.

The metrics that could not be mapped to the best practices are listed in the next section on identified gaps.

F To be Findable: a software can be found and unequivocally identified.		To be Interoperable: a software can be integrated with other tools in the users computational workflow.	
F1	The software has a proper, unique and persistent identifier.	I1	Input/output data types and formats are documented
	F1.1 The software has a unique name to identify it.		I1.1 Input and output data types are formally specified and related to accepted ontologies.
	F1.2 A scheme is used to uniquely and properly identify the software version.		I1.2 APIs (Rest, libraries) are documented in a standard framework (OpenAPI, WES..)
F2	Software is described with rich metadata, including scientific applicability.		I1.3 Input/output data are specified using verifiable schemas (e.g. XDS, Json schema, ...)
	F2.1 Metadata is adjusted to specific metadata formats.		I1.4 The software allows users to choose among various input/output data formats, or provide the necessary tools to convert other common formats into the supported ones.
	F2.2 Metadata is described using accepted ontologies.		I1.5 The software provides provenance information according to accepted standards (PROV)
F3	The software can be found.	I2	The software can be deployed in a format to be included in pipelines.
	F3.1 The software is included in the main software registries.		I2.1 The software has API /library versions to be included in users' pipelines.
	F3.2 The software can be found in any of the major software repositories e.g. GitHub, GitLab, SourceForge.		I2.2 The software can be deployed in e-infrastructures (e. g. Galaxy).
	F3.3 The software can be found in specialised literature services e.g. EuropePMC, PubMed, Journals Site, bioArxiv.	I3	A proper documentation on the software's dependencies as well as mechanisms to obtain them is available.
			I3.1 The software includes details about dependencies.
			I3.2 The software includes its dependencies or mechanisms to access them.
			I3.3 The software is distributed via a dependencies aware system.
A To be Accessible: it is possible to access a usable form of a software.		R To be (Re)Usable: the software can be properly used and/or contributed to.	
A1	A working version of the software can be accessed/downloaded/built.	R1	The software provides adequate usage documentation.
	A1.1 A working version of the tool can be accessed through an API or web.		R1.1 The software user guides are provided.
	A1.2 A working version of the software can be accessed/downloaded/built, including the generation of a software container.		R1.2 Examples of use cases are provided.
	A1.3 A set of instructions and other necessary information the user can follow to build the software is available.	R2	A clear and accessible usage licence is provided.
	A1.4 Test data is available.		R2.1 Terms of Use are stated.
	A1.5 Source code of the software is available.		R2.2 Conditions of installation and usage are stated.
A2	Code and metadata are available even when the software is no longer in use	R3	A contributors policy exists.
	A2.1 Metadata of previous versions is available.		R3.1 A document stating the contributors policy exists.
	A2.2 Previous versions are available.		R3.2 Credit for contributions is provided.
A3	No restrictions exist to access the software	R4	Provenance is available.
	A3.1 The software can be used without registration.		R4.1 The software follows a version-control system.
	A3.2 The software can be used in a free operating system.		R4.2 The software follows a defined and documented release policy.
	A3.3 Versions of the software for several operative systems are available.		R4.3 Metadata of previous versions is available.
	A3.4 The software can be used in a free e-infrastructure		
	A3.5 The software can be used in several e-infrastructure		

(FAIRsoft indicators for the FAIR principles applied to Research Software (Del Pico et al., 2022))

Accessible

A1. A working version of the software can be accessed/downloaded/built

- Open Data Access: Provide open access to climate data through APIs (Application Programming Interfaces) to encourage collaboration, innovation, and the development of third-party applications.
 - **In order to fulfil this FAIR requirement not only the data access software/API's should be accessible but also software involved in the processing of the data.**

Interoperable

I1. Input/output data types and formats are documented

- Standardised Data Formats: Use standardised data formats and protocols to facilitate interoperability and seamless integration with other climate data sources.
- Documentation: Document APIs thoroughly to guide developers in using and integrating climate data into their applications.
 - **An important additional guideline is that, for each software/APIs, at least the input and output data should be clearly documented**

I2. The software can be deployed in a format to be included in pipelines

- Collaboration Platforms: Facilitate interdisciplinary collaboration by integrating collaboration platforms or tools that allow researchers, policymakers, and other stakeholders to work together seamlessly.
 - **The FAIR metric is in this case a more concrete guideline specific for the software involved and that this by itself should be able to be deployed in other pipelines (meaning it should be interoperable in technical sense, input/output, etc).**

Reusable

R1. The software provides adequate usage documentation

- Comprehensive Documentation: Provide comprehensive documentation, including user manuals, API documentation, and tutorials to assist users in understanding and effectively using the software.
- Training Programs: Offer training programs to educate users on the features and capabilities of the software.
- Documentation: Document APIs thoroughly to guide developers in using and integrating climate data into their applications.

R2. A clear and accessible usage licence is provided

- Data Security: Implement robust security measures to protect sensitive climate data from unauthorised access or manipulation.
 - **Security may be seen as an aspect of the usage licence, but the usage licence itself also should be clearly provided in the metadata.**

2.2.3.3 Identified Gaps

It can be seen that not all FAIR metrics are covered by current best practices. This is probably due to the difference in focus, where the FAIR metrics focus on the technology side and the machine to machine use, whereas the best practices from the survey are more general and also focus on the end-user and other stakeholders. We will come back to this conclusion. For now, we can see that the following metrics are not covered yet, and might need further attention and integration into the Best Practices for climate services software.

Findable

F1. The software has a proper, unique and persistent identifier

The uniqueness of an identifier is a necessary condition to unambiguously refer to that resource and that resource alone. The name is commonly used as the main identifier of a software. Each tool should have a unique name to avoid ambiguities. A version scheme is necessary to refer to a specific release of a software and keep track of the incrementally different versions of the software. Different versions of the same software should share a name, but if substantial modifications in the algorithm are done, the identifier should change for the new piece of software.

F2. Software is described with rich metadata, including scientific applicability

Metadata makes it possible to find tools through search engines and enables a potential user to decide whether a tool is of interest. Metadata should be adjusted to specific metadata formats and described using accepted ontologies. This is important, because if each tool is being described with different terminology, with non-specified meanings, it makes it more difficult to interpret metadata and it makes searching slower and more difficult.

F3. The software can be found

There are multitude of possibilities for scientists to find specific software. Software registries are the main resource scientists use when searching for software. Software repositories such as GitHub, GitLab, SourceForge etc. can be an additional resource used by scientists when looking for software. Specialised literature like EuropePMC, PubMed etc. is a good reference to find software, especially to discover new software.

A2. Code and metadata are available even when the software is no longer in use

This metric governs whether there is available code and metadata even when the software is no longer in use, in order to be able to match software provenance with analysed data provenance. Even if the code/functionality is missing, the existence of metadata for a given version of a software provides evidence of its previous existence and relevant details should still be available, to be able to reproduce analyses with the original software versions still available.

A3. No restrictions exist to access the software

Registration is a barrier for accessibility, since users may not want to identify themselves. Non-free operating systems are an economical barrier for software accessibility. Having to run a software in a different operating system than the one used by a user in their research implies making a big effort. Sometimes, only one operating system is available in research (if a cluster or supercomputer is needed, for instance). The more supported operating systems, the more accessible the software.

I3. A proper documentation on the software's dependencies as well as mechanisms to obtain them is available

Dependencies of a software are absolutely necessary to run a software. A dependencies statement allows the user to know what additional software they must install to get a software running. Even if dependencies are stated, the process of downloading and installing them can be incredibly hard, except if the dependencies are directly provided with software or the mechanisms to obtain them are provided. Dependencies-aware systems like 'bioconda' are the easiest and most straightforward way to obtain and install all the dependencies a software needs.

R3. A contributors policy exists

The lack of a policy or guide stating how to contribute to a software prevents external developers willing to contribute to it. The credits for contributions should be stated, as contributors hold the copyright.

R4. Provenance is available

Provenance allows us to understand where and how the software has been used and the applications it may have been involved with. Version control ensures the software changes are recorded and searchable. Without a release policy, users cannot know the differences between versions of a software. Even if the code/functionality is missing, the existence of metadata for a given version of a software, used in a published research the user is interested in, for instance, enables them to be certain about the past existence of a software and relevant details about it (see also Deliverable 2.2 on provenance).

2.2.3.4 Main Conclusions

As mentioned above, there are several Climate Services Best practices that do not map to the FAIR software metrics:

- **Understanding User Needs:** Conduct user research to understand the specific needs and requirements of climate service users, including scientists, policymakers, and the general public.
- **Data Quality Assurance:** Implement mechanisms for data quality assurance, validation, and error handling to ensure the accuracy and reliability of the information provided.
- **Intuitive Interface:** Design software with a user-friendly and intuitive interface to enhance accessibility for users with varying levels of technical expertise.
- **Scalability and Performance:**
 - **Scalable Architecture:** Design software with a scalable architecture to handle increasing amounts of data and users over time.
 - **Performance Optimization:** Optimise code and algorithms to ensure efficient data processing and analysis, especially for large datasets.
- **Visualisation and Analysis Tools:**
 - **Graphical Representation:** Include visualisation tools that allow users to easily interpret and analyse climate data through charts, graphs, maps, and other graphical representations.
 - **Customizable Dashboards:** Enable users to create customizable dashboards to tailor the display of information based on their specific interests and needs.
- **Machine Learning and Predictive Modelling:**
 - **Integration of Models:** Incorporate machine learning algorithms and predictive modelling techniques to enhance the accuracy of climate predictions and analysis.
 - **Real-Time Forecasting:** Implement real-time forecasting capabilities to provide timely and relevant climate information.
- **User Privacy:** Adhere to privacy regulations and best practices to safeguard user data and ensure compliance with privacy standards.
- **Accessibility and Inclusivity:**
 - **Multi-language Support:** Provide multi-language support to cater to a diverse user base.
 - **Accessibility Features:** Design software with accessibility features to ensure that individuals with disabilities can use and benefit from the climate services.
- **Community Engagement:**

- Feedback Mechanisms: Establish feedback mechanisms to gather input from users and stakeholders, allowing continuous improvement of the software based on user experiences and needs, which can be enhanced by fostering collaboration within the climate science and software development communities.

When checking this list we see that most are focused on e.g. the end-user visualisation, inclusivity, user interface and not focused on the Findability, Accessibility, Interoperability and Reusability as defined under the FAIRness for software. It needs to be discussed if there are some practices that should actually be included after all (expanding the FAIR metrics for software in climate services), and the rest still needs to be kept of course, but maybe as e.g. “Guidelines for climate service development”.

At the same time the FAIR metrics for software could be discussed by a CE2 working group in the context of climate service software, reformulated/tuned/stricter defined where needed and published as “FAIRness guidelines for climate services software”.

2.2.4 Software and Services Licensing and Citation

Software licensing establishes the terms and conditions under which software can be copied, distributed and modified by granting legal permission. Licences are assigned by the software owner or copyright holder. An overview over existing licences with their legal conditions is available at <https://spdx.org/licenses/>. Open licences allow the reuse of the software by all stakeholders. Open software licensing is encouraged by the FAIR for research software (FAIR4RS) initiative. Imposed restrictions limit the reuse and recombination of the software or software component. As climate services often use software components from multiple copyright holders, the handling of these imposed restrictions, which might be incompatible, might curtail the publication of a modified software product.

The RDA FAIR4RS WG recommends an as unrestrictive as possible licence to and choosing a widely used and recognized licence (Chue Hong et al., 2022). Examples are ISC, Apache 2.0, MIT, GPL, AGPL, and BSD2 licences. This licence must comply with the licences of the software’s dependencies so that the software can be legally combined. Furthermore, software as different functions, for which Software Heritage (<https://www.softwareheritage.org>) names different typical open source software licences:

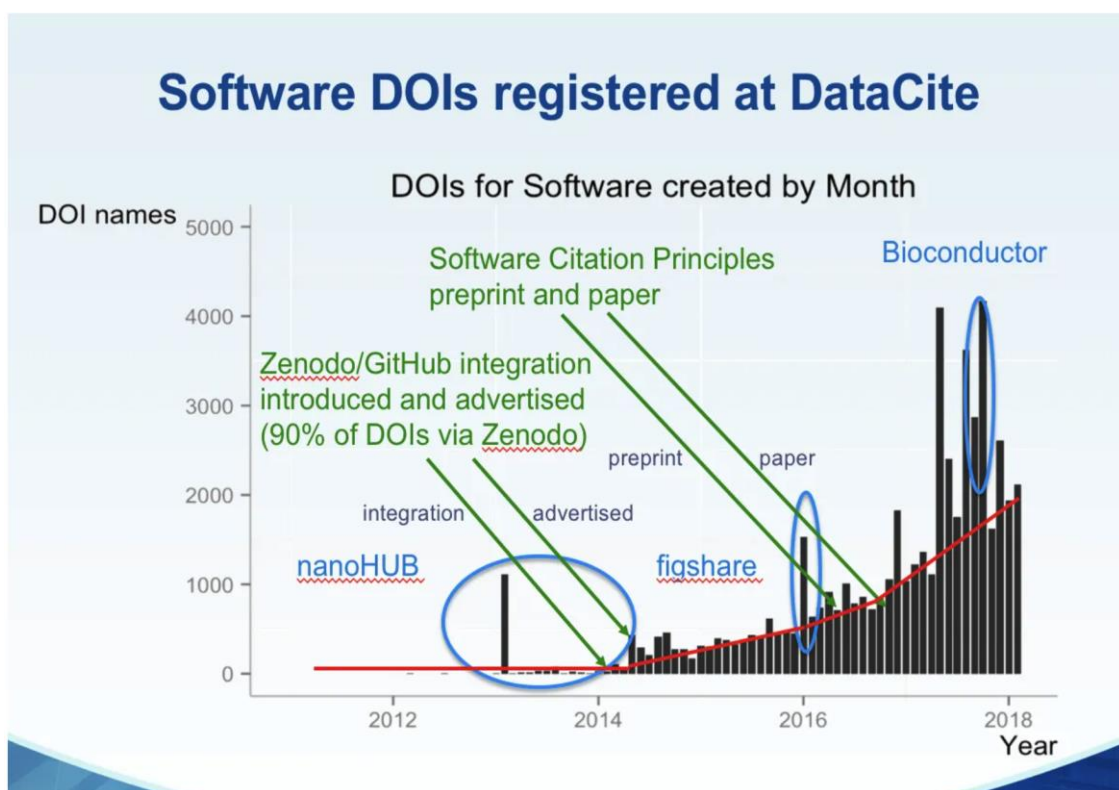
- AGPLv3+ for public-facing services on the Internet
- GPLv3+ for backend code and services
- Apache 2 for Puppet code and recipes

Whether the adoption of a copyleft licence such as GPL and AGPL should be recommended versus a more permissive one, such as Apache 2, is ultimately dependent on the envisaged impact of the software on the data and the user community, and on the need for transparency as a fundamental principle, which is applicable to every software that has been derived from the original source code. More specifically, any source code that produces derived climate products starting from a single or multiple raw data sources, should be made available under GPL or AGPL. The latter when this is offered as a service. Thereby guaranteeing transparency and fostering reproducibility throughout future adaptations and improvements by third parties. A more permissive approach is acceptable instead for software that serves merely as an operational element of an infrastructure. For instance, this applies to libraries and services used for the orchestration and management of their underlying storage and computational resources, not having a direct impact on the content delivered.

Licences should be human and machine-accessible. Many licences require to describe modifications made to the original software, which also supports the transparency and traceability of the climate service product.

Good scientific practice asks for giving appropriate credit to the software (component) developers. In science credit is commonly given by citation. Software Citation is an important step in having research software engineers have their work recognized and cited. The FORCE11 principles of software citation are described by Smith et al. (2016). The follow-up FORCE11 Software Citation Implementation Working Group has derived concrete guidelines for software citation (Stall et al., 2023). DOIs are an important component in provenance information and for connecting scholarly information via PIDs to build PID graphs.

According to the FAIR4RS recommendation, software should be identified using a DOI. Several software platforms exist, which allow to provide a data reference like the [CFF file in github](#). Some platforms offer DOIs on software and software versions, like the popular [github-Zenodo link](#), where Zenodo registers DOIs through DataCite. This approach was launched in 2014 and has resulted in a significant increase of DOIs on software (Fenner, 2018).



Software DOIs registered at DataCite

In September 2016, the FORCE11 Software Citation Principles ([@https://doi.org/10.7717/peerj-cs.86](https://doi.org/10.7717/peerj-cs.86)) were published, the Zenodo/GitHub integration was upgraded ([@http://help.zenodo.org/whatsnew/](http://help.zenodo.org/whatsnew/)), and in October 2016 the [GitHub Guide to Making your Code Citable](#) was updated. There appears to be a change of in the rate of growth around this time as well.

(source: Fenner, 2018)

A paper on persistent identification and citation of software using DOIs by Jones et al. (2017) was published in 07/2017, based on earlier work from 2015 (<http://purl.org/net/epubs/work/24058274>), and the DataCite Metadata 4.1 schema focussing on software citation was released in 09/2017 (DataCite Metadata Working Group, 2017; Starr, 2017). The most current version of DataCite Metadata schema is now 4.5 (DataCite Metadata Working Group, 2024a and 2024b). The mandatory properties of the DataCite schema are shown below.

Table 1: DataCite Mandatory Properties

ID	Property	Obligation
1	Identifier	M
2	Creator	M
3	Title	M
4	Publisher	M
5	PublicationYear	M
10	ResourceType	M

(source: *DataCite Metadata Working Group, 2024*)

DataCite recommends the inclusion of a data citation statement on the DOI landing pages, which should be used by the climate service providers to cite the used software (components). Software run as a climate service should be as open as possible. Open software sharing is encouraged for transparency and product traceability together with a consistent versioning concept and the use of the github-Zenodo link to make the different versions uniquely referenceable and citable.

2.3. Organisational Standards

Data and software/service standards are implemented and maintained by the climate service providers. Thus, organisational standards and certificates that define how these technical standards are supported by the organisation in its workflows and employee training, as well as the corporate identity and culture of the organisation, play an important role. Existing international standards target different aspects, such as data quality and provenance.

The Quality Management System (QMS) approach of the ISO 9000 family focuses on the organisation's processes to ensure consistently high quality of customer-oriented products. The inclusion of the supply chain is becoming increasingly important for ISO 9001 certification.

The [European Framework for Audit and Certification of Digital Repositories](#) focuses on trustworthy repositories and the compliance of their data management practices with common general data management principles such as FAIR for data, CARE (Collective benefit, Authority to control, Responsibility, Ethics) for indigenous data governance, and TRUST (Transparency, Responsibility, User focus, Sustainability, Technology) for digital repositories. The three certification instruments reflect an increasing level of complexity, ranging from a self-assessment of the repository to a formal audit designed to encourage a continuous improvement of good practices.

The ISO 9001 certification is more popular in the public and private sectors, while the certifications associated with the trustworthy repository framework are limited to institutions hosting a research data repository.

2.3.1 Quality Management System (QMS)

ISO 9000 is defined as a series or family of international standards on quality management and quality assurance developed to help companies effectively document the quality system elements needed to maintain an efficient quality system. It helps organisations of all sizes and sectors to improve their performance, meet customer expectations and demonstrate their commitment to quality.

The ISO 9000 family contains following four standards of interest for CS development (<https://www.iso.org/standards/popular/iso-9000-family>):

- ISO 9000:2015: Quality Management Systems - Fundamentals and Vocabulary (definitions)
Standard describes the fundamentals of QMSs and specifies the terminology used in ISO 9000.
- ISO 9001:2015: Quality Management Systems - Requirements
Requirements are generic in nature and can be adopted by almost any organisation. Those standards of quality are internationally recognized. It is the only standard in the ISO family against which organisations can be certified (or registered), through a third party audit process.
- ISO 9004:2018: Quality Management - Quality of an Organization: Guidance to Achieve Sustained Success (continuous improvement)
This standard focuses on achieving sustainable success by meeting the needs and expectations of customers and other stakeholders. It promotes self-assessment as an important tool, but it is not a substitute for a third-party audit process, which is fundamental to ISO 9001.
- ISO 19011:2018: Guidelines for Auditing Management Systems
Standard provides guidelines for auditing management systems, including quality management systems (ISO 9001) and environmental management systems (ISO 14001). It outlines the principles of auditing, managing audit programs, and conducting management system audits. It is a critical standard for any organisation looking to manage external audits.

The ISO 9001 (<https://www.iso.org/standard/62085.html>) standard specifies requirements for the establishment, maintenance, and continuous improvement of a quality management system, covering a wide range of topics including context of the organisation, leadership, planning, support, operation, performance evaluation and improvement. All these topics are explained in detail in ISO 9001. Overall, ISO 9001 helps organisations address all aspects of managing the quality of their products and services, with an emphasis on customer satisfaction and constant enhancement of the quality management system.

One of the advantages of introducing ISO 9001 standard is the improvement of internal efficiency. Through the process of introducing standards, focusing on processes in the organisation, procedures and responsibilities are clearly defined, which leads to increased efficiency and reduces the loss of time and resources. The existence of standards indicates the existence of an established quality management system, which increases user confidence in the quality of products or services. Having the ISO 9001:2015 certificate, the organisation can distinguish itself from its competitors, thereby attracting new users and increasing business opportunities, which gives it a competitive advantage. The standard encourages the organisation to continuously improve, to constantly evaluate its performance and identify opportunities for improvement. A significant benefit of this standard is better risk management because it introduces a systematic approach to risk management. Implementing standards can be a financial challenge, especially for small organisations that do not have the resources to hire external consultants, which represents increased costs.

The implementation of the standard can take up to three years, which requires the time and commitment of employees, and this can affect their normal daily routine. Another disadvantage of introducing the ISO 9001 standard is the excessive documentation that sometimes creates additional overhead for organisations and causes them to be overextended. This depends only on the organisation itself and its understanding of the need for documentation within the process. In order to properly implement the standard, it is necessary to provide training to employees so that they

understand and implement the requirements of the standard. After obtaining the certificate, the organisation must continuously maintain the quality management system in order to maintain its status, which can be a challenge in the long run.

The transition to a new version of the standard, like any transition, has its own obstacles and requires time and dedication, especially if a new concept is introduced, such as in the latest version of ISO 9001:2015, risk-based thinking. For many users, risk-based thinking also brought the most benefits compared to the previous version of the standard. Although there are some potential disadvantages of implementing the ISO 9001:2015 standard, most organisations that have implemented this standard find that its advantages are much more important and beneficial to their business. More about the advantages and disadvantages in the implementation of ISO 9001 as well as the transition through different versions of the standard can be found in Bakator and Cockalo (2018), Fonseca et al. (2019) and Bravi et al. (2019).

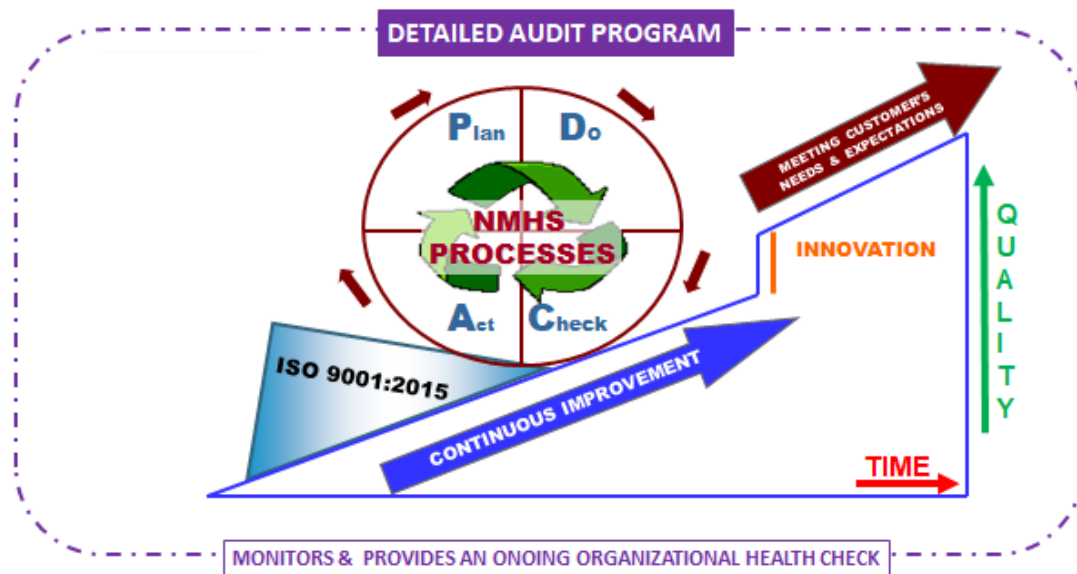
Although the standard is most often associated with traditional companies, its principle of quality management can also be applied in innovative companies by:

- adjustment of requirements - a company can adapt the requirements of the standard to suit the needs of the innovation environment;
- focus on processes - this involves the identification and documentation of key processes, as well as the continuous monitoring and improvement of those processes to ensure the quality of products or services;
- introducing a culture of continuous improvement - this involves identifying potential opportunities for improvement, implementing changes and monitoring their impact on service quality and user satisfaction;
- ensuring customer satisfaction - this may include gathering feedback from users, analysing their requirements and adapting/creating innovative solutions to meet their needs.

Within ISO 9001:2015 there is a section that describes in detail the requirements related to the innovation and development part of the company (clause 8. Operation, 8.3 Design and development of products and services).

The documents WMO Guide to the Implementation of Quality Management Systems for National Meteorological and Hydrological Services and Other Relevant Service Providers (WMO-No. 1100; WMO, 2017a) and Guidelines on Quality Management in Climate Services (WMO-No. 1221; WMO, 2018a) provide detailed recommendations for the introduction of ISO 9001:2015 in NHMSs and especially in climate services. As NHMSs are responsible for conducting measurements and observations, requirements are defined here for that segment as well. Climate services that do not have measurements and observations need to pay special attention to the choice of data providers. The quality of data, which can be ensured by clear requirements towards providers, directly affects the quality of services and thus the satisfaction of users.

Recommendations for the necessary knowledge and skills of employees of NHMS climate services are given in the documents WMO Compendium of WMO Competency Frameworks (WMO-No. 1209; WMO, 2019) and Guidelines for the Assessment of Competencies for Provision of Climate Services (WMO-No. 1285; WMO, 2022).



Implementation of QMS in NHMS

The WMO QMS references documenting:

- *Quality control of climate data:*

All observations should be appropriately flagged to indicate the level of quality control applied to the data. The quality flag should indicate the different stages of quality control that a given value went through. This includes differentiating whether it went through extended quality control procedures or only basic ones, or no quality control at all. Corrections or estimations should be clearly flagged before being entered into the database. However, the original data must also be retained in a database. Quality control procedures and algorithms should be documented in detail for each stage of data processing, from observation through to archiving.

- *Homogenize data:*

Many homogenization methods are widely used by various international research institutions. Recommendations for homogenization of monthly, seasonal and annual data are summarised in the Guidelines on Climate Metadata and Homogenization (WMO, 2003), which describe the requirements for station metadata and homogeneity and the recommended methodologies. Processes for the availability of updated, homogenised climate datasets with complete and relevant metadata should be documented.

- *Ensure traceability of products:*

It is a requirement for nearly all products produced by NMHSs that each product must be traceable to the version of the data from which it was created, along with details of items such as station location and time of data receipt/product issuance. This also means keeping detailed records of the product history. A product can be identified by providing a serial number, but effective traceability means that it would be possible to provide the history of that product and how it was produced (see also Deliverable 2.2 on provenance for more details).

WMO Publication No. 1131 Climate Data Management System Specifications provides additional guidance with respect to establishment of a Climate Data Management System. The publication establishes a framework defining the functionality required within a climate data management system (CDMS). A CDMS is an integrated computer-based system that facilitates the effective archival, management, analysis, delivery and utilisation of a wide range of integrated climate data. The framework comprises a set of interrelated building blocks called components. Each component describes a specific functional requirement of a CDMS and contains, where appropriate, references to further information. The section on Climatology Policy (3.1.6) reviews policies that ensure appropriate climate metadata are maintained to facilitate a better understanding of climate data, metadata on observations, discovery and data provenance.

In accordance with their activities and processes, climate services can define specific competencies for their employees. It should be kept in mind the focus on users, their needs, as well as maintaining the quality of services and continuous improvement, even when defining the competencies of employees. Continuous improvement of the knowledge and skills of employees through additional education can significantly contribute to maintaining the quality and maintenance of climate services.

2.3.2 Avio - The International Civil Aviation Organization (ICAO) standards

The importance of weather and climate for aviation and aerodrome siting has led to adoption of the Standards and Recommended Practices related to meteorological service to air navigation, that are provided in the Annex 3 to the Convention on International Civil Aviation (ICAO, 2010). These standards and recommendations are in a continuous improvement process, following close collaboration of the International Civil Aviation Organization (ICAO) and World Meteorological Organization (WMO)(ICAO, 1963) with the main purpose of improving the safety, regularity and efficiency of international air navigation. Specification of the needs of international civil aviation concerning meteorological service is reviewed and developed by ICAO, while technical regulations, guidance and other reference materials regarding the manner in which the required meteorological service is provided are reviewed and developed by WMO.

Reorganisation of regulatory material, in order to increase efficiency and avoid duplication while pertaining compatibility with quality management system principles, is taking place. As a result, two-stage discontinuation of the Technical Regulations (WMO-No. 49; WMO, 2018), Volume II, Meteorological Service for International Air Navigation, has been agreed between ICAO and WMO. At the first stage, Parts I (International Standards and Recommended Practices: Core Standards and Recommended Practices) and Part II (International Standards and Recommended Practices: Appendices and Attachments) of WMO-No. 49 Volume II are withdrawn from publication on 31 December 2023, while in the second stage Parts III (Aeronautical Climatology) and Part IV (Format and Preparation of Flight Documentation) are to be withdrawn from publication when any material of continuing relevance has been transferred to ICAO Procedures for Air Navigation Services – Meteorology (PANS-MET) (provisionally in 2026).

Climate services for aviation required for the planning of flight operations are prepared in the form of aerodrome climatological tables and aerodrome climatological summaries, defined by technical specifications covering processing, exchange and content of aeronautical climatological information.

2.3.3 Laboratory Standards

In accordance with the CS taxonomy developed in the EU H2020 MARCO (Cortekar et al., 2020) and adopted in the D4.1 and D4.2, one of the types of CS are Measurements - Instruments and technologies for measurement and calibration.

International Organization for Standardization (ISO) regulates the general requirements for the work of laboratories for calibration and testing of instruments, including sampling and routine observations through ISO/IEC 17025 General requirements for the competence of testing and calibration laboratories (<https://www.iso.org/publication/PUB100424.html>). This standard specifies the requirements for the competence, impartiality and consistent operation of laboratories and is applicable to all organisations performing laboratory activities, regardless of the number of personnel. The standard is of an accreditation nature, i.e. it is used to assess compliance and award accreditation. The assessment of compliance is carried out by accreditation bodies. Once accreditation is established, there is an ongoing periodic external audit, which provides additional proof that standards have been maintained. Besides these external controls from the certifying agency, employees at the laboratory are obliged to perform regular and extraordinary checks of the correctness of the measuring instruments.

The latest version of the standard ISO/IEC 17025:2017 is aligned with ISO 9001 and standards for conformity assessment (ISO/IEC 17000 series standards: Conformity assessment — Vocabulary and general principles) as well as with other standards (user requirements, tenders...).

In addition to general requirements (impartiality and confidentiality), the standard defines technical and management requirements. An essential component of the technical requirements is the development of uncertainty analyses for each of the measurement processes, including documented and verified traceability to international metrology standards. More specifically they include:

- Structure requirements - a laboratory is required to define and document the scope of laboratory activities for which it is compliant with ISO/IEC 17025 which excludes externally procured laboratory activities on an ongoing basis (define accredited calibration methods).
- Resource requirements - a laboratory is required to have available personnel, facilities, equipment, systems and support services (like metrological traceability and/or outsourced products and services) necessary to manage and perform laboratory activities.
- Process requirements - these requirements define the work flow and include: review of requests, offers and contracts; selection, verification and validation of methods; sampling; handling of test or calibration items; technical records; evaluation of measurement uncertainty; ensuring the validity of the results; reporting of results; objections (request for "another person" and dealing with complaints); nonconforming business; and data and information management.

For management system requirements, a laboratory must establish, document, implement and maintain a management system that is able to support and demonstrate consistent implementation of the requirements and to ensure the quality of the laboratory's results (i.e. implementation of ISO 9001). All laboratory processes must be approved, verified and validated in a suitable manner to meet the requirements.

While ISO/IEC 17025 is general standard for any kind of laboratory, the Commission for Instruments and Methods of Observations (CIMO) of the WMO is responsible for matters relating to international standardisation, compatibility and sustainability of instruments and methods of observation of meteorological, climatological, hydrological, marine, and related geophysical and environmental variables. The Commission is responsible for developing the guidelines and recommendations to assess the quality of the observational data provided by various systems and shared through metadata, and underpins all observations within the WMO Integrated Global Observing System (WIGOS).

One of the WMO guidelines, that is the authoritative reference for all matters related to instrumentation and methods of observation in the context of WIGOS, is the Guide to Meteorological Instruments and Methods of Observation (WMO-No. 8; WMO, 2018b). This Guide is a key resource that provides a

description, procedures and the basic capabilities of most instruments, systems and techniques in regular use in order to assist NHMSs and other interested users operating observing. The Guide restricts standardisation to the essential requirements only, and confines recommendations to the general features most common to various configurations of a given instrument or measurement system, thus enabling wide areas for further development. Volume V of this Guide is specifically dedicated to Quality assurance and management of observing systems where previously mentioned standards are described as well as their implementation and process for accreditation.

Accreditation of the laboratories requires additional processes and documentation to the QMS and, most importantly, evidence that laboratory staff have been trained and have mastered the processes and methods to be accredited. The engineers who design or maintain complex instrumentation systems require a particularly high order of theoretical and practical training. The WMO competency frameworks for meteorological observations, instrumentation, calibration and observing programme and network management provide a more detailed description of the job responsibilities and tasks, and required instrument knowledge and skills for practising professionals.

2.3.4 The role of Education

There are significant differences between academic qualifications and competencies. WMO Technical Regulations, Volume I (WMO-No. 49; WMO, 2017b), describes qualifications as the minimum core knowledge, usually acquired through education required to enter a profession, while competency is defined as the knowledge, skills and personal attributes or behaviours required to perform specific tasks in the fulfilment of a job responsibility. Acquiring competency in specific areas will require ongoing job-specific education and training throughout an individual's career.

WMO establishes competency requirements that become part of the WMO Technical Regulations (in the form of standard or recommended practices) for service areas and related personnel for which international harmonisation and standardisation is considered important, but not every service area will need international competency requirements. For service areas where harmonisation of competency is considered desirable and beneficial, such requirements may be defined as "recommendations".

On the other hand, for service areas that are subject to stringent international regulatory frameworks, such as aeronautical meteorological services, competency requirements may be defined as "standards". In these types of services, the implementation of a QMS is mandated by international regulatory bodies (i.e. the International Civil Aviation Organization (ICAO)). Implementation of a QMS is a key driver to have demonstrably competent staff across an organisation's service area(s).

Competency assessment requirements as well as the interval between competency assessments will vary depending on the service area and the structure of the organisation providing the service. In some countries, Aeronautical Meteorological Forecasters are reassessed (obtaining a licence extension) on a six-month or annual basis, while in others, in-depth reassessment may be undertaken only every five years with less intensive assessment in the interim period. The reassessment period should be documented and included in the organisation's QMS.

Generally, for services implementing a QMS, qualification and competency records are required for auditing purposes and should be continually kept up to date and retained with appropriate privacy systems in place. This is good practice for all services.

Provision of climate services requires competencies for the transformation of climate data into climate products and services. The WMO developed the competency framework for the provision of climate services to help NMHSs and other institutions to deliver high-quality climate services in compliance with WMO standards and regulations, specifically those defined by the Standing Committee on

Climate Services (SC-CLI) and the Global Framework for Climate Services (GFCS)(WMO-No. 1285; WMO, 2022).

The competency framework for the provision of climate services is divided into the five top-level competencies, with related performance criteria and suggested learning outcomes:

1. Create and manage climate data sets;
2. Derive products from climate data;
3. Create and interpret climate forecasts, climate projections and model output;
4. Ensure the quality of climate information and services;
5. Communicate climatological information to users.

The list of the competencies to be met and the associated performance criteria would be determined by the infrastructural and human capacity of each institution. The competencies falling in the areas of quality of climate information and services as well as communication of climatological information with users are considered cross-cutting and should be met, at least at basic levels, by all institutions providing climate services.

All the WMO competency frameworks can be accessed through Compendium of WMO Competency Frameworks (WMO-No. 1209; WMO, 2019).

The American Meteorological Society (AMS, <https://www.ametsoc.org>) has established Certified Consulting Meteorologist Program (CCM): high standards of technical competence, character, and experience for those who would like to provide consultations in meteorology. The CCM certification is awarded by the AMS to highly qualified meteorologists providing research and services to a wide variety of users of weather and climate information. Certification enables users of meteorological and climate services to select consultants with greater confidence in the quality and reliability of the products and services they will receive. It has been nationally recognized as the standard in meteorological certification since 1957.

The CCM certification is granted only to those who demonstrate a broad background in meteorology and/or oceanography together with a detailed knowledge in a particular field of specialisation. As a minimum, applicants must meet one of the educational requirements and have a minimum of five years of work at the professional level in meteorology or a related field. To earn the CCM certification, a meteorologist must make an application, be recommended by three associates, pass a stringent written examination, and pass an oral examination before a national board of examiners. Certification is for a period of one year, and can be renewed annually.

2.3.5 Trustworthy Digital Repository Framework

The [European Framework for Audit and Certification of Digital Repositories](#) consists of a sequence of three levels with increasing trustworthiness:

1. [CoreTrustSeal \(CTS\)](#) as basic level certification based on a self-assessment expiring after 5 years;
2. [Nestor-Seal/DIN 31644](#) as extended-level certification based on a plausible self-assessment without a defined expiry date; and
3. [ISO 16363:2012](#) (Audit and certification of trustworthy digital repositories) as formal-level certification based on an external audit process.

These three groups signed a Memorandum of Understanding in 2010 to collaborate in setting up an integrated framework for auditing and certifying digital repositories. The framework is intended to guide repositories in the continuous improvement of their best practice.

The initial set of criteria for the CTS was the result of a harmonisation of the Data Seal of Approval (DSA) criteria and the World Data System (WDS) criteria for a Regular Membership. The work was coordinated by the [RDA/WDS Certification of Digital Repositories IG](#) and the resulting catalogue of common procedures for certification was approved by RDA (Rickards et al., 2016). CoreTrustSeal's Assembly of Reviewers consists of representatives of CTS-certified repositories. Therefore, the review process can be regarded as a community review. The Assembly of Reviewers elects the members of the CTS governance board. The current CTS criteria for 2023-2025 can be grouped in the categories: organisational infrastructure, digital object management, and IT and security (CoreTrustSeal, 2022).

The CTS has become the minimum standard for research repositories supported by funders and scholarly publishers. An example for this is the community-driven [Enabling FAIR Data Project](#), which is supported by the Coalition for Publishing Data in the Earth and Space Sciences (COPDESS). It has developed criteria and the Repository Finder tool for an easy selection of a trustworthy repository. CTS certification is an important prerequisite for recommended repositories. Very few research repositories have achieved the higher certification level of a Nestor Seal or successfully passed the ISO 16363 audit process.

3 User Engagement in Developing Standards

For a standard to be suitable for its intended purpose, practical to implement, and widely adopted by potential users, it is vital that the users are not seen simply as passive consumers. Instead, they should be invited to engage and even participate throughout the process of standards development to ensure the end result meets their needs and expectations.

An example of this type of user engagement is the CF community, comprising all users of the CF metadata conventions who choose to participate. Typically, these are scientists intending to write data, software developers writing tools to work with the conventions, and data stewards needing to apply standards to enable searching, access and processing of data from multiple sources. Individuals may ask for advice from other members of the community, propose additions to the controlled vocabularies, propose an extension to the conventions document to cover a new use case, or contribute to the effort of maintaining the document and website.

Another example of user engagement on the development of standards, at a smaller scale, is what is currently being done in the Destination Earth project where there have been active exchanges between the contract responsible for the climate simulations (DE340⁷) and the contract of the use cases (DE370) to design the standards for the climate simulations used in services.

3.1 CF Community Engagement

The cf-conventions.org GitHub organisation is key to achieving active community participation. In addition to containing all the documentation pertaining to the standard, the CF website explains the community processes for proposing, discussing and agreeing additions and enhancements. The underlying philosophy behind the development of CF is that it is impossible to foresee all future requirements, so enhancements are only added when a use case is shown to exist. To date, all additions to the conventions have been made in a backwards compatible manner, so as not to invalidate data produced in compliance with earlier versions of the document.

Anyone with a GitHub account may open an issue in one of the CF GitHub repositories to begin a conversation. The “[discuss](#)” repository is often the first point of contact for someone new to the community. It is used as a forum to ask questions about how to apply the conventions, to propose additions to the controlled vocabularies, and to hold preliminary discussions which may lead to a more formal proposal for an enhancement to the CF conventions. Once a proposal is made for either a standard name or a conventions enhancement, the CF process requires a public discussion in which any interested person may participate. The aim of the discussion is to evaluate the necessity of the change and agree on precise wording. A moderator (usually but not exclusively a member of the CF committee) oversees the discussion to ensure that all points raised are responded to by the proposer and steers the conversation towards a community consensus. There are community [rules](#) governing such matters as what constitutes a consensus, and the timescale on which a proposal can be regarded as being fully agreed and ready for publication. Transparency is a vital part of the process and all contributions to discussions are given equal consideration, regardless of whether they come from a longstanding community member or someone entirely new.

3.2 CMIP Community Engagement

A second example for user engagement is the approach of the [Coupled Model Intercomparison Project](#) (CMIP), a project of the World Climate Research Programme (WCRP) providing climate projections to understand past, present, and future climate changes. Under the guidance and at the direction of the Working Group on Coupled Modelling (WGCM), all CMIP activities are overseen by a coordinated pair of subcommittees with their objectives outlined in Terms of References: the CMIP Panel and the WGCM Infrastructure Panel (WIP). CMIP involves international modelling centres as

⁷ <https://stories.ecmwf.int/provider-page-climate-adaptation-digital-twin/index.html>

participants and provides information and data for a wide range of users, organisations and applications for example IPCC and Copernicus Climate Data Store (CDS). CMIP utilises a set of different standards like the Data Reference Syntax (DRS), a Controlled Vocabulary (CV; https://github.com/WCRP-CMIP/CMIP6_CVs) and NetCDF/CF and best practices in e.g. data versioning, documentation and citation. Therefore stakeholder engagement is complex.

1. Temporary [CMIP7 Tasks Teams \(TTs\)](#) are set up to investigate specific scientific and/or technical aspects of CMIP. They are composed of members of the Panels and experts from the community, which have applied through open calls. Thus in the TTs experts of specific fields support CMIP to shape the next phase CMIP7. CMIP7 TT Co-Leads are ex-officio members of the CMIP Panel and report one or both of Panels depending on their subjects.

2. A second set of CMIP community engagement activities are consultations with related organisations within WCRP like the Earth System Modelling and Observations (ESMO) and the Coordinated Regional Climate Downscaling Experiment (CORDEX) and outside of WCRP like IPCC and the [Earth System Grid Federation \(ESGF\)](#). The aim of these consultations are general (timeline) coordination and harmonisation/coordination of their goals and approaches. The ongoing revision of the WIP's Terms of Reference aims to establish a continuous exchange with these groups through liaison members.

3. The third set of user engagement activities is gathering feedback from modelling centres participating in CMIP and CMIP users through surveys and drop-in sessions organised by CMIP's International Project Office (CMIP-IPO). The CMIP-IPO supports the Panels and the TT, coordinates across these groups and with the wider CMIP community apart from documenting CMIP's progress and procedures more transparently than in the past.

Technically, the user engagement is supported by e.g. github features such as issue tracker, pull requests with integrated review step or the discussion boards. A practical example for a user engagement process is the baseline variable and data request development. In the first step a set of core or baseline variables was developed based on experience and information from CMIP6, IPCC (variable usage within the IPCC WGI AR6 provided by the DDC at DKRZ), the CMIP6 CDS and a user survey. The derived set of variables together with a suggestion of the procedure to maintain this list was developed. This initial list was discussed with the Panels and revised accordingly before feedback from the community was gathered through a drop-in session. The resulting baseline variable list has been taken up by the CMIP7 Data Request TT for integration into the Data Request development.

4 Gaps, Challenges and Opportunities

4.1 Sectorial Standardisation Approaches

Within climate services, a very large and diverse number of communities and sectors use data and software as described in the previous sections: historic and real-time observational (including both in-situ and remotely sensed data), climate projection (i.e. climate model data), reanalysis, and so forth at different space and time scales. Even if some commonalities can be found in the standardisation approaches, the diversity of usage leads to a need for a more sectorial standardisation approach, with a hierarchisation of the standard by community such as, for example, starting with a top level standardisation on file formats and very basic set of metadata for files (e.g. Dublin Core <https://www.dublincore.org/specifications/dublin-core/dces/>) to more advanced and specialised data and metadata per community (e.g. encoding of time in climate predictions). Having this “general to specific levels of standardizations” might help non-experts or new data users and producers obtain the relevant information for the standardisation more readily.

4.2 Identified Gaps and Challenges

Despite the very large number of conventions and standards, a gap identified is that a standard for validation of satellite observations is currently missing, which means that standards for other types of observation are instead being substituted for satellite observations, which, depending on the case, may lead to a lack of precision or adequacy.

Further, if the formats, vocabularies, metadata standards and other ontologies for data evolve, their evolution is usually not backwards compatible. However, service standards are still evolving quite fast, which makes it complicated to adhere to unique standards especially for a machine to machine communication between services.

Regarding FAIRness, several gaps (and associated recommendations) have been identified and are detailed in [sections 2.1.5](#) and [2.2.3](#)) but they can be summarised by mentioning that:

- FAIRness should apply to the whole life cycle of the data and software, and that preservation of the metadata (even if the data and software are not preserved) is essential;
- the link between data and metadata must be always preserved and this is not always the case when having the metadata stored separately from the files containing the data;
- readability of the metadata and FAIRness must be ensured both for humans and machines.

5 Conclusion

Climate services are about the generation, translation, transfer and use of climate information and knowledge (Vaughan and Dessai, 2014). These functions require data, software and infrastructure components that need to be standardised and persistent in order to provide reliable information to a range of different stakeholders. This makes a subset of climate service providers to be data providers, data (product) and software creators and service providers. Technical standards support the whole workflow from input data to the delivered data product. These include aspects of data and software management, service maintenance. Institutional commitments ensure service reliability and sustainability.

The preliminary deliverable on good practices for data, software and technical standards reflects this complexity but also shows the diversity of good practices applied. Practices vary from sector to sector and depend on the climate services provider, e.g. whether it is based in a public authority or a research institution. Some commonalities have been identified across the diversity of practices:

1. FAIR (Findable, Accessible, Interoperable and Reusable) principles for data and software have become guiding principles outside the technical data and software developer community. They have been complemented by implementation guidelines emphasising the roles of PIDs and metadata standards.
2. WMO Standards are central for climate service providers, providing guidelines on many different aspects ranging from data specifications in the GRIB guideline to the Guidelines on Quality Management in Climate Services.
3. WMO standards are complemented by bottom-up developments of good practices within the climate research community such as in WCRP CMIP and by EU regulations such as the INSPIRE directive. Technical standards and recommendations from technical authorities such as RDA, OGC and DataCite but also good scientific practices in terms of citation and licensing of data and software are taken up for the evolution of these good practices for climate service providers.
4. Quality management standards and institutional commitments are important for a reliable climate service delivering high-quality products. User engagement is essential for the development of climate services that meet the evolving needs of different sectors.

The current large number of communities with their evolving practices in data and software can be prioritised in a hierarchical approach, starting with the implementation of general standards, which can be gradually improved to more detailed and sectorial-specific standards. Standards often lack backwards compatibility, which poses a challenge to the operational use of standards. Therefore, user engagement in a collaborative development of standards is crucial to improve usability and practicality, in addition to taking into account new developments in data science. For this reason, the authors welcome feedback and experiences from climate service providers with standards for inclusion in the revised final version of this task on data and software standards for climate services.

References

Bakator M and Cockalo D (2018): Improving business performance with ISO 9001: A review of literature and business practice. *The European Journal of Applied Economics*. 15. 83-93. [10.5937/EJAE15-16145](https://doi.org/10.5937/EJAE15-16145).

Barker, M., Chue Hong, N. P., Katz, D. S., Lamprecht, A. L., Martinez-Ortiz, C., Psomopoulos, F., ... & Honeyman, T. (2022). Introducing the FAIR Principles for research software. Scientific Data, 9(1), 622.
<https://www.nature.com/articles/s41597-022-01710-x>

Brase, Jan (2019). 10 years of DataCite – How it all began. <https://doi.org/10.5438/g4x3-r450>.

Bravi L, Murmura F, and Santos G (2019): The ISO 9001: 2015 quality management system standard: Companies' drivers, benefits and barriers to its implementation. *Quality Innovation Prosperity*, 23(2), 64-82.

Christophersen, Allan. Colón-Marrero, Elena. Dietrich, Dianne. Falcao, Patricia. Fox, Claire. Hanson, Karen. Kwan, Allen. McEniry, Matthew. (2023, October 17). Software Metadata Recommended Formats Guide. Software Preservation Network. <https://doi.org/10.5281/zenodo.10001787>

Chue Hong, N. P. et al. (2022). FAIR Principles for Research Software (FAIR4RS Principles) (1.0). Zenodo. <https://doi.org/10.15497/RDA00068>

Cortekar J, Themessl M, Lamich K, 2020: Systematic analysis of EU-based climate service providers. *Climate Services*, Volume 17, 2020, 100125, ISSN 2405-8807, <https://doi.org/10.1016/j.cliser.2019.100125>.

CoreTrustSeal Standards and Certification Board. (2022). CoreTrustSeal Requirements 2023-2025 (V01.00). Zenodo. <https://doi.org/10.5281/zenodo.7051012>

DataCite Metadata Working Group. (2017). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data. Version 4.1. DataCite e.V. <http://doi.org/10.5438/0014>.

DataCite Metadata Working Group. (2024a). DataCite Metadata Schema Documentation for the Publication and Citation of Research Data and Other Research Outputs. Version 4.5. DataCite e.V. <https://doi.org/10.14454/g8e5-6293>

DataCite Metadata Working Group. (2024b). DataCite Metadata Schema for the Publication and Citation of Research Data and Other Research Outputs. Version 4.5. DataCite e.V. <https://doi.org/10.14454/g8e5-6293>

Data Citation Synthesis Group (2014): Joint Declaration of Data Citation Principles. Martone M. (ed.) San Diego CA: FORCE11; 2014 <https://doi.org/10.25490/a97f-egyk>

del Pico, E. M., Gelpí, J. L., & Capella-Gutiérrez, S. (2022). FAIRsoft-A practical implementation of FAIR principles for research software. *bioRxiv*, 2022-05. <https://www.biorxiv.org/content/10.1101/2022.05.04.490563v2.full.pdf>
https://docs.google.com/document/d/1OcGxtX-LqloePSeEE-S1v3kgKXe6wiCgo_VMTF6idhM/edit

Eaton B., Gregory J., Drach B., Taylor K., Hankin S., Blower J., Caron J., Signell R., Bentley P., Rappa G., Höck H., Pamment A., Juckes M., Raspaud M., Horne R., Whiteaker T., Blodgett D.,

Zender C., Lee D., Hassell D., Snow A.D., Kölling T., Allured D., Jelenak A., Soerensen A.M., Gaultier L., Herlédan S., Manzano F., Barring L., Barker C., Bartholomew S., 2023: NetCDF Climate and Forecast (CF) Metadata Conventions, version 1.11. <https://cfconventions.org/Data/cf-conventions/cf-conventions-1.11/cf-conventions.html>.
<https://cfconventions.org/Data/cf-conventions/cf-conventions-1.11/cf-conventions.pdf>.

Fenner, Martin (2018). DOI Registrations for Software. <https://doi.org/10.5438/1nmy-9902>

Fonseca LM, Domingues JP, Baylina-Machado P and Harder D. (2019): ISO 9001:2015 adoption: A multicountry empirical research. *Journal of Industrial Engineering and Management*, 12(1), 27-50.

Huard, D., Pirani, A., Chen, R., Gutiérrez, J. M., Juckes, M., Krey, V., Spinuso, A., & Stockhause, M. (2022). IPCC Data and Code Licensing Guidelines (1.1). Zenodo. <https://doi.org/10.5281/zenodo.7431834>

ICAO, 1963: Working Arrangements Between ICAO and WMO

ICAO, 2010: Annex 3 to the Convention on International Civil Aviation, Meteorological Service for International Air Navigation

Jones, C.M.; Matthews, B.; Gent, I.; Griffin, T.; Tedds, J. (2017) Persistent Identification and Citation of Software. *Int. Journal of Digital Curation*, Vol.11 No.2. <https://doi.org/10.2218/ijdc.v11i2.422>

Lamprecht, A. L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., ... & Capella-Gutierrez, S. (2020). Towards FAIR principles for research software. *Data Science*, 3(1), 37-59. DOI: [10.3233/DS-190026](https://doi.org/10.3233/DS-190026)

Laurence, B. N., Balaji, V., Bentley, P., Callaghan, S., DeLuca, C., Denvil, S., Devine, G., Elkington, M., Ford, R. W., Guilyardi, E., Lautenschlager, M., Morgan, M., Moine, M.-P., Murphy, S., Pascoe, C., Ramthun, H., Slavin, P., Steenman-Clark, L., Toussaint, F., Treshansky, A., and Valcke, S.: Describing Earth system simulations with the Metafor CIM, *Geosci. Model Dev.*, 5, 1493–1500, <https://doi.org/10.5194/gmd-5-1493-2012>, 2012.

Yann Le Franc, Luiz Bonino, Hanna Koivula, Jessica Parland-von Essen, & Robert Pergl. (2022). D2.8 FAIR Semantics Recommendations Third Iteration (V1.0). Zenodo. <https://doi.org/10.5281/zenodo.6675295>

Magagna, B., Moncoiffé, G., Devaraju, A., Stoica, M., Schindler, S., Pamment, A., & RDA I-ADOPT WG. (2022). Interoperable Descriptions of Observable Property Terminologies (I-ADOPT) WG Outputs and Recommendations (1.1.0). Zenodo. <https://doi.org/10.15497/RDA00071>

Matentzoglou, Nicolas, James P Balhoff, Susan M Bello, Chris Bizon, Matthew Brush, Tiffany J Callahan, Christopher G Chute, William D Duncan, Chris T Evelo, Davera Gabriel, John Graybeal, Alasdair Gray, Benjamin M Gyori, Melissa Haendel, Henriette Harmse, Nomi L Harris, Ian Harrow, Harshad B Hegde, Amelia L Hoyt, Charles T Hoyt, Dazhi Jiao, Ernesto Jiménez-Ruiz, Simon Jupp, Hyeongsik Kim, Sebastian Koehler, Thomas Liener, Qinqin Long, James Malone, James A McLaughlin, Julie A McMurry, Sierra Moxon, Monica C Munoz-Torres, David Osumi-Sutherland, James A Overton, Bjoern Peters, Tim Putman, Núria Queralt-Rosinach, Kent Shefchek, Harold Solbrig, Anne Thessen, Tania Tudorache, Nicole Vasilevsky, Alex H Wagner, Christopher J Mungall, A Simple Standard for Sharing Ontological Mappings (SSSOM), Database, Volume 2022, 2022, baac035, <https://doi.org/10.1093/database/baac035.0> .

Moine, M.-P., S. Valcke, B. N. Lawrence, C. Pascoe, R. W. Ford, A. Alias, B. Balaji, P. Bentley, G. Devine, S.A.Callaghan, and E. Guilyardi. Development and exploitation of a controlled vocabulary in support of climate modelling. *Geosci. Model Dev.*, 7, 479-493, (2014) doi: 10.5194/gmd-7-479-2014

Pascoe, C. L., Dankers, R., Muñoz, Á.G. (2024), Preliminary best practices in climate uncertainty quantification and communication, D2.3 of the Climateurope2 project.

Pascoe, Charlotte & Lawrence, Bryan & Guilyardi, Eric & Jukes, Martin & Taylor, Karl. (2020). Documenting numerical experiments in support of the Coupled Model Intercomparison Project Phase 6 (CMIP6). *Geoscientific Model Development*. 13. 2149-2167. 10.5194/gmd-13-2149-2020.

Petrie, R., Denvil, S., Ames, S., Levavasseur, G., Fiore, S., Allen, C., Antonio, F., Berger, K., Bretonnière, P.-A., Cinquini, L., Dart, E., Dwarakanath, P., Druken, K., Evans, B., Franchistéguy, L., Gardoll, S., Gerbier, E., Greenslade, M., Hassell, D., Iwi, A., Jukes, M., Kindermann, S., Lacinski, L., Mirto, M., Nasser, A. B., Nassisi, P., Nienhouse, E., Nikonov, S., Nuzzo, A., Richards, C., Ridzwan, S., Rixen, M., Serradell, K., Snow, K., Stephens, A., Stockhouse, M., Vahlenkamp, H., and Wagner, R.: Coordinating an operational data distribution network for CMIP6 data, *Geosci. Model Dev.*, 14, 629–644, <https://doi.org/10.5194/gmd-14-629-2021>, 2021.

Rickards, L., Vardigan, M., Diepenbroek, M., Dillo, I., Genova, F., L'hours, H., Hugo, W., Khodiyar, V., Li, G., Minster, J.-B., Trilsbeek, P., Wyborn, L., Edmunds, R., Mokrane, M., & Panagou, E. (2016). DSA–WDS Partnership Working Group Catalogue of Common Procedures for Certification (1.2). Zenodo. <https://doi.org/10.15497/RDA00019>

Smith AM, Katz DS, Niemeyer KE, FORCE11 Software Citation Working Group. 2016. Software citation principles. *PeerJ Computer Science* 2:e86 <https://doi.org/10.7717/peerj-cs.86>

S. Stall, G. Bilder, M. Cannon, N. Chue Hong, S. Edmunds, C. C. Erdmann, M. Evans, R. Farmer, P. Feeney, M. Friedman, M. Giampoala, R. B. Hanson, M. Harrison, D. Karaikos, D. S. Katz, V. Letizia, V. Lizzi, C. MacCallum, A. Muench, K. Perry, H. Ratner, U. Schindler, B. Sedora, M. Stockhouse, R. Townsend, J. Yeston, T. Clark, *Journal Production Guidance for Software and Data Citations, Scientific Data* v.10, 656, 2023. <https://doi.org/10.1038/s41597-023-02491-7>

J. Starr (2027). New DataCite Metadata Updates Support Software Citation. <https://doi.org/10.5438/nzhx-xx96>.

Karl E. Taylor, Martin Jukes, V. Balaji, Luca Cinquini, Sébastien Denvil, Paul J. Durack, Mark Elkington, Eric Guilyardi, Slava Kharin, Michael Lautenschlager, Bryan Lawrence, Denis Nadeau, and Martina Stockhouse (2018) CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's. Version: 10 September 2018 (v6.2.7). <https://goo.gl/v1drZl>

Vaughan, Catherine; and Dessai, Suraje (2014). Climate services for society: origins, institutional arrangements, and design elements for an evaluation framework. *WIREs Clim Change* 2014, 5:587–603. <https://doi.org/10.1002/wcc.290>

WMO (2016). "The sixteenth session of the Commission for Basic Systems (CBS-16) (Guangzhou, China, 23-29 November 2016) decided to remove the GRIB edition 1 from the Manual on Codes, Volume I.2 (WMO-No. 306) and to ensure it remains available as follows (Decision 7 (CBS-16)). ... Decides to remove GRIB edition 1 from the Manual on Codes (WMO-No. 306), Volume I.2; Requests the Secretary-General to ensure that the specification of FM 92-XI Ext. GRIB (GRIB edition 1) remains permanently available so that archived data can still be referenced." <https://community.wmo.int/en/activity-areas/wis/grib-edition-1>.

WMO (2022). Manual on Codes, Volume I.2 – International Codes. Annex II to the WMO Technical Regulations. Part B – Binary Codes, Part C – Common Features to Binary and Alphanumeric Codes. World Meteorological Organization (WMO), 2019 edition, updated in 2022. WMO-No.306 ISBN 978-92-63-10306-2. <https://library.wmo.int/idurl/4/35625>.

WMO, 2017a: Guide to the Implementation of a Quality Management System for National Meteorological and Hydrological Services (WMO-No. 1100). Geneva.

WMO, 2017b: Technical Regulations Volume I – General Meteorological Standards and Recommended Practices (WMO-No. 49), Geneva

WMO, 2018a: Guidelines on Quality Management in Climate Services (WMO-No. 1221). Geneva.

WMO, 2018b: Guide to Instruments and Methods of Observation, Volume V – Quality Assurance and Management of Observing Systems (WMO-No. 8), Geneva

WMO, 2018c: Technical Regulations Volume II – Meteorological Service for International Air Navigation (WMO-No. 49), Geneva

WMO, 2019: Compendium of WMO Competency Frameworks (WMO-No. 1209), Geneva.

WMO, 2022: Guidelines for the Assessment of Competencies for Provision of Climate Services WMO-No. 1285), Geneva.