Climateurope2

Common procedures for data verification, certification and quality management

Deliverable 2.9

Authors: Xavier Domingo, Fanny Adloff, Daniel Arteaga, Anna-Lena Flügel, Tjerk Krijger, Aleksandra Krzic, Julien Lehnhardt, Dragan Mihic, Nakiete Msemo, Christian Pagé, Alessandro Spinuso

Contributors from D2.4: Francisco J. Doblas-Reyes, Javier Diez-Sierra, Fai Fung, Birgit Hassler, Lina Sitz

Contributors from D2.9: Victòria Agudetse, Ilaria Gallo, Katherine Grayson, Nils Hempelmann, Andrea Lammert, Guillaume Levavasseur, Ed Pechorro, Rob Roebeling, Francesc Roura, Marlies van der Schee, Paul van Schayck, Peiliang Shi, Gerard van der Schrier, Christine Träger

Reviewers: Carlo Buontempo (D2.4), Andre Obregon (D2.9)

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	2022/09/01
RELATED WORK PACKAGE	WP2 Data and processes
RELATED TASK(S)	T2.3
LEAD ORGANISATION	BSC
AUTHORS	Xavier Domingo, Fanny Adloff, Daniel Arteaga, Anna-Lena Flügel, Tjerk Krijger, Aleksandra Krzic, Julien Lehnhardt, Dragan Mihic, Nakiete Msemo, Christian Pagé, Alessandro Spinuso
SUBMISSION DATE	2026-04-30
DISSEMINATION LEVEL	PU

History

DATE	SUBMITTED BY	REVIEWED BY	VISION (NOTES)
2026-03-31	Xavier Domingo	André Obregon	Internal review
2026-03-28	Xavier Domingo	All WP2 partners	All comments from the internal review were addressed. The document is ready to be submitted.

Please cite this report as: Domingo, X., Adloff, Arteaga, D., Flügel, A., Krijger, T., Krzic, A., Lehnhardt, J., Mihic, M., Msemo, N., Pagé, C. & Spinuso, A. (2026) Common procedures for data verification, certification and quality management. Deliverable D2.9 of the Climateurope2 project.

Disclaimer 1: 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.

Disclaimer 2: LLM models were used to refine the flow and readability of some sections in this report. The content was always double-checked and edited by authors.

Table of Contents

List of tables

List of figure

Acronyms

Acronym	Definition
ACDD	Attribute Conventions for Data Discovery
AFT	Assessment Fast Track
AOA	Area of Applicability
API	Application Programming Interface
AR	Assessment Report
C3S	Copernicus Climate Change Service
C3S-CDS	Copernicus Climate Change Service - Climate Data Store
CF	Climate and Forecast
CI	Continuous Integration
CIM	Climate Information Model
CMEC	Coordinated Model Evaluation Capabilities
CMIP	Coupled Model Intercomparison Project
CMOR	Climate Model Output Rewriter
CORDEX	Coordinated Regional Climate Downscaling Experiment
CRAN	The Comprehensive R Archive Network
DECK	Diagnostic, Evaluation and Characterization of Klima
DOI	Digital Object Identifier
DSMM	Data Stewardship Maturity Matrix
ECV	Essential Climate Variable
EOSC	European Open Science Cloud
ES-DOC	Earth System Documentation
ESGF	Earth System Grid Federation
ESM	Earth System Model
FAIR	Findable, Accessible, Interoperable, and Reusable

Acronym	Definition
FIP	FAIR Implementation Profiles
GCM	General Circulation Model
HVD	High-Value Datasets
IEC	International Electrotechnical Commission
ILAMB	International Land Model Benchmarking
INSPIRE	Infrastructure for Spatial Information in the European Community
IOMB	International Ocean Model Benchmarking
IPCC	Intergovernmental Panel on Climate Change
ISO	International Organization for Standardization
LLM	Large Language Model
NetCDF	Network Common Data Form
OGC	Open Geospatial Consortium
OOS	Out-of-sample
P2P	Peer-to-Peer
PID	Persistent IDentifier
PMP	PCMDI Metrics Package
PROV-DM	Provenance Data Model
QA	Quality Assurance
QAE4O	Quality Assurance Framework for Earth Observation
QC	Quality Control
RCM	Regional Climate Model
RCP	Representative Concentration Pathways
SAF	Satellite Application Facilities
SSP	Shared Socio-economic Pathway
TIB	Technische Informationsbibliothek

Acronym	Definition
UNFCC	United Nations Framework Convention on Climate Change
URL	Uniform Resource Locator
W3C	World Wide Web Consortium
WCRP	World Climate Research Programme
WDCC	World Data Center for Climate
WG	Working Group
WIGOS	WMO Integrated Global Observing System
WIS	WMO Information System
WMO	World Meteorological Organization

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 and appropriate. A large variety of activities to support the communities involved in European Climate Services will also be organised.

Executive Summary

This report presents a set of **recommendations** about verification, certification and quality management that aim to inform **the standardisation process of Climate Services**. The background knowledge used to write these recommendations was collected through an extensive literature review, interviews with representatives of both the public and private sector, and collaborators who willingly shared their expertise with the writing team. The recommendations presented in this report are broadly consistent with the WMO Guidelines on Quality Management in Climate Services (WMO-No. 1221), and thus represent an added value to the existing international guidelines.

The final recommendations, which are described in Section 2 in the formal language of international standards are:

- R1: A systematic user requirements process should be established to ensure service fitness-for-purpose
- R2: Quality Assurance and Quality Control processes shall be embedded across the entire lifecycle of Climate Services
- R3: Certification standards serve as essential instruments for quality, trust, and transparency
- R4: Quality communication shall be precise and transparent
- R5: Robust Feedback Loops and Support Lifecycles should be established
- R6: Quality metadata shall be provided
- R7: Traceability shall be integrated into QA/QC workflows for data and model evaluation
- R8: A community driven FAIRness implementation and verification tool should be developed
- R9: All quality controlled data and tools shall be accompanied by a fully reproducible validation report
- R10: The area of applicability shall be reported
- R11: A set of minimum quality criteria should be standardised to democratise discoverability

While the focus of this document has been on the technical quality of data and tools, we have provided some guidance on other aspects like scientific quality and fitness-for-purpose. Both of these aspects are essential and deserve special attention when it comes to standardisation and certification. Therefore, we would like to encourage future dedicated investigations around them.

Additionally, the presented list of case studies from the public community could largely benefit from including operational Climate Services by National Meteorological and Hydrological Services.

Keywords

Climate Services, Quality Control, Quality Assurance, Communication, Quality Management, Verification, FAIR principles

1 Introduction

The aim of this report is to inform a potential standardisation process for Climate Services by providing recommendations related to **verification, certification and quality management**. Climate Services involve the provision of Climate information in such a way as to assist decision-making. They include appropriate engagement from users and providers, are based on scientifically credible information and expertise, have an effective access mechanism and respond to user needs (Hewitt et al., 2012). A schematic of the Climate Services' lifecycle is represented in Figure 1. All steps should be understood as interconnected, and that since Climate Services are created in an iterative fashion, the elements shown in the schematic are sometimes not followed in a linear sequence.

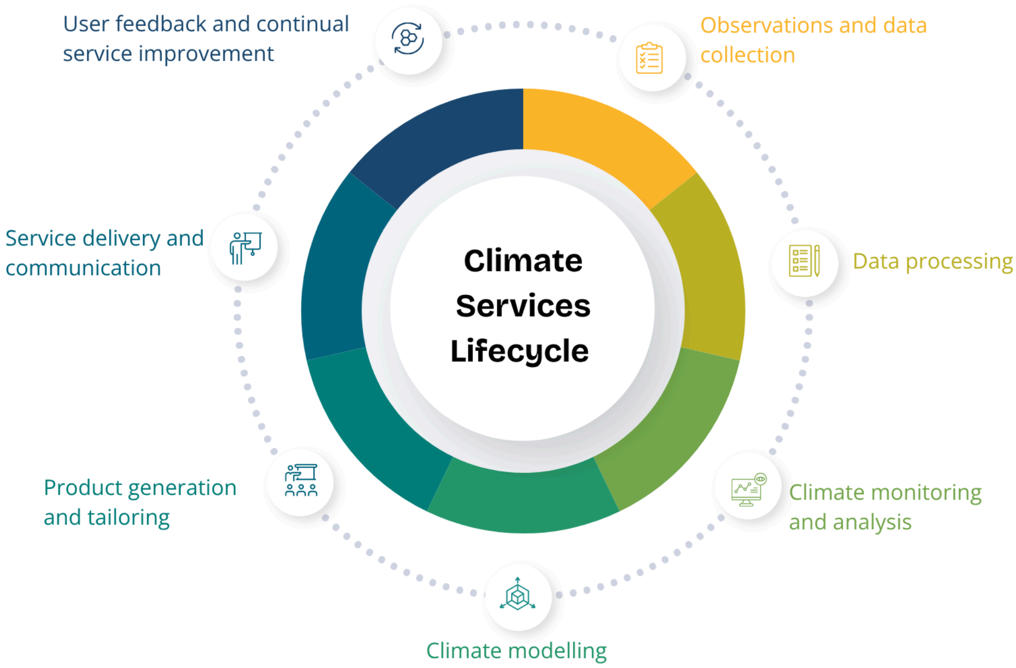


Figure 1: Schematic representation of the Climate Services' lifecycle. All items have implicit QA/QC of the process and products within.

Following ISO 9000:2015 while **Quality Assurance** (QA) is the part of quality management focused on providing confidence that quality requirements will be fulfilled, **Quality Control** (QC) is focused on fulfilling quality requirements. In other words, while QA involves setting up standards, choosing the right tools, and designing workflows to ensure that mistakes don't happen in the first place and quality requirements are met, QC refers to the checks and tests used to catch any errors or non-compliance that slipped through the QA process. Together, they are known as QA/QC processes.

The **highlight of this report** is the list of recommendations presented in **Section 2**. The rest of the report provides background material collected by the writing team, and is here listed for the interested reader to explore further :

- **Section 3** presents the **FAIR principles,** what they are and why they are important in data and software management. However, the main focus of this section is to discuss the vagueness of the FAIR principles, and the necessity for agreed implementations and metrics, so that FAIRness can be provided and verified consistently.

- **Section 4** consists of a collection of **case studies** from the public sector within the Climate Services community that illustrate different QA/QC methodologies.

- **Section 5** deals with how **quality information is stored, traced and communicated**, including the interaction between users and developers.

- **Section 6** addresses the **certification** landscape of quality within the Climate Services community. It includes institutional and personnel certifications, and labels.

- **Section 7** features the insights collected through a set of **interviews** with a set of representatives from both the public and private sector.

The recommendations presented in this report are broadly consistent with the WMO Guidelines on Quality Management in Climate Services (WMO-No. 1221). In particular, the emphasis on verification, traceability, documentation, user communication, and continual improvement reflects key quality management principles relevant to the provision of Climate Services. This linkage helps situate the report's recommendations within a recognised international framework for quality management in Climate Services.

The information presented here is far from complete and it keeps evolving, but we hope it can already trigger interesting conversations in the standardisation domain.

Finally, it should be mentioned that this report is part of a set of Deliverables presented by the Work Package 2 team of the Climateurope2 project, and these are hereby listed as reference for the interested reader:

- **Deliverable D2.6**: Guidelines of good practices for vocabularies, formats, metadata and technical standards

- **Deliverable D2.7**: Good practices in provenance of processing methodologies and ways to communicate it in a user-friendly manner

- **Deliverable D2.8**: Guidelines for Climate uncertainty quantification and communication

- **Deliverable D2.9**: Common procedures for data verification, certification and quality management

- **Deliverable D2.10**: Guidelines on how to enhance data, tools and knowledge for Climate impacts, extremes and risks through standardisation of Climate Services

2 Recommendations

This section presents the recommendations from the Work Package 2 Task 2.3 team on verification, certification and quality management. The language used here follows the practices in international standards: “shall” is used to express a strict mandatory requirement, “should” indicates a strong recommendation and “can” states a possibility or an action that is permissible.

R1: A systematic user requirements process should be established to ensure service fitness-for-purpose

Quality in Climate Services is a multi-dimensional concept that shall be defined collaboratively by all actors involved in the service chain, i.e. via co-production. To ensure that technical outputs translate into actionable insights, providers should implement a structured requirements management lifecycle. This process ensures the service remains aligned with evolving user needs and maintains its relevance over time. Some key points are:

- **Requirement Elicitation**: Providers should regularly engage with a representative user base to identify specific needs.
- **Co-definition of Quality Metrics**: Technical quality benchmarks should be defined in partnership with end-users. These metrics shall reflect not only scientific accuracy but also the operational characteristics required for the user’s specific application.
- **Iterative Validation and Co-Design**: The lifecycle should include formal validation checkpoints to verify that the service’s outputs are understandable and fit for purpose.
- **Traceability of Features**: All core functionalities and datasets within the service should be traceable to a documented user requirement.
- **Requirement Prioritisation and Governance**: For long-term climate services, a formal mechanism for co-governance should be established, so that users can help prioritise the development roadmap based on shifting requirements.
- **Fitness-for-Purpose Statement**: Providers should publish a fitness-for-purpose statement co-authored with representative users to describe the specific context in which the service is reliable and any scenarios where the data is not intended for use.

R2: Quality Assurance and Quality Control processes shall be embedded across the entire lifecycle of Climate Services

Quality Assurance and Quality Control (QA/QC) processes shall not be treated as end-stage validation steps, but rather embedded systematically throughout the entire lifecycle of Climate Services, from the technical implementation to the scientific validation and user testing in operational environments. Keeping QA/QC processes up to date and well-maintained is essential for developing sustainable and scalable services that evolve with user needs, remain fit-for-purpose, and comply with current standards and best-practices. This end-to-end perspective is also consistent with the WMO quality management approach for Climate Services, which emphasises process management across the full chain from observations and data handling to service delivery and user engagement

In addition, applying QA/QC at all stages of the life-cycle enables quality traceability, which is invaluable to efficiently identify sources of non-compliance and correct them.

QA/QC processes should be properly documented. This includes keeping track of what requirements QA/QC is aiming to fulfill, what metrics are being used to measure success, and how these are implemented. Moreover, a versioning system should be integrated to keep track of how the QA/QC and all this information evolves in time.

R3: Certification standards serve as essential instruments for quality, trust, and transparency

By guaranteeing compliance with recognised best practices, certifications increase both the robustness and long-term sustainability of a service. For instance, institutional standards like ISO 9001 (quality management) and ISO 27001 (information security) allow providers to demonstrate a high level of care for user needs and data. However, these institutional certifications often impose significant financial and time burdens on candidates.

Beyond the institutional level, personnel certifications validate individual competence and adherence to professional standards. These credentials not only support career development and professional recognition but also connect experts to trusted networks for collaboration.

Within the European context, establishing a personnel certification for weather and Climate information providers is essential. By formalising the requirements for those professionals delivering Climate information, this certification would bolster user trust and foster a community of highly-skilled, accredited professionals.

Finally, specialised labels can be obtained for specific tools, repositories or climate service components. While these lack the heavy overhead of institutional certifications, they remain highly valuable within communities that recognise and endorse them.

R4: Quality communication shall be precise and transparent

Vague language in quality reporting may lead to indecision, create confusion, and erode trust. Buzzwords aren't just misleading to users, they also create an unfair playing field for competitors. To ensure accountability, quality shall be communicated through a framework of evidence, not intuition, and in a manner that responds to the needs and expectations of the intended users.

- **Definition**: All core concepts and metrics shall be explicitly defined in a findable, referenceable and accessible glossary.
- **Honesty (no overpromising)**: Applicability and limitations shall be as prominent as features, and displayed accordingly.
- **Quantification**: Subjective descriptors (e.g., "fast", "robust", "accurate") shall be anchored to quantitative benchmarks.
- **Verifiability**: Comparisons with competitors or internal baselines shall be quantitative, scope-matched, and reproducible by an independent third party.

R5: Robust Feedback Loops and Support Lifecycles should be established

Co-production is an essential aspect of building Climate Services, and co-production can't happen without effective communication channels between providers and users to ensure that unforeseen errors are identified, user needs are met, and the limitations of a tool are understood in practice.

The requirements for a robust feedback loop and support lifecycle are:

- **Bi-directional channels**: Providers must establish clear, accessible feedback mechanisms (e.g., helpdesks, issue trackers, or dedicated contact points). These channels must be explicitly linked within the dataset or tool metadata to ensure they are findable at the point of use.
- **Knowledge Retention**: All user interactions, bug reports, and resolutions should be documented and archived. This information should be maintained as long as the tool is in use to prevent the repetition of errors and to serve as a searchable knowledge base for other users.
- **Support Transparency:** The level of support (e.g., "active development," "maintenance only," or "community supported") must be clearly stated.
- **End-of-Life Protocol:** When a tool or dataset reaches its sunset phase, providers must give advance notice. This communication should include the date support will cease and, if possible, recommendations for alternative tools or "legacy" usage instructions.
- **Service Level Expectations**: Providers should state expected response times. Even if a fix isn't immediate, an acknowledgement of the issue helps maintain user trust.
- **Shared Bug Tracker**: Whenever possible, known issues and their status (Open/Resolved) should be accessible within the community of users. This prevents redundant support requests and allows users to self-correct based on known tool limitations.
- **Closing the Loop**: Providers should proactively communicate how user feedback has been implemented in new versions or updates, demonstrating that the quality control process is active and responsive.
- **Privacy and Data Protection**: Data retention policies for user interactions must comply with relevant privacy and security regulations (e.g., GDPR, ISO 27001).

R6: Quality metadata shall be provided

Quality metadata is not an optional add-on; it is a fundamental component of the data itself. When metadata is detached, the data loses its context, leading to potential misuse and the inability to post-process the data. To support the creation of reliable derived products, quality information must remain persistently linked to its parent dataset. Providers shall:

- **Couple metadata persistently**: Metadata shall be stored within the same file as the data (e.g., in file headers) or via strictly linked sidecar files to prevent accidental separation.
- **Follow standardised practices and conventions**: Use industry-standard for content, format and controlled vocabularies to ensure cross-platform interoperability. For example, ISO 19157 on data quality for geographic information for metadata content, and the Climate and Forecast convention for controlled vocabularies.

- **Methodological Traceability**: All quality derivations shall be fully documented and referenced within the metadata using Persistent Identifiers (PIDs), such as DOIs, to ensure the "recipe" for quality is always findable and accessible.

R7: Traceability shall be integrated into QA/QC workflows for data and model evaluation

Reliable quality assessment and sustained quality monitoring depend on the ability to trace how data are produced, transformed, evaluated, and maintained over time. Without traceability, quality claims remain opaque, difficult to verify, and hard to compare across datasets, organisations, or processing histories. Effective traceability provides the evidential backbone that links measurements to reference standards, documents the assumptions and decisions embedded in processing workflows, and records how quality characteristics evolve through reprocessing and stewardship activities. It enables users to judge fitness for purpose, supports interoperability and reuse, and allows data providers to monitor, diagnose, and improve quality in a systematic way. In this sense, the scope of traceability includes all elements inside a pipeline that are necessary to reproduce it, for example, data sources, model versions, and processing workflows.

Developers shall ensure that traceability is integrated into QA/QC workflows. This information should be open to increase transparency, and if possible, embedded inside the metadata. However, if proprietary knowledge restrictions don't allow it, detailed access to such information shall be regulated by additional policies, so that pipelines can be easily audited when necessary.

R8: A community driven FAIRness implementation and verification tool should be developed

The FAIR principles stand for Findable, Accessible, Interoperable and Reusable, and they currently represent an essential guidance on how to manage software and data.

Given the vague nature of the FAIR principles, implementation tools form the foundation of FAIRness as they translate principles into technical solutions. Climate service providers should embed **FAIR implementation tools** directly into their workflows to ensure that structured metadata, provenance information, persistent identifiers, machine-accessible interfaces and more are implemented during data and software production.

Unfortunately, currently there is no agreed complete solution within the Climate Services community on how to implement these principles. Nevertheless, some solutions are published openly as FAIR Implementation Profiles (FIP). For Interoperability and Reusability in particular, it is important for the Climate service community to agree on controlled vocabularies, common metadata schemas, provenance standards and versioning practices for datasets and software.

FAIR verification tools are used to monitor FAIR compliance through automatic assessments. However, they are limited by the choice of implementation they verify and the degree of "machine-readability" of the information they analyse, thus making generalisable verification tools currently not possible.

R9: All quality controlled data and tools shall be accompanied by a fully reproducible validation report

To ensure trust, quality controlled data and tools should be distributed alongside a fully reproducible validation report[1]. To meet this recommendation, the validation report **shall**:

- **Prioritise Transparency**: Ensure that assumptions, limitations, and disclaimers are prominently disclosed and easily findable. These should include a meta-analysis of the validation itself.
- **Bridge the Expertise Gap**: Be structured for dual-readability, offering a clear narrative for a broad audience while maintaining exhaustive technical depth for specialists.
- **Guarantee Reproducibility**: Include all necessary datasets and diagnostic tools required for a third party to independently replicate the validation results.

And **should**:

- **Be published openly**: While proprietary restrictions can limit data transparency, they should never compromise the integrity of the validation process. By publishing validation reports alongside the datasets and tools needed for reproduction, private actors can participate in transparent benchmarking and peer analysis. This allows companies to build community trust without sacrificing their core proprietary knowledge.

R10: The area of applicability shall be reported

Models (or ensembles of them) are inherently limited by their training datasets and parameterisation. These characteristics determine what the model can learn. Thus, a key analysis of a model is defining the Area of Applicability (AOA), that is the range and combination of features that the model has been trained on, and thus under which it is expected to provide valid results. This way, when the model is applied to new data, trust on its prediction can be quantified based on where the new data lies with respect to the AOA. In the literature, this is also referred to as analysing data novelty or dissimilarity with respect to data used in training.

Given the limitations of historical data in terms of observed features and patterns, some models are also trained on simulated data. Independently of the origin of the training data or how well the model performs on the training dataset, the reliability of a model is expected to degrade when encountering new data not statistically represented during training, i.e. out-of-sample (OOS) data.

Regarding models, to ensure transparency and build trust, developers shall:

- **Define the Area of Applicability:** Clearly state the statistical boundaries of the training data, both of features individually and their combinations. In addition, report what metric is used to define the AOA, its interpretation and its limitations
- **Warn Users**: Explicitly warn users when a model is applied to OOS data. This includes "black swan" or extreme events.
- **Quantify Generalisation**: In case the model is meant to be used for OOS datasets, include OOS performance tests to measure error margins as data deviates from the training set. These tests should reflect real-world operational conditions.

In turn, the applicability of data products is influenced by the underlying assumptions of processing algorithms and/or existing uncertainties in the raw data. Developers shall report the limitations,

[1] Even though this terminology is more linked to satellite Earth observation datasets, this recommendation is applicable to all tools and datasets related to Climate Services.

caveats and intended use of datasets, thus enabling users to make informed decisions about their fitness-for-purpose in their specific application context.

R11: A set of minimum quality criteria should be standardised to democratise discoverability

As the number of services increases, more tools are being developed to streamline the discovery process. Large Language Models (LLMs), for instance, have become a primary means of identifying datasets and exploring new research avenues.

However, these tools face two significant challenges:

1. **Discoverability**: Resources are not uniformly accessible, leading search engines (both AI-driven and traditional) to rank them in an opaque order. Furthermore, users lack the capacity to analyse hundreds of resources simultaneously, typically limiting their review to a small, preselected subset.

2. **Comparability**: Providers currently document resource quality inconsistently. In the absence of a common standard, datasets cannot be compared automatically by machines—or even objectively by human experts.

To democratise dataset discovery, a set of minimum quality criteria should be established. Ideally, these criteria should be machine-readable and follow standard vocabularies to enable automated processing. While individual organisations could adopt these recommendations internally, achieving truly democratic discovery and comparability within the European context—spanning both the public and private sectors—requires a transparent and standardised framework for quality metrics, and perhaps even a centralised catalogue of services.

3 FAIR principles

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.

The impact of Climate Services depends on the trustworthiness of its output datasets including, for example, being able to trace back the full input data flow, its quality, as well as processing software and its settings, by the end-user. This last point covers the provenance of the data and is especially important for re-use of data. Many metadata elements are required for this, but especially some of the specific attributes checked under the Interoperability and Reusability. The sections below cover the FAIRness principles in more detail and explain how they could be met in practice via recommendations & implementation tools and tested via verification tools.

3.1 FAIR metrics

As described in D2.1 (https://zenodo.org/records/10912719), 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/). These principles define what it means for data to be Findable, Accessible, Interoperable and Reusable but do **not** prescribe solutions.

The path to FAIR data can therefore 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. The FAIR principles themselves can be used as the first generic metrics to evaluate the FAIRness of data from the Climate Services and help to have a common target and means to validate how far a data provider is in the process.

However, in order to evaluate progress in a consistent and comparable way, several community initiatives have developed related FAIR metrics and assessment frameworks that translate the principles into criteria. One of such examples is the "FAIR Data Maturity Model", which makes the output of **FAIR verification tools** comparable by delivering a set of indicators with priorities and guidelines that provide a 'lingua franca' (Bahim et al., 2020). For Climate Services, such frameworks are valuable, since they allow for evaluating the FAIRness of datasets and identify areas where domain-specific agreements are still required.

While FAIR metrics and maturity models assess the level of FAIRness, practical implementation requires concrete technical and organisational solutions. **FAIR implementation tools** help data providers apply standards, create structured metadata, assign persistent identifiers, document provenance, and provide machine-accessible access, thereby translating the FAIR principles into operational practice.

3.2 FAIR implementation tools

FAIR implementation tools help to make data more FAIR and guide the providers in creating and publishing FAIR data. They are 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 Profile (FIP)[2] via which a community can determine the current state of FAIRness and suggest and agree on implementation choices for improvement.

The FIP is a collection of FAIR implementation choices made by a community of practice for each of the FAIR Principles. Ready-made and well-tested FAIR Implementation Profiles created by trusted communities can find widespread reuse among other communities, and vastly accelerate convergence onto well-informed FAIR implementations.

The creation of the FIP is done via the FIP Wizard[3]. This version is by now a stable and robust tool. However, as the FIP Wizard generates and relies on information that is open to the community at large, it requires some expert guidance in its use in order to ensure the quality of the output. This means that anyone who wishes to create a FIP is recommended to first participate in a FIP Introduction workshop.

As an impression, making a FIP will typically involve three activities summarised below:

- Completing the FIP Questionnaire: Answer the questions of the FIP Wizard questionnaire by selecting for each FAIR Principle, the FAIR Enabling Resources from the drop-down lists, setting the implementation status, and stating your considerations behind that choice.

- Minting New Nanopublications: the FIP-Wizard offers implementation choices (solutions, standards, technology) to fulfil a certain FAIR metric. These are stored as a so-called nano publication. When a new community starts to work in the FIP-Wizard there can/will be a need to add community agreed solutions not present in the wizard yet.

- Publish the FIP.

After this, the participants will work on the steps to improve FAIRness based on the selected planned solutions, and after some time a new FIP version will be created to track the process. In this way the FIP wizard offers very valuable guidance in the work.

The creation of nanopublications is also echoed by to the FAIR Connect initiative[4], which aims to be a global hub for sharing, documenting, and giving academic legitimacy to tools and practices that make data FAIR. Within FAIR Connect, foundational elements such as FAIR Enabling Resources are complemented by practical tools, workflows, and materials that assist users in applying FAIR principles. These practical assets, referred to as FAIR Supporting Resources (FSRs), include Data Management Plan templates, metadata examples, training materials, policies, and guidelines. Together, they represent the operational resources that help communities, institutions, and individuals steward data effectively.

3.3 FAIR verification tools

The implementation of the FAIR principles in practice requires evaluation of the FAIR level of data to assess the degree to which datasets comply with these principles. As FAIR was intentionally designed as a set of high-level guidelines, its practical implementation requires translation into measurable indicators and verifiable criteria. This need has led to the development of multiple FAIR assessment tools, both manual and automated, designed to evaluate FAIR compliance across domains and infrastructures[5].

[2] https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/
[3] https://fip.fair-wizard.com/wizard
[4] https://fairconnect.pro
[5] https://www.thehyve.nl/articles/evaluation-fair-data-assessment-tools

FAIR verification tools support benchmarking, and quality improvement processes. They are increasingly used by repositories, research infrastructures, and large-scale data platforms to monitor compliance and data stewardship practices. In particular, platforms hosting substantial volumes of datasets often employ automated assessment using these kinds of tools to systematically verify that exposed metadata meets agreed FAIR criteria.

These tools check for the metadata that is typically embedded in dataset landing pages in machine-readable formats (e.g. JSON-LD or RDFa), or exposed via APIs or catalogue services. This exposure enables automated validation by tools such as F-UJI or FAIR-Checker. Their outputs are often quantitative (pass/fail), which makes them particularly useful at later stages of FAIR implementation, when metadata standards have been agreed and seek to validate compliance. It must be noted that these tools are still under development, while others are still emerging.

3.3.1 Challenges

Despite increasing efforts to create new FAIR assessment tools, methodological challenges persist. As discussed by Candela et al. (2024) the interpretation of FAIR metrics varies across tools. Translating the FAIR guidelines into operational and measurable criteria inevitably involves subjective interpretative decisions. This has resulted in multiple metric interpretations and multiple FAIR verification tools with different methodological purposes.

There are two main kinds of FAIR verification tools; manual and automated tools. From one side, manual tools such as [FAIR Data Self Assessment Tool] are more focused as a guideline in the form of questionnaires or indications of how the different FAIR guidelines could be met. The tool is not scalable and requires user intervention to obtain results but are good guidelines to improve the current state of a dataset in terms of FAIRness. Additionally some metrics design lack clear documentation which increases the subjectivity.

Automated tools, by contrast, enhance scalability, repeatability, and objectivity in execution. However, they are not free from subjectivity. Their metrics and scoring logic reflect the design choices of their developers leading to different outcomes. This was supported by the comparative analysis conducted by Sun et al. (2022), which demonstrated that three automated tools (F-UJI, FAIR-Checker and FAIR-Evaluator) produced different FAIR scores when applied to identical datasets, highlighting the necessity of greater harmonisation, standardisation and transparency in metric design. Similarly, Krans et al. (2022) observed that automated tools scale effectively but may lack contextual awareness. They also note that only a limited number of tools provide actionable recommendations to improve FAIR compliance.

3.3.2 Applicability and Limitations

Most automated tools assess FAIRness by programmatically accessing a dataset's resolvable identifier (e.g. DOI or URL), retrieving the associated landing page and analysing the metadata that is machine-readable and externally accessible. They typically evaluate structured metadata embedded in HTML (e.g. JSON-LD or RDFa), metadata retrievable via content negotiation, or metadata exposed through APIs and catalogue services. It is very important to note that automated tools generally do not download or inspect the full content of data files or compressed archives during the evaluations. This means that the assessment is limited to what is discoverable at the web interface level.

The previous has important implications. Repositories that explicitly expose structured metadata on dataset landing pages can benefit from automated FAIR assessment at scale. Verification tools can systematically analyse metadata across large collections and support FAIR compliance monitoring. By contrast, when metadata is only available with the downloadable files (for example, embedded inside compressed archives or domain-specific data formats), automated tools are unable to access the metadata. In such cases, the dataset could contain a comprehensive and high-quality metadata, however, the FAIR evaluation tool will score the dataset with a low FAIRness score.

It is therefore essential to distinguish between:

- **Intrinsic metadata quality**: What exists within the dataset or files.
- **Machine-exposed metadata**: What is accessible to automated systems.

Automated FAIR evaluation tools interpret FAIR principles through predefined, machine-executables tests. While these tools provide valuable and scalable assessments, some metrics only inspect general aspects of the criteria to be evaluated as the tests are designed to be broadly applicable to data/metadata characteristics, they may not capture discipline-specific practices, contextual nuances, or metadata that is not exposed in a machine-readable format.

As a result, certain aspects specified in FAIR principles could benefit from some level of human mediation and interpretation[6] The following table provides a comparison of how F-UJI and FAIR-Checker tools approach FAIR principles and their limitations.

Table 1. Comparison of two implementation approaches to the FAIR principles and their limitations. The approaches being compared are F-UJI and the FAIR-Checker. **Note:** Empty cells (—) indicate that the tool does not currently provide a specific test or explicit metric for that FAIR principle..

FAIR Principle	GO-FAIR (F-UJI)	FAIR-Checker	Key Limitation
F1: (Meta)data are assigned globally unique and persistent identifiers	*PID syntax, persistence, resolvability, and PID authority registration*	*Identifier is reachable URL; stronger checks require Schema.org/DCTerms properties*	If PID exists but is not machine-readable tools may fail or obtain a lower score
F2: Data are described with rich metadata	Essential citation elements needed for basic findability (creator, title, object_identifier, publication_date, publisher, object_type)	Verifies that at least one RDF triple can be found in metadata and specific structured properties. Additionally, at least one used ontology class is used in major ontology registries (OLS, BioPortal, LOV)	Field presence does not ensure domain acceptability, completeness, or descriptive quality. These tools are starting to create domain specific assessments.
F3: Metadata clearly and explicitly include the identifier of the data they describe	Metadata contains a PID or URL which indicates the location of the downloadable data content	—	Link existence does not guarantee version correctness or whether the link resolves to the exact described object
F4: (Meta)data are registered or indexed in a searchable resource	*Checks Structured data embedded in the landing page and content negotiation with a PID provider*	—	Tools test indexability in the landing page, not actual discoverability, check
A1: (Meta)data are retrievable by their identifier using a standardised communication protocol	*Checks protocol type, resolvability, authentication support*	Verifies that the resource is accessible via an open protocol, for instance the protocol needs to be HTTP.	Tools cannot bypass authentication barriers (leading to false negatives) and only check connectivity.

[6] F-UJI - Automated FAIR Data Assessment Tool

A2: Metadata should be accessible even when the data is no longer available	—	—	Tools cannot reliably assess long-term metadata preservation
I1: (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation	Machine-readable, structured metadata embedded directly in the landing page and content negotiation, through standardised machine-accessible mechanisms	Verifies that at least one RDF triple can be found in metadata	Structured syntax does not guarantee correct semantic, ontology or logical consistency
I2: (Meta)data use vocabularies that follow the FAIR principles	—	Verifies that at least one used ontology class or property are known in major ontology registries (OLS, BioPortal, LOV)	Registered vocabularies may be inconsistently or incorrectly applied
I3: (Meta)data include qualified references to other (meta)data	Qualified (PROV-O or DataCite) machine-readable references	Multiple external URL authorities	Links do not ensure meaningful scientific contextualisation
R1: (Meta)data are richly described with a plurality of accurate and relevant attributes	Checks structured technical data properties, license and provenance elements Detects domain standards via namespace	Looks for license/provenance properties in RDF Detects ontology registry presence	Namespace detection confirms reference to a recognised standard but does not ensure full compliance with the domain-specific metadata requirements or best practices.

4 Case studies from the public community

This section presents a set of case studies from the public community to illustrate multiple Quality Assurance and Quality Control methodologies applied to different contexts. These case studies are presented in the form of structured one-pagers, because we decided to trade readability for detailed completeness. We tried to follow the same structure in all case studies as long as the information was available, and this has been divided in five sections:

- **Context**: provides a high-level description of the initiative presented in the case study.
- **Challenge**: addresses the problem at hand that is required to be monitored.
- **QA/QC actions**: introduces the actions taken to monitor the challenge described above.
- **QA/QC Results and their Communication**: highlights how the QA/QC results are communicated.
- **Feedback**: includes methodologies used in the co-production between users and developers.

We tried to present trustable information by checking the content of the case studies with experts that had directly worked on them. This was possible for most of the case studies. However, we recommend the interested reader to directly contact the organisation behind the initiative to obtain more detailed information.

4.1 Copernicus Climate Change Service - Climate Data Store

Context

The Copernicus Climate Change Service (C3S) is a European Union initiative that provides authoritative, quality-assured data on the past, present, and future Climate. This information is primarily accessed through the Climate Data Store (CDS), a centralised digital platform that offers free, open-source tools, application services and datasets to help users analyse Climate drivers and impacts.

Challenge

Previous efforts established a strong technical basis for quality information—including data maturity assessments, integrity checks, and diagnostic outputs. The primary lesson learned was the need to present quality information in a more user-centric way, focusing on the practical questions stakeholders ask about data quality and fitness for purpose (Dee et al., 2024).

QA/QC actions Technical Quality Scientific Quality

The goal of the **Evaluation and Quality Control** is to develop precise statements about data quality that pertain to well-identified use cases (Dee et al., 2024). It has three layers:

- **Quality assurance (technical)**
  Serves to inform users that data, metadata, and documentation comply with a well-defined set of verifiable technical requirements compatible with the FAIR principles, which are routinely checked independently from the data producers. Users should expect that all technical requirements are met. The **FAIRness** of CDS datasets can be checked through the FAIRsFAIR F-UJI assessment page[7].
- **Quality assessment (scientific)**
  The goal is to provide science-based information about the accuracy, uncertainties, strengths, and weaknesses of a dataset in the context of realistic use cases. Assessments are carried out by experts and support users in understanding how the dataset performs, where it is reliable, and where caution is needed..
- **Fitness for purpose**
  Brings together the outcomes of technical quality assurance and scientific quality assessment, and translates them into practical guidance on whether, how, and under which limitations the dataset can be used for a particular purpose.

QA/QC Results and their Communication

EQC information is exposed to users through the CDS and associated EQC documentation. This includes dataset-level quality assurance checklists, fitness-for-purpose information in factsheet-style summaries, and detailed scientific quality assessments in a dedicated EQC Jupyter Book.

Feedback channels

C3S employs a transparent, traceable process to gather user requirements through workshops, training events, surveys, and direct stakeholder interactions. Each requirement is logged in a **database** with its technical details, context, and source. These entries are then continuously analysed and clustered, allowing C3S to prioritise implementation based on feasibility, user benefit, and cost-effectiveness within its current capacity.

[7] https://www.f-uji.net/

4.2 Couple Model Intercomparison Project (CMIP)

4.2.1 CMIP5

Context

CMIP5 was the fifth phase of CMIP[8]. It ran roughly between 2010 and 2014 to provide a framework for coordinated Climate change experiments to inform the IPCC Fifth Assessment Report (AR5). Data was distributed through the Earth System Grid Federation (ESGF), an international Peer-to-Peer (P2P) enterprise system to handle large-scale data management for worldwide distribution. Parts of CMIP5 data are long-term archived at the World Data Center for Climate (WDCC) at DKRZ.

Challenge

In contrast to previous CMIP phases, CMIP5 faced an unprecedented scale and distributed nature of data. Because data was generated by over 20 modeling groups worldwide and stored across decentralised repositories, maintaining consistency required a complex, tiered system of automated and human checks. In particular, the aim was that every file adhered to the Climate and Forecast (CF) convention and the Climate Information Model (CIM) before it could be "frozen" and formally published with a persistent identifier.

QC/QA actions Technical Quality

The quality control (QC) consist of 3 levels[9]:

- **QC Level 1: Basic ESGF Conformance Checks**
  Uses the Climate Model Output Rewriter (CMOR) to ensure all data follows strict structural and naming standards, like the correct "data reference syntax" (DRS) and validity against the CF conventions, ensuring coordinate axes and variable names are globally consistent. Technical checks verified file format, size, readability and data discovery. Simultaneously, the Climate Information Model (CIM) is followed to document the numerical requirements and lineage of the simulations. This information is collected with the METAFOR Questionnaire, and stored in XML files.
- **QC Level 2: WDCC Conformance Checks and Subjective Quality Control**
  File and metadata consistency is checked. In addition, physical variables are compared with specified ranges.
- **QC Level 3: STD-DOI Data Publication Process**
  After the final review of the authors, the data is published following the Publication and Citation of Scientific Primary Data (STD-DOI) project. This creates a citable frozen version of data that can be accessed through a persistent identifier.

QA/QC Results and their Communication

Quality control results are stored jointly with the data and metadata.

Feedback channels

[8] CMIP4 was skipped to keep the CMIP version number in sync with the IPPC AR version number.
[9] https://redmine.dkrz.de/documents/4

The ESGF provides dedicated support for both modeling centers and data users. For producers and authors, the process includes a mandatory review stage: once a dataset reaches Level 3, authors are notified to provide final approval before official publication.

4.2.2 CMIP6

Context

Phase 6 of CMIP was initiated around 2016 with the data release between 2019 and 2021 to support the IPCC Sixth Assessment Report (AR6). While CMIP5 focused on Representative Concentration Pathways (RCPs), CMIP6 introduced a more complex set of Shared Socio-economic Pathways (SSPs) that combined Climate forcing with socio-economic narratives (Eyring et al., 2016a). As CMIP5, CMIP6 data was facilitated by the ESGF.

Challenge

Data continued to grow between CMIP5 and CMIP6 reaching tenths of petabytes, and while CMIP5 struggled with basic data consistency, CMIP6 focused on scientific validation and ensemble reliability. The early citability of the data and thus the credit for the data authors became more of a focus resulting in a PID Service for data files and a DOI service for data collections.

QC/QA actions Technical Quality Scientific Quality

Technical verification for ESGF nodes continued throughout CMIP6. For long-term archiving at WDCC (IPCC-DCC), data underwent rigorous quality assurance, including format and checksum validation, as well as cross-referencing metadata between citation records, ESGF entries, and file headers. The software tools used for these exercises were QA-DKRZ, nctime and PrePARE. Only the last one was systematically implemented in the ESGF publication workflow to ensure all published datasets were CMIP6 compliant.
This phase introduced the DECK (Diagnostic, Evaluation and Characterization of Klima). It is a small set of four mandatory experiments that every model must complete before it is allowed to participate in any other specialised project. The experiments were:

- **AMIP**: an atmosphere-only run.
- **Pre-industrial control**: a long run with stable 1850s-era CO2 to see if the model drifts over time.
- **Abrupt 4xCO2**: A shock test where CO2 is suddenly quadrupled to see how fast the model warms up.
- **1% CO2 per year**: A steady increase in CO2 to measure the model's Climate sensitivity.

Because every model was now running the exact same DECK experiments, scientists could build automated tools like **ESMValTool** (Earth System Model Evaluation Tool; Eyring et al., 2016b) or **PCMDI Metrics Package** (PMP; Gleckler et al., 2016) to perform objective comparisons between models and observations.

QA/QC Results and their Communication

Results of the Quality Assurance at WDCC are published with the data.

Feedback channels

Integrated support for users is provided through the ESGF and WDCC helpdesks and a newly implemented Errata Service.

4.2.3 CMIP7

Disclaimer: the infrastructure supporting CMIP7 is currently under development. Thus, information provided here could not be fully consistent with the final implementation.

Context
Phase 7 of CMIP is currently in its active implementation phase with a central goal of supporting the IPCC Seventh Assessment Report (AR7).
Challenge
Traditionally, Climate modeling centers have documented model performance only after finalising development and completing major simulations. According to Hoffmann et al. (2025), this retrospective approach creates a bottleneck for IPCC assessments due to slow publication cycles, fragmented data across various journals, and a lack of standardised metrics between different research groups.
QC/QA actions Technical Quality Scientific quality
While CMIP7 maintains the metadata standards and controlled vocabulary checks of CMIP5, and the **DECK** from CMIP6 as the mandatory "entry card", it also expands DECK to include new experiments focused on CO2 emission-driven scenarios and updated historical forcings that extend to at least 2021. For CMIP7, a unified QA/QC framework has been developed as an IOOS Compliance Checker plugin, consolidating the legacy validation tools used for CMIP6 into a single, robust, and long-term maintainable framework. This new plugin provides a highly configurable validation engine, enabling consistent enforcement of CMIP requirements across modelling centres. It is strongly recommended for use by CMIP7 data producers and will be systematically applied by ESGF publishers. In addition, ESGF indexing will rely on mandatory server-side metadata validation through a JSON STAC schema derived directly from the newly centralised and machine-readable CMIP7 controlled vocabulary. This represents a significant procedural shift: metadata validation at publication time is no longer advisory but enforced, ensuring strict compliance and strengthening the overall integrity and interoperability of CMIP7 datasets. To ensure that the most critical data is ready for IPCC AR7, the Model Benchmarking Task Team (MBTT) designed the **Assessment Fast Track (AFT)**, which is a subset of Climate model experiments aimed at meeting the tight 2026-2027 deadlines of the AR7 (Dunne et al., 2024). The **Rapid Evaluation Framework** (REF) is the technical engine designed to evaluate that as soon as it is published on ESGF by the originating modelling centres (Hoffman et al., 2025). The REF leverages multiple benchmarking tools (ESMValTool, PMP, ILAMB, IOMB, CMEC[10], ...), and the best-available observational datasets collected in the project obs4MIPS (Wallser et al., 2020) to provide assessments of model fidelity across a collection of diagnostics, which were identified and finally selected with community involvement and consultation. Rather than directly ranking models, the REF is primarily concerned with providing objective measures of model performance to allow the wider community to make informed decisions about models that are most appropriate for their specific needs.

[10] https://cmec.llnl.gov/

4.3 WIS and WIGOS

Context

WMO Integrated Global Observing System (WIGOS, 2007) and **WMO Information System** (WIS, 2011) serve as the operational twin-pillars of the World Meteorological Organization (WMO). WIGOS provides a unified framework for observations, while WIS is the automated infrastructure used to exchange that data globally. In other words, WIGOS generates the data, and WIS ensures it is discoverable and accessible.

Challenge

The primary challenge is maintaining consistent and enforceable standards across nearly 200 member stations with vastly different technical and financial resources. This "digital divide" often results in data gaps or varying levels of precision, which can weaken the accuracy of global Climate models. In addition, ensuring that metadata remains accurate and synchronised is a constant battle against technical obsolescence.

QC/QA actions Technical quality Scientific quality

The key measures defined in the Manual on WIGOS (WMO-No. 1160) are:

- **WIGOS Data Quality Monitoring System (WDQMS)**: This tool monitors the availability, timeliness, and quality of data from all observing components in real-time. It links performance data from monitoring centres with station metadata in the OSCAR database to identify network issues.
- **Regional WIGOS Centres (RWCs)**: These centres act as the bridge for quality assurance, following up on issues identified by WDQMS and assisting WMO Members with remedial actions at the national level.
- **Automatic Quality Control (QC)**: WMO provides Guidelines on QC Procedures for automatic weather stations, including checks on plausible values (gross error detection), time consistency (rate of change and persistence tests) and internal consistency (between related variables).
- **Traceability and Calibration**: The Instruments and Methods of Observation Programme (IMOP) sets technical standards and ensures measurement traceability to international standards through specialised Regional Instrument Centres.

Regarding WIS:

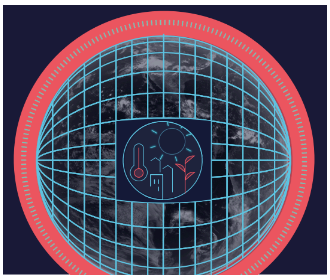
- **Standardised Metadata**: Quality assurance in WIS relies on the WMO Core Metadata Profile (WCMP) and WIGOS Metadata Standard (WMO-No. 1192) to ensure data is discoverable and its lineage is fully documented.
- **Regulatory Compliance**: Operational standards are mandated by the Manual on the WMO Information System (WMO-No. 1060), which ensures uniformity in communication practices and data representation formats.
- **WIS 2.0 Monitoring**: The transition to WIS 2.0 introduces enhanced monitoring through **Global Brokers** and **Global Caches**, which track data notifications and delivery reliability to maintain high service levels.

Data Stewardship Maturity Assessment (see Section 5.4.1 for more information): WMO is further developing approaches to assess the maturity of data stewardship practices through maturity matrices. These matrices assess how well datasets are documented, quality assured, and curated across their lifecycle, considering aspects such as metadata completeness, traceability of the data origin and processing history, and uncertainty information that informs users about dataset reliability. **The WMO Expert Team on Data Development and Stewardship (ET-DDS)** contributes to the adaptation of stewardship maturity matrices for Climate datasets.

4.4 Destination Earth — Climate Digital Twin

Context

The DestinE Digital Twin for Climate change adaptation is a new framework to provide Climate information based on global Climate projections from 1990 to ~2050 at the very high-spatial resolution of 5 to 10 km by leveraging European supercomputing facilities and cutting-edge Earth-system models. Its goal is to resolve phenomena inaccessible to Global Circulation Models, which usually operate at a scale 10 to 20 times coarser than the 10km and 5km DestinE simulations (Doblas-Reyes et al., 2026). The Climate DT is being developed as part of Destination Earth, through contracts procured by ECMWF.



ECMWF - DestinE factsheets

Challenge

Since there is no trivial way of assessing when a simulation has good quality or not, there is the need for a flexible tool that accommodates the different needs and questions that researchers might have by providing multiple diagnostics. In addition, given the nature of high-resolution Climate simulations, the tool needs to be able to efficiently handle large datasets and support various data formats.

QC/QA actions Scientific quality

The Application for QUality Assessment (AQUA), is an open-source code package written in Python for running a wide range of diagnostics on high-resolution Climate models[11].
There are two main types of diagnostics:

- **State of the art diagnostics** (link)
  With the aim of saving computational resources, they are run on a data-reduced version of the simulation, specifically at 1 degree in longitude and latitude and with monthly data. The diagnostics include comparisons with observational datasets and pre-existing simulations.
- **Frontier diagnostics** (link)
  These are run at the native spatial and temporal resolution of the simulations. The goal of these diagnostics is not to check the quality of the simulations, but rather provide new insights using high-resolution data.

QA/QC Results and their Communication

The tool offers many plotting options and also the possibility of feeding external dashboards with automatic analytics.

Feedback channels

Users of AQUA can provide feedback about coding issues or suggestions through the Github repository. In addition, scientific analysis using AQUA can be published as long as the developers are properly cited.

[11] For more information see the Documentation, the GItHub package and a presentation by Silivia Caprioli and the AQUA team.

4.5 High-Value Datasets and the RODEO project

Context

RODEO logo

The INSPIRE Directive (EC) 2007/2 requested to make spatial data available for environmental and policy purposes. Then the EU Open Data Directive (EU) 2019/1024 broadened the scope of data sharing by requiring public authorities to freely publish all public sector data — including geospatial information — to foster innovation and economic growth. The Open Data Directive also introduced the concept of High-Value Datasets (HVDs), referring to datasets with significant social and economic potential. Subsequently, the Implementing Regulation (EU) 2023/138 mandated that these datasets be made available free of charge, in machine-readable formats, and accessible through APIs.

Challenge

The mandate to create APIs did not specify how they should be implemented. Consequently, nine National Meteorological Services joined forces to form a consortium and develop a common API approach within the framework of the RODEO project. Their proposed solution extends the OGC API Environmental Data Retrieval standard and is referred to as MetOcean EDR, which is openly available on GitHub. A derivative challenge lies in ensuring that the APIs developed by each National Meteorological Service comply with the consortium's agreed implementation.
Ensuring compliance is essential because all APIs are integrated into a single data hub called Climate Aggregator, that serves as a unique access point to the data of all participating Met Services. This data hub follows the CDM-Obs-Core data model proposed by Copernicus.

QC/QA actions Technical quality

To ensure that all APIs follow the agreed implementation and produce outputs in a consistent format (CovJSON), the RODEO Work Package 5 team developed two Python libraries to support developers in building their own APIs (see presentation by Paul van Schayck):

- OGC Environmental Data Retrieval (EDR) API Pydantic
- CoverageJSON Pydantic

Both libraries are freely available on GitHub. The "Pydantic" part of the title refers to a Python library that is used and helps compare an object to what a standard says it should be, i.e. ensures compliance.

QA/QC Results and their Communication

If the API returns something that doesn't match the expectations, Pydantic will raise a validation error.

Feedback channels

GitHub

4.6 European Open Science Cloud

Context

The European Open Science Cloud (EOSC) is a European project, whose objective is to provide researchers and innovators in Europe with an open and trusted multi-disciplinary environment where they can publish, find and reuse data, tools and services for research and innovation. It's funded by the European Commission, national contributions, in-kind contributions, and project-based calls.

Challenge

EOSC seeks to promote the adherence to best practices in three areas: software, services and FAIR data. For each area it has defined a specific set of quality criteria (see previously attached links). The FAIR specification follows the FAIR Data Maturity Model Specification and Guidelines provided by the RDA.

QC/QA actions Technical quality

The Software Quality Assurance as a Service (SQAaaS) is a platform with a web interface that has two modules:

- **Pipeline as a Service** (link)
  It helps compose CI/CD pipelines based on the feedback of the user, and a means to distribute them so that they can be added to the target code repository. The checks that pipeline provides are limited to predefined software criteria.
- **Quality Assessment & Awarding (QAA)** (link)
  It uses a selection of open source tools to evaluate each of the criteria. For example, the FAIR principles are evaluated using the FAIR EVA (Evaluator, Validator & Advisor). In practice, one could think of QAA as a wrapper of these open source tools.

Alternatively, tech-savvy users are offered a coded version of SQAaaS called the Jenkins Pipeline Library (JePL).

QA/QC Results and their Communication

The QAA provides two main outputs:

- a **quality report** with the results of the assessment
- a **digital badge** certifying the achievements of the evaluated software.

There are three levels of compliance: bronze, silver, and gold. With regards to sharing the generated badges within code and data repository platforms, SQAaaS offers two options: using the badge image or through the shields.io common badge solution.

Feedback channels

Feedback can be provided through the corresponding GitHub repositories.

4.7 Open Geospatial Consortium

Context

The Open Geospatial Consortium (OGC) is a global community of experts dedicated to creating open standards for geospatial data and services. Founded in 1994, the organisation has grown to include more than 400 members (as of 2025) and has published over 150 freely available standards. For example, the OCG API - Environmental Data Retrieval standard, which provides a modern web-friendly approach for accessing environmental data

OGC compliance logo

Challenge

The use of OGC standards is widely recognised and endorsed across the Climate Services community. However, new versions are released frequently, and the community continues to grow. At the same time, OGC standards are technically detailed and comprehensive, making manual verification of compliance increasingly impractical. As a result, there is a clear need for an automated verification system capable of assessing compliance with the diverse and evolving suite of OGC standards. Publicly demonstrating compliance is also strategically valuable, because it ensures interoperability with other solutions and alignment with established best practices.

QC/QA actions Technical quality

The OGC validator is a software tool that assesses whether web services conform to OGC standards. . It is freely available on GitHub, and the list of available test suites—i.e., the standards for which compliance can be assessed—is published on the OGC validator website.

QA/QC Results and their Communication

Products that successfully pass the validation process can be acknowledged or registered in the OGC Implementation Database. For organisations seeking official certification and the right to display the OGC compliance logo, the OGC Compliance Program which grants an OGC Trademark License. The certification involves a fee, which varies based on the product's gross annual revenue and the OGC membership status of the organisation that owns the product.

Feedback channels

Issues related to the OGC validator, can be reported directly via GitHub.

4.8 EUMETSAT & Climate Anomalies and Normals-Service

Disclaimer: Information regarding the CAN-service is preliminary and subject to change, as the service is currently in active development. Official launch is anticipated for mid-to-late 2026.

Context

EUMETSAT is the European organisation for the exploitation of meteorological satellites. This includes the development and provision of data on geophysical parameters retrieved from satellite observations and the creation of long-term homogeneous Climate Data Records (CDRs) based on satellite observations. Until now EUMETSAT has not been providing normal nor anomaly products nor other Climate indices from its CDRs. Users had to calculate them on their own, and it's a complex process, reprocessing needs to be done first to make the data comparable over the whole time period. As a consequence, Climate Services from member states requested EUMETSAT to provide pre-calculated normals and anomalies. In response, EUMETSAT is developing the **Climate Anomalies and Normals** (CAN)-service to operationalise their provision.

Challenge

While the challenges are numerous, the development of the CAN-service includes: developing a **common data model**, establishing a scientific process to ensure consistency among the variables provided via the CAN-service within the context of the service, and determining a good strategy for communicating inconsistencies the team might discover among CAN products and in relation to C3S products (ERA5-reanalysis mainly). To address these challenges, a dedicated EUMETSAT Working Group is currently defining a set of "**good practices**" that need to be followed by all providers of CAN-products, currently only EUMETSAT's Satellite Application Facilities (SAFs).

QC/QA actions Technical quality Scientific quality

The scientific scientific consistency assessment process evaluates two primary dimensions:

- **Internal Consistency**: Aligning two datasets of the same variable from different sources (e.g., EUMETSAT versus another relevant provider) or from different instruments (e.g. geostationary vs. polar orbiting satellites).
- **External Consistency**: Correlating datasets of different variables where a physical relationship is expected.

This process begins with an automated "light" consistency check, followed by a more rigorous, manual analysis. Note that inconsistencies between data-records can be helpful to improve both the observing system and earth system models.

QA/QC Results and their Communication

When inconsistencies among different CAN-products are identified, the CAN-service team collaborates directly with the relevant developers to investigate and resolve the root cause. A procedure still needs to be worked out for the case when inconsistencies with products of other Climate data providers (e.g. C3S) are identified.

Feedback channels

User feedback for the CAN-service will be managed through a dedicated helpdesk and ticketing system. On the technical side,

4.9 European Space Agency's Climate Change Initiative

Context

The European Space Agency's Climate Change Initiative (ESA CCI) is a long-term research programme whose main objective is to generate, validate, and maintain satellite-based Climate data records. It encompasses 33 Essential Climate Variable (ECV) projects, such as soil moisture, fire, and land cover, and over 50 CCI projects overall when accounting for cross-ECV, observations and modelling, tipping elements, and Paris Agreement support. The CCI is funded by ESA and supports the UNFCCC and IPCC. Global accessibility of data is granted through the CCI Open Data Portal. In order to ensure openness and consistent production among datasets, ESA CCI has developed a Data Policy, a Data Standard and an Ontology. The Data Standard requires conformance to the netCDF format, the CF metadata convention, ACDD, and INSPIRE.

Challenge

There are two main challenges:

- Ensuring the compliance with the aforementioned Data Policy & Standards.
- Providing trust on the generated ECV data records from satellite observations.

QC/QA actions Technical quality Scientific quality

- **Technical compliance to Data Policy & Standards**
  The Open Data Portal provides help and advice on data formatting and support for checking compliance A liaison between CCI Projects and the portal is encouraged.
- **Product validation and intercomparison reports**
  International expert groups provide a Product Validation Plan (PVP) and a Product Validation and Intercomparison Report (PVIR) for each project. The validation described in the PVP is performed by comparison to in-situ observations.

QA/QC Results and their Communication

Validation reports — see this PVIR example for the Soil Moisture project — are available in the documentation site of each ECV project, and attached to each dataset provided in the Open Data Portal. The CCI programme includes a dedicated "CCI Knowledge Exchange" project, which addresses Article 12 of the Paris Agreement.

Feedback channels

Aside from providing the contacts to the international teams behind the production of the datasets, so that issues can be reported, ESA CCI created the **Climate Modelling User Group (CMUG).** The CMUG is an international forum through which the Earth observation data community and the Climate modelling and reanalysis community work closely together to build a bridge between satellite earth observation data records and the needs from the Climate science community.

4.10 ESMValTool

4.10.1 Data

Context

Earth System Model Evaluation Tool (ESMValTool) is an open-source community-developed diagnostics and performance metrics tool for the evaluation and analysis of Earth System Models (ESMs). ESMValTool allows for a comparison of single or multiple models against predecessor versions and observations. The aim of ESMValTool is to take model evaluation to the next level by facilitating analysis of many different ESM components, providing well-documented source code and scientific background of implemented diagnostics. Traceability and reproducibility of the results are ensured.

Challenge

The goal is to develop a benchmarking and evaluation tool that produces well-established and consistent analyses, mainly from model output from CMIP simulations. The suite of diagnostics needs to be broad enough to reflect the diversity and complexity of Earth System Models, but must also be robust enough to be run routinely or semi-operationally.

QC/QA actions Technical quality Scientific quality

The QC is done through the CMOR check and fixes preprocessor. The CMOR check module checks for compliance of the netCDF input data and metadata against the CF metadata convention and CMOR standards. It checks for the most common dataset problems. The format checks and fixes based on the CMOR tables (can be user-defined).
The adoption of Iris, with strict requirements of CF compliance for input data, generally requires the implementation of fixes for a larger number of datasets. Although from a user's perspective this makes the reading of some datasets more demanding, the stricter standards enforced ensure their correct interpretation and reduce the probability of unintended behavior or errors. No QA is provided. The code responsible for those checks is available.

QA/QC Results and their Communication

When the ESMValTool QC does not pass the CMOR check preprocessor, a CMORCheckError is issued by the ESMValCore and the processing is terminated. The user is informed with a specific error message. The processing can be either terminated after the first error, or terminated at the end of the preprocessing to inform the user about all errors, depending on user configuration. There are also different levels of possible error reporting: debug, warnings, normal, critical.

Feedback channels

Issues can be reported to the corresponding data producer of the input files. In the case of CMIP datafiles and CF compliant data, the contact organisation is mentioned in the metadata of the NetCDF files. Regarding the software itself, issues must be reported directly on the github repositories of ESMValGroup.

4.10.2 Software

Challenge

The goal is to develop a benchmarking and evaluation tool that produces well-established and consistent analyses, mainly from model output from CMIP simulations. To be robust the code development is going through a quite strict quality control and checking, to ensure proper written code.

QC/QA actions Technical quality Scientific quality

The QC is done through several tools that run either locally on each contributor's code or on the hosting code platform GitHub (with its embedded tools). Pre-commit checks are performed before each commit is added to a local development branch, and block the commit if any check has not passed successfully. The checks rely on the software Ruff that integrates various linting and code formatting rules, notably following PEP8. Additional checks available through GitHub (pre-commit-hooks) are also performed simultaneously. On the code hosting platform GitHub, other quality checks are performed for each commit on opened pull requests. The tool Codacy identifies issues relating to code style, best practices, code complexity, or documentation among others. A report on the test coverage of any new code is also provided to guide the design or update of tests. Concurrently, CircleCI runs a series of selected tests. Both of these checks need to pass successfully before the pull request can be reviewed by a member of the technical team. Routine tests are also run automatically every day to ensure the stability and consistency of the software.

Results and their communication

Outputs from the pre-commit checks are shown on the terminal, and provide detailed feedback on where and which rule break has been identified. The results from Codacy and CircleCI are provided through links to their own platforms, and offer an overview of all current issues and test logs.

Feedback channels

Any software issues can be reported on the GitHub platforms following the provided issue template, and are regularly monitored by members of the technical team. When contributing code, the technical team provides feedback during the technical code review which is a mandatory step before any new code can be merged into the software repository.

4.11 AtMoDat project

Context

The Atmospheric Model Data (AtMoDat) project, funded by the German Federal Ministry of Education and Research (BMBF), began in mid-2019 with the goal of enhancing the overall FAIRness of atmospheric model data. The consortium includes four partners: two from Climate research — the Universities of Hamburg and Leipzig — and two from infrastructure and operations — the German Climate Computing Center (DKRZ) and the Technische Informationsbibliothek (TIB)

AtMoDat logo

Challenge

The consortium developed a FAIR implementation protocol called the ATMODAT standard. This standard defines requirements for comprehensive metadata, DOI assignment, the structure and content of DOI landing pages, the use of controlled vocabularies, data formats, licenses, and other key aspects. Developing such an implementation was a significant achievement in itself. The next challenge lies in enabling the automatic verification of datasets' compliance with the ATMODAT standard.

QC/QA actions Technical quality

The ATMODAT Data Checker is a Python-based tool designed to automatically assess compliance with the ATMODAT standard. It is available as open-source software on GitHub. The tool builds upon the IOOS Compliance Checker — a Python utility for validating netCDF files against the CF convention — and extends its functionality to cover all requirements defined in the ATMODAT standard.

QA/QC Results and their Communication

If a dataset does not meet the ATMODAT standard, the Data Checker provides a list of warnings and errors. Conversely, datasets that are fully compliant may display the EARth SYstem DAta Branding (EASYDAB) logo on their landing page as a mark of compliance. To use the logo, the repository must first enter into an agreement with the TIB.

Feedback channels

For general inquiries or questions about the ATMODAT project, contact information is provided. Specific issues related to the ATMODAT Data Checker, however, can be reported directly via GitHub.

4.12 The Comprehensive R Archive Network

Context

CRAN is a public repository of code packages of publication quality built for the R programming language, which is maintained by the CRAN team. The packages are built by the community and are freely accessible by anyone. New packages can be submitted through a web form. Regarding funding, CRAN operates through volunteer efforts, institutional support and donations.

R-project logo

Challenge

CRAN has a Repository Policy, which submitted packages need to comply with. This policy includes topics such as metadata content, licenses, documentation, package structure, maintainer information, etc. In addition, packages need to be operable in different operative systems, like different distributions of Linux and Windows. One of the challenges of CRAN is to continuously check the compliance of new and old repositories with the Repository Policy.

QC/QA actions Technical quality

The R CMD Check tool is an automatic quality control tool that is used to check the compliance of packages with the Repository policy. In addition, it also runs tests defined for each specific package by their respective developers. CRAN continuously re-checks all packages, especially when new R versions or updates from common packages are released.

QA/QC Results and their Communication

The results of the automatic tests are published in the website CRAN Package Check Results. There are multiple levels of compliance: OK, NOTE, WARNING and ERROR. The first two are acceptable for compliant packages, but WARNINGs indicate problems that could affect functionality. Packages with any ERRORs are rejected, while WARNINGs or NOTEs may lead to questions or rejection depending on severity.

Feedback channels

When a package is flagged as non-compliant, CRAN asks the maintainer to fix it, otherwise the package is archived, meaning that access to the package is temporarily blocked. In addition, CRAN gives package developers the option to provide a link for bug reports from users (e.g. a link to the package GitHub page). From the developers' side, they can reach the CRAN team via various e-mail channels.

5 Quality information storage, traceability and communication

This section addresses the storage, traceability and communication of information about quality. Aside from pulling information from literature review and our own expertise, we also used the content from the previous section on case-studies. This work has been mainly constrained to the communication of technical quality, and we welcome future initiatives to expand upon it to include other essential aspects of Climate Services like scientific quality, fitness for purpose and coproduction.

5.1 Automatic monitoring and communication

To manage the massive volume of data and software provided by modern Climate institutions, automatic monitoring tools have become indispensable. These systems ensure that strict technical requirements are met at every stage of the data lifecycle. However, the initial hurdle isn't just building the tools—it is defining the precise requirements that need to be monitored in the first place.

The process begins at the code level with checker libraries that allow developers to run compliance tests and receive immediate feedback. For instance, the Pydantic library is frequently used to ensure data matches the exact shape and type required by an API. In more specialised Climate contexts, tools like the Climate Model Output Rewriter (CMOR) check processor verify that model data adheres to CMIP standards, while AtMoDat tools ensure compliance with specific AtMoDat standards.

Once these tools are integrated into operational workflows, they often feed into publicly available dashboards. These platforms provide a "single source of truth" for both providers and users. For example, the Comprehensive R Archive Network (CRAN) publishes automatic compliance results for software packages, ensuring that users know exactly which tools meet the required standards before they download them.

Finally, when these monitoring tools are embedded within complex processing chains, they are supported by robust notification systems. These systems act as an early-warning mechanism, alerting developers to non-compliance or errors so they can be rectified before the data moves further downstream. Crucially, these notifications also serve as a communication bridge to the end-user, providing direct updates on any newly discovered issues that might impact data reliability. By linking automated checks with transparent communication, institutions can maintain high levels of trust in their Climate Services.

5.2 Feedback between users and providers

To build a service that is both useful and sustainable, user feedback must be integrated into the development process from the very beginning. Ideally, service creation should stem directly from verified user needs, which are often uncovered through participatory activities like workshops. Once a foundation is built, a "beta" version allows providers to observe how users navigate the service in a real-life setting.

To manage this influx of feedback efficiently, many teams adopt Continuous Integration and Continuous Delivery (CI/CD)[12]. By automating the building, testing, and deployment of code, CI/CD creates a tight feedback loop between developers and stakeholders. This DevOps practice ensures

[12] https://about.gitlab.com/topics/ci-cd/

that user insights from workshops and beta tests are translated into features and fixes rapidly, allowing the service to evolve in lockstep with user expectations while maintaining high software quality. As shown in the case studies, developers primarily use GitHub to collaborate on code, as it ensures the traceability of both software versions and issues.

As a service gains traction, managing the increasing volume of user requirements and development actions becomes a significant challenge, thus overwhelming user support. However, this complexity is a positive signal of a healthy, growing user base. To manage this at scale, the Copernicus Climate Change Service (C3S) serves as a prime example: they log every requirement into a centralised database, capturing technical details, context, and source. By continuously analysing and clustering these entries, C3S can objectively prioritise new features based on feasibility, user benefit, and cost-effectiveness within their current capacity.

This systematic approach to information management is particularly critical for organisations pursuing ISO 9000 (Quality Management) certification. At this level, the focus expands beyond user needs to include the tracking of quality standards from external providers and the documentation of continuous improvement measures.

Finally, because gathering direct feedback often involves handling personal or sensitive proprietary data, the process must be underpinned by a robust legal and ethical framework. Ensuring strict compliance with GDPR, and where applicable, the ISO 27001 information security standard, is not just a legal necessity—it is the foundation of the user trust required for long-term service sustainability.

5.3 Storage of quality information and metadata

As the preceding case studies demonstrate, many institutions store quality control and assurance information alongside their datasets. However, the sheer diversity of formats—ranging from exhaustive technical reports to metadata—often makes it difficult for users and machines to locate and interpret the information they need.

A primary effort to standardise the provision of quality information is ISO 19157-1:2023 Geographic information — Data quality. ISO 19157 provides a comprehensive workflow for reporting quality and verification results. The standard dictates that data quality should be reported as structured metadata, which enhances interoperability and enables automation. This is often supplemented by a standalone, human-readable quality report. To maintain a "single source of truth", the structured metadata should include a direct reference to this report, ideally through Persistent Identifiers (PIDs).

Some key principles within the ISO framework are:

- **Hierarchical Inheritance**: Quality specifications defined at an upper level (e.g., a full dataset) automatically apply to lower levels (e.g., individual subsets) unless otherwise specified.
- **Structured Quality Units**: To ensure that quality reports are consistent and easy to compare, ISO 19157 organises information into "Quality Units." Each Quality Unit is structured in two parts, with the latter having subcategories:
    - **The Scope**: "*What is the specific boundary of the evaluation?*"
      It could be the entire global dataset, or it could be narrowed down to a specific region or variable.
    - **The Quality Elements**

    - **The measure**: "*What kind of quality are we measuring?*"
      The standard requires you to pick a specific category of quality, such as, completeness, positional or thematic accuracy, temporal quality, etc.
    - **The Evaluation Method**: "*How did we check it?*"
    - **The result**: "*What did we find?*"
- **Metaquality**: Metaquality elements can be added to evaluate the quality assessment itself, effectively measuring the reliability of the result.
- **Temporal Traceability**: All evaluations must include the specific dates they were performed to ensure the assessment remains relevant as the service evolves.

While ISO 19157 provides the governance framework for what should be reported, the Climate and Forecast (CF) Convention provides the controlled vocabularies and technical mechanism for how that information is embedded directly into the data files.

In the CF framework, quality is handled primarily through ancillary variables and status flags. This allows for a direct, machine-readable link between a "physical variable" (such as Sea Surface Temperature) and its corresponding quality status. For example, a status flag can instantly inform a user—or an automated processing system—whether a specific data point is "good," "suspect," or "bad."

5.4 Provenance: Relationship between traceability, QC/QA and data stewardship maturity

Reliable quality assessment and sustained quality monitoring depend on the ability to trace how data *are produced, transformed, evaluated, and maintained over time*. Without traceability, quality claims remain opaque, difficult to verify, and hard to compare across datasets, organisations, or processing histories. Effective traceability provides the evidential backbone that links measurements to reference standards, documents the assumptions and decisions embedded in processing workflows, and records how quality characteristics evolve through reprocessing and stewardship activities. It enables users to judge fitness for purpose, supports interoperability and reuse, and allows data providers to monitor, diagnose, and improve quality in a systematic way.

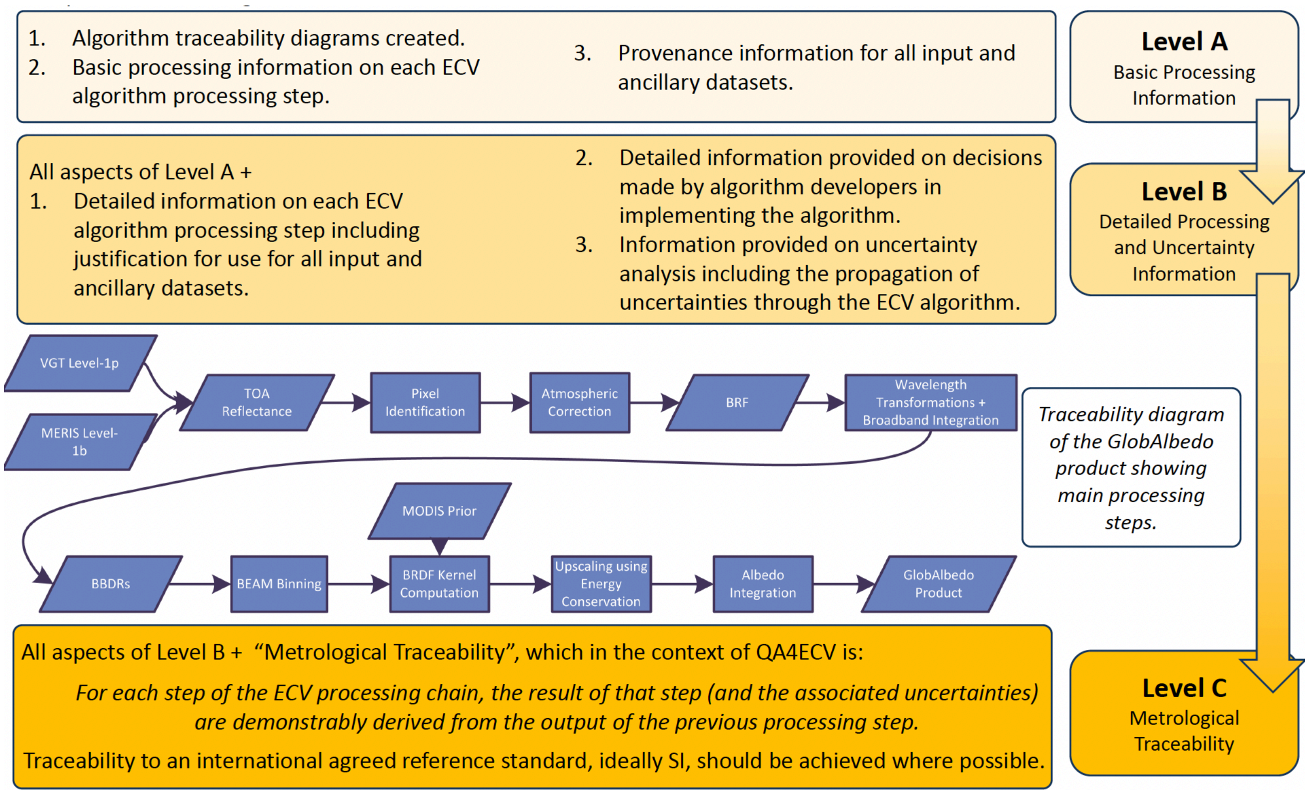
For instance, the ESMValTool and the CMIP Rapid Evaluation Framework (REF), mentioned in the CMIP case study sections, represent coordinated efforts to enable traceability and reproducibility in Climate model evaluation. ESMValTool provides an open-source workflow for diagnostics and performance evaluation of Earth System Model outputs and offers an initial implementation that captures detailed provenance information on datasets, processing steps, software versions, and contributors. Building on similar principles, REF applies automated diagnostics and benchmarking to CMIP simulations to support assessments such as those used in IPCC reports, with strong requirements for traceability. As already discussed in D2.2 and D2.7, these developments are still work in progress, particularly regarding the adoption of common data models and interoperable provenance approaches that could support traceability more systematically across Climate research infrastructures within a unified solution.

Other general purpose frameworks such as Quality Assurance Framework for Earth Observation (QA4EO) and the Data Stewardship Maturity Matrix (DSMM) offer complementary but distinct approaches to structuring, implementing, and evaluating traceability in support of robust quality assessment and long-term data stewardship. While QA4EO focuses on how traceability should be constructed within processes, DSMM focuses on how well traceability is actually implemented and sustained at the dataset level, making it a management and governance instrument as much as a technical one.

5.4.1 Quality Assurance Framework for Earth Observation (QA4EO)

QA4EO[13] defines traceability as the *foundational mechanism by which data quality becomes demonstrable, comparable, and trustworthy*. Its core principle is that *all Earth observation data and derived products must have a documented, fully traceable Quality Indicator (QI)*, where traceability is explicitly tied to evidence, uncertainty, and, where possible, internationally agreed reference standards. Traceability is required *end-to-end across the data lifecycle*, from measurement and calibration, through processing and dissemination, with quality information propagated through every step of the chain. In effect, QA4EO frames traceability as a structured audit trail that enables users to independently assess "fitness for purpose," and allows quality claims to be justified rather than asserted.

While the core QA4EO *principles document* dates to 2010 and remains the foundational reference, the framework has been further developed through projects[14], best-practice guides, and practical implementations. For instance, elements of QA4EO guidance (especially on *provenance, uncertainty propagation, and traceability*) have been incorporated into related Earth observation quality initiatives such as QA4ECV (Quality Assurance for Essential Climate Variables), which uses QA4EO principles for uncertainty and validation workflows for Climate-oriented products (Scanlon et al., 2015 and 2016). In the QA4ECV approach (see Figure X), data provenance and lineage are combined with metrological traceability according to a multilevel framework.



[13] https://qa4eo.org/

[14] www.gfz.de/en/section/remote-sensing-and-geoinformatics/projects/qa4eo2-quality-assurance-framework-for-earth-observation

Figure 2: QA4ECV Traceability framework, showing the diagram of the GlobAlbedo products and the description of the three traceability levels applied.

5.4.2 Data Stewardship Maturity Matrix (DSMM)

DSMM reframes traceability from a principle into a *measurable, progressive stewardship capability.* It treats *Transparency/Traceability* as one of nine distinct maturity components and evaluates it on a five-level scale from ad hoc to optimal, recognising partial and incremental implementation. Traceability in DSMM is explicitly evidence-based: maturity ratings must be justified by documented practices, applied standards, metadata, reports, and artefacts, rather than by intent or policy statements.

Maturity Scale / Key Component	Level 1 Ad Hoc Not Managed	Level 2 Minimal Managed Limited	Level 3 Intermediate Managed Defined, Partially Implemented	Level 4 Advanced Managed Well-Defined, Fully Implemented	Level 5 Optimal Level 4 + Measured , Controlled , Audit
Data Quality Assurance	Data quality assurance (DQA) procedure unknown or none	Ad Hoc and random DQA procedure not defined and documented	DQA procedure defined and documented and partially implemented	DQA procedure well documented, fully implemented and available online with master reference data Limited data quality assurance metadata	Level 4 + DQA procedure monitored and reported Conforming to community quality metadata & standards External review
Data Quality Control/ Monitoring	None or Sampling unknown or spotty Analysis unknown or random in time	Sampling and analysis are regular in time and space Limited product-specific metrics defined & implemented	Level 2 + Sampling and analysis are frequent and systematic but not automatic Community metrics defined and partially implemented Procedure documented and available online	Level 3 + Anomaly detection procedure well-documented and fully implemented using community metrics, automatic, tracked and reported Limited quality monitoring metadata	Level 4 + Cross-validation of temporal & spatial characteristics Physical consistency check Conforming to community quality metadata & standards Dynamic providers/users feedback in place
Data Quality Assessment	Algorithm/method/model theoretical basis assessed (methods and results online)	Level 1 + Research product assessed (methods and results online)	Level 2 + Operational product assessed (methods and results online)	Level 3 + Quality metadata assessed Limited quality assessment metadata	Level 4 + Assessment performed on a recurring basis Conforming to community quality metadata & standards External ranking
Transparency /Traceability	Limited product information available Person-to-person	Product information available in literature	Algorithm Theoretical Basis Document (ATBD) & source code online Dataset configuration managed (CM) Unique Object Identifier (OID) assigned (dataset, documentation, source code) Data citation tracked (e.g., utilizing Digital Object Identifier (DOI) system)	Level 3 + Operational Algorithm Description (OAD) online, OID assigned, and under CM	Level 4 + System information online Complete data provenance online

Figure 3: DSMM, showing different levels of maturity Levels combining aspects of Quality and Traceability (Peng, 2015).

Crucially, DSMM operationalises traceability as a monitoring tool over time, through baselining, reassessment, versioned maturity reports, and integration into dataset-level metadata, so that improvements or regressions in quality practices can be tracked, compared, and communicated to both human and machine users.

The DSMM is under refinement and adoption by WMO with particular focus on Climate Data (SMM-CD), with a condensed version with reduced complexity (SMM-CD_NRP), to address the National and Regional Products developed by smaller institutions with limited data management resources. The WMO SMM CD and SMM CD NRP go beyond FAIR assessments by evaluating broader data stewardship practices aligned with the ISO OAIS Reference Model, which also underpins CoreTrustSeal (see Section 6.3) requirements. By capturing dataset level stewardship maturity, they support key TRUST principles, particularly transparency, responsibility, and user focus, while recognising that individual datasets may still be hosted in repositories that do not themselves meet CoreTrustSeal or TRUST standards.

According to the implementation guidelines[15] provenance information in the SMM-CD (and SMM-CD_NRP) should include details of any *quality control or homogenisation processes*, details of disaggregation or infilling, or any other changes made to the dataset. More specifically, in the SSM-CD, quality management covers QA, QC, and quality assessment, including how dataset reliability is communicated. At lower maturity levels, monitoring and correction are weak or inconsistently applied, while higher levels involve a well-managed, end-to-end QA process with full traceability, including retrospective QC of historical records. It also encompasses uncertainty analysis and data integrity checks to ensure data remain consistent and uncorrupted throughout their lifecycle.

In the WMO SMM, quality control relies heavily on Contextual Metadata to assess the reliability and fitness-for-purpose of data. Contextual Metadata and provenance are closely related elements of data stewardship matrix adopted by WMO. Their role spreads across many of the conceptual aspects of the framework, depending on increasing maturity levels. At lower maturity levels, contextual metadata is often limited to basic information such as where and when data were collected, while provenance remains implicit and undocumented. As maturity increases, contextual metadata becomes more complete and structured, and provenance starts to explicitly link data to their sources and key processing steps.

At higher maturity levels, contextual metadata and provenance are fully integrated. Contextual metadata captures authoritative, versioned information about data collection or generation, while provenance records the sequence of transformations applied to the data over time. Together, they support reproducibility, auditability, and fitness-for-purpose assessments, forming a solid foundation for quality assurance and trustworthy use of Climate data and services.

5.5 Scientific discussions within the community

The quality of Climate Services isn't only about technical compliance, but also includes other aspects like scientific quality, which Climate Services are fundamentally rooted in; consequently, scientific advancements serve as the primary engine for service innovation. This evolution relies heavily on transparent technical discussions occurring across a spectrum of platforms—from informal social media exchanges and private dialogues to formal conferences and peer-reviewed literature. Because the community of actors is so diverse, these spaces are essential for forging agreements.

However, "agreement" in science is rarely a binary state. It is a nuanced spectrum representing the level of endorsement an idea has reached versus the threshold required to claim a true consensus. Reaching this point requires immense effort: the first hurdle is a universal openness to debate, and the second is the difficult task of establishing a shared technical language.

This process is seldom linear. What is accepted as fact today may be refined or disproven tomorrow. Unfortunately, despite the pursuit of objectivity, scientific endorsement is not always dictated solely by the presentation of data, but is prone to cognitive biases. Furthermore, transparency—the lifeblood of scientific verification—is often hindered not just by a lack of intent, but by insufficient resources or structural complexities.

[15] https://library.wmo.int/idurl/4/68238

The growth of the private market in Climate Services introduces another layer of difficulty. Private firms must protect proprietary knowledge, which often includes their core scientific methodologies. While some private actors publish in peer-reviewed journals, these contributions are rarely fully transparent and seldom represent the true state-of-the-art. Given that public findings often trigger heated debates, the fact that private methodologies remain shielded from external review suggests a critical need for standardised transparency requirements.

Finally, one should note that despite the good will of many scientists, inconsistencies and healthy professional disagreements are sometimes weaponised politically to undermine the credibility of the field. This external pressure creates a paradox: it makes open scientific discussion a sensitive, high-risk endeavor, even though that very openness is the mechanism that ensures quality and builds long-term public trust in Climate Services.

5.6 Labels, badges and logos

While the terms labels, badges, and logos are often used interchangeably in the marketplace, they all serve the same core purpose: acting as a visual shorthand for compliance with a specific set of requirements. These markers provide a highly recognisable and efficient method of quality communication, offering immediate "at-a-glance" assurance to those familiar with their underlying standards.

However, the effectiveness of these visuals depends heavily on the user's prior knowledge. For those unaware of the specific criteria behind a mark, the landscape can become confusing—particularly when multiple organisations issue different labels to convey similar concepts. Without standardisation or clear messaging, this proliferation can lead to "label fatigue," where the sheer volume of badges diminishes the perceived value of each individual one.

Practical applications of these labels within the community are detailed in the following sections:

- EOSC, OGC, and AtMoDat: See the case studies in Section 4.
- CoreTrustSeal: See Section 6.3 for a deep dive into this specific quality mark.

5.7 Climate Services provision through Large Language Models

Recently, several major actors, like C3S, have started advocating for Large Language Models to solve two major challenges within the community of Climate Services (Dee et al., 2024):

1. **Resource discoverability**: With increased time and effort invested in Climate research, the volume of available resources grows, making it increasingly difficult for users to identify relevant datasets for their specific applications.

2. **Understanding**: Climate information is often a highly technical resource, rendering a deep understanding of its origin, applicability, and limitations inaccessible to many users.

The combination of these two factors—amount of resources and lack of capacity to deep-dive—has led some actors to start considering the use of chatbots. Chatbots are highly interactive tools that accommodate the knowledge of the user, and deliver customised responses. In a way, chatbots represent an attempt to provide democratic access to “experts” for the general public. Some examples of such chatbots include: ClimSight (Koldunov and Jung, 2024; Kutnetsov et al., 2025), ChatClimate

(Vaghefi et al., 2023), ClimateGPT (Thulke et al., 2024), ClimateQ&A (De la Calzada et al., 2024), askWMO[16], and CHAT IPCC[17].

The development of these chatbots is not exempt from challenges. Aside from generic ones related to LLMs, here we list some that are particularly relevant in this context:

- **Having an updated knowledge-base:** Currently, chatbots access the latest information through two primary mechanisms: (a) direct internet access and (b) Retrieval-Augmented Generation (RAG). RAG couples the chatbot with an updated database, which it searches to contextualise its responses (Lewis et al., 2020).

- **Providing accurate responses**: Chatbots are powered by Large Language Models (LLMs) which, at their core, are statistical engines that predict text based on a given input. This predictive process is susceptible to errors—not only grammatical but also scientific—also called hallucinations. A primary challenge for developers lies in training models to provide scientifically valid answers, particularly as scientific consensus is complex and continuously evolving. Developing a tool that can identify established concepts while recognising nuance, or even acknowledging its own limitations in discerning an answer, is a significant undertaking. Recently, this has led to the rise of 'agentic AI'—complex LLM architectures that are capable of writing and testing code, thus enabling them to go through complex custom analysis.

- **Getting agreeable responses or being influenced by them**: Agreeability reflects the fact that chatbots tend to agree with the statements given by the user. This is particularly critical if chatbots are supposed to substitute experts to guide users in decision-making. On the other hand, users can also be heavily influenced by chatbots with or without justification (Fang et al., 2025; Krügel et al., 2023; Krügel et al., 2025). In combination, agreeability and influence can create dangerous feedback loops of ideas that may lead to isolated perspectives. A situation which is very distant from the scientific philosophy of critical thinking.

- **Making quality differences understandable and actionable**: In order for LLMs to discern if a dataset is fit-for-purpose, they need to be trained to do so, and for that the quality of datasets needs to be determined, contextualised, and documented in a way that multiple datasets can be compared. As reported by C3S and EUMETSAT, the solution to this challenge is not trivial, and requires a major effort since quality reporting needs to be homogenised among all datasets.

- **Establishing unbiased discoverability strategies**: Since LLMs are statistical tools, some fear that resources that have a higher presence during the training of LLMs or simply are engineered to be more easily discovered in the internet will be recommended more frequently than others which appear less often. This is not only a problem of chatbots, but also of search engines, like Google Search. Ideally, all resources should be equally “discoverable” but this is not the case.

- **Complying with regulations:** The European AI Act (Regulation (EU) 2024/1689) establishes a risk-based framework for AI systems, where a system's classification determines its specific documentation requirements. While chatbots generally fall under the "limited risk" category, certain use cases can elevate them to "high risk." This escalation occurs if a chatbot automates critical decision-making or impacts essential infrastructure, such as water or electricity. Furthermore, because these systems often process sensitive personal data to provide customised responses, developers must look beyond the AI Act to ensure compliance with the GDPR and ISO/IEC 27001:2022 standards for information security.

[16] https://www.chatwmo.app/
[17] https://ipcc.chat/

6 Certification storylines

This section addresses the topic of certification within the community of Climate Services. It starts with an analysis of institutional certifications based both on literature review and insights coming from interviews. It then presents several certifications for personnel, and it finishes with a brief section on labels.

6.1 Certification for institutions

In order to support the delivery of quality-assured and trustworthy Climate information, Climate Services need to possess an adequate underlying organisation and infrastructure. Certification standards for institutions provide the relevant framework for Climate Services to build trust, credibility, and transparency. The need for certification can also arise from regulatory compliance and alignment with existing international standards (e.g. WMO, ICAO), or through the growing importance of certification requirements in procurement and tenders. More generally, certification contributes through standardised processes and documentation, risk management, service robustness, and long-term sustainability of the activity.

Beyond formal certification, institutional quality also depends on governance and maturity arrangements, including clear mandates, coordination mechanisms, and service delivery frameworks. In this respect, effective coordination between NMHSs and sectoral users, as well as structures such as National Framework for Climate Services (NFCS) where relevant, can provide an important foundation for the delivery of quality-assured and user-oriented Climate Services.

Certification standards relevant for Climate Services can cover different aspects of the Climate service chain, and thus present relative interests depending on the relative positioning of the certified institution in the provision chain. We will highlight in this section the three following ISO standards which were identified during the research process and during the interviews performed as part of this deliverable:

- **ISO 9001:2015 - Quality management systems**[18]: mainly covers process standardisation, customer satisfaction, and continuous improvement. Its applicability mainly concerns Climate service provision, project management, and organisation governance. A new version will be released in September 2026.

- **ISO/IEC 27001:2022 - Information security, cybersecurity and privacy protection — Information security management systems — Requirements**[19]: focuses on information security, and the confidentiality and integrity of data. This certification is highly relevant for Climate data, critical infrastructure, and client-sensitive information.

- **ISO/IEC 17025:2017 - General requirements for the competence of testing and calibration laboratories**[20]: is mostly relevant for certifying the competence of calibration laboratories, and guaranteeing the traceability of calibration and measurements. It is by design aimed towards meteorological and calibration laboratories.

While they represent some of the main certification standards, they do not represent in its entirety the existing certification ecosystem relevant for Climate Services.

---

[18] https://www.iso.org/standard/62085.html
[19] https://www.iso.org/standard/27001
[20] https://www.iso.org/ISO-IEC-17025-testing-and-calibration-laboratories.htmlClimate

6.1.1 Certification lifecycle and common mechanisms

While they pertain to different aspects and processes of Climate Services, the certification standards highlighted here share common facets regarding their lifecycle, ranging from the preparation for the certification to the internal and external audits, further emphasising certain general benefits they can bring to institutions. In the following, we provide a synthesis of common certification mechanisms applicable to Climate Services.

Throughout the certification process and its preparation, the institution might be faced with human and organisational challenges. As the organisational changes are focused around transforming practices from informal to standardised, staff resistance is not uncommon. A successful certification is dependent on top management engagement as resource and capacity constraints need to be addressed pre-emptively in order to avoid reticence and workload excess. Scaling progressively through a gradual rollout of processes across units can help alleviate some of these challenges.

A common aim of these certifications is to structure the institution around shared organisational blocks through formalised management systems, defined responsibilities, and clear process ownership. In practice, this requires the institution not only to document its procedures, but also to assign specific processes to designated process owners responsible for their implementation, monitoring, and quality performance. This helps ensure that responsibilities and authority for QA/QC, verification, corrective actions, and management review are clearly defined across the organisation. In addition, institutions need to establish recording and traceability mechanisms and to coordinate internal audits and management reviews. Decision-making processes, alongside corrective and preventive workflows, further help streamline and standardise management procedures and ensure their resilience and responsiveness.

An important contributing block of certification standards for institutions is the regular organisation of audits (internal and external). However, through the certification process, the audits go beyond a compliance check and embody a valuable learning tool. The crucial role of internal audits is to detect early any deviation, while external audits provide credibility through an independent assessment of the institution. Organised along a predefined cycle, the audits set the stage for a continuous improvement of the institution and its collaborators.

At the core of these standards also lies an essential element which backs up a great share of the implementation of the standard: documentation. In documentation reside the prerequisites for the transparency and reproducibility of processes. In the medium to longer term, documentation supports organisational memory and staff continuity. However, its implementation is not exempt of risks from being too separate from practice or due to documentation overload. More generally, in order to effectively help the standardisation process, documentation should aim to stay functional and maintainable.

It is not uncommon to call upon external expertise to analyse gaps, support training and design. External consultants can be a great resource to provide actionable interpretation of the standard requirements in the operational context. During the whole process, the transition can then be done from external support to internal ownership.

Lastly, the certification can represent a great opportunity as a management instrument. It provides a framework to evaluate performance and underpins evidence-based prioritisation. It creates a common language for quality and governance across teams, and can be embedded in the work culture rather than being a stand-alone label given to the institution.

6.1.2 Certification storylines and particularities

Even though the certification processes converge on some general aspects as highlighted previously, the three certification standards we draw attention to here have their own particularities with respect to the processes and topics they cover.

- The **ISO 9001** standard defines what a Quality Management System (QMS) should encompass. It focalises on laying down organisation-wide quality management practices by focusing on the consistency of processes, continuous improvement, and customer satisfaction. This standard is commonly required for national services in order to be aligned with recommendations from international bodies such as the World Meteorological Organization (WMO) or the International Civil Aviation Organization (ICAO), aiming for long-term quality and risk reduction for the Climate Services provided by the certified institutions. In national contexts, these certification approaches can also be understood in relation to broader Climate Services implementation structures, including NFCS and other institutional coordination mechanisms. Such structures help clarify mandates, roles, and responsibilities across the Climate Services value chain, including how QA/QC, verification, user engagement, and service delivery are organised in practice. In this sense, certification can reinforce not only internal institutional performance, but also the wider coordination arrangements that underpin the effective delivery of Climate Services at national level.

  On the other hand, this certification can help private companies build client trust and credibility, conform to requirements laid down by some actors (e.g. the European Space Agency or the European Organisation for the Exploitation of Meteorological Satellites), and form an advantage when competing in tenders. In practice, the implementation of this standard can materialise in different ways: mapping processes across operations and support, constructing a centralised QMS documentation that is linked to each ISO requirements, setting up performance indicators, tracking deviations and their resolution, or establishing review cycles. In general, the certification infuses clearer structure for activities and responsibilities in the Climate service, facilitates resource allocation, systematically integrates user feedback, and delivers consistency across services and projects.

- The **ISO 27001** standard targets information security, risk management, and data integrity and confidentiality. It is a requirement from institutional or private clients in order to handle their sensitive data and to protect their critical assets. Its implementation ranges from risk assessment routines to the definition and implementation of security controls. To ensure the resilience and preparedness of the organisation, incident management and planning are set in place. Security requirements coming from different clients need to be addressed, and it is common to rely on non-disclosure agreements during the audits to shield any access to sensitive data. For Climate Services, this standard positively impacts client confidence and trust, and further raises internal awareness on information security risks. It is also a key element for the Climate service to be eligible for sensitive contracts.

- The **ISO 17025** standard is primarily aimed at calibration and meteorological laboratories as highlighted previously. It focuses on the calibration of instruments and processes, and the validation and traceability of measurements. It is fundamental for national meteorological and calibration laboratories in order to demonstrate their technical competence and credibility in the calibration of meteorological instruments and in supporting international or national observation networks. In practice, it entails building technical documentation for methods, validation, and uncertainty quantification of the measurements. Calibration records need to be put in place alongside traceability chains. Proof of staff competence and training (cf. following section) needs to be demonstrated, and audits are performed with technical evaluators. Naturally, this standard influences the scientific and technical reputation of the organisation. It contributes to the confidence in the measurement and calibration results, furthering the international and national recognition of the Climate service.

6.2 Certification for experts

Several professional certifications are available worldwide for meteorologists, Climate service providers and data managers. Certifications demonstrate competence, credibility, and adherence to high ethical standards, while enhancing professional recognition and supporting career development. They also connect professionals to trusted networks for collaboration, knowledge exchange, and shared expertise. Periodic renewal requirements encourage continuous professional development and engagement with the latest research, tools, and best practices. The subsections below provide details on the **issuing organisations**, **eligibility requirements**, and **application processes** for these certifications.

In addition to the professional certifications, there are also defined competences related to ISO standards:

- **ISO 14064:2018 Greenhouse gases:** This standard includes an internationally recognised framework that guides organisations in measuring, reporting, and verifying GHG emissions and removals. Certified Lead Verifiers provide independent and impartial verification in accordance with ISO 14064 and relevant international frameworks. ISO 14064 Greenhouse Gas Lead Verifier certification is issued by an accredited certification body after successful completion of the course.

- **ISO/IEC 17025 — Testing and calibration laboratories**: Personnel in an ISO/IEC 17025-accredited laboratory are required to have appropriate education, demonstrate technical competence, and undergo internal or external training (e.g., laboratory techniques, calibration, standard testing methods). The laboratory maintains records of staff qualifications to meet ISO/IEC 17025 requirements. Individuals do not receive an internationally recognised ISO/IEC 17025 certificate or license. However, they receive certificates for completing external training (e.g., on knowledge of the standard, internal auditing), while for mandatory internal training, individuals are granted authorisation upon successful completion of the exam.

Additionally, there are numerous courses that provide professional certification after completion. These courses are aimed for working professionals seeking continuing education in various areas related to Climate change, such as Climate change management[21], Climate change and health[22], Climate risk, valuation, and investing[23], sustainability reporting[24] and more.

[21] s://extension.harvard.edu/academics/programs/managing-for-climate-change-graduate-certificate/#what-you-ll-learn
[22] https:/Climateale.edu/cchcert/
[23] https://www.cfainstitute.org/programs/climate-investing-certificate
[24] https://sustainability.lehigh.edu/professional-sustainability-certifications

6.2.1 Certified Consulting Meteorologist (CCM)

About

Certified Consulting Meteorologist (CCM) Program was established by American Meteorological Society (AMS; USA) to provide qualified consultants for those who require advice and consultation in meteorology, ranging from specialised forecasts to engineering design support and expert testimony on weather-related court cases.

The CCM designation is granted only to those who demonstrate a broad background in meteorology together with a detailed knowledge in a particular field of specialisation. The list of CCMs across multiple specialties within the weather, water, and Climate enterprise is provided in a Weather and Climate Service Provider's Directory.

There is also an established Association of Certified Meteorologists (CCMs), a professional association of private sector consulting meteorologists, to better represent and market the diverse skillset of consulting members.

In addition to CCM, AMS offers three more certifications for: Certified Broadcast Meteorologist (CBM), Certified Digital Meteorologist (CDM), and Certified AMS Teacher Program (CAT).

Eligibility requirements

As a minimum, applicants must meet one of three requirements for background education and have a minimum of five years of work at the professional level in meteorology or a related field.

Application process

To earn the CCM credential, an expert must submit an application, obtain recommendations from three associates, pass a rigorous written examination, and successfully complete an oral examination before a national board of examiners.

Certification is valid for one year and may be renewed annually. To maintain certification, the expert must achieve an acceptable level of continuing professional development during the five-year period following initial certification or the completion of the most recent professional development cycle.

There is an application and annual renewal fee.

Validity

AMS certifications have been USA nationally recognised as the standard in meteorological certification since 1957.

6.2.2 Registered Meteorologist (RMet) and Chartered Meteorologist (CMet)

Context

The aim of professional accreditation is to provide professional qualifications in meteorology that assure clients, employers, and the public that individuals have attained and continue to maintain a specified level of professional knowledge and experience. The Royal Meteorological Society (RMetS; UK) is the body able to award Registered or Chartered status for Meteorologists.
Registered Meteorologist (RMet) accreditation can obtain someone with practical knowledge and experience in the field of meteorology. It was established in 2014.
Chartered Meteorologist (CMet) is the highest level of accreditation and it is awarded to meteorologists with specific skills, specialist knowledge and a wealth of experience. CMet was established in 1994.

Eligibility requirements

There are two distinct accreditation schemes, depending on whether the applicant seeks RMet or CMet status, with differing requirements in terms of scientific background, knowledge of meteorological science and practice, experience, and competence. However, both require the ability to communicate clearly in English, professional integrity, and a commitment to continuing professional development.
The schemes conform to the principles specified in the International Standard ISO/IEC 17024:2012 - *Conformity assessment — General requirements for bodies operating* certification of persons concerning the operation of certification bodies, as well as follow WMO standards.
Membership in the Society is required to apply for either scheme.
The Society also welcomes applicants from outside the UK who meet the requirements.

Application process

There are four stages when completing the first application (Completing the application, Society Evaluation, Workplace assessment/Interview process, and Decision), followed by a renewal every two years.
The Accreditation Board is responsible to the Council of the Royal Meteorological Society for recommending acceptance or rejection of applicants. The Board arrange:

- Interviewing Panels (IPs) and Work Place Assessment Panels (WPAPs) as appropriate to examine individual applicants, and;
- A Course Evaluation and Continuous Professional Development Panel (CE/CPD Panel) to examine courses claimed to satisfy the various criteria of the schemes and to audit all CPD returns submitted by RMets and CMets.

There is an application fee and annual renewal fee.

Validity

The RMetS is recognised in European and UK law as both the Competent Authority and the Regulatory Body for Meteorology in the UK.

Although it is not legally regulated internationally, these accreditations hold reputational value within the professional community.

6.2.3 The Certified Environmental Practitioner for Climate Change (CEnvP-CC)

About

The Certified Environmental Practitioner (CEnvP) Scheme is the certification program for environmental professionals across Australia and Aotearoa New Zealand. The scheme has 9 certifications, one of which is for Climate Change. It is externally audited and compliant with ISO 17024:2012.

This certification can also be understood within a broader quality management perspective, in which competency frameworks, training, and continuing professional development are essential to maintaining the quality and credibility of Climate Services.

Mandatory Climate reporting frameworks require organisations to provide comprehensive, transparent, and auditable disclosures of Climate-related risks and impacts. To meet these requirements, CEnvP offers certification in Climate Change defined as professional competency in the scientific principles, risk assessment methodologies, adaptation strategies, regulatory frameworks and greenhouse gas accounting practices related to environmental aspects of Climate change.

Eligibility requirements

The applicants have to fulfil educational qualifications and work experience in lieu of educational qualifications.

To be certified as a Climate Change Specialist, applicants must also hold CEnvP General certification.

Application process

Applicants have to provide referee statements, written demonstration and documentary evidence of key proficiencies and professional competency, as well as to document past and future commitment to professional development. As part of the application process, an online interview with a panel of three certified practitioners is required.

Successful applicants receive an official CEnvP certificate, a personalised CEnvP seal to use and are entitled to use the title 'Certified Environmental Practitioner' and the post-nominal letters 'CEnvP-CC' as follows: Mr/s NN BSc CEnvP-CC.

Evidence of Continuing Professional Development (CPD) must be provided on a two-year basis.

There is an application and annual renewal fee.

Validity

The CEnvP certification is primarily recognised in Australia and New Zealand as a professional standard, but it is not legally mandated in any country.

6.2.4 Certified Climate Change Professional (CC-P®)

About

Association of Climate Change Officers (ACCO) is the professional development organisation and cross-sector community of practice for individuals addressing Climate change within their organisations' operations and missions.

ACCO is issuing CC-P® certificates.

The CC-P® credential is designed for senior and mid-level professionals and practitioners across sectors—including business, government, academia, and non-profits—whose roles focus on or significantly intersect with Climate change, sustainability, resilience, and adaptation.

The credential signifies expertise in assessing Climate risks, reducing greenhouse gas emissions and resource vulnerabilities, navigating legal and policy frameworks, driving organisational change and stakeholder engagement, and building strong business cases for effective Climate response strategies.

Eligibility requirements

CC-P® applicants must have one of three defined combinations of education and professional working experience - prerequisites.

Application process

Applicants must demonstrate mastery of core Climate change competencies, and the ability to address operational, strategic, and financial implications of Climate change.

Applicants apply through the ACCO online portal, submit documentation of relevant education and professional experience, complete required training, pass the CC-P® examination, and receive approval upon review.

The CC-P® credential is valid for six years and requires annual continuing education and usage requirements.

There is an application and maintenance fee.

Validity

The CC-P® credential is internationally applicable and recognised across sectors, but it is not a legally regulated license in any specific country.

6.2.5 Sustainability & Climate Risk (SCR®) Certificate

About

Global Association of Risk Professionals® (GARP) is the international professional association for risk managers, dedicated to the advancement of the profession through education, research, and the promotion of best practices.

GARP's SCR Certificate is developed by risk, Climate, and sustainability experts. It provides knowledge on sustainability, Climate and nature risk, and transition planning.

Eligibility requirements

-

Application process

Applicants need to pass an exam. They are also encouraged to participate in GARP's Continuing Professional Development (CPD) program to stay up-to-date with evolving standards and practices in sustainability and Climate risk management.

There is an application fee.

Validity

-

6.2.6 The Certified Data Management Professional (CDMP®)

About

Global Association for Data Management Professionals (DAMA) is a globally recognised framework that defines the core principles, best practices, and essential functions of data management.

DAMA® established the Certified Data Management Professional® (CDMP®) certification. Certification validates expertise in data management across disciplines like: governance, metadata and reference data, quality, architecture, integration, interoperability, security, modeling and design. It is based on the DAMA Body of Knowledge (DMBOK®), a framework that defines the core principles, best practices, and essential functions of data management.

Eligibility requirements

There are three levels of certification (Associate, Practitioner or Master) depending on the experience.

Application process

After online registration, applicants have to pass the exams depending on the level of requested certification.

DAMA International provides a digital badge system that serves as proof of certification. The digital badge includes metadata detailing the certification's issuance date, level achieved, and the skills validated.

The certification is valid for three years, and for the renewal, applicants have to fulfill Continued Professional Development activities.

There is an application and an annual maintenance fee.

Validity

Certification is globally recognised.

6.3 Labels, badges and logos

This section expands upon the examples that we already presented in the case studies section coming from EOSC, OGC and AtMoDat.

6.3.1 CoreTrustSeal

CoreTrustSeal is an international, community-based, non-governmental, and non-profit organisation promoting sustainable and trustworthy data infrastructures. Their certificate, launched in 2017, defines requirements and offers core level certification for Trustworthy Data Repositories holding data for long-term preservation, based on the Core Trustworthy Data Repositories Requirements[25].

The CoreTrustSeal (CTS) resulted from an effort between the Data Seal of Approval (DSA) and the World Data System of the International Science Council (WDS), under the umbrella of the Research Data Alliance (RDA), to harmonise their data repository certifications. The certificate is awarded for three years according to regularly updated requirements. The path to successful certification involves the preparation of a self-assessment by following the applicable requirements[26]. This assessment is then evaluated by two independent reviewers and discussed by the CoreTrustSeal Board. To manage its finances, CoreTrustSeal is a legal entity under Dutch law governed by a Standards and Certification Board composed of 12 elected members representing the Assembly of Reviewers. The administrative fee is 3,000 EUR for each three-year certification.

From the perspective of repositories, CTS certification brings some advantages. On the one hand, the self-assessment allows for an in-depth look at one's own documentation and services, thereby promoting transparency. This offers a benchmark for comparison and helps to determine the strengths and weaknesses of a repository. In addition, there are also external benefits, such as building stakeholder confidence, enhancing the reputation of the repository, and demonstrating that the repository is following good practices. National and international funders are increasingly likely to mandate open data and data management policies that call for not only long-term , but also trustworthy storage and accessibility of data.

The CoreTrustSeal certification is envisioned as the first step in a global framework for repository certification which includes the extended level certification (nestor-Seal DIN 31644) and the formal level certification (ISO 16363). Ultimately, CoreTrustSeal plans to also provide core level certification for other research entities such as data services and software.

[25] https://doi.orClimate81/zenodo.17660463
[26] https://doi.org/10.5281/zenodo.17660463

7 Insights from interviews with the Climate Services community

This section presents the synthesised insights from a set of interviews with providers from both the public and private sector. The interviews were specifically conducted to inform the final deliverables from Work Package 2. Particularly, we would like to acknowledge the participation of representatives from EUMETSAT, KNMI, NebboWeather, Lobelia, MeteoClim del Grupo WDNA, Geoskop, and Eoliann. Their insights were gathered in separate one-to-one interviews, synthesised, compared, and mixed, so that they could be presented here anonymously. Note that the primary focus in some of the interviews was different than quality management, but are included here because they still provided valuable insights.

Disclaimer — The opinions presented here do not necessarily reflect the views of all companies and public institutions in the sector, may evolve over time, and should not be interpreted as the position of the authors of this document. In addition, given the limited scope of this work, the results should not be considered statistically representative of the market as a whole.

Nevertheless, we believe that this effort contributes to bridging the persistent gap between the private and public sectors, which rarely exchange views despite the clear need to do so in order to build an ecosystem that benefits all.

---

1. **High certification costs drive a preference for pragmatic best practices**
   Strict adherence to formal standards (e.g., ISO, AFNOR) can be prohibitively expensive in terms of time and resources. Consequently, many practitioners favor the adoption of recognised data management frameworks, such as DAMA-DMBOK, over formal accreditation.

2. **Ambiguous principles create implementation gaps that hinder adoption**
   High-level conventions like the FAIR principles are often perceived as undescriptive and open to interpretation. This lack of precision requires significant additional effort from providers to reach implementable and common technical solutions. This also applies to undescriptive standards and regulations.

3. **Institutional capacity limits compliance more than a lack of willingness**
   Adherence to WMO guidelines is often a matter of resource availability rather than intent. As Climate service standardisation moves forward, adoption may be hindered unless the sector addresses the underlying lack of technical and financial capacity.

4. **Standardisation could clarify technical differences between service providers**
   Beyond defining common baselines, standardisation efforts should strive to make the unique differences between providers transparent. This ensures that both technical and non-technical stakeholders can make informed comparisons.

5. **Tiered certification systems help users navigate diverse market needs**
   A system of certification with multiple quality levels enables users to find the specific offer that best suits their requirements. This approach supports price diversification and helps new providers establish credibility in a competitive market.

6. **General quality and security standards remain essential for sensitive data**
   In some markets, ISO 9001 (quality management) and ISO 27001 (information security) are the most frequently requested certifications, particularly when services involve the

processing of sensitive or proprietary client data.

7. **Unified uncertainty and validation frameworks could shift market focus toward technical performance**
   Establishing a common framework for uncertainty quantification and performance validation allows data from multiple providers to be compared objectively. This transition makes technical and scientific performance more important than the visual "aesthetics" of a service.
8. **Multi-institutional consortia provide the most credible quality guidelines**
   Guidelines developed by open consortia are less biased than those from private companies. Collaborative authorship is essential when seeking agreement between institutions operating across different sectors and markets.
9. **Major public initiatives act as de facto standardisation bodies**
   Large-scale actors like CMIP, the IPCC, and C3S effectively set industry standards. When these entities recommend a specific data format or tool, they drive widespread community uptake, particularly among those who post-process public data.
10. **For private providers, FAIR principles improve internal efficiency despite low external demand**
    Private-sector clients rarely request FAIR-compliant data due to the need for closed, customised access. However, these principles are highly valued internally for building the robust software infrastructure required for efficient operations.
11. **Operational efficiency necessitates strategic methodological simplifications**
    Scaling a service for operational use requires making assumptions about its applicability to avoid excessive costs. Providers may choose a globally viable methodology even if specialised methods exist for specific regions or timeframes.
12. **Client budgets directly influence the selection of scientific methodologies**
    The high cost of computational power means that methodological rigor is often dictated by the client's budget. For example, resource-intensive dynamic downscaling is frequently priced higher than statistical alternatives.
13. **Proprietary software remains the primary competitive edge for private providers**
    Private firms are often reluctant to provide open access to their software, as their unique code and processing logic represent the core value and differentiating factor of their commercial service.
14. **Structured disclaimers manage user risks and explain technical trade-offs**
    Providers use disclaimers to ensure users acknowledge the risks of misusing Climate data. These tools also clarify necessary trade-offs, such as the balance between rapid data delivery and the thoroughness of quality checks.
15. **Public institutions might struggle to communicate data quality effectively**
    While public bodies are not restricted (in general) by proprietary concerns, they often lack the capacity to maintain up-to-date, user-friendly platforms. This makes it difficult to communicate quality information to diverse audiences.
16. **Providers must balance data digestibility with scientific nuance**
    Users often lack the capacity or willingness to engage with highly detailed data. While providers strive to offer "digestible" information, they face the constant challenge that simplicity is highly subjective and varies by user.
17. **Institutional reputation often serves as a proxy for technical traceability**

Data provided by recognised authorities is often trusted without detailed documentation. For many users, the provider's status is a sufficient reason for trust, especially when they lack the resources to verify the data themselves.

18. **Knowledge gaps make decision-makers vulnerable to provider bias**
Decision-makers are rarely experts in Climate science and often seek data without defined quality requirements. This asymmetry makes them vulnerable to the opinions of professionals and leads to frequent misunderstandings.

19. **Limited methodological transparency protects private-sector commercial viability**
Private providers often cannot be fully transparent about their methodologies without risking their business model. While confidentiality agreements provide some flexibility, the "black box" nature of private services remains a significant debate.

20. **Peer-reviewed publications build market trust and signal technical excellence**
Private companies use scientific journals and conferences to validate their methodologies. Even if the published research differs from their internal state-of-the-art, it serves as a public acknowledgement that enhances their reputation.

21. **Subseasonal free trials allow users to verify performance before purchase**
The shorter timeframes of subseasonal scales, as opposed to decadal scales, allow providers to offer free trials. This enables customers to test products in real-time and verify performance before committing financially.

22. **Information decay limits the understanding of data limitations downstream.**
Downstream players often report a lack of clear information regarding the assumptions and limitations of the data they receive. This gap restricts the depth of quality information they can then pass on to the end-user.

23. **Comparative dataset analysis requires careful communication to avoid misuse.**
While comparing datasets is essential for scientific growth, communicating inconsistencies is a sensitive task. If handled poorly, these differences can be misused to spread disruptive or anti-scientific messages.

24. **Harmonised terminology improves data discoverability and comparability**
Standardising nomenclature helps users navigate the market and allows providers to communicate more efficiently. Common terms ensure that products are presented in a way that supports findability and interoperability.

25. **Defining "fit-for-purpose" metrics is essential for AI-driven discovery**
As investment in chatbots grows, the challenge remains to define data quality in a machine-readable way. Objective "fit-for-purpose" classifications are required to train large language models for accurate recommendations.

26. **Provider-agnostic chatbots ensure equitable access to Climate data**
To maintain a fair market, discovery tools like chatbots must remain unbiased and represent all available data. It is essential to develop systems that cannot be influenced to favor one dataset over another based on the provider.

8 References

Bahim, C., Casorrán-Amilburu, C., Dekkers, M., Herczog, E., Loozen, N., Repanas, K., ... & Stall, S. (2020). The FAIR data maturity model: An approach to harmonise FAIR assessments. Research Data Alliance.

Calzada, N. D. L., Costa, T. A. D., Blangero, A., & Chesneau, N. (2024). ClimateQ&A: Bridging the gap between Climate scientists and the general public. arXiv. https://arxiv.org/abs/2403.14709

Dee, D., Obregon, A., & Buontempo, C. (2024). Are our Climate data fit for your purpose? Bulletin of the American Meteorological Society, 105(9), E1723–E1733. https://doi.org/10.1175/BAMS-D-23-0295.1

Doblas-Reyes, F., Kontkanen, J., Sandu, I., Acosta, M., Al Turjmam, M., Alsina-Ferrer, I., Andres-Martínez, M., Anerdi, C., Arriola, L., Axness, M., Batlle Martín, M., Bauer, P., Becker, T., Beltrán, D., Beyer, S., Bockelmann, H., Bretonniere, P.A., Cabaniols, S., Caprioli, S., Castrillo, M., Chandrasekar, A., Cheedela, S., Correal, V., Danovaro, E., Davini, P., Enkovaara, J., Frauen, C., Fruh, B., Gaya \`Avila, A., Ghinassi, P., Ghosh, R., Ghosh, S., González, I., Grayson, K., Griffith, M., Hadade, I., Haine, C., Hartick, C., Haus, U.U., Hearne, S., Järvinen, H., Jiménez, B., John, A., Juchem, M., Jung, T., Kegel, J., Kelbling, M., Keller, K., Kinoshita, B., Kiszler, T., Klocke, D., Kluft, L., Koldunov, N., Kölling, T., Kolstela, J., Kornblueh, L., Kosukhin, S., Lacima-Nadolnik, A., Leal Rojas, J., Lehtiranta, J., Lunttila, T., Luoma, A., Manninen, P., Medvedev, A., Milinski, S., Mohammed, A., Müller, S., Naryanappa, D., Nazarova, N., Niemelä, S., Niraula, B., Nortamo, H., Nummelin, A., Nurisso, M., Ortega, P., Paronuzzi, S., Pedruzo-Bagazgoitia, X., Pelletier, C., Peña, C., Polade, S., Pradhan, H., Quintanilla, R., Quintino, T., Rackow, T., Räisänen, J., Rajput, M., Redler, R., Reuter, B., Rocha Monteiro, N., Roura-Adserias, F., Ruppert, S., Sayed, S., Schnur, R., Sharma, T., Sidorenko, D., Sievi-Korte, O., Soret, A., Steger, C., Stevens, B., Streffing, J., Sunny, J., Tenorio, L., Thober, S., Tigerstedt, U., Tinto, O., Tonttila, J., Tuomenvirta, H., Tuppi, L., Van Thielen, G., Vitali, E., Hardenberg, J., Wagner, I., Wedi, N., Wehner, J., Willner, S., Yepes-Arbós, X., Ziemen, F., & Zimmermann, J. (2026). The Destination Earth digital twin for Climate change adaptation. Geoscientific Model Development, 19(7), 2821–2848. https://gmd.copernicus.org/articles/19/2821/2026/

Dunne, J. P., Hewitt, H. T., Arblaster, J. M., Bonou, F., Boucher, O., Cavazos, T., Dingley, B., Durack, P. J., Hassler, B., Juckes, M., Miyakawa, T., Mizielinski, M., Naik, V., Nicholls, Z., O'Rourke, E., Pincus, R., Sanderson, B. M., Simpson, I. R., & Taylor, K. E. (2025). An evolving Coupled Model Intercomparison Project phase 7 (CMIP7) and Fast Track in support of future Climate assessment, 18, 6671–6700. https://doi.org/10.5194/gmd-18-6671-2025

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. Geoscientific Model Development, 9(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Eyring, V., Righi, M., Lauer, A., Evaldsson, M., Wenzel, S., Jones, C., Anav, A., Andrews, O., Cionni, I., Davin, E. L., Deser, C., Ehbrecht, C., Friedlingstein, P., Gleckler, P., Gottschaldt, K.-D., Hagemann, S., Juckes, M., Kindermann, S., Krasting, J., ... Williams, K. D. (2016). ESMValTool (v1.0) – a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. Geoscientific Model Development, 9(5), 1747–1802. https://doi.org/10.5194/gmd-9-1747-2016

Fang, C. M., Liu, A. R., Danry, V., Lee, E., Chan, S. W. T., Pataranutaporn, P., Maes, P., Phang, J., Lampe, M., Ahmad, L., & Agarwal, S. (2025). How AI and human behaviors shape psychosocial effects of extended chatbot use: A longitudinal randomized controlled study. arXiv. https://arxiv.org/abs/2503.17473

Gleckler, P. J., Doutriaux, C., Durack, P. J., Taylor, K. E., Zhang, Y., Williams, D. N., Mason, E., & Servonnat, J. (2016). A more powerful reality test for Climate models. Eos, 97. https://doi.org/10.1029/2016EO051663

Hewitt, C., Mason, S. & Walland, D. (2012) The Global Framework for Climate Services. Nature Clim Change 2, 831–832 . https://doi.org/10.1038/nClimate1745

Hoffman, F. M., Hassler, B., Swaminathan, R., Lewis, J., Andela, B., Collier, N., Hegedűs, D., Lee, J., Pascoe, C., Pflüger, M., Stockhause, M., Ullrich, P., Xu, M., Bock, L., Chun, F., Gier, B. K., Kelley, D. I., Lauer, A., Lenhardt, J., ... Dunne, J. P. (2025). Rapid evaluation framework for the CMIP7 Assessment Fast Track. EGUsphere. https://doi.org/10.5194/egusphere-2025-2685

Koldunov, N., & Jung, T. (2024). Local Climate Services for all, courtesy of large language models. Communications Earth & Environment, 5, 13. https://doi.org/10.1038/s43247-023-01199-1

Krans, N. A., Ammar, A., Nymark, P., Willighagen, E. L., Bakker, M. I., & Quik, J. T. K. (2022). FAIR assessment tools: Evaluating use and performance. NanoImpact, 27, 100402. https://doi.org/10.1016/j.impact.2022.100402

Krügel, S., Ostermaier, A., & Uhl, M. (2023). ChatGPT's inconsistent moral advice influences users' judgment. Scientific Reports, 13, 4569. https://doi.org/10.1038/s41598-023-31341-0

Kruegel, S., Ostermaier, A., & Uhl, M. (2025). ChatGPT's advice drives moral judgments with or without justification. arXiv. https://arxiv.org/abs/2501.01897

Kuznetsov, I., Jost, A. A., Pantiukhin, D., et al. (2025). Transforming Climate Services with LLMs and multi-source data integration. npj Climate Action, 4, 97. https://doi.org/10.1038/s44168-025-00300-y

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., tau Wen-Yih, Rocktäschel, T., Riedel, S., & Kiela, D. (2021). Retrieval-augmented generation for knowledge-intensive NLP tasks. arXiv. https://arxiv.org/abs/2005.11401

Ordonez, A., Ullrich, P., & Lee, J. (2025). cmecmetrics/cmec-driver: v1.1.9-release (v1.1.9-release) [Software]. Zenodo. https://doi.org/10.5281/zenodo.15548748

Peng, G., Privette, J. L., & Kearns, E. J., et al. (2015). A unified framework for measuring stewardship practices applied to digital environmental datasets. Data Science Journal, 13, 231–253. https://doi.org/10.2481/dsj.14-049

Scanlon, T., Nightingale, J., Fox, N., & Muller, J. P. (2015). QA4ECV: Developing metrological traceability through ECV products. https://doi.org/10.13140/RG.2.1.3123.2484

Scanlon, T., Muller, J. P., & Kharbouche, S. (2016). QA4ECV—Developing metrological traceability through the QA4ECV albedo product through the determination of Level 1 uncertainties. https://doi.org/10.13140/RG.2.2.27717.29920

Sun, C., Emonet, V., & Dumontier, M. (2022, January). A comprehensive comparison of automated FAIRness evaluation tools. In 13th International Conference on Semantic Web Applications and Tools for Health Care and Life Sciences (pp. 44–53).

Thulke, D., Gao, Y., Pelser, P., Brune, R., Jalota, R., Fok, F., Ramos, M., van Wyk, I., Nasir, A., Goldstein, H., Tragemann, T., Nguyen, K., Fowler, A., Stanco, A., Gabriel, J., Taylor, J., Moro, D., Tsymbalov, E., de Waal, J., … Erasmus, D. (2024). ClimateGPT: Towards AI synthesizing interdisciplinary research on Climate change. arXiv. https://arxiv.org/abs/2401.09646

Vaghefi, S. A., Stammbach, D., Muccione, V., et al. (2023). ChatClimate: Grounding conversational AI in Climate science. Communications Earth & Environment, 4, 480. https://doi.org/10.1038/s43247-023-01084-x

Waliser, D., Gleckler, P., Ferraro, R., Taylor, K., Ames, S., Biard, J., Bosilovich, M., Brown, O., Chepfer, H., Cinquini, L., Durack, P., Eyring, V., Mathieu, P., Lee, T., Pinnock, S., Potter, G., Rixen, M., Saunders, R., Schulz, J., & Tuma, M. (2020). Observations for Model Intercomparison Project (Obs4MIPs): Status for CMIP6. Geoscientific Model Development, 13(6), 2945–2958. https://doi.org/10.5194/gmd-13-2945-2020