



Toward Interoperable and Globally Integrated Open Science in Antarctic Research through Standards and Identifiers

Sofiia Zhrebchuk¹, Andriy Fedchuk¹, Yulia Kryvytska¹, Olha Fedchuk¹,
Iryna Tsybenko²

¹National Antarctic Scientific Center of Ukraine, Kyiv, Ukraine

²State Scientific and Technical Library of Ukraine, Kyiv, Ukraine

zerebcuksofia@gmail.com, andriyf@gmail.com,
juliyakrivickaya@gmail.com, olgav2002@ukr.net, tsyb2012@ukr.net

Abstract

This paper examines the role of standards, persistent identifiers, and interoperable infrastructures in enabling open science within Antarctic research, with a particular focus on the National Antarctic Scientific Center (NASC) of Ukraine. Antarctic science depends on long-term, multi-decadal datasets that are essential for understanding global climate processes, environmental change, and ecosystem dynamics. However, the sustainability and usability of these datasets are challenged by technological obsolescence, fragmentation, and operational constraints inherent to polar research.

The study analyses international practices in Antarctic data governance and highlights the importance of metadata standards, FAIR principles, and globally recognised identifiers such as DOIs, ORCID, and ROR. It demonstrates that open science is not merely a normative ideal but a structural necessity for ensuring data continuity, interoperability, and reproducibility across generations of researchers and national programmes. Drawing on comparative international models, the paper proposes a conceptual framework for implementing integrated data management systems at NASC, including the adoption of standardised metadata architectures, persistent identifiers, and alignment with global research infrastructures.

The results show that the systematic adoption of open standards enhances data discoverability, strengthens international collaboration, and secures long-term preservation of Antarctic datasets. The paper concludes that embedding open science principles into institutional practices is critical for increasing the global visibility and scientific impact of Ukrainian Antarctic research, while contributing to a shared, interoperable knowledge system supporting global environmental governance.

Keywords: Antarctic research; open science; interoperability; metadata; identifiers; FAIR principles; data governance.

1 Introduction

Antarctic research relies fundamentally on the continuous accumulation, preservation, and interpretation of diverse scientific data. These datasets ranging from long-term meteorological observations and cryospheric measurements to biological monitoring, geophysical surveys, and environmental assessments form the empirical foundation upon which present understanding and future projections of Antarctic systems depend. In Ukraine, the National Antarctic Scientific Center (NASC) plays a central role in ensuring that such information is systematically collected, curated, and made accessible to the research community. As the operator of Ukraine's long-standing Antarctic programme and station, the Centre is responsible not only for generating new data but also for maintaining their scientific integrity across decades.

The distinctive nature of Antarctic science places exceptional demands on how data are managed. Many research questions in polar regions can be addressed only through continuous observational series spanning forty to fifty years or more. Climate dynamics, ocean ice interactions, ecological shifts, and atmospheric processes reveal meaningful patterns only when examined over multi-decadal timescales. The longevity and uniqueness of these datasets create a requirement for robust mechanisms that can safeguard them against technological obsolescence, data fragmentation, and loss of provenance. Moreover, the global relevance of Antarctic science and its implications for climate models, biodiversity assessments, and international environmental governance necessitates data practices that facilitate transparency, reproducibility, and international comparability.

In this context, standards, persistent identifiers, and interoperable information infrastructures emerge as essential components rather than optional enhancements. Metadata standards ensure that datasets remain interpretable across generations of researchers and software systems. Globally recognised identifiers such as DOIs for datasets, ORCID identifiers for researchers, and ROR identifiers for institutions support persistent attribution, traceability, and accurate linkage between data, publications, people, and research facilities. Interoperable infrastructures, built on open protocols and harmonised schemas, enable Antarctic data produced by NASC to be efficiently integrated into regional and global repositories, incorporated into cross-disciplinary models, and reused in collaborative international research.

The aim of this paper is to articulate the conceptual principles and practical considerations that should guide NASC in adopting such standards, identifiers, and open-science infrastructures. By outlining a cohesive vision for data interoperability and long-term stewardship, the article seeks to contribute to the establishment of a sustainable, globally aligned framework that will enhance the resilience, accessibility, and scientific value of Ukraine's Antarctic research outputs.

2 Methodology

This study employs a qualitative research design combining conceptual analysis, comparative review, and practice-informed reflection. The methodological approach is grounded in the assumption that Antarctic data governance cannot be adequately examined through purely quantitative measures, as it is shaped by institutional practices, international frameworks, and long-term infrastructural dynamics.

The research is based on a systematic review of academic literature, policy documents, and technical guidelines related to Antarctic data management and open science. Sources include peer-reviewed publications, reports from international organisations (e.g., SCAR, ATCM, COMNAP), and documentation of national Antarctic data centres. Particular attention was given to materials addressing metadata standards, FAIR principles, persistent identifiers, and data governance models.

The literature was analysed using a thematic approach, allowing the identification of recurring patterns, gaps, and best practices across different national contexts.

In parallel, a comparative analysis of established international infrastructures was conducted, focusing on selected national Antarctic data centres (such as those of Australia, the United States, and the United Kingdom). These cases were examined to identify common organisational, technical, and policy elements that support long-term data stewardship, interoperability, and open access. The comparison emphasised practical mechanisms, including data policies, repository architectures, metadata frameworks, and identifier systems.

An important component of the study is practice-based insight derived from the authors' direct involvement in Antarctic research coordination and data management processes within the National Antarctic Scientific Center. This insider perspective enabled a more nuanced understanding of institutional constraints, operational workflows, and ongoing efforts to align national practices with international standards. Such experiential knowledge was used to contextualise the findings of the literature review and to assess the feasibility of proposed solutions.

The final stage of the research involved the synthesis of analytical and empirical insights into a conceptual model for implementing standards, persistent identifiers, and interoperable infrastructures at NASC. The proposed framework reflects both international best practices and the specific institutional and geopolitical conditions within which Ukrainian Antarctic research operates.

3 Antarctic Research as a Long-Term Global Resource

Antarctic science is distinguished by its reliance on sustained observational programmes extending over multiple decades. These long-term records represent one of the few continuous baselines available for assessing variability and change within the Earth's climate system. Their value lies in enabling detection of low-frequency signals that are otherwise obscured by short-term variability and observational noise (Abram et al., 2013; Stauffer, 2000).

Meteorological monitoring networks, including long-established stations such as Vernadsky, Amundsen–Scott, Halley, and Mawson, provide uninterrupted measurements of key atmospheric variables such as temperature, pressure, wind regimes, precipitation, and radiative fluxes. When integrated across national programmes, these datasets form a quasi-continuous circumpolar archive that supports the analysis of interannual to multi-decadal variability and atmospheric circulation dynamics (Hughes et al., 2021; Rogers et al., 2019). Such coordinated observational systems are increasingly recognised as essential infrastructure for climate system attribution studies and model validation.

Equally critical are ice-core archives, which preserve stratified records of atmospheric composition and climatic conditions spanning hundreds of thousands of years. Through the analysis of stable isotopes and trapped gas inclusions, these records enable the reconstruction of palaeoclimate variability and provide essential context for modern anthropogenic change (Lorius, 1992; Jong et al., 2022; Srivastava et al., 2023). Recent high-resolution ice-core datasets from East Antarctica further enhance temporal precision in accumulation-rate reconstructions, improving understanding of regional mass balance dynamics and surface–atmosphere interactions (Jong et al., 2022; Srivastava et al., 2023).

Long-term atmospheric chemistry observations, particularly of ozone, greenhouse gases, and aerosol loading, constitute another fundamental pillar of Antarctic science. These records have been instrumental in identifying anthropogenic perturbations of the stratosphere and in documenting the evolution and partial recovery of the Antarctic ozone hole. More broadly, they provide empirical constraints for radiative forcing estimates and atmospheric transport models, thereby improving global climate projections (Stauffer, 2000; Hughes et al., 2021; Abram et al., 2013).

In parallel, oceanographic and ecosystem-based observations, often coordinated through international frameworks such as CCAMLR, capture variability in sea-ice dynamics, biogeochemical cycles, and biological productivity across the Southern Ocean. These integrated datasets underpin ecosystem-based management strategies and support assessments of biodiversity resilience under climate forcing (Rogers et al., 2019; Gutt et al., 2022). Recent synthesis efforts increasingly emphasise the coupling between physical and biological systems, particularly in relation to rapid cryospheric change and its cascading ecological impacts.

The scientific value of Antarctic datasets is fundamentally rooted in their temporal continuity. Multi-decadal observational series enable the detection of gradual system transitions, threshold behaviours, and episodic extremes that cannot be resolved through short-term campaigns alone (Abram et al., 2013; Hughes et al., 2021). Phenomena such as shifts in sea-ice extent, atmospheric circulation regimes, and stratospheric temperature trends only become statistically robust when evaluated across extended temporal windows (Srivastava et al., 2023; Rogers et al., 2019).

A prominent example of the power of sustained observation is the discovery of the Antarctic ozone hole. Its identification and subsequent characterisation were only possible through the integration of long-term, multi-agency datasets and open data exchange practices (Stauffer, 2000; Hughes et al., 2021). This case remains a paradigmatic demonstration of how interoperable data infrastructures and international collaboration can enable transformative scientific discovery in Earth system science.

However, maintaining continuity in Antarctic observational records remains structurally challenging. High personnel turnover associated with seasonal deployment cycles introduces methodological discontinuities and limits institutional memory retention (Hughes et al., 2021). Similarly, instrument replacement, calibration drift, and station relocation can introduce subtle but significant biases in long-term time series if not rigorously documented and homogenised (Stauffer, 2000; Gutt et al., 2022). These challenges highlight the importance of standardised metadata practices and persistent identifiers to ensure dataset traceability.

Beyond logistical constraints, Antarctic science is also shaped by broader governance and funding structures. Variability in national research priorities can affect the continuity of long-term monitoring programmes, particularly those requiring sustained multi-decadal commitment (Hughes et al., 2021; Rogers et al., 2019). In this context, international coordination mechanisms and data governance frameworks such as those promoted by SCAR and related data management initiatives play a critical role in ensuring interoperability, openness, and long-term preservation of polar datasets (SCAR, n.d.; SCADM et al., 2025; SOOS DMSC, 2022). Strengthening alignment with FAIR data principles is increasingly recognised as a strategic requirement for the resilience of polar research infrastructures (Polar Data Catalogue, n.d.; United States National Science Foundation, n.d.).

The Antarctic observational system represents a globally significant scientific asset. Its continued effectiveness depends not only on technological advancement, but also on institutional stability, coordinated governance, and sustained international collaboration across political and disciplinary boundaries (Secretariat of the Antarctic Treaty, n.d.; Walton, 2011).

4 The Antarctic Treaty as the First International Model of Open Science

The intellectual and geopolitical landscape of the 1950s was characterised by a profound restructuring of scientific production, circulation, and governance. In the aftermath of the Second World War, scientific research became increasingly embedded within national security architectures, with strategic disciplines such as nuclear physics, geophysics, and atmospheric science frequently subjected to classification regimes and restricted dissemination protocols (Walton, 2011; SCADM et

al., 2025). This epistemic fragmentation reflected broader Cold War dynamics, in which knowledge itself became a form of geopolitical capital. As a result, scientific openness was often subordinated to state security imperatives, generating asymmetries in access to data and limiting transnational epistemic integration.

Simultaneously, however, the post-war period also witnessed a countervailing trend toward institutionalised international cooperation in science. Large-scale collaborative initiatives, most notably the International Geophysical Year (1957–1958), marked a pivotal shift toward coordinated global observation systems and standardised data exchange practices (Walton, 2011). These efforts reflected an emerging recognition that certain Earth system processes, particularly those related to polar dynamics, atmospheric circulation, and ocean–cryosphere coupling, could not be adequately understood within isolated national research frameworks. Instead, they required synchronised, transboundary observational infrastructures and shared analytical protocols.

Within this dual context of scientific fragmentation and emergent cooperation, Antarctica acquired exceptional epistemic and geopolitical significance. The continent’s geographical remoteness, absence of an indigenous population, and extreme environmental conditions rendered it both strategically neutral and scientifically invaluable. It functioned simultaneously as a natural laboratory for global system processes and as a politically ambiguous space in which competing territorial claims could be reconfigured into cooperative arrangements. This convergence of scientific and diplomatic interests ultimately facilitated the negotiation of the Antarctic Treaty (1959), which institutionalised a novel governance regime centred on peaceful use and scientific exchange (Secretariat of the Antarctic Treaty, n.d.; Walton, 2011).

Article III of the Antarctic Treaty represents one of the earliest formalised international commitments to structured scientific openness. It obliges Parties to exchange scientific observations, results, and personnel “to the greatest extent feasible and practicable,” thereby establishing a normative framework for systematic knowledge circulation (Secretariat of the Antarctic Treaty, n.d.). Importantly, this provision extends beyond finalised research outputs to include observational data and methodological transparency, effectively embedding data-sharing obligations within a legally binding international treaty. In doing so, it anticipates later developments in open science governance by several decades.

The requirement for the exchange of scientific personnel among expeditions further strengthens this framework by enabling epistemic convergence through direct field-based collaboration. Such mobility facilitates methodological harmonisation, reduces observational divergence, and enhances inter-calibration of instruments and protocols. Additionally, the obligation to provide advance notification of research activities introduces an early form of coordinated research planning, effectively functioning as a precursor to modern research coordination frameworks and distributed observatory systems (Secretariat of the Antarctic Treaty, n.d.; Scientific Committee on Antarctic Research, n.d.).

Beyond its procedural provisions, Article III also had profound structural consequences for the organisation of polar science. It catalysed the emergence of transnational scientific governance institutions, most notably the Scientific Committee on Antarctic Research (SCAR), which plays a central role in coordinating research priorities, standardising metadata practices, and promoting interoperability across national programmes (Scientific Committee on Antarctic Research, n.d.). Alongside SCAR, specialised bodies such as the Standing Committee on Antarctic Data Management (SCADM) have further institutionalised data stewardship practices, reinforcing long-term preservation, discoverability, and cross-platform integration of polar datasets (SCADM et al., 2025).

From an analytical perspective, the Antarctic Treaty system can be interpreted as an early experiment in depoliticised knowledge governance. By insulating scientific activity from territorial disputes and geopolitical competition, the Treaty effectively constructed a semi-autonomous epistemic domain in which data circulation is normatively decoupled from sovereignty claims. This

represents a significant deviation from prevailing Cold War knowledge regimes, where scientific information was often treated as strategically sensitive and selectively disclosed (Walton, 2011).

This depoliticisation has had enduring epistemological consequences. Antarctic science increasingly operates under a norm of epistemic collectivism, where the validity and utility of data are maximised through openness, replication, and integration across national datasets. In this sense, the Treaty system can be interpreted as an early institutional precursor to contemporary open science frameworks, including FAIR data principles, which emphasise findability, accessibility, interoperability, and reusability of scientific outputs (SCADM et al., 2025; Polar Data Catalogue, n.d.).

However, the Antarctic system does not merely reflect scientific idealism; it also functions as a pragmatic mechanism for stabilising international cooperation in a strategically sensitive region. By embedding scientific collaboration within a legally binding treaty, the system transforms knowledge production into a diplomatic instrument that reinforces peace through shared epistemic dependency. Scientific infrastructure thus becomes both a research asset and a governance technology that sustains long-term international engagement (Secretariat of the Antarctic Treaty, n.d.; Walton, 2011).

In contemporary discourse, Antarctica is increasingly conceptualised as an “open science commons” a globally shared epistemic infrastructure in which data, observations, and research platforms are collectively maintained for planetary benefit (Polar Data Catalogue, n.d.; United States National Science Foundation, n.d.). This framing emphasises the role of Antarctic datasets as global public goods, particularly in relation to climate system monitoring, biodiversity assessment, and cryosphere–ocean interactions (Rogers et al., 2019; Gutt et al., 2022). The continent’s observational records constitute one of the most critical empirical foundations for understanding Earth system variability and long-term environmental change.

At the same time, the commons metaphor highlights inherent governance tensions. While data are formally designated for open exchange, their production remains dependent on national funding structures, logistical capacities, and institutional priorities. This creates asymmetries in data generation, access, and analytical capacity, particularly between technologically advanced and resource-limited research systems (United States National Science Foundation, n.d.; Southern Ocean Observing System Data Management Sub-Committee, 2022). As a result, the Antarctic data ecosystem operates through a hybrid governance model combining legal openness, institutional coordination, and uneven material capability.

Ultimately, the Antarctic Treaty system demonstrates that open science is not merely a technical or ethical construct, but a deeply political achievement requiring continuous institutional maintenance. Its durability over more than six decades suggests that scientific cooperation can function as a stabilising force in international relations, even under conditions of broader geopolitical tension. The Antarctic case thus provides a rare empirical example of how shared epistemic infrastructures can generate long-term diplomatic resilience while simultaneously advancing global scientific understanding (Secretariat of the Antarctic Treaty, n.d.; Walton, 2011; Scientific Committee on Antarctic Research, n.d.).

5 Open Science is a Structural Necessity for Antarctic Research

Open science principles in Antarctic research should not be understood merely as an aspirational or normative framework; rather, they constitute a foundational infrastructural condition that enables the long-term functioning of the entire polar scientific system. The defining characteristic of Antarctic science is its temporal extensivity: observational programmes often span multiple decades, outlasting individual projects, institutional mandates, and even technological paradigms. Consequently, datasets

are routinely transferred across generations of researchers who were not involved in their initial production. This creates a structural dependency on transparency in documentation, methodological traceability, and sustained accessibility, without which observational continuity would be epistemically compromised (Clarke, 1993).

In this context, openness functions as a mechanism of intergenerational knowledge transfer. Unlike laboratory-based sciences, where experiments can be replicated under controlled conditions, Antarctic observations are often unique, non-repeatable, and environmentally contingent. This makes the preservation of full contextual metadata—instrumentation details, calibration procedures, spatial-temporal resolution, and processing workflows—essential for ensuring scientific validity over time. Without such transparency, the interpretability of long-term climate records would degrade significantly, undermining their analytical value for trend detection and model validation (Clarke, 1993; Jong et al., 2022).

Operational realities further reinforce the necessity of open science practices. Antarctic research stations rely on highly mobile scientific personnel, with frequent rotations and limited continuity of on-site expertise. As a result, institutional memory is inherently fragmented and cannot serve as the primary mechanism for knowledge retention. In such a system, openness is not merely a dissemination strategy but a substitute for physical continuity: it externalises scientific memory into shared infrastructures that persist beyond individual careers and organisational cycles (Clarke, 1993; Polar Data Catalogue, n.d.).

A second structural justification for open science arises from the heterogeneity of Antarctic research infrastructures. The continent is characterised by a distributed network of national programmes that differ significantly in technological capacity, observational design, calibration standards, and disciplinary priorities. This heterogeneity produces inherent risks of epistemic fragmentation, where datasets become difficult to integrate due to inconsistencies in metadata structures, sampling protocols, or terminological conventions. In this regard, open science operates as a coordinating mechanism that mitigates systemic incoherence by enforcing interoperability and encouraging standardisation across institutional boundaries (Graiff et al., 2023; Southern Ocean Observing System Data Management Sub-Committee, 2022).

Recent analyses of Antarctic data ecosystems highlight that the absence of harmonised metadata standards remains one of the principal barriers to large-scale data synthesis (Graiff et al., 2023). Initiatives coordinated through international frameworks, including SOOS data governance structures, explicitly address these challenges by promoting unified data policies and encouraging adoption of interoperable formats (Southern Ocean Observing System Data Management Sub-Committee, 2022). In this sense, open science functions not only as a principle of transparency but as an infrastructural condition for epistemic integration at the continental scale.

Environmental and logistical fragility further strengthens the argument for openness. Antarctic field operations are exposed to extreme climatic variability, infrastructural isolation, and high operational risk. Equipment failure, power disruption, logistical constraints, or station decommissioning can result in irreversible data loss if local storage systems are not redundantly backed up. Under these conditions, distributed digital repositories play a critical role as resilience mechanisms, ensuring that scientific records are preserved independently of local infrastructural stability (United States National Science Foundation, n.d.; Polar Data Catalogue, n.d.). Open access architectures therefore function as a form of epistemic redundancy, safeguarding data integrity against environmental and operational uncertainty.

Beyond preservation, Antarctic data possess substantial global relevance, extending their significance far beyond regional scientific inquiry. Observational outputs contribute directly to international climate assessments, Earth system modelling, and ocean–atmosphere coupling studies, as well as to biodiversity monitoring and ecosystem-based management frameworks in the Southern Ocean. These applications depend fundamentally on unrestricted data accessibility, as computational modelling systems require large-scale, standardised, and machine-readable datasets to generate

reliable projections. In this sense, restricted or fragmented data environments would not only reduce scientific efficiency but also undermine the epistemic robustness of global environmental governance systems (Rogers et al., 2019; Graiff et al., 2023).

Within this evolving landscape, the FAIR data principles (Findable, Accessible, Interoperable, Reusable) provide a critical conceptual and technical framework for sustaining long-term data usability. FAIR is particularly significant in the Antarctic context because it addresses the temporal dimension of scientific value: datasets must remain interpretable not only at the moment of their production but across multiple technological and institutional transitions. Persistent identifiers ensure stable referencing across evolving digital infrastructures, while rich metadata schemas preserve contextual integrity. Interoperability standards, in turn, enable integration across heterogeneous data systems, supporting cross-disciplinary synthesis and large-scale modelling efforts (SCAR Standing Committee on Antarctic Data Management et al., 2025; Southern Ocean Observing System Data Management Sub-Committee, 2022).

FAIR principles should be understood not merely as a technical specification but as a governance logic that extends the logic of open science into operational practice. They ensure that data retain epistemic coherence across time, space, and institutional boundaries, thereby transforming raw observations into durable scientific assets. In Antarctic research, where environmental monitoring spans decades and relies on globally distributed infrastructures, this long-term compatibility is essential for sustaining cumulative knowledge production (SCAR Standing Committee on Antarctic Data Management et al., 2025; Southern Ocean Observing System Data Management Sub-Committee, 2022).

Open science in Antarctica is best conceptualised as a systemic requirement rather than a policy preference. It emerges from the intersection of environmental extremity, institutional fragmentation, and the global dependence on polar data for climate system understanding. In this context, openness functions simultaneously as a preservation strategy, a coordination mechanism, and a form of epistemic infrastructure that enables Antarctic science to operate as a coherent, long-term, and globally integrated system (Clarke, 1993; Rogers et al., 2019; SCAR Standing Committee on Antarctic Data Management et al., 2025).

6 Conceptual Model for Implementing Standards and Identifiers at the National Antarctic Scientific Center

For the National Antarctic Scientific Center, the creation of a coherent system of persistent identifiers represents the first necessary step towards a modern, interoperable information infrastructure. Assigning DOIs to datasets, observational series, expedition reports, and other research outputs stabilises their identity in the scientific ecosystem and allows them to be cited unambiguously in future work. Using ORCID to identify individual researchers, technicians, and expedition participants ensures that contributions are accurately recorded and linked to personal scholarly profiles. The adoption of ROR identifiers for institutional affiliations introduces consistency into the organisational metadata and facilitates automated linking across external systems. Together, these identifiers establish a structured network of relationships among people, institutions, datasets, and publications.

To achieve deeper semantic and structural coherence, this network must be embedded within a standardised metadata architecture. The CERIF model, widely used in European research information systems, offers a flexible structure capable of capturing the full complexity of Antarctic research. Its entity-relationship framework enables the representation of not only datasets and publications but also instruments, research activities, field events, and logistical components. Through controlled

vocabularies, geographic descriptors, and typological classifications, CERIF ensures that metadata remain consistent across the entire system. This approach makes it possible to trace the entire research sequence from a specific expedition using a particular piece of equipment, to the observations collected, the datasets curated, and eventually the publications derived from them.

Such a system gains its full potential when connected to the wider ecosystem of international open infrastructures. Integration with OpenAIRE enhances the visibility and discoverability of NASC's outputs at the European and global level. Collaboration with DataCite ensures high-quality DOI registration and facilitates metadata alignment. The use of ORCID and ROR provides authoritative identity and affiliation records that interoperate with platforms used by publishers, repositories, and funding agencies. Through engagement with SCAR, COMNAP, and related disciplinary networks, NASC positions its data within domain-specific frameworks that consolidate Antarctic research internationally. Participation in international CRIS ecosystems strengthens interoperability further, enabling cross-institutional exchange of metadata and enhancing analytical capacities.

The convergence of these elements produces a range of benefits that extend far beyond administrative efficiency. The reduction of manual work frees researchers from repetitive metadata entry and allows greater emphasis on scientific tasks. Full traceability of data provenance and relationships supports better reproducibility, quality assessment, and long-term stewardship. Enhanced metadata structures facilitate analytical functions such as expedition planning, instrument tracking, and programme evaluation. Ultimately, by adopting internationally recognised standards and identifiers, NASC strengthens its scientific visibility and ensures that its research outputs participate fully in the global knowledge system.

7 International Models for Antarctic Data Governance: A Comparative Overview

International experience demonstrates that effective Antarctic data governance is not the product of isolated institutional effort, but rather the outcome of a multi-layered system combining coordinated national infrastructures, clearly defined open data mandates, and sustained participation in global disciplinary networks (Clarke, 1993; Graiff et al., 2023; Southern Ocean Observing System Data Management Sub-Committee, 2022). In this configuration, governance is distributed across national, institutional, and international levels, each contributing distinct but interdependent functions within a broader epistemic infrastructure.

At the national level, several countries have developed highly structured and mature Antarctic data centres that function as authoritative custodians of scientific outputs. Notable examples include the Australian Antarctic Data Centre (AADC), the US Antarctic Program Data Center (USAP-DC), and the UK Polar Data Centre (UK PDC), all of which serve as centralised nodes for the aggregation, curation, preservation, and dissemination of polar datasets (United States National Science Foundation, n.d.; Polar Data Catalogue, n.d.). These institutions operate not merely as repositories but as long-term stewardship organisations, ensuring that observational records remain accessible, interpretable, and technically usable across extended temporal horizons.

A key feature of these centres is their dual mandate: they manage both digital datasets and the associated scientific documentation produced through Antarctic field campaigns, while in some cases extending their scope to physical collections such as ice cores, geological samples, biological specimens, and instrument logs (SCAR Standing Committee on Antarctic Data Management et al., 2025). This integrated approach reflects a recognition that Antarctic science is fundamentally hybrid

in nature, requiring the co-preservation of both digital and material evidence to maintain epistemic completeness.

A defining characteristic of contemporary Antarctic data governance is the formalisation of Open Data policies embedded within national research systems. These policies articulate binding expectations regarding data lifecycle management, including acquisition, documentation, publication, and long-term archiving. Importantly, they are often linked directly to research funding mechanisms, meaning that compliance is structurally enforced through grant conditions rather than left to voluntary practice (Southern Ocean Observing System Data Management Sub-Committee, 2022). Within this framework, Data Management Plans (DMPs) serve as operational instruments that structure research workflows from project inception, ensuring alignment with international standards and facilitating downstream interoperability.

The increasing institutionalisation of Data Transfer Agreements (DTAs) further reinforces governance consistency. These agreements define the responsibilities of researchers regarding metadata completeness, data quality assurance, intellectual property conditions, and timing of public release. In doing so, they transform data submission from an informal practice into a regulated procedural pipeline that standardises the transition from field collection to archival preservation (United States National Science Foundation, n.d.; Polar Data Catalogue, n.d.). Empirical assessments of polar data infrastructures indicate that such formalisation has significantly reduced fragmentation and improved cross-project comparability, particularly in long-term environmental monitoring programmes (Clarke, 1993; Graiff et al., 2023).

At the technical level, persistent identifiers and standardised metadata frameworks constitute the backbone of interoperability across Antarctic data systems. DOI registration through DataCite ensures long-term citability and stable referencing of datasets, while ORCID and ROR systems provide persistent identification of researchers and institutions. Complementing these, SCAR-aligned metadata standards and ISO-based schemas establish semantic consistency across heterogeneous datasets, enabling integration across disciplinary and national boundaries (SCAR Standing Committee on Antarctic Data Management et al., 2025). This standardisation is particularly critical in Antarctic science, where data are generated under diverse methodological regimes but must ultimately be synthesised into coherent circumpolar analyses.

The cumulative effect of these infrastructures is the creation of a globally interoperable data ecosystem capable of supporting large-scale Earth system modelling and climate synthesis. By ensuring structural compatibility between datasets, these standards enable the integration of observations across temporal scales, spatial domains, and disciplinary frameworks, thereby increasing the scientific utility of national data holdings within global research infrastructures (Southern Ocean Observing System Data Management Sub-Committee, 2022; Graiff et al., 2023).

At the international level, governance coordination is facilitated by institutions such as the Scientific Committee on Antarctic Research (SCAR), the Council of Managers of National Antarctic Programs (COMNAP), and the Standing Committee on Antarctic Data Management (SCADM). Rather than acting as repositories, these bodies function as meta-governance platforms responsible for harmonising standards, promoting interoperability, and disseminating best practices across national systems (SCAR, n.d.). SCADM in particular plays a central role in consolidating technical guidance on metadata, data publication workflows, and digital preservation strategies, thereby ensuring convergence in system design across geographically and institutionally diverse programmes (SCAR Standing Committee on Antarctic Data Management et al., 2025).

This layered governance architecture illustrates a key structural feature of Antarctic data management: its reliance on distributed coordination rather than centralised control. The system functions effectively because it aligns national autonomy with shared international standards, producing a form of “soft integration” in which interoperability is achieved through voluntary

alignment rather than hierarchical enforcement (SCAR Standing Committee on Antarctic Data Management et al., 2025; Southern Ocean Observing System Data Management Sub-Committee, 2022).

An additional but often underemphasised dimension of Antarctic data governance is the stewardship of physical scientific collections. Material archives such as ice-core repositories, geological sample banks, and biological specimen collections constitute an essential complement to digital datasets. Facilities such as national polar rock repositories and ice-core libraries ensure the long-term preservation of irreplaceable physical evidence, which often serves as the only means of validating or recalibrating historical observational records (United States National Science Foundation, n.d.). Increasingly, these physical collections are integrated with digital persistent identifier systems, enabling bidirectional linkage between material samples, datasets, and published analyses.

This integration of physical and digital infrastructures reflects a broader epistemic shift toward hybrid data stewardship models, in which scientific knowledge is maintained through interconnected material-digital ecosystems rather than isolated repositories. Such systems enhance traceability, improve reproducibility, and strengthen the long-term resilience of scientific evidence in polar research contexts.

Taken together, international Antarctic data governance models demonstrate a high degree of institutional maturity and methodological sophistication. They illustrate that effective stewardship is not achieved through technology alone, but through the coordinated interaction of policy frameworks, institutional mandates, technical standards, and international governance structures (Clarke, 1993; Graiff et al., 2023; Southern Ocean Observing System Data Management Sub-Committee, 2022). Crucially, they also highlight that long-term data sustainability depends on continuous organisational commitment, stable funding environments, and the maintenance of interoperable infrastructures capable of adapting to evolving scientific and technological landscapes.

8 Practical Recommendations for NASC Based on International Experience

A comparative assessment of established Antarctic data infrastructures indicates that effective data governance is not achieved through isolated technical solutions, but through the coordinated implementation of policy, infrastructure, and standardisation mechanisms. For the NASC, this implies the need for a structured, multi-level approach rather than incremental or ad hoc improvements.

At the policy level, the absence of a formalised and enforceable Data Policy remains a critical gap. International practice shows that such policies are most effective when they explicitly define obligations regarding data submission timelines, metadata completeness, and conditions of reuse. In this context, the introduction of mandatory Data Management Plans (DMPs) at the project approval stage would serve not merely as a formal requirement but as an operational instrument shaping data workflows from the outset. Empirical evidence from other national programmes suggests that early-stage planning significantly reduces data loss, inconsistencies in documentation, and delays in publication.

At the infrastructural level, the central challenge lies in the fragmentation of data storage and the lack of a unified access point. The establishment of a dedicated repository either as a national platform or as a federated node within an existing international infrastructure would address this issue by consolidating datasets, standardising metadata, and enabling persistent access. However, the effectiveness of such a repository depends on clearly defined ingestion workflows. This includes the

specification of mandatory metadata schemas (e.g., ISO-compliant or domain-specific standards), version control procedures, and quality assurance protocols. The introduction of Data Transfer Agreements (DTAs) at the point of submission is particularly important, as it formalises responsibilities for data integrity and ensures that deposited datasets meet minimum interoperability requirements.

The implementation of persistent identifiers should be understood not as a technical add-on but as a core structural component of the data ecosystem. The combined use of DOIs for datasets, ORCID identifiers for individual contributors, and ROR identifiers for institutional affiliations enables the construction of a relational metadata environment in which datasets, publications, projects, and personnel are systematically interconnected. This significantly enhances traceability and supports automated data exchange across platforms, including integration with CRIS systems and international repositories.

A further area requiring explicit regulation is data accessibility. International experience demonstrates that ambiguity in access conditions often leads to either excessive restriction or uncontrolled data release. To avoid this, NASC should define a differentiated access model that includes clearly specified embargo periods, conditions for restricted datasets (e.g., sensitive environmental or geopolitical data), and standardised procedures for data requests. Such a framework not only facilitates reuse but also reduces transaction costs in international collaboration by providing predictable and transparent access rules.

Equally important is the alignment with international coordination mechanisms. Participation in initiatives led by SCAR, COMNAP, and related data management bodies should not be limited to formal membership but should involve active adoption of recommended vocabularies, metadata profiles, and interoperability protocols. Comparative analysis shows that institutions that closely follow these frameworks achieve significantly higher levels of data integration and visibility within the global research landscape.

From a strategic perspective, the transition to standardised and interoperable data practices has direct implications for the positioning of NASC within the international Antarctic research system. In an environment where data infrastructures are increasingly interconnected, the ability to provide well-documented, machine-readable, and persistently identifiable datasets becomes a key determinant of institutional relevance. This is particularly important for integration into European and global research infrastructures, where compliance with FAIR principles, structured metadata, and identifier systems is a prerequisite for participation.

The long-term dimension of Antarctic research further amplifies the importance of these measures. Observational datasets extending over several decades require governance mechanisms that ensure not only their preservation but also their interpretability over time. Without standardised metadata, persistent identifiers, and controlled vocabularies, the contextual information necessary for reuse is gradually lost, reducing the scientific value of the data. Conversely, well-structured data systems enable continuity across generations of researchers and support cumulative knowledge production.

Finally, the adoption of open and interoperable standards has a demonstrable effect on international collaboration. Data that are easily discoverable, properly documented, and reliably linked to their creators are more likely to be reused in comparative studies, incorporated into global models, and cited in high-impact publications. This, in turn, increases the visibility of national research programmes and facilitates their inclusion in multinational projects.

In this sense, the implementation of standards and identifiers should be seen not only as a technical or administrative reform, but as a strategic investment in the integration of Ukrainian Antarctic science into the global knowledge system.

9 Conclusions

Antarctic research operates within a uniquely demanding scientific environment defined by long temporal horizons, distributed infrastructures, and strong dependence on international collaboration. The analysis presented in this study demonstrates that under such conditions, openness, interoperability, and long-term data stewardship are not supplementary principles but foundational requirements for the functioning of the entire research system. Multi-decadal observational datasets central to understanding climate dynamics, atmospheric processes, and ecosystem change retain their value only when supported by robust data governance frameworks. Without structured metadata, persistent identifiers, and standardized documentation practices, these datasets risk becoming fragmented, uninterpretable, or effectively lost for future use.

The findings further indicate that the transition to FAIR-aligned data practices represents a critical shift from locally managed data collections to globally integrated knowledge infrastructures. Standards, identifiers, and interoperable formats enable the transformation of isolated datasets into components of a larger, machine-readable ecosystem in which data, publications, researchers, and institutions are systematically linked. This transition is essential for ensuring reproducibility, facilitating large-scale comparative analyses, and enabling the integration of Antarctic data into global climate and environmental models.

For the NASC, the adoption of such an approach has both scientific and strategic implications. At the operational level, the implementation of standardised metadata frameworks, persistent identifier systems (DOI, ORCID, ROR), and repository-based data management will significantly improve data traceability, reduce fragmentation, and support long-term preservation. At the institutional level, alignment with international standards and infrastructures will enhance the visibility, credibility, and usability of Ukrainian Antarctic research outputs within the global scientific community.

Importantly, the study highlights that integration into international data ecosystems is no longer optional. Participation in contemporary research infrastructures particularly within European and global contexts requires compliance with established norms of data management, including the use of structured metadata, open access policies, and formalised data management planning. Institutions that fail to meet these requirements risk marginalisation, regardless of the scientific quality of their data.

At the same time, the adoption of open science practices creates tangible opportunities. Interoperable and well-documented datasets are more readily reused in international projects, incorporated into interdisciplinary research, and cited in high-impact publications. This increases not only the scientific visibility of NASC but also its capacity to engage in multinational initiatives and contribute to global environmental governance.

Finally, the long-term perspective is of particular importance. Antarctic datasets are inherently intergenerational assets, and their preservation requires governance mechanisms that extend beyond individual projects, funding cycles, or institutional changes. By embedding open science principles into its data management practices, NASC can ensure the continuity, reliability, and cumulative value of its scientific outputs over decades. In doing so, it strengthens Ukraine's position as an active and credible participant in the global Antarctic research system and contributes to the development of a shared, resilient, and interoperable scientific knowledge base.

References

1. Abram, N. J., Gillett, N. P., England, M. H., Sigman, D. M., & LeGrande, A. N. (2013). A review of sea ice proxy information from polar ice cores. *Quaternary Science Reviews*, 79, 55–74. <https://doi.org/10.1016/j.quascirev.2013.01.011>
2. Antarctic and Southern Ocean Coalition. (n.d.). Antarctic governance. Retrieved from <https://www.asoc.org/learn/antarctic-governance/>
3. Clarke, D. (1993). Antarctic data management. *Antarctic Science*, 5(3), 237. <https://doi.org/10.1017/S0954102093000318>
4. Graiff, A., Braun, M., Driemel, A., Ebbing, J., Grossart, H. P., Harder, T., et al. (2023). Big data in Antarctic sciences: current status, gaps, and future perspectives. *Polarforschung*, 91, 45–57. <https://doi.org/10.5194/polp-91-45-2023>
5. Gutt, J., Barnes, D. K. A., Brewin, P., & Griffiths, H. (2022). A framework to observe, understand, and project long-term changes in Antarctic marine ecosystems: the case for an Integrated East Antarctic Marine Research observatory. *Biogeosciences*. Preprint. Retrieved from <https://bg.copernicus.org/preprints/bg-2022-110/>
6. Hughes, K. A., Convey, P., & Turner, J. (2021). Developing resilience to climate change impacts in Antarctica: an evaluation of Antarctic Treaty System protected area policy. *Environmental Science & Policy*, 124, 12–22. <https://doi.org/10.1016/j.envsci.2021.05.023>
7. Jong, L. M., Etheridge, D. M., van Ommen, T. D., Curran, M. A. J., & Legrand, M. R. (2022). 2000 years of annual ice core data from Law Dome, East Antarctica. *Earth System Science Data*, 14, 3313–3340. <https://doi.org/10.5194/essd-14-3313-2022>
8. Lorius, C. (1992). The ice core record: past archive of climate and environmental parameters. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 98(1–2), 39–49.
9. Polar Data Catalogue. (n.d.). International polar data repositories. Retrieved from <https://www.polardata.ca>
10. Rogers, A. D., Yesson, C., Gravestock, P., & Sumaila, U. R. (2019). Antarctic futures: an assessment of climate-driven changes in Antarctic marine ecosystems. *Annual Review of Marine Science*, 11, 1–24.
11. SCAR Standing Committee on Antarctic Data Management (SCADM) et al. (2025). A brief guide to publishing data for the polar research community. Retrieved from <https://zenodo.org/records/15846462>
12. Scientific Committee on Antarctic Research. (n.d.). SCAR overview. Retrieved from <https://scar.org/>
13. Secretariat of the Antarctic Treaty. (n.d.). The Antarctic Treaty. Retrieved from <https://www.ats.aq>
14. Southern Ocean Observing System Data Management Sub-Committee. (2022). SOOS Data Policy 02/2022. <https://doi.org/10.5281/zenodo.6041433>
15. Srivastava, R., Sengupta, D., & Ghosh, P. (2023). Apparent ice accumulation rate in East Antarctica: relation with temperature and thinning pattern. *arXiv*. Retrieved from <https://arxiv.org/abs/2309.14373>
16. Stauffer, B. (2000). Long-term climate records from polar ice. *Space Science Reviews*, 94, 321–336. <https://doi.org/10.1023/A:1026791811601>
17. United States National Science Foundation, Office of Polar Programs. (n.d.). Data repositories and resources. Retrieved from <https://www.nsf.gov/geo/opp/data-repositories>
18. Walton, D. W. H. (2011). The Scientific Committee on Antarctic Research and the Antarctic Treaty. In P. A. Berkman, M. A. Lang, & O. R. Young (Eds.), *Science diplomacy: Antarctica, science, and the governance of international spaces* (pp. 75–88). Smithsonian Institution.