



Australia's National  
Science Agency

# International Strategic Advisor Report to ROIS

Dr Jens Klump

January 2026

CSIRO Mineral Resources Research Unit

#### Citation

Klump J (2026) International Strategic Advisor Report to ROIS. CSIRO, Australia.

#### Copyright

© Commonwealth Scientific and Industrial Research Organisation 2026. To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

#### Important disclaimer

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

CSIRO is committed to providing web accessible content wherever possible. If you are having difficulties accessing this document, please contact [csiro.au/contact](https://csiro.au/contact).

#### Acknowledgement of Country

CSIRO acknowledges the Traditional Owners of the lands, seas and waters of the area that we live and work on across Australia and pays its respects to Elders past and present. CSIRO recognises that Aboriginal and Torres Strait Islander peoples have made, and will continue to make, extraordinary contributions to Australian life, including in cultural, economic, and scientific domains.

# Introduction

As part of their International Advisory Programme, the Research Organisation of Information and Systems (ROIS) invited Dr Jens Klump to visit the Joint Support Centre for Data Science Research (ROIS-DS) during the period of 1-12 December 2025. Dr Klump is a Senior Principal Research Scientist and Leader of the Exploration Through Cover Research Group at the Commonwealth Scientific and Industrial Research Organisation (CSIRO), Director of the AuScope Virtual Research Environment, and President of the International Generic Sample Number Organization (IGSN e.V.).

The purpose of the visit was to gain an overview of the activities at ROIS-DS from an external perspective, and to provide the centre with recommendations and suggest benchmark institutions, to help guide the future development of the centre. This report focuses on the Polar Environment Data Center (PEDSC), but will also cover some of the other systems associated with ROIS-DS.

## Visit Schedule and Events

The visit was organised in two parts, starting with the 16<sup>th</sup> Symposium on Polar Science in Week 1, followed by a ROIS-DS Special Seminar and several individual meetings in Week 2. The symposium, and in particular, the Special Session of the symposium, helped put the developments in PEDSC into a broader perspective of NIPR's science strategy.

### Schedule Week 1

- Introduction to ROIS-DS
- 16<sup>th</sup> Symposium on Polar Science (Special Session [S], Polar Data Science [ID], Polar Geosciences [OG], Antarctic Meteorites [OA])

### Schedule Week 2

- Overview of ROIS-DS systems
- Visit to SHRIMP lab, computer server room, meteorite collection, ice core sample storage
- ROIS-DS Special Seminar
- Discussion of an AI-enhanced access to polar data
- Discussion of social science data
- Discussion on the interpretation of seismological data from the Orkney Mine, Witwatersrand, South Africa



Figure 1 Overview of data systems associated with PEDSC.

## Current Status of ROIS-DS and PEDSC

ROIS-DS, as well as PEDSC, have grown over the years, based on the needs of scientific projects and research programmes, and are currently engaged in (Kadokura, 2025):

- Building an integrated database of polar science research
- Upgrading existing database systems in PEDSC and enhancing their interoperability
- Archiving and sharing of Antarctic time series data
- Archiving and sharing of physical samples data
- Data publication through the NIPR Polar Data Journal
- Training of Data Scientists
- Collaboration programs

With these ongoing activities, the accompanying data systems have evolved with time into 41 separate systems related to PEDSC.

The data systems are developed and operated by separate disciplinary groups. Some of the infrastructure components that are served by these systems are nearing the end of their deployment, e.g., the PANSY radar system. The replacement of systems and the deployment of new infrastructure will require an update of the supporting IT infrastructure. At present, the computing centre is based on a heterogeneous technology stack, with legacy components from existing projects. ROIS-DS is taking first steps in the direction of moving services to the cloud.

In 2024, ROIS-DS released the *Advanced Multidisciplinary Integrated Database for Exploring new Research* (AMIDER) as a single point of entry to the PEDSC system. To assist navigation, Dr Zhang Qi is exploring the use of knowledge graphs to help researchers find data across domains that are held in different sub-systems of PEDSC.

NIPR also holds large volumes of legacy data in its geophysical data archive. Some of them are analogue materials like aerial photography and plots of geophysical data. Some of the materials are in the process of

digitisation. The current interest in digitalisation of aerial photography for geomorphology research. NIPR has also contacted the electronics manufacturer NEC for tape readers to be able to transfer geophysical data stored on magnetic media to modern storage media. Since the holdings go back many decades, there is a great opportunity to make data from the pre-satellite era accessible for reuse.

The rock sample collection has an online catalogue that makes the collection and its holdings findable. The system seems to have placeholders for IGSN IDs, but no IDs have been registered. The introduction of IGSN IDs would make it easier for NIPR to connect samples to data and track how materials are used and cited in the literature.

The meteorite collection is a world-class research collection with an online catalogue. Similar to the rock collection, it might be investigated whether it is possible to introduce IGSN IDs.

In both cases, the rock collection and the meteorite collection, there would be no need to introduce new IDs to mint IGSN IDs; the existing IDs could be reused.

The ice core storage facility is a world-class facility. Given its specific workflows and sampling protocols, the benefits of cross-linking would lie in linking sample materials to data and publications.

Polar research at NIPR is expanding from its current research stations to broader regions related to the core study regions, e.g., understanding the evolution of the Kaapvaal Craton as an analogue to Proterozoic terranes in Antarctica.

Outside of PEDSC, the facility for viewing social science data matches international requirements for similar facilities.

## Recommendations

The aim of the data systems hosted as parts of the PEDSC is to:

- Provide access to fundamental data of the Arctic and Antarctic region
- Provide access to research data from projects associated with NIPR
- Support priority research projects through data management and data provision

NIPR is preparing the Japanese contributions to IPY5 2032-2033, and Phase XI of the Antarctic research programme will serve as the ramp-up to IPY5. This large, internationally connected research programme will bring new requirements to the research data infrastructures at NIPR, including PEDSC. In addition, there will be significant changes to the Japanese Antarctic research infrastructure. Icebreaker Shirase will be retired in 2033, and a follow-on is unclear. Other systems will reach the end of life too and will need to be replaced, e.g., PANSY. These changes will require adaptations of the supporting data systems.

In addition to established research topics, new topics are being added. Understanding the processes and effects of basal melting on processes associated with the Antarctic margin will require more disciplinary crossover of data in the future. Some research will even be conducted outside the polar regions, like the study of the Kaapvaal Craton as an analogue to once

neighbouring Proterozoic terranes in Antarctica. These developments will, again, require changes to the supporting data systems.

Because of the organic growth of the systems that constitute PEDSC today, the system has, most likely, accumulated significant technical debt. To be able to meet the data requirements of IPY5, the system should be modernised and the hardware upgraded.

A primary aim should be to follow the FAIR Data Guidance Criteria (Devaraju et al, 2021; Wilkinson et al, 2016) to enhance the findability, accessibility, interoperability, and reuse of the data. The current work on semantic interoperability and AI-supported tools for exploring the data is already moving in the right direction, and will be necessary to support cross-disciplinary work in future programmes.

Data collection and curation are often expensive and labour-intensive, and should be streamlined and automated as much as possible (Koppe et al, 2015). This makes the collected data valuable for future use. In addition to WDS-CTS criteria (CoreTrustSeal Standards and Certification Board, 2022), ROIS-DS should also investigate the resilience of its systems against disruption (Gum et al, 2025).

Over the coming years, we can expect a rapid development of AI-tools that can be used to support data curation and documentation, as well as data exploration and analysis. Here, ROIS-DS should conduct its own research, as it already does, and collaborate with other institutions in Arctic and Antarctic research (see Benchmark Institutions).

Until now, the Japanese Polar research programme has only worked in unpopulated regions. This might change in the future, in particular in the context of IPY5. ROIS-DS already has significant experience in dealing with human subject data. These two strands, PEDSC and the Centre for Social Data Structuring, could work together once human subject data becomes relevant. In addition, it would be advisable to seek input from organisations like Polar Knowledge Canada on how to manage data and data collection when Indigenous people are involved (Jennings et al, 2025).

The engagement of ROIS-DS with the World Data System (WDS) is one of the great success stories. This effort has been able to build a remarkable network of science collaboration across Asia and Oceania. ROIS-DS should continue and grow its ongoing engagement with WDS. In addition, ROIS-DS might consider engaging with the activities of the Polar Data Forum for alignment with international data practices (Janssen et al, 2023).

The organisational structure of PEDSC follows the disciplinary domains served by the system. This makes sense, as the strength of disciplinary data repositories lies in the support they can offer because of the disciplinary knowledge of their operators. To reduce the current duplication of technical systems, ROIS-DS should consider creating a digital technologies team to support the technical operation of the systems. This would also make it easier to adopt solutions developed in other groups, e.g., the Center for Data Assimilation Research and Applications, or elsewhere in Japan and abroad.

# Benchmark Institutions

With the current perspective of working towards IPY5, ROIS-DS and PEDSC should seek to build connections with other polar research data centres to enhance data exchange and interoperability, and for the alignment of best practices among polar research data centres.

## AWI

Over the past years, the Alfred Wegener Institute, Helmholtz Center for Polar and Marine Research (AWI) has developed data management practices that are well integrated with their research activities. O2A is a generic and sustainable framework for enabling the seamless flow of sensor observations to archives (Koppe et al, 2015). This framework builds on international OGC standards for metadata and data interoperability and is designed to assist scientists in developing enhanced data products and facilitating data reuse. AWI works closely with PANGAEA for data dissemination.

## PANGAEA

The information system PANGAEA provides targeted support for research data management as well as long-term data archiving and publication (Felden et al, 2023). PANGAEA is operated as an open-access library for archiving, publishing, and distributing georeferenced data from earth and environmental sciences. It focuses on observational and experimental data. PANGAEA is a pioneer of FAIR and open data infrastructures to enable data-intensive science and an integral component of national and international science and technology activities. In addition, PANGAEA is a leader in technical innovation for research data repositories.

## Australian Antarctic Data Centre

The Australian Antarctic Data Centre (AADC) was established in 1996 to provide long-term management of Australia's Antarctic data. The AADC team blends data management and scientific domain expertise with skills in computing, application development, geographic information systems, modelling, and statistical analyses. Data held in the AADC are qualified with metadata and discoverable through the Catalogue of Australian Antarctic and Subantarctic Metadata (CAASM). Some data are also delivered through customised applications on the AADC website. Data delivery systems, where appropriate, use international standards for data exchange and data discovery. AADC staff assist in the capture of data at sea and on land, organise satellite tasking to support science campaigns, and provide advice on data collection methods.

## British Antarctic Survey

The UK Polar Data Centre (UK PDC) is the focal point for Arctic and Antarctic environmental data management in the UK. It is part of the Natural Environment Research Council's (NERC) Environmental Data Service and based at the British Antarctic Survey. The UK PDC coordinates the preservation and management of polar data from UK-funded research and supports researchers in

complying with national and international data legislation and policy. As a data centre based primarily on geographical extent rather than scientific discipline, the UK PDC works closely with data scientists at the other NERC data centres and cooperates within the organisation on many of its own data management activities.

## Polar Data Catalogue

The Polar Data Catalogue is a CoreTrustSeal Certified Repository of metadata and data that describes and provides access to diverse data sets generated by Arctic and Antarctic researchers. The Polar Data Catalogue is a member of the Canadian Polar Data Consortium (CPDC), formerly known as the Canadian Consortium for Arctic Data Interoperability (CCADI). The Polar Data Catalogue provides a variety of online tools, available to the public and researchers alike, that facilitate the contribution and discovery of essential Arctic and Antarctic research.

## References

- CoreTrustSeal Standards and Certification Board** 2022 CoreTrustSeal Requirements 2023–2025. CoreTrustSeal. Available at <https://doi.org/10.5281/zenodo.7051012> [Last accessed 8 August 2024].
- Devaraju, A, Mokrane, M, Cepinskas, L, Huber, R, Herterich, P, Vries, J de, Akerman, V, L'Hours, H, Davidson, J, and Diepenbroek, M** 2021 From Conceptualization to Implementation: FAIR Assessment of Research Data Objects. *Data Science Journal*, 20(1): 4. DOI: <https://doi.org/10.5334/dsj-2021-004>
- Felden, J, Möller, L, Schindler, U, Huber, R, Schumacher, S, Koppe, R, Diepenbroek, M, and Glöckner, F O** 2023 PANGAEA - Data Publisher for Earth & Environmental Science. *Scientific Data*, 10(1): 347. DOI: <https://doi.org/10.1038/s41597-023-02269-x>
- Gum, J, Duerr, R, Blake, R E, Garretson, A, Wyborn, L, Choi, J, Nurnberger, A, Jenkyns, R, Koren, G, and Cluster, E S D M** 2025 Repository Crisis Scorecards. Earth Science Information Partners. Available at <https://doi.org/10.5281/zenodo.15122045> [Last accessed 26 November 2025].
- Janssen, A R, Bricher, P, Payne, K, Badhe, R, Biebow, N, Bruin, T de, Duerr, R, Elshout, P, Gaylord, A, Godøy, Ø, Gorringer, P, Guihen, D, Kool, J, Larsen, J R, Nolan, J, Novellino, A, Manley, W, Marouchos, A, McCammon, M, Murray, M, Parrott, J, Pearlman, J, Peat, H, Pulsifer, P, Symons, L, Tacoma, M, Theocharides, S, Tronstad, S, Verhey, C, and Putte, A V de** 2023 Polar Data Forum IV - An Ocean of Opportunities. *Data Science Journal*, 22(18): 1–13. DOI: <https://doi.org/10.5334/dsj-2023-018>
- Jennings, L, Jones, K, Taitingfong, R, Martinez, A, David-Chavez, D, Alegado, R 'Anolani, Tofighi-Niaki, A, Maldonado, J, Thomas, B, Dye, D, Weber, J, Spellman, K V, Ketchum, S, Duerr, R, Johnson, N, Balch, J, and Carroll, S R** 2025 Governance of Indigenous data in open earth systems science. *Nature Communications*, 16(1): 572. DOI: <https://doi.org/10.1038/s41467-024-53480-2>
- Kadokura, A** 2025 *Activities of the Polar Environment Data Science Center of ROIS-DS*.
- Koppe, R, Gerchow, P, Macario, A, Haas, A, Schäfer-Neth, C, and Pfeiffenberger, H** 2015 O2A: A generic framework for enabling the flow of sensor observations to archives and publications. In: *OCEANS 2015 - Genova*. Presented at the OCEANS 2015 - Genova. Genoa, Italy: IEEE. pp. 1–6. DOI: <https://doi.org/10.1109/OCEANS-Genova.2015.7271657>
- Wilkinson, M D, Dumontier, M, Packer, A L, Gray, A J G, Mons, A, Gonzalez-Beltran, A, Waagmeester, A, Baak, A, Brookes, A J, Evelo, C T, Mons, B, Persson, B, Goble, C, Schultes, E,**

**Mulligen, E van, Aalbersberg, Ij J, Appleton, G, Boiten, J-W, Dillo, I, Grethe, J S, Heringa, J, Strawn, G, Velterop, J, Bouwman, J, Lei, J van der, Kok, J, Zhao, J, Wolstencroft, K, Santos, L B da S, Roos, M, Thompson, M, Martone, M E, Crosas, M, Swertz, M A, Axton, M, Blomberg, N, Dumon, O, Groth, P, 't Hoen, P A C, Wittenburg, P, Bourne, P E, Rocca-Serra, P, Schaik, R van, Finkers, R, Hooft, R, Kok, R, Edmunds, S, Lusher, S J, Sansone, S-A, Slater, T, Sengstag, T, Clark, T, and Kuhn, T** 2016 The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3: 160018. DOI: <https://doi.org/10.1038/sdata.2016.18>

**As Australia's national science agency, CSIRO is solving the greatest challenges through innovative science and technology.**

CSIRO. Creating a better future for everyone.

**Contact us**

1300 363 400  
+61 3 9545 2176  
[csiro.au/contact](http://csiro.au/contact)  
[csiro.au](http://csiro.au)

**For further information**

**Mineral Resources**

Dr Jens Klump  
+61 8 6436 8828  
[Jens.klump@csiro.au](mailto:Jens.klump@csiro.au)  
[csiro.au/en/work-with-us/industries/mining-resources/Exploration](http://csiro.au/en/work-with-us/industries/mining-resources/Exploration)