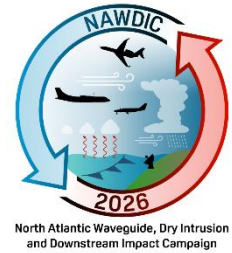# NAWDIC Data Policy

Version 1.1 (24 October 2025)

Authors (in alphabetical order):

Bastian Kirsch (KIT), Katharina Loewe (KIT), Julian Quinting (University of Cologne), Corinna Rebmann (KIT)

1. Introduction

The "North Atlantic Waveguide, Dry Intrusion and Downstream Impact Campaign" (NAWDIC[1]) is a WWRP (World Weather Research Programme) endorsed project which aims to improve the understanding and modelling of mesoscale processes along the dry intrusion (DI) air stream leading to high-impact weather (HIW) in the North Atlantic region in winter. The core observation period of NAWDIC is scheduled for January and February 2026 and involves research aircraft and ground-based measurements from several European institutions. Overall, NAWDIC serves as an umbrella for an international measurement effort with strong links to AR Recon[2] and NURTURE[3] led by partners from the United States.

This Data Policy acts as a written agreement between the NAWDIC projects (see Appendix A for the terminology used in this document) to commit to an open and respectful use of data, which we consider good scientific practice. It summarizes all provisions regarding the storage, exchange, and publication of NAWDIC data. This Data Policy has been developed in line with the World Meteorological Organization (WMO) guidelines[4] which state that the exchange of data products should be free, timely, and unrestricted, and follow the FAIR principles[5] (data sets should be findable, accessible, interoperable and reusable). More specifically, this Data Policy aims to:

- Ensure the FAIR use of the data
- Encourage an orderly and timely analysis, provision and publication of the data
- Uphold the rights of all involved researchers and have them treated equitably
- Ensure the visibility and the integrity of the project
- Define a central, open access point of the data to be published
- Encourage the rapid dissemination of the scientific results

---

[1] https://www.nawdic.kit.edu/
[2] https://cw3e.ucsd.edu/arrecon_overview/
[3] https://espo.nasa.gov/nurture
[4] https://community.wmo.int/resolution-40
[5] https://www.nature.com/articles/sdata201618

2. Data categories

Data collected in NAWDIC originate in a variety of different sources and therefore requires appropriate approaches in data archiving to ensure traceability and reproducibility of scientific results. Broadly the data can be divided into three different categories:

- *Primary Data* are raw observational data without further post-processing or quality assurance procedures applied to them. Depending on the instrument type, primary data may be raw measurement signals or (meteorological) variables derived from them.
- *Secondary Data* are derived from primary data through post-processing and quality assurance procedures.
- *Model Data* are the data required to conduct numerical simulations and result from model runs that may involve primary or secondary data as input.

3. Data management

3.1 Near real-time data exchange
During the campaign, operational weather forecast products need to be accessible for the NAWDIC investigators for campaign decisions. In addition, NAWDIC investigators need to provide information about the instrument status as well as quick looks for the planning of further instrument deployments. Sharing of these data within the NAWDIC community during the observation campaign can be limited to images (e.g., quick looks). These products can be made accessible to the NAWDIC community via MS Teams. Additionally, an SFTP server hosted by KIT will be provided for near real-time exchange of data during the campaign. Depending on the data size, data may only be stored on the SFTP server for a limited time.

3.2 Data exchange for data assimilation purposes
NAWDIC investigators interested in data assimilation are responsible for finding a solution for transfer and formatting of observational data according to their needs in coordination with the responsible data provider. A central solution will not be provided.

3.3 Long-term data storage and access
Secondary data and model data need to be stored long-term and made accessible to the NAWDIC investigators and the wider scientific communities via approved repositories. Version control is necessary to allow the publication of updated datasets as new calibrations or data retrievals become available, or potential bugs are discovered. DOIs need to be assigned to published datasets according to requirements of funding agencies and scientific journals.

The **Earth Data Portal**[6] (EDP) will act as the central access point to the repositories that store data sets generated by the NAWDIC investigators. The EDP eases the findability and accessibility of the individual data sets by demanding standardized keywords, metadata, and data formats. To be integrated into the EDP, data repositories need to fulfil one of the following criteria:

---

[6] https://earth-data.de/

- The repository is already integrated in EDP, meaning that datasets published at the respective repository are already harvested by and accessible from the EDP (Appendix B).
- The repository is not yet integrated in EDP but provides an OAIPMH[7] (Open Archives Initiative Protocol for Metadata Harvesting) standard for metadata.

In addition, predefined metadata need to be used to assign the respective dataset to NAWDIC on the EDP (Appendix C).

If NAWDIC investigators must use institutional repositories or follow specific requirements provided by the respective funding agencies, the use of curated data repositories is strongly encouraged to increase the quality, consistency, and visibility of NAWDIC datasets. A list of recommended data repositories is given in Appendix B.

3.4 Data provision and publication
Both, to respect the efforts of NAWDIC investigators in collecting, processing, and providing data and to ensure a rapid dissemination of data, the process towards data publication is divided into certain time periods following definitions of the HALO Data Policy [8].

3.4.1 Upload period for data provision
An upload period exists for **six months** after the end of the campaign. By the end of the upload period, data providers should have **provided** their preliminary data sets with interested members of the NAWDIC community. The data should contain qualifiers of the status of the analysis and corresponding documentation. Corrections and amendments to preliminary data must be communicated to all NAWDIC investigators as soon as possible. The KIT-hosted SFTP server (see Section 3.1) may be used for sharing the data with the NAWDIC community.

3.4.2 Quality assurance period for data publication
A quality assurance period exists for **24 months** after the end of the campaign. Data providers are required to **publish** a first version of their quality-controlled and/or processed data product no later than the end of the quality assurance period, ensuring that the selected data repository is compatible with the EDP and that all guidelines regarding keywords, etc. are followed to ensure that the dataset is included as a NAWDIC dataset in the EDP. The processing and applied quality control must be specified in detail in the accompanying metadata. Data providers are encouraged to choose a data repository that assigns a DOI to the dataset to facilitate referencing, even if not explicitly required by their funding agency. Data providers are further responsible for their own storage of raw, unprocessed data. Furthermore, data providers may set embargo periods for specific datasets to safeguard ongoing PhD projects; however, these periods must not extend beyond the submission date of the doctoral thesis.

3.4.3 Publication of model data
All model data required to run simulations must be published at the time of publishing corresponding papers. This includes model namelists, information regarding model version, model code in the case of code modifications, and input data or links to input data if they are publicly available (e.g., ERA5). To keep data storage requirements manageable, model output

---

[7] https://www.openarchives.org/pmh/
[8] https://halo-spp.de/Data/

does not need to be published but is encouraged for simulations that are likely of further use to other NAWDIC investigators.

3.4.4 Data acknowledgements

Data providers should always provide an acknowledgement following the example in Section 3.7 that can be used to cite the respective dataset.

3.5 Data formats

All datasets should be published in **netCDF[9] format**, unless there are justified exceptions (e.g., model code or namelists). While not strictly enforced, data providers are encouraged to use the Climate and Forecast[10] (CF) netCDF standard, the standard recommended by Unidata, who developed and maintains the netCDF libraries. Adhering to a common standard will increase the usability of datasets across the NAWDIC community, thus increasing the visibility of datasets through their use in multiple publications. The CF standard contains a list of recommended attributes to describe the file content, including, for example, variable units, missing value information, and information about the data provider and instrument. Attribute and variable names should follow the metadata guidelines for environmental data[11] developed at the Institute für Meteorology und Climate Research of KIT. A list of standard attributes for NAWDIC projects is provided in Appendix C.

3.6 Data licensing

All data must be provided under a **Creative Commons (CC) license**[12]. The data provider can select the license from the available CC license options.

3.7 Data use

Before using a NAWDIC data set (independent of its processing stage), data users are strongly encouraged to contact the respective instrument PI or data provider to discuss potential limitations and pitfalls of the data set as well as restrictions of its usage (see contact information in Appendix D). Any use of NAWDIC data products must include an acknowledgement (i.e., citation) of the source and the data provider. Funding information on NAWDIC projects can be found in Appendix D. The following acknowledgement is recommended:

> *The [description of the data product] was collected/produced in the [name of project] project [name of funding agency, information about grant] under the umbrella of NAWDIC and provided by [name of the data provider, institution of data provider]. [name of the data provider]'s contribution to NAWDIC was funded by [name of the funding agency, information about grant]. The data are archived at [name of data repository] and are accessible at [URL/DOI of data product].*

Data users must inform the respective data providers if a data product is to be shared with other parties via journal articles, presentations, and research proposals. If the data product constitutes a substantial part of the work (scientific article, conference contribution, or public meeting), the data provider should be offered co-authorship and the opportunity to collaborate

---

[9] https://www.unidata.ucar.edu/software/netcdf/
[10] http://cfconventions.org/Data/cf-conventions/cf-conventions-1.7/cf-conventions.html
[11] https://hgf-ee-netcdf-conventions-draft.readthedocs.io/en/latest/
[12] https://creativecommons.org/share-your-work/cclicenses/

(both during and after the quality assurance period). In the case of co-authorship, an additional acknowledgement of the data provider is not required. Commercial use of NAWDIC data products is prohibited, unless authorized by the data provider.

No further public release of a NAWDIC data product (sharing with colleagues, conference presentations, publications, commercial and media use, etc.) is allowed without permission of the data provider.

4. Commencement of Data Policy
This Data Policy will take effect on 1 December 2025. Any objections or requests for modification must be communicated to the NAWDIC data management team no later than 15 November 2025. In the absence of such communication, the Data Policy shall be considered accepted by all NAWDIC investigators.

Appendix A – Glossary

- *Data repository:* any storage infrastructure that is used to host data sets, to ensure long-term data availability, promote scientific reproducibility, and typically provide persistent identifiers like DOIs for proper citation
- *NAWDIC project:* a research initiative that has been endorsed by NAWDIC and addresses one or more of the goals of NAWDIC. There are two types of NAWDIC projects: (1) NAWDIC core projects (NAWDIC-HALO, NAWDIC-DICHOTOMI, NAWDIC-KITcube, NAWDIC-iso) and (2) NAWDIC related projects (i.e., those addressing NAWDIC objectives and conducting complementary observations or simulations outside the main study region or main study period). For the NAWDIC core projects, the data collected must be shared with the NAWDIC community. For the NAWDIC-related projects, it is requested (but not required) that the data are made available to the NAWDIC community.
- *NAWDIC investigator:* any scientist committed to be involved in a NAWDIC project
- *NAWDIC community*: all NAWDIC investigators
- *Data provider:* any individual, group of individuals, or institution in charge of an instrument used to collect data or responsible for a data retrieval or numerical simulation
- *Data user:* any individual, group of individuals, or institution accessing NAWDIC data during or after the quality assurance period through temporary servers or data repositories
- *Data publication:* a published data set or a scientific journal article with the data set appended to it
- *Quick look:* visualization of preliminary data that is produced shortly after the respective data set was created

Appendix B – Recommended data repositories

Data repository	Web link	Data curation	Embargo period	Data volume limits	Notes
HALO Database	https://halo-db.pa.op.dlr.de/	yes	yes		Registration required; mandatory for HALO data sets
RADAR	https://radar.products.fiz-karlsruhe.de/en				For German institutions Service; agreement between institution and FIZ Karlsruhe required annual fee
zenodo	zenodo.org	no	yes	50 GB/record (higher quotas can be requested)	No curation
CEDA	archive.ceda.ac.uk	yes (level depending on anticipated data re-use)			NERC-funded projects
World Data Centre for Climate	www.wdc-climate.de/ui	yes		8 GB/file 1000 files/dataset	Approval required prior to data submission

The above list is not complete and contains only repositories that provide services to more than just a single institution. A full list of repositories that have been already harvested by the EDP can be found under https://earth-data.de/data.

Appendix C – Attribute conventions

Mandatory and optional global attributes in the published netCDF files are listed under https://hgf-ee-netcdf-conventions-draft.readthedocs.io/en/latest/global_attributes.html.

The following NAWDIC-specific attribute values should be used:

Valid *project* values:

- NAWDIC
- NAWDIC-HALO
- NAWDIC-DICHOTOMI
- NAWDIC-KITCUBE
- NAWDIC-ISO

Valid *platform* values:

- HALO
- ATR-42
- CESSNA-F406
- KITCUBE

Appendix D – Contact & funding information

The following table provides contact information of principle investigators of NAWDIC core projects and related instruments:

Project name	Contact person	Institution	Role/Instrument
NAWDIC-HALO	Julian Quinting	Uni Cologne	PI
	Andreas Schäfler	DLR-PA	PI
	Annika Oertel	KIT	PI
	Martin Wirth	DLR-PA	WALES
	Benjamin Witschas	DLR-PA	HEDWIG
	Andreas Wieser	KIT	KITsonde
	Daniel Kunkel	JGU Mainz	UMAQS
	Anja Stallmach	LMU Munich	SpecMACS
	Christian Rolf	FZ Jülich	FISH
	Andreas Zahn	KIT	FAIRO
	Jennifer Haase	Scripps/UC San Diego	ARO
NAWDIC-DICHOTOMI	Julian Quinting	Uni Cologne	PI
	Gwendal Rivière	LMD Paris	PI
	Philipp Gasch	KIT	AIRflows
NAWDIC-KITcube	Andreas Wieser	KIT	Science Director
NAWDIC-iso	Franziska Aemisegger	Uni Bern	PI/Picarro
	Harald Sodemann	Uni Bergen	PI
	Iris Thurnherr	ETH Zurich	PI

The following table lists information on the funding of NAWDIC core projects:

Project name	Funding agency	Grant no.
NAWDIC-HALO	German Science Foundation (DFG)	316646266, 552644175
NAWDIC-DICHOTOMI	German Science Foundation (DFG) & French National Research Agency (ANR)	545764561
NAWDIC-KITcube	Helmholtz Association	-