# AMIDER as an Integrated Publication System for Research Data

Masayoshi Kozai[1], Yoshimasa Tanaka[1], Shuji Abe[1], Yasuyuki Minamiyama[2] and Atsuki Shinbori[3]

[1]*Polar Environment Data Science Center, Joint Support-Center for Data Science Research, Research Organization of Information and Systems (PEDSC/ROIS-DS)*
[2]*Research Center for Open Science and Data Platform, National Institute of Informatics (RCOS/NII)*
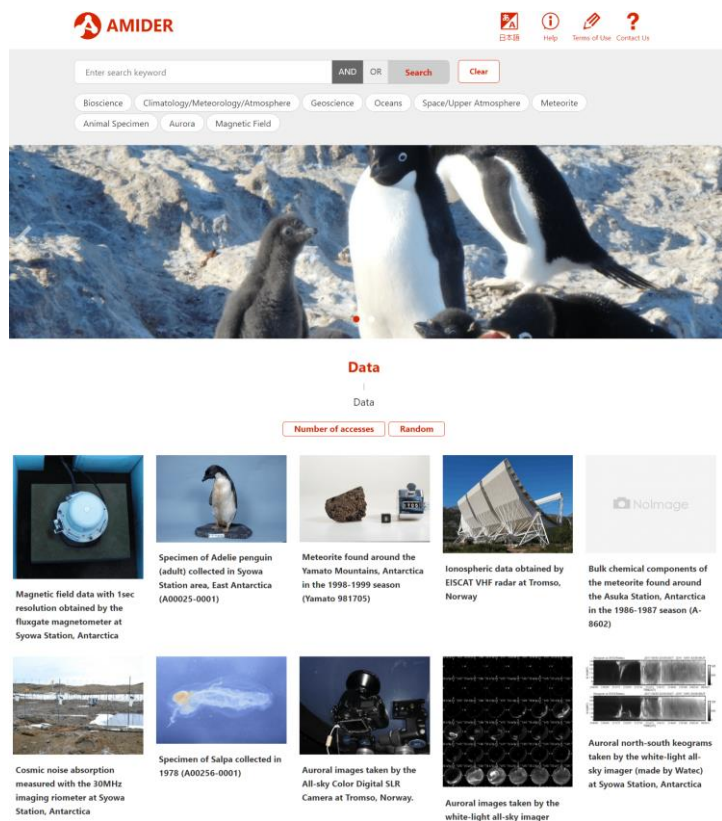[3]*Institute for Space-Earth Environmental Research (ISEE), Nagoya University*

Sharing research data, such as observation data in geospace science, forms a basis for studying complex systems in nature and is capable of promoting interdisciplinary research. The Polar Environment Data Science Center (PEDSC) of the Joint Support-Center for Data Science Research (DS), Research Organization of Information and Systems (ROIS), has been supporting the management, publication, and use of research data mainly in polar science. AMIDER is one of the pillar projects of PEDSC and will propose a model case of the research data publication in the next generation.

The catalog view on the AMIDER website displays research data in diverse scientific fields, e.g., geomagnetic observation, meteorites, and animal specimens. Its structure with thumbnails and snippets follows a common design in EC (electric commerce) or streaming sites, achieving a user-friendly interface that induces access from a wide range of users. The individual dataset page provides data downloading, visualized data (e.g., data plots), and metadata table in a uniform format. The AMIDER website has started a test operation and will be public in the near future.

We are starting further advanced attempts for the open science, utilizing the AMIDER as a test field. The key concepts are NLP (natural language processing) or text mining, networking between datasets, and DX (digital transformation) of data curation. The metadata attached to each research data enables us to connect the research data to outside data repositories in various categories, such as researchers, publications, and projects. Text mining of these big data will visualize the network of scientific activities. The data curation, including metadata creation, is a non-negligible load for researchers to publish research data. NLP is expected to also resolve this problem, supporting the creation and quality assurance of metadata. These studies will demonstrate the new role of the data publication system beyond its current primary function of sharing data within a specific scientific field.

AMIDER project is conducted by geospace researchers, while actively collaborating with researchers in information science such as library science. This scheme enables interaction between practical perspectives from data production activities, such as experimental projects, and advanced attempts utilizing data science which is rapidly developing.

Catalog view of AMIDER