

PROGRAMME and ABSTRACTS

11 – 15 December 2023



International Symposium on Data Science

DSWS - 2023

– Building an Open Data Collaborative Network
in the Asia-Oceania Area –



Science Council of Japan

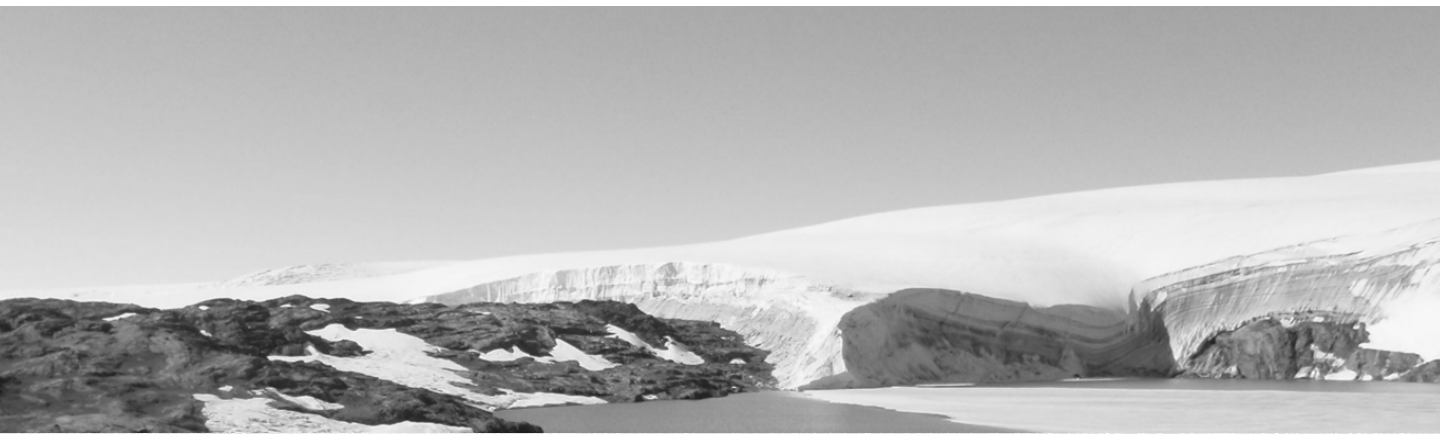
Roppongi, Minato-ku, Tokyo, Japan



International Symposium on Data Science

DSWS – 2023

– Building an Open Data Collaborative Network
in the Asia-Oceania Area –



PROGRAMME and ABSTRACTS

Science Council of Japan

11 – 15 December 2023

International Symposium on Data Science (DSWS-2023)

- Building an Open Data Collaborative Network in the Asia-Oceania Area -

11 – 15 December 2023

Science Council of Japan
Roppongi, Minato-ku, Tokyo, Japan
(hybrid type conference)

Background and Scope:

One of the objectives of the Open Science community is to develop and support new research and technologies based on the vast quantities of data being produced from a broad variety of scientific domains. The importance of the FAIR (Findable, Accessible, Interoperable, and Reusable) Principles for data have been recognized for data-oriented activities in this era of open science, leading to efforts to develop Open Data infrastructure, including: enhanced and integrated metadata catalogues, metadata standards for research data management, and certification of data repositories, among others. In addition, the use of persistent identifiers (PIDs; e.g., Digital Object Identifiers) for people, places, and other entities is becoming best practice in preservation and provision of data produced by their research activities.

Further effort is still needed to resolve various challenges relating to scientific research data, particularly the sharing and reuse of such data. Although the importance of multidisciplinary data integration has been widely advocated, data reuse by scientists within and across disciplines is still not easy from the point of view of the FAIR Principles. For example, there may be difficulties in discovering and accessing the data or insufficient information on the data to enable easy analysis. Often additional information is needed to assist when sharing data with a general audience, including policymakers. To improve the situation, it is important to stimulate collaborations among scientists from various disciplines, and to establish systems that facilitate the interactions between the users and providers of research data.

In response to the above challenges, this session welcomes submissions (either for “Research” or “Practice” presentations) on the following topics involving data-oriented activities in Asia and Oceania area among global scientific and/or non-scientific associations/initiatives. For example, sophisticated data sharing platforms have been planned and developed in the Asia-Oceania area through discussions held as part of previous conferences hosted by the World Data System (WDS). The goal of this session is to build consensus on various aspects of research data management by stakeholders in alignment with open research policies and FAIR principles. The session will explore

new ways of promoting interdisciplinary and collaborative research, data management platforms, and efficient data reuse under different scientific disciplines based on evidence and feedback from the Asia and Oceania communities.

It is expected that this conference will lead to a better mutual understanding of various aspects of data management by different stakeholders, and will open new paths for pursuing activities in different fields of science. The activities will play a central role in the promotion of inter-disciplinary sciences and new collaborative research paths based on multi-disciplinary data and directly contribute to global data activities based on the facilities provided by the "Joint Support-Center for Data Science Research (DS)" of the "Research Organization of Information and Systems (ROIS)".

Session Themes :

The session topics to be discussed in the symposium are:

- **Opening session, Keynote talks:** Conveners: *Masaki Kanao, Johnathan Kool, Juanle Wang
 - ✓ Opening addresses from conference organizers.
(Local Advisory Committee, International Advisory Committee, International Science Council)
 - ✓ Keynote talks by invited speakers.
 - ✓ Introduction of the symposium content, housekeeping issues, etc.

- **Challenges of data systems and networks:** Conveners: *Johnathan Kool, Toshihiko Iyemori, Chantelle Verhey, Susumu Goto, Jens Klump, Takashi Watanabe, Jung-Ho Um, Masaki Kanao
 - ✓ Sharing information on current status and activities of for creating and operating individual and organisational data centres and FAIR data repositories.
 - ✓ Discussions on international initiatives and alliances on collaborative disciplinary data systems and institutional data centres/repositories.
 - ✓ Sharing information on practices and practical challenges for stable, sustainable and/or transparent operation and development of data systems and repositories in a realistic environment and regulations.
 - ✓ Sharing information on regional activities regarding research data networks and discussing opportunities for promoting collaboration.

- **Open Science and the FAIR Principles: Why, what, and how:** Conveners: *Rorie Edmunds, Estelle Cheng, Hideaki Takeda, Kazuhiro Hayashi, Toshihiro Ashino
 - ✓ The FAIR Principles describe how research outputs, both physical and digital, should be findable, accessible, interoperable, and reusable. These concepts go hand-in-hand with ensuring the research endeavour is trustworthy and transparent under Open and Global Research Infrastructure. But, what can organizations and individuals do to ensure their research outputs are FAIR and Open? What methods, tools, and infrastructures are available to help them? What guidance exists for them to follow? With a focus on initiatives in the Asia–Oceania region, this session will present practice papers highlighting current trends for implementing Openness and FAIRness from perspectives such as FAIR-enabling frameworks and infrastructure, Open Science policy, and technologies that support multidisciplinary use of research outputs. Significant time will be devoted to discussion among (in-person and virtual) speakers and participants.

- **Lessons learned from COVID-19 data:** Conveners: *Tomoya Baba, Tadahiko Maeda, Mari Minowa, Elaine Faustman
 - ✓ Sharing information on COVID-19-related data across disciplines, covering social, genomic and biological aspects in the Asia and Oceania area.

- ✓ Exploring the global impact of COVID-19, the role of research collaboration and data sharing in mitigating effects of future pandemics. As this session has a broad international interest, all COVID-19-related data issues are welcome.
- **Recent developments in data science:** Conveners: *Asanobu Kitamoto, Kassim S. Mwitondi
 - ✓ Presentations on the applications of advanced data technologies for open science, such as big-data analysis, applications of machine-learning techniques, Artificial Intelligence, and data assimilation.
 - ✓ Reporting on new research projects leveraging multi-disciplinary data sources and interdisciplinary data-led technology innovations.
 - ✓ Developing recommendations to support data activities addressing open science, data-centric research and over-arching societal challenges across the Sustainable Development Goals (SDGs).
- **Involvement of early career researchers and scientists:** Conveners: *Shuji Abe, Akira Kadokura, Masahito Nose, Toshihiro Ashino, Susumu Goto, Takashi Watanabe, Maja Dolinar
 - ✓ A forum for young-generation researchers and scientists in the Asia-Oceania area to identify current and future societal challenges. This session is led by the WDS Early Career Researchers (ECR) Network Reports.
 - ✓ Developing proposals on training and activities aimed at increasing engagement of young researchers and scientists in the broad data science community.
- **Asia-Oceania data forum:** Conveners: *Yasuhiro Murayama, Toshihiko Iyemori, Toshihiro Ashino, Takashi Watanabe, David Castle, Juanle Wang
 - ✓ The forum will be an attempt of networking of data repositories and data sciences in the Asia-Oceania area, including WDS, CODATA and other international activities.
 - ✓ Discussions will be important on current status and future of open data/science infrastructures in the Asia-Oceania region, as well as in context of international global science data platform.
- **GEO variables and data mapping for Cold Regions:** Conveners: *Yubao Qiu, Hiroyuki Enomoto, Masaki Kanao, Lanhai Li
 - ✓ Global warming threatens the world's cold regions, while the essential variables of the Group on Earth Observations (GEO) provide mapping schemes for the data in the current EO system.
 - ✓ In this session, the GEO Cold Regions Initiative (GEO CRI) calls for a gathering of the GEO variables for cold regions and mapping existing or emergency data products for tackling the challenge of the warming world.
 - ✓ The topics could be essential variables, GEO data system, and data application, and pilot services using the data stream, and cases studies over cold regions by GEO.
- **Strategic discussion:** Conveners: *Takashi Watanabe, Juanle Wang, David Castle, Johnathan Kool, Noorsaadah Abd Rahman
 - ✓ The organizers provide opportunities for conducting individual committee meetings, data networks, and interdisciplinary research groups to discuss future activities in the Asia and Oceania area.
 - ✓ Focus on specific data-driven applications in the region, institutional collaborations and operational requirements and resources.

Invited Speakers:

Johnathan Kool (Australian Antarctic Division, Data Center)

Yubao Qiu (GEO Cold Regions Initiative (GEOCRI), International Research Center of Big Data for Sustainable

Development Goals)

Juanle Wang (Institute of Geographic Sciences and Natural Resources Research, China Academy of Science)

Kassim S. Mwitondi (Sheffield Hallam University)

Chantelle Verhey (WDS-International Technology Office)

Jens Klump (Mineral Resources, CSIRO)

Lanhai Li (Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences)

David Castle (University of Victoria)

Jung-Ho Um (Korea Institute of Science and Technology Information)

Noorsaadah Abd Rahman (Malaysian Open Science Alliance)

Pei-shan Liao (Research Center for Humanities and Social Sciences, Academia Sinica)

Maja Dolinar (University of Ljubljana)

Kannan Palavesam (Information and Library Network Centre, India)

Estelle Cheng (ORCID)

A.P.Dimri (Indian Institute of Geomagnetism)

Tyng-Ruey Chuang (Institute of Information Science, Academia Sinica)

Madiareni Sulaiman (University College London)

Conference Week Schedule:

Monday 11 December 2023; Registration, Ice breaker dinner

Tuesday 12 December 2023; Public lectures (in Japanese), Reception

Wednesday 13 December 2023; Symposium (Sessions; day 1), Banquet

Thursday 14 December 2023; Symposium (Sessions; day 2), Working dinner

Friday 15 December 2023; Symposium (Sessions; day 3), Strategic discussion

Conference Venue:

Science Council of Japan

Roppongi, Minato-ku, Tokyo, Japan

<https://www.scj.go.jp/en/index.html>

Followings are reference websites for the access to Roppongi (near SCJ) from the Airports (Haneda & Narita)

<https://www.asiacenter.or.jp/eng/location/> (Hotel Asia Center of Japan)

<https://www.grips.ac.jp/en/about/access/> (National Graduate Institute For Policy Studies)

<https://en.tokyo-midtown.com/access/#airport> (Tokyo Midtown)

Conference Programme (hybrid type conference) :

(Time; JST = UTC+9h)

Dates	Morning (10:00-12:00)	Lunch Time (12:15-14:00)	Afternoon-1 (14:00-15:30)	Afternoon-2 (16:00-17:30)	Dinner Time (19:00-)
11 DEC (MON)			Registration	Registration	Ice breaker dinner
12 DEC (TUE)	Registration		Public Lectures (in Japanese)	Public Lectures (in Japanese)	Reception

13 DEC (WED)	Opening session, Keynote talks	Poster Session (Lightning talks 1; 11:50-12:05) (Core time 1; 13:00-14:00)	Session 1: Challenges of data systems and networks	Session 2: Open Science and the FAIR Principles	Banquet
14 DEC (THU)	Session 3: Lessons learned from COVID-19 data	Poster Session (Lightning talks 2; 12:00-12:15) (Core time 2; 13:00-14:00)	Session 4: Recent developments in data science	Session 5: Involvement of early career researchers and scientists	Working dinner
15 DEC (FRI)	Session 6: Asia-Oceania data forum	Poster Session (Lightning talks 3; 12:00-12:15) (Core time 3; 13:00-14:00)	Session 7: GEO variables and data mapping for Cold Regions	Strategic discussion, Closing remark	

Presentations:

Oral Presentations:

All oral presentations will be made via the online Meeting Application (Zoom Webinar).

Presenters can attend the "Zoom Webinar" by either "in person at the Science Council of Japan" or by "online".

Presenters are allocated for 10-30 min. of talks depending on the programme of each session.

Detail of the link information on the Zoom meeting room for all sessions is available later.

Poster Presentations:

All poster presentations will be made in the virtual chat space (e.g., Spatial Chat).

Presenters can attend the "Spatial Chat" only by "online" and display their poster files during three days in 13-15 DEC.

Core presentation times in the "Spatial Chat" are allocated for the corresponding sessions in each day (13:00-14:00 JST, 13-15 DEC).

Lightning talk times are allocated by 1 minute for individual presenters just after the morning sessions in each day (13-15 DEC).

Detail of the link information on the Spatial Chat for all sessions is available later.

Registration:

Pre-conference registration is available from the following "Google Forms".

No registration fee is required to attend the conference.

["Pre-conference registration Form"](#)

For the participants who are staying in places/countries where the above Google Form cannot be accessed for any reasons, please use the "Offline pre-registration form (MS-excel file)". The "Offline pre-registration form" should be emailed to LOC by the end of November.

Onsite registration desk will be open on 11 December at the Science Council of Japan.

Accommodation and Dining:

As there are many hotels around central Tokyo area, participants are recommended to make their own reservations for their accommodation.

The LOC is ready to reserve several rooms for invited speakers and foreign participants in "[Hotel Asia Center of Japan](#)" located about 10 minutes' walk in north-west direction from the Science Council of Japan.

When invitation letter and/or VISA application form to come Japan are required, please inform to the LOC.

There are many restaurants around Nogizaka & Roppongi area and the Science Council of Japan. Lunch will not be served at the conference venue; therefore, delegates of the conference are recommended to use the restaurants in the vicinity of the Science Council of Japan.

A good resource for identifying the “halal” restaurants in Tokyo can be found here:

<https://www.halalgourmet.jp/?fbclid=IwAR2YEpC7LBB3U3hn-j2wL9MgwSz9CfM-Fp9nGkPJ5olaNzaE4CbcMLPgZh0>

It also seems to have an option for scanning for vegetarian menu when necessary.

Social Events:

There will be several social events as follows (public lectures, icebreaker dinner, reception, banquet, working dinner, etc.). All events can be registered from the "[Pre-conference registration Form](#)". (Times are written by JST in the followings)

(* formal events; reception & banquet)

- **Ice breaker dinner** (11 DEC 19:00–21:00). This is a town hall type casual dinner and will take place at a [Japanese-type restaurant](#) near the Science Council of Japan. Participation fee is approximately 5,500 JPyen. Vegetarian menu is not provided.
- **Public Lectures** (12 DEC 13:30–17:30). The public lectures are held for the general public, and it will be delivered in Japanese at the lecture hall of the Science Council of Japan. Final programme can be downloaded here.
- ***Reception** (12 DEC 19:00–21:00). This is a welcome party for the conference participants and will take place at a restaurant in "[Hotel Asia Center of Japan](#)" near the Science Council of Japan. Participation fee is approximately 6,000 JPyen. Vegetarian menu can be provided by buffe style dinner.
- ***Banquet** (13 DEC 19:00–21:00). This is a format event of the international symposium and will take place at a [Chinese-type restaurant](#) near the Science Council of Japan. Participation fee is approximately 7,000 JPyen. Vegetarian menu is not provided as the course menu, basically, but you may order them by specific plates.
- **Working dinner** (14 DEC 19:00–21:00). This is a town hall type casual dinner and will take place at a [Western-type restaurant](#) near the Science Council of Japan. Participation fee is approximately 6,000 JPyen. Vegetarian menu is not provided.

All payments for attending the foregoing events must be made at the registration desk of the conference, or at the desks for individual events.

More detailed information will be announced at the conference venue.

Special Issue:

The organizing committee is planning to launch the “Special Collection” in the Data Science Journal (<https://datascience.codata.org/>) as a special issue volume of the DSWS-2023 symposium.

Title of the Special Collection:

Building an Open Data Collaborative Network in the Asia-Oceania Area

- **Scope:**

This special collection derives from the International Symposium on Data Science (DSWS-2023; https://ds.rois.ac.jp/article/dsws_2023) that was held in Tokyo, Japan (11-15 December 2023). The symposium was organized by the Joint Support-Center for Data Science Research, Research Organization of Information and Systems (ROIS-DS) in collaboration with the Committee of International Collaborations on Data Science and the Science Council of Japan (SCJ). The event was also strongly supported and facilitated by the global data community, led by the World Data System (WDS) and the Committee on Data (CODATA) of the International Science Council (ISC). It aimed to facilitate information exchange regarding the archiving, publication, and utilization of diverse data relating to societal and global challenges such as COVID-19, information proliferation, global warming, extreme weather events, regional conflicts, etc., and their impact on the Asia-Oceania region.

The symposium was organized in several interdisciplinary scientific sessions involving international data activities in the Asia-Oceania region and beyond. They included various aspects of accreditation schemes and their benefits, individual international initiatives, data centres and networks, data management planning, data policies, legacy data, historical data, data sharing, citation and publication across disciplines.

Over 80 presentations were made, triggering fruitful discussions that focused on forming international collaborative networks related to open data in the region and establishing concrete cooperation frameworks within the global framework. The goal of the symposium was to build consensus on various aspects of research data management by stakeholders in alignment with open research policies and FAIR principles. The conducted scientific sessions could potentially lead to new ways of promoting interdisciplinary and collaborative research, data management platforms, and efficient data reuse under different scientific disciplines, based on evidence and feedback from the Asia and Oceania communities.

This special collection targets articles that outline best practices for attaining the foregoing goal. In particular, it seeks to publish research articles that relate to developing data systems and data analysis procedures from a multidisciplinary viewpoint. Contributions are not restricted to presentations made at the symposium, and so the editors would welcome submissions from any authors, globally, whose research and practical interests align with the symposium themes.

Further inquiries regarding the Special Issue can be directed to the Guest Editors.

- **Guest Editors:**

Tomoya Baba (Research Organization of Information and Systems)

David Castle (University of Victoria)

Masaki Kanao (Research Organization of Information and Systems)

Johnathan Kool (Australian Antarctic Division)

Kassim S. Mwitondi (Sheffield Hallam University)

Yubao Qiu (GEO Cold Regions Initiative)

Juanle Wang (China Academy of Science; Editorial Board of the Data Science Journal)

- **Expression of Interest (Eoi):**

Please input your Expression of Interest for the "Special Collection" to the following Google Form; <https://docs.google.com/forms/d/e/1FAIpQLSeENmAbBSBeHgLow5tybfljYnsudjZZaiKUi1i7Aoz9AR4MhA/viewform>

You are required to input the information on "Author(s), Affiliation(s), Contact Address and Tentative Article Title(s)". This Eoi Form will be closed by 29 February 2024.

- **Deadline of Article Submission: 31 July 2024**

- **Final Publishing Online for All Articles: 31 March 2025 (provisional)**

International Advisory Committee (IAC) : (* Chair)

- * Johnathan Kool (Australian Antarctic Division, Data Center)
- David Castle (University of Victoria)
- Estelle Cheng (ORCID)
- A. P. Dimri (Director, Indian Institute of Geomagnetism)
- Maja Dolinar (University of Ljubljana)
- Rorie Edmunds (DataCite)
- Elaine M. Faustman (University of Washington)
- Meredith Goins (WDS-International Program Office)
- Simon Hodson (Executive Director, CODATA)
- Jens Klump (Mineral Resources, CSIRO)
- Lanhai Li (Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences)
- Pei-shan Liao (Research Center for Humanities and Social Sciences, Academia Sinica)
- Claudia Bauzer Medeiros (Institute of Computing, University of Campinas)
- Kassim S. Mwitondi (Sheffield Hallam University)
- Mark Parsons (CODATA Data Science Journal, Editor in Chief)
- Yubao Qiu (GEO Cold Regions Initiative (GEOCRI), International Research Center of Big Data for Sustainable Development Goals)
- Noorsaadah Abd Rahman (Malaysian Open Science Alliance)
- Jung-Ho Um (Korea Institute of Science and Technology Information)
- Chantelle Verhey (WDS-International Technology Office)
- Juanle Wang (Institute of Geographic Sciences and Natural Resources Research, China Academy of Science)

Local Advisory Committee (LAC) : (* Chair)

- * Hiroyuki Araki (Joint Support-Center for Data Science Research, ROIS)
- Toshihiro Ashino (Toyo University)
- Hiroyuki Enomoto (National Institute of Polar Research, ROIS)
- Susumu Goto (Database Center for Life Science, DS, ROIS)
- Kazuhiro Hayashi (National Institute of Science and Technology Policy)
- Asanobu Kitamoto (Center for Open Data in the Humanities, DS, ROIS)
- Yuji Kohara (Database Center for Life Science, DS, ROIS)
- Tadahiko Maeda (Center for Social Data Structuring, DS, ROIS)
- Hideki Noguchi (Center for Genome Informatics, DS, ROIS)
- Masahito Nose (Nagoya City University)
- Hideaki Takeda (National Institute of Informatics, ROIS)
- Seiji Tsuboi (Japan Agency for Marine-Earth Science and Technology)
- Genta Ueno (Center for Data Assimilation Research and Applications, DS, ROIS)

Local Organizing Committee (LOC) : (* Chair)

- * Masaki Kanao (Polar Environment Data Science Center, DS, ROIS)
- Shuji Abe (Polar Environment Data Science Center, DS, ROIS)
- Tomoya Baba (Data Science Promotion Section, DS, ROIS)

Mamoru Ishii (National Institute of Information and Communications Technology)
Toshihiko Iyemori (Kyoto University; International Union of Geodesy and Geophysics; WDS-SC)
Akira Kadokura (Polar Environment Data Science Center, DS, ROIS)
Kumiko Kanekawa (Data Science Promotion Section, DS, ROIS)
Yashuhisa Kondo (Research Institute for Humanity and Nature)
Masayoshi Kozai (Polar Environment Data Science Center, DS, ROIS)
Mari Minowa (Data Science Promotion Section, DS, ROIS)
Yasuhiro Murayama (National Institute of Information and Communications Technology)
Akihiko Nomizu (Data Science Promotion Section, DS, ROIS)
Jun'ichi Okuno (Polar Environment Data Science Center, DS, ROIS)
Kunio Takahashi (Polar Environment Data Science Center, DS, ROIS)
Yoshimasa Tanaka (Polar Environment Data Science Center, DS, ROIS)
Takashi Watanabe (WDS Japan)

Organized by:

Joint Support-Center for Data Science Research (DS), Research Organization of Information and Systems (ROIS)
Committee of International Collaborations on Data Science, Science Council of Japan (SCJ)

Supported by:

National Institute of Information and Communications Technology (NICT)
Research Organization of Information and Systems (ROIS)
DNA Data Bank of Japan (DDBJ)
World Data System (WDS) of the International Science Council (ISC)
Committee on Data (CODATA) of the International Science Council (ISC)
Science Council of Japan (SCJ)

Important Dates:

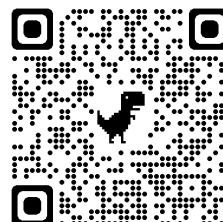
Registration & Abstract submission open: 1 June 2023
Fixing the speakers & program of the oral sessions: 31 August 2023
Abstract file submission deadline: 30 September 2023
Program booklet online: 1 November 2023
Registration deadline: 30 November 2023
Symposium date: 11-15 December 2023

Contact Address:

dsws.loc-2023@nipr.ac.jp

Workshop Website :

https://ds.rois.ac.jp/article/dsws_2023/



International Symposium on Data Science

DSWS – 2023

– Building an Open Data Collaborative Network
in the Asia-Oceania Area –



PROGRAMME

Science Council of Japan

11 – 15 December 2023

PUBLIC LECTURES (in Japanese)

International Symposium on Data Science (DSWS-2023)

- Building an Open Data Collaborative Network in the Asia-Oceania Area

12 December 2023

Science Council of Japan, Tokyo (hybrid type conference)

(Time; JST = UTC+9h)

Tuesday 12 December 2023

Public Lectures (13:30-17:30 JST)

Facilitators: *Akira Kadokura, Kumiko Kanekawa

13:30–14:05 **Opening addresses** (Chair: Akira Kadokura)

- **Hiroyuki Araki** (Chair of Local Advisory Committee; Director-General, Joint Support-Center for Data Science Research (DS), ROIS) (10')
- **Yasuhiro Murayama** (Chair of Committee of International Collaborations on Data Science, Science Council of Japan; National Institute of Information and Communications Technology) (10')
- **Sakae Shibusawa** (Secretary General, Science Council of Asia) (15')

14:05–14:15 **Break**

14:15–15:55 **Part 1: Keynote talks** (Chair: Toshihiro Ashino)

Keynote 1: Scientific data activities in Southeast Asian studies

Shoichiro Hara (Center for Southeast Asian studies, Kyoto University) (20')

Keynote 2: Scientific Data on Air Pollution Studies in South Asia – Perspectives on the Aakash Project Activities

Sachiko Hayashida (Nara Women's University) (20')

Keynote 3: Open Science Initiatives in Malaysia

Noorsaadah Abd Rahman (University of Malaysia) (20')

Keynote 4: Convention on Biological Diversity (CBD) and Open Science

Masanori Arita (DNA Data Bank of Japan) (20')

Keynote 5: New Technologies and Emerging Opportunities in Data Management

Johnathan Kool (Australian Antarctic Division) (20')

15:55–16:05 **Break**

16:05–17:25 **Part 2: Panel discussion** (Chair: Yasuhisa Kondo)

- **Mamoru Ishii** (National Institute of Information and Communications Technology) (10')
- **Tadahiko Maeda** (Center for Social Data Structuring, DS, ROIS) (10')
- **Mari Minowa** (Data Science Promotion Section, DS, ROIS) (10')
- **Jun Matsumoto** (Tokyo Metropolitan University), **Shigeko Haruyama** (Mie University) (10')
- **Yasuhisa Kondo** (Research Institute for Humanity and Nature) (10')

Discussion (30')

17:25–17:30 **Concluding Remarks** (Akira Kadokura)

PROGRAMME

International Symposium on Data Science (DSWS-2023)

- Building an Open Data Collaborative Network in the Asia-Oceania Area

13 – 15 December 2023

Science Council of Japan, Tokyo (hybrid type conference)

(Time; JST = UTC+9h)

Wednesday 13 December 2023

Opening session, Keynote talks (10:00-12:00 JST)

Conveners: *Masaki Kanao, Johnathan Kool, Juanle Wang

10:00–10:40 **Opening addresses** (Chair: Masaki Kanao)

- **Hiroyuki Araki** (Chair of Local Advisory Committee; Director-General, Joint Support-Center for Data Science Research (DS), ROIS) (5')
- **Johnathan Kool** (Chair of International Advisory Committee; Data Center Manager, Australian Antarctic Division) (5')
- **Motoko Kotani** (Chair of the Committee of Science Planning (CSP), Vice President for Science and Society, international Science Council; Tohoku University) (10')
- **David Castle** (Chair of Scientific Committee of World Data System; University of Victoria) (10')
- **Masaki Kanao** (Chair of Local Organizing Committee), outline of symposium & practical information (5')

10:40–10:50 **Coffee break**

10:50–11:50 **Keynote talks** (Chairs: Johnathan Kool, Juanle Wang)

Keynote 1: Challenges of Open science data governance and China's practice

Juanle Wang (Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences) (15')

Keynote 2: Malaysian Open Science Platform (MOSP)

Noorsaadah Abd Rahman (Malaysian Open Science Alliance) (15')

Keynote 3: GEO Cold Regions Initiative (GEOCRI)

Yubao Qiu (Aerospace Information Research Institute, Chinese Academy of sciences) (15')

Keynote 4: International Collaborative Data Management – Challenges and Opportunities

Johnathan Kool (Australian Antarctic Division) (15')

11:50–12:05 **Lightning Talks 1 (Poster Sessions)**

12:05–12:15 **Group photo @ SCJ in person & Zoom screen**

12:15–14:00 **Lunch Time**

13:00–14:00 **Core Time 1 (Poster Sessions)**

Wednesday 13 December 2023

Session 1: Challenges of data systems and networks (14:00-15:30 JST)

Conveners: *Johnathan Kool, Toshihiko Iyemori, Chantelle Verhey, Susumu Goto, Jens Klump, Takashi Watanabe, Jung-Ho Um, Masaki Kanao

14:00– 15:30 (Chairs: Chantelle Verhey, Jens Klump, Jung-Ho Um)

- *Modernisation of the AuScope Virtual Research Environment – Design and lessons learnt*
Jens Klump (Mineral Resources, Commonwealth Scientific and Industrial Research Organization) (15')
- *DataON: platform for discovery and analysis of research data*
Jung-Ho Um (Korea Institute of Science and Technology Information) (15')
- *Enhancing Social Science Research in East Asia: An Introduction to SRDA and NASSDA*
Pei-shan Liao (Research Center for Humanities and Social Sciences, Academia Sinica, Taiwan) (15')
- *BigFlow: The End-to-End Workflow Framework for Cross-Center Scientific Data Analysis*
Xiaojie Zhu (Computer network Information Center, Chinese Academy of Sciences) (15')
- *Polar data Discovery and Mobilization Pathways*
Chantelle Verhey (WDS-International Technology Office) (15')

Discussion (15')

15:30–16:00 **Break**

Wednesday 13 December 2023

**Session 2: Open Science and the FAIR Principles:
Why, what, and how (16:00-17:30 JST)**

Conveners: *Rorie Edmunds, Estelle Cheng, Hideaki Takeda, Kazuhiro Hayashi, Toshihiro Ashino

16:00—17:30 (Chairs: Rorie Edmunds, Estelle Cheng)

- *You Can't Make Science Open Without Committing to PIDs*
Estelle Cheng (ORCID)
Rorie Edmunds (DataCite) (15')
- *Open Science and FAIR Principles in Indonesia*
Hermin Triasih & Seno Yudhanto (National Research and Innovation Agency, Indonesia) (12')
- *IRINS: Building FAIR Research information System in India*
Kannan Palavesam (Information and Library Network Centre, India) (12')
- *An Advanced identifier architecture for enabling FAIR principle: Global Open Identifier*
Jia Liu (Computer Network Information Center, Chinese Academy of Sciences) (12')

- *Open Science Policy Development in Japan: Importance of an intermediate sector balancing between Top-down and Bottom-up*
Kazuhiro Hayashi (National Institute of Science and Technology Policy, Japan) (12')
- *Making FAIR Scale*
Erik Schultes (GO FAIR Foundation & Leiden Academic Center for Drug Research) (12')

Discussion (15')

Thursday 14 December 2023

Session 3: Lessons learned from COVID-19 data (10:00-12:00 JST)

Conveners: * Tomoya Baba, Tadahiko Maeda, Mari Minowa, Elaine Faustman

10:00–12:00 (Chairs: Tomoya Baba, Tadahiko Maeda)

- *Reviewing COVID-19 Management in Practice by a comparative study of India, Taiwan, USA and Japan through the Perspective of TQM and Preparing for Future pandemics*
Kazuyuki Suzuki (University of Electro-Communications, Japan) (20')
- *Towards a global and integrated wastewater surveillance system: a review of historical and latest activities*
Noriko Endo (Kyoto University, Japan) (20')
- *Wastewater Surveillance Data: Tracking Imported SARS-CoV-2 Variants*
Jatuwat Sangsanont (Chulalongkorn University, Thailand) (20')
- *Harnessing Advanced Data Science and Machine Learning in Smart Wastewater-Based Epidemiology*
Guangming Jiang (University of Wollongong, Australia) (20')
- *Airborne surveillance of bioaerosol in built environment for future pandemic preparedness: Beyond COVID-19*
Irvan Luhung (Nanyang Technological University, Singapore) (20')
- *Issues of Digital Sequence Information (DSI) emerged from the COVID-19 pandemic*
Masanori Arita (DNA Data Bank of Japan, National Institute of Genetics, Japan) (20')

12:00–12:15 **Lightning Talks 2 (Poster Sessions)**

12:15–14:00 **Lunch Time**

13:00–14:00 **Core Time 2 (Poster Sessions)**

Thursday 14 December 2023

Session 4: Recent developments in data science (14:00-15:30 JST)

Conveners: *Asanobu Kitamoto, Kassim S. Mwitondi

14:00–15:30 (Chairs: Asanobu Kitamoto, Kassim S. Mwitondi)

- *A Recursive Method for Identifying Glacial Landforms in Hillshaded Areas Using Digital Elevation Data and CNN Models*

Kassim S. Mwitondi (Sheffield Hallam University) (15')

- *Synergies of Satellite Datasets, Modeling and Data Analytic Approaches in Pollutant Retrieval*
Hugo Wai Leung MAK (The Chinese University of Hong Kong;
The Hong Kong University of Science and Technology) (15')
- *NSSDC's Data Governance Practices for AI for Space Science*
Zou Ziming (National Space Science Center, Chinese Academy of Sciences) (15')
- *OpenCSDB: Enabling Knowledge-Driven Scientific Data Linking and Integration*
Han Fang (Computer Network Information Center, Chinese Academy of Sciences) (15')
- *A Robust Data-Driven Method for Quantifying Methane Emission from Cattle*
Tebogo Sebake (African Earth Observation Network, Nelson Mandela University) (15')

Discussion (15')

15:30–16:00 **Break**

Thursday 14 December 2023

Session 5: Involvement of early career researchers and scientists (16:00-17:30 JST)

Conveners: *Shuji Abe, Akira Kadokura, Masahito Nose, Toshihiro Ashino, Susumu Goto, Takashi Watanabe, Maja Dolinar

16:00–17:30 (Chairs: Shuji Abe, Susumu Goto, Masahito Nose)

- *Empowering the Future of Data Science: Introducing the Activities of the WDS-ECR Network*
Maja Dolinar (Slovenian Social Science Data Archives, Faculty of Social Sciences, University of Ljubljana; Co-Chair of the World Data System ECR Network) (15')
- *CODATA Connect Early Career and Alumni initiative*
Shaily Gandhi (CEPT Research and Development Foundation, India) (15')
- *When Graduates Increase, So Must Faculty Members: Project for Training Experts in Statistical Sciences*
Yoshinori Kawasaki (The Institute of Statistical Mathematics) (15')
- *Introduction of Data Science Activities for Young Researchers in IUGONET*
Shun Imajo (Data Analysis Center for Geomagnetism and Space Magnetism,
Graduate School of Science, Kyoto University) (15')
- *Navigating the Research Data Management Compliance Maze as an Indonesian Early-Career Researcher: Unravelling Policies Across Borders*
Madiareni Sulaiman (National Research and Innovation Agency, Indonesia;
University College London) (15')
- *DAEPO Projects of National Astronomical Data Center (NADC) in China*
Shanshan Li (National Astronomical Observatories, Chinese Academy of Sciences) (15')

Friday 15 December 2023

Session 6: Asia-Oceania data forum (10:00-12:00 JST)

Conveners: *Yasuhiro Murayama, Toshihiko Iyemori, Toshihiro Ashino, Takashi Watanabe, David Castle, Juanle Wang

10:00–12:00 (Chairs: Yasuhiro Murayama, Juanle Wang, Tyng-Ruey Chuang, Toshihiko Ashino)

- *International Policy and Practices of Open Science-based Data Sharing*
Yasuhiro Murayama (National Institute of Information and Communications Technology) (5')
- *A Survey of current Status of Open Science oriented Activities in the Asia-Oceania Area*
Takashi Watanabe (National Institute of Information and Communications Technology) (10')
- *The Action Plan and Priorities of the World Data System*
David Castle (Chair of Scientific Committee of World Data System; University of Victoria) (15')
- *Toward a Reproducible Research Data Repository*
Cheng-Jen Lee & Tyng-Ruey Chuang (Institute of Information Science,
Academia Sinica, Taiwan) (15')
- *Overview of Research Activities at Indian Institute of Geomagnetism*
A. P. Dimri (Indian Institute of Geomagnetism) (10')
- *The role of data publishing activities at universities and research institutes on Asia and Oceania*
Shaliza Ibrahim (University of Malaya) (10')
- *National Scientific Repository (RIN)*
Sjaeful Afandi (National Research and Innovation Agency (BRIN of Indonesia)) (10')
- *Towards Open Science: Chinese National Spatial Science Data Centre Activities*
Xu Qi (National Space Science Center, Chinese Academy of Sciences) (10')
- *Open Government Data (OGD) in Bangladesh NSO*
Chandra Shekhar Roy (Bangladesh Bureau of Statistics) (10')
- *The Global Microbiology Open Science Data Cooperation*
Junca Ma (Institute of Microbiology, Chinese Academy of Sciences) (10')

Discussion (15')

12:00–12:15 **Lightning Talks 3 (Poster Sessions)**

12:15–14:00 **Lunch Time**

13:00–14:00 **Core Time 3 (Poster Sessions)**

Friday 15 December 2023

Session 7: GEO variables and data mapping for Cold Regions (14:00-15:30 JST)

Conveners: *Yubao Qiu, Hiroyuki Enomoto, Masaki Kanao, Lanhai Li

14:00– 15:30 (Chairs: Yubao Qiu, Hiroyuki Enomoto)

- *Snow Properties Survey and Snow Avalanche Hazards in Tianshan Mountains*
Lanhai Li (Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences) (15')
- *Long term records of glacier evolution and associated proglacial lakes on the Tibetan Plateau 1976 - 2020*
Lhakpa Dolma (Tibet Institute of Plateau Atmosphere and Environmental Sciences)
Yubao Qiu (Aerospace Information Research Institute, Chinese Academy of sciences) (15')
- *Data Management of the Arctic and Antarctic Region*
Hironori Yabuki (National Institute of Polar Research) (15')
- *Utilization of Snow and Ice Disaster Prevention Information created by National Research Institute for Earth Science and Disaster Resilience (NIED), Japan*

Satoru Yamaguchi (National Research Institute for
Earth Science and Disaster Resilience) (15')

- *Digital Arctic Environment Information Service*

Yubao Qiu (Aerospace Information Research Institute, Chinese Academy of sciences) (15')

Discussion (15')

15:30–16:00 **Break**

Friday 15 December 2023

Strategic discussion, Closing remark (16:00-17:30 JST)

Conveners: *Takashi Watanabe, Juanle Wang, David Castle,
Johnathan Kool, Noorsaadah Abd Rahman, A. P. Dimri

16:00–17:30

(Chairs: Takashi Watanabe, Juanle Wang)

1. *Background and Scope of the Strategic discussion*

Takashi Watanabe (National Institute of Information and Communications Technology) (5')

2. *Summary of the DSWS-2023 individual sessions*

Lead-conveners of all sessions (35')

- *Session 1: Challenges of data systems and networks (5')*
- *Session 2: Open Science and the FAIR Principles: Why, what, and how (5')*
- *Session 3: Lessons learned from COVID-19 data (5')*
- *Session 4: Recent developments in data science (5')*
- *Session 5: Involvement of early career researchers and scientists (5')*
- *Session 6: Asia-Oceania data forum (5')*
- *Session 7: GEO variables and data mapping for Cold Regions (5')*

3. *Supplementary comments and discussions (20')*

4. *Action plan (10')*

5. *Summary (5')*

(Chair: Masaki Kanao)

Closing remark (15')

Johnathan Kool (Chair of International Advisory Committee; Australian Antarctic Division) (5')

Noorsaadah Abd Rahman (Malaysian Open Science Alliance) (5')

Masaki Kanao (Chair of Local Organizing Committee) (5')

Poster Sessions

Abstracts will be included in the conference “Programme and Abstracts” online booklet.

Wednesday 13 December 2023

Lightning Talks 1; 11:50-12:05 JST (@Zoom)

Core time 1; 13:00-14:00 JST (@Spatial Chat)

(Chairs: Toshihiko Iyemori, Susumu Goto)

- *Advancing Data Stewardship: Join the World Data System and Contribute to Scientific Data Excellence*
Vincent Bernabeo & Meredith Goins (World Data System International Program Office) (S1)
- *Building the Web Site of the WDS Asia-Oceania network*
Takashi Watanabe (National Institute of Information and Communications Technology) (S1)
- *Data and metadata sharing among Asian countries*
Masaki Kanao (Joint Support-Center for Data Science Research (DS), ROIS) (S1)
- *Concept of NIES Environmental Research Hub: A Data Infrastructure to support/promote Environmental Research*
Tomoko Shirai (National Institute for Environmental Studies) (S1)
- *Survey for policy of research data openness in South Korean government-funded research institutes*
Jung-Ho Um (Korea Institute of Science and Technology Information) (S2)
- *Challenges of Open Research Data in Indonesia: Case Study National Scientific Repository*
Seno Yudhanto (National Research and Innovation Agency, Indonesia) (S2)
- *MosquitoDB: A comprehensive-electronically-based entomological surveillance system for the control and elimination of mosquito-borne diseases*
Victor Mero & Janice Maige (Ifakara Health Institute, Environmental Health, and Ecological Sciences Department, Tanzania) (S2)
- *Further improve accessibility of JAMSTEC research data - DOI assignments to cruise reports -*
Ayumi Sumikawa (Japan Agency for Marine-Earth Science and Technology) (S2)
- *Scientometric Analysis and Mapping of Research Publications of Central University of Gujarat*
Kotrayya Agadi (Central University of Gujarat, Gandhinagar, India) (S2)
- *Indian Language Open Educational Resources for the Implementation of the National Education Policy 2020: A Comprehensive Evaluation*
Kotrayya Agadi (Central University of Gujarat, Gandhinagar, India) (S2)
- *Effects of information on behavioral change during Covid-19 pandemic: An International comparative study between Japan, UK, and Taiwan*
Naoko Kato-Nitta (J.F. Oberlin University, ROIS-DS and The Institute of Statistical Mathematics) (S3)

Thursday 14 December 2023

Lightning Talks 2; 12:00-12:15 JST (@Zoom)

Core time 2; 13:00-14:00 JST (@Spatial Chat)

(Chairs: Elaine Faustman, Mari Minowa)

- *Genomic Surveillance of SARS-CoV-2 based on comparison among Japanese cities*
Takashi Abe (Niigata University, Japan) (S3)
- *Unsupervised AI analysis of oligonucleotide usage in a wide range of variants of SARS-CoV-2*
Toshimichi Ikemura (Nagahama Institute of Bio-Science and Technology, Japan) (S3)
- *Host cell-dependent genomic changes in SARS-CoV-2*

- **Yuki Iwasaki** (Nagahama Institute of Bio-Science and Technology, Japan) (S3)
SARS-CoV-2 HaploGraph: visualization of SARS-CoV-2 haplotypes spread in Japan
- **So Nakagawa** (Tokai University, Japan) (S3)
Urban Microbiome Monitoring in Japan: Collaborative Research with MetaSUB Consortium
- **Haruo Suzuki** (Keio University, Japan) (S3)
Comparison of microbial diversity and resistant determinants in the built environment in Japan using shotgun metagenomic sequencing
- **Dewa A.P. Rasmika Dewi** (International University of Health and Welfare, Japan) (S3)
Evaluation of rRNA depletion methods for capturing the RNA virome from environmental surface
- **Yuh Shiwa** (Tokyo University of Agriculture, Japan) (S3)
Wastewater-Based Epidemiology for Efficient SARS-CoV-2 Surveillance in Three Japanese Cities
- **Ryo Honda** (Kanazawa University, Japan) (S3)
Implementation of wastewater-based epidemiology for viral infectious diseases in Japan
- **Masaaki Kitajima** (Hokkaido University, Japan) (S3)
Investigation of experimental procedures for detecting SARS-CoV-2 from airborne low-density RNA samples
- **Jun Uetake** (Hokkaido University, Japan) (S3)
Educational Activities for Young Researchers in Kyusyu University's MAGDAS Project
- **Shuji Abe** (Polar Environment Data Science Center,
Joint Support-Center for Data Science Research (DS), ROIS;
International Research Center for Space and Planetary Environmental Science,
Kyushu University) (S5)

Friday 15 December 2023

Lightning Talks 3; 12:00-12:15 JST (@Zoom)

Core time 3; 13:00-14:00 JST (@Spatial Chat)

(Chairs: Akira Kadokura, Masahito Nose)

- *Analysis on characteristics of surface-observed and model-based cloud cover for historical weather reconstruction using data assimilation*
Xiaoxing Wang (Center for Open Data in the Humanities,
Joint Support-Center for Data Science Research (DS), ROIS) (S4)
- *Explainable AI Framework for Interpreting the Electrochemical Oxygen Evolution Activity of (Ni-Fe-Co-Ce)Ox Catalysts*
Paul Rossener Regonia (University of the Philippines Diliman) (S4)
- *Predictive Modelling of Under-Five Mortality Determinants Using Machine Learning Techniques*
Rakesh Kumar Saroj (Jawaharlal Nehru University, India) (S4)
- *How Can Digital Innovation Transform People's Behaviour in Response to the COVID-19 Restrictions in Indonesia*
Kiki Adhinugraha (La Trobe University, Australia) (S4)
- *Hyper basic oxygen*
Norihan Abdelaziz Elgazzar (Tanta University, Egypt) (S4)
- *Data publication and collaborative research infrastructure of the Antarctic Syowa MST/IS (PANSY) radar*
Taishi Hashimoto (National Institute of Polar Research) (S4)
- *An overview of metadata in Research Data Clouds towards Global Open Research Commons*
Mikiko Tanifuji (National Institute of Informatics) (S4)
- *Progress of the AMIDER project*
Masayoshi Kozai (Polar Environment Data Science Center,
Joint Support-Center for Data Science Research (DS), ROIS) (S5)

- *Conducting a Lesson on Earth's Global Sea Ice Distribution Using a Handmade Globe*
Takeshi Sugimura (National Institute of Polar Research) (S5)
- *Educational practices using polar sea ice data for high school science and mathematics classes*
Yoshihiro Niwa (National Institute of Polar Research) (S5)
- *Application of Independent Component Analysis to Satellite Gravity Data for Antarctic Research*
Tianyan Shi (Institute for Geothermal Sciences, Kyoto University) (S7)

International Symposium on Data Science

DSWS – 2023

– Building an Open Data Collaborative Network
in the Asia-Oceania Area –



ABSTRACTS

Science Council of Japan

11 – 15 December 2023

Challenges of Open science data governance and China's practice

Juanle Wang^{1*}

^{1*} *Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences*
Email: wangjl@igsnr.ac.cn

Summary. Open data is a crucial component of open science, and its data governance is a global challenge. The FAIR principles (Findable, Accessible, Interoperable and Reusable) were proposed and quickly spread across multiple domains in data governance, which provided an effective way for scientific data governance. The study firstly analyzed the progress and characteristics of FAIR principles in different disciplinary fields based on a bibliometric method. On this basis, we put forward ten recommendations for scientific data governance. Combining the effectiveness and experience in terms of open data governance of National Science Data Centers and World Data System Science Data Centers in China, we analyzed common challenges and then proposed strategies for scientific data governance. Moreover, combining research practices of the Deep Time Digital Earth Science Program (DDE), we also put forward to the specific FAIR governance preliminary framework for Geoscience.

Keywords. open science, data governance, FAIR principles, data sharing, suggestions.

Malaysian Open Science Platform (MOSP)

Noorsaadah Abd. Rahman^{1*}

^{1*} *Malaysian Open Science Alliance*

Email: noorsaadaharahman@gmail.com

Summary. The 4th National Policy for Science, Technology and Innovation (2021-2030) or NPSTI for Malaysia commenced in 2021. The promotion of and role of Open Science in supporting this new policy are indicated in some of its strategies to revitalise the nation's STI ecosystem. Thus, Malaysia embarked on the Open Science movement through the launching of the Malaysia Open Science Platform (MOSP). MOSP is aimed to gather and consolidate Malaysia's research data in a platform that allows accessibility and sharing of these research data. Managed by the Academy of Sciences Malaysia through the Malaysia Open Science Alliance, MOSP which is a two and a half year pilot project (2020 to 2023) involving five Research Universities in the country. The task is to look into the initial three main areas, namely:

- (i) Guidelines for Open Science;
- (ii) Capacity Building and Awareness; and
- (iii) Open Science Infrastructure.

The output and challenges faced while implementing this program is discussed.

Keywords. Malaysian Open Science, MOSP, research data platform.

GEO Cold Regions Initiative (GEOCRI)

Yubao Qiu^{1, 2*}, **Massimo Menenti**³, **Hiroyuki Enomoto**⁴

^{1*} *International Research Center of Big Data for Sustainable Development Goals (CBAS)*

² *Aerospace Information Research Institute, Chinese Academy of Sciences, China*

³ *Delft University of Technology, Netherland*

⁴ *Arctic Environment Research Center (AERC) of the National Institute of Polar Research (NIPR), Japan*

Email: qiuyb@aircas.ac.cn

Summary. Under a global warming scenario, the high elevation and high latitude cold regions, dominated by the cryosphere elements, are inherently fragile to the environment, where changes in the phase of water and the induced result to the environment affect billions of human lives there and the downstream area. Societal and economic development has been leading to a growing dependence on natural, ecosystem, and environmental resources. The warming reshaped the cryosphere and its embed regions, influencing the societal benefits of water availability in the downstream areas, transportation in opening Northern Sea Routes (NSR), the infrastructure and road stabilities in permafrost-rich areas, food and agriculture strategy by blooming ecosystem, climate and weather forecasting service, and challenges addressing and assessment to the sustainable development of cold regions.

Timely and accurate information on the cryosphere elements, like snow, glaciers, permafrost, freshwater ice, sea ice, and even solid precipitation, is necessary to protect fragile ecosystems and the environment, facilitate sustainable exploitation of environmental resources, provide forcing data to hydrometeorological services, support the safe use of the land and ocean facilitates, and thus evaluate and foster addressing the sustainable development goals.

GEOCRI brings together the efforts of different science and industry communities' activities currently and stakeholders in the world's cold regions. The core interest of the GEOCRI is to bring fruitful information, gathered continuously by the national and multi-national, growing infrastructures of diverse and complementary Earth observations, to users on a global scale. The contributors to the objectives of GEOCRI are currently operating observational and data infrastructures with high-performance data streaming processing capabilities with open data principles on an international platform. Likewise, data systems have been developed and are hosting rich data assets.

We expect the initiative to generate continuous data streams on Essential Cold Regions Variables (ECRVs), and provide pilot services on the water availabilities in the cold mountain area, safety transportation for the land and northern sea routes, emerging cryosphere disaster mitigation, and assessment supporting the UN Sustainable Development Goals (UN SDGs), etc.

Keywords. Cold Regions, Earth observations, Variables and Products, Information services, Data Stream.

International Collaborative Data Management – Challenges and Opportunities

Johnathan Kool ^{1*}

^{1*} *Australian Antarctic Division / Standing Committee on Antarctic Data Management*

Email: johnathan.kool@aad.gov.au

Summary. Addressing major global challenges typically demands greater effort than a single research group or country can provide. Instead, it is important to make effective use of limited resources through collaboration, particularly when sharing information. Large-scale data sharing has often been hindered by limited discovery and access, lack of standardization, and IP and attribution concerns. Information security is also a pervasive concern. New cloud-oriented technologies are helping to break down these barriers, leading to new opportunities for integrating large, disparate, and diverse collections of information. Machine learning/artificial intelligence can be leveraged to identify patterns and perform analysis, however there are also new prospects for using these new capabilities for data management, including auto-generating more complete and consistent metadata.

Keywords. FAIR Data, Data Management, Distributed Data, Cloud, Machine Learning.

Modernisation of the AuScope Virtual Research Environment - Design and Lessons Learnt

Jens Klump^{1*}, **Vincent Fazio**¹, **Alex Hunt**¹,
Pavel Golodoniuc¹, **Ulrich Engelke**¹

^{1*} *Mineral Resources, Commonwealth Scientific and Industrial Research Organization*

Email: jens.klump@csiro.au

Summary. AuScope provides research infrastructure to Australia's Earth and geospatial science community. Its strategic objectives include developing the Downward Looking Telescope (DLT), symbolising an integrated system for Earth Science research. The AuScope Virtual Research Environment (AVRE) program supports the DLT by ensuring continuous access to necessary data and tools from various sources. AVRE evolved from a portal introduced in 2008 offering access to advanced data products, and, over time, Virtual Laboratories were added. As requirements and technologies evolved, the portal did, too. By 2021, AVRE recognised the need for modernisation to enhance user interaction, storage capacity, computational processing options, notebooks, containerisation, API access, and AI techniques. AVRE aims to accommodate diverse and evolving research needs and enable exchange with other data facilities. In this presentation, we will discuss our consultations with AVRE users and the co-design process leading to a re-engineered AVRE platform to meet the diverse needs of AVRE users.

Keywords. Virtual Research Environment, FAIR Data, Open Access, User Experience Design.

DataON: platform for discovery and analysis of research data

Jung-Ho Um^{1*}, **YoungHo Shin**¹, **Mi-Nyeong Hwang**¹, **SangBae Park**¹,
SeungBock Lee¹, **Na-eun Han**¹, **Hea Lim Rhee**¹, **Hyung-Jun Yim**¹

^{1*} *Research Data Sharing Center, Korea Institute of Science and Technology Information*
Email: jhum@kisti.re.kr

Summary. Recently, efforts to solve global social problems such as COVID-19 and abnormal climate through global data sharing are increasing. It means that research based on open science and the sharing of research data are required. Therefore, DataON has been developed to support facilitate data sharing, utilization and analysis, contributing to enhance scientific research. DataON is an integrated research data storage and analysis platform that provides a one-stop solution for collecting, publishing, searching, and analyzing research data. For introducing a real case of data usage, we would like to show DataON's ability to find research data by the institution through keyword or map searches, using case studies in the fields of weather/climate, sensors, and scientific images. As further research, we will continue to develop and expand the functionalities of DataON to support the resolution of global social and scientific problems.

Keywords. Research data, platform, open science, data analysis, open data.

Enhancing Social Science Research in East Asia: An Introduction to SRDA and NASSDA

Pei-Shan Liao^{1*}

^{1*} *Research Center for Humanities and Social Sciences, Academia Sinica*
psliao@gate.sinica.edu.tw

Summary. This presentation highlights the vital role of the Survey Research Data Archive (SRDA) and the Network of Asian Social Science Data Archives (NASSDA) in advancing social science research in East Asia. Established in 1994, SRDA systematically acquires, organizes, preserves, and disseminates academic survey data for researchers. In addition, SRDA explored researchers' attitudes towards openly sharing research data and identified motivating factors for Taiwanese researchers to engage in open data sharing. We strive to encourage researchers to embrace open data practices. Furthermore, NASSDA, founded in 2016, promotes collaboration and data sharing among social science data archives in Asia. Currently, members of NASSDA include SRDA, SSJDA, KOSSA, and CSDA. This collaboration empowers scholars with diverse datasets for cross-disciplinary research while reinforcing regional research opportunities, international cooperation, and academic partnerships.

Keywords. Survey Research Data Archive, Network of Asian Social Science Data Archives, social science research, open data, data management.

BigFlow: The End-to-End Workflow Framework for Cross-Center Scientific Data Analysis

XiaoJie Zhu^{1*}, HuaJin Wang¹, ZhiHong Shen¹

^{1*} *Computer Network Information Center, Chinese Academy of Sciences*
Email: xjzhu@cnic.cn

Summary. The rapid development of big data and artificial intelligence technologies has led to the transformation of research paradigms. The convergence research paradigm emphasizes collaborative analysis of scientific data from different domains to achieve scientific discoveries, where analysis task types are diverse and analysis processes span across different scientific data centers. Existing process-based analysis frameworks are difficult to support end-to-end cross-center scientific data analysis requirements due to the lack of analysis process expression capabilities, heterogeneous computing framework integration capabilities, and cross-center job scheduling capabilities. This paper proposes a software framework for performing end-to-end cross-center analysis of scientific data, supporting the construction of heterogeneous workflows across centers, transparent data transfer across frameworks, optimized scheduling of cross-center jobs, and data ownership and traceability. This framework effectively addresses the challenges of cross-center scientific data analysis, providing researchers in different scientific fields with powerful tools for comprehensive and collaborative analysis. Finally, we validate the advancement and feasibility of this framework by combining it with real-world scenarios.

Keywords. scientific research paradigm, analysis workflow, scientific data center, cross-center computing.

Polar Data Discovery and Mobilization Pathways

Chantelle Verhey^{1*}

^{1*} *World Data System - International Technology Office*
Email: cverhey@oceannetworks.ca

Summary. Enhancing data discovery and reuse continues to be at the forefront of discussions in the polar research community. Over the past 3 years, the POLDER working group has worked on enhancing the interoperability of repository holdings. The Polar Federated Search (PFS) (<https://www.search.polder.info>) is now available and compiles metadata from 23 repositories in a single user interface by indexing repository landing pages that have been enriched with the addition of semantic markup. This process marries the best practices from the web publishing world with domain specific metadata. Benefits of this technique include broader discovery, notably by Google Dataset Search, and better reuse of data repository holdings. We have advanced the ability of researchers to find polar data, using open source software components and a well described ontology that is built on top of commonly used metadata terms found across the web that benefits data managers and researchers alike. This presentation will highlight the developments of the PFS, various avenues for repositories to be indexed, semantic technologies and how the PFS fits in the international federated search landscape.

Keywords. Semantics, Polar, Federated Search, Data, Discovery.

Advancing Data Stewardship: Join the World Data System and Contribute to Scientific Data Excellence

**Vincent Bernabeo^{1*}, Meredith Goins¹,
World Data System International Program Office staff**

^{1*} World Data System International Program Office (WDS-IPO)
Email: vbernabe@utk.edu

Summary. Discover the benefits of becoming a member of the World Data System (WDS) and contribute to advancing global data stewardship. The WDS is a leading international organization that promotes trustworthy data services, open science, and data sharing principles. As a WDS member, you gain increased visibility, trustworthiness, and performance within the international data community. Showcase your organization's commitment to open science, enhance data discovery and exchange, and facilitate data citation. Join us in this poster presentation to learn about the membership process, different membership categories, and the impactful role you can play in shaping the future of data stewardship.

Keywords. World Data System, Data Stewardship, Trustworthy Data Services, Open Science, Data Sharing Principles.

Building the Web Site of the WDS Asia-Oceania Network

Takashi Watanabe^{1*}

^{1*}*National Institute of Information and Communications Technology*

Email: watanabeggl@gmail.com

Summary. The World Data System (WDS) is the interdisciplinary body of the International Science Council, whose mission is to support the ISC's vision by promoting long-term stewardship of, and universal and equitable access to, quality-assured scientific data and data services. To establish a regional network of WDS-oriented activities in the Asia-Oceania (A&O) area, this WDS Asia-Oceania (A&O) Network Website is being under preparation by an imprecation group of the Japanese WDS National Committee of the Science Council of Japan. Principal contents of the site are information on current WDS Members and regional WDS-oriented activities in the area. We would welcome to receive your comments and suggestions on this website.

URL: <https://takashiwatanabe.wixsite.com/wds-asia-oceania>

Keywords. Research Data, Data Network, Asia-Oceania.

Data and metadata sharing among Asian countries

Masaki Kanao ^{1*}

^{1*} *Joint Support-Center for Data Science Research,
Research Organization of Information and Systems
Email: kanao@nipr.ac.jp*

Summary. The Polar Environmental Data Science Center (PEDSC) of the Joint Support-Center for Data Science Research (DS), the Research Organization of Information and Systems (ROIS) has a responsibility to manage and publish the data involving Japanese research activities as one of a National Antarctic Data Center (NADC). At the International Polar Year (IPY2007-2008), a significant number of multi-disciplinary data have been compiled. These collected data/metadata have a tight collaboration with the Global Change Master Directory (GCMD), the Polar Information Commons (PIC), as well as several data centers belonging to the World Data System (WDS). In terms of data activities in polar communities of the Scientific Committee on Antarctic Research (SCAR) and the International Arctic Science Committee (IASC), tighter linkages of data/metadata sharing within the Asian Forum for Polar Sciences (AFoPS) countries has been discussed and should be further promoted by the involved Asian countries, in particular China, India, South Korea, Malaysia and Japan.

Keywords. Polar Environmental Data Science Center, data sharing, polar communities, AFoPS.

Concept of NIES Environmental Research Hub: A Data Infrastructure to support/promote Environmental Research

**Tomoko Shirai^{1*}, Hisashi Yashiro¹, Isao Murakami¹,
Ariya Fujita¹, Yoko Fukuda¹**

^{1*}National Institute for Environmental Studies
Email: tshirai@nies.go.jp

Summary. NIES Environmental Research Hub is a next-generation data infrastructure needed to support and promote environmental research and the use of environmental data under situations where data volume is increasing, data distribution becomes more active, and data-driven research progresses at an accelerating pace. The hub consists of a research platform and an integrated database. The research platform is a large-scale computation and analysis infrastructure that enables climate/environmental simulations/analysis, with high-level data security. The integrated database serves as a repository that disseminates environmental research data widely in accordance with the needs of the public, government, and research communities. It can host multiple data providers in the institute and promote open-data sciences in related fields by collaborating with other repositories. Not only hardware, but the development of human resources and software is also included in this concept. The hub contributes to making environmental research data FAIR (Findable, Accessible, Interoperable, and Reusable).

Keywords. Environmental data, Research Data infrastructure, Simulation platform, Data Repository, FAIR.

You Can't Make Science Open Without Committing to PIDs

Chieh-Chih Estelle Cheng^{1*}, Rorie Edmunds²

^{1*} ORCID

² Data Cite

Email: ludy001@brin.go.id

Summary. Data Cite and ORCID are global communities with a shared mission to ensure that the outputs of research can be discovered, accessed, and repurposed, and that the people who contribute to such outputs are correctly acknowledged for their work. They strive to build transparent and trustworthy linkages among researchers, resources, and organizations. Data Cite and ORCID are thus helping to advance knowledge by connecting research and researchers.

Both communities enable the creation and application of globally unique and persistent identifiers (PIDs), as well as integrating services to improve research workflows. Digital Object Identifiers (DOIs) registered under Data Cite services can be applied to many types of research outputs coming from any discipline. ORCID iDs enable researchers to connect themselves to their contributions and affiliations as they engage in research, scholarship, and innovation activities.

PIDs, and the metadata that underpin them, are vital components of making Open Science and the FAIR Data Principles a reality. When registering a DOI or ORCID iD, one supplies a PID metadata record under a CC0 licence, as well as more in-depth, structural information about the output or person—via a permanently resolvable URI—in a findable and accessible digital location (i.e., a landing page or an ORCID record). The provision of such rich metadata enhances the interoperability and reusability of an output associated with a uniquely identified individual, as well as facilitating crosslinking and citation.

This presentation will update the community on strategic initiatives and services led and offered by Data Cite and ORCID that assist the Openness and FAIRness of entities in the research ecosystem by increasing trust, diversity, and attribution and credit.

Keywords. ORCID, Data Cite, Persistent Identifier (PID), Metadata, Communities.

Open Science and FAIR Principles in Indonesia

Hermin Triasih¹, Sjaeful Afandi¹, Seno Yudhanto^{1*}, Madiareni Sulaiman²

¹National Research and Innovation Agency (BRIN) of the Republic of Indonesia

²University College London, UK

Email: seno.yudhanto@brin.go.id

Summary. Indonesia has implemented a range of strategic initiatives to promote the open science (OS) movement, particularly within the framework of BRIN Indonesia. This organisation has played a crucial role in (1) establishing the necessary infrastructure, (2) conducting outreach and advocacy activities, and (3) formulating regulations and guidelines at the national level. BRIN has developed essential infrastructure, specifically the National Scientific Repository (RIN), which serves as a repository for data and scholarly works. RIN facilitates the effective management of primary research data and output, employing the FAIR principle to ensure that research data and output generated at BRIN can be efficiently managed and utilised in the future. The support includes encompassing storage capabilities as well as integration and collaboration with diverse research support systems, including those related to funding, ethics, and laboratory facilities. Hence, implementing the OS movement at BRIN poses significant challenges, as the provision of infrastructure alone is insufficient to effectively raise researchers' awareness regarding the significance of research data management (RDM) and OS. BRIN undertakes extensive outreach and advocacy efforts to enhance researchers' understanding of the advantages associated with RDM and OS through seminars, trainings, and podcasts. Numerous individuals with a shared interest also convene in OS forums to effectively identify advocates and champions as well as foster a climate of transparency and knowledge. Therefore, some works on OS have been delivered through various regulatory and guideline documents, such as: RIN Policy, RIN User Manual, Data Curation, and Ethics Clearance in order to provide guidance to stakeholders and communities in Indonesia, facilitating their understanding and adoption of OS practices and FAIR principles.

Keywords. Open Science, FAIR Principles, Repositori Ilmiah Nasional, Indonesia.

IRINS: Building FAIR Research Information System in India

Kannan Palavesam^{1*}

^{1*} *Information and Library Network (INFLIBNET) Centre, Gandhinagar, Gujarat, India*

Email: kannan@inflibnet.ac.in

Summary. According to the All-India Survey of Higher Education 2020-2021, the Indian higher education system comprises 1113-degree awarding institutes, 43796 colleges and 11296 stand-alone institutes and a diverse set of stakeholders involved in the massive higher education system, including funding agencies/policymakers, academic institutes, scholarly community. Unfortunately, the existing digital ecosystem cannot interconnect with all the stakeholders to retrieve scholarly communication activities. The INFLIBNET Centre developed the project called Indian Research Information Network System (IRINS) to interconnect the scholarly community and their scholarly activities through academic identities. The scholarly data sets collected through this project bring more value and are discoverable and widely used by academicians, funding agencies and policymakers in various aspects. There are 1000 + institutes, including academic and R&D organizations, that have been using the IRINS and 158 thousand faculty members connected through this project. More than 2.4 million publication metadata were retrieved from various sources using academic identities. This presentation discusses various challenges to developing and executing the IRINS projects for higher education institutions in India. The paper also discusses applying four foundational principles such as Findable, Accessible, Interoperable and Reusable in the IRINS Project and its implications.

Keywords. Research Information Management System; RIM, CRIS, IRINS.

An Advanced identifier architecture for enabling FAIR principle: Global Open Identifier

Jia Liu^{1*}, Xuezhi Wang¹, Zhihong Shen¹, Yuanchun Zhou¹

¹ *Computer Network Information Center, Chinese Academy of Sciences*
Email: liujia@cnic.cn

Summary. The identifier is one of the essential parts of the "Open science infrastructures" and aims to make science and technology resources been found and interworked easily through managing a unique digital identity for each. Extensive works conducted to promote the realization FAIR Principle and create enormous benefits to the whole academia, such as DOI, ORCID and CSTR. Most existing identifiers mainly works on development in terms of its own scenarios. The new challenges have been raised as the barriers among different identifiers bring significant work with high repetitive. Therefore, we propose a novel architecture from a global perspective named "Global open identifier infrastructure", adaptively active the interoperability between various identifiers across mechanism, management policy, technology and information systems. The specific analysis of the challenges as well as the description and demonstration of the global open identifier will be mainly included in the presentation.

Keywords. Global open Identifier, interoperability, FAIR principle, identifier architecture.

Open Science Policy Development in Japan: importance of an intermediate sector balancing between Top-down and Bottom-up

Kazuhiro Hayashi^{1*}

^{1*} *National Institute of Science and Technology Policy (NISTEP)*

Email: khayashi@nistep.go.jp

Summary. Open Science is a complex issue to transform Science, Technology and Society. It includes some transformation of organization or sector, represented by Libraries. So that, it is important that instead of local optimization in each sector, we need to try to optimize the whole towards Open Science Paradigm. NISTEP, as a national institution for ST Policy making, has been contributed to Open Science Policy from various points of view by balancing between Top-down and Bottom-up. From Top-down approach, NISTEP has contributed to Open Science policy development for both of international (UNESCO, G7, OECD) and domestic (Cabinet Office, MEXT, and some funding agencies). From Bottom-up approach, NISTEP has committed to some new initiatives such as Japan Open Science Summit, Research Data Utilization Forum (RDUF) and RDA, as well as committed to some established scholarly activities such as Science Council of Japan, World Data System, and CODATA. It is hard for a sector to transform drastically by itself. NISTEP is in good position between policy sector and academic sector, facilitating the transformation of those robust and established sectors with including mind-change. This kind of intermediate sectors and activities would be crucial towards Open Science paradigm.

Keywords. Open Science, Open Access, Research Data, Sharing, Policy Development, FAIR principle.

Making FAIR Scale

Erik Schultes^{1*}, Barbara Magagna¹

^{1*} GO FAIR Foundation & Leiden Academic Center for Drug Research
Email: eriks@gofair.foundation

Summary. Making FAIR data is an intricate socio-technical challenge that involves the orchestration of a range of human expertise and a large pallet of potential information technologies. Currently, many approaches to FAIRification simplify the task by focusing on specific use cases. Although this engenders practical outcomes, it also leads to implementations that are often over-specified on particular domains. This ‘creolisation’ of applications will likely fail to achieve broadly generalised interoperability and risks further fragmentation of the data landscape. The GO FAIR Foundation has been exploring decision support structures that guide stakeholders in making well-informed implementation choices that are, as much as possible, technology and use case agnostic. The approach is referred to as the Three-Point FAIRification Framework (3PFF) as it defines three focal areas: 1) building overall FAIR Implementation Profiles that document essential technology choices; 2) creation of domain-relevant community-based FAIR metadata standards; 3) deployment of infrastructures that mediate FAIR operations. By encouraging the use of minimal open standards, the 3PFF accelerates the FAIRification process, lowers the cost of committing to FAIR and provides guarantees on cross-disciplinary interoperability, all of which contribute to the scalability of FAIR data and services.

Keywords. GO FAIR Foundation, Three-Point FAIRification Framework, FAIR Implementation Profile, Metadata for Machines Workshop.

Survey for policy of research data openness in South Korean government-funded research institutes

Jung-Ho Um^{1*}, **YoungHo Shin**¹, **Mi-Nyeong Hwang**¹, **SangBae Park**¹,
SeungBock Lee¹, **Na-eun Han**¹, **Hea Lim Rhee**¹, **Hyung-Jun Yim**¹

^{1*} *Research Data Sharing Center, Korea Institute of Science and Technology Information*
Email: jhum@kisti.re.kr

Summary. Since 2018, the Ministry of Science and ICT in S. Korea has created a national research data sharing and utilization strategy to promote a policy of collecting research data for projects that produce data among R&D projects performed in S. Korea. After the policy was created, a survey was conducted targeting 24 government-funded research institutes under the National Research Council for Science and Technology to recognize the current status of government-funded research institutes' awareness of research data sharing. This survey consists of questions such as awareness, organization, and systems for research data sharing, and reasons why disclosure is difficult. 19 organizations responded to the survey. About 14% of research institutes have an organization for research data, and about 53% of institutions have a research data management system, and security issues were chosen as the first reason for difficulty in disclosing. Currently, a legislation related to research data is being prepared by the Ministry of Science and ICT, and a consultative body for government-funded research institutes has been established to continue discussions on research data. Through the survey conducted earlier, we plan to discuss together the difficulties of disclosing research data and seek ways to overcome them.

Keywords. Open science, data policy, survey, research data openness, research data management.

Challenges of Open Research Data in Indonesia: Case Study National Scientific Repository

Seno Yudhanto^{1*}, **Ludya Arica Bakti**¹, **Ria Ariani**¹

^{1*} *National Research and Innovation Agency*
Email: ludy001@brin.go.id

Summary. The openness of research data is vital and required in research so that it may be reproduced and used in a sustainable way. Moreover, the availability of research data can also prevent the double funding for research. However, Indonesia' condition has not fully supported that concept. Therefore, the purpose of this study is to identify challenges associated with open research data in Indonesia, particularly in the National Scientific Repository (RIN). This study uses a qualitative approach with a case study method. Data was gathered through semi-structured interviews and literature reviews. Meanwhile, coding was used to analyze the data and the results using verbatim. This study is expected to be able to identify challenges and solutions to the open research data implementation in Indonesia, through RIN utilization, so that this study has implications for promoting the open research data. Besides, RIN also can be a reliable system for managing open research data.

Keywords. open data, research data, data repository, Repositori Ilmiah Nasional (RIN), National Scientific.

MosquitoDB: A comprehensive-electronically-based entomological surveillance system for the control and elimination of mosquito-borne diseases

Victor Mero^{1,4*}, Dickson Msaky¹, James Michael¹, Henry Lyamuya¹, Gerald Kiwelu¹, Janice Maige^{1,2}, Njire Choba¹, Prosper P. Chaki^{1,4}, Silas Majambere¹, Nicodemus J. Govella¹, Samson Kiware^{1,3,4}

^{1*} Ifakara Health Institute, Environmental Health, and Ecological Sciences Department, Tanzania

² University of Dar es Salaam, College of Information & Communication Technology, Tanzania

³ The Nelson Mandela, African Institution of Science and Technology, The School of Life Science and Bio-engineering (LISBE), Tanzania

⁴ The Pan-African Mosquito Control Association (PAMCA), KEMRI Headquarters, Kenya
Email: jmaige@ihi.or.tz

Summary. Most National Disease Control/Elimination Program and researchers lack a robust entomological surveillance system that can manage field and laboratory-based mosquito data leading to the inability to make timely informed decisions on the deployment of vector control tools. We have developed a mosquito database management system (MosquitoDB) that can manage diverse entomological studies to control and eliminate vector-borne diseases. The system is freely and securely accessible online with the data collection application available in the Google Play Store – capable of validating and recording geolocation mosquito data even in the absence of an internet connection. The key functionalities include but are not limited to customization of variables by forms, proper linkage of field and laboratory data, data sharing capabilities in standardized formats, multi-language support, and access to linked datasets with summarized reports in different formats. Also, an interactive dashboard to support making informed decisions on where, when, how, and which vector control tool(s) should be implemented. MosquitoDB can easily be linked to other databases (e.g., epidemiological) with the ability to push key entomological indicators to DHIS2. Current users include researchers and some national malaria control programs focusing on malaria vectors. MosquitoDB is an effective and comprehensive electronic-based entomological surveillance system that can support national diseases control with entomological routine surveillance and researchers with complex diverse entomological studies.

Keywords. Mosquito, malaria, entomological, vector borne diseases, database, surveillance.

Further improve accessibility of JAMSTEC research data
- DOI assignments to cruise reports –

Ayumi Sumikawa^{1*}, Kazuyo Fukuda¹, Miwa Hishiki¹

^{1*} Japan Agency for Marine-Earth Science and Technology
Email: ayumi_sumikawa@jamstec.go.jp

Summary. Japan Agency for Marine-Earth Science and Technology (JAMSTEC) has assigned Digital Object Identifiers (DOIs) to research cruises since 2017 as an attempt for open data, and the number has reached 2500 as of September 2023. In addition, we have been publishing cruise reports as preliminary results of its research cruises. A total of 1,227 reports on research cruises undertaken from 1992 to 2023 have been published and accessible in the “Cruise reports & DATA BOOK Catalog (https://www.godac.jamstec.go.jp/cr_catalog/)”. In this poster, we will introduce 1) citation tracking results of cruise reports without DOIs and the issues, 2) Construction of the Data DOI Information Management System for assigning DOIs and publishing landing pages and 3) efforts to assign DOIs to cruise reports. Consequently, JAMSTEC has assigned DOIs to 99% of all cruise reports. These efficient efforts are expected to promote more accurate report citation and further utilization of data.

Keywords. cruise report, citation tracking, DOIs, research cruise, marine science.

Scientometric Analysis and Mapping of Research Publications of Central University of Gujarat

Kotrayya Agadi^{1*}

^{1*} Central University of Gujarat, Gandhinagar, India

Email: kbagadi@cug.ac.in

Summary. The study aims to examine the research and scholarly output produced by the Central University of Gujarat (CUG) between 2010 and 2022. The bibliographic data for this study was extracted from the Dimensions database. This paper analyses 1921 research papers published by the Central University of Gujarat during the period under study. The study examines the year-wise research publications, authorship pattern, degree of collaboration, types of documents, preferred journals for publication, prolific authors, h-index, most collaborating institutions and countries, etc. Additionally, in this research paper, different software was used, e.g., 'MS-Excel', 'bibliometrix', and 'biblioshiny' of R- Package, and VOSviewer was applied for the analysis and data visualisations. The research findings indicate that the growth pattern of literature is linear, and journal articles are the most preferred form of publication by researchers to communicate their research. Degree of Collaboration (DC) indicates that the majority of research work is jointly carried out by research scholars at the Central University of Gujarat (CUG). The research method used in this study was scientometric analysis. The main research area at the Central University of Gujarat was Chemistry. The most preferred journal for publication was the Journal of Molecular Liquids, and the most prolific author was Prof. Man Singh. The present study revealed several scientometric indicators like citation pattern, collaborative institutions, most prolific authors, etc.

Only one paper was two-authored and received the highest number of citations (439), while the other nine research papers were shown to be multiple-authored. Collaborating Institutions like the USA, Australia, South Africa, the UK, and Canada were the top five most collaborative countries. The subject of chemistry was the most prominent research area during the study period. The Central University of Gujarat has to make more effort to promote research, develop quality research, and develop better policies to enhance and enrich research performance.

Keywords. Central University of Gujarat (CUG), Scientometric, Bibliometric, Science Indicators, Mapping of Scholarly publications.

Indian Language Open Educational Resources for the Implementation of the National Education Policy 2020: A Comprehensive Evaluation

Kotrayya Agadi^{1*}, Sagender Singh Parmar¹

^{1*} *Central University of Gujarat, Gandhinagar, India*
Email: kbagadi@cug.ac.in

Summary. The New Education Policy (NEP) 2020 has emphasised two main aspects: first, education should be accessible to all children, and second, multilingualism and instruction in the mother tongue. OERs have the potential to transform education by facilitating access to a variety of learning resources. NEP 2020's emphasis on the use of indigenous languages for instruction, the availability of OERs in Indian languages becomes of the utmost importance. The study examines the readiness of OERs in Indian languages for the effective implementation of NEP 2020, including an evaluation of Indian Language OER websites using the CRAAP (Currency, Relevance, Authority, Accuracy, and Purpose) test. The findings will shed light on the status of existing OERs in Indian languages, identifying improvement areas and potential bottlenecks. Implementing NEP 2020 requires a comprehensive repository of OERs in Indian languages. The study emphasises the need to assess the preparedness of such resources using the CRAAP test, ultimately contributing to the improvement of education quality and accessibility in accordance with the vision of the policy.

Keywords. Open Education Resources, National Education Policy, Indian Languages.

Reviewing COVID-19 Management in Practice by a comparative study of India, Taiwan, USA and Japan through the Perspective of TQM and Preparing for Future Pandemics

Kazuyuki Suzuki^{1*}

^{1*} *University of Electro-Communications, Tokyo*
Email: suzuki@uec.ac.jp

Summary. When comparing the number of deaths, the number of cases and the fatality rate per million people per day in India, Taiwan, USA, and Japan concerning each of the five epidemic virus strain types such as Wuhan strain, α strain, δ strain, Omicron BA1+BA2, and Omicron BA5, the major differences between the four countries are revealed as numerical values along with changes in virus strains. How to analyse the factors that give this difference? What lessons should be learned? How to prepare for the next pandemic? For example, the fatality rates for the USA and India have similar trends for each virus strain, with a high rate of 1.4-1.9% in the early stage and a low rate of 0.4-0.6% in Omicron. On the other hand, the fatality rate for Taiwan and Japan were suppressed to 1/3 (0.15 to 0.18%) of those of the United States and India in Omicron. However, Japan and Taiwan have sharp increases in death tolls in Omicron. Focusing on and analysing the factors behind the difference such as PPE preparation/ vaccination/ demographics and deaths by age group/ mobility restrictions/ mask wearing/ popularization of online communication/ ventilation/ hand sanitizer/ the use of digital technology, this presentation proposes the measures for future virus epidemics using the perspective of TQM (Total Quality Management) created by Japan.

Keywords. fatality rate for each virus strain type, vaccination rate, mobility, mask wearing rate, demographics and deaths by age group.

Towards a global and integrated wastewater surveillance system: a review of historical and latest activities

Noriko Endo^{1*}

^{1*} *Research Center for Environmental Quality Management, Department of Environmental Engineering, Graduate School of Engineering, Kyoto University*
Email: endo.noriko.3p@kyoto-u.ac.jp

Summary. Wastewater surveillance has emerged as an effective tool for monitoring community spread of Covid-19 and other diseases during the Covid-19 pandemic. The history of development and implementation of wastewater surveillance has witnessed significant international and interdisciplinary cooperation since 2020. Researchers around the world developed methods quickly and shared know-how extensively via an *ad hoc* Slack group or online webinars. Governments, academia, and industry worked together to implement and expand wastewater surveillance in many countries. There are even cross-governmental efforts to share the surveillance data. However, many challenges remain to achieve more global and integrated wastewater surveillance: expansion of detection targets, data integration with other public health data, and data ownership agreement to allow international data sharing.

Keywords. Wastewater surveillance, international cooperation, data integration, data sharing.

Wastewater Surveillance Data: Tracking Imported SARS-CoV-2 Variants

Jatuwat Sangsanont^{1,2*}, **Chaw su Thin**³, **Kwanrawee Sirikanjana**⁴

^{1*} *Department of Environmental Science, Faculty of Science, Chulalongkorn University, Thailand*

² *Water Science and Technology for Sustainable Environmental Research Group, Chulalongkorn University, Thailand*

³ *International Program in Hazardous Substance and Environmental Management, Graduate School, Chulalongkorn University, Thailand*

⁴ *Research Laboratory of Biotechnology, Chulabhorn Research Institute, Thailand*

Email: Jatuwat.s@chula.ac.th

Summary. Wastewater-based epidemiology (WBE) has the potential to bolster disease surveillance efforts. The central inquiry revolves around the utilization of WBE. This study examined the ability of WBE in tracking the emergence and spread of SARS-CoV-2 variants in Thailand. Samples were collected from domestic and airport wastewater treatment plants in the major cities of Bangkok and Phuket. The WBE data revealed the replacement of older variants by newer ones, transitioning from Delta to BA.1 and BA.2 during the COVID-19 pandemic. This data aligns with the extensive clinical surveillance conducted within the country and pinpoints Phuket as the initial point of introduction for the new variants. Furthermore, by monitoring variants at the airport, evidence of new variants being imported from other countries was observed. The findings underscore the benefits of WBE, showcasing its ability to provide early warning about the emergence of new variants and track the shifts in SARS-CoV-2 sub-variants.

Keywords. wastewater surveillance, SARS-CoV-2, variants.

Harnessing Advanced Data Science and Machine Learning in Smart Wastewater-Based Epidemiology

Guangming Jiang^{1*}

^{1*} *School of Civil, Mining, Environmental & Architectural Engineering, University of Wollongong, Australia*
Email: gjiang@uow.edu.au

Summary. The conventional wastewater-based epidemiology (WBE) back-calculation is usually impractical due to the dynamic viral shedding, unknown decay/partitioning in sewers, and a single output with high uncertainty. Many wastewater-based epidemiology studies struggled to obtain accurate and reliable back-calculation. Our recent studies on COVID-19 explored the potential of machine learning models by integrating wastewater-based epidemiology data with sewer catchment, weather, clinical testing, and vaccination data. The developed artificial neural network model can provide accurate prediction of COVID incidence, prevalence, and effective reproduction rate in Utah and Wisconsin, USA. This demonstrated the huge potential of advanced data analytics by harnessing the power of artificial intelligence and big data. The wastewater-based epidemiology data can inject into various actionable strategies like nonpharmaceutical interventions, instead of just providing infection case numbers. The smart wastewater-based epidemiology can lead to novel data-driven and evidence-based pandemic management that will minimize viral transmission and societal disruption during an outbreak.

Keywords. COVID-19, wastewater-based epidemiology, SARS-CoV-2, artificial neural network, prevalence.

Airborne surveillance of bioaerosol in built environment for future pandemic preparedness: Beyond COVID-19

Irvan Luhung^{1*,2}, **Stephan C. Schuster**¹

^{1*} *Singapore Centre for Environmental Life Sciences Engineering, Nanyang Technological University*

² *School of Chemical and Environmental Engineering, Yale University*

Email: irvan.luhung@yale.edu

Summary. Environmental monitoring of biological aerosols is often challenging due to the ultra-low biomass nature of the air samples, as well as the temporal and spatial variability of the air microbiome. In the case of SARS-CoV-2, direct air sampling had emerged as one of the most utilized surveillance approaches in addition to wastewater monitoring. There was, however, relatively limited effort in understanding how different factors from the chosen methodological pipeline, e.g., air sampling (flowrate, placement of samplers, sampling media), sample processing (extraction techniques) and analysis (qPCR, ddPCR, sequencing) could affect the interpretations of the results. This study highlights important indoor environmental conditions and the nucleic acid analysis approaches which had profound impacts to our environmental air surveillance of SARS-CoV-2. Beyond the current COVID-19 pandemic, it is also important to comprehend the dynamics between indoor/outdoor bioaerosols and building ventilation system operation in shaping bioaerosol monitoring standards for future pandemic preparedness.

Keywords. Bioaerosols, Air sampling parameters, metagenomics and metatranscriptomics, qPCR and ddPCR, indoor environmental quality.

Issues of Digital Sequence Information (DSI) emerged from the COVID-19 pandemic

Masanori Arita^{1*}

^{1*} *Bioinformatics and DDBJ Center, National Institute of Genetics*
Email: arita@nig.ac.jp

Summary. The issue of benefit sharing arising from digital sequence information (DSI) has been a subject in various international forums, including the Convention on Biological Diversity (CBD), the treaty on marine areas beyond national jurisdiction (BBNJ), the World Health Organization (WHO) Pandemic Treaty, and the international Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA). Each of these forums contributes to the discourse with distinct policies and strategies. Currently, each forum tends to establish specialized data repositories, exemplified by the GISAID database in the context of COVID-19. While commendable for its global coverage and data integrity, the restrictions on access and usage have sparked controversy.

One prominent apprehension pertains to the sustainability of such specialized repositories, given that COVID-19 is not an isolated pandemic event. The repository's adaptability to future pandemics needs verification and another critical issue is the controversy concerning quality control and access restriction. In this talk, I introduce the practical challenges associated with managing multiple repositories, particularly from an informatics perspective. Emphasis will be placed on a FAIR (Findable, Accessible, Interoperable, and Reusable) and harmonized system across diverse international fora to foster collaboration, and to ensure equitable benefit sharing across the globe.

Keywords. Data repository, Convention on Biological Diversity (CBD), Digital Sequence Information (DSI), Global Initiative on Sharing Avian Influenza Data (GISAID).

Effects of information on behavioral change during Covid-19 pandemic: An International comparative study between Japan, UK, and Taiwan

Naoko Kato-Nitta^{1*}, Yusuke Inagaki², Tadahiko Maeda²

^{1*} *J.F. Oberlin University, ROIS-DS and The Institute of Statistical Mathematics*

² *ROIS-DS and The Institute of Statistical Mathematics*

Email: kato_n02@obirin.ab.jp

Summary. In the midst of the Covid-19 pandemic, the government and experts frequently disseminated information on the number of new cases per day, the level of bed congestion, and the rate of vaccination. Citizens have decided whether or not to take action to prevent the spread of infection, taking into account such information and their own circumstances (e.g., whether they live with elderly people or alone, whether they have a job that allows remote work, whether they have already been vaccinated, etc.). Although many experts are interested in how these information and circumstances influence people's behavior change, there are countless combinations of information and circumstances that are difficult to demonstrate. In this study, we attempted an approximate solution by extracting factors that are considered important and applying a factorial experimental survey (conjoint method) based on the L18 orthogonal design. This presentation will outline the results of a statistical analysis of international comparative survey data in Japan, the U.K., and Taiwan, conducted from February to March 2022 using the same survey design.

Keywords. Covid-19, Behavioral Change, Social Survey, Science Literacy, Vaccine Hesitancy.

Genomic Surveillance of SARS-CoV-2 based on comparison among Japanese cities

Ryoya Kawai¹, **Koshi Nakano**¹, **Takashi Abe**^{1,2*}

^{1*} *Information Engineering, Niigata University*

² *Bioinformation and DDBJ Center, National Institute of Genetic*

Email: takaabe@ie.niigata-u.ac.jp

Summary. The SARS-CoV-2 has spread rampantly worldwide and pandemic variants continue to emerge. Genomic surveillance is being conducted at worldwide to understand the details of this pandemic. It is important not only to investigate the pandemic variants worldwide, but also to investigate the epidemic variants in within a country. We investigated genomic surveillance using haplotype network and geophylogenetic analyses to understand the epidemic trends among Japanese cities. The delta variants that were predominant strain in Japan was AY.29 strain with mutations that first appeared in Japan. In addition, the omicron variants, regional epidemic variants that appeared for the first time in Japan, such as BA.2.3.1, BA.2.3.11, BF.22, and BQ.1.1.70, were detected. Differences in the trends of variants among Japanese cities are observed, and they also play an important role in understanding the epidemic trends in Japan.

Keywords. Genomic surveillance, Haplotype network analysis, Regional epidemic variants.

Unsupervised AI analysis of oligonucleotide usage in a wide range of variants of SARS-CoV-2

Toshimichi Ikemura^{1*}, **Yuki Iwasaki**¹, **Takashi Abe**²

^{1*} *Nagahama Institute of Bio-Science and Technology*

² *Niigata University*

Email: t_ikemura@nagahama-i-bio.ac.jp

Summary. We conducted time-series analysis of oligonucleotide usage in SARS-CoV-2 genomes, revealed distinctive mutations rapidly expanding their intrapopulation frequency, and developed a sequence alignment free method that extensively searches for advantageous mutations for efficient reproduction in humans (doi.org/10.1016/j.gene.2020.100038). We previously established an unsupervised AI, a BLSOM (batch-learning SOM) for oligonucleotide usage to obtain novel knowledge from genomic sequences (doi.org/10.1101/gr.634603), which can analyze five million genomic sequences simultaneously. When only oligonucleotide usages in a massive number of SARS-CoV-2 genomes were given, the obtained clusters corresponded primarily to known main clades and internal divisions in the main clades. Since the BLSOM is explainable-type AI, it reveals oligonucleotide usage patterns responsible for clade clustering (doi.org/10.1186/s12866-022-02484-3). To distinguish advantageous mutations from hitchhiking-type neutral mutations, we focus on mutations that appear to occur independently in different lineages and expand in frequency in a convergent evolutionary manner, by analyzing Omicron strains with BLSOM (doi.org/10.1371/journal.pone.0273860).

Keywords. Oligonucleotide usage, SARS-CoV-2, Unsupervised AI, Clade clustering.

Host cell-dependent genomic changes in SARS-CoV-2

Yuki Iwasaki^{1*}, **Takashi Abe**², **Toshimichi Ikemura**¹

^{1*} *Nagahama Institute of Bio-Science and Technology*

² *Niigata University*

Email: y_iwasaki@nagahama-i-bio.ac.jp

Summary. Various host factors are involved in viral reproduction, and host environment-specific adaptation and avoidance of host antiviral factors are important for their efficient reproduction. When a virus that has been endemic in a non-human population changes its host to humans, the virus should adapt to the new host environment for its efficient reproduction. Here, we investigated how the genome sequence of SARS-CoV-2, which is currently causing a pandemic worldwide, changes in the human cellular environment. Genomic sequences of coronaviruses isolated from humans and bats, including SARS-CoV-2, were obtained from NCBI, and the frequency of oligonucleotide in their genomic sequences were analyzed. Interestingly, the oligonucleotide compositions of SARS-CoV-2 changed towards those of coronaviruses that have been prevalent in humans and away from those of bat coronaviruses. Thus, the frequency of oligonucleotide in SARS-CoV-2 genomes have changed in a host-dependent manner (Iwasaki et al., BMC Microbiol 2021. <https://doi.org/10.1186/s12866-021-02158-6>).

Keywords. Oligonucleotide frequency, SARS-CoV-2, Viral adaptation, Zoonotic virus.

SARS-CoV-2 HaploGraph: visualization of SARS-CoV-2 haplotypes spread in Japan

**So Nakagawa^{1,2*}, Toshiaki Katayama³, Lihua Jin⁴, Jiaqi Wu², Kirill Kryukov¹,
Rise Oyachi², Junko S Takeuchi⁵, Takatomo Fujisawa¹, Satomi Asano¹,
Momoka Komatsu⁶, Jun-ichi Onami⁷, Takashi Abe⁶, Masanori Arita¹**

^{1*} *Bioinformation and DDBJ Center, National Institute of Genetics*

² *Department of Molecular Life Science, Tokai University School of Medicine*

³ *Database Center for Life Science*

⁴ *Genomus Co., Ltd.*

⁵ *Center for Clinical Sciences, National Center for Global Health and Medicine*

⁶ *Smart Information Systems, Faculty of Engineering, Niigata University*

⁷ *Research Center for Open Science and Data Platform, National Institute of Informatics*

Email: so@tokai.ac.jp

Summary. Since the early phase of the coronavirus disease 2019 (COVID-19) pandemic, a number of research institutes have been sequencing and sharing high-quality severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) genomes to trace the route of infection in Japan. To provide insight into the spread of COVID-19, we developed a web platform named SARS-CoV-2 HaploGraph to visualize the emergence timing and geographical transmission of the SARS-CoV-2 haplotypes. Using data from the GISAID EpiCoV database as of June 4, 2022, we created a haplotype naming system by determining the ancestral haplotype for each wave and showed prefectural or region-specific haplotypes in each of the four epidemic waves in Japan. The SARS-CoV-2 HaploGraph allows for interactive tracking of virus evolution and geographical prevalence of haplotypes and aids in developing effective public health control strategies during the global pandemic. The code and the data used for this study are publicly available at: <https://github.com/ktym/covid19/>. A customized version of the SARS-CoV-2 HaploGraph is deployed at the COVID-19 Data Portal Japan: <https://covid19dataportal.jp/>.

Keywords. SARS-CoV-2 haplotype, Genomic surveillance, Data visualization, Open data.

Urban Microbiome Monitoring in Japan: Collaborative Research with MetaSUB Consortium

Haruo Suzuki^{1*}, **Dewa A.P. Rasmika Dewi**², **Yuh Shiwa**³

^{1*} *Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan & Faculty of Environment and Information Studies, Keio University, Fujisawa, Kanagawa, Japan*

² *School of Medicine, International University of Health and Welfare, Narita, Japan & Faculty of Medicine and Health Sciences, Udayana University, Bali, Indonesia*

³ *Department of Molecular Microbiology, Tokyo University of Agriculture, Tokyo, Japan & NODAI Genome Research Center, Tokyo University of Agriculture, Tokyo, Japan*

Email: haruo@sfc.keio.ac.jp

Summary. The MetaSUB (Metagenomics & Metadesign of Subways & Urban Biomes) International Consortium was founded in 2015. In this presentation, we provide an overview of our collaborative research with MetaSUB and highlight our ongoing efforts. We have been collecting samples from urban built environments, including transit systems such as subway stations and bus stops, as well as university campuses in multiple cities across Japan. This sampling effort commenced on June 21, 2016, during the Global City Sampling Day (gCSD), and has continued during the COVID-19 pandemic and throughout the Olympics and Paralympics in Tokyo. We have refined and standardized procedures for sample collection and sample processing (DNA/RNA extraction and shotgun sequencing). All sequence data and metadata related to the projects have been deposited in public databases to facilitate access by anyone interested in the resource. Our goal is to establish the ultimate monitoring infrastructure for infectious diseases through comprehensive sequence analysis of environmental samples.

Keywords. Urban microbiome, MetaSUB Consortium, metagenomics, metatranscriptomics, collaborative research.

Comparison of microbial diversity and resistant determinants in the built environment in Japan using shotgun metagenomic sequencing

Dewa A.P. Rasmika Dewi^{1*}, **Haruo Suzuki**², **Tetsuya Matsumoto**³

^{1*} *School of Medicine, International University of Health and Welfare, Narita, Chiba, Japan & Faculty of Medicine and Health Sciences, Udayana University, Bali, Indonesia*

² *Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan & Faculty of Environment and Information Studies, Keio University, Fujisawa, Kanagawa, Japan*

³ *School of Medicine, International University of Health and Welfare, Narita, Chiba, Japan & Graduate School of Medicine, International University of Health and Welfare, Narita, Chiba, Japan*
Email: d-dewi@iuhw.ac.jp

Summary. Urban microbiomes are specific and unique for each city. In recent years, the urbanization rate has been increasing, and this contributes to the dynamic of city microbes, which has influenced the health of its inhabitants. We investigated the diversity of urban microbiomes and resistant determinants in Japanese cities by performing a shotgun metagenome sequencing of various surface samples collected in the built environment during the COVID-19 pandemic, between 2021 and 2022. As part of the MetaSUB project, samples were collected in train stations, bus stops, and university campuses using standardized sampling kits provided by the MetaSUB International Consortium. The shotgun metagenomic sequencing was performed using DNBSEQ-G400RS, and after sequence reads quality checking, the relative abundance of microbial taxa and resistant genes in the metagenome were estimated using computational tools. Our study revealed the variation of microbial diversity and resistant determinants over the collection periods in different cities.

Keywords. Microbial diversity, resistant determinants, built environment, shotgun metagenome.

Evaluation of rRNA depletion methods for capturing the RNA virome from environmental surfaces

Yuh Shiwa^{1,2*}, **Tomoya Baba**^{3,4}, **Haruo Suzuki**^{5,6}

^{1*} *Department of Molecular Microbiology, Tokyo University of Agriculture, Tokyo, Japan*

² *NODAI Genome Research Center, Tokyo University of Agriculture, Tokyo, Japan*

³ *Advanced Genomics Center, National Institute of Genetics, Mishima, Japan*

⁴ *Joint Support-Center for Data Science Research, Research Organization of Information and Systems, Tokyo, Japan*

⁵ *Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan*

⁶ *Faculty of Environment and Information Studies, Keio University, Fujisawa, Japan*

Email: y3shiwa@nodai.ac.jp

Summary. There are several approaches aimed to detect RNA viruses such as SARS-CoV-2 from urban environments using metatranscriptomics, including the METACoV project by the MetaSUB International Consortium. Shotgun sequencing analysis of RNA viromes on built-environment surfaces is hampered by low targeted RNA yields and high abundance of ribosomal RNA (rRNA) from various cellular organisms, including eukaryotes (humans) and prokaryotes (bacteria). Therefore, we evaluated the quality of libraries, efficiency of rRNA depletion, and viral detection sensitivity using low RNA inputs (<5 ng) derived from a mock community or surfaces such as melamine-coated tables with a standard library preparation kit. Good-quality RNA libraries were obtained successfully from 0.1 ng of RNA by changing the adapter concentration and number of PCR cycles. Differences in the eukaryotes or prokaryotes of the rRNA depletion method affected the community composition and sensitivity of virus detection. We established methods of metatranscriptome analysis for RNA viromes from very small amounts of targeted RNA using a standard library preparation kit.

Keywords. rRNA depletion, virus detection, environmental metatranscriptome analysis, virome, library preparation.

Wastewater-Based Epidemiology for Efficient SARS-CoV-2 Surveillance in Three Japanese Cities

Ryo Honda^{1*}, **Yui Araki**¹, **Momoko Kaji**², **Md. Alamin**², **Satoshi Ezaki**³,
Ryotaro Okazaki³, **Akihiko Hata**⁴

^{1*} Faculty of Geosciences and Civil Engineering, Kanazawa University, Japan

² Graduate School of Natural Science and Technology, Kanazawa University, Japan

³ Kubota Corporation, Japan

⁴ Department of Environmental and Civil Engineering, Toyama Prefectural University, Japan

Email: rhonda@se.kanazawa-u.ac.jp

Summary. Wastewater-based epidemiology by monitoring of SARS-CoV-2 at wastewater treatment plants (WWTPs) is expected to enable efficient epidemic surveillance of a community. In this study, we compared the concentration of SARS-CoV-2 RNA in influent wastewater at WWTPs with clinically confirmed cases in three Japanese cities. The viral titers observed in wastewater surveillance exhibited a strong correlation with newly confirmed cases reported within one week after the last sampling. Notably, occasional peaks in viral titers suggested the presence of infected individuals among attendees of sporting events in the target city. These findings underscore the utility of wastewater-based epidemiology in comprehensively assessing the epidemic status of a community. As a result of these findings, Komatsu City in Japan has implemented a weekly notification system to inform its citizens of the epidemic status based on wastewater monitoring through a social network service.

Keywords. Public health, epidemic surveillance, wastewater treatment plant.

Implementation of wastewater-based epidemiology for viral infectious diseases in Japan

Masaaki Kitajima^{1*}

^{1*} *Division of Environmental Engineering, Hokkaido University, Japan*
Email: mkitajima@eng.hokudai.ac.jp

Summary. Wastewater-based epidemiology (WBE) is an efficient, non-invasive, and objective means of pathogen surveillance that does not rely on clinical specimens. It can be used to determine the prevalence of infectious agents at the population level. We have developed a highly sensitive virus detection technique in wastewater, capable of detecting viruses even during low epidemic periods when the number of newly infected persons per 100,000 population is 1 person/day. This technology was implemented in local governments and the Tokyo 2020 Olympic and Paralympic Village, where it was used as one of the indicators of infection status and contributed to infection control measures. It is expected that this technology will be widely used in society in the future as a surveillance method to complement clinical testing and as a comprehensive pathogen surveillance method important in preparing for the next pandemic.

Keywords. COVID-19, SARS-CoV-2, influenza, PCR, public health.

Investigation of experimental procedures for detecting SARS-CoV-2 from airborne low-density RNA samples

Jun Uetake^{1*}, Tomoya Baba², Shunichi Matsuura³, Masanori Arita²

^{1*} *Field Science Center for Northern Biosphere, Hokkaido University*

² *National Institute of Genetics, Research Organization of Information and Systems*

³ *National Institute of Advanced Industrial Science and Technology*

Email: jun.uetake@fsc.hokudai.ac.jp

Summary. The infection of the novel coronavirus (SARS-CoV-2) is suspected to transfer through the air, and many localized outbreaks would be attributed to airborne transmission of the virus in the form of aerosolized particles. Therefore, it is urgent to understand the mode of SARS-CoV-2 transmission through both respiratory droplets and aerosolized particles. This study aims to optimize experimental procedures for the detection of SARS-CoV-2 in samples from environments with extremely low virus concentrations. The impact of reagents and sampling conditions on the recovery of RNA is compared for two aerosol collection methods that have suggested SARS-CoV-2 detection capability. The results revealed significant variations in RNA recovery rates, particularly due to differences of buffers required for storage and elution. Even commonly used virus transport media could significantly reduce RNA recovery. This suggests that improving the experimental procedures could potentially enhance detection sensitivity, not only in the air but also in low RNA concentration environmental samples.

Keywords. SARS-CoV-2, bioaerosol, airborne transmission.

A Recursive Method for Identifying Glacial Landforms in Hillshaded Areas Using Digital Elevation Data and CNN Models

Kassim S. Mwitondi^{1*}, **Robert D. Storrar**²

^{1*} *Sheffield Hallam University, Department of Computing*

² *Sheffield Hallam University, Department of the Natural and Built Environment*

Email: k.mwitondi@shu.ac.uk

Summary. Data randomness presents a major challenge to scientific communities in all fields, and it is often exacerbated by strong correlations among data attributes across sectors. Consider, for instance, the non-orthogonality of the 17 United Nations Sustainable Development Goals (SDG), in which the warming of the Arctic (SDG#13 - Climate Action), is associated with human and natural phenomena elsewhere. Apparently, the warming of Arctic is having a considerable impact on glacier retreat and as glaciers retreat, they leave behind landforms that record elements of both their historic and contemporary behaviour. These landforms can be used to better understand how glaciers have changed in the past, and therefore how they may continue to change in the future under increased warming. More specifically, the areas where ice sheets used to be present, can help us understand how water moved underneath the ice. Long, winding hills made of sand and gravel, known as eskers, form a record of meltwater drainage and hence measurements of eskers can be used to understand how much water and sediment moved beneath ice sheets. Such knowledge can improve our understanding of how water is likely to be moving beneath the modern ice sheets. We propose an automated method for mapping esker locations to facilitate detailed morphometric study of their form, using Convolutional Neural Network (CNN) models. High resolution Digital Elevation Models (DEM) data were obtained from a test area in Canada, consisting of 1041 esker positive JPEG files and 37000 esker negative JPEG files. The DEM data files - rasters (grid squares) where the pixel value represents elevation, were processed into a hillshade (a model of shadows based on an illumination angle, with pixel values representing shadows), from which eskers can be identified and mapped. Multiple CNN models were trained and tested via a specially-designed Sample-Measure-Assess (SAM) algorithm, with a built-in recursive mechanism for model optimisation. To address data randomness, and attain model optimisation, the algorithm compares performances across samples and model architectures. We demonstrate that tackling this problem space requires interdisciplinary, timely and recursive adaptations to applications beyond the realms of SDG#13. Further, we highlight the potential for recursively adapting this strategy to an integrated cloud-based data modelling system for developing a consilient interdisciplinary environment for Big Data Modelling of SDG research and prediction of how Arctic glaciers may evolve in the future.

Keywords. Artificial Intelligence, Big Data Modelling, Cloud Computing, Convolutional Neural Networks, Data Recycling, Digital Elevation Models, Eskers, Hillshading, Machine Learning, SMA Algorithm.

Synergies of Satellite Datasets, Modeling and Data Analytic Approaches in Pollutant Retrieval

Hugo Wai Leung MAK^{1, 2*}

^{1*} *Department of Mathematics, The Chinese University of Hong Kong, Shatin,
New Territories, Hong Kong, China*

^{2*} *Department of Mathematics, The Hong Kong University of Science and Technology,
Clear Water Bay, Kowloon, Hong Kong, China*
Emails: hwlmak@math.cuhk.edu.hk; hwlmak@ust.hk

Summary. In recent decades, air pollution has become a global pressing issue, thus it is of paramount importance to obtain accurate fine-scale spatial and temporal distribution of major pollutants that could affect health qualities, so that relevant precautionary measures can be implemented. This talk first focuses on exploring how detailed high-resolution satellite products, numerical modeling and geo-processing framework, and data analytic techniques can be properly synergized for retrieving spatial maps of fine particulate and nitrogen dioxide concentrations. The process of laying down the machine learning framework for retrieval and prediction will be explored, while validation against raw measurement datasets and associated statistical assessments were conducted, with the aim of extending aforementioned approaches for larger-scale spatial retrievals. Observations of spatial or temporal pollutant trends and features, the identification of crucial physical parameters, as well as some possible room for enhancing the existing big data geoprocessing framework will also be discussed.

Keywords. Satellite Remote Sensing, Statistical Machine Learning Methods, Atmospheric Modeling, Environmental Data Analysis, Ground-based Pollutant Retrieval.

NSSDC's Data Governance Practices for AI for Space Science

Zou Ziming^{1,2*}, **Hu Xiaoyan**^{1,2}, **Xu Jiyao**¹, **Tong Jizhou**^{1,2}, **Xu Qi**^{1,2}

^{1*} *National Space Science Center, Chinese Academy of Sciences*

² *National Space Science Data Center*

Email: mzou@nssc.ac.cn

Summary. Artificial intelligence enables a new era in the space science research data application and research, offering new opportunities as well as challenges, such as the absence of data governance theories and standards, data quality to be improved, and insufficient supply of intelligence-driven data application tools. To make good use of large-scale space science data quickly and accurately, NSSDC has carried out a series of data governance practices for AI for space science. On the theoretical side, NSSDC constructs conceptual models of scientific data governance and develops metadata and data management standards. For the data application tool, NSSDC develops a set of intelligent analysis tools and systems for the space scientific data life cycle, based on the research of key technologies such as automatic event identification and feature extraction, correlation analysis and modelling, and data fusion management. Furthermore, in exploring ways to address the supply of high-quality domain scientific data, NSSDC is developing domain datasets for AI-ready, with a prospect of promoting data analysis and unleashing the value of space science data.

Keywords. AI for space science, intelligence-driven data application tools, space science data application, domain data governance.

OpenCSDB: Enabling Knowledge-Driven Scientific Data Linking and Integration

Han Fang^{1*}, Feng Liu¹, ZhiHong Shen¹, Kun Chen¹

^{1*} *Computer Network Information Center, Chinese Academy of Sciences*
Email: hanfang@cnic.cn

Summary. Inspired by the AI for Science research paradigm, there is an urgent imperative to shift traditional data sharing services towards knowledge-driven data solutions. Leveraging semantic linking technology and large-scale language models (LLMs) to organize, link, and uncover insights within vast scientific datasets is the fundamental strategy to confront these challenges. This paper introduces the OpenCSDB framework for semantic linking of scientific data, built upon extensive research into global data linking technology and LLMs trends and advancements in various application fields. OpenCSDB achieves breakthroughs in pivotal areas like structured data linking and publication, cross-endpoint data path discovery, semantic retrieval via LLMs, and data integration services. It has developed relevant tools and achieved initial success in linking, publishing, and semantically fusing scientific data. Implemented in data centers spanning chemistry, plant science, microbiology, and beyond, OpenCSDB has delivered favourable service outcomes, further enhancing data centers' capacity to offer semantic, knowledge-based services. The establishment of a scientific data linking network by diverse field-specific data centers is envisioned as a novel data infrastructure, aptly catering to the evolving demands of emerging research paradigms.

Keywords. linked data, LLMs, semantic retrieval, cross-endpoint, data centers.

A Robust Data-Driven Method for Quantifying Methane Emission from Cattle

Tebogo Sebake^{1*}, David Waswa²

^{1*} African Earth Observation Network, Nelson Mandela University, Port Elizabeth South Africa

² Centre for Broadband Communication, Nelson Mandela University, Port Elizabeth South Africa

Emails: tebogo.sebake@gmail.com, david.waswa@mandela.ac.za

Summary. Livestock production is a major contributor to methane emissions, with enteric fermentation and manure management accounting for approximately 20% of total emissions each year. Globally, livestock occupy an estimated 30 million square kilometres of grazing land - more than the land used for human food production. Despite this, the number of livestock is still increasing, with around 1.5 billion cattle currently being raised to meet the growing demand for meat-rich diets in developing countries. Methane is a potent greenhouse gas that contributes to climate change and has 23 times more global warming potential than carbon dioxide, making it crucial to measure accurately. Various methods for measuring methane emission rate have been proposed, such as the use of permeation tube containing SF₆ that relies on the SF₆ release rate and other parameters. However, most of these proposed techniques are susceptible to randomness due to data sampling. We propose a novel sensor scheme for robust and accurate quantification of methane emissions from cattle. The scheme relies on multiple data readings from each of the sampled animal at different time points and states and it is implemented as follows. First, a methane sensor is developed and mounted to a halter, placed next to the mouth and neck of the animal. This sensor is used to establish the concentration of methane in one burp and the number of burps per day for 60 days. The variations in the daily number of burps, the state of the animal as well as the concentrations are captured for comparative analysis. The novel methane sensor is calibrated using a Picarro, a field deployable gas spectrometer. Lastly, a vacuum-sealed tube is used to collect gas belched by the animal and sent to the lab for analysis of isotopic signatures. This has a detection range of 200 to 10000 ppm of methane and a response time of around 10 seconds, providing big data input models when aggregated over time. It is expected that this work will positively contribute to accurate quantification of carbon emissions, hence help in stewarding our environment to desired concentrations of greenhouse gases. It will also significantly contribute towards understand the impact of human activities and work towards reducing carbon footprint.

Keywords. Big Data, Comparative Analysis, Data Accuracy, Data Randomness, Data Visualisation, Gas Sensors, Gas Spectrometer, Greenhouse Gas Concentration, Isotopic Signature, Livestock Production, Methane Emission, Methane Sensor, Quantification.

Analysis on characteristics of surface-observed and model-based cloud cover for historical weather reconstruction using data assimilation

Xiaoxing Wang^{1*}, Mika Ichino¹

^{1*} *Center for Open Data in the Humanities, Joint Support-Center for Data Science Research, ROIS*
Email: wangxx@ism.ac.jp

Summary. Old diaries precisely recorded locations and dates of weather events during the historical period. Cloud cover amount can be converted from these descriptive weather records; it remains objective, unaffected by recorder subjectivity. Previous studies have confirmed the feasibility of assimilating cloud cover to reconstruct historical weather. Nevertheless, inconsistencies between surface-observed and model-simulated cloud cover may introduce additional errors in the assimilation process. This study examined the discrepancy between cloud cover converted from surface-observed “general weather conditions” and cloud cover from the reanalysis data over a five-year period starting from 1995. Results indicated that in the winter season, stations showing large discrepancy are distributed on the Sea of Japan side. The error originates from that model simulation could not reproduce the high frequency when cloud cover is greater than 90%. The attributes in other seasons are currently under examination and will be elucidated during the presentation.

Keywords. Data assimilation, historical documents, descriptive weather records, cloud cover distribution, reanalysis data.

Explainable AI Framework for Interpreting the Electrochemical Oxygen Evolution Activity of (Ni-Fe-Co-Ce)O_x Catalysts

Paul Rossener Regonia^{1*}, **Christian Mark Pelicano**²

^{1*} Department of Computer Science, College of Engineering, University of the Philippines Diliman, Philippines

² Department of Colloid Chemistry, Max Planck Institute of Colloids and Interfaces, 14476 Potsdam, Germany
Email: pdregonia@up.edu.ph

Summary. This study presents an Explainable AI (XAI) framework for predicting oxygen evolution activities of (Ni-Fe-Co-Ce)O_x electrocatalysts, which are essential for renewable energy storage and conversion. The XAI framework employs a Random Forest Regression (RFR) model, which outperforms previous models ($R^2=0.82$ vs. SVR $R^2=0.72$, ANN $R^2=0.57$, KNN $R^2=0.68$) in predicting overpotentials of (Ni-Fe-Co-Ce)O_x compounds. Despite its high predictive power, RFR remains partially explainable because its massive decision trees obscure the feature interactions and decision boundaries. To enhance the model's interpretability, we used SHapley Additive exPlanations (SHAP) to generate explanations for feature importance and interactions. The results shed light on the contribution of each metal composition to the overpotential value prediction. With this XAI framework, we promote the responsible use of AI in materials research and development by enabling machine learning models to be more transparent without compromising their performance.

Keywords. Explainable AI, OER, Machine Learning, Random Forest Regression, Water Oxidation Catalyst.

Predictive Modelling of Under-Five Mortality Determinants Using Machine Learning Techniques

Rakesh Kumar Saroj^{1*}

^{1*} *Jawaharlal Nehru University*
Email: rakeshsaroj@jnu.ac.in

Summary. Background: Under-five mortality is one of the major public health issues and directly influences the population's health, social development, and economic status of countries. Thus, early detection is essential to find what efficient prevention can take to save it. Therefore, this study will explain how machine-learning techniques can be helpful in predicting the important determinants of under-five mortality in India. Methods: This study used data from the National Family Health Survey-V of India. We performed the tenfold cross-validation to assess the model's capability in the dataset. The decision tree, random forest, logistic regression, neural network, ridge regression, k-nearest neighbor, and naive Bayes machine learning models were used in under-five mortality data, and metrics like confusion matrix, accuracy, recall, precision, F1-score, Cohen kappa, and the area under receiver operative characteristics (AUROC) were used to assess the predictive power of the models. The chi-square scores, recursive feature elimination, extra tree classifier, random forest importance, sequential feature selector, and traditional logistic regression were used to predict the important features(factors) of under-five mortality. All computational algorithms were done with the help of SPSS-27 and Jupiter notebook (inbuilt python 3.3) software. Result: The result reveals that the random forest model was the best predictive model compared to other ML models for under-five mortality. The Random Forest model's precision was estimated to be 98.88% for all factors and 96.25% for important selected variables. After that, neural network accuracy was 96.52%, and accuracy was 94.83% with important variables. Traditional logistic regression accuracy was 93.99% and 93.51%, respectively. The number of living children, breastfeeding status, birth in the last five years, children ever born, time, antenatal care, region, size of children, number of household members, and birth order, were important factors of under-five mortality after using the feature selection methods. Conclusions: The random forest model predicted the most important factors with the highest accuracy of under-five mortality, however, the neural network also showed decent accuracy. We conclude that machine learning algorithms could find unique risk variables for under-five mortality compared to traditional logistics regression models.

Keywords. Logistic regression, Machine Learning, Mortality, Accuracy, AUROC.

How Can Digital Innovation Transform People's Behaviour in Response to the COVID-19 Restrictions in Indonesia

Kiki Adhinugraha^{1*}, Rebriarina Hapsari¹, Vincentia Rizke Ciptaningtyas¹

^{1*} *La Trobe University, Australia*
Email: k.adhinugraha@latrobe.edu.au

Summary. Digital innovations and data generation continuously impinge on our daily lives and behaviour. In this project, we aimed to monitor the behaviour of people in Indonesia using low-cost, low-complexity BLE (Bluetooth Low Energy) technologies in response to the COVID-19 safety restriction. We record how the people in Indonesia responded to the digitalised COVID-19 safety protocols, using IoT basic IoT technology. We monitored the activities in three different locations, which are the ER department, hospital administration office and a boarding school, an ad hoc choice of locations to reflect societal representativeness. In this observation, we applied several scenarios with three different devices, including smartphone utilisation and two types of low-cost BLE devices. Our observation found that the smartphone is not suitable to be used to monitor the people movement due to impracticality, especially in the ER department. The first type of BLE bracelet (Moko W5) which lack features but has very long battery life tended to be the most preferable device, compared with the second type of BLE bracelet (MT58) which offers smart watch features with active proximity sensors that keep reminding the users to keep the safe distance. Due to the short battery life and annoying active proximity reminder, this device is considered as the least preference device. During this observation, the data is collected through BLE beacon scanner MKGW1, and we managed to collect around 4M records for 3-months observation. Our experiment shows that it is possible to utilise the low-cost, low-complexity BLE device to monitor people's movements indoors. The W5 costs less than \$20 each, and the MT58 cost is \$25 each. These devices are quite affordable to be implemented in a developing country such as Indonesia. Interestingly, people do not like to have active monitoring devices, where it might distract their activities. They prefer to have passive monitoring devices without fancy features if it has minimal impact to their work behaviour. More intensive study is required to learn about the characteristic of these devices to ensure effective deployment. Furthermore, in order to ensure the safety distance rule is obeyed, the authorities should use a certain approach that might be suitable to be applied in Indonesia or other countries as well.

Keywords. Behaviour Monitoring, COVID-19, Digital Innovation, IoT, Low-cost BLE.

Hyper baric oxygen

Norihan Abdelaziz Elgazzar ^{1*}

^{1*} Tanta university

Email: wafaagad2015@gmail.com

Summary. Indication and contraindications of HBOT. Physical therapy is still the first line of rehabilitation in CP that must be started as early as possible to improve gross and fine motor dysfunction. HBOT has minor role in improving motor delay and disabilities of CP. HBOT may be used as complementary line with physical therapy, speech therapy due to its effect in improving speech problems and cognitive problems. HBOT is better to be received as early as possible in CP before 5years. HBOT is important in rehabilitation of mildly affected cases than severely affected.

Keywords. HBOT.

Data publication and collaborative research infrastructure of the Antarctic Syowa MST/IS (PANSY) radar

Taishi Hashimoto^{1*}, **Masaki Tsutsumi**¹, **Masashi Kohma**², **Koji Nishimura**³,
Kaoru Sato²

^{1*} *National Institute of Polar Research*

² *Department of Earth and Planetary Science, Graduate School of Science, The University of Tokyo*

³ *Research Institute for Sustainable Humanosphere, Kyoto University*

Email: hashimoto.taishi@nipr.ac.jp

Summary. The Program of the Antarctic Syowa MST/IS radar (the PANSY radar) is a large VHF-band atmospheric radar located at the Syowa Station, Antarctic (69.01°S, 39.59°E). The PANSY radar first started its standard operation with the full capability in late 2015. Since then, we have been conducting 24/365 operation of measuring fine 3-D wind vectors in Troposphere, Stratosphere, and Mesosphere. Now that we have accumulated enough amount of data and experiences in operating an atmospheric radar in Antarctic region, we have shifted the PANSY radar to a collaborative research infrastructure early in 2023. We have started two new services: a peer-reviewed time-sharing program of radar operations and a web site for data publication. In this presentation, we will introduce these efforts and invite applications for observing time and data sharing in 2024.

Keywords. Antarctic, Atmospheric radar, time-sharing, data publication, peer-reviewed program.

An overview of metadata in Research Data Clouds towards Global Open Research Commons

Mikiko Tanifuji^{1*}, Kazutsuna Yamaji¹

^{1*} *National Institute of Informatics*
Email: tanifuji@nii.ac.jp

Summary. With the promotion of open science, there is a growing demand for reliability and sustainability in research infrastructures as a place to store, analyse and utilise research data across data infrastructures and services and to practise algorithms. We have realised through numerous research projects that expectations for the NII Research Data Cloud (NII RDC), one of our research infrastructures in Japan, are increasing. Through practices of national projects and international collaborations, we have been trying to find a way to ensure the sustainable provision of a management and publication infrastructure, and to find a balance between quality and usability that can be managed under data policy in each domains. The Research Data Alliance (RDA), a leading community in research data, has been conducting real practices for open science on the balance that is an important counterpart: it encompasses a wide range of research data issues, from infrastructure to services, and addresses common challenges through working group activities, in particular the Global Open Research Commons (GORC) has become one of the key issues in recent years. This paper introduces the concept of globally interoperable metadata descriptions and metadata attributes based on the Global Open Research Commons International Model Working Group (GORC-WG) in the RDA. We attempt to map the suggested concepts in Standards & Conventions and discussing the operational perspective of the Research Data Cloud (NII RDC) in Japan, for pursuing towards the FAIR.

Keywords. Research data cloud, data repository, NII RDC, RDA GORC, FAIR.

Empowering the Future of Data Science: Introducing the Activities of the WDS-ECR Network

Maja Dolinar^{1*}

^{1*} *Slovenian Social Science Data Archives, Faculty of Social Sciences, University of Ljubljana; Co-Chair of the World Data System ECR Network*
Email: maja.dolinar@fdv.uni-lj.si

Summary. Acknowledging the significance of Early Career Researchers (ECRs) in promoting best practices in data management, analysis, and sharing, the World Data System has established the WDS-ECR Network. The WDS-ECR Network serves as a platform for self-defined ECRs dedicated to advancing data stewardship and best practices in data management. Its primary objective is to connect ECRs across scientific communities and disciplines, fostering collaboration and active participation in the development of innovative approaches to data integration and sharing. This network provides ECRs with opportunities for skill development, knowledge sharing, and networking with experienced leaders within the WDS and beyond. Through its various initiatives, the network aims to empower ECRs, enhance their professional development, and encourage their active participation in shaping the future of research data management. This presentation aims to present the network, its mission, and its goals to encourage active participation and membership among ECRs.

Keywords. Early Career Researchers, World Data System, data management, collaboration, professional development.

CODATA Connect Early Career and Alumni initiative

Shaily Gandhi^{1*}, **Felix Emeka Anyiam**²

^{1*} *CEPT Research and Development Foundation, India*

² *Centre for Health and Development, University of Port Harcourt, Nigeria*

Email: shaily.gandhi@cept.ac.in

Summary. The CODATA Connect Working Group exists to ensure a structural and sustained integration of the activities of the alumni of CODATA-RDA Summer Schools, of the CODATA China and Trieste Training Workshops and other early career researchers and data professionals involved in the CODATA community. The mission of the CCWG is to build a global community of early career individuals working on scientific data to connect people from various disciplines for cooperation towards impactful cross-domain research projects and to open opportunities for global collaboration. CODATA Connect focus on growing the role within the wider CODATA community for developing future membership and leadership by encourage young researchers to disseminate their research projects and findings towards the overall good of the community. The objectives of the CODATA-Connect Working Group are to Propagate the added value of data sharing to early career professionals worldwide. The CCWG is fundamentally inclusive and democratic. The following principles govern membership. The group is concerned with many aspects of data science, analysis and use, data policy, data management and the education, training and career development of data professionals and those in the allied sciences interested in being data professionals. Since 2019 till date there have been various activities which have been undertaken by CODATA Connect working group which involves 3 webinar series on Smart and Resilient Cities, Research Skills and Research Skills Enhancement. The group has hosted 3 podcast series on Data for Disaster Risk Reduction, 'Data-Knowledge-Action for Urban Systems and 'Data for Resilient Cities'. Two essay writing competition has been hosted in 2020 and 2021. A Workshop on Introduction to Spark with R and Lazy evaluation was also conducted online. This initiative has come from the Early career researchers in the group and they have taken the lead in conducting the activities.

Keywords. Early Career Researchers, Data Skills, Data Science, cross-domain research, capacitybuilding.

When Graduates Increase, So Must Faculty Members: Project for Training Experts in Statistical Sciences

Yoshinori Kawasaki^{1*}

^{1*} *The Institute of Statistical Mathematics,
Research Organization of Information and Systems
Email: kawasaki@ism.ac.jp*

Summary. As the demand for data scientists increases in society, a number of universities are establishing data science faculties and related new departments. Statistics is a fundamental discipline of data science, but there is a complete lack of statistics faculty in the newly established department. Therefore, a national project named "Project for Training Experts in Statistical Sciences" was started in 2021 to provide short-term (2 years) training to young faculty members dispatched from consortium universities who did not necessarily specialize in statistics, and to have them return to their universities as "statistical experts" to teach statistics to undergraduate and graduate students and supervise their research. This presentation gives a brief overview of the project and its results to date.

Keywords. Human Resource Development, Data Science Education, Statistical Experts, University Consortium.

Introduction of Data Science Activities for Young Researchers in IUGONET

Shun Imajo^{1*}, **Atsuki Shinbori**², **Shuji Abe**^{3,4}, **Satoru UeNo**⁵, **Masahito Nosé**⁶,
Yoshimasa Tanaka³

^{1*} *Data Analysis Center for Geomagnetism and Space Magnetism, Graduate School of Science, Kyoto University*

² *Institute for Space-Earth Environmental Research (ISEE), Nagoya University*

³ *Joint Support-Center for Data Science Research (DS), Inter-university Research Institute Corporation Research Organization of Information and Systems (ROIS)*

⁴ *International Research Center for Space and Planetary Environmental Science(i-SPES), Kyushu University*

⁵ *Astronomical Observatory, Graduate School of Science, Kyoto University*

⁶ *Nagoya City University*

Email: imajo.shun.4k@kyoto-u.ac.jp

Summary. IUGONET (Inter-university Upper atmosphere Global Observation NETwork) is a research project since 2009 to understand the mechanism of the long-term variation in the upper atmosphere by analyzing various data obtained from the global ground-based network observation. We provide metadata schema, metadata database, and analysis software to manage and analyze the above-mentioned variety of data. One of our action plans is the international contribution and capacity building. We contribute to the training of student and young researchers in the solar terrestrial physics through hands-on workshops. We also promote the international collaborative research in cooperation with researchers in Asia, Africa, and other regions. In this poster, we give our activities related data science for young researchers.

Keywords. Open science, Research data, Metadata database, Analysis tool, Interdisciplinary study.

Navigating the Research Data Management Compliance Maze as an Indonesian Early-Career Researcher: Unravelling Policies Across Borders

Madiareni Sulaiman^{1*}

^{1*} *University College London, UK and BRIN Indonesia*
Email: madiareni.sulaiman.22@ucl.ac.uk

Summary. Research Data Management (RDM) policies, which encompass Open Science policies, are formulated by diverse communities, institutions, regions, and countries. Various approaches are employed, contingent upon the specific context, field of study, capabilities, and even existing policies, such as those pertaining to ethics, data protection, publishing, copyright, intellectual property and more. As individuals engaged in professional research, they are expected to possess a certain level of familiarity with the aforementioned concepts. However, it remains uncertain whether early career researchers possess the same level of knowledge and understanding. Researchers are required to navigate through a multitude of paperwork in order to comply with diverse layers of policies that have been established. Can it generate a complex maze or labyrinth consisting of a thousand interconnected points? How can researchers enhance their skills within this particular ecosystem, especially Indonesian who work collaboratively and internationally? Three key insights can be derived from personal experiences: the importance of assuming the role of a data champion, the value of networking, and the significance of collaboration. Active participation in their own domain is crucial for the acquisition and reinforcement of knowledge. While layer policies of RDM may have been implemented, it is crucial that they align with the continuous development of the skills and abilities of not only researchers but also the individuals providing research support.

Keywords. Research data management, open science, policies, compliance, early career researchers.

DAEPO Projects of National Astronomical Data Center (NADC) in China

Shanshan Li^{1*}

^{1*} *National Astronomical Observatories, Chinese Academy of Sciences*

Email: lishanshan@nao.cas.cn

Summary. Mega-science astronomical projects such as DESI, EUCLID, FAST, GAIA, JWST, LAMOST, LSST, SDSS, and SKA, located around the world, generate massive amounts of data every day, as do large scale astronomical simulations. While a large part of this data is designed and used by astronomers, physicists, and computing experts, there is also a growing interest in using this data for education and public outreach (EPO) activities. This has led to the development of numerous data-driven astronomy education and public outreach (DAEPO) projects, which benefit from the development of Big Data and Internet technology and are characterized by their flexibility and diversity. National Astronomical Data Center of China (NADC) has also made efforts in this regard and has achieved certain results. I will briefly introduce the DAEPO projects currently being carried out by our NADC.

Keywords. Education, Data driven, citizen science, public outreach, science data.

Educational Activities for Young Researchers in Kyushu University's MAGDAS Project

Shuji Abe^{1,2*}, **Akihiro Ikeda**³, **Akiko Fujimoto**⁴, **Teiji Uozumi**⁵, **Yuki Obana**², **Kirolosse Girgis**², **Moe Hayashi**⁶, **Kumi Takayama**⁶, **Takuya Ijuin**⁶, **Akimasa Yoshikawa**^{7,2}

^{1*} *Joint Support-Center for Data Science Research, Inter-university Research Institute Corporation
Research Organization of Information and Systems*

² *International Research Center for Space and Planetary Environmental Science, Kyushu University*

³ *Kagoshima National College of Technology*

⁴ *Kyushu Institute of Technology*

⁵ *Laboratory of Lean Satellite Enterprises and In-Orbit Experiment, Kyushu Institute of Technology*

⁶ *Department of Earth and Planetary Sciences, Graduate School of Science, Kyushu University*

⁷ *Department of Earth and Planetary Sciences, Faculty of Science, Kyushu University*

Email: abeshu@i-spes.kyushu-u.ac.jp

Summary. International Research Center for Space and Planetary Environmental Science (i-SPES), Kyushu University is the research institute aiming for further developing space weather/solar-terrestrial physics, and space and planetary environmental science by expanding their time and spatial axes. As a core research infrastructure of this center, we operate the global geomagnetic field observation network, named MAGnetic Data Acquisition System (MAGDAS). The magnetometer system of this project and data are used not only for academic research by professional researchers, but also for educational activities for the general people and/or students who aspire to become researchers. We had been holding some lectures for instrument operation, data analysis, our research field itself when we visited observation sites. In addition, we invited our collaborators of foreign observation sites to attend lectures in Japan. In this paper, we will introduce our educational activities with some examples.

Keywords. Research data, Geomagnetism, Capacity building.

Progress of the AMIDER project

Masayoshi Kozai^{1*}, **Yoshimasa Tanaka**¹, **Shuji Abe**¹,
Yasuyuki Minamiyama², **Atsuki Shinbori**³

^{1*} *Polar Environment Data Science Center, Joint Support-Center for Data Science Research, Research Organization of Information and Systems*

² *National Institute of Informatics, Research Organization of Information and Systems*

³ *Institute for Space-Earth Environmental Research, Nagoya University*

Email: kozai.masayoshi@nipr.ac.jp

Summary. AMIDER is a cross-disciplinary scientific data catalog designed to promote cross-disciplinary studies and uses by non-specialized users. It provides inclusive browsing and access to data in diverse scientific fields such as bioscience, geoscience, and space science. In this presentation, we will introduce the AMIDER project and report on the preparation for a test operation that will start this year focusing on the polar science field.

Keywords. Open science, Research data, Database, Cross-disciplinary research.

Conducting a Lesson on Earth's Global Sea Ice Distribution Using a Handmade Globe

Takeshi Sugimura^{1*}, Hironori Yabuki¹

^{1*} *National Institute of Polar Research*
Email: sugimura.takeshi@nipr.ac.jp

Summary. We constructed and launched a website that generates "Handmade Globes," a tailored educational tool for elementary school classes, aiming to enhance students' understanding of Earth's conditions through origami. Students utilize AMSR2 satellite data to visualize sea ice concentration and sea surface temperatures, creating origami globes from printed paper reflecting this data. This hands-on method enables students to observe seasonal sea ice fluctuations and regional distinctions, fostering environmental awareness. We conducted a global environmental awareness class at a Kagoshima elementary school using this resource. Remarkably, students displayed newfound interest in the sea ice melting phenomenon linked to global warming's impact. This initiative successfully encourages students to explore and appreciate our planet's environment.

Keywords. Origami, Handmade Globe, elementary school class, global environmental awareness class, AMSR2 satellite data.

Educational practices using polar sea ice data for high school science and mathematics classes

Yoshihiro Niwa^{1*}

^{1*} *International Polar and Earth Environmental Research Center, National Institute of Polar Research*
Email: niwa.yoshihiro@nipr.ac.jp

Summary. As an example of data-based climate change education, we conducted two classroom practices using polar sea ice data for high school students. The first educational practice was conducted for first-grade high school students in their mathematics class on data analysis. Each student operated a computer to analyze 30-year time series data of Arctic and Antarctic ice extent. The students observed the contrast in the decreasing trend between Arctic and Antarctic sea ice extent. Then they had a free discussion about the polar climate, the statistical validity of the trend difference, etc. The second educational practice was conducted for first-grade high school students in their science class on scientific inquiry activity. Each student used the online tool ADS (Arctic Data Archive System) of NIPR, accessible at <https://ads.nipr.ac.jp/vishop/>, to repeatedly and closely observe the seasonal variation of Arctic sea ice distribution. As a result, the students could discover and learn about various formation and melt processes of Arctic sea ice.

Keywords. High school education, Climate change education, Polar sea ice data, ADS.

International Policy and Practices of Open Science-based Data Sharing

Yasuhiro Murayama^{1*}

^{1*} *National Institute of Information and Communications Technology*
Email: murayama@nict.go.jp

Summary. This paper will discuss international policy regarding sharing and management of research data, e.g., as discussed in G7, UN/UNESCO, and ISC bodies, which will be influential for the data strategies in the Asia and Oceania data centers/repositories. Those top-down approaches have been effectively working to improve leadership in related ministries and countries, for example, defining national strategy and/or national academies' policies and so forth. On the other hand, community-based activity is absolutely important and critical for making the global vision reality. The author attempts to introduce a few examples of community practices for research data management (RDM), and research data sharing (RDS), compliant with the Open Science as well as FAIR (findable, accessible, interoperable and reusable) principles.

Keywords. Open Science, international data policy.

A Survey of current Status of Open Science oriented Activities in the Asia-Oceania Area

Takashi Watanabe^{1*}, **Toshihiko Iyemori**², **Yasuhiro Murayama**¹, **Masaki Kanao**³,
Yasuhisa Kondo⁴

^{1*}*National Institute of Information and Communications Technology*

²*Kyoto University*

³*Joint Support-Center for Data Science Research, Research Organization of Information and Systems*

⁴*Research Institute for Humanity and Nature*

Email: watanabegl@gmail.com

Summary. Based on a general understanding of discussions held in the previous Asia-Oceania (A&O) oriented symposia to promote regional activities on data (Kyoto in 2017, Beijing in 2019, and Tokyo in 2021), the WDS National Committee (a sub-committee) of the Science Council of Japan has initiated to operate a Web site “WDS Asia-Oceania Network*” including information on the current status of Open-Science oriented activities in the A&O area. As a new trend, several South-Asian countries making an effort to establish their own data networks. Collaborations with these regional networks will be one of our essential works. Another remarkable movement seen in the A&O area is that almost all countries in the area have established their own platform to promote Open-Science activities in their countries. It will be important to build a collaborative network to establish systems for preservation and provision of science data being produced by scientific activities in these countries. Although every country in this area is operating well-established governmental data repositories, involvement of these repositories will be our next challenge.

*<https://takashiwatanabe.wixsite.com/wds-asia-oceania>

Keywords. Research Data, Data Network, Asia-Oceania.

The Action Plan and Priorities of the World Data System

David Castle^{1*}, **Meredith Goins**²

^{1*} *University of Victoria, Canada*

² *University of Tennessee, Knoxville, USA*

Email: dcastle@uvic.ca

Summary. World Data System (WDS) Chair of the Scientific Committee and the executive director for the International Program Office share information and updates on the WDS Action Plan (2022-2024). The plan has four objectives: (1) Provide services and support to existing and new WDS members, (2) Develop value narratives for WDS members, (3) Provide global leadership and agenda setting, and (4) Enhance access, quality, and accessibility of data worldwide. An overall aim on increasing regional participation in Africa, Latin and South America, as well as Southeast Asia will focus the discussion on priorities and implementation.

Keywords. Repositories, regions, engagement, access.

Toward a Reproducible Research Data Repository

Cheng-Jen Lee^{1*}, **Chia-Hsun Ally Wang**¹, **Ming-Syuan Ho**²,
Tyng-Ruey Chuang^{1,2,3*}

^{1*} *Institute of Information Science, Academia Sinica, Taiwan*

² *Research Center for Information Technology Innovation, Academia Sinica, Taiwan*

³ *Research Center for Humanities and Social Sciences, Academia Sinica, Taiwan*

* Email: {cjlee, trc}@iis.sinica.edu.tw

Summary. The depositar (<https://data.depositar.io/>) is a research data repository at Academia Sinica (Taiwan) open to researchers worldwide for the deposit, discovery, and reuse of datasets. The depositar software itself is open source and builds on top of CKAN. CKAN, an open source project initiated by the Open Knowledge Foundation and sustained by an active user community, is a leading data management system for building data hubs and portals. In addition to CKAN's out-of-the-box features such as JSON data API and in-browser preview of uploaded data, we have added several features to the depositar, including sourcing from Wikidata for dataset keywords, a citation snippet for datasets, in-browser Shapefile preview, and a persistent identifier system based on ARK (Archival Resource Keys). At the same time, the depositar team faces an increasing demand for interactive computing (e.g., Jupyter Notebook) which facilitates not just data analysis, but also for the replication and demonstration of scientific studies. Recently, we have provided a JupyterHub service (a multi-tenancy JupyterLab) to some of the depositar's users. However, it still requires users to first download the data files (or copy the URLs of the files) from the depositar, then upload the data files (or paste the URLs) to the Jupyter notebooks for analysis. Furthermore, a JupyterHub deployed on a single server is limited by the machine's processing power which may lower the service level to the users. To address the above issues, we are integrating the BinderHub into the depositar. BinderHub (<https://binderhub.readthedocs.io/>) is a kubernetes-based service that allows users to create interactive computing environments from code repositories. Once the integration is completed, users will be able to launch Jupyter Notebooks to perform data analysis and visualization without leaving the depositar by clicking the BinderHub buttons on the datasets. In this presentation, we will first make a brief introduction to the depositar and BinderHub along with their relationship, then we will share our experiences in incorporating interactive computation in a data repository. We shall also evaluate the possibility of integrating the depositar with other automation frameworks (e.g., the Snakemake workflow management system) in order to enable users to reproduce data analysis.

Keywords. BinderHub, CKAN, Data Repositories, Interactive Computing, Reproducible Research.

Overview of Research Activities at Indian Institute of Geomagnetism

A. P. Dimri^{1*}

^{1*} *Indian Institute of Geomagnetism, Mumbai, India*
Email: apdimri@hotmail.com

Summary. Indian Institute of Geomagnetism (IIG) is one of the important institutes working on the geomagnetism which covers its studies from lithosphere-atmosphere-ionosphere-magnetosphere. Within in these spheres there are strong projects seen for the coupling among them. It has a unique data record from 1816 onwards of Colaba-Alibaug observatory. In this presentation, an overview of research activities of IIG will be introduced.

Keywords. IIG, geomagnetism, coupling of multi-spheres, Indian data center, geophysical observatory.

The role of data publishing activities at universities and research institutes on Asia and Oceania

Shaliza Ibrahim^{1*}

^{1*} *Institute of Ocean and Earth Sciences, University of Malaya, Kuala Lumpur*
Email: shaliza@um.edu.my

Summary. Data publishing plays an important role in shaping the research culture in Asia and Oceania. In Malaysia the launch of the Malaysia Open Science Platform sees at least the five research universities piloting data repositories connected to a central hub, which is in turn linked to research institutes and agencies also participating in the initiative. Data sharing is also expandable regionally and globally. Several positive impact of successful data publishing include allowing for better interactions among researchers or data originators, opportunities for collaborations, optimal use of resources, and better strategies and research data management at institutional, national and international levels. Data publishing will instil good research practices that pay attention to data quality, integrity, validity and interoperability, if the objective of creating vibrant scholarly interactions is to be realized.

Keywords. data integrity, interoperability, research data management, Asia and Oceania, research culture.

National Scientific Repository (RIN)

Sjaeful Afandi ^{1*}

^{1*} *National Research and Innovation Agency (BRIN)*

Email: sjaeful.afandi@brin.go.id

Summary. The government of Indonesia has ratified a regulation that accommodates Mandatory Deposit of Primary Data and the Output of Research Results through Law Number 11 of 2019 concerning the National System of Science and Technology, specifically in article 40. Furthermore, the National Research and Innovation Agency (BRIN) has also issued BRIN Regulation Number 18 of 2022 concerning Mandatory Deposit of Primary Data and the Output of Research Results in BRIN. In accordance with these regulations, all parties in Indonesia involved in data management, storage, and exchange must have an integrated and trusted repository management mechanism. Following these regulations, the National Research and Innovation Agency (BRIN), an Indonesian research institution, created a data repository system known as the National Scientific Repository (RIN). The National Scientific Repository (RIN) is an open source-based management system for primary data, based on a dataverse platform. The RIN system was developed to facilitate researchers, academics and engineers in storing and managing research data. RIN was created in response to the needs of researchers to store, preserve, and share research data.

Keywords. deposit, regulation, data, scientific, repository.

Towards Open Science: Chinese National Spatial Science Data Centre Activities

Xu Qi^{1*, 2}, **Hu Xiaoyan**^{1,2}, **Tong Jizhou**^{1,2}, **Zou Ziming**^{1,2}

^{1*} *National Space Science Center, Chinese Academy of Sciences*

² *National Space Science Data Center*

Email: xuqi@nssc.ac.cn

Summary. In the context of open science, data repositories are expected to play a more important role in effective data governance and maintaining the value of data. This presentation will introduce the activities implemented by the Chinese National Space Science Data Center (NSSDC) as a data facility and a national-level data center under the current trend. NSSDC is implementing specific activities around data governance, including open data action of research projects, quality assurance, clear ownership, trustworthy repository and community focusing. NSSDC is expected to improve the state of open data and application of space science data, respond to the needs of domain communities in data management, data discovery, data access, data analysis and mining application in these activities.

Keywords. space science data management, open data, domain data governance activities, space science data application, data repository.

Open Government Data (OGD) in Bangladesh NSO

Chandra Shekhar Roy^{1*}, Suranjit Mandal², Md Kamal Hossain³

^{1*} *Bangladesh Bureau of Statistics, Ministry of Planning*

² *Centre for Research and Information*

³ *Nano Information Technology*

Email: csroy1@gmail.com

Summary. To implement open government data (OGD), the NSO shall implement several principles; there are standard data, metadata, interoperability, and reference code or master data. All open data are not reused. The open data initiatives as well as users' intent for open data are changing continuously and today, in line with AI, IoT and smart Bangladesh trends, real-time data and sensor-generated data have higher interest for users. These "smarter" open (government) data are also considered to be one of the crucial drivers for the sustainable economy, and might have an impact on ICT innovation and become a creativity bridge in developing a new ecosystem in 5IR and Society 5.0. The presence of "smarter" data, their level of accessibility, availability, velocity and timeliness, as well as support for users, are analyzed. The OGD portals react to users' needs and updating of data for their further potential reuse, which is essential in the digital data-driven society.

Keywords. Society5.0, Metadata, Sustainability, OGD, Timelines.

The Global Microbiology Open Science Data Cooperation

Juncai Ma^{1*}

^{1*} *Institute of Microbiology, Chinese Academy of Sciences*
Email: ma@im.ac.cn

Summary. Established in the 1960s, World Data Centre for Microorganisms, hereinafter referred to as WDCM (<http://www.wdcm.org/>), is the most important physical resource data platform in the field of microorganisms throughout the world. Getting hosted by the Institute of Microbiology, Chinese Academy of Sciences in 2010, WDCM is the first world data center in the field of life sciences in China, which has long been committed to facilitating the application of cutting-edge information technology to improve the interoperability of microbial data, promote the access and use of data and information, and coordinate international open science and cooperation between culture collections, scientists and other user communities. Culture Collections Information Worldwide (CCINFO) is a world directory of culture collections maintained by WDCM, which contains data on the organization, management, services and scientific interests of the collections. At present, there are 843 international culture collections from 79 developed and developing countries joined in CCINFO.

To help plenty of culture collections that cannot make their data available online as well as advocate open science and global cooperation, WDCM launched the Global Catalogue of Microorganisms (GCM) (<http://gcm.wdcm.org/>) project in 2012. GCM program aims to provide a globally uniform data warehouse for valuable microbial resources scattered in various culture collections and in the hands of the scientists around the world, providing effective data support for all aspects of physical resources of microorganisms such as gathering, collection, transnational transfer, academic and commercial applications and benefit sharing, and offering the most important support to the implementation and enforcement of the Convention on Biological Diversity in the field of microorganisms. Up to now, GCM has become one of the largest data portals for public service microbial collections and several international culture collection networks, providing data retrieval, analysis, and visualization system for microbial resources. Furthermore, GCM gradually developed into a knowledge base linking taxonomy, phenotype, omics data as well as relative scientific papers and patents with its catalogue information, which currently has aggregated 529,902 strains and other holdings (plasmids and antibodies) deposited in 151 collections from 51 countries.

On the basis of GCM international cooperation program, WDCM launched GCM2.0 international cooperation program--Global Microbial Type Strain Genome and Microbiome Sequencing Project--for complete coverage of microbial genomes, and established cooperation network for genome sequencing and function exploring of microbial resources covering 30 major culture collections in more than 20 countries. globally authoritative reference database and data analysis platform of microbiome. Focused on exploring the genomic information of microorganisms, this project has planned to sequence all uncovered prokaryotic type strains together with select eukaryotic type strains, construct a database for genomics data sharing based on open science, and also provide online data mining environment. Working groups responsible for selecting bacterial and fungal strains, drafting SOP, managing

intellectual property right and legal issues and constructing database have already embarked on the pioneer stage of GCM 2.0. The project will establish an open science network for type strain sequencing and functional mining, and complete genome sequencing of over 10000 species of microbial type strains in five years.

Keywords. Microbiology, Data, Open science, Global Catalogue of Microorganisms.

Snow Properties Survey and Snow Avalanche Hazards in Tianshan Mountains

Lanhai Li^{1,2*}, **Jiansheng Hao**³

^{1*} *Xinjiang Key Lab of Water Cycle and Utilization in Arid Zone, Urumqi, Xinjiang, China*

² *Tianshan Snowcover and Avalanche Observation and Research Station of Xinjiang, Xinyuan, Xinjiang, China*

³ *Institute of Geographic Sciences and Natural Resources Research Chinese Academy of Sciences, Beijing, China*

Email: lilh@ms.xjb.ac.cn

Summary. Conducting a systematic and standardized field survey of snow properties in Tianshan Mountains (TSs) provides reliable and comprehensive snow properties information for remote sensing inversion, climate change impact assessment, water resource management, and snow disaster prevention. Snow properties survey was conducted with three scales: station, line, and polygon. Three field surveys were conducted in three periods within a snow season: snow accumulation period, stable period, and melting period, and lasted two snow seasons (2017-2018, 2018-2019). The results indicated that snow snow properties during the three snow season periods was different. Snow depth increases with elevation in all three periods, while Snow density and deep hoar increases with elevation in both snow accumulation and stable periods, and they showed no significant difference in melting period. Snow avalanche types vary with seasonal change due to snow properties heterogeneity. Through collecting and analyzing data related to avalanche, snowpack, and meteorology during four snow seasons (from 2015 to 2019), this study quantifies the hazards of different types of avalanches and describes their situation over time in the continental snow climate region. The findings highlighted that surface-layer dry snow avalanches were characterized by large sizes and high release frequencies. The overall avalanche hazard showed a single peak pattern over time during a snow season, and the greatest hazard occurred in the second half of February when the snowpack was deep and the temperature increased. The findings help the disaster and emergency management departments rationally arrange avalanche relief resources and develop avalanche prevention strategies.

Keywords. Snow property, Snow avalanche, hazards, snow seasons, Tianshan Mountains.

Long term records of glacier evolution and associated proglacial lakes on the Tibetan Plateau 1976 - 2020

Lhakpa Dolma^{1*}, **Yubao Qiu**²

^{1*} *Tibet Institute of Plateau Atmospheric and Environmental Sciences, Tibet Meteorological Bureau*

² *Aerospace Information Research Institute, China Academy of sciences*

Email: DG1827016@smail.nju.edu.cn

Summary. The glaciers on the Tibetan Plateau (TP) constitute critical sources of water for the proglacial lakes and many rivers found downstream. To better understand the evolution of glaciers and the impact of this on proglacial lakes, seven glaciers corresponding to continental, subcontinental, and marine climate types that are influenced by westerlies and the Indian summer monsoon were selected for study. The evolution of the edges of these glaciers and their associated proglacial lakes were identified based on the visual interpretation of Landsat TM/ETM+/OLI images. A dataset covering the period 1976 – 2020 that included the glacier and proglacial lake edge vectors was then created. The relative errors in the areas of the individual glaciers were less than 3%, and for the proglacial lakes these errors were in the range 0%–7%. The dataset was used to effectively compare the changes in glaciers and proglacial lakes that have occurred over the past four decades. The most striking changes that were found were the retreat of glaciers and the formation of small proglacial lakes. This dataset could also be used as a proxy to support research on changes in mountain glaciers, particularly their response to climate change and water resources. This response is of great scientific significance and is important in many applications, including assessments of the ecological problems caused by melting glaciers.

Keywords. Glacier and proglacial lake, long-term changes, Landsat, Tibetan Plateau.

Data Management of the Arctic and Antarctic Region

Hironori Yabuki^{1*}, **Takeshi Sugimura**¹

^{1*} *National Institute of Polar Research*

Email: yabuki.hironori@nipr.ac.jp

Summary. In the National Institute of Polar Research, perform the integration and sharing of data across a multi-disciplinary such as atmosphere, ocean, snow and ice, land, ecosystem, model, for the purpose of cooperation and integration across disciplines, we build the Arctic Data archive System (ADS). On the other hand, the National Institute of Polar Research to operate the ADS, has acquired a wide variety of polar science data through Japanese Antarctic research expedition (JARE) and Arctic research project. The data has been published through "JARE date reports" and "NIPR Arctic Data Reports". NIPR has published a new data journal "Polar Data Journal" for the purpose of distributing quality-controlled actual data. ADS play a role as a main data repository for "Polar Data Journal". PDJ publication and ADS collaboration will contribute to the promotion of open science.

Keywords. Arctic, Antarctic, Data Management, Open science.

Utilization of Snow and Ice Disaster Prevention Information
created by National Research Institute for Earth Science and
Disaster Resilience (NIED), Japan

Satoru Yamaguchi^{1*}

^{1*} *Snow and Ice Research Center,
National Research Institute for Earth Science and Disaster Resilience
Email: yamasan@bosai.go.jp*

Summary. More than half of Japan's land area is covered by heavy snowfall in winter. Every winter, avalanches, blizzards, and other snow and ice disasters threaten the safety and security of residents. In recent years, abnormal weather conditions caused by global warming have changed the face of snow and ice disasters. As a result, the number of unusual snow and ice disasters that cannot be dealt with by conventional methods is increasing. The National Research Institute for Earth Science and Disaster Resilience (NIED) has been developing information by combining various observation data and numerical model simulation techniques in order to realize efficient and effective response to snow and ice disasters. The generated information has been used in actual disaster response on a trial basis during the winter season, based on co-creation with stakeholders. In this presentation, those activities by NIED will be introduced.

Keywords. snow and ice disaster, creating information, disaster response, observation data, numerical simulation.

Digital Arctic Environment Information Service

Yubao Qiu^{1,2*}, **Yang Li**^{1,2}, **Andrea Marinoni**³, **Lijuan Shi**^{1,2}, **Paola De Salvo**⁴,
Lanhai Li⁵, **Shuwen Yu**^{1,2}, **Massimo Menenti**⁵

^{1*} *Aerospace Information Research Institute, Chinese Academy of Sciences, China*

² *International Research Center of Big Data for Sustainable Development Goals (CBAS)*

³ *Dept. of Physics and Technology, the Arctic University of Norway, Norway*

⁴ *Group on Earth Observations, Switzerland*

⁵ *Xinjiang Institute of Geography and Ecology, Chinese Academy of Sciences, China*

⁶ *Delft University of Technology, Netherland*

Email: qiuyb@aircas.ac.cn

Summary. The processes of the environment exert an immense and profound influence on all human productions and life. The undeniable effects of global warming on international and regional environments, such as water, land and atmosphere, energy style, and economic activities, call for innovative science and technology to mitigate its influences.

The Earth's environment is sensitively impacted by and responds to climate actions, especially the snow-ice-covered plateaus, and the Earth's poles. The scientific and decision-making concerns in these areas include natural physical environmental changes; the interactions between climate actions and the environment; the strategies for climate actions to construct a future society–environment nexus; and how digital technologies, e.g., remote sensing, big data, communication, navigation, and ICT can assist the implementation of environmental assessments.

The world is currently undergoing sci-tech revolutions and industrial transformations on a larger scale and in greater depth than ever before. As a frontier of science and technology, digital technology is key to meeting the challenges of society, environment, and economy, which is also a vital element for achieving the United Nations Sustainable Development Goals (UN SDGs). Digital technologies, such as earth observations, data science, big data, and artificial intelligence, are integrating into every aspect of society.

While, under an increasing warming trend for the Arctic regions, the changing temperature inducing the snow and ice melting faster than ever, which poses new challenge for the society activities, the changes of snow, and ice brings new adaption for the people there, the observations are needed for these understanding and services. Recently, fast development of the information data from the space- or air- borne observations provides new insight into the services for these kinds of adaption.

The Digital Environment Technology (DET) employs frontier digital technology, including Big Earth Data, artificial intelligence, aerospace technology, big data, and data science to address the challenges raised by the changes in the Earth's environment. In this talk, the developed data products will be introduced for the Arctic environment, information service for the open seaway by Earth observations valued data products, and data-based AI forecast model, for example, sea ice forecast model.

Keywords. Arctic, Digital Environment Technology (DET), Sea Ice, Information Service, Earth Observations and Forecast.

Application of Independent Component Analysis to Satellite Gravity Data for Antarctic Research

Tianyan Shi^{1*}, **Yoichi Fukuda**², **Koichiro Doi**³, **Jun'ichi Okuno**³

^{1*} *Institute for Geothermal Sciences, Kyoto University*

² *Kyoto University*

³ *National Institute of Polar Research*

Email: shi.tianyan.5w@kyoto-u.ac.jp

Summary. With the accumulation of over two decades of time-variable gravity data from Gravity Recovery and Climate Experiment (GRACE) and GRACE Follow-On (GRACE-FO) satellite missions, a foundational dataset has been established, facilitating the comprehension of the processes driving mass changes in Antarctica. To gain a deeper insight into the complex Antarctic system, we applied independent component analysis (ICA), a statistical blind source separation method, to extract signals from different physical sources. The results reveal that the overall continental-scale mass changes can be effectively separated into several patterns that are potentially dominated by different physical processes, such as glaciers, snowfall, and the solid Earth. These patterns also include the influence of extreme snowfall events and the impact of Glacial Isostatic Adjustment (GIA). ICA demonstrated a distinctive and valuable approach for understanding time-variable gravity data, and along with other statistical methods, deserves further exploration.

Keywords. GRACE/GRACE-FO, Antarctica, ICA, GIA.

