

PROGRAMME

International Workshop on Data Science - Present & Future of Open Data & Open Science -

12 – 15 November 2018

Mishima Citizens Cultural Hall & Joint Support-Center for Data Science Research,
Mishima, Shizuoka, Japan

Monday 12 November 2018

15:30–17:30 Public Lecture (for general public, in Japanese)

16:00–19:30 Registration @ Mishima Citizens Cultural Hall

18:30–19:30 **Icebreaker party @ Mishima Citizens Cultural Hall** (Facilitator: Tomoya Baba)

Tuesday 13 November 2018

09:05–09:30 **Opening Remarks** (Chair: Masaki Kanao)

- **Asao Fujiyama** (Director-General, Joint Support-Center for Data Science Research (DS), ROIS) (15')
- **Masaki Kanao** (LOC Chair), agenda of workshop & practical information (10')

09:30–12:30 **Session A: International and National data activity**

(Session chair: Akira Kadokura, Seiji Tsuboi)

Keynote 1: *The International Seismological Centre (ISC): 50-Year Miracle*

Dmitry A. Storchak (International Seismological Centre) (30')

- *Data Management at Korea Polar Data Center, Korea Polar Research Institute*
Dongchan Joo (Korea Polar Research Institute) (20')
- *Activities of World Data Center for Geomagnetism, Kyoto*
Satoshi Taguchi (Data Analysis Center for Geomagnetism and Space Magnetism, Kyoto University) (20')
- **Coffee Break** (20')

Keynote 2: *World Data Center for Microorganisms: The global cooperation on microbial big data*

Juncai Ma (Chinese Academy of Sciences) (30')

- *DIAS Platform Contributing to Open Science in Earth Environmental Informatics*
Masaki Yasukawa (Earth Observation Data Integration and Fusion Research Initiative, the University of Tokyo) (20')
- *Data Management at Polar Research Institute of China*
Lizong Wu (Polar Research Institute of China) (20')
- *Activities of Polar Environment Data Science Center*
Akira Kadokura (Polar Environment Data Science Center, DS, ROIS) (20')
- **Discussion**

12:30–14:00 **Group Photo & Lunch**

14:00–17:20 **Session B1: Data science and Inter-disciplinary science**

(Session chair: Susumu Goto, Koji Nishimura)

Keynote 1: Amenability of the United Nation's Sustainable Development Goals to Big Data Modelling

Kassim S. Mwitondi (Sheffield Hallam University) (30')

- *Life Science Database Integration and Its Application for Medical Science*
Susumu Goto (Database Center for Life Science, DS, ROIS) (20')
- *Current Situation of Data Archiving for Japanese Official Statistics*
Shinsuke Ito (Chuo University) (20')
- *Importance of semantics and rated challenges in science and humanities according to the Humboldtian ideal*
Bernd Ritschel (Kyoto University) (20')
- **Coffee Break** (20')

Keynote 2: SPEDAS (Space Physics Environment Data Analysis Software): Multi-mission Heliophysics Data Management, Analysis, Visualization, and Collaboration

Jim Lewis (Space Sciences Laboratory, University of California) (30')

- *Application of deep learning and large scale simulation to the Earth science*
Seiji Tsuboi (Japan Agency for Marine-Earth Science and Technology) (20')
- *Domestic and international activities of DOI-minting to solar-terrestrial physics data and their citation in publication*
Masahito Nose (Nagoya University) (20')
- *Data Treatment Strategies and Some Science Projects of PANSY Rader*
Koji Nishimura (Polar Environment Data Science Center, DS, ROIS) (20')
- **Discussion**

17:20–18:20 **Poster Session**

18:30–20:00 **Reception @ Mishima Shoukou-Kaigi-Sho** (Facilitator: Masaki Kanao)

- **Welcome Remark: Ryoichi Fujii** (President, Research Organization of Information and Systems (ROIS))

Wednesday 14 November 2018

09:05–11:20 **Session B2: Data science and Inter-disciplinary science**

(Session chair: Hideaki Takeda, Masaki Kanao)

Keynote 3: Data Archives for Social Science in Korea, Challenges and Opportunities

Hearan Koo (Seoul National University) (30')

- *Exploring public attitudes toward scientific research with visitor surveys and nationally representative surveys*
Naoko Kato-Nitta (Center for Social Data Structuring, DS, ROIS) (20')

Keynote 4: Age-Period-Cohort Analysis of Data Obtained from Repeated Social Surveys Such as the Surveys on the Japanese National Character

Takashi Nakamura (Center for Social Data Structuring, DS, ROIS) (30')

- *Extended Cell Suppression Problem Towards Better Data*
Kazuhiro Minami (The Institute of Statistical Mathematics, ROIS) (20')
- *Estimation of age-specific reporting ratio of sentinel influenza surveillance using seroprevalence data*
Masaya Saito (The Institute of Statistical Mathematics, ROIS) (20')

- **Coffee Break** (15')

11:20–13:20 **Session C: Legacy data, Historical data, Industry-academia collaboration**

(Session chair: Takeshi Watanabe, Naoko Kato-Nitta)

Keynote 1: *The ISC-GEM Global Earthquake Catalogue: Making Good Use of Historical Observations*

Dmitry A. Storchak (International Seismological Centre) (30')

- *Reframing Scholarly Communication by Persistent Identifier*
Hideaki Takeda (National Institute of Informatics, ROIS) (20')

Keynote 2: *Resource and Environment Scientific Data Sharing and Disaster Risk Reduction Knowledge Service*

Juanle Wang (Chinese Academy of Sciences) (30')

- *Multidisciplinary Study of the Earth's Environment in 18th-19th Centuries - A Trial to find an Approach to the Open Data and Open Science*
Takashi Watanabe (International Program Office, ICSU World Data System) (20')
- *Research Data Management in Industry-Academia Collaboration*
Yuko Toda and Hodaka Nakanishi (Teikyo University) (20')
- **Discussion**

13:20–14:50 **Lunch & Poster**

14:50–17:50 **Visit to NIG & DBCLS**

18:00–19:30 **Banquet @ NIG Lecture Hall** (Facilitator: Hideki Noguchi)

- **Welcome Remark:** **Isao Katsura** (Director-General, National Institute of Genetics, ROIS)

Thursday 15 November 2018

09:05–12:30 **Session D: Education and capability building** (presented by Japanese, using English slides)

(Session chair: Masanori Arita, Yoshimasa Tanaka)

Keynote 1: *Health - Prediction of Infectious Disease System Dynamics using Machine Learning and Mathematics*

Shailza Singh (National Centre for Cell Science, India) (30')

- *Open genome analysis in the post-genomic era*
Masanori Arita (National Institute of Genetics, ROIS) (20')
- *Education Programs for AI/IoT/BigData Skills Development at the Medical and Drug Discovery Data Science Consortium*
Eri Kaminuma (Tokyo Medical and Dental University) (20')
- *Kyoto University Academic Data Innovation Unit - bottom-up promotion for various research data infrastructure*
Takaaki Aoki and Shoji Kajita (Kyoto University) (20')

- **Coffee Break** (15')

Keynote 2: *From Human Genome Project to Genome Cohort Study*

Atsushi Shimizu (Iwate Tohoku Medical Megabank) (30')

- *Genetic relationship between tea plant (*Camellia sinensis*) and ornamental camellia (*C. japonica*)*
Kazumi Furukawa (National Institute of Technology, Numazu College) (20')
- *Research outline with regional theme: Examples using machine learning, geographic information system, and social big data*
Shizuo Suzuki (National Institute of Technology, Numazu College) (20')

- *Useful Tools for Education and Capacity Building about Solar Terrestrial Physics Study*
Yoshimasa Tanaka (Polar Environment Data Science Center, DS, ROIS) (20')
- **Panel discussion** (10')

12:30–14:00 **Lunch & Poster**

14:00–17:20 **Session E: Future on data science**

(Session chair: Asanobu Kitamoto, Hironori Yabuki)

Keynote 1: *Embedding Machine Learning in the Data Life Cycle - An Example from Minerals Exploration*

Jens Klump (Mineral Resources, CSIRO in Australia) (30')

- *Toward universal information access on the digital object cloud*
Kei Kurakawa (National Institute of Informatics, ROIS) (20')
- *Development of Japanese Research Data Discovery Service*
Fumihiko Kato (National Institute of Informatics, ROIS) (20')
- *AI and Data Science - Machine Learning as Digital Catalyst for Data Curation*
Asanobu Kitamoto (Center for Open Data in the Humanities, DS, ROIS) (20')

- **Coffee Break** (20')

Keynote 2: *Open Research Data Progress in Korea*

Myung-Seok Choi (Korea Institute of Science and Technology Information) (30')

- *The Cooking Recipes without Border Data set: FAIR challenges*
Frederic Andres (National Institute of Informatics, ROIS) (20')
- *Arctic Data Archive System (ADS)*
Hironori Yabuki (Polar Environment Data Science Center, DS, ROIS) (20')
- *Perspective of Open Science in Japan driven by Integrated Innovation Strategy*
Kazuhiro Hayashi (National Institute of Science and Technology Policy) (20')
- **Discussion**

17:20–17:30 **Closing Remarks** (LOC Chair)

Poster Presentation @ outside the Conference Hall

Core Time 1; Tuesday 13 November 17:20–18:20, Wednesday 14 November 13:50-14:50

- *P-1: Multivariate analysis of the occupations of rental rooms by using the housing information website data*
Hayafumi Watanabe (Center for Social Data Structuring, DS, ROIS)
- *P-2: Useful Tools for Education and Capacity Building about Solar Terrestrial Physics Study*
Yoshimasa Tanaka (Polar Environment Data Science Center, DS, ROIS)
- *P-3: Arctic Data Archive System (ADS)*
Hironori Yabuki (Polar Environment Data Science Center, DS, ROIS)
- *P-4: Data Processing and Archive System for the Antarctic PANSY Rader*
Koji Nishimura (Polar Environment Data Science Center, DS, ROIS)
- *P-5: "Polar Data Journal" : A new data publishing platform for polar science*
Akira Kadokura (Polar Environment Data Science Center, DS, ROIS)
- *P-6: A decade of history for polar data management in Japan*
Masaki Kanao (Polar Environment Data Science Center, DS, ROIS)
- *P-7: Data and metadata sharing among AFoPS countries*
Masaki Kanao (Polar Environment Data Science Center, DS, ROIS)
- *P-8: Development of Metadata, Conversion and Archiving of the Time Series Data of the Completed Censuses and Surveys of the BBS*
Chandra Shekhar Roy (Bangladesh Bureau of Statistics)
- *P-9: An attempt for the thermal transport modelling of fusion plasmas based on the statistical approach*
Masayuki Yokoyama (National Institute for Fusion Science)

Core Time 2; Thursday 15 November 13:00-14:00 (for Session D)

- *P-10: A Python library for parallelised particle filters*
Shinya Nakano (The Institute of Statistical Mathematics, ROIS)
- *P-11: A Study on Speaker Discrimination Using Machine Learning*
Wataru Murata and ManYong Jeong (National Institute of Technology, Numazu College)
- *P-12: A Study on Speaker Discrimination by Deep Learning*
Shinpei Mine and ManYong Jeong (National Institute of Technology, Numazu College)
- *P-13: A Study on Tapping Diagnosis of Structures Using Machine Learning*
Ken Higuchi and ManYong Jeong (National Institute of Technology, Numazu College)
- *P-14: Development of Traffic Flow Data Measurement System by Drive Recorder*
Hironori Kanazawa and ManYong Jeong (National Institute of Technology, Numazu College)
- *P-15: Optimization of Pipe Support Arrangement by Machine Learning*
Hisashi Sakashita and ManYong Jeong (National Institute of Technology, Numazu College)
- *P-16: Genomic analysis in somatic embryogenesis of tea plant(Camellia sinensis)*
Kaito Yamaki and Kazumi Furukawa (National Institute of Technology, Numazu College)
- *P-17: Preliminary estimation of the aesthetic value of cultural ecosystem services by mapping geotagged photos from social media data around Mt. Fuji and Izu Peninsula*
Naoto Unno and Shizuo Suzuki (National Institute of Technology, Numazu College)
- *P-18: Preliminary classification of Japanese cedar forests with aerial photographs using machine learning approach*
Ochiai Ryosuke and Shizuo Suzuki (National Institute of Technology, Numazu College)
- *P-19: Preliminary estimation of tea leaf quality using remote sensing techniques in Numazu city*
Tomoaki Okatsu and Shizuo Suzuki (National Institute of Technology, Numazu College)
- *P-20: Preliminary application of results in Japanese historical character recognition by machine learning*

to locally historical documents

Masuhiko Yamaji and Shizuo Suzuki (National Institute of Technology, Numazu College)