Opening Up IIASA's Research

Workshop "Big Data and Systems Analysis"
24 February 2020 - 25 February 2020
IIASA, Laxenburg, Austria

Michaela Rossini, IIASA library

First steps

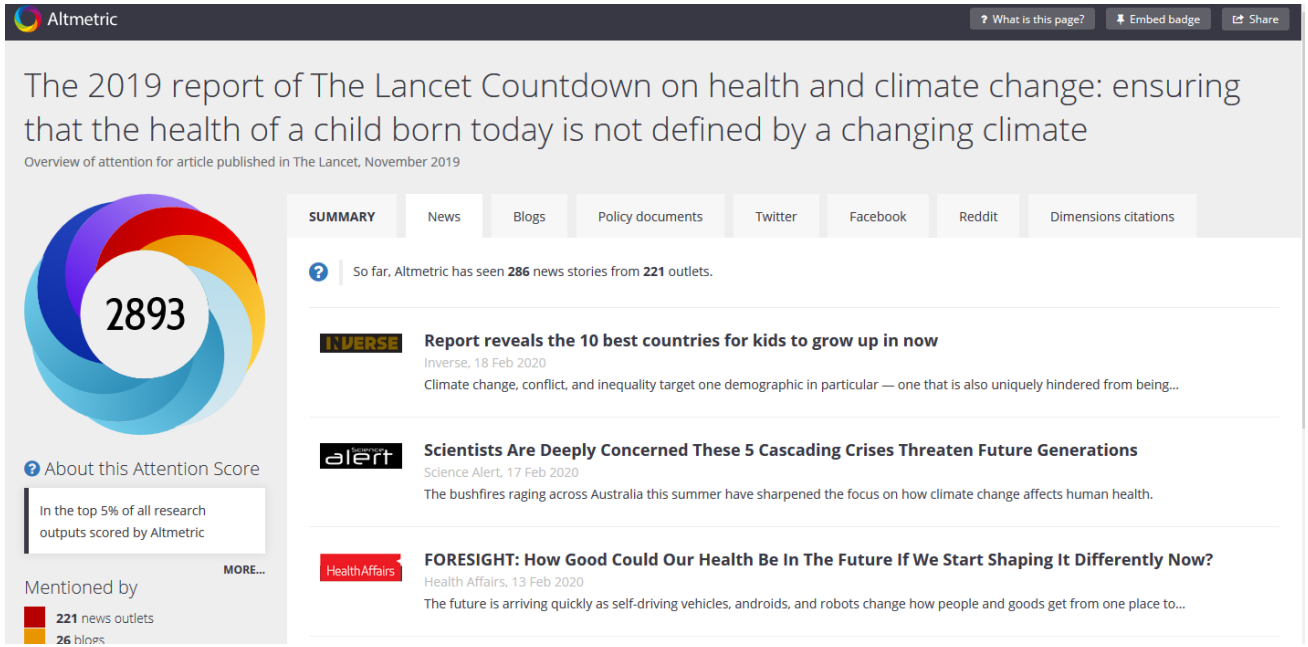
2007: Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities signed

Commitment to providing unrestricted and cost-free online access to scientific publications and data for all users and researchers.

To ensure good data management and open access to all IIASA research, the institute maintains both an institutional publications repository (PURE) launched in 2016 and a data repository (DARE) launched in 2019 for research produced by IIASA affiliated researchers.

Open Access to Publications

- 2016: Migration of +10,000 entries (legacy data)
- 2020: over 15,000 entries, over 1.1 M downloads
- Compliance with open access policy (as of 2016): about 80%
- Altmetric widget to track impact online attention



Open Access to Data

- 2016: Open Access to Data Task Force (OADTF)
- Catalogue of Data, Models and Tools
- Open Access to Data Management Policy
- IIASA Data Repository (DARE) - 34 Datasets since 2019

Welcome to IIASA DARE

IIASA DARE is the Data Repository of the International Institute of Applied Systems Analysis (IIASA), provided as a service by the IIASA Library.

Latest Additions

1. Folberth C, Khabarov N, Balkovic J, Skalsky R, & Obersteiner M (2020). Supplementary Data 1 for: The global cropland sparing potential of high-yield farming. [Data Collection]
2. Poledna S (2020). Economic Forecasting with an Agent-Based Model. [Data Collection]
3. See L (2017). A global reference database of crowdsourced cropland data collected using the Geo-Wiki platform. [Data Collection]
4. Steffen F, See L, Perger C, McCallum I, Schill C, Schepaschenko D, Duerauer M, Karner M, et al. (2012). A global dataset of crowdsourced land cover and land use reference data (2011-2012). [Data Collection]
5. Tramberend S, Burtscher R, Burek P, Kahil T, Fischer G, Mochizuki J, Kimwaga R, Nyenje P, et al. (2020). Data related to research project: East Africa Water Scenarios to 2050. [Data Collection]
6. Leclère D, Obersteiner M, Butchart SHM, Chaudhary A, De Palma A, DeClerk FA, Di Marco M, Doelman JC, et al. (2019). Reversing terrestrial biodiversity declines due to habitat loss: a model ensemble approach. [Data Collection]

More..

IIASA DARE supports OAI 2.0 with a base URL of `http://dare.iiasa.ac.at/cgi/oai2`

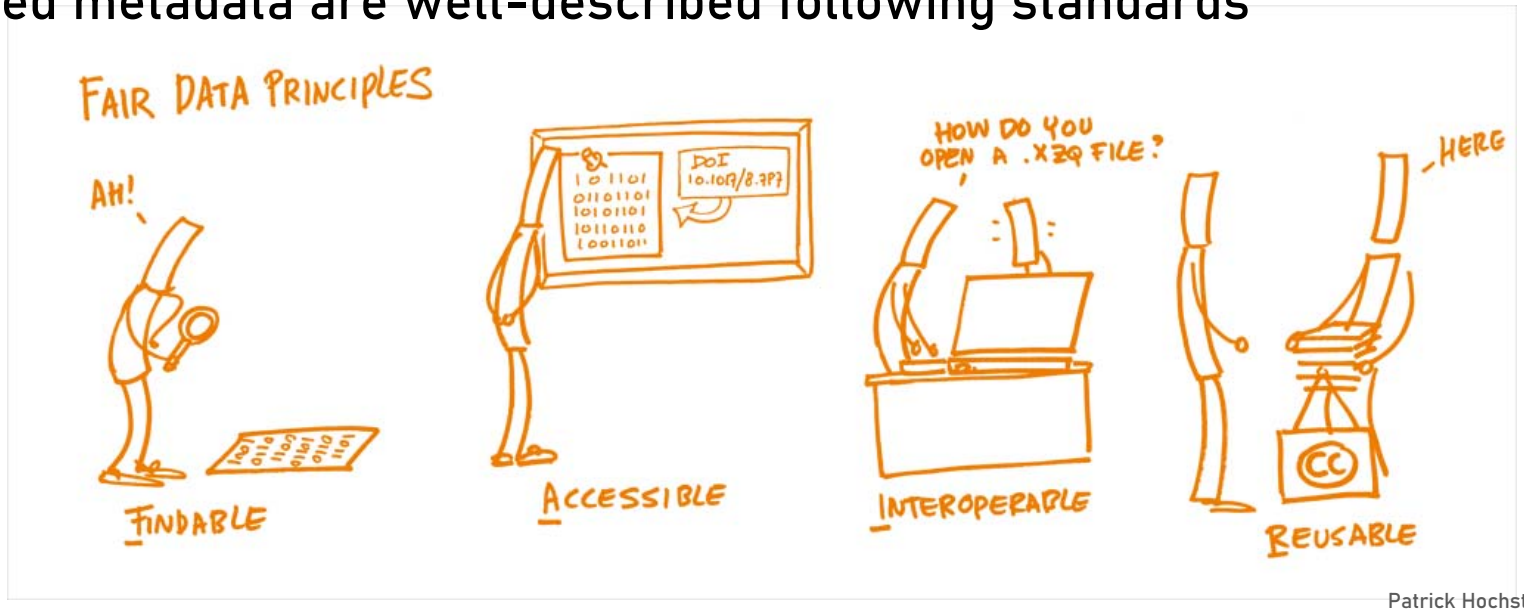
Making Data FAIR

Findable: ensuring data can be found by both humans and machines

Accessible: users need to know how they can get access to the data

Interoperable: ensuring data can be integrated with other data and utilised by applications or workflows for analysis, storage, and processing

Reusable: maximising potential reuse by making sure that data and related metadata are well-described following standards



Integration & Compliance

- PURE and DARE are registered repositories (default CC-BY-NC license, others possible)
- fully machine actionable and regularly harvested by both data harvesters and OpenAIRE for their metadata
- DARE customized to specifically operate as a data repository, so the metadata fields reflect those used in other machine actionable data repositories

ROARMAP
Registry of Open Access Repository Mandates and Policies

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

OpenDOAR
The Directory of Open Access Repositories

ROAR- Registry of Open Access Repositories

DOAJ DIRECTORY OF OPEN ACCESS JOURNALS

OpenAIRE

Levels Of Openness

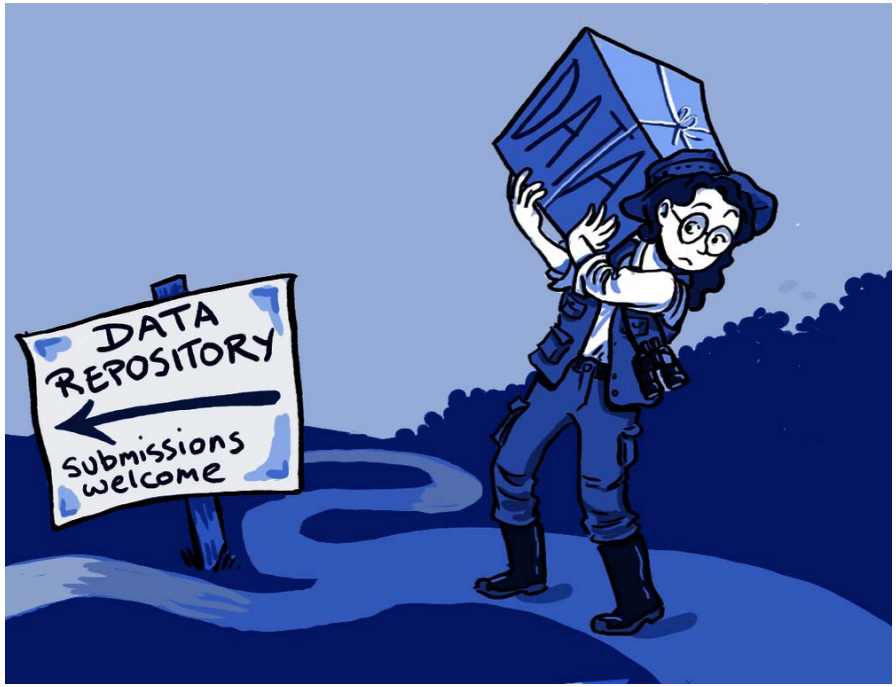
Open data – data that anyone can access, use and share. Must be licensed to make clear what on can do with the data, incl. transforming, combining, and sharing it with others, even for commercial purposes.

Shared data - may be made widely accessible but could have some conditions such as non-commercial reuse or reuse with attribution. It is important to note that not all shared data has to be available to anyone. Sometimes shared data is only made available to specific groups such as peers from another institution or collaborators.

Closed data – e.g. highly sensitive data or data that cannot be shared for certain purposes (working data, contractual obligations etc.)

Responsible Research Data Management

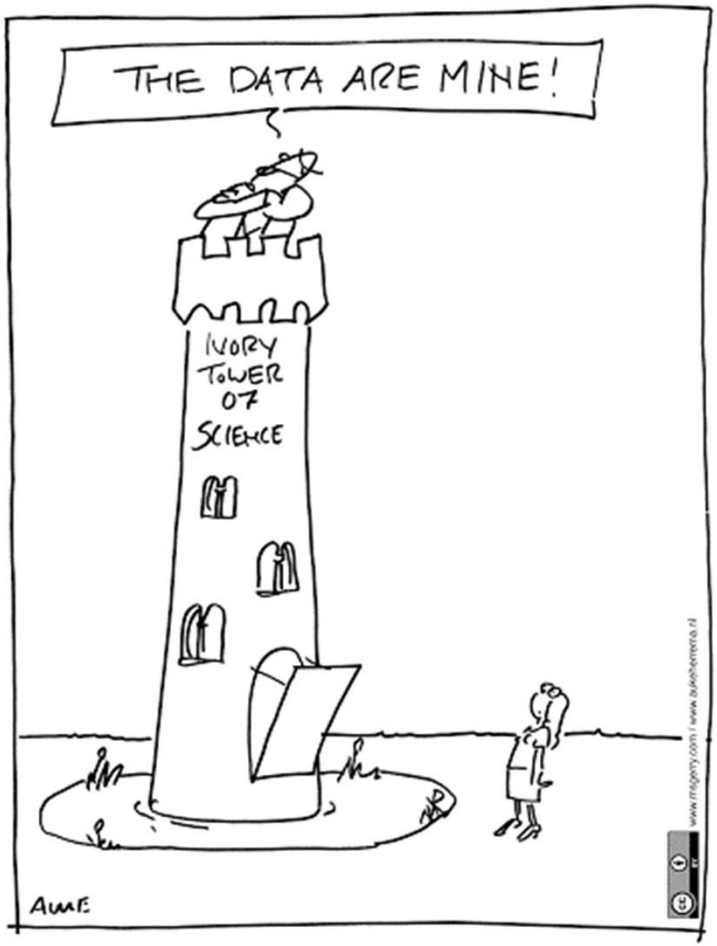
- Guidance: Understand the key elements in the process, as well as roles and responsibilities to support researchers meet expectations and requirements
- Compliance: Understand the key points of the funders' requirements by checking adherence with funder policies, at both ends of the funding process (pre-award and end-of-project)
- Selection: supporting researchers in making decisions regarding what the institution will want to keep /share, under what conditions, and for how long



Open Science - more than just open access to publications!

- Create an environment that enables others to collaborate and contribute
- Make publications, research data, workflows and other research processes freely available, under terms that enable reuse, redistribution and reproduction of the research and its underlying data and methods
- Improve transparency and reproducibility
- Maximize the impact of IIASA's research and provide the foundations for others to build upon
- Train and educate the research community to use, apply and develop IIASA's tool & models further

The Path Forward



SCENE FROM THE PAST?

- Initiate a new process focused on best practice alignment and a more comprehensive Open Science strategy for IIASA including models and tools driven by scientific leadership
- Align efforts regarding sharing models and tools with international best practices and establish principles to ensure the quality of our models and tools
- Engagement with key players like CODATA and other international bodies
- Efficient use of workshop with CODATA practitioners to receive input and expertise

Key actions needed

- Develop guidelines, policies, workflows
- Pay attention to legal aspects
- Provide infrastructure for storage, access and sharing
- Raise awareness of benefits of OA to research data
- Monitor requirements of different stakeholders (e.g. compliance with funders)
- Promote exchange of best practice
- What else should we look at to advance RDM and OA to data?