

FAIRNESS OF MICROMETEOROLOGICAL DATA & RRI: AN OPEN FRAMEWORK FOR CLIMATE RESEARCH

1Branislava Lalic, 2Ivan Koci, 3Mark Roantree

Faculty of Agriculture, University of Novi Sad, Novi Sad, Serbia
Forecasting and Reporting Service for Plant Protection of the Republic of Serbia (PIS)
Insight Centre for Data Analytics, School of Computing, Dublin City University, Ireland



CA20108 - FAIR NETWORK OF MICROMETEOROLOGICAL MEASUREMENTS (FAIRNESS)



Funded by
the European Union



www.cost.eu

OUTLINES

RRI concept

FAIR principles

- provide **Open Access (OA)** to R&I results
- pursue high-quality **Science Education (SE)**;
- develop equal opportunity research environments regardless of gender (**Gender Equality**), race, nationality, or religion;
- provide whole **Public Engagement (PE)** in research and SE by maintaining communication with the general and specialized public and supporting Citizen Science (CS) initiatives;
- overcome all ethical issues (**Ethics**) by constructive discussions with all interested parties;
- participate in **Governance (GOV)** and convince policymakers that research results should be fully considered while designing strategic, long-term, and short-term plans

“The total amount of data created, captured, copied, and consumed globally is forecast to increase rapidly, reaching 64.2 zettabytes in 2020. Over the next five years up to 2025, global data creation is projected to grow to more than 180 zettabytes. In 2020, the amount of data created and replicated reached a new high” (Statista, 2022).

“The total amount of data created, captured, copied, and consumed globally is forecast to increase rapidly, reaching **64.2 zettabytes** in 2020. Over the next five years up to **2025**, global data creation is projected to grow to more than **180 zettabytes**. In 2020, the amount of data created and replicated reached a new high” (Statista, 2022).

The lack of data's findability, accessibility, interoperability, and reusability (FAIR) costs Europe a minimum of € 10.2bn per year - approximately 78% of the Horizon 2020 budget per year (PwC EU Services, 2018)

“The total amount of data created, captured, copied, and consumed globally is forecast to increase rapidly, reaching 64.2 zettabytes in 2020. Over the next five years up to 2025, global data creation is projected to grow to more than 180 zettabytes. In 2020, the amount of data created and replicated reached a new high” (Statista, 2022).

The lack of data's findability, accessibility, interoperability, and reusability (FAIR) costs Europe a minimum of € 10.2bn per year - approximately 78% of the Horizon 2020 budget per year (PwC EU Services, 2018)

Define a minimal set of related but independent and separable guiding principles and practices, which enable both machines and humans to find access, interoperate and reuse research data and metadata" (PwC EU Services, 2018).

FAIR principles

Findability - data could be found by humans or machines

Accessibility - humans or machines can gain access to data under specific conditions or restrictions

Interoperability - data and metadata conform to recognized formats and standards in order to be combined and exchanged.

Reusability - data and metadata are licensed, conforming to community norms and allowing users to know what kinds of reuse are permitted

DATA SOURCE

Micrometeorological data measured by the Forecasting and Reporting Service for Plant Protection of the Republic of Serbia (PIS) AWS network

FAIRNESS VALIDATION METHODOLOGY

Simplified methodology designed by Jones and Grootveld (2017)

DATA SOURCE

Micrometeorological data measured by the Forecasting and Reporting Service for Plant Protection of the Republic of Serbia (PIS) AWS network

FAIRNESS VALIDATION METHODOLOGY

Simplified methodology designed by Jones and Grootveld (2017)

Table 1. Findability – Data can be found by humans or machines

Ref	Metric	Test	Current	CA20108 Enhancement Plan
F1	A persistent identifier is assigned to data	True	AWS ID	DOI for every dataset
F2	There are rich metadata describing the data	Partial	Scientific paper (where available)	Metadata Description
F3	The metadata are online in a searchable resource	Partial	Scientific paper (where available)	Searchable Metadata Repository
F4	The metadata record specifies the persistent identifier	True	Paper DOI	Permanent Metadata Repository.

Table 2. Accessibility – humans or machines can gain access to data under specific conditions or restrictions

Ref	Metric	Test	Current	CA20108 Enhancement Plan
A1	Following the persistent ID will lead to data or associated metadata	True	AWS ID	DOI for every dataset & searchable metadata repository
A2	The protocol by which data can be retrieved follows recognized standards	Partial	http for most data with remaining data in csv format.	Online searchable repository
A3	The access procedure includes authentication and authorization steps	Partial	Mixed Username/password in open access catalogue	Registration required for Cost Action online repository.
A4	Metadata accessible where possible (even when data are not)	True	Paper DOI	Searchable Metadata Repository with access to data.

Table 3. Interoperability – data and metadata conform to recognized formats and standards in order to be combined and exchanged.

Ref	Metric	Test	Current	CA20108 Enhancement Plan
I1	Data is provided in commonly understood and preferably open formats	Partial	Partly JSON, partly (interactively) csv, Excel.	WMO GAMP with downloadable data & metadata in CSV format.
I2	The metadata provided follows relevant standards	True	WMO GAMP	WMO GAMP
I3	Controlled vocabularies, keywords, thesauri or ontologies are used where possible	Partial	According to best practice	WMO GAMP
I4	Qualified references and links are provided to other related data	Partial	Aggregated with other tools	CA20108 metamodel facilitates embedded links.

Table 4. Reusability – data and metadata are licensed, conforming to community norms and allowing users to know what kinds of reuse are permitted.

Ref	Metric	Test	Current	CA20108 Enhancement Plan
R1	Data are accurate, well described with many relevant attributes	Partial	Basic quality control checks.	Automated Quality Control checks including data imputation.
R2	The data have a clear and accessible data usage license	True	Data is free to use	Data is free to registered researchers.
R3	It is clear how, why and by whom the data have been created and processed	True	Fully documented.	Part of the Metadata specification.
R4	The data and metadata meet relevant domain standards	Partial	Both data and metadata meet standards for micrometeorological measurements in a field/canopy layer.	WMO GAMP.