



# BEST PRACTICES TO REACH FAIRNESS

Francesca Martines – INAF -Osservatorio Astronomico di Palermo  
*From Science Gateways to Papers, Palermo, 23-26 May 2022*



FAIRness is a target,  
but...

... how to reach it?

Questa foto di Autore  
sconosciuto è concesso in  
licenza da CC BY-NC

## LET'S START FROM THE BEGINNING...

“The FAIR guiding principles for scientific data management and stewardship”, Wilkinson et al., *Scientific Data*, 2016

*Scientific Data* is a peer-reviewed open-access journal for descriptions of datasets and research that advances the sharing and reuse of research data. (...) These principles are designed to align with and support the [FAIR Principles](#) for scientific data management and stewardship, which declare that research data should be **Findable, Accessible, Interoperable** and **Reusable**.

## THIS GIVES US MANY USEFUL POINTERS:

### To be **Findable**:

- F1. (meta)data are assigned a globally **unique and persistent identifier**
- F2. data are described with **rich metadata** (defined by R1 below)
- F3. metadata clearly and explicitly **include the identifier of the data it describes**
- F4. (meta)data are registered or indexed in a **searchable resource**

THIS GIVES US MANY USEFUL POINTERS:

To be **Accessible**:

- A1. (meta)data are retrievable by their identifier using a **standardized communications protocol**
- A1.1 the protocol is **open, free, and universally implementable**
- A1.2 the protocol allows for an **authentication and authorization** procedure, where necessary
- A2. **metadata are accessible**, even when the data are no longer available

THIS GIVES US MANY USEFUL POINTERS:

To be **Interoperable**:

- I1. (meta)data use a **formal, accessible, shared, and broadly applicable language** for knowledge representation.
- I2. (meta)data use **vocabularies** that follow **FAIR** principles
- I3. (meta)data include qualified **references to other (meta)data**

THIS GIVES US MANY USEFUL POINTERS:

To be **Reusable** :

- R1. meta(data) are richly **described** with a plurality of accurate and relevant **attributes**
- R1.1. (meta)data are released with a clear and accessible **data usage license**
- R1.2. (meta)data are associated with **detailed provenance**
- R1.3. (meta)data meet domain-relevant **community standard**



## WHICH IS THE SITUATION IN ASTRONOMY/INAF?

Essential components of the research work of most astronomers:

- a. (Observational) data, so called «**Raw Data**»
- b. The data obtained as a result of reductions, simulations, filtering, but also the software used for these operations: we will call all these «**Derived Data**»
- c. The **publications**, based on the above data and processes



## RAW DATA

- They are structured in a (more or less) standard way (e.g. FITS)
- Preservation is generally guaranteed
- They are (moderately) findable but... must be cited!
- They are provided with metadata
- They are accessible (more or less open)

## DERIVED DATA

- Are stored on the PCs of individual researchers (no guaranteed preservation)
- They are not findable
- They are not provided with metadata
- They are not accessible

# PUBLICATIONS

- Are kept in open access repositories, so preservation is guaranteed
- Are findable
- Are provided with metadata
- Are accessible



THERE IS CLEARLY A BIG WEAK LINK IN THIS CHAIN...

- **Raw Data** and **Publications** are generally preserved and available (telescope archives, OA repositories, etc.) as well as “FAIR enough”
- **Derived Data** are generally stored on individual researchers' PCs, or on generic storage spaces or repositories such as Zenodo (space limitations) or on 'public' platforms that do not ensure preservation of objects (not to mention FAIRness)

## BEST PRACTICES

Best Practices for Data Publication in the  
Astronomical Literature

<https://iopscience.iop.org/article/10.3847/1538-4365/ac6268>

Tracy X. Chen *et al* 2022 *ApJS* **260** 5

# BEST PRACTICES FOR DATA PUBLICATION IN THE ASTRONOMICAL LITERATURE: GENERAL RULES

- *Define symbols, acronyms, and abbreviations*
- *Provide uncertainty and confidence level when reporting a new measurement*
- *Present the appropriate number of significant figures for numerical measurements and uncertainties*
- *Report the units for measurements if present, and adopt commonly used ones*
- *Indicate preferred values if applicable*

# BEST PRACTICES FOR DATA PUBLICATION IN THE ASTRONOMICAL LITERATURE: SECTIONS

- 2. Astrophysical Data Guidelines: Nomenclature, Astrometry, Photometry, Time, Redshift/velocity, Classifications, Orbital parameters
- 3. Data Presentation: Tables, Figures,
- 4. Data Archiving and Access
- 5. Citations and Credits: Literature citations, Facility credits, Software credits, Digital object identifiers
- 6. Data Content Keywords



## DATA ARCHIVING AND ACCESS (SECTION 4):

- (a) Append small data sets as part of the publication.
- (b) Deposit large or complex data at a long-term archive most appropriate for your data. Adhere to the specific format requirements from the archives.
- (c) Provide a complete list of metadata.
- (d) Include a Data Availability Statement if required by the journal.
- (e) Do not publish data sets at URLs lacking long-term support.

# WHAT DO WE NEED TO ACHIEVE FAIRNESS? (1)

- ***Spread the FAIR culture!***
- Have an institutional ***FAIR Data Policy*** (telling researchers that their data must be FAIR and why)
- Create a ***Data Management Plan*** (which gives specific and operational indications on how to manage data)

## SO WHAT DOES IT TAKE TO ACHIEVE FAIRNESS? (2)

- Offer a **storage space with the possibility of attributing a DOI** to "objects" selected by researchers: the DOI request process involves the implementation of a whole series of metadata (from a FAIR perspective). This way, derived data are deposited, their preservation is ensured and they can be made FAIR.
- Have **dedicated staff**.

## SO WHAT DOES IT TAKE TO ACHIEVE FAIRNESS? (3)

- Last but not least, have ***data citation guidelines*** (for raw and derived data).
- ✓ There are ***no universal standards*** but there are ***many indications*** from institutions such as RDA, UK DCC and many others. These indications generally ***converge on the same points***.
- ✓ In addition, there are Best Practices.



# Thank you!

Best Practices to reach FAIRness - Francesca Martines , INAF OAPa  
From Science Gateways to Papers – Palermo, 26/05/2022