Best Practices for Data Publication in the Astronomical Literature

Marion Schmitz & Tracy Chen (Caltech/IPAC-NED)

Motivation

- Better represent and support science results
- Enable better data sharing
- Improve reproducibility
- Enhance the reusability of data
- Help astronomical archives to streamline the extraction, preservation, integration, and cross-linking of valuable data in literature

Intended audience

- Authors
- Referees
- Journal editors
- Publishers
- International Astronomical Union (IAU)
- International Virtual Observatory Alliance (IVOA)
- Anyone involved in the publication of astronomical nomenclature, data, and metadata

Download

https://ned.ipac.caltech.edu/uri/Docs::BPDP

Reference

1. Cook, D. O., Lee, J. C., Adamo, A., et al. 2019, MNRAS, 484, 4897, doi: 10.1093/mnras/stz331

Best Practices (Checklist)

1. General rules
a) Define all symbols, acronyms, and abbreviations at first use. b) Provide uncertainty and confidence level when reporting a new measurement. c) Present the appropriate number of significant figures for numerical measurements and uncertainties. d) Adopt commonly-used units whenever possible. e) Indicate preferred values if applicable.

2. Nomenclature
a) Provide the complete name. b) Include the "J" in names based on J2000 coordinates. c) Insert spacers between a catalog name and the identifiers within the catalog. d) Distinguish between part of an object and the object itself. e) Do not use the same name for different objects. f) Always assign a name and verify it is unique. g) Keep the appropriate number of significant figures in coordinate-based names. h) Use established names for known objects. i) Confirm the names and positions for cross-identifications. j) Cross-match the same objects in different tables within the same article.

3. Astrometry
a) Provide the best available coordinates. b) Specify the celestial reference system and/or frame. c) Indicate the equinox and epoch of observation when necessary. d) State the wavelength range from which astrometry is obtained.

4. Photometry
a) State the facility, telescope and instrument used. b) Describe the method used to estimate photometry. c) Use standard passband/filter identifiers. d) Clarify the magnitude system. e) Specify spectral transitions completely.

5. Time
a) Provide the time of observation and exposure time. b) Favor full Julian Dates over abbreviated or offset Julian Dates. c) Include phase timing measures along with reported periods when relevant. d) State when observations from multiple missions are executed simultaneously.

6. Redshift/velocity
a) Define the method of redshift measurements and give references. b) Specify the reference frame of the redshift measurements. c) State whether a published redshift or recessional velocity is based on observed frequency or wavelength shifts. d) Provide the wavelength range of the measurement. e) Indicate the quality of the measurement when possible.

RA, Dec = (131.32134587°, 0.01243229°) would imply a positional accuracy of 0.00001", which is not obtainable by most current telescopes.

By SDSS J1441+0948, do you mean
SDSS J144157.24+094859.1, or
SDSS J144156.97+094856.5, or
SDSS J144157.26+094853.7?

Cook et al. (2019)
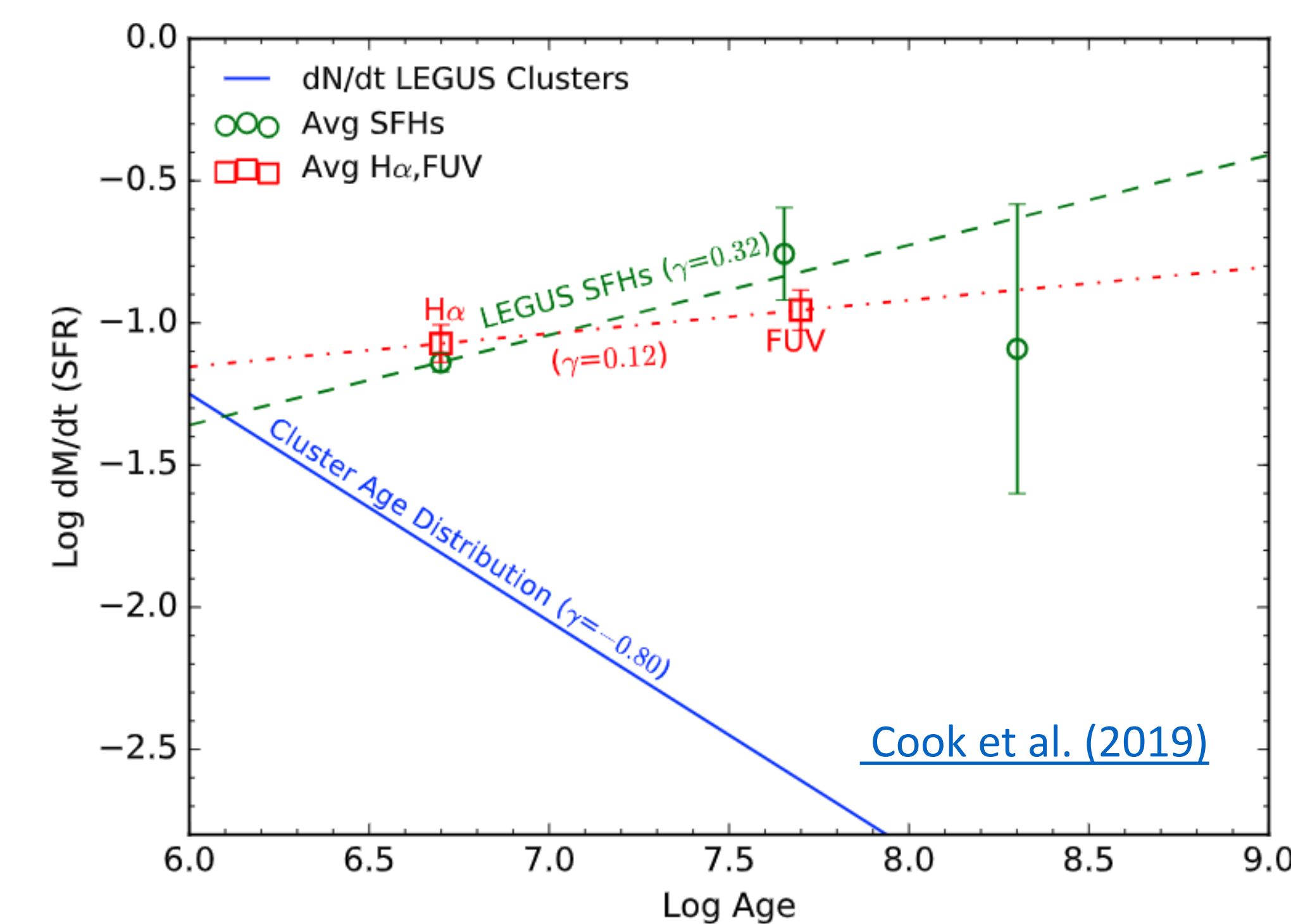
"We adopted a heliocentric redshift of 1.234 (Smith et al. 2012)", where "Smith et al. 2012" should be listed correctly in the bibliography section.

Indicate "Johnson B" or "Cousins B" instead of just "B"; use "2MASS Ks" instead of just "K".

Use JD 2456789.123 instead of MJD 6789.123. MJD is formally defined as the Julian Date minus 2400000.5. This has not always been followed.

Clearly define the facility used. "ARO" could mean
Astronomical Research Observatory
Arizona Radio Observatory
Abbey Ridge Observatory
Algonquin Radio Observatory

Spectroscopic? Photometric? machine learning?

The data is available at MAST:
\dataset[10.17909/T9XG63]{\doi{10.17909/T9XG63}}.

Subject keyword examples:
AGN host galaxies, Exoplanet astronomy

7. Classifications
a) Utilize established classifications as available. b) Define new classifications clearly.

8. Orbital parameters
a) Avoid using "longitude of periapsis" in place of "argument of periapsis". b) Be explicit about which body's orbit a longitude or argument of periapsis refers to. c) Include time of periapsis as appropriate.

9. Tables
a) Provide a clear title and unambiguous labels for columns. b) Explain the contents of each column, including symbols and flags. c) Use the same explicitly defined null for missing values throughout. d) Prepare standard ReadMe files for machine-readable tables.

10. Figures
a) Provide clear caption, legend and axis labels for each figure. b) Design the graphics to be accessible. c) Make public data files used to create the figures.

11. Data archiving and access
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. c) Provide a complete list of metadata. d) Include a Data Availability Statement if required. 3) Do not publish data sets at URLs lacking long-term support.

12. Literature citations
a) Cite the original references. b) Use preferred citations by the authors. c) Provide full provenance of the data. Credit the originator of archival data. d) Include all references in the bibliography section. e) Distinguish original data in your article and data from other work.

13. Facility credits
a) Indicate the facilities involved, such as telescopes, instruments, and databases. b) Use standard keywords when possible. c) Indicate facility's own statement if available.

14. Software credits
a) List the software and version used in the production of the article.

15. Digital object identifiers (DOIs)
a) Use DOIs to cite data sets, software and services if available.

16. Data content keywords
a) Tag articles with relevant data content keywords from the UAT.