



APRENDE Data management

KICK-OFF MEETING

Data management

- Task 7.3: Data management - Start: M1, End: M48. Lead: CIEMAT, Other participants: LGI
 - This task focuses on the guiding principle for data management.
 - The Data Management Plan (DMP) D7.1 and its updated version D7.2 will give an overview of what data will be gathered and processed in the project, accordingly to the EC FAIR Data Management principle making data findable, accessible, interoperable and reusable

Data Summary

- Will you re-use any existing data and what will you re-use it for? State the reasons if re-use of any existing data has been considered but discarded.
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

FAIR Data

- Making data findable
 - Will data be identified by a persistent identifier?
 - What disciplinary or general standards will be followed?
- Making data accessible
 - Repository: will the data be deposited in a trusted repository?
 - Data: Will all data be made openly available? If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
 - Metadata: Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement?
- Making data interoperable
 - What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines?
- Increase data re-use
 - How will you provide documentation needed to validate data analysis and facilitate data re-use? Will your data be made freely available in the public domain to permit the widest re-use possible?

Types of data

All these types of data and metadata need to become publicly available at the end of APRENDE, either at the final public repositories or in the intermediate APRENDE repository (Web + Teams folders).

I. Experimental data

- **Differential data.** Ultimately available in EXFOR. **Metadata:** facility, setup, resolution(s), various uncertainties, background, normalisation, ...
- **Integral data.** Ultimately available in ICSBEP, IRPHE, SFCOMPO or SINBAD. **Metadata:** facility, setup, uncertainties, geometric description, ...
- **Raw data, correlated data.** Large files. Storage centres at the experimental facilities?

II. Evaluated data

- **Differential data.** Available in ENDF format at IAEA, NEA/OCDE and incorporated to evaluations (JEFF?). **Metadata:** data sets used, comments on the covariance analysis, normalisation, validation, explanations of why was done what...
- **Decay data.** Available in ENSDF and incorporated to ENDF (JEFF?). **Metadata:** data used, uncertainties, normalisations, efficiencies...
- **Sensitivity profiles.** NDAST database?

Types of data (cont)

Codes

- Software developments / codes / scripts promised as deliverables. Ultimately available in public software repositories (GitHub, NEA/OCDE, IAEA...)

Experimental setups

- Detailed description of the development and its performance. Typically, an OA publication / report / theses

Reports, theses, OA papers, OA proceedings, presentations

- Available in OA document servers.

Actions: identify which types of **data** and which **metadata** need to be provided. The complementary deliverables (reports) should have all the details for making the data usable in the future.

Other research outputs

- In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).
- Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

Allocation of resources

- What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?
- How will these be covered?

Data security

- What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?
- Will the data be safely stored in trusted repositories for long term preservation and curation?

Ethics

- Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).
- Will informed consent for data sharing and long-term preservation be included in questionnaires dealing with personal data?

THANK YOU!

Reach out for more information!

Contact us:

 6 Cité de l'Ameublement, 75011, Paris

 +33 (0)1 84 16 30 73

 contact@lgi-consulting.com

Follow us on:

 @LGI

 @LGI_Consulting

 @lgi_sustainableinnovation

