Fair for Fusion – Open Access for fusion data in Europe

- Non-(big)-experimental data
  - i.e. data not already considered in the project
  - consider modelling data in the context of this meeting

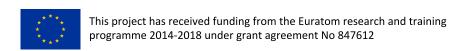David Coster, WP2 Leader, Fair for Fusion

Type of project: Coordination and Support Action
Overall budget: € 1 987 960
Project duration: Sept 2019-Aug 2021

Would like to increase FAIRness of data

- How interested is the data / code owner in increasing FAIRness of data?

- What resources are available or would be needed to improve FAIRness?

- How much support is there from

  - the research units to support this?

  - EUROfusion to support this?

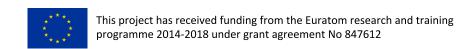 - other sources (local, national, EU) to support this?

Would like to increase FAIRness of data

- Need to identify what data is out there
  - volume (#cases, # files, # GB)
  - format of the data (ascii, netcdf, hdf5, MDSplus, own-binary, …)
  - persistence of the data (how long will the data be kept)
  - how well documented is the code and data?
  - is the code in a version control system (externally accessible)?
  - are PIDs (or similar) used?
  - how open should the data be?
  - licensing?
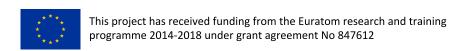 - is the data already using IDS's?

What can F4F do to help?

- Current idea is to use the summary IDS to
  - help people find data
  - give a quick (limited) look at key results
- Need to think about making the actual data and metadata available
  - Would examples help?
  - Would tools to capture the provenance help?
  - Would a data storage facility help?

Characterization of non-experimental data

- Task 2.3: (M12-18: MPIPP, CTH, CEA, NCSRD)
  - Identify the owners/producers of non-experimental data and formulate the requirements that would be needed in a future effort to bring their data into the developed open data framework. Here the challenges are likely to include: (1) a significantly larger number of contributors; (2) no well developed infrastructure for storing the data and producing metadata; (3) a much larger range of data formats which might not fit into a single data model.
- D2.4: (M18: MPIPP)
  - Report on what non-experimental data (e.g. from modelling) is being produced, and what steps would be necessary to bring it into the developed open data framework.