

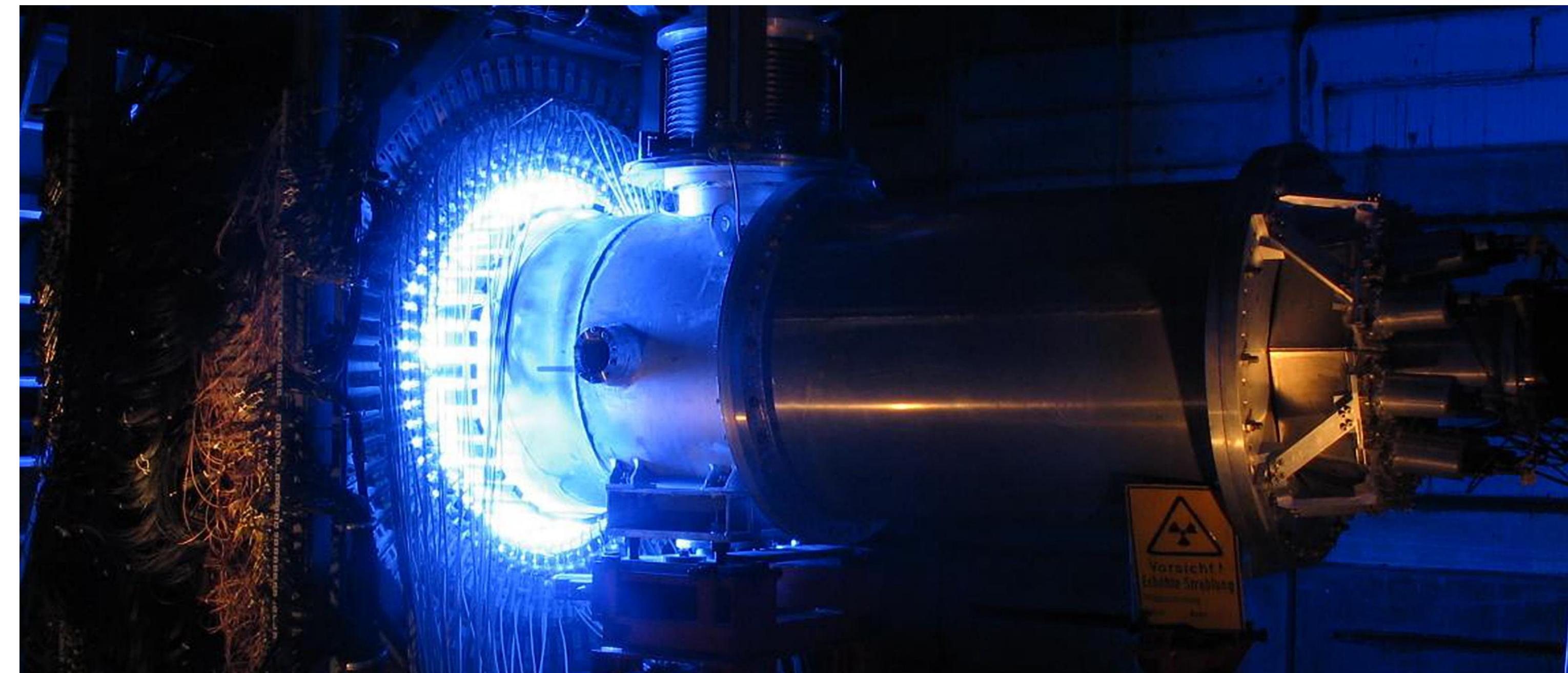
The A4 Experiment – from Data to Data Publication.

Gerrit Günther¹ // Boxing Gou² // Frank Maas³ // Oonagh Mannix¹ // Kilian Schwarz⁴ // Vivien Serve¹ // Malte C. Wilfert³

¹ Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin // ² Chinese Academy of Sciences, Lanzhou, 730000, China // ³ Johannes Gutenberg-Universität Mainz, Staudingerweg 18, D-55099 Mainz // ⁴ DESY, Notkestraße 85, 22607 Hamburg

1. Motivation

The particle physics experiment A4 at MAMI produced a stream of valuable data for many years which already released scientific output of high quality and still provides a solid basis for future publications. Here, we report from our approach to make the data hoard of a dismounted project sustainable and publishable according to the FAIR principles which, finally, would make the data intelligible to, both, humans and machines.



2. Starting Point: Prototypical Dataset

- Hierarchical folder structure creates vague context
- Heterogenous file formats
- Minimal metadata provided
- Keeping structure and naming for convenience

3. Convert & Enrich Metadata

- Convert to machine-readable xml format
- Add metadata through json configuration file (e.g. column header and units)
- RDF annotations provide machine-readable definitions (e.g. using ontology qudt)

4. Make Data Interoperable

- Implement customized code in ROOT files (object serialization)
- Version control of entire ROOT software system through container technology (e.g. Docker)
- Work in progress

5. Select Content: e-Lab Notebook

- Creates valuable context of higher level, especially between datasets
- Use of white-/blacklist to select published content (e.g. removing personal data)
- Provenance information: make changes visible (e.g. indicate where content is removed)

6. Create Context: Readme File

- Central file to create context, add bibliographic data, and, potentially, instruct how to read data
- Human- (html) and machine-readable (xml, DataCite)