Open Data
– a personal view –

Uli Katz
Big Data Workshop, Aachen
20 February 2018

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

The next 12+3 minutes

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- The challenge
- Scientific objectives
- Implementation
- Implications for science organisation and governance

Disclaimer:
Personal views …
More questions than answers …
KM3NeT background …

The challenge

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- Political requirement: ”Make data public”
  - Imposed by funding agencies, national governments, EU, …
    But (hopefully) also by science !
  - Requires work and resources
  - **How can it be done/organised to the best of science?**
- Who will use the open data, and for what?
  - Depends critically on content, access method and reliability of data
  - Important: Documentation and support
  - **Can we define common standards and generate synergies?**
- Activities in different contexts, at different places
  - **How can we coordinate them?**
    Is a loose, temporary project like ASTERICS enough?

Scientific objectives

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- Clearly multi-wavelength/multi-messenger studies
  - Cooperation of collaborations (MoUs, common analyses, …)

    → Is this a target of Open Data?

    → Do we expect the “end of proprietary data”?
  - Studies by individual astrophysicists, e.g. theorists

    → Can we provide usable, sustainable environment(s)?
- Use of ancillary data (atmospheric, deep-sea, …)
  - Who are the users and what are their requirements?
  - (How) can we unite efforts?
- Big data applications?
- Science outreach
  - How can we profit from open data?

Implementation

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- FAIR data: findable, accessible, interoperable, re-usable
  - Required for EU-funded projects (very reasonable!)
  - Imposes a certain complexity & cooperation/coordination
- What level of data to be publicised? Raw, high-level, meta?
- What about simulation data?
- New vs. existing tools, data formats, storage, access, …?
  - Virtual Observatory, eCOMMONS, … (ASTERICS!)
  - Data processing, storage, access (data centres → A. Haungs' talk)
  - Support & documentation ?
- Resources
  - Substantial in terms of HR and computing.
  - What is needed for a sustainable open data culture?
  - Who pays for it?

Implications

- If taken to the extreme, Open Data will change the way astroparticle physics is organised, performed and planned
  - Separation of RI construction and usage → Different communities
  - From Collaboration to Consortium (cf. ERIC contracts)
  - Similar to astrophysics, where this is already reality (since long)
  - **Do we want this? How can we shape this process?**
- Are our experiments suited for this development?
  If not, what are the alternatives?
- So, what does Open Data mean for (each of) us?