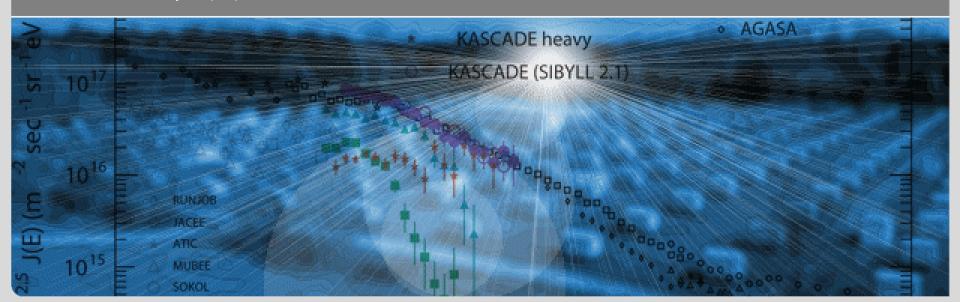


KCDC – publishing research data from the KASCADE experiment

Benjamin Fuchs

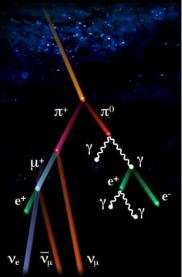


Institute for Nuclear Physics (IKP)



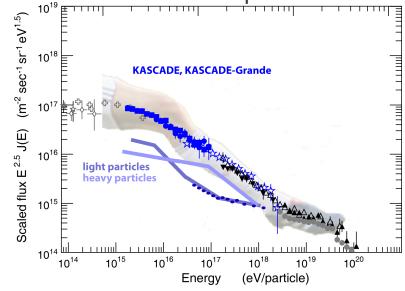
KASCADE - KArlsruhe Shower Core and Array DEtector





- High energy cosmic ray measurement via the detection of extensive air showers
- Experiment decommissioned in 2012 after 16 years of successful data acquisition
- Investigation of sources, acceleration and propagation of cosmic rays
- Investigation of the knee feature in the energy spectrum
- Search for anisotropies in arrival directions
- Investigation of hadronic interactions in the atmosphere





KASCADE - KArlsruhe Shower Core and Array DEtector



- World-wide collaboration of 11 institutions from 7 countries
- 50 collaborative publications over 16 years (amongst others in Nature, PRL, ...)
- 53 Ph.D. thesis
- 83 diploma thesis
- 450mio events for first data release

Open Data and KCDC



KASCADE was financed by taxes like most modern science experiments in physics

→ collected KASCADE data is public property

Idea of Open Data:

- The access and reuse of the data should be free
- Computerized processing of the data has to be possible

Helmholtz Open Access Workshop, DESY Hamburg

The data has to be usable today and in the future

Open Data – the concept



Berliner Erklärung on Open Access

Our mission of disseminating knowledge is only half complete if the **information** is not made **widely and readily available to society**. New possibilities of knowledge dissemination not only through the classical form but also and increasingly through **the open access paradigm via the Internet** have to be supported. We define open access as a comprehensive source of human knowledge and cultural heritage that has been approved by the scientific community.

In order to realize the vision of a global and accessible representation of knowledge, the future Web has to be sustainable, interactive, and transparent. Content and software tools must be openly accessible and compatible.

Open Data – challenges



- Making specialist knowledge available and usable for a general audience (also non-scientists)
- Providing interactive access to data, to analysis software and to meta-information
- Preserving the accumulated data and knowledge for future generations
- Open Data compliant licences for publishing scientific data

KCDC - what is it?



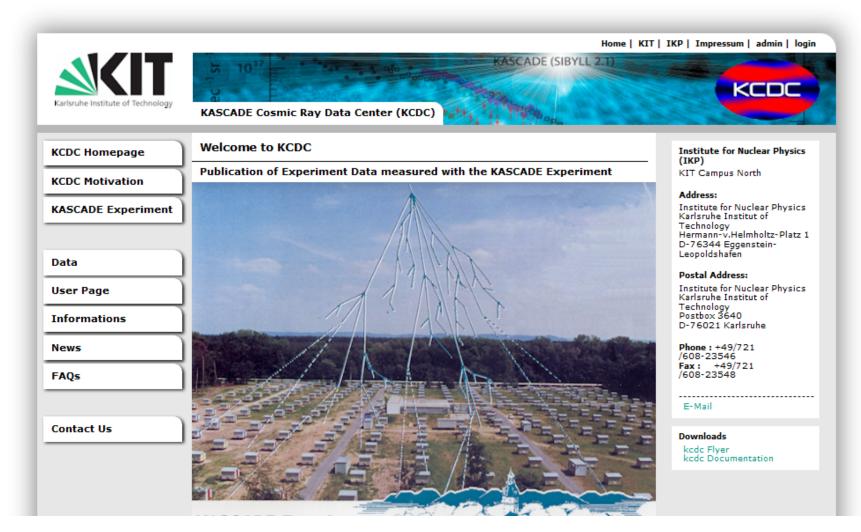
■Web portal → providing a modern software solution for publishing KASCADE data for a general audience

KCDC as a web portal offers:

- Software as a service without installation by user
- world wide access via web browser to data and associated processing facilities
- Consistent presentation of data and meta information on e.g. the detector, available data, example analysis, ...
- In a second step: Release the software as Open Source for free use by other experiments

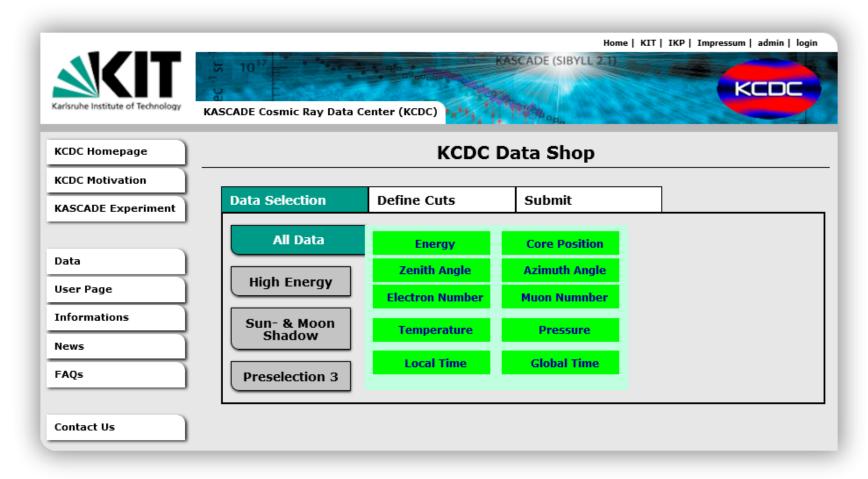
KCDC - Frontpage





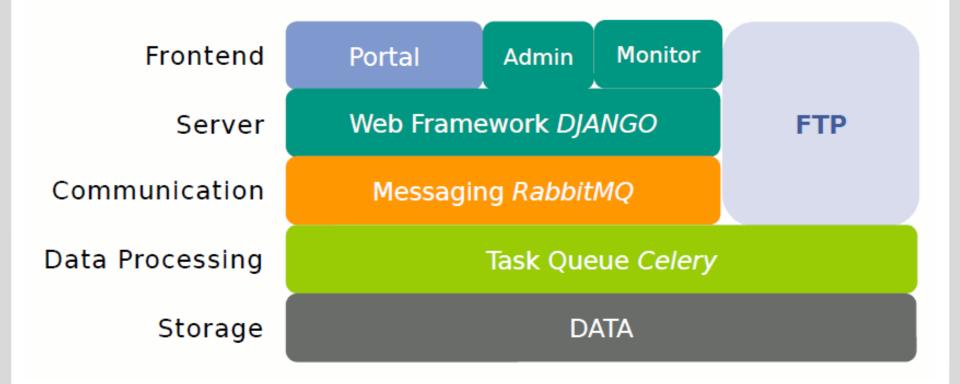
KCDC – data access and retrieval system





KCDC – Information flow





KCDC – reaching non-scientists

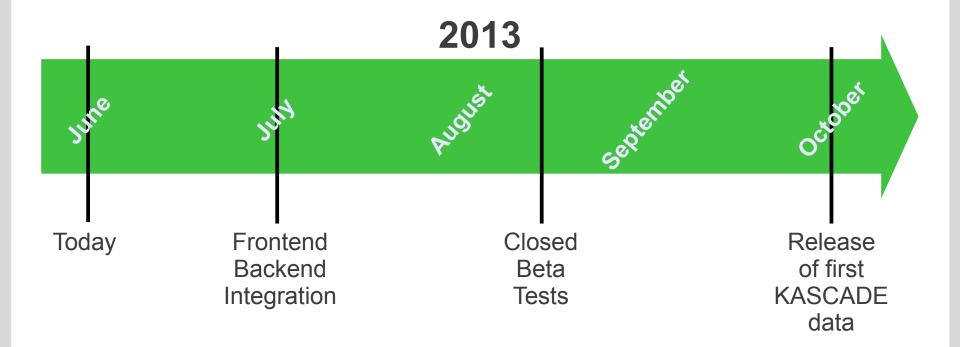


KCDC will provide different information levels:

- Meta-Information online and as PDF download (e.g. on the experiment, calibration, data, ...)
- Detailed description for example data analyses (e.g. as general tutorials and as Python/C++ scripts)
- Special data selections for dedicated analyses with accompanying descriptions
- Special data analysis tutorials for the use in schools are also planned

KCDC – The near future





1st Release: End of October Web portal online, first 450mio events are public accessible

2nd Release: Beginning 2014 Publishing KCDC software as open source

KCDC – legal issues and intellectual property



Open Questions:

- Which licence to choose?
- Can Open Access licences be applied?
- How to handle the redistribution of data?
- ■How to publish the software under Open Source?

How to address legal issues inside a web portal:

- The user has to agree explicitly to the licence
- Shipping data includes the licence

KCDC – Outlook



- Release of 450mio KASCADE events (October 2013)
- Release of the software beginning 2014
- KCDC is a software as a service solution for Open Data compliant publication of scientific data
 - Consistent presentation of information and data
 - Meta-information and step-by-step analysis tutorials

Helmholtz Open Access Workshop, DESY Hamburg

- Simple data access via web browser
- Database and computing cluster interfaces for data processing
- Open Data compliant licences

14