

Curation in ACTRIS

Context of Curation in ACTRIS

Complete ACTRIS report on Curation available at: <https://envriplus.manageprojects.com/projects/requirements/notebooks/470/pages/36/comments/389/attachments/607/download>

Summary of ACTRIS requirements for curation

Detailed requirements

Background

The ACTRIS Data Centre consists of three topical data repositories archiving the measurement data, which are all linked through the ACTRIS data portal to provide a single access point to all data. Hence, the ACTRIS Data Centre is founded on 3 topical data repositories:

- Near-surface aerosol and trace gas data are reported to **EBAS**
- Aerosol profile data are reported to the **EARLINET** Data base
 - **LIDAR** data (acronym for Light Detection And Ranging), which is a **remote sensing** technology that measures distance by illuminating a target with a **laser** and analyzing the reflected light.
- Cloud profile data are reported to the **CLOUDNET** data base

The ACTRIS data center is shared among the 3 components. However, the quality insurance will be share among the project soon.

There are a lot of curation activities in the database. The operational database contains always the full dataset: Database + Dataset + Operation environment. It is important to mention that historical data is still accessible through the database.

The policy on retaining/discarding datasets, software, etc ... depends on the component we are referring to. For example, **EARLINET** if a new version of a previous data arrives, both versions are stored (new and old), but by the website, only the new version is accessible. However, by using the database directly (special rights), both versions are available. ACTRIS uses internally a svn repository, so they could reproduce everything at every stated (based on released). They also have a back track system for error. Due to the new EU requirements, they archive all previous documents (probably for 10 years).

ACTRIS policy for having data accessible for long term is keep the system running. As soon as they have the system running, data will be accessible. Recently, ACTRIS got findings to have their system running for several years. The ideal situation for ACTRIS is to become an ERIC. We have to consider that the datacenters are currently funded by the project

For tracking data curation activities, ACTIS uses a log system associated with the elements in the database, at least for Cloudnet and Near surface components (not so sure about Earlinet)

ACTRIS plans to use OAI-PMH [1] and ISO9115 wmo profile as metadata standards. It allows for discovering metadata, privacy, data policy, owner of the data, rights.

ACTRIS preserves the software it by using back tracking system and svn repository. For curating data, ACTRIS has a central processing system with have different levels, each of them to process data to the next level. For each data level, each data process will have a version, with a specific description. And it curates the operating environment by scheduling backups for the 3 repositories/databases. Finally, ACTRIS considers necessary to have a portal for data submission, that provides automatically immediately feedback and outliers.

References:

1. <https://www.openarchives.org/pmh/>

Formalities (who & when)

Go-between	(rosafilgueira)
RI representative	Markus Fiebig and Paul Eckhardt
Period of requirements collection	November 2015
Status	Finished