ARCHIVER: archiving and preservation of large scientific datasets

Matthew Addis, Arkivum
matthew.addis@arkivum.com
orcid.org/0000-0002-3837-2526

In this lightning talk, we would like to present recent developments by Arkivum in the ARCHIVER project where we have been working on new approaches to very large-scale digital preservation of scientific datasets. The European Commission supported ARCHIVER project (Archiving and Preservation for Research Environments) aims to "introduce significant improvements in the area of archiving and digital preservation services, supporting the IT requirements of European scientists and providing end-to-end archival and preservation services, cost-effective for data generated in the petabyte range with high, sustained ingest rates, in the context of scientific research projects". Arkivum, along with other contractors in the ARCHIVER project, have been building prototype LTDP solutions in Phase 2 of the project. Initial results of a working prototype for ARCHIVER are already available. Further results of Phase 2 will be ready in time for presentation at iPRES. The focus of the talk will be to highlight how new cloud computing models such as serverless computing allows for LTDP systems to be constructed and operated in a way that is economically and environmentally sustainable as well as to preserve and provide access to very large datasets – not just for scientific data, but for all forms of digital content.

Notes:

Our understanding is that other ARCHIVER contractors will also be submitting lightning talks with a view for them to be combined into a joint session including a very short presentation from CERN. We would be very happy to support this approach and to adapt the lightning talk to best fit whatever the iPRES committee thinks would work well for a combined session.

Arkivum has submitted a full paper proposal about our work in ARCHIVER (submission 45). This submission for a lightning talk is in addition to the paper submission.