

Federated and Intelligent Datacubes

C. Atherton (GÉANT Association), Peter Löwe (Deutsches Institut für Wirtschaftsforschung e.V. (DIW Berlin)), and Torsten Heinen (Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR))

Background

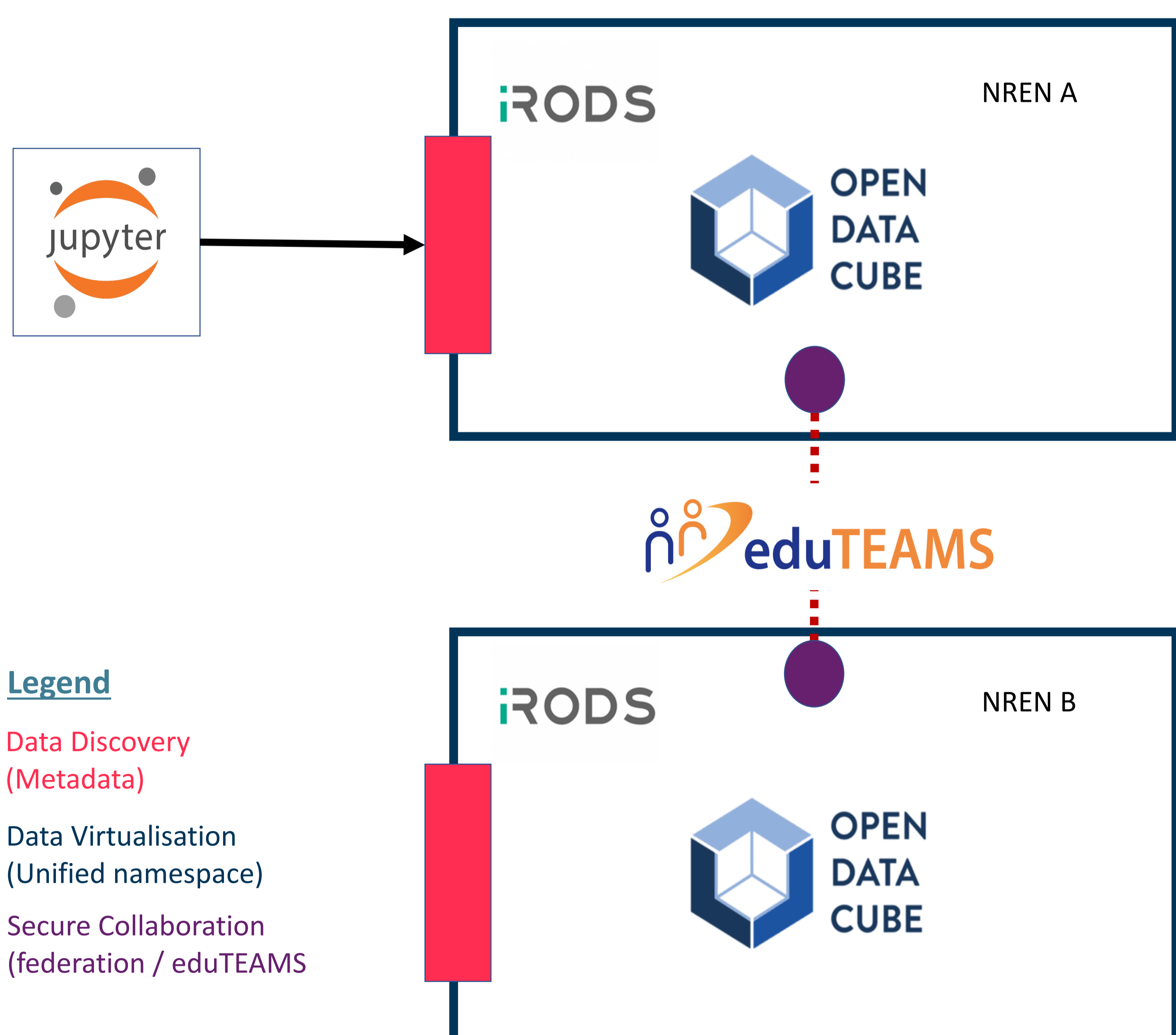
- Societal challenges are quickly threatening our existing way of life.
- Leveraging large data sets ('Big Data') for research is seen as the way to understand and move towards addressing societal challenges.
- Society needs research to accelerate but there are huge challenges in utilising big data sets at scale for individual researchers, as well as a high barriers of entry for new entrants into big data research fields.
- Datacubes have emerged as a technology for storing pre-processed imagery data which makes it easier and quicker for research to be done at scale.
- Datacubes are currently limited in number, size and scope due to financial and knowledge constraints.

Problems to Solve

- How do you scale up the use of Datacubes on a continental scale?
- How do you maintain redundancy and resiliency across datacubes?
- How to you provide access to nationally sensitive or sovereign data sets while still conforming to FAIR principles?

Proposed Solution

- National and Research Education Networks (NRENs) provide computer network and Authentication and Authorization infrastructures to research and education communities at a national level. In some case they also provide storage, cloud and HPC infrastructures.
- NRENs are starting to deploy Datacubes as a national infrastructure (Armenia).
- NRENs have an established working and trust relationship built up over many years to provide pan-continental services.
- iRODS (integrated Rule Orientated Data System) could be used as an abstraction layer between a nationally hosted Datacube and underlying existing infrastructure (namespace), and a student/researcher using a web accessible compute platform such as a jupyter notebook.
- iRODS allows trusted namespaces to be federated, allowing for the sharing of . This could be achieved using eduTEAMS AAI to allow scalable addition of other national Datacubes as well as definition of trusted groups and roles.
- iRODS can be used to broker access between namespaces, allowing assigned read and write permissions to be allocated to authenticated users from another trusted namespace.



Legend

Data Discovery
(Metadata)

Data Virtualisation
(Unified namespace)

Secure Collaboration
(federation / eduTEAMS)