Vision

Network and services

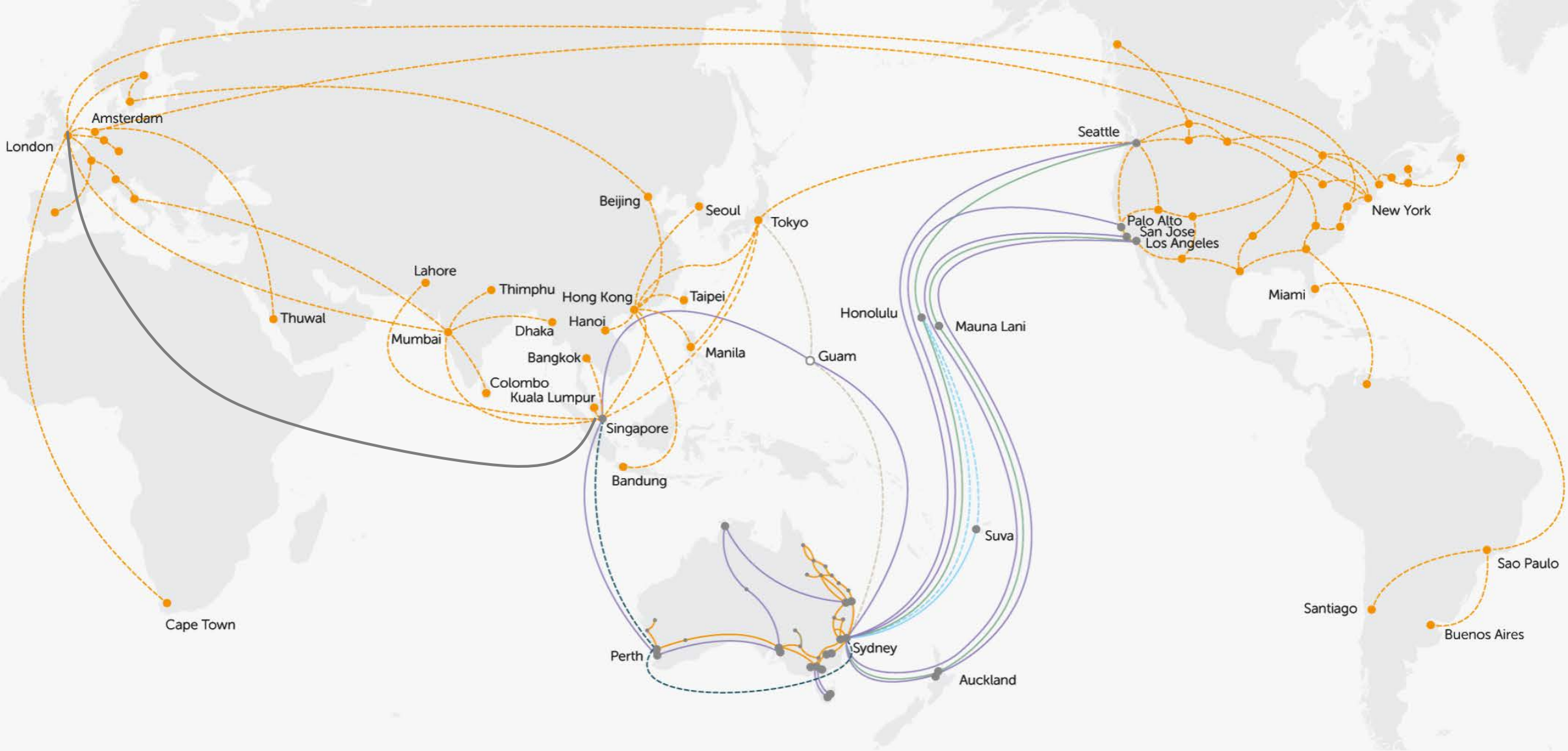
Tools and standards

Projects

AARNET'S VISION IS OF A GLOBALLY NETWORKED DATA-SHARING ECOSYSTEM

Accelerating knowledge creation and innovation to benefit future generations of Australians.

aarnet Connect. Collaborate. Innovate.


Amsterdam
London
Beijing
Seoul
Tokyo
Seattle
New York
Palo Alto
San Jose
Los Angeles
Lahore
Thimphu
Hong Kong
Taipei
Thuwal
Mumbai
Dhaka
Hanoi
Manila
Honolulu
Mauna Lani
Miami
Bangkok
Guam
Colombo
Kuala Lumpur
Singapore
Bandung
Suva
Sao Paulo
Santiago
Buenos Aires
Cape Town
Perth
Sydney
Auckland
AARNET POP
OTHER R&E POP
R&E NETWORK
5 GBPS
5 GBPS BACKUP
10 GBPS
100 GBPS
N* 100G WDM
INDIGO PROJECT
JGA PROJECT

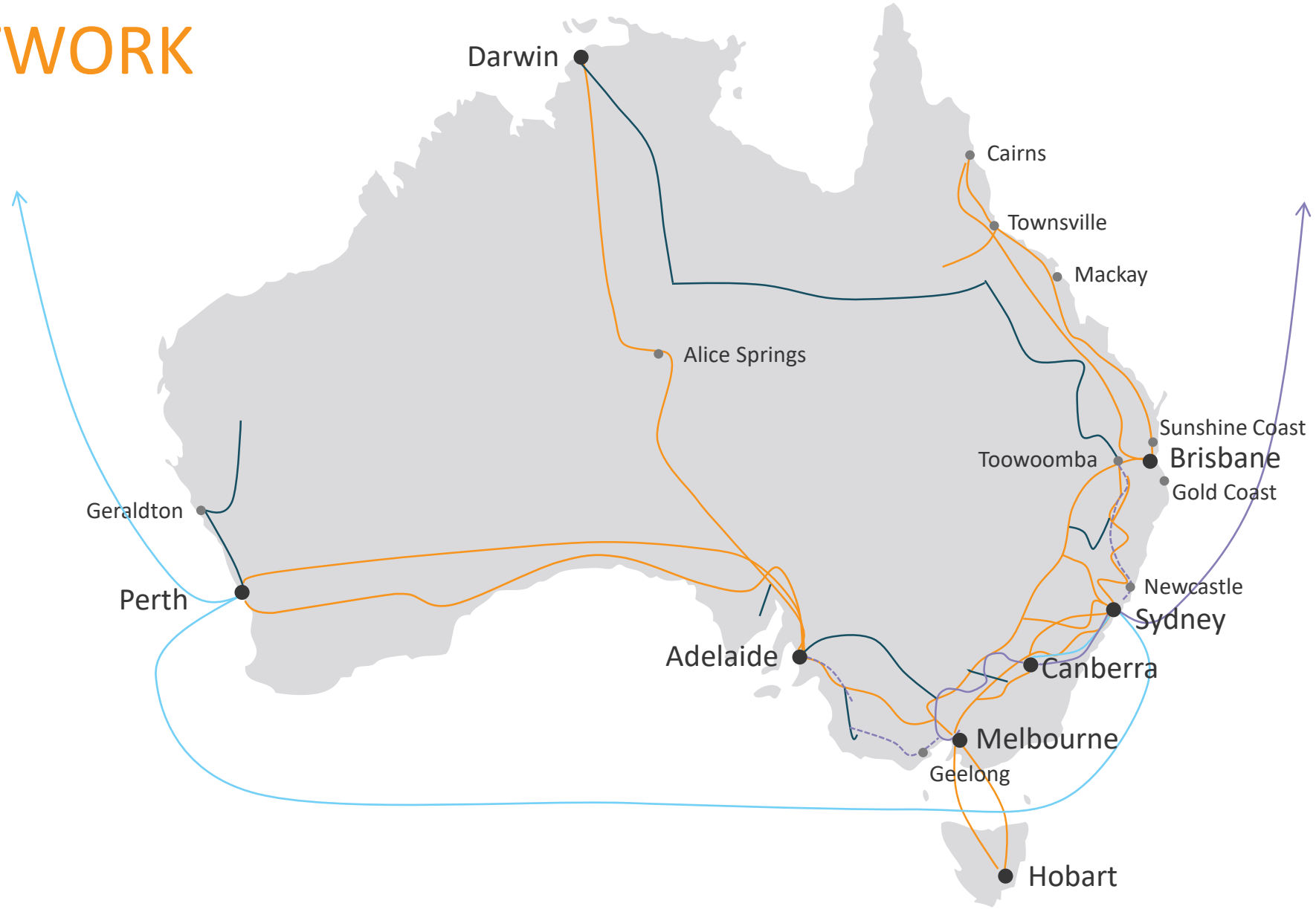
NATIONAL NETWORK

IRU's & Capacity Swaps

Blackspots (RBBP) & Tails

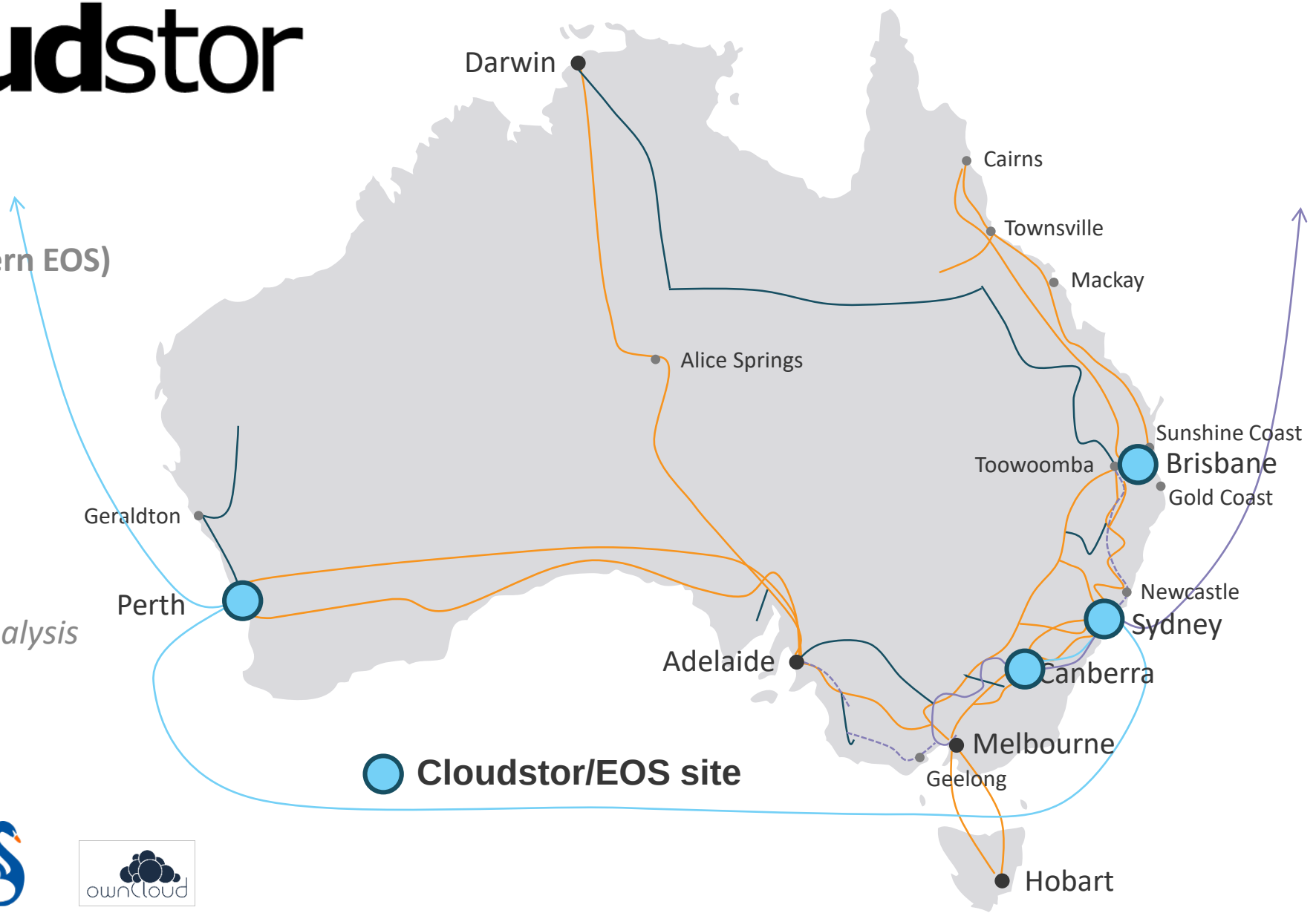
AARNet Network 2019

AARNet Network 2020+



cloudstor

- **Distributed object storage (Cern EOS)**
- **S3 gateways**
- **Cloudstor**
  - File sync & share
  - Collaboration platform
  - Collections packaging
  - SWAN
    Service for Web based Analysis
    (Jupyter Notebooks)



AARNET'S CLOUD SERVICE PORTFOLIO TODAY:

Peer

- Google
- Akamai

Connect

- aws
- Azure

Partner

- zoom
- Panopto

Build

- cloudstor
- eduroam

DIGITAL PRESERVATION

The actions to ensure that digital information of continuing value remains accessible and usable

The management of activities that will allow the data to be discovered, accessed, rendered, deemed reliable and re-used over many years and even decades.

F

TO BE FINDABLE:
Assign a persistent identifier, described with rich metadata, indexed in a search engine.

A

TO BE ACCESSIBLE:
Make the data open and online. If it can't be open then access conditions are clearly defined.

I

TO BE INTEROPERABLE:
Use community agreed formats, language and vocabularies. Link to related information using identifiers.

R

TO BE RE-USABLE:
Give the data a clear machine readable license. Data is associated with their provenance. Data meets standards to support re-use. Data richness is maintained.

ARCHIVER REQUIREMENTS

ARCHIVER LAYERS	DESCRIPTION	EXAMPLE SERVICES
LAYER4 Advanced services	High level services: visual representation of data, reproducibility of scientific analysis, deep learning; K8 and ML; Emulation	Jupyter, TensorFlow, Watson, GrayMeta, EaaSI
LAYER 3 Baseline user services	Search, discover, share, indexing, data removal. Access under AAF	OwnCloud, ElasticSearch, Solr, Omeka, IIIF, Repositories
LAYER 2 Preservation	OAIS conformant services: data readability, normalization, obsolescence monitoring, files fixity, authenticity checks, data packaging ISO 14721/16363, 26324 and related standards	Preservica, Archivematica, Bagger, Fixity tools, Arkisto
LAYER 1 Storage Basic archiving Secure backup	Data integrity/security; cloud/hybrid deployment; data volume in the PB range; high, sustained ingest data rates. IRAP and ISO certification: 27000, 27040, 19086 and related. Archives connected to AARNet network	EOS, Cern Tape Archive, Minio, Globus, Rucio, Veeam, Restic, StorNext

Adapted from CERN ARCHIVER-PROJECT archiver-project.eu

STORAGE

- Automate data packaging (Archival Information Package)
- Reliable long-term bit preservation
- Redundancy
- Geographically dispersed data storage model
- Exit strategy

digipres.org/tools

aarnet

DATA PACKAGING

- Application-independent
- Ensure objects remain fixed over time
- Robustness against errors, corruption, and migration
- Promote long-term access
- Human and machine readable
- Help rebuild a repository after disaster or obsolescence

ocfl.io

aarnet

FIXITY

Fixity, in the preservation sense, means the assurance that a digital file has remained unchanged, i.e. fixed.

Fixity of files can established and monitored through the use of checksums.

Checksums have three main uses:

- To know that a file has been correctly received from a content owner or source and then transferred successfully to preservation storage
- To know that file fixity has been maintained when that file is being stored.
- To be given to users of the file in the future so they know that the file has been correctly retrieved from storage and delivered to them

www.dpconline.org/handbook/technical-solutions-and-tools/fixity-and-checksums

CERN Tape Archive

- In production at CERN – 358 PB on tape now. Next LHC run will double size
- Not an HSM – it's an archive – mean recall time is under a minute
- Time is the critical dimension for archival storage
- Tape is still the cheapest storage, well known and formats understood
- Easy to mark immutable
- Easy to physically take offline

eoscta.docs.cern.ch/cta

EMULATION-AS-A-SERVICE INFRASTRUCTURE (EAASI)

Australian Research Council Linkage project *Play it Again: Preserving Australian videogame history of the 1990s*

Lead by Professor Melanie Swalwell, Swinburne University of Technology

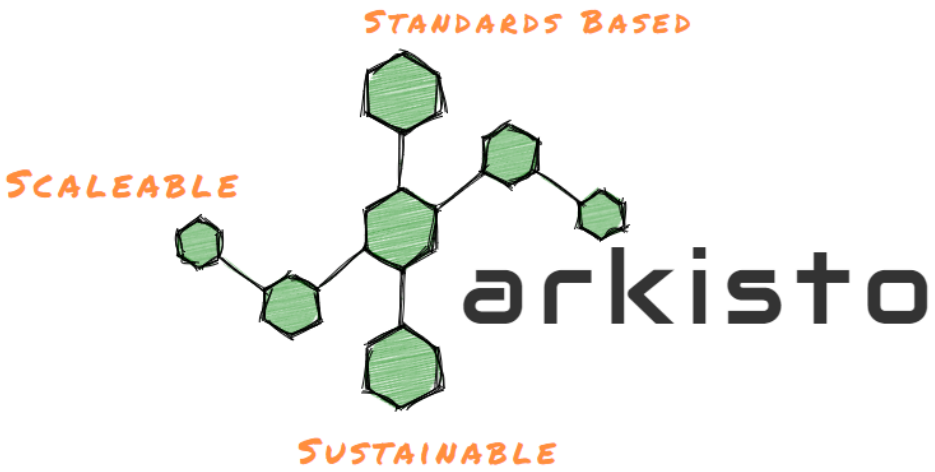
Collaborators include ACMI, RMIT and AARNet.

EaaSI is a major component

softwarepreservationnetwork.org, tmt.edu.au/projects-2-2/play-it-again/

aarnet

ARKISTO: A SCALABLE, STANDARDS BASED PLATFORM FOR SUSTAINABLE DATA



- Builds pathways from existing data formats to capture encoded knowledge
- Makes this knowledge explicit by keeping metadata together with the data
- Built on top of RO-Crate and OCFL
- Long-term preservability of well-described data is *always* the first consideration
- Data on an Arkisto deployment is always available on disc (or object storage) with a complete description *independently* of any services such as websites or APIs

arkisto-platform.github.io

aarnet

AUSTRALASIA PRESERVES

Community of practice est. 2018

#AusPreserves
#DigitalPreservation
#DigiPres

www.australasiapreserves.org

Illustration: Matthew Burgess, SLNSW

aarnet

DIGITAL PRESERVATION COALITION

Our worldwide membership is working towards a secure digital legacy, through:

Community Engagement

Advocacy

Workforce Development

Capacity Building

Good Practice and Standards

Management and Governance

info@dpconline.org // www.dpconline.org // @dpc_chat

DigitalPreservationCoalition

Adam.Bell@aarnet.edu.au