



ELAR

Endangered Languages Archive

**Part of Hans Rausing Endangered
Languages Project**

School of Oriental and African Studies

University of London



Overview

- **Funding ~£900K over 10 years**
- **Staff**
- **Scope of operations**
 - *data preparation and preservation*
 - *advice and training*
 - *contributions to academic developments in documentation, archiving etc*
 - *mobilisation, projects and frameworks*



Storage

- *current: 4TB FC SAN array, 1 media/web server*
- *to be backed up by 1 x 2 LTO2 drive tape lib (8~14TB)*
- *SAN fully replicated at Vernon Sq campus of SOAS*
- *currently in various stages of installation, dependencies on Dell, SOAS IT*
- *agreement with OTA for parallel accession of ELAR data, with storage in OUCS HFS and offsite; to serve as Disaster Recovery; also joint work proposed on data and cataloguing encoding methods*



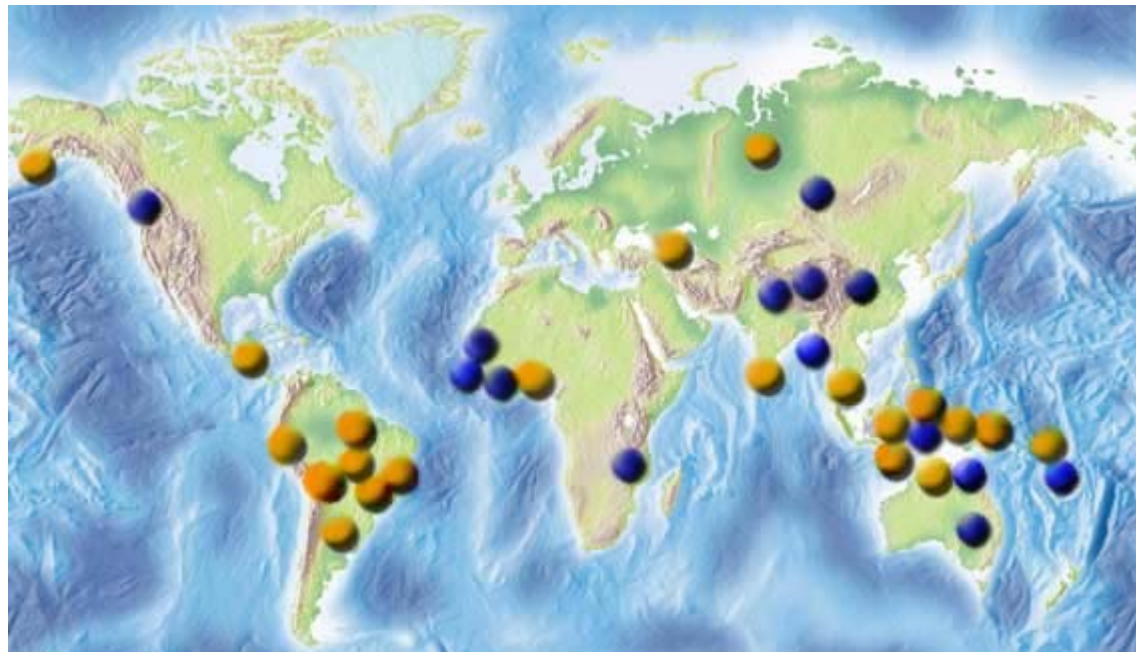
Expected deposits

within 12 months:

- *1st cohort of ELDP grantees, together with incremental archiving of 2nd, 3rd cohort*
- *each project total data estimate (early rough estimate, average – depends on size and emphasis of project)*
- *at 20 projects per year, this represents a total of approx 0.7 TB per year*
- *together with other (non-ELDP) deposits, we estimate ingesting about 1TB per year.*



ELDP projects as at Jan 2005





Policy

Version: Draft 0.1 Do not distribute This document expresses ELAR's aims, values and priorities. It should form the basis of ELAR's documents and procedures

1–2	Collection
3–4	Access
5–6	Technologies
7–8	Mobilisation/repurposing
9–10	Relationships and services

- **1. Collection.** Build and grow a collection of digital endangered languages materials, focussed on documentation material and output from ELDP funded projects. Deposits will be accepted according to published criteria, and will not be removed except under extraordinary circumstances.
- **2. Storage and management.** Store and manage deposited materials to ensure integrity, preservation, future accessibility, and acceptable usage.
- **3. Ownership and protocol.** Respect and implement copyright, intellectual property, sensitivities, protocols and restrictions placed on deposited materials, and encourage ethical practice in collecting such protocol information.
- **4. Access.** Provide and mediate access to materials through relationships between language providers, language communities, depositors, other stakeholders, ELAR, and users. These relationships to be ongoing and managed through ICT and formal agreements. Language providers, their families, and language communities have priority for access.



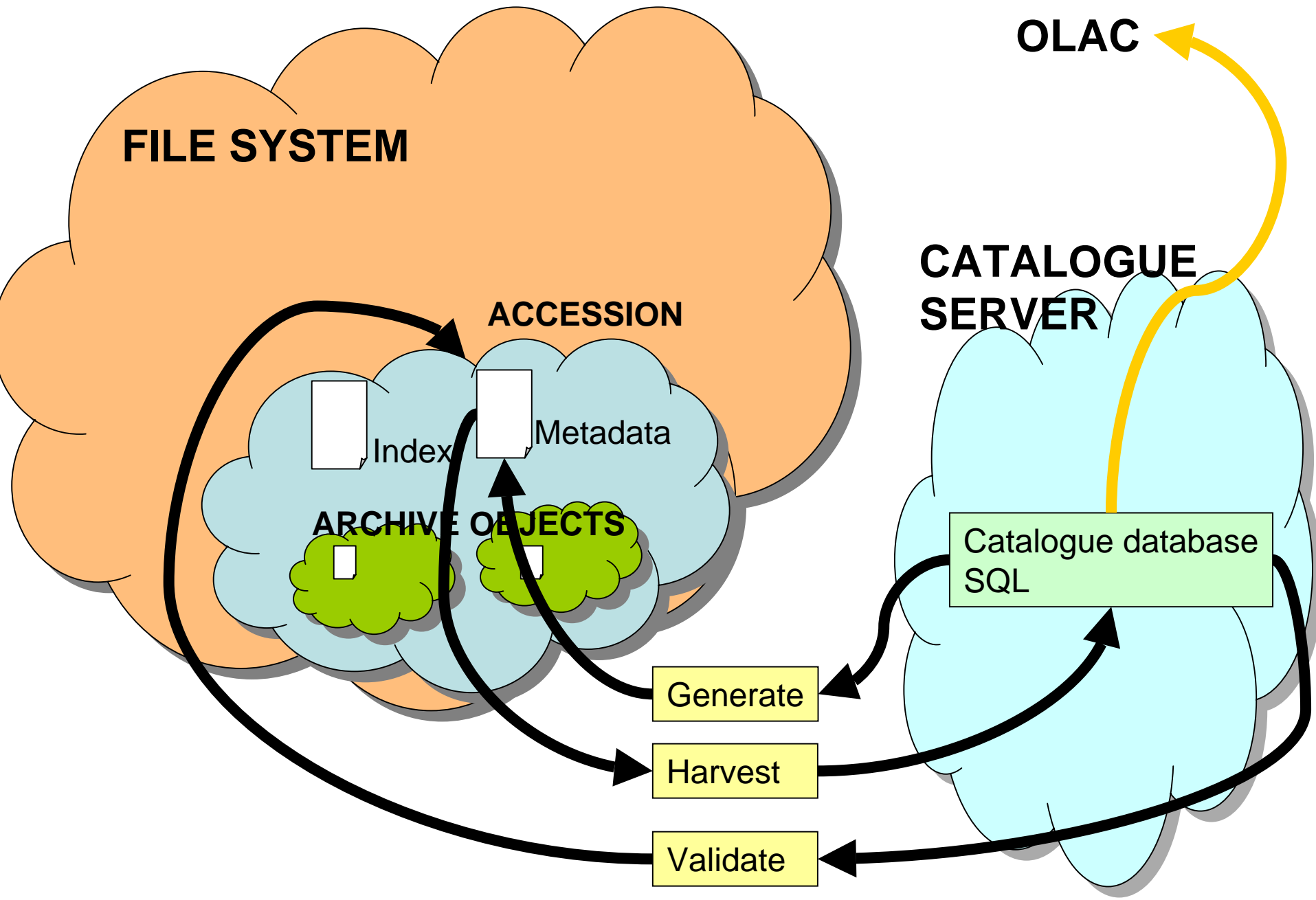
Policy

- **5. Standards. Use, contribute to and encourage use of accepted standards in electronic data management and associated processes**
- **6. Keeping up to date. Keep up to date with changing technologies, linguistic practices, and roles of digital archives**
- **7. No prejudice to usage. Not prevent or reduce capability of materials to be used in support of language strengthening or maintenance**
- **8. Mobilisation. Play an active role in mobilising archive materials for use in support of endangered languages**
- **9. Collaboration. Develop and maintain partnerships with other archives and institutions to strengthen long term preservation of materials and archive practices**
- **10. Services. Provide advice and other services to depositors**



Ingestion

- **Assume majority of data comes as digital data**
- **Limited digitisation will be available**
 - *provision of equipment*
 - *technician services*
 - *paid services*
 - *investigate student employment*
- **Ingestion and accession systems yet to be created – some basic outline: see cloud diag.**





Metadata

- **We will favour but not require OLAC metadata**
- **Goal will be**
 - *ensuring best possible metadata informational content supplied and maintained by depositors*
 - *encouraging good practice*
 - *monitoring changes in understanding and practice*
 - *flexible systems to maintain and implement metadata content*



Access and dissemination

- **Clear distinction between preservation and access**
- **Catalogue to be open access on www**
- **All resources by default have restricted access**
- **Conditions for access to reflect a channel of information flow that starts with recording events and should be maintained up to date if possible**



Other activities

- **Training: 2 x grantee training, 2 x multimedia training, lectures on data management, XML etc**
- **Multimedia Development – *Disappearing Sounds of the World's Endangered Languages, Hearing Voices***
- **Exhibition – *Endangered Voices***
- **Academic papers and publishing, incl LAN, journal and book articles, conferences etc**
- **HRELP website**
- **Services to ELDP etc**