

ÖAW

AUSTRIAN
ACADEMY OF
SCIENCES



ARCHE



- ▶ ARCHE = A Resource Centre for the HumanitiEs
- ▶ A digital archive for the humanities
- ▶ Implements the OAIS Reference Model for an Open Archival Information System
- ▶ arche.acdh.oeaw.ac.at

WHAT IS AN ARCHIVE?

- ▶ Preserves information for access and use
- ▶ Committed to long-term preservation!
- ▶ Should adhere to the OAIS reference model (Open Archival Information System)

WHAT IS AN ARCHIVE?

- ▶ Preserves information for access and use
- ▶ Committed to long-term preservation!
- ▶ Should adhere to the OAIS reference model (Open Archival Information System)
 - ▶ solid technical basis for storing and preserving
 - ▶ metadata schema for discovery and dissemination
 - ▶ user interface for search and reuse
 - ▶ policies and documentation for transparency

HISTORY

- ▶ ARCHE supersedes CLARIN Centre Vienna / Language Resources Portal (<https://clarin.oeaw.ac.at>)
- ▶ It was focused on digital language resources
- ▶ It was active from 2014-2017
- ▶ Data from LRP was migrated to ARCHE
- ▶ ARCHE is designed for a broader spectrum of data from all humanities fields

- ▶ ARCHE preserves
 - ▶ Plain text, text documents, PDF documents, images, spreadsheets, databases, audio, video, 3d data, GIS and much more

- ▶ ARCHE preserves
 - ▶ Plain text, text documents, PDF documents, images, spreadsheets, databases, audio, video, 3d data, GIS and much more
- ▶ ... as long as files are provided in formats suitable for long-term preservation

METADATA

- ▶ Why?
 - ▶ understand, discover and share data
 - ▶ How? Who? What?
 - ▶ FAIR Data Principles

METADATA

- ▶ Why?
 - ▶ understand, discover and share data
 - ▶ How? Who? What?
 - ▶ FAIR Data Principles
Findable, Accessible, Interoperable & Reusable

METADATA

- ▶ Why?
 - ▶ understand, discover and share data
 - ▶ How? Who? What?
 - ▶ FAIR Data Principles
Findable, Accessible, Interoperable & Reusable
- ▶ Challenges:
 - ▶ heterogeneity of data
 - ▶ support multiple metadata schemas

METADATA

- ▶ Why?
 - ▶ understand, discover and share data
 - ▶ How? Who? What?
 - ▶ FAIR Data Principles
Findable, Accessible, Interoperable & Reusable
- ▶ Challenges:
 - ▶ heterogeneity of data
 - ▶ support multiple metadata schemas
- ▶ On GitHub: [acdh-oeaw/repo-schema](https://github.com/acdh-oeaw/repo-schema)

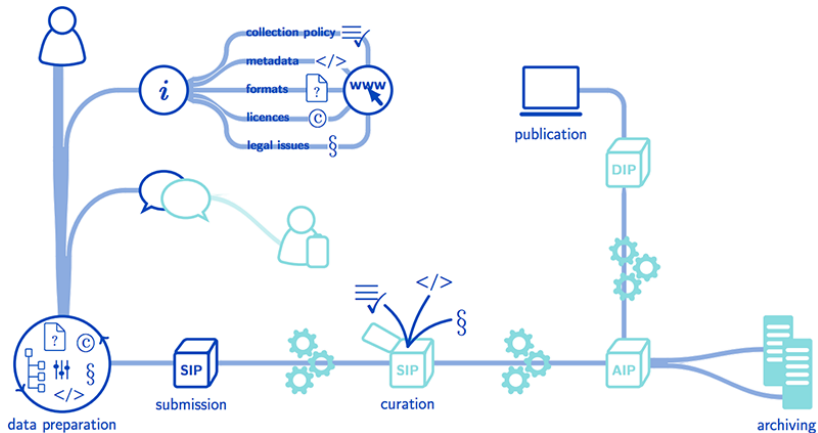
METADATA

- ▶ Metadata can be harvested via OAI-PMH.
- ▶ Resources in ARCHE can be disseminated in additional channels
- ▶ E.g. Virtual Language Observatory, CLARIN's central metadata catalogue

USER INTERFACE

- ▶ Live Demo!
- ▶ arche.acdh.oeaw.ac.at

DEPOSITION PROCESS



DEPOSITION PROCESS - EXAMPLE

- ▶ TuniCo
- ▶ <https://id.acdh.oeaw.ac.at/tunico-data>
- ▶ Collection of audio files with transcriptions in XML and a XML file with dictionary

DEPOSITION PROCESS - EXAMPLE

- ▶ TuniCo
- ▶ <https://id.acdh.oew.ac.at/tunico-data>
- ▶ Collection of audio files with transcriptions in XML and a XML file with dictionary
- ▶ Check file naming
- ▶ Check formats
- ▶ Enrich metadata
- ▶ Ensure no sensitive data is contained or consent was given
- ▶ Send it to ARCHE

MORE

- ▶ arche.acdh.oeaw.ac.at
- ▶ OAI: <https://public.ccsds.org/Pubs/650x0m2.pdf>
- ▶ www.oeaw.ac.at/acdh
- ▶ Twitter: [#acdhARCHE](https://twitter.com/acdhARCHE) and [@ACDH_OeAW](https://twitter.com/ACDH_OeAW)
- ▶ YouTube: [ACDH_OeAW](https://www.youtube.com/ACDH_OeAW)
- ▶ We have a video!
www.youtube.com/watch?v=mmRfc7fNNSA

TECHNICAL BASIS

- ▶ Uses Fedora Commons, version 4
- ▶ Blazegraph as external triplestore metadata storage
- ▶ Custom PHP components:
 - ▶ repo-php-util
 - ▶ doorkeeper
 - ▶ OAI-PMH endpoint
 - ▶ Browser based on Drupal
 - ▶ Data type specific dissemination services
- ▶ Drupal for Browsing with custom dissemination services
- ▶ On GitHub: [acdh-oeaw](#)

SYSTEM ARCHITECTURE

