

# Integrating domain repository for archaeological data into the National Data Infrastructure: a case study of fruitful synergy

David Novák

Institute of Archaeology of the Czech Academy of Sciences, Prague



**Archaeological  
Information System**  
of the Czech Republic



INSTITUTE OF ARCHAEOLOGY  
OF THE CAS PRAGUE

# Research Data Day & EOSC National Tripartite Event 2025

Integrating domain repository for archaeological data

- **AIS CR** (<https://www.aiscr.cz/>) = Archaeological Information System of the Czech Republic
- **large research infrastructure** listed in the national roadmap (<https://www.vyzkumne-infrastruktury.cz/>) approved by government
- consortium of the **two Institutes of Archaeology** of the Czech Academy of Sciences
- **centralised** team, data, services, hardware facilities, software applications, know-how, collaborations, standards and methodologies



**Archaeological  
Information System**  
of the Czech Republic



**LARGE RESEARCH  
INFRASTRUCTURES**



INSTITUTE OF ARCHAEOLOGY  
OF THE **CAS** PRAGUE



Institute of Archaeology  
Czech Acad Sci, Brno



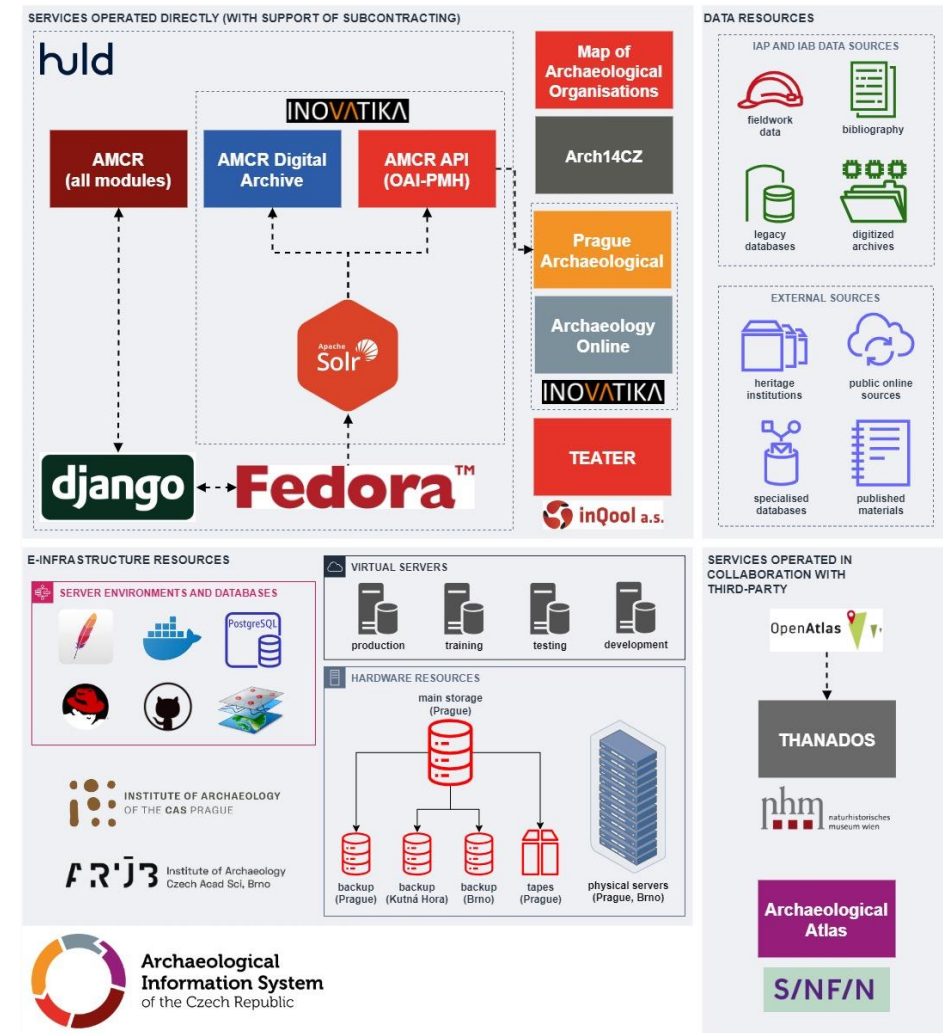
---

## What is AIS CR?

# Research Data Day & EOSC National Tripartite Event 2025

## Integrating domain repository for archaeological data

- Provides **digital workflow** to handle legal responsibilities in the field of archaeology
  - developers (announcing projects)
  - archaeologists (conducting fieldwork & reporting)
  - heritage managers (oversight and curation)
  - general public (data access and collaboration)
- **Archives data** on archaeological fieldwork
- Develops related **SW** and run needed **HW**
- Supports and represents the **community**
- Develops and implements **standards**
- Ensures long term **strategy and sustainability**
- **Integrates** disparate resources



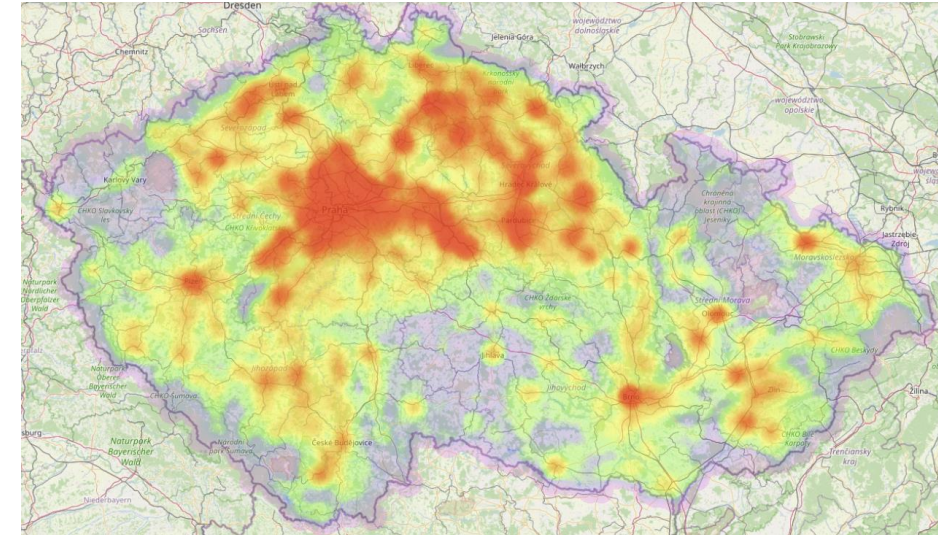
What AIS CR does?

# Research Data Day & EOSC National Tripartite Event 2025

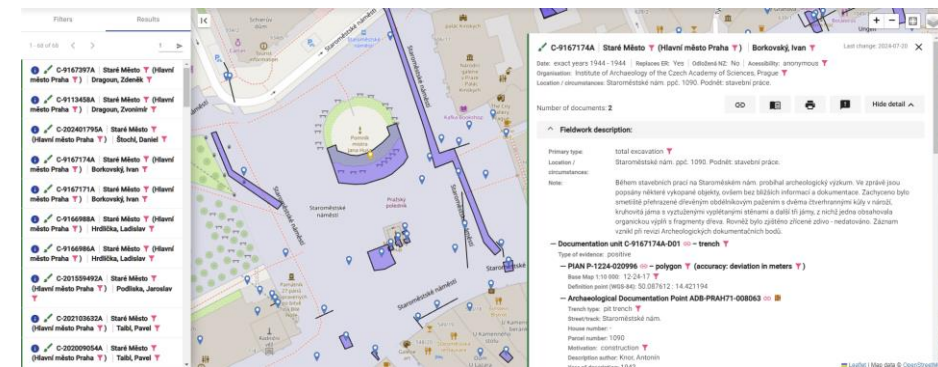
## Integrating domain repository for archaeological data

- **212k projects** as announced by developers since 2008
- **220k fieldwork events** (what, where, when data) recorded retrospectively since 19th century
- **11k systematically described sites** (deserted villages, castles, hillforts, barrow cemeteries etc.)
- **12k individual finds** recorded by amateur collaborators (AMCR-PAS system)
- **204k documents** (reports, photographs, maps...) on fieldwork events  
(about the same number digitised but still not published)
- **10k bibliographic resources** related to fieldwork events
- **40+** standardised controlled **vocabularies**
- **2.6k C14 dates** with detailed metadata

(and more – 3D models, specialised datasets, etc.)



<https://digiarchiv.aiscr.cz/>



What data we handle?



```
▼<OAI-PMH xmlns="http://www.openarchives.org/OAI/2.0/OAI-PMH.xsd">
  <responseDate>2024-08-28T16:25:29Z</responseDate>
  <request verb="GetRecord" metadata="http://www.openarchives.org/OAI/2.0/metadata.xsd">
    ▼<GetRecord>
      ▼<record>
        ▼<header>
          <identifier>https://api.ais.cz/api/ais/record/2024-07-20T05:50:00Z</identifier>
          <timestamp>2024-07-20T05:50:00Z</timestamp>
          <setSpec>archeologicky_zaznam</setSpec>
        </header>
        ▼<metadata>
          ▼<amcr:amcr xmlns:xsi="http://www.openarchives.org/OAI/2.0/metadata.xsd" xsi:schemaLocation="https://api.ais.cz/api/ais/record/2024-07-20T05:50:00Z https://api.ais.cz/api/ais/record/2024-07-20T05:50:00Z">
            ▼<amcr:archeologicky_zaznam>
              <amcr:ident_cely>C-9167</amcr:ident_cely>
              <amcr:stav>3</amcr:stav>
              <amcr:okres xml:lang="cs">Praha</amcr:okres>
              <amcr:pristupnost xml:lang="cs">Otevřen</amcr:pristupnost>
              ▼<amcr:chranene_udaje>
                <amcr:hlavní_katastr>1</amcr:hlavní_katastr>
              </amcr:chranene_udaje>
              ▼<amcr:akce>
                <amcr:typ>N</amcr:typ>
                <amcr:je_nz>true</amcr:je_nz>
                <amcr:odlozena_nz>false</amcr:odlozena_nz>
                <amcr:hlavní_vedoucí>1</amcr:hlavní_vedoucí>
                <amcr:organizace xml:lang="cs">AMCR</amcr:organizace>
                <amcr:specifikace_data>1</amcr:specifikace_data>
                <amcr:datum_zahajeni>2024-07-20T05:50:00Z</amcr:datum_zahajeni>
                <amcr:datum_ukonceni>2024-07-20T05:50:00Z</amcr:datum_ukonceni>
                <amcr:hlavní_typ xml:lang="cs">Přístup</amcr:hlavní_typ>
              </amcr:chranene_udaje>
              <amcr:lokalizace_okr>1</amcr:lokalizace_okr>
              <amcr:souhrn_upresneni>Zachyceno bylo smeti organickou výplň s f</amcr:souhrn_upresneni>
            </amcr:archeologicky_zaznam>
          </amcr:amcr>
        </metadata>
      </record>
    </GetRecord>
  </request>
</OAI-PMH>
```

## AMCR-PAS

The Portal of Amateur Collaborators and Register of Individual Finds is an AMCR module aimed at managing information about individual finds collected during a surface surveys or sampling. It forms the framework for effective cooperation between professional public and amateur collaborators.

WEB

## Archaeological Atlas of the CR

The Archaeological Atlas presents archaeological monuments in the landscape in a popular way to the general public. It is an accompanying service of the related printed publications, expanding their contents in a virtual environment. It offers full content in English.

WEB

## Prague- Archaeological (MADP)

A web portal with the task to provide access to archaeological sources on the Capital City of Prague for professionals and the public. The website enables to visit the underground of Prague and to follow in the footsteps of Prague's history.

WEB

## THANADOS (MEDCEM)

The service deals with the collection and presentation of medieval cemeteries. All information can be explored via a digital catalogue and within an interactive map. Cartographic visualisations as well as charts and plots are created dynamically based on real research data.

WEB

## Arch14CZ

The purpose of the Czech Archaeological Radiocarbon Database is to gather all radiocarbon (14C) dates from archaeological contexts in Czechia. The service offers continuously updated data also in machine readable format for download and harvesting.

WEB

## Thesaurus of Archaeological Terminology (TEATER)

TEATER makes archaeological terminology available via an web interface, with emphasis on library applications. It consists of standard subject terms, supplemented by additional terms and descriptions.

WEB

## Map of Archaeological Organisations

In the Czech Republic there are more than 100 organisations licensed to carry out archaeological fieldwork under the State Heritage Protection Act. This application provides an overview of these organisations, including the territorial jurisdictions.

WEB

## 3D Library

AMCR 3D Library serves as a space for the systematic recording of archaeologically relevant three-dimensional models and textures. Its aim is not to store individual models, but to allow for their uniform expert description and to ensure their searchability in the AMCR Digital Archive environment.

WEB

www.openarchives.org/OAI/2.0/OAI-

https://api.ais.cz/api/ais/record/2024-07-20T05:50:00Z https://api.ais.cz/api/ais/record/2024-07-20T05:50:00Z .2.1/gml.xsd">

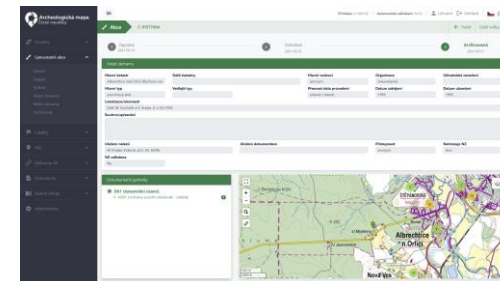
šších informací a dokumentace. ři jámy, z nichž jedna obsahovala \_upresneni>

# Research Data Day & EOSC National Tripartite Event 2025

## Integrating domain repository for archaeological data

### Archaeological Map of the CR (AMCR)

AMCR unites the archaeological information systems in the CR and offers tools to register fieldwork projects, collect information on fieldwork and their results, serves as digital repository and provides data for the AMCR Digital Archive and API.



<https://aiscr-webamcr.readthedocs.io/>



[https://github.com/ARUP-CAS/aiscr-webamcr/blob/dev/webclient/xml\\_generator/definitions/amcr.xsd](https://github.com/ARUP-CAS/aiscr-webamcr/blob/dev/webclient/xml_generator/definitions/amcr.xsd)



[https://github.com/ARUP-CAS/aiscr-webamcr/blob/dev/webclient/xml\\_generator/definitions/metadata.xml](https://github.com/ARUP-CAS/aiscr-webamcr/blob/dev/webclient/xml_generator/definitions/metadata.xml)



<https://ocfl.io/>

# Research Data Day & EOSC National Tripartite Event 2025

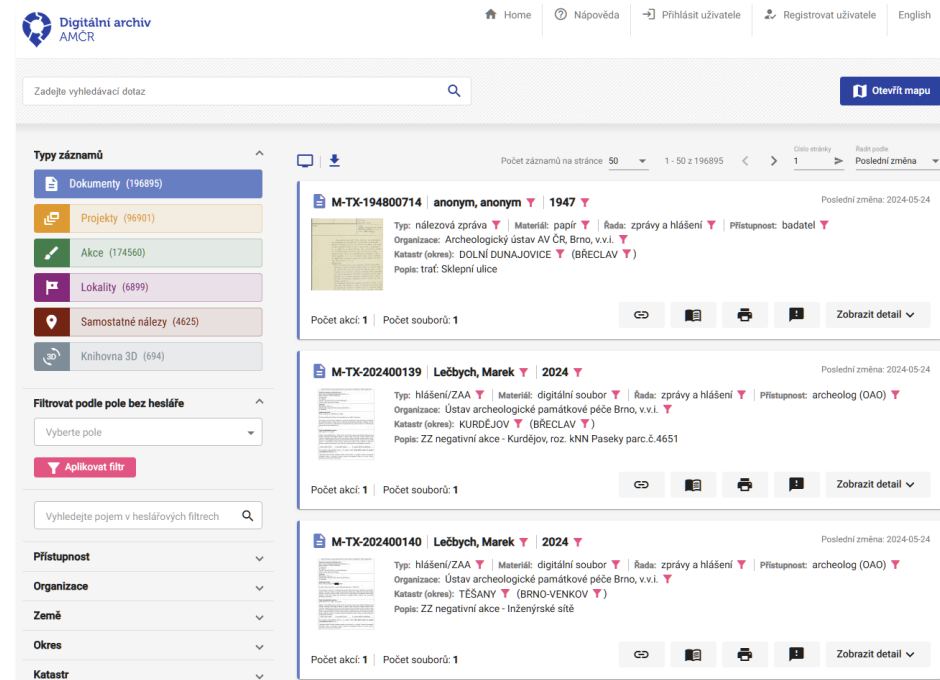
## Integrating domain repository for archaeological data

### AMCR Digital Archive

Digital Archive is a web application used for browsing data stored in the AMCR. It is fully connected to the AMCR and makes available all the information collected and stored via AMCR and its individual modules. It offers English interface and vocabularies.



```
https://fedora-test.aiscr.cz/rest/fcr:search?condition=fedora_id%3DAMCR%2Frecord%2F*&condition=modified%3E%3D2023-12-15T00%3A00%3A00.000Z
```



# Research Data Day & EOSC National Tripartite Event 2025

## Integrating domain repository for archaeological data

### AMCR API

The API offers a Data Provider service that provides metadata records from the AMCR database using the OpenArchives Initiative Protocol for Metadata Harvesting (OAI-PMH). OAI-PMH implementation supports representation in native AMCR XML and Dublin Core.

Solr



GET / HTTP/1.1

request

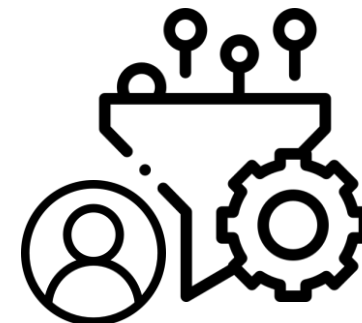
HTTP/1.1 401 Unauthorized

WWW-Authenticate: Basic realm="Access to the staging site"

response

oai\_amcr

oai\_dc



GET / HTTP/1.1

Authorization: Basic KAxhq06Lrc0984908wiosdk

request

HTTP/1.1 200 OK

OR

HTTP/1.1 401 Forbidden

response



# Research Data Day & EOSC National Tripartite Event 2025

Integrating domain repository for archaeological data

**F**   
indable

AMCR in re3data  
and **National  
Repository  
Catalogue**

PIDs <-> **National  
DOI Centre**

rich metadata <->  
**National Metadata  
Directory**

**A**   
ccessible

- persistent landing page in the AMCR Digital Archive
- HTTPS / OAI-PMH
- AMCR AAI <-> **e-Infra AAI**
- preservation and versioning using Fedora (OCFL)

**I**   
nteroperable

- AMCR data model <-> **Dataspecer <-> CCMM**
- AMCR vocabularies (SKOS version planned)
- **ORCID, Wikidata** and ROR references

**R**   
eusable

- standard domain-specific metadata
- CC-BY-NC 4.0 licence (to be reworked)
- basic provenience tracking
- **FIP specification (FIP Wizard)**



**xxx** – NDI tool

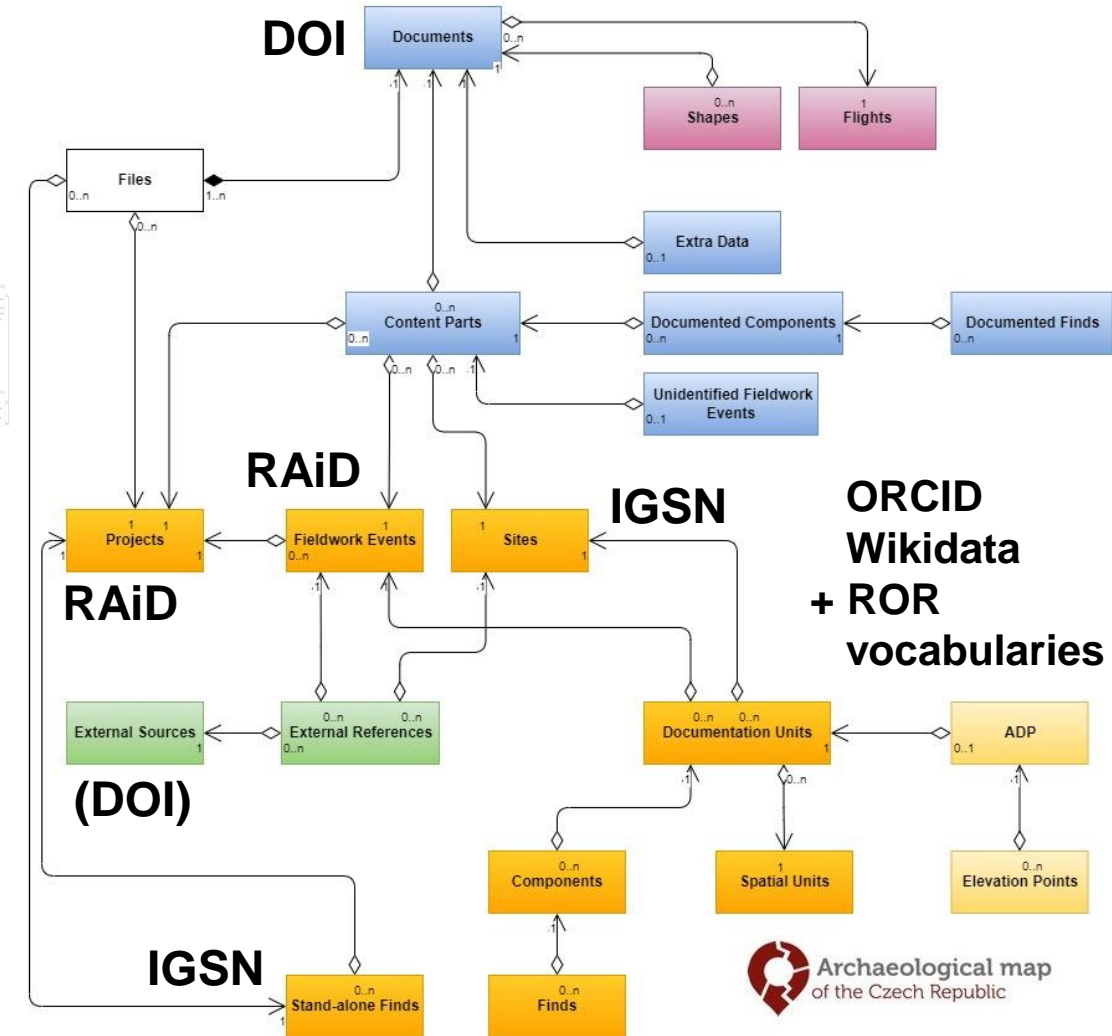
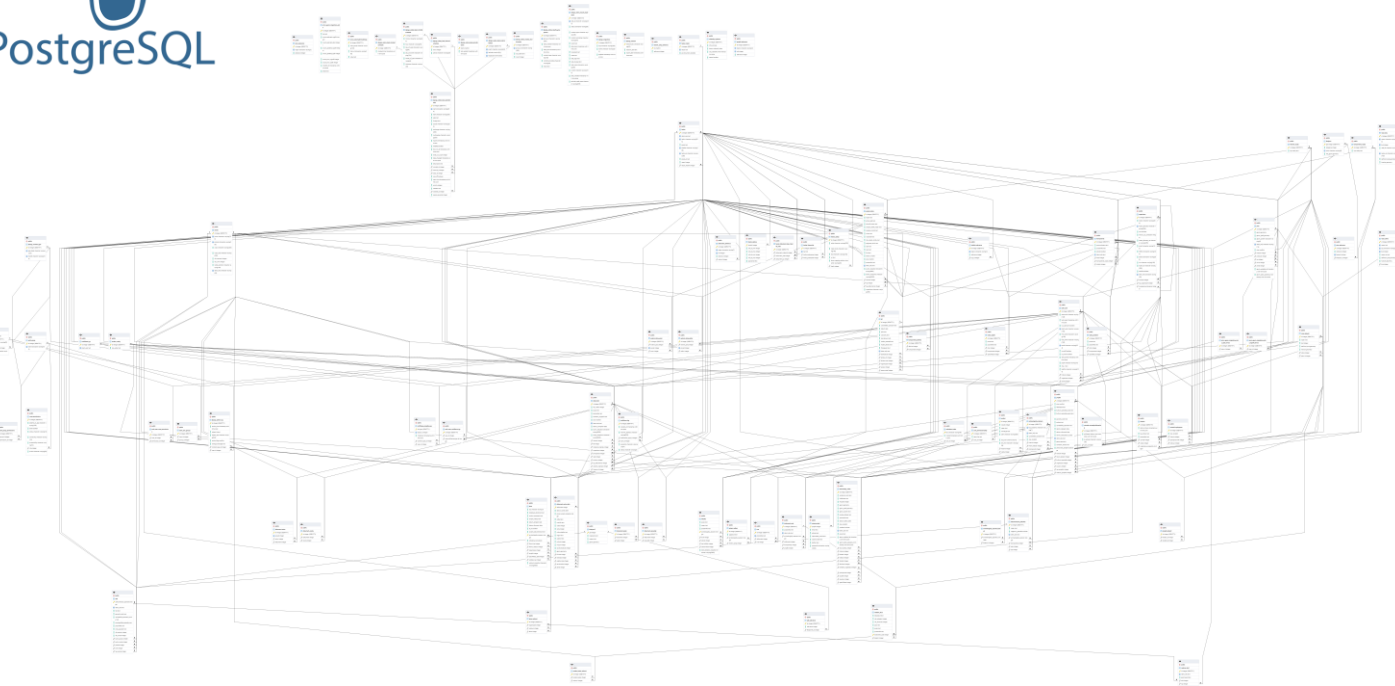
**xxx** – expected implementation (OSII project)

# Research Data Day & EOSC National Tripartite Event 2025

## Integrating domain repository for archaeological data

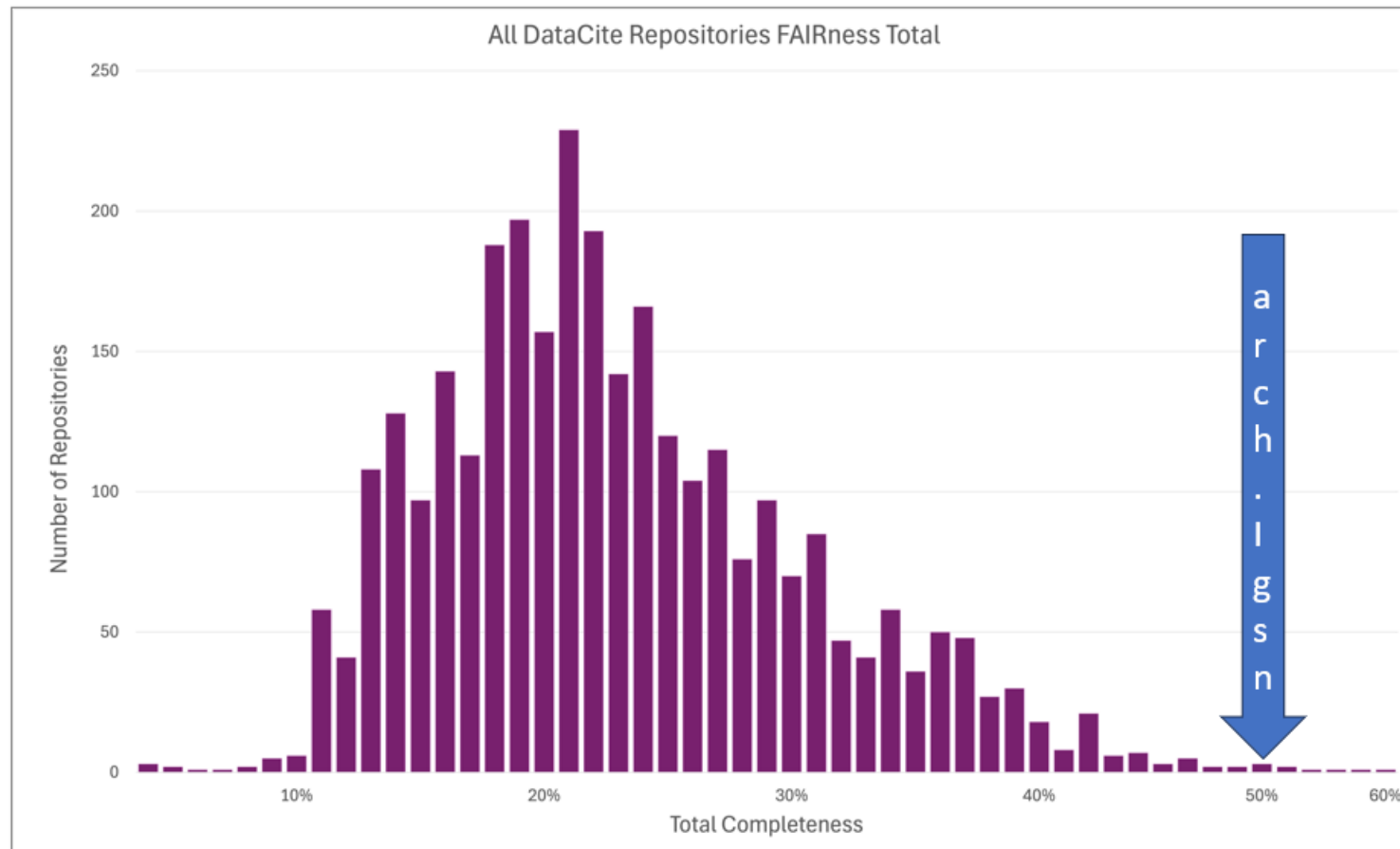


PostgreSQL



Model of the AMCR database and PID usage

### Distribution of FAIR Completeness



This slide shows the distribution of FAIR completeness scores across all DataCite repositories. The average is 23.5%.

**THANK YOU!**

Follow AIS CR community at Zenodo: <https://zenodo.org/communities/aiscr>



INSTITUTE OF ARCHAEOLOGY  
OF THE CAS PRAGUE



Institute of Archaeology  
Czech Acad Sci, Brno