Mapping Archaeological Heritage in South Asia: open-source technologies and sustainability

Junaid Abdul Jabbar
Mapping Archaeological Heritage in South Asia (MAHSA)

• Documentation of the endangered archaeology and cultural heritage of the Indus River Basin in an open access online database using the Arches platform

• Provide an open access mapping resource and research repository

• Collaborative research output with local heritage professionals

• Concurrent programme of collaborative development and training programmes
Data Sources

- **Historical Maps**
- **Legacy Data**
- **Remote Sensing**
- **Field Surveys**

MAHSA map sheet georeferencing: Progress to date over thousand one-inch to a mile maps
Data Sources

Machine Learning

Berganzo-Besga et al (2023)
Arches

- Open-source web-based, geospatial information system for cultural heritage inventory and management.

- Developed by the Getty Conservation Institute and World Monuments Fund.

https://www.archesproject.org/
Arches

- CIDOC-CRM ontology and works by mapping the data to CRM classes and properties, using a special graph structure referred to as resource models.

- Nodes

- Branches

Arches

• Several data types including concept i.e., controlled vocabularies (can be hierarchical as well)

• Relate resources across database with defined relationships
Arches

• User can make custom resource models as per requirement

• Complex nested multiples e.g., one site with multiple conditions and each condition assessment with multiple disturbance, threats etc.
## Arches - Data Template

<table>
<thead>
<tr>
<th>ResourceID</th>
<th>Name</th>
<th>Name Type</th>
<th>Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Shaheeda Abai</td>
<td>Primary Name</td>
<td>Built/Historical Heritage</td>
</tr>
<tr>
<td>2</td>
<td>Mohammadi Dherai</td>
<td>Primary Name</td>
<td>Archaeological Site</td>
</tr>
<tr>
<td>2</td>
<td>Khazana Dherai</td>
<td>Alternate Name</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Kaga Darra</td>
<td>Primary Name</td>
<td>Archaeological Feature</td>
</tr>
</tbody>
</table>
Arches - Data Template

1- Indicates whether one site can have more than one set of the defined attribute

2- Data type

3- Attribute Name

4- Dropdown for controlled vocabularies

5- Colored columns for easy identification of attribute groups

6- Unique MAHSA_ID
Arches - Releases

• Several releases (minor and major updates) since October 2020

• Version 6 released with a new method to import data i.e., SQL ETL Methods or Relational Schema

• Version 7 released with internationalisation capability
Arches - Relational Schema
Setting up Centralised Database

• PostgreSQL relational database with all the attribute fields from the resource models

• Simple data types as compared to ‘jsonb’ data type in the Arches relational views

• Configuring SSL for PostgreSQL

• Automated backup using cron job
QGIS Project Setup

• Load all layers
• Discover relations
• Configure edit settings
• Design data entry form
• User Constraints & Validations
• Set layer visibility
QGIS Form for Centralised Database
```sql
-- inserting name data
with name_type (valueid, value) as (
    select valueid, value
    from __arches_get_labels_for_concept_node(
        __arches_get_node_id_for_view_column(
            'heritage_location_resource_model_v6', 'name', 'name_type'
        )
    )
)
insert into heritage_location_resource_model_v6.name (tileid, name_type, name, resourceinstanceid)
(
    select uuid_generate_v4(), (
        select valueid
        from name_type
        where value = name_type
        limit 1
    ),
    when name is null then null
    else jsonb_build_object('en', jsonb_build_object('value', name, 'direction', 'ltr'))
end,
    resourceinstanceid from hl_name
);
```
Arches - Imported Data

Results: 31

1. Punjab No. 39 J 16, Formal title
   Maps created by the Survey of India Offices to help survey and consolidate the territories of the British East India Company and then the British Raj.

2. N.W. Frontier Province & Punjab No. 43 C 01_1909, Formal title
   Maps created by the Survey of India Offices to help survey and consolidate the territories of the British East India Company and then the British Raj.

3. N.W. Frontier Province & Punjab No. 43 C 01_1912, Formal title
   Maps created by the Survey of India Offices to help survey...
Arches Collector

• Companion mobile data collection app for the Arches data management platform

• Variable pace of development for Arches and Arches Collector

• When updated, would provide an integrated system for data collection and validation

https://www.archesproject.org/collector/
Open Data Kit (ODK)

- ODK is an open-source mobile data collection platform.
- Offline Data Collection
- User scalability – stakeholder training
- Multilingual
MAHSA Data Life Cycle

Mapping Archaeological Heritage of South Asia (MAHSA) - Data Life Cycle

Data Sources
- Historical Maps
- Legacy Data
- Remote Sensing
- Machine Learning
- Field Surveys

Data Development
- Relational GeoPackage & QGIS
- OpenDataKit Survey Submissions
- Structured CSV Files
- SharePoint Storage
- Formatting & Quality Assurance Checkpoints – Collaborators and MAHSA Team

Publicly Hosted Arches Database
- Different levels of user access
- Multiple languages – English, Hindi, Urdu
- Rich search scenarios
- Directed IIIF images for historical maps
- Data export options

Centralized Data Development
- PostgreSQL

Decentralized Data Development
Further Steps

• A new sustainability group

• Collaborative sustainability group meetings for multiple Arcadia funded projects

• Sustainability consultant for Arches

• Active community engagement
Acknowledgements

- Arcadia Fund for generously supporting the MAHSA project
- Staff at the British Library and Cambridge University Library for their guidance and advice
- IIIF Maps community for developing the standards, use cases and demos
- Teams at our sister Arcadia projects MAEASaM, CAAL, EAMENA, MarEA for their openness and support
- Teams from the *Land, Water & Settlement* and *TwoRains* projects for making a fundamental contribution to the development of many of the underlying methods being used by the MAHSA project
Thank you!

www.mahsa.arch.cam.ac.uk