



UNIVERSITY OF
CAMBRIDGE

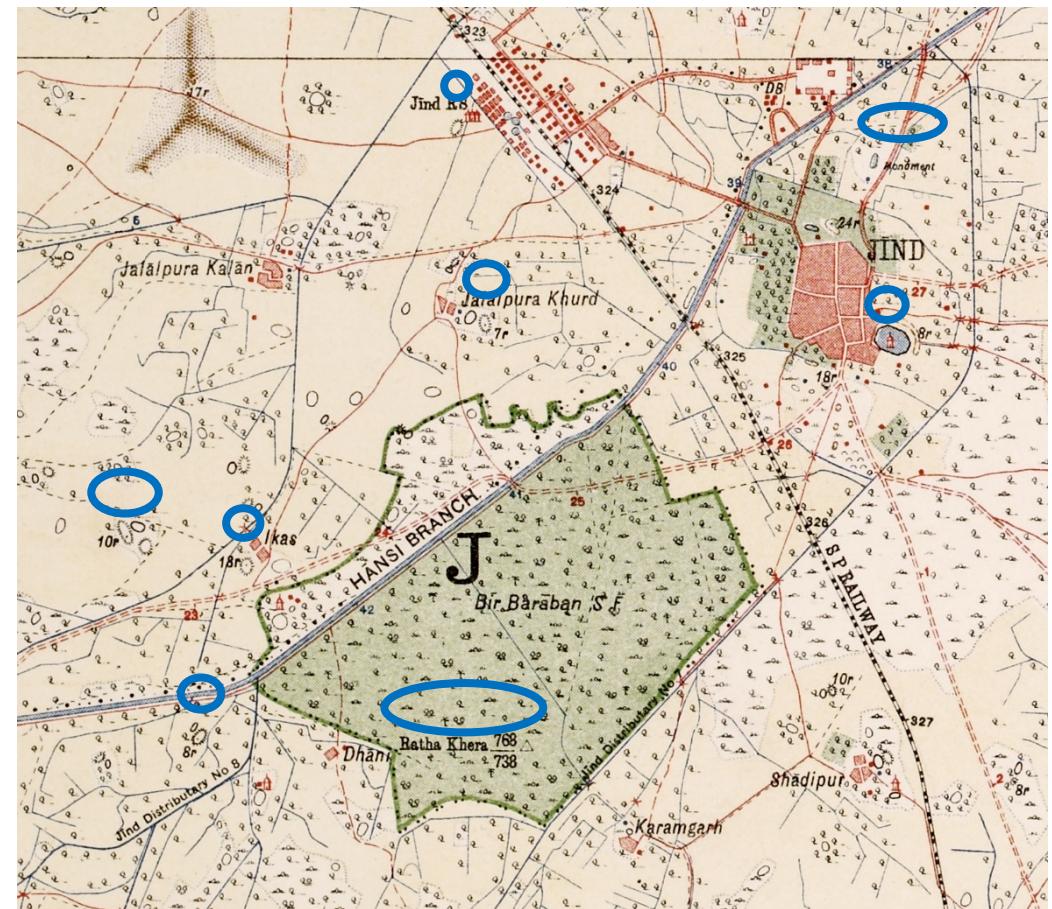


Mapping Archaeological Heritage in South Asia: open-source technologies and sustainability

Junaid Abdul Jabbar

Mapping Archaeological Heritage in South Asia (MAHSA)

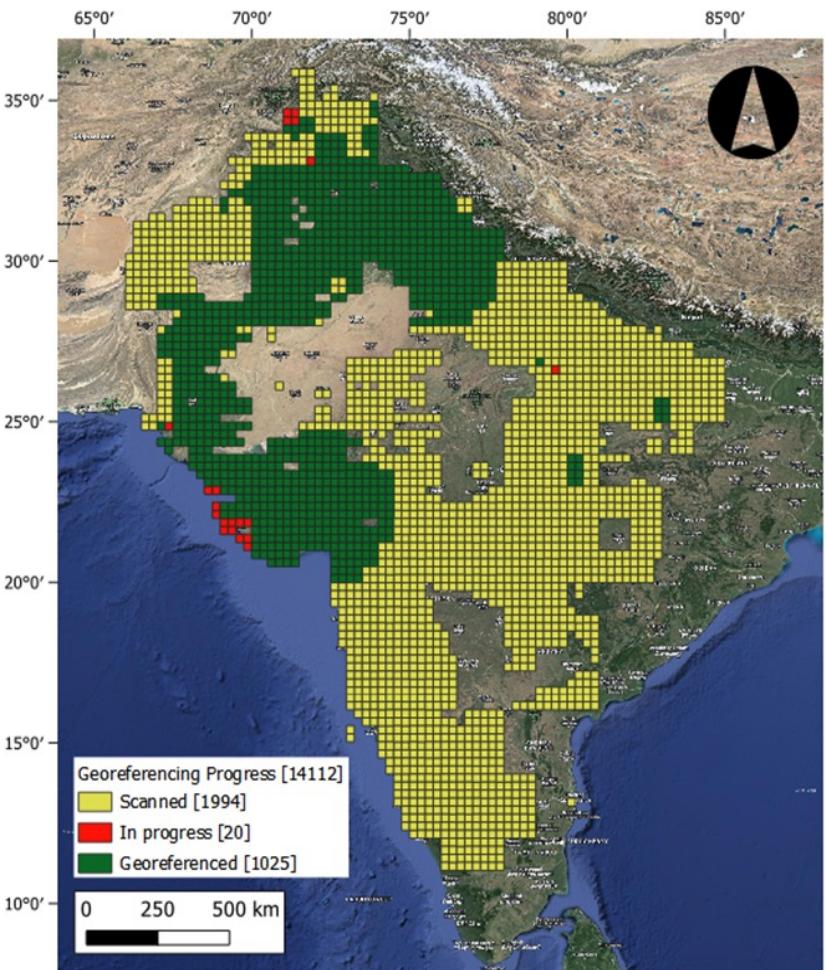
- Documentation of the endangered archaeology and cultural heritage of the Indus River Basin in an open access online database using the Arches platform
- Provide an open access mapping resource and research repository
- Collaborative research output with local heritage professionals
- Concurrent programme of collaborative development and training programmes



Data Sources



Historical Maps



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



Legacy Data



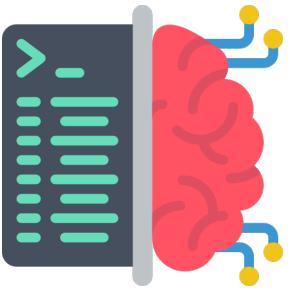
Remote Sensing



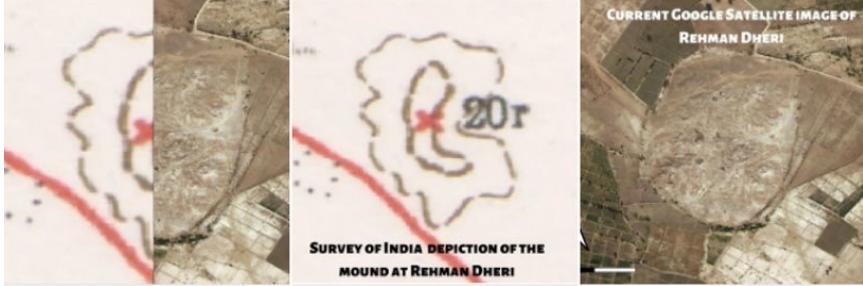
Field Surveys



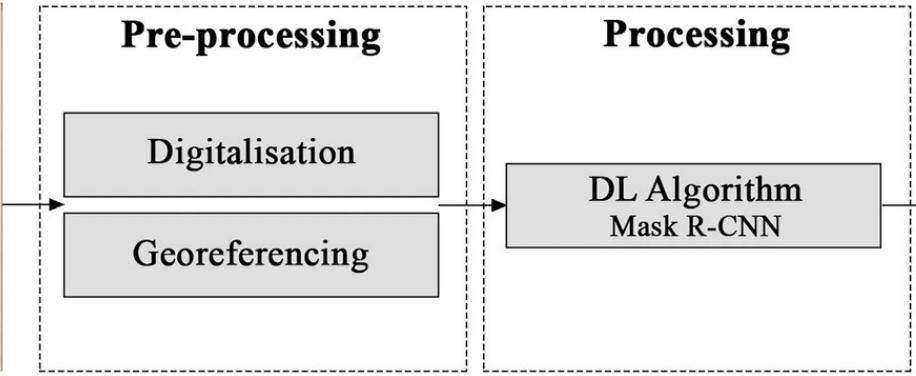
Data Sources



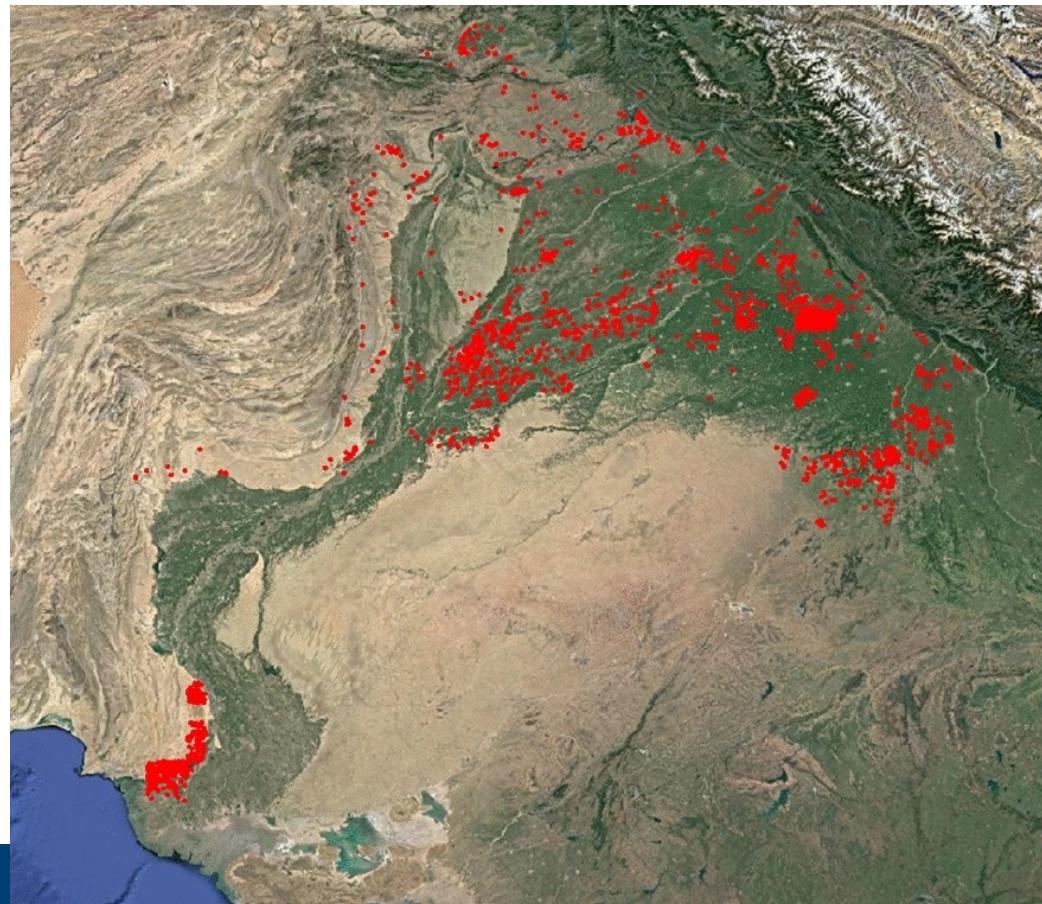
Machine Learning



Historical Map



Detected distribution



Berganzo-Besga et al (2023)



UNIVERSITY OF
CAMBRIDGE

Arches

- Open-source web-based, geospatial information system for cultural heritage inventory and management.
- Developed by the Getty Conservation Institute and World Monuments Fund.

The screenshot displays the Arches web application interface. On the left, there is a sidebar with various icons and a 'DEBUG' section indicating 'Arches 7.4.0'. The main area features a search bar at the top with dropdown menus for 'Historic Maps Resource Model' and 'All (*)'. Below the search bar, it shows 'Results: 31' and a 'Clear Filters' button. The results list contains three items, each with a thumbnail, title, and a brief description. The first item is 'Punjab No. 39 J 16, Formal title' with the description: '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.' The second item is 'N.W. Frontier Province & Punjab No. 43 C 01_1909, Formal title' with the same description. The third item is 'N.W. Frontier Province & Punjab No. 43 C 01_1912, Formal title' with the same description. At the bottom of the results list is a pagination control with buttons for 1, 2, 3, 4, ..., 7, and >. To the right of the search results is a map view showing a satellite-style map of a region. Blue square markers are placed on the map, corresponding to the locations of the resources listed in the results. The map includes a 'Map Filter' section with checkboxes for 'Advanced' and 'Related', and a 'Details' button. The bottom right corner of the map area includes copyright notices for 'Mapbox' and 'OpenStreetMap'.

<https://www.archesproject.org/>

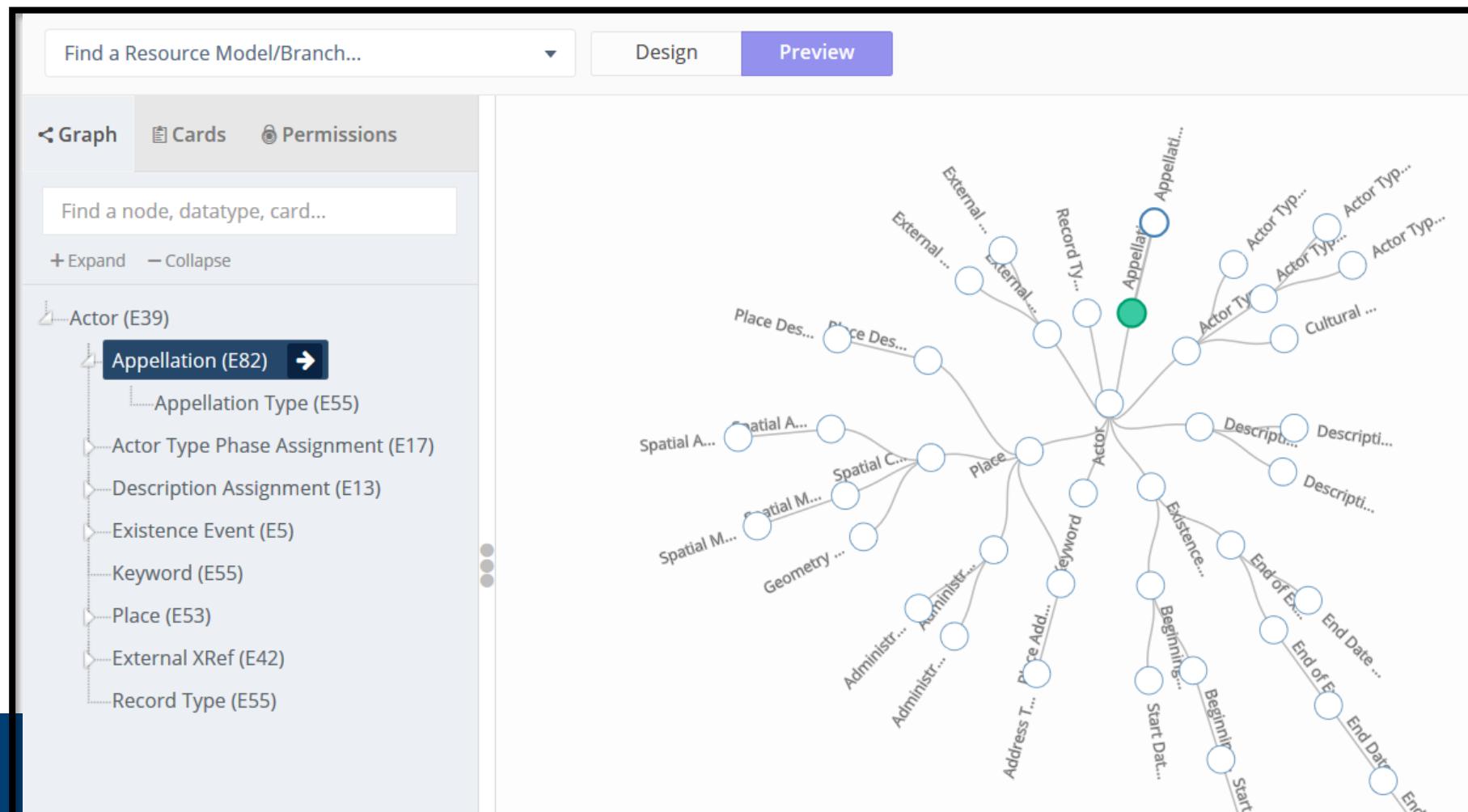
Arches

- 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

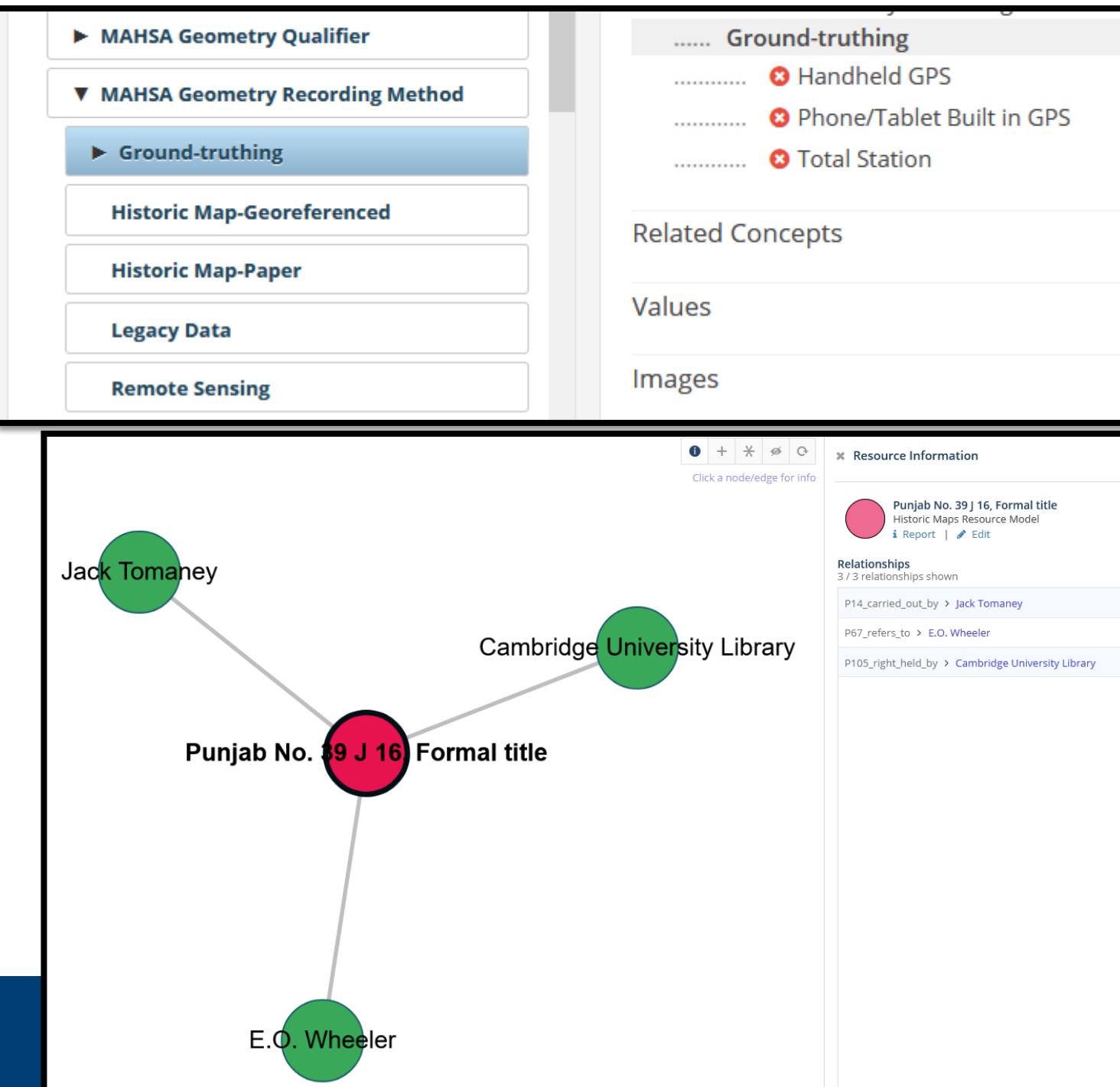
- Branches

<https://arches.readthedocs.io/en/stable/administering/designing-the-database/>



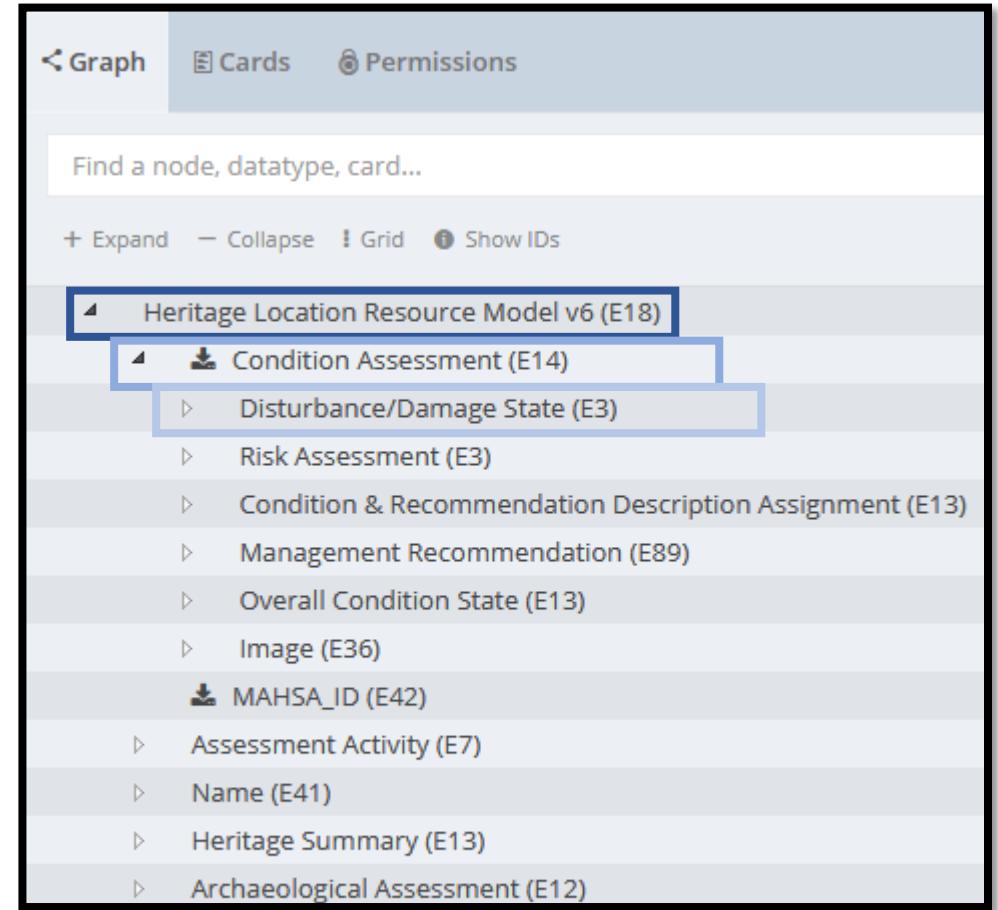
Arches

- 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

ResourceID	attribute 1	attribute 2	attribute 3
1	attr. 1 value	attr. 2 value	attr. 3 value
2	attr. 1 value	attr. 2 value	attr. 3 value
2		attr. 2 additional value	
3	attr. 1 value	attr. 2 value	attr. 3 value

ResourceID	Name	Name Type	Classification
1	Shaheeda Abai	Primary Name	Built/Historical Heritage
2	Mohammadi Dherai	Primary Name	Archaeological Site
2	Khazana Dherai	Alternate Name	
3	Kaga Darra	Primary Name	Archaeological Feature

Arches - Data Template

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

3- Attribute Name

2- Data type

5- Colored columns for easy identification of attribute groups

6- Unique MAHSA_ID

4- Dropdown for controlled vocabularies

Site Name - Repeat Group

String	Concept
Name	Name Type
	Primary Name
	Primary Name
	Alternate Name(s)
	Village Name
	Survey/Reference ID

Assessment Activity - Repeat Group

ResourceID	Assessment Type	Date of Assessment	Proposed Completion Activities
Same as MASHA_ID	Concept-list (can select multiple)	Must be in YYYY-MM-DD form	Concept-list (can select multiple)
ResourceID	Assessment Type	Date of Assessment	Proposed Completion Activities
HL-LD-PO0004-000001	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy
HL-LD-PO0004-000001	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy
HL-LD-PO0004-000002	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy
HL-LD-PO0004-000002	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy
HL-LD-PO0004-000003	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy
HL-LD-PO0004-000003	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy
HL-LD-PO0004-000003	Legacy Data Digitisation/Interpretation	2022-07-27	Condition Assessment Using Remote Sensing Methods, Location Accuracy

Site Name - Repeat Group

String	Concept	Concept	Concept	Concept	String
Name	Name Type	Location Classification	Archaeological Certainty	Overall Morphology	Previous Research Activities
	Primary Name	Archaeological Site	High	Unknown	(A.R.A.S.I., F.C. 1916-17; A.P. 11.9)
	Alternate Name(s)				
	600	Survey/Reference ID			
	Takhabad	Primary Name	Archaeological Site	High	Unknown
	601	Survey/Reference ID			(S.N. Khan 1993)
	Tajrai Dherai	Primary Name	Archaeological Site	High	Unknown
	602	Survey/Reference ID			(A.R.A.S.I., F.C. 1916-17; 35)
	Shera Ghund	Alternate Name(s)			

Heritage Summary

R	S	T	U	V	W
---	---	---	---	---	---



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

Sr. No.	Version	Release Date
1	7.4.0	June 2023
2	7.3.0	January 2023
3	7.2.1	December 2022
4	7.2.0	November 2022
5	7.1.1	September 2022
6	7.1.0	September 2022
7	7.0	August 2022
8	6.2.3	April 2023
9	6.2.2	February 2023
10	6.2.1	October 2022
11	6.2.0	October 2022
12	6.1.2	November 2022
13	6.1	June 2022
14	6.0	October 2021
15	5.1.4	March 2021
16	5.1.3	March 2021
17	5.1.1	March 2021

Arches - Relational Schema

The screenshot shows the Arches Relational Schema interface with the following sections:

- Views (44):** A list of views including address_directions, administrative_subdivision, archaeological_assessment, archaeological_description, assessment_activity, built_component, condition_assessment, condition_recommendation_description_assignment, cultural_period_assignment, cultural_period_inference_making, description_assignment, disturbance_damage_state, eligibility_requirement_type, environmental_assessment, evaluation_assignment, evaluation_criteria_type, evaluation_time_span, existence_event, external_xref, finds, finds_assessment, finds_cultural_period_inference_making, and finds_image.
- Query Editor:** A code editor containing a PostgreSQL query:

```
60 select __arches_create_resource_model_views(graphid)
61 from graphs
62 where isresource = true
63 and replace((name->'en')::text,'','','') != 'Arches System Settings'
64 and graphid = '51b5d8b8-e57f-11ec-8903-fa163e2afc92';
65
66
67
68
69
70
71
72
73
```
- Data Output:** A table showing the results of the query:

	__arches_create_resource_model_views
1	schema "heritage_location_resource_model_v6" created for resource model.

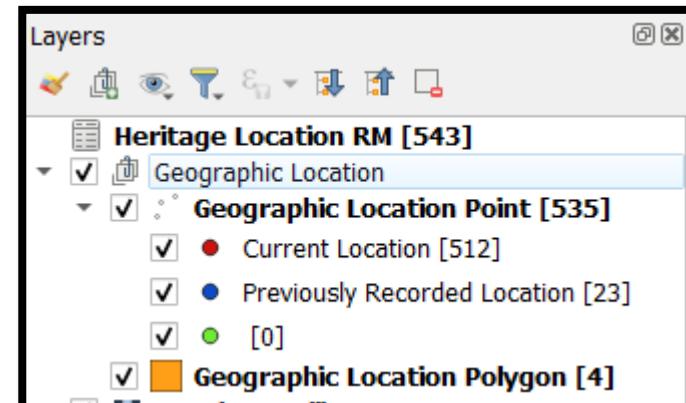
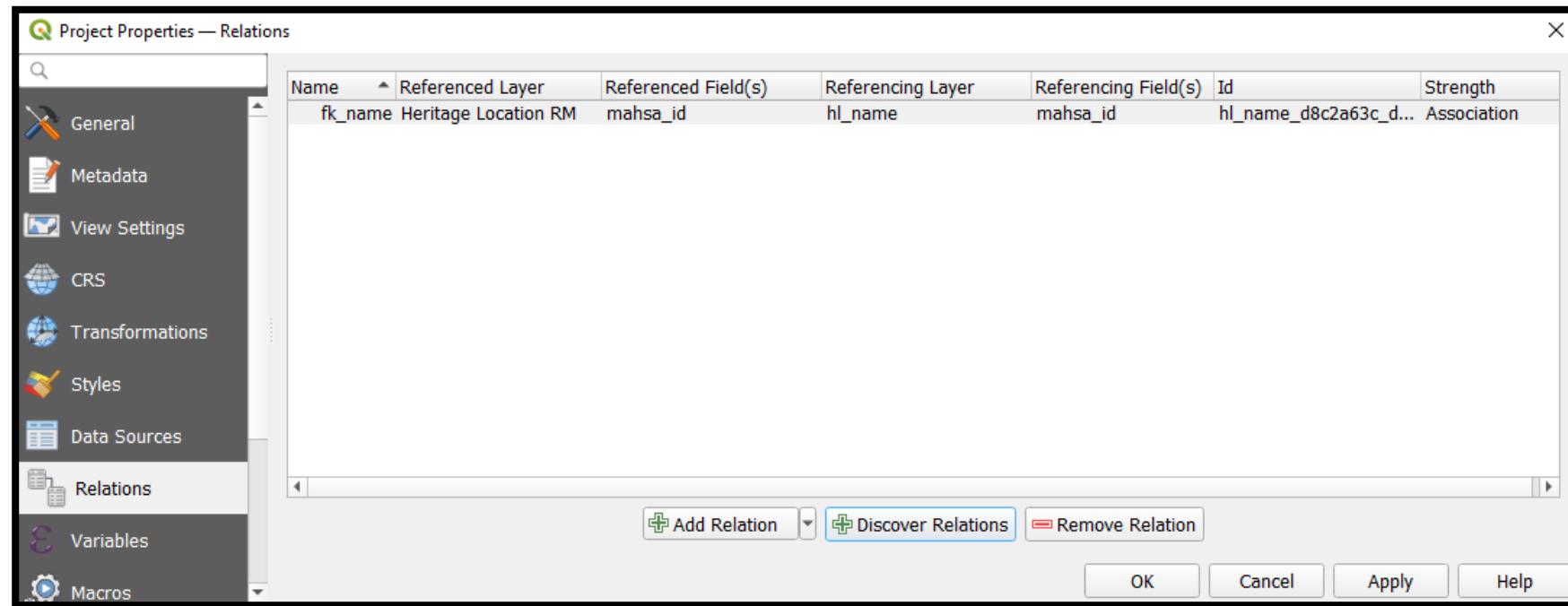
Setting up Centralised Database

- 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

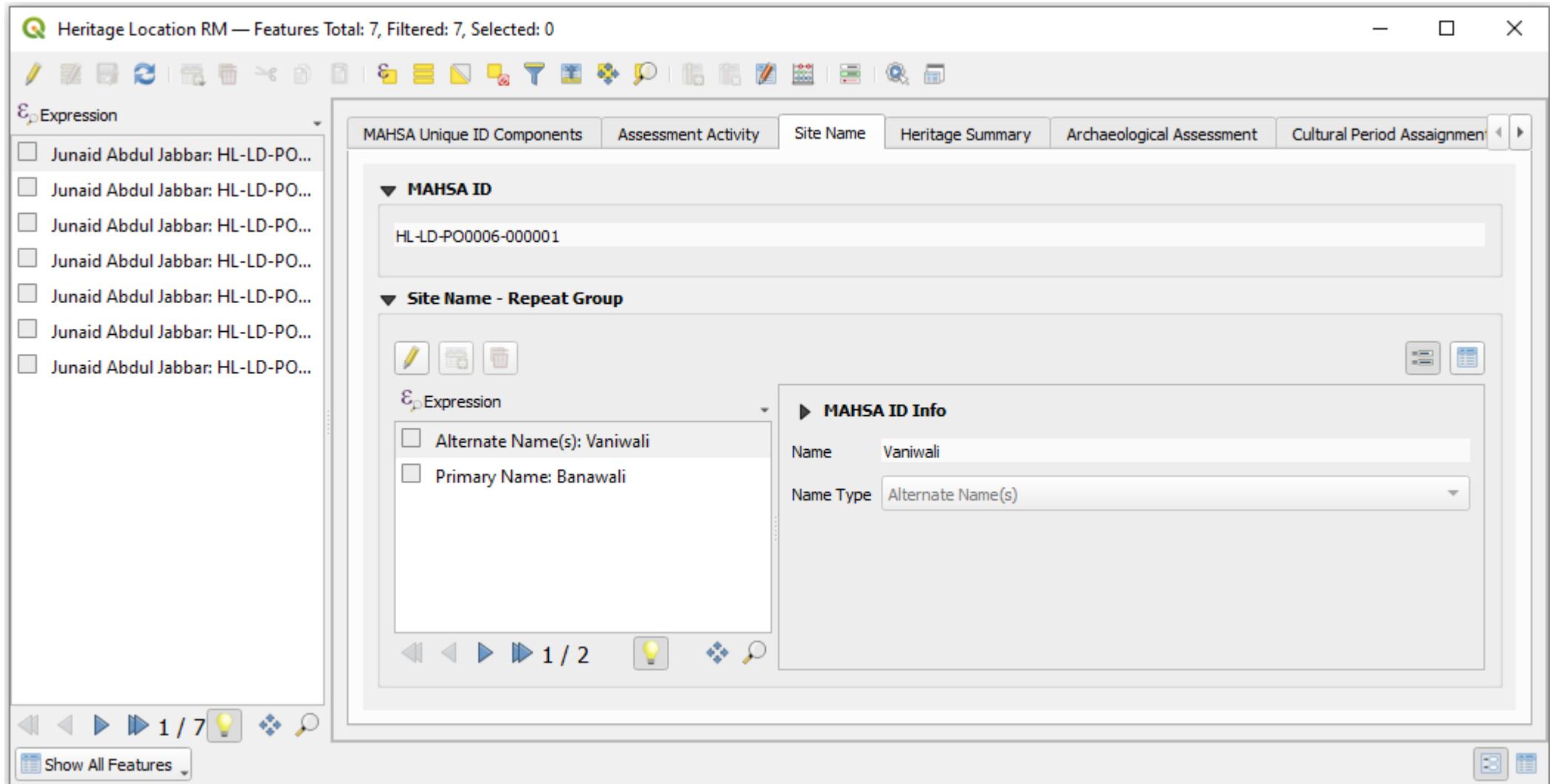


QGIS Project Setup

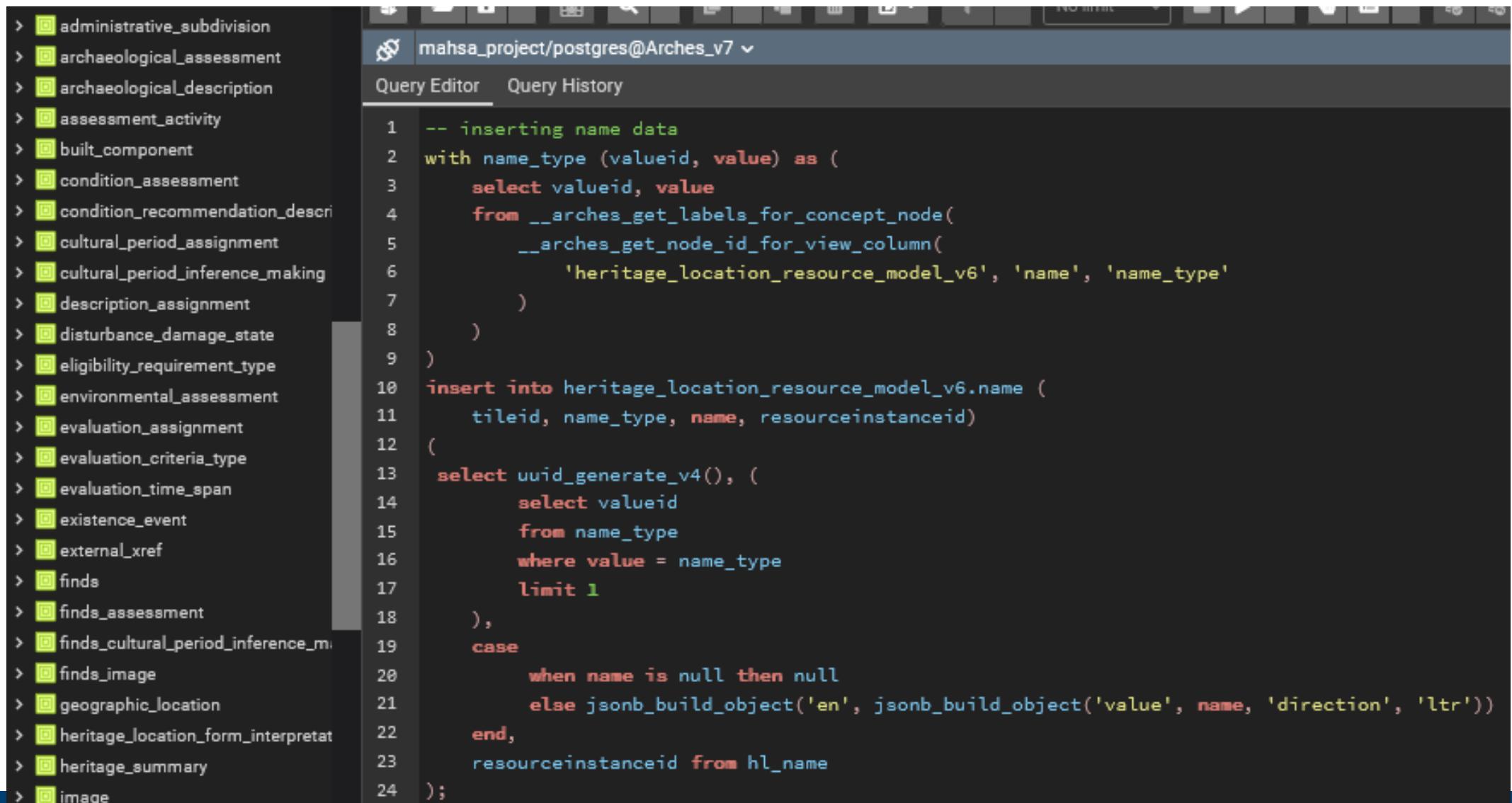
- Load all layers
- Discover relations
- Configure edit settings
- Design data entry form
- User Constraints & Validations
- Set layer visibility



QGIS Form for Centralised Database



Import Data to Arches



The screenshot shows a pgAdmin interface with a sidebar containing a tree view of database objects and a main window titled 'Query Editor'.

The Query Editor contains the following PostgreSQL code:

```
-- 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
    ),
    case
        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

Search

Historic Maps Resource Model

All (*)

QA Type Resource Type

Results: 31 Clear Filters Ascending

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.

[Details](#) [Edit](#) [Map](#) [Related Resources](#)

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.

[Details](#) [Edit](#) [Map](#) [Related Resources](#)

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

« 1 2 3 4 ... 7 »

Welcome, admin English (en) ?

Map Filter Advanced Related Details

Search

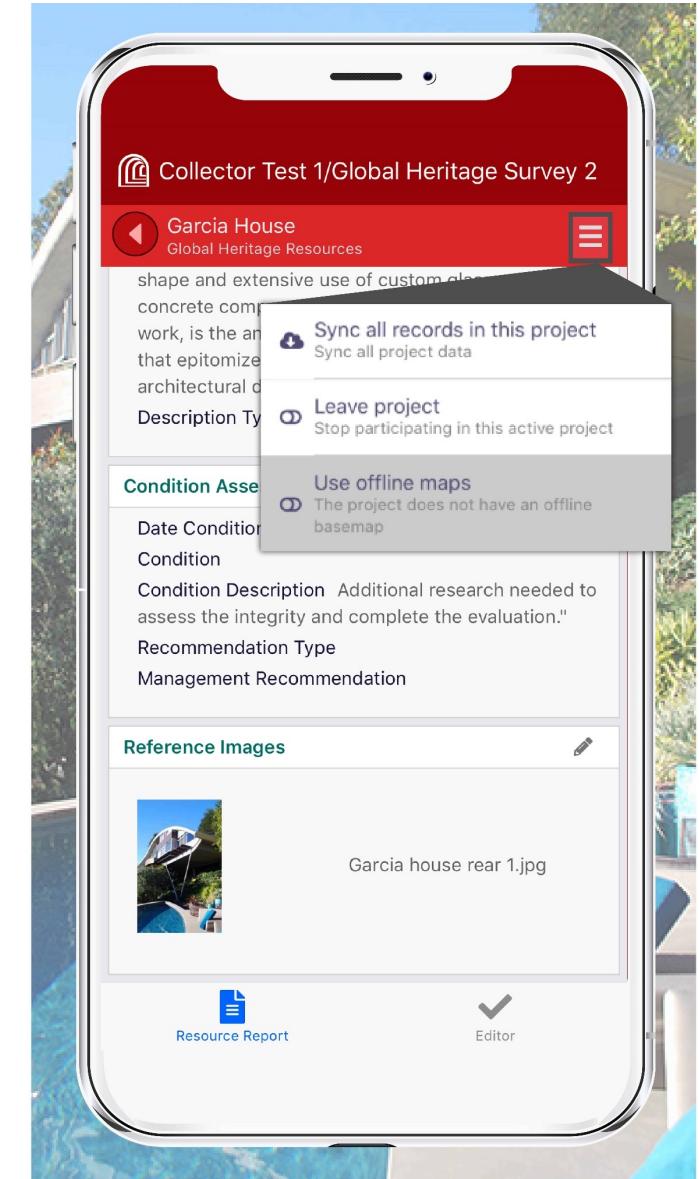
Mapbox © Mapbox © OpenStreetMap Improve this

Filter Basemap Overlays Legend



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

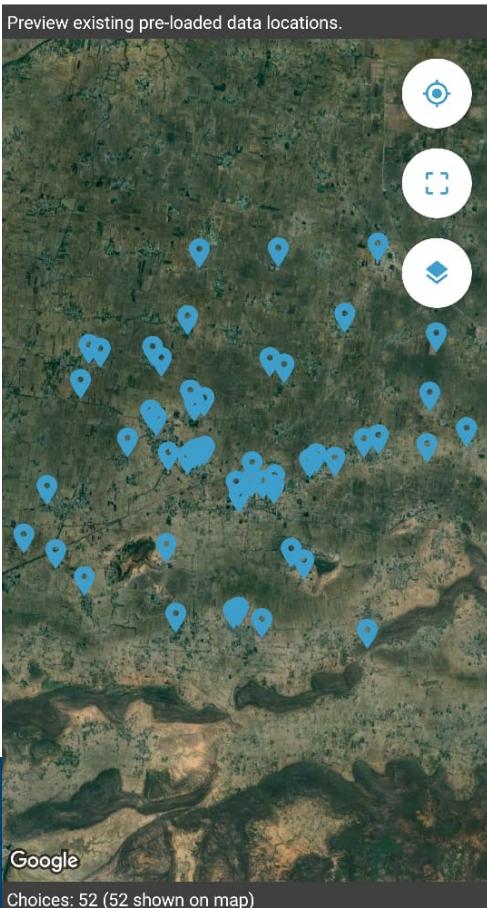


Open Data Kit (ODK)

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



UNIVERSITY OF
CAMBRIDGE



11:27 ⊖

87%

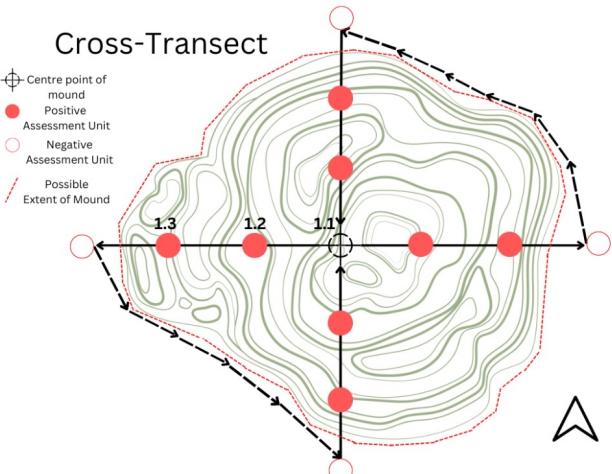
MAHSA Site Survey Fo...

Transect.

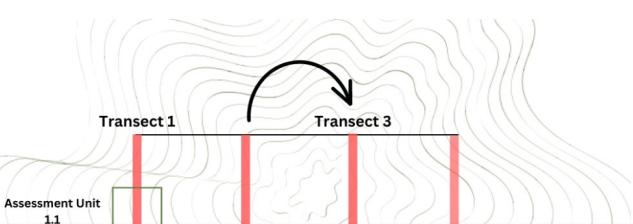
* Transect Questions

What type of transect or assessment are you conducting?

Cross Transect



Parallel Transect



< BACK

NEXT >

11:10 ⊖

91%

MAHSA Site Survey Fo...

Site Condition Information.

Assess the overall condition and current state of the site.

Excellent/Undisturbed

Good

Partially Damaged/Poor

Heavily Damaged

No Feature/Mound Visible

Unknown/Not Clear

Record any past or current disturbances to the site.

Select Answer ▾

What percentage of the site is affected by damage?

1-10%

11-30%

< BACK

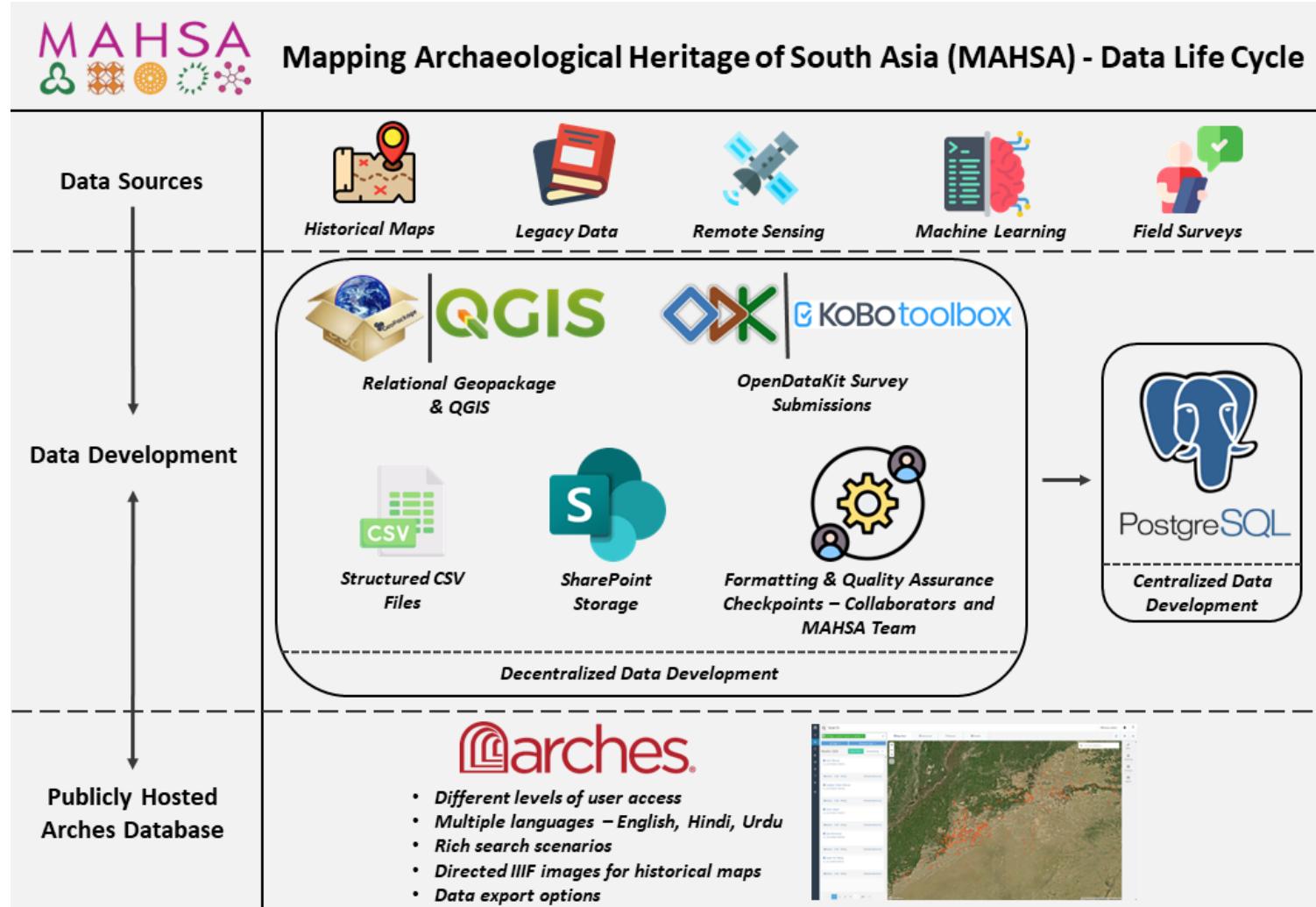
NEXT >

ODK



UNIVERSITY OF
CAMBRIDGE

MAHSA Data Life Cycle



Further Steps

- A new sustainability group
- Collaborative sustainability group meetings for multiple Arcadia funded projects
- Sustainability consultant for Arches

The screenshot shows the homepage of the Arches Project Community Forum. At the top, there is a navigation bar with links to archesproject.org, about, documentation, and code of conduct. On the far right of the navigation bar are icons for search, menu, and user profile. Below the navigation bar, a red banner displays the text "Welcome to the Arches Project Community Forum!". A blue header bar contains a notification message asking if the user wants live notifications when people reply to their posts, with an "Enable Notifications" link and a close button. Below this, there is a search bar and a row of filters: "all categories", "all tags", "Categories" (which is highlighted in red), "Latest", "New (3)", "Unread (54)", and "Top". To the right of these filters is a "+ Open Draft" button. The main content area features a table with columns for "Category", "Topics", and "Latest". The first row of the table is titled "News & Announcements". At the bottom of the page, there is a footer bar with a gear icon, the text "Welcome to the Arches Project", and a "0" indicating zero notifications.

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



McDonald Institute for
Archaeological Research



Universitat
Pompeu Fabra
Barcelona



ARCADIA



UNIVERSITY OF
CAMBRIDGE

Thank you!

www.mahsa.arch.cam.ac.uk

