Welcome!
Intro to the Arches Platform
Part 3 of 4 – Data Technology Concepts

Thursday, January 21, 2021
Webinar Presenter and Q&A Moderators

Annabel Lee Enriquez
Getty Conservation Institute

David Myers
Getty Conservation Institute

Dennis Wuthrich
Farallon Geographics Inc
Intro to the Arches Platform Webinars

Part 1: Capabilities
Part 2: Data Management Tools
Part 3: Data Technology Concepts
Part 4: Information Architecture
Webinar Repository – www.archesproject.org/videos

Arches Videos

The following are videos that feature Arches. If you have put together a video about Arches and would like it included here (with attribution, of course), please let us know via email (contact@archesproject.org) or via our Discussion Forum.

Webinars

These are videos that are related to Arches Webinars. When available, transcripts and slide decks will be included.

- Arches Developer Meeting: Project Presentations by Arches Community (December 2020)
  View Full Transcript (PDF).

- Arches Community Webinar: Documentation by Adam Cox (November 2020)
  View Full Transcript (PDF).

- Arches Version 5.0 Demo + Q&A by the Arches Team (August 2020)
  View Full Transcript (PDF).
  View Slide Deck (PDF).
Webinar Outline

- About Arches
- Arches Platform: Data Technology Concepts
  - What is data modeling in Arches?
  - Data modeling goals & concepts
  - Data modeling documentation and resources
- Q&A Session
About Arches

Arches is an open-source software platform for cultural heritage data management, originally developed by the Getty Conservation Institute in partnership with World Monuments Fund.
About Arches (cont’d)

- Open Source and free – no licensing fees and unlimited users.
- Enterprise-level software hosted on a server
- Independent deployment by an organization or institution
- Supported by a growing community of heritage conservation and technology professionals
Arches Overview

Arches Demo & Update

Project/Task Management

Data Management

Data Discovery & Visualization

• User & group secure access management
• Granular (node level) permission control
• Detailed audit of changes to data
• Provisional (unpublished) data; data review
• Semantic & thesauri-enhanced searches
• Searchable annotations
• Support for custom indexing
• Export search results
• Geospatial search
• Saved and Advanced searches
• Integrates GIS functionality
• Seamless management of Arches data in external GIS – e.g., esri
• Changeable basemaps – e.g., historical maps
• Overlays supported
• Satellite imagery
• Explore relationships between resources – e.g., people, materials, activities, historic events, objects, iconography
• Discovery of previously unknown connections / new knowledge
• Temporal search
• Fuzzy date support
• Visualization of time distribution in entire dataset – Timewheel

Reports, Viewers & Annotations

Workflows

• Comprehensive, flexible workflow (i.e., process) management
• Task tracking
• Notifications
• Project Status plugin
• Correspondence management and tracking
• Letter templates with auto completion

Mobile app: on/off-line collection
• Robust import/export with notifications

Semantic, self-describing, & sustainable data structure
• Dynamic UI-generation
• Multiple ontologies – e.g., CIDOC CRM
• Cloud/server deployable
• API & linked data support
• Multiple data types supported

Thesauri management — Reference Data Manager (RDM)

Collector & Cloudserver WebApp

A comprehensive information and task management platform for heritage institutions
archesproject.org
A comprehensive information and task management platform for heritage institutions

**ARCHES DESIGNER**

**REFERENCE DATA MANAGER (RDM)**
What is data modeling in Arches?

Creation of **Resource Models** (data models) using Arches Designer to define:

- database structure
- data entry interface
- overall organization of information in your Arches implementation.
Why does Arches include advanced data modeling capabilities?

To give implementers the ability to ensure that their DATA is:

- complete
- high quality
- discoverable
- shareable
- persistent
Modeling for COMPLETE DATA...

Resource Models encompass all project requirements, defining what data is managed in Arches.
<table>
<thead>
<tr>
<th>ID</th>
<th>Name</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>00001</td>
<td>Mark Twain (pen name); Samuel Clemens (birth name)</td>
<td>Writer</td>
<td>Samuel Clemens (November 30, 1835 – April 21, 1910), known by his pen name Mark Twain, was an American writer, humorist, entrepreneur, publisher, and lecturer. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00002</td>
<td>Frank Lloyd Wright</td>
<td>Architect</td>
<td>Frank Lloyd Wright (June 8, 1867 – April 9, 1959) was an American architect, designer, writer, and educator. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00003</td>
<td>Elizabeth II; Elizabeth Alexandra Mary Windsor (full name)</td>
<td>Monarch</td>
<td>Elizabeth II (born 21 April 1926) is Queen of the United Kingdom and 15 other Commonwealth realms. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00004</td>
<td>Vincent Van Gogh</td>
<td>Artist</td>
<td>Vincent Willem van Gogh (30 March 1853 – 29 July 1890) was a Dutch post-impressionist painter who posthumously became one of the most famous and influential figures in the history of Western art. (WIKIPEDIA)</td>
</tr>
<tr>
<td>ID</td>
<td>Name</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------------------</td>
<td>-------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>00001</td>
<td>Mark Twain (pen name); Samuel Clemens (birth name)</td>
<td>Writer</td>
<td>Samuel Clemens (November 30, 1835 – April 21, 1910), known by his pen name Mark Twain, was an American writer, humorist, entrepreneur, publisher, and lecturer. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00002</td>
<td>Frank Lloyd Wright</td>
<td>Architect</td>
<td>Frank Lloyd Wright (June 8, 1867 – April 9, 1959) was an American architect, designer, writer, and educator. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00003</td>
<td>Elizabeth II; Elizabeth Alexandra Mary Windsor (full name)</td>
<td>Monarch</td>
<td>Elizabeth II (born 21 April 1926) is Queen of the United Kingdom and 15 other Commonwealth realms. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00004</td>
<td>Vincent Van Gogh</td>
<td>Artist</td>
<td>Vincent Willem van Gogh (30 March 1853 – 29 July 1890) was a Dutch post-impressionist painter who posthumously became one of the most famous and influential figures in the history of Western art. (WIKIPEDIA)</td>
</tr>
<tr>
<td>ID</td>
<td>Name</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>00001</td>
<td>Mark Twain (pen name); Samuel Clemens (birth name)</td>
<td>Writer</td>
<td>Samuel Clemens (November 30, 1835 – April 21, 1910), known by his pen name Mark Twain, was an American writer, humorist, entrepreneur, publisher, and lecturer. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00002</td>
<td>Frank Lloyd Wright</td>
<td>Architect</td>
<td>Frank Lloyd Wright (June 8, 1867 – April 9, 1959) was an American architect, designer, writer, and educator. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00003</td>
<td>Elizabeth II; Elizabeth Alexandra Mary Windsor (full name)</td>
<td>Monarch</td>
<td>Elizabeth II (born 21 April 1926) is Queen of the United Kingdom and 15 other Commonwealth realms. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00004</td>
<td>Vincent Van Gogh</td>
<td>Artist</td>
<td>Vincent Willem van Gogh (30 March 1853 – 29 July 1890) was a Dutch post-impressionist painter who posthumously became one of the most famous and influential figures in the history of Western art. (WIKIPEDIA)</td>
</tr>
<tr>
<td>ID</td>
<td>Name</td>
<td>Type</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-----------------------------------------------------------</td>
<td>-----------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>00001</td>
<td>Mark Twain (pen name); Samuel Clemens (birth name)</td>
<td>Writer</td>
<td>Samuel Clemens (November 30, 1835 – April 21, 1910), known by his pen name Mark Twain, was an American writer, humorist, entrepreneur, publisher, and lecturer. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00002</td>
<td>Frank Lloyd Wright</td>
<td>Architect</td>
<td>Frank Lloyd Wright (June 8, 1867 – April 9, 1959) was an American architect, designer, writer, and educator. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00003</td>
<td>Elizabeth II; Elizabeth Alexandra Mary Windsor (full name)</td>
<td>Monarch</td>
<td>Elizabeth II (born 21 April 1926) is Queen of the United Kingdom and 15 other Commonwealth realms. (WIKIPEDIA)</td>
</tr>
<tr>
<td>00004</td>
<td>Vincent Van Gogh</td>
<td>Artist</td>
<td>Vincent Willem van Gogh (30 March 1853 – 29 July 1890) was a Dutch post-impressionist painter who posthumously became one of the most famous and influential figures in the history of Western art. (WIKIPEDIA)</td>
</tr>
</tbody>
</table>
ARCHES DESIGNER INTERFACE
Simple Resource Model for Person

In Arches Designer:

Graphical Representation:
Simple Resource Model for Person

In Arches Designer:

- **Type**: "Writer"
- **Name**: "Frank Lloyd Wright"
  - **Name Type**: "Primary Name"
- **Identifier**: "00001"
  - **Identifier Type**: "System ID"
- **Description**: "Frank Lloyd Wright…"
  - **Description Type**: "Biography"
  - **Description Source**: Wikipedia URL

Graphical Representation:
Simple Resource Model for Person

- **Type**
- **Name**
  - Name Type
- **Identifier**
  - Identifier Type
- **Description**
  - Description Type
  - Description Source

Sophisticated Resource Model for Person

```xml
<Person (E21)>
  <Statement (E33)>
    <Statement_type (E55)>
    <Statement_content (xsd:string)>
    <Statement_name (E33)>
      <Statement_name_language (E56)>
      <Statement_name_label (xsd:string)>
      <Statement_name_source (E33)>
      <Statement_name_type (E55)>
      <Statement_name_content (xsd:string)>
    </Statement_name>
    <Statement_language (E56)>
    <Statement_source (E33)>
    <Statement_label (xsd:string)>
  </Statement>
  <Contact Point (E42)>
    <Contact Point_label (xsd:string)>
    <Contact Point_type (E55)>
    <Contact Point_content (xsd:string)>
    <Contact Point_source (E33)>
  </Contact Point>
  <Nationality (E55)>
    <Nationality_Meta_Type (E55)>
  </Nationality>
</Person>
```
Modeling for COMPLETE DATA...

Resource Models may correspond to a data content standard, such as:

- MIDAS Heritage
- International Core Data Standard for Archaeological and Architectural Heritage
- US SHPO requirements
- Dublin Core
Modeling for HIGH QUALITY DATA...

Resource Models define the data parameters, such as:

• Number of entries
• Data types
• Terminology used i.e. controlled vocabularies
Data Type Examples

- **String** (short or long alphanumeric text)
- **Number**
- **Boolean**
- **Date**
- **EDTF** (Extended Date Time Format)
- **Geo-json feature collection** (geospatial data)
- **Concept** (references controlled vocabulary)
- **Resource-instance** (references another Resource Model)
- **File-list** (for upload of digital files)
- **Annotation** (supports image annotations using a IIIF manifest)
Data Type Examples

- **String** (short or long alphanumeric text)
- **Number**
- **Boolean**
- **Date**
- **EDTF** (Extended Date Time Format)
- **Geo-json feature collection** (geospatial data)
- **Concept** (references controlled vocabulary)
- **Resource-instance** (references another Resource Model)
- **File-list** (for upload of digital files)
- **Annotation** (supports image annotations using a IIIF manifest)
## Original Date Data

<table>
<thead>
<tr>
<th>Built Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 million BCE</td>
</tr>
<tr>
<td>1745 - 1753</td>
</tr>
<tr>
<td>Autumn 1825</td>
</tr>
<tr>
<td>Definitely 1753, likely June 1753</td>
</tr>
<tr>
<td>March 22, 1751</td>
</tr>
</tbody>
</table>

**Fuzzy dates** are incomplete or approximate dates.
<table>
<thead>
<tr>
<th>Original Date Data</th>
<th>Standardized Date Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Built Date</strong></td>
<td><strong>Built Date</strong></td>
</tr>
<tr>
<td>5 million BCE</td>
<td>--</td>
</tr>
<tr>
<td>1745 - 1753</td>
<td>1749</td>
</tr>
<tr>
<td>Autumn 1825</td>
<td>1825</td>
</tr>
<tr>
<td>Definitely 1753, likely June 1753</td>
<td>1753</td>
</tr>
<tr>
<td>March 22, 1751</td>
<td>1751</td>
</tr>
<tr>
<td>Original Date Data</td>
<td>Standardized Date Data</td>
</tr>
<tr>
<td>--------------------</td>
<td>------------------------</td>
</tr>
<tr>
<td><strong>Built Date</strong></td>
<td><strong>Built Date</strong></td>
</tr>
<tr>
<td>5 million BCE</td>
<td>--</td>
</tr>
<tr>
<td>1745 - 1753</td>
<td>1749</td>
</tr>
<tr>
<td>Autumn 1825</td>
<td>1825</td>
</tr>
<tr>
<td>Definitely 1753, likely June 1753</td>
<td>1753</td>
</tr>
<tr>
<td>March 22, 1751</td>
<td>1751</td>
</tr>
</tbody>
</table>
Data Type Setting in Arches Designer

Node Data Type and Configuration

Data type
- date

Date Format
- CE Year (YYYY)

Node Settings

- Expose to Advanced Search
  Activate to use this node in Advanced Search.

- Required
  Activate to require that data be collected for this node when a card value is edited

- Export via Search Results
  Provide a field name for shapefiles. Name will be truncated to 10 characters to meet shapefile requirements.
  - shapefile fieldname

- Place node(s) in a separate card
  Data from this node and downstream nodes will be collected in a single form section
Fuzzy Date for construction of Tomb of Nefertari (QV 66):
approximately 1279 BCE to approximate 1212 BCE

Date expressed using EDTF:
-1279~/-1212~

How Arches encodes the approximate date

1 year padding to account for uncertainty
(setting can be changed in Arches Designer)
Fuzzy Date for construction of Tomb of Nefertari (QV 66):
approximately 1279 BCE to approximate 1212 BCE

Date expressed using EDTF:
-1279~/-1212~
Search: EDTF Date Format
Modeling for DISCOVERABLE DATA...

Resource Models optimize search by employing:
• normalized structures
• controlled vocabularies to standardize terminology.
RDM and Search: Controlled Vocabularies
Modeling for SHAREABLE and PERSISTENT DATA...

Resource Models enable data that is interoperable and help to ensure data longevity by encoding semantic metadata.

Semantic metadata corresponds to a semantic ontology or standard, such as the CIDOC-CRM.
ARCHES RESOURCE MODEL
A person is identified by name which has a type.
A person (E21) is identified by (P1) a name (E41) which has a type (P2) name type (E55)
A person (E21) is identified by (P1) a HANDLE (E41) which has a type (P2) HANDLE type (E55)
Arches Designer: Embedding Semantic Metadata
Arches Resource Models

• Data models created in Arches Designer
• Define the data structure and general organization of information in Arches implementation
• Generate the data entry forms for end users
Arches Resource Models enable...

- Data that is complete, high quality, discoverable, shareable and persistent
- A data management platform that is tailored to your use case and end users
Data Modeling requires...

• In-depth knowledge of your data and use case
• Experience or training in knowledge organization
Arches Resource Model Documentation and Resources
Arches Resource Model Working Group (ARM-WG)

- Working group of Arches community members with expertise in data modeling with the aim of generating and managing:
  - Arches Package Library, including Resource Models
  - Arches Modeling Documentation w/suggested methodology
  - Modeling discussion in Arches community
Arches Package Library

- **Package**: a configuration bundle that can be loaded into Arches. Includes resource models, branches, vocabularies, ontologies, sample data, and other configuration settings.

- **Library**: web page resource with links to and information on submitted packages that can be used and adapted by anyone in the Arches community.
Arches Package Library

The Arches Package Library is a resource for Arches implementers who are looking for entire packages (or their constituent branches and resource models) to help to form the foundation for their own Arches instance.

It is a work-in-progress and is currently being managed by the Arches Resource Model Working Group.

What is a package?

Each package included in this library is a bundle of information and core components critical to an Arches instance. A package can include any combination of branches and resource models, thesauri and collection files, business data, html files and other files that configure an Arches instance. The goal behind making these packages available is to offer a means of inspecting resource models that already exist in an Arches instance and providing a starting point for creating a package for your instance.

How to use this library

Below, you will find links to the information pages for the available Arches packages in this library. Each page includes information on what's included in each package and the required Arches version. Each page also includes a link to the Arches instance that the package was initially created for as well, if applicable, as well as a link to the GitHub repository where the package can be cloned or downloaded. These packages can be used as-is or can be modified and adapted to fit your own data requirements.

For information on how to load a package, click here.

List of Available Packages

Arches v5 Demo Site

The complete branch and resource model diagrams and downloads for the Arches Project v5 Demo site (released 2020).

Access the demo here: http://v5demo.archesproject.org/
Arches Modeling Documentation

- Foundational information on structured data and data modeling
- Suggested Arches modeling methodology
- Additional modeling resources
Arches Modeling Documentation

The following documentation presents information, compiled by the Arches Resource Model Working Group, on using and creating Arches Resource Models and Branches for use in Arches implementations. Each guide will help Arches users to understand the basic concepts behind modeling in Arches, the ARM WG methodology for Arches Resource Models, and the benefits of adopting the ARM WG methodology, as well as other resources for more information.

This documentation is a WORK-IN-PROGRESS, and content will be continually added. Last update: November 17, 2020.

BASIC CONCEPTS

METHODOLOGY (coming soon!)

BENEFITS

GLOSSARY

ADDITIONAL LINKS

Arches Modeling Documentation: www.archesproject.org/arm-wg-documentation
Community Involvement

- Community-based discussion around modeling in Arches
- Contributions to the package library
- Participation in enhancing the modeling documentation
Coming Next!
## Intro to the Arches Platform Upcoming Webinars

### Part 1: Capabilities
- Data Management
- Data Discovery & Visualization

### Part 2: Data Management Tools
- Arches Collector – Mobile App
- Resource Manager
- Arches Designer
- Reference Data Manager

### Part 3: Data Technology Concepts
- Data Standards
- Semantics/Ontologies (incl. CIDOC CRM)
- Controlled Vocabularies
- Fuzzy Dates

### Part 4: Information Architecture
- Software Stack
- Software Standards
Arches Q&A

Annabel Lee Enriquez
Getty Conservation Institute

David Myers
Getty Conservation Institute

Dennis Wuthrich
Farallon Geographics Inc
- Thank you -

Share your feedback: contact@archesproject.org

Community Support - Arches Forum:  www.groups.google.com/archesproject
Webinar repository – www.archesproject.org/videos