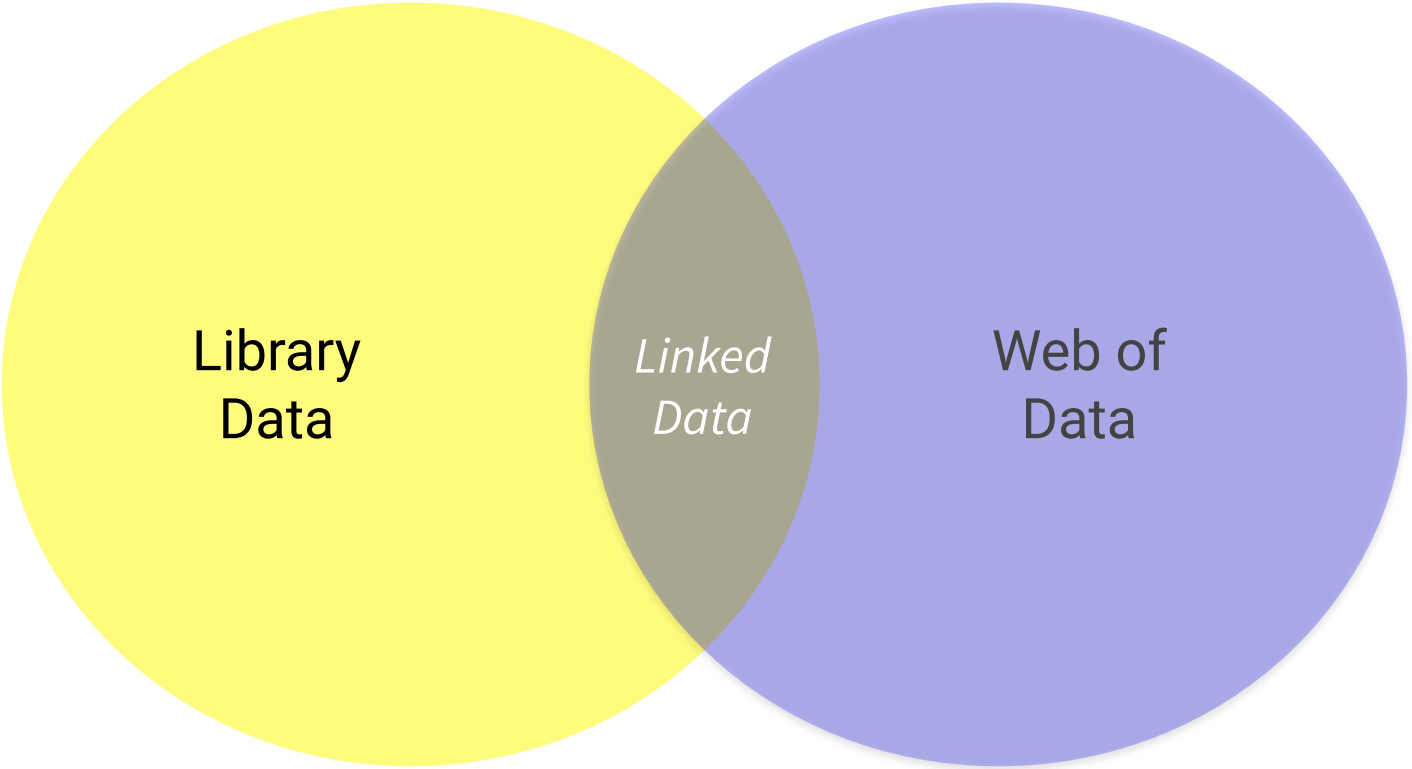
Discovery Futures: Linked Data & AI

Tom Cramer
Associate University Librarian
Stanford University
@tcramer

CNI "San Diego"
May 21, 2020

Stanford | LIBRARIES

Linked Data & Libraries



Linked Data is the gateway between library data and the Web.

- Get library data onto the web
- Get web data into the library

Discovery is where

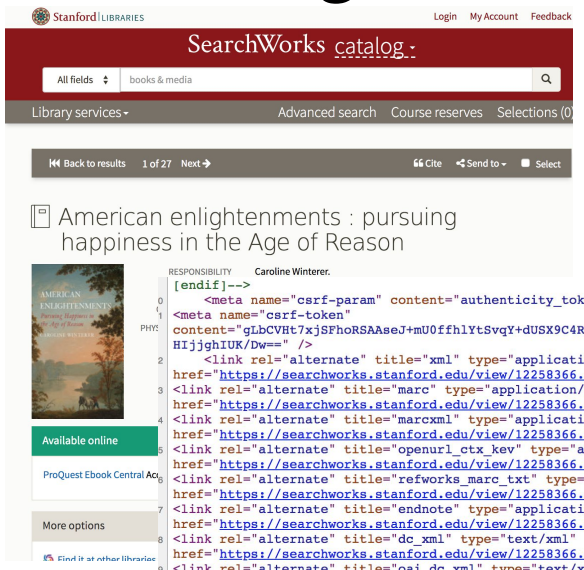
...we will see tangible benefits of LOD

Linked Data & Library Discovery

1. Schema.org (et al.) mark up
2. Knowledge Panels
3. Semantic Browse
4. Semantic Search
5. Type Ahead / Auto-Suggest

Schema.org (et al.) Mark Up

Catalog Item



Page Source

Schema.org Mark Up

Book

Canonical URL: http://schema.org/Book

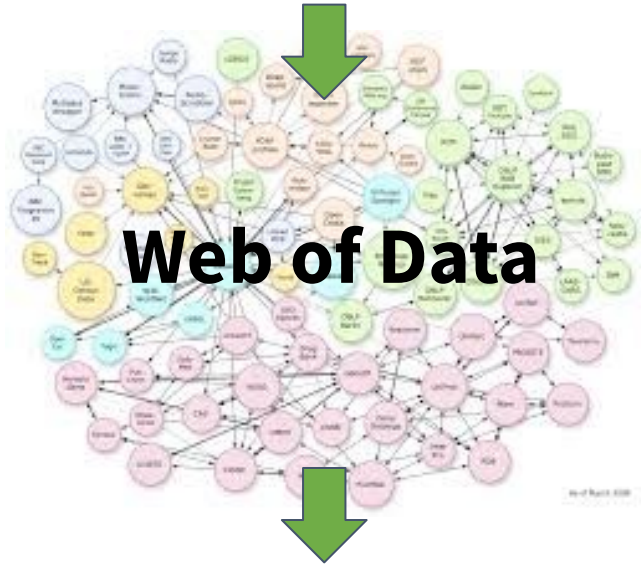
Thing > CreativeWork > Book

A book.

Usage: Between 10,000 and 50,000 domains

Property	Expected Type
Properties from Book	
abridged	Boolean
bookEdition	Text
bookFormat	BookFormatType

Harvesters



Web of Data

SEO

Knowledge Panel

LIBRARIES

UW ASK LOGIN GIVE

SEARCH
University of Wisconsin-Madison Libraries

ARTICLES CATALOG DATABASES JOURNALS UWDC

This item is also held by

UW-La Crosse

Publication Details

Cite/Export

CREATOR	by Gertrude Stein
FORMAT	Books
PUBLICATION	London : B.T. Batsford, Ltd., [1938]
PHYSICAL DETAILS	vii, 1 unnumbered page, 55 pages, 1 unnumbered page : color frontispiece, illustrations, plates (part color) portraits (part color) ; 23 cm
OCLC	ocm00477052

Data from other sources

Gertrude Stein Expand

Pablo Picasso Expand

Library Staff Details

Staff view

Data from other sources

Gertrude Stein Collapse

Brief Biography

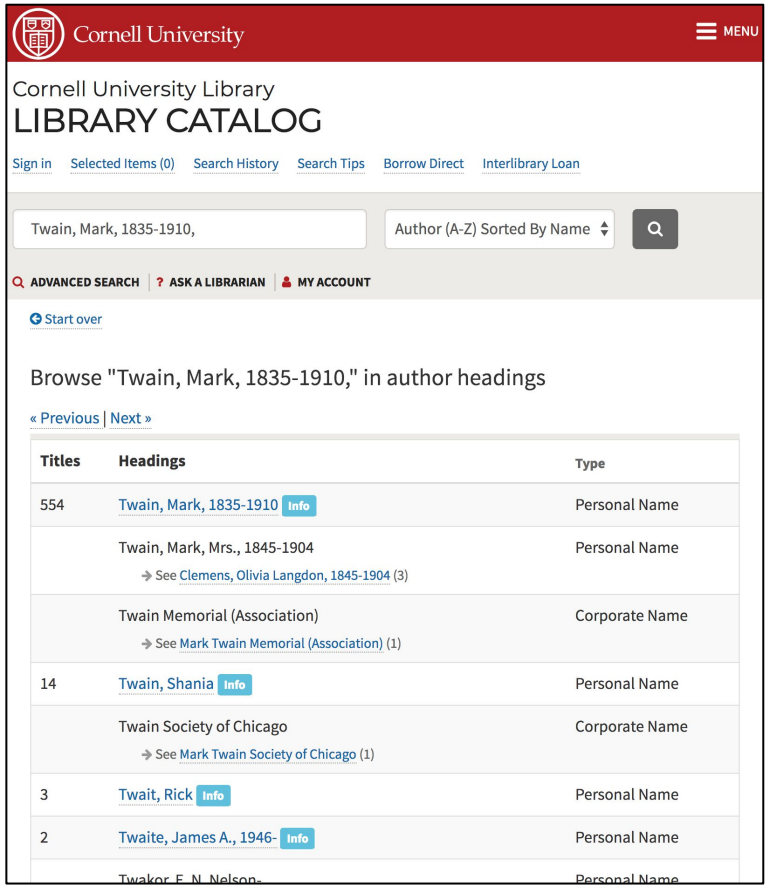
Description from Getty Linked Open Data: American avant-garde writer, patron, and collector of art. Together with her brother Leo, the Steins created a an important collection of works by Henri Manguin, Pierre Bonnard, George Braque, Pablo Picasso, Paul Cézanne, Pierre-Auguste Renoir, Honoré Daumier, Henri Matisse, and Henri de Toulouse-Lautrec. This collection was later split between the two of them. The salon frequented by artists and writers at the Stein residence at 27 rue de Fleurus in Paris is thought to have helped define Modernism in the early 20th century. Gertrude lived with her companion Alice B. Toklas from 1909 until her death in 1946. (Sources: Grove Art Online (2008-); Encyclopedia Britannica Online (2002-))

http://vocab.getty.edu/ulan/500273319

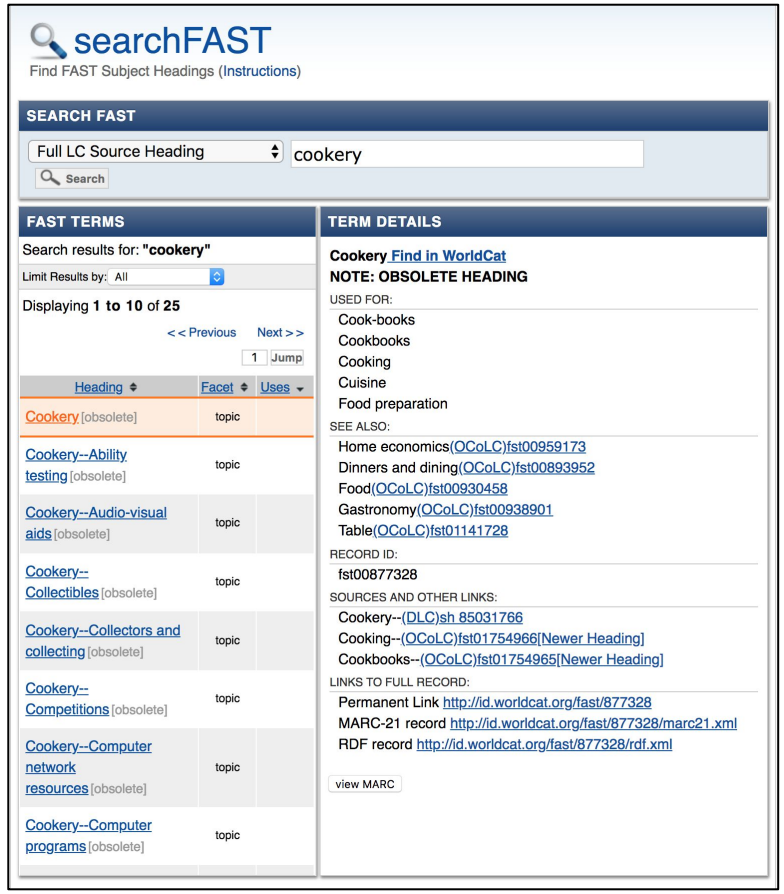
Abstract from DBPedia: Gertrude Stein (February 3, 1874 – July 27, 1946) was an American novelist, poet, playwright, and art collector. Born in the Allegheny West neighborhood of Pittsburgh, Pennsylvania, and raised in Oakland, California, Stein moved to Paris in 1903, and made France her home for the remainder of her life. She hosted a Paris salon, where the leading figures in modernism in literature and art, such as Pablo Picasso, Ernest Hemingway, F. Scott Fitzgerald, Sinclair Lewis, Ezra Pound, and Henri Matisse would meet. In 1933, Stein published a kind of memoir of her Paris years, The Autobiography of Alice B. Toklas, written in the voice of her life partner, Alice B. Toklas. Toklas was an American-born member of the Parisian avant-garde. The book became a literary bestseller and vaulted Stein from the relative obscurity of the cult literarure scene into the limelight of mainstream attention. Two quotes from her works have become widely known: "Rose is a rose is a rose is a rose" and "there is no there there", with the latter often taken to be a reference to her childhood home of Oakland, California. Her books include Q.E.D. (Quod Erat Demonstrandum) (1903) about a lesbian romantic affair involving several of

- BibCard implementation from University of Wisconsin
- https://search-ld.library.wisc.edu/
- “Gertrude Stein on Picasso”

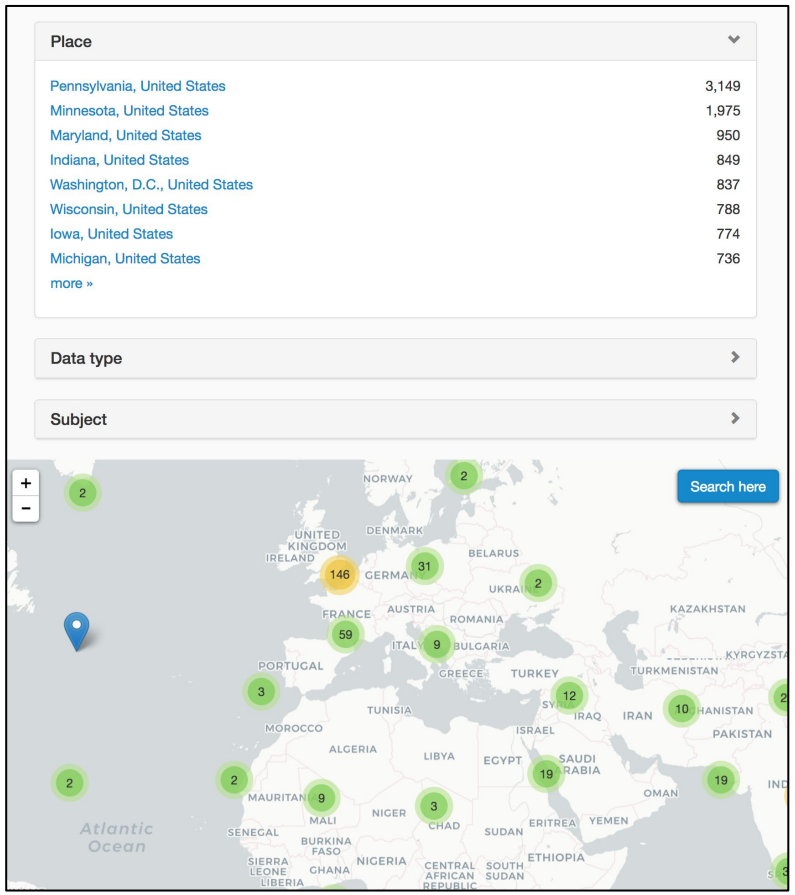
Semantic Browse



Author



Subject



Place

Semantic Search

NIH U.S. National Library of Me

MeSH

Search | Tree View | MeSH on Demand NEW

About MeSH Browser | Contact Us

heart attack

- All Terms
  - Main Heading (Descriptor) Terms
  - Qualifier Terms
  - Supplementary Concept Record Terms
- MeSH Unique ID

NIH U.S. National Library of Medicine

MeSH

Search | Tree View | MeSH on Demand NEW | MeSH 2017 | MeSH Suggestions

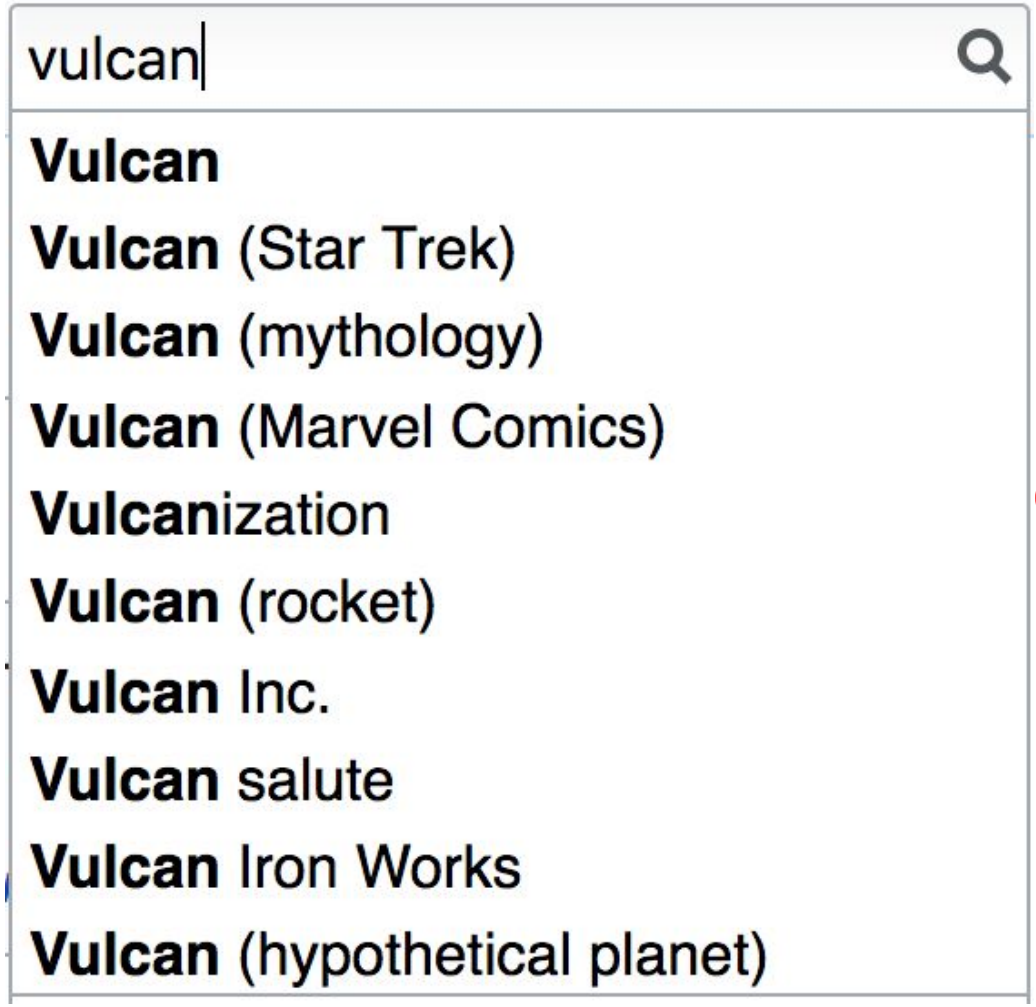
About MeSH Browser | Contact Us

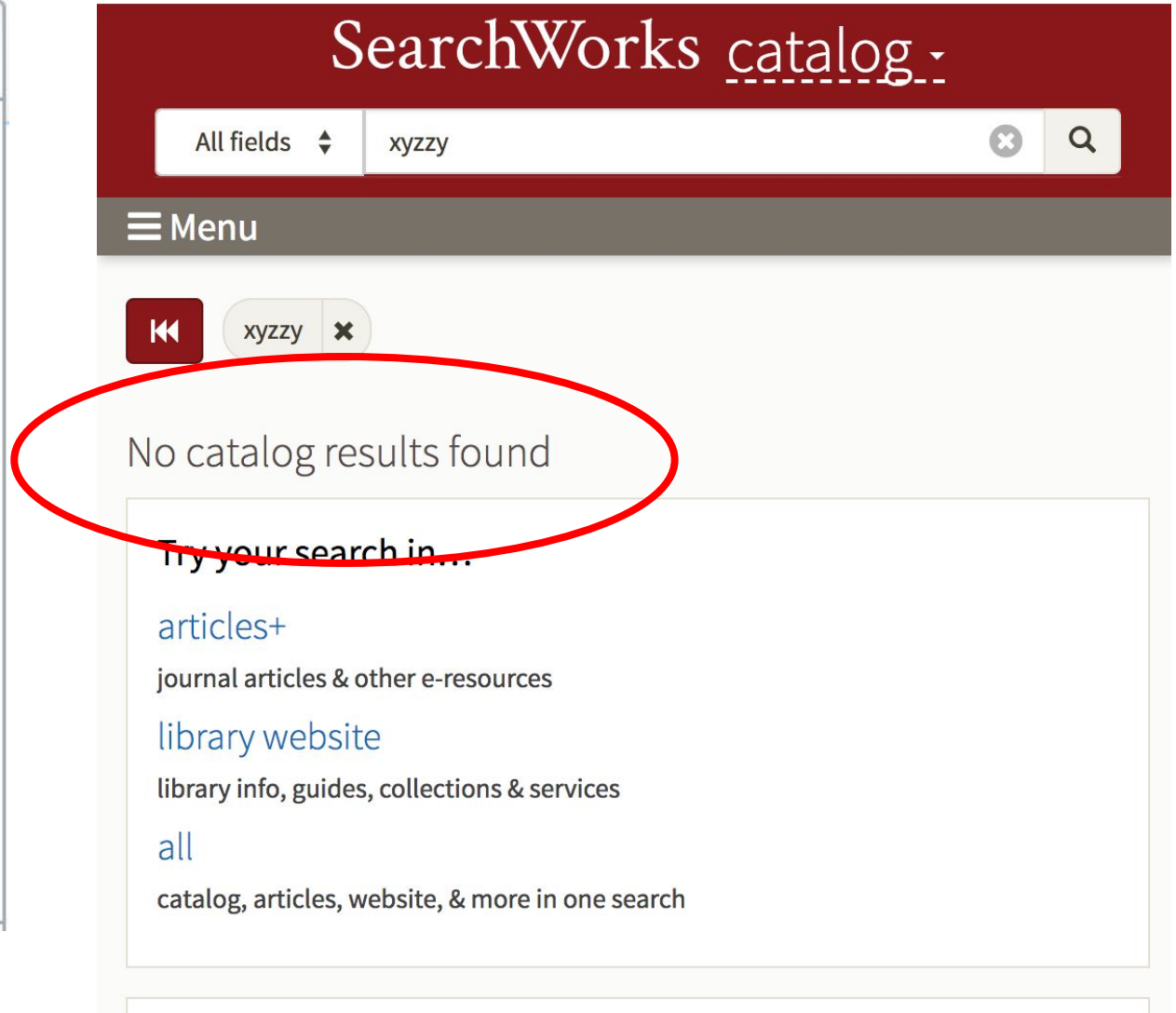
Myocardial Infarction MeSH Descriptor Data 2018

Details | Qualifiers | MeSH Tree Structures | Concepts

MeSH Heading	Myocardial Infarction
Tree Number(s)	C14.280.647.500 C14.907.585.500
Unique ID	D009203
Annotation	do not coordinate with ACUTE DISEASE for "acute infarct"
Scope Note	NECROSIS of the MYOCARDIUM caused by an obstruction of the blood supply to the heart (CORONARY CIRCULATION).
Entry Term(s)	Cardiovascular Stroke Heart Attack Myocardial Infarct
NLM Classification #	WG 310
See Also	Heart Rupture, Post-Infarction

Type Ahead / Auto-Suggest





LD4 Knowledge Panel Recipe

How to doc

DRAFT (see original outline in separate doc)

Table of Contents (proposed new structure)

Topics Covered

Link

https://bit.ly/ld4-kp-recipe

LD4 Discovery Affinity Group

Dashboard / Linked Data for Production: Pathway to Implementation (LD4P2) / Discovery (WP4)

LD4 Discovery Affinity Group Charge

Created by Simeon Warner, last modified by Huda Khan on May 12, 2020

Rationale

There is a pressing need to provide concrete demonstrations of the benefits of linked data in order to help motivate the transition of library catalogs from MARC. Within LD4P2, work to enhance the discovery of library resources is the focus of WP4: Discovery. The partners have committed to working with the Blacklight framework to demonstrate enhanced discovery features such as:

1. knowledge panel in search results to present contextual information powered by linked data;
2. browsing based on authority files and links to related entities in external data;
3. semantic search, which could include laying out alternative terms in a "no results" page, suggesting semantically related terms in a type-ahead, or providing richer, geo-based browse by leveraging URIs for places; and
4. microdata on item pages to enable machine crawling.

In all of these areas there is opportunity for community input on and discussion of the best approaches and the desired outcomes. The LD4 Discovery Affinity Group will provide a forum for these discussions, with the following goals:

- identify exemplar systems/features the demonstrate the benefits of linked data and could be applied to library resources
- identify opportunities for enhanced discovery of library resources using linked data that can be implemented in the short to medium term
- document connections between cataloging practices and discovery outcomes
- document effective assessment approaches
- provide advice to LD4P2 partners and cohort members for WP4: Discovery

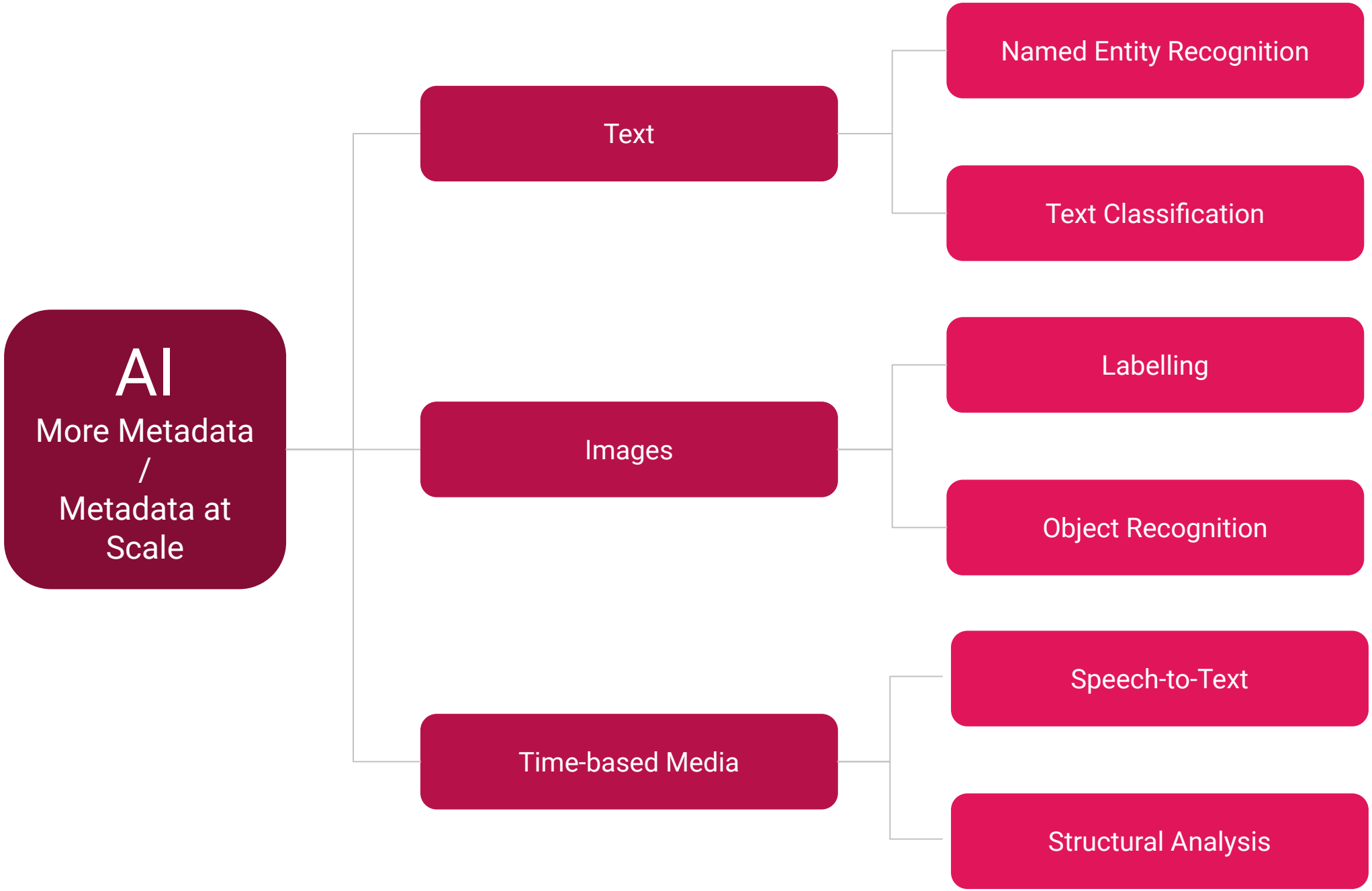
- Biweekly open calls
- Extensive knowledge base

https://wiki.lyrasis.org/display/LD4P2/LD4+Discovery+Affinity+Group+Charge

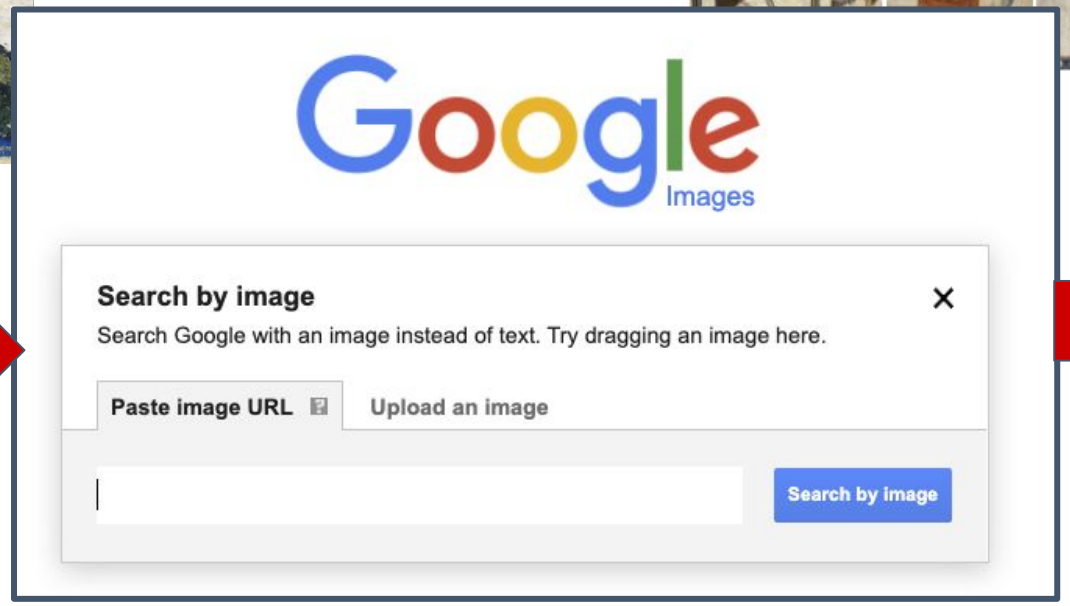
AI & Discovery

AI: Two Big Opportunities

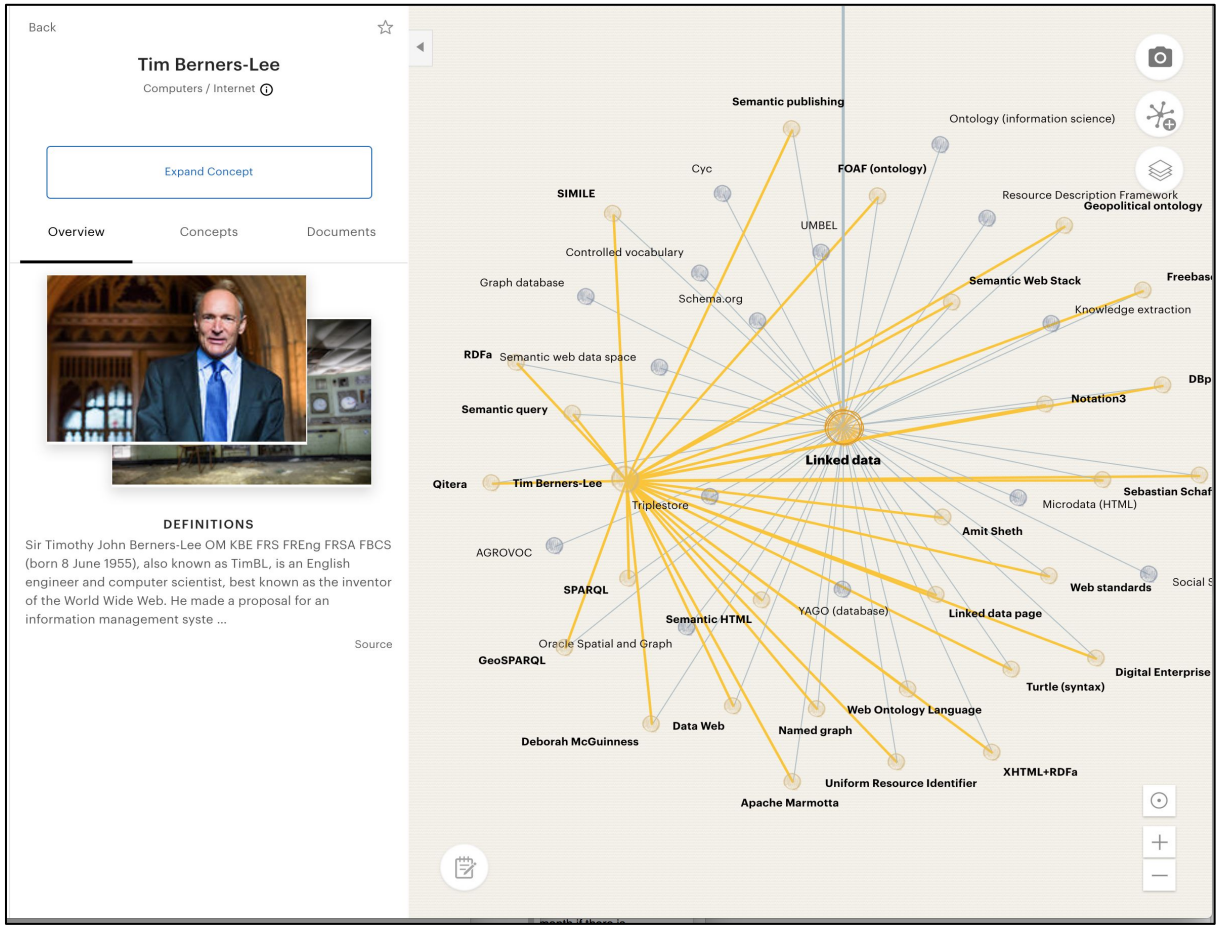
1. More (meta)data
2. New interfaces



New Interfaces

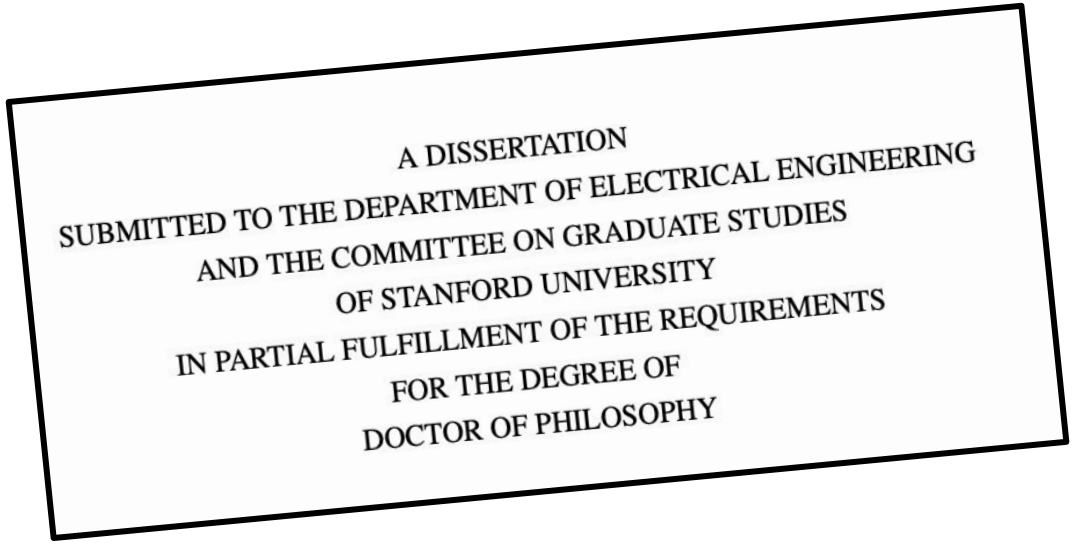


New Interfaces



Yewno Concept Explorer

Pathfinder Projects at Stanford


A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Enhanced Cataloging for Theses & Dissertations

Question: Can we use AI to *automatically* add keywords & subjects to catalog records?

Models / Engines: Yewno, ANNIF, BERT

Key Points: Integration with traditional cataloging workflow & ILS

Computer Vision for Archaeological Photos

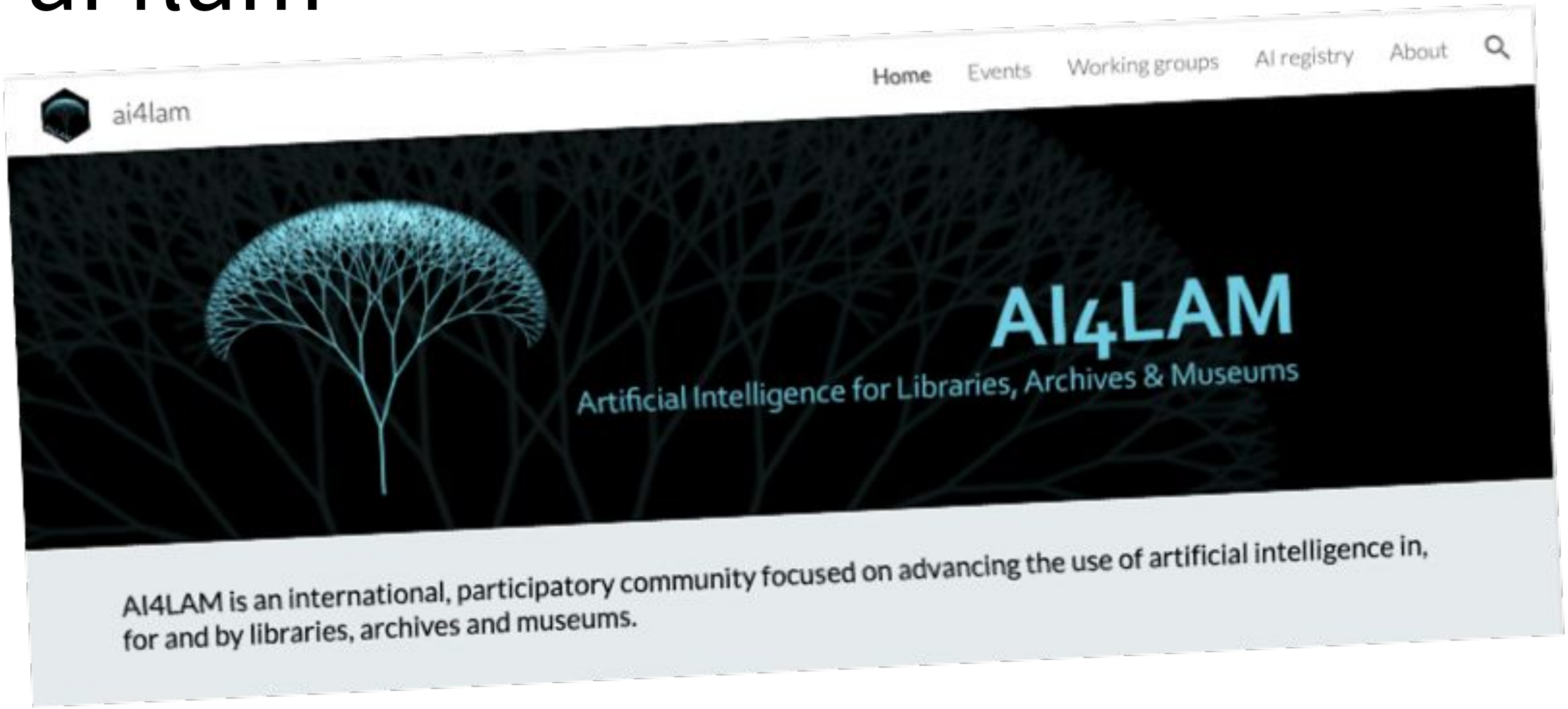
Question: Can we usefully classify & mine under-described photos from a 20 year dig?

Models & Engines: Claif.ai, Google CV, AutoML

Key Points: Comparison of commercial vs. custom models, integration with dig records

http://catalhoyuk.com/research/imageai

ai4lam



https://ai4lam.org