Open Data Perspectives and Problems: An Archaeological Perspective

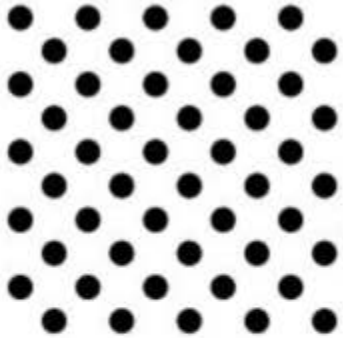
Presented in:
Sharing is Caring:
Challenges and Opportunities in Producing Open Data

Lynne Goldstein
Professor of Anthropology
Michigan State University

Archaeology and Big Data

BIG DATA
VOLUME
DATA SIZE
VELOCITY
SPEED OF CHANGE
VARIETY
DIFFERENT FORMS OF DATA SOURCES
VERACITY
UNCERTAINTY OF DATA

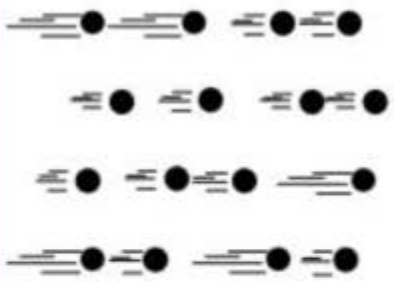
Volume



Data at Rest

Terabytes to exabytes of existing data to process

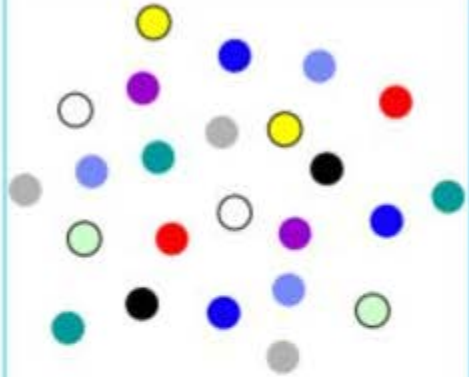
Velocity



Data in Motion

Streaming data, milliseconds to seconds to respond

Variety



Data in Many Forms

Structured, unstructured, text, multimedia

Veracity*

Data in Doubt

Uncertainty due to data inconsistency & incompleteness, ambiguities, latency, deception, model approximations

Volume:

Archaeological data represent a large volume, but not when compared to physics or medicine

Velocity:

Archaeologists do relatively little data streaming, so velocity is not one of our problems

Variety:

In this area, we beat almost every other discipline

Veracity:

Another big area for archaeologists because we use “messy” data that is not structured and often not comparable

Legacy Data

- Messy
- Not disciplinarily structured
- Long time depth – data collected 150 years ago can be very significant

Data Sensitivity: A Huge Issue

- Most people unaware that archaeological data are sensitive
- Site files exempt from Freedom of Information Act requests
- Compliance work
- Ownership of data
- Legal issues

Are There Ways to Share or Open Archaeological Data?

Examples

Preservation Depositories

Find & share research

As a repository for digital data, tDAR contains a vast amount of resources and tools that can enrich your ongoing research.

- Search our expansive database of documents, data sets, images and more.
- Integrate your data with data from other projects.
- Collaborate with other researchers

LEARN MORE

Preserve & protect data

The core mission of tDAR is to support better stewardship of the documents and data produced during research.

- Manage all of your archaeological data in one place.
- Organize your resources.
- Download reports or bibliographies.

LEARN MORE

Extend your classroom

As a discipline-specific archive of archaeological documents, tDAR is dedicated to the teaching of archaeology.

- Use reports, documents, data sets and images in classroom assignments and lectures.
- Find reports and grey literature from investigations from around the world.
- Compare different data sets and analyses.

LEARN MORE

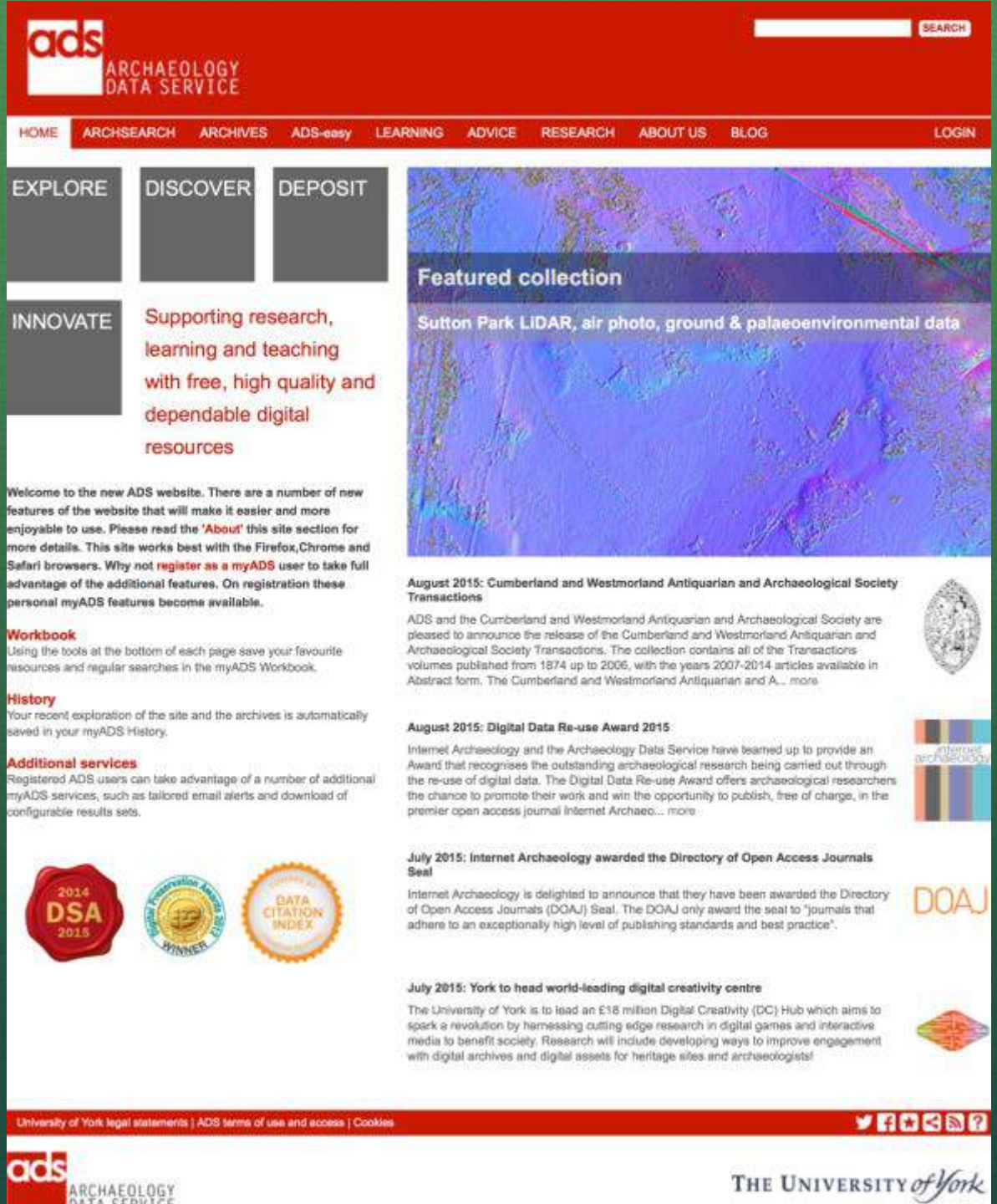
Meet Your Legal Obligations

Use tDAR to meet legal obligations to preserve and provide access to digital archaeological files, which mandate that data generated by federal agencies must be deposited in a repository capable of:

- Providing long-term digital curation;
- Ensuring preservation of the digital files; and
- Facilitating accessibility to qualified users.

LEARN MORE

TDAR: The Digital Archaeological Record



ADS: Archaeology Data Service, Great Britain

Examples

Open Data

Home About Projects Browse Lightbox Tables Weblog Search

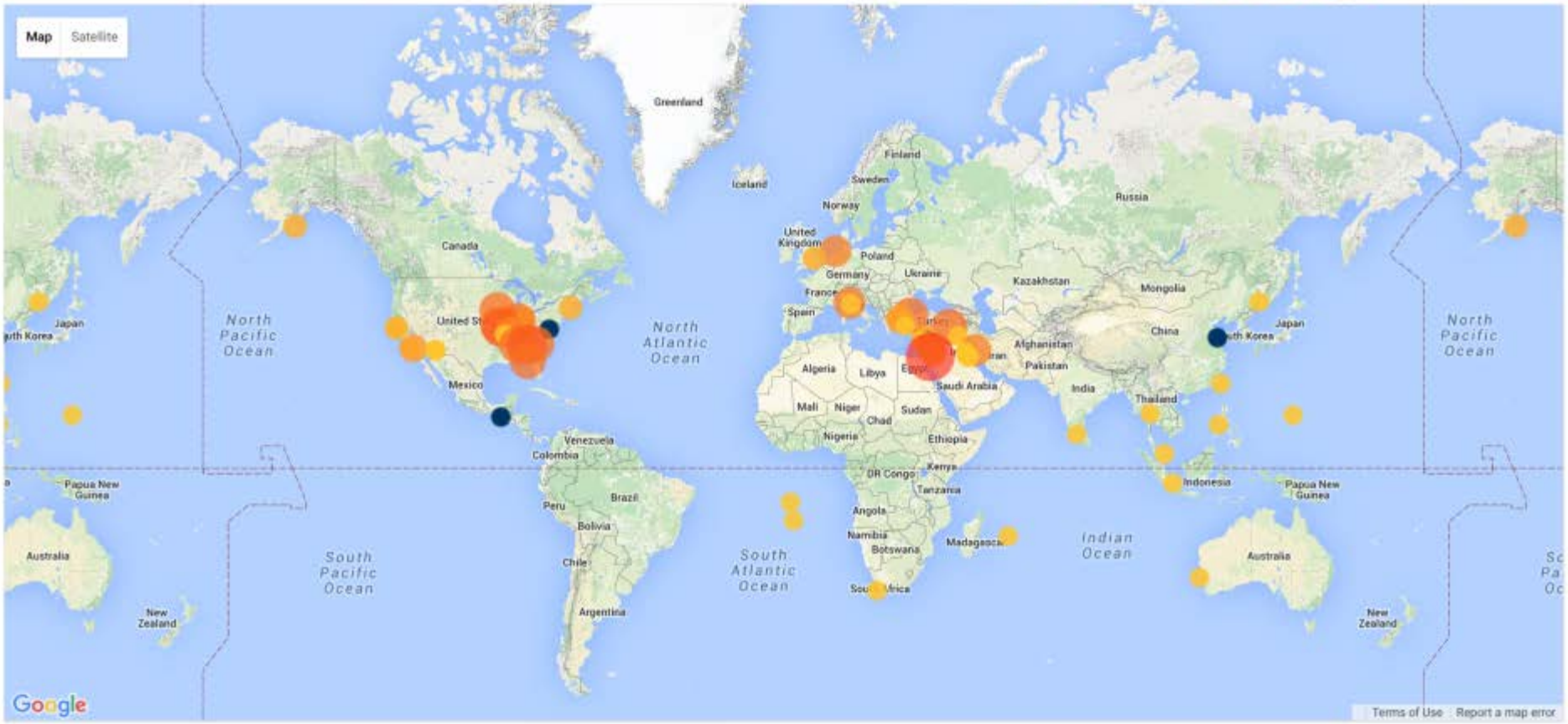
Welcome to Open Context

Open Context reviews, edits, and publishes archaeological research data and archives data with university-backed repositories, including the California Digital Library.

Donate

Make a donation or volunteer services

Content Map and Timeline



Example

DINAA:
Digital Index of North American Archaeology

Digital Index of North American Archaeology (DINAA)

The Digital Index of North American Archaeology (DINAA) is a multi-institutional undertaking to create interoperability models for archaeological site databases in the eastern United States. The 2-year project, which is funded by the National Science Foundation, commenced September 1, 2012. View the technical proposal.

This page provides ongoing information about the progress of the project. Visit the links below to learn more about this project. We also broadcast updates to the community via the NASA (National Association of State Archaeologists) listserv. We welcome your questions, comments, and other participation in this project, so please comment on the blog posts or contact us directly by emailing project PIs David Anderson and Josh Wells.

- Project Overview (Aug. 6, 2012)
- Initial Announcement (Aug. 7, 2012)
- Initial Communication with the National Association of State Archaeologists (NASA) (Aug. 24, 2012)
- Sensitive Data Security Measures and SHPO Collaboration (Aug. 24, 2012)
- Data Transfer Instructions (Oct. 31, 2012)
- Second Communication with the National Association of State Archaeologists (NASA) (Nov. 2, 2012)
- DINAA Annual Report (October 2013)
- Florida and Georgia Site Files Launch DINAA Project (December 2013)
- Program for the DINAA Workshop at the University of Tennessee, Knoxville (March 19-20, 2014)
- **Explore a draft release of the data**, updated in August 2014 with Alabama and Virginia site file records

Additional Project News and Updates:

DINAA @ SAA

The DINAA project had a very successful reception at the 79th Annual Society for American Archaeology conference in Austin, TX last month. With several presentations spanning traditional conference papers, posters, and a lightning talk at the Digital Data Interest Group meeting, conference-goers were provided with no shortage of opportunities to learn about multiple aspects of […]

SUBSCRIBE

Site RSS Feed

About Heritage Bytes

Unless otherwise noted, the Heritage Bytes Weblog by The Alexandria Archive Institute is licensed under a Creative Commons Attribution 3.0 Unported License.
Please cite: ux.opencontext.org.

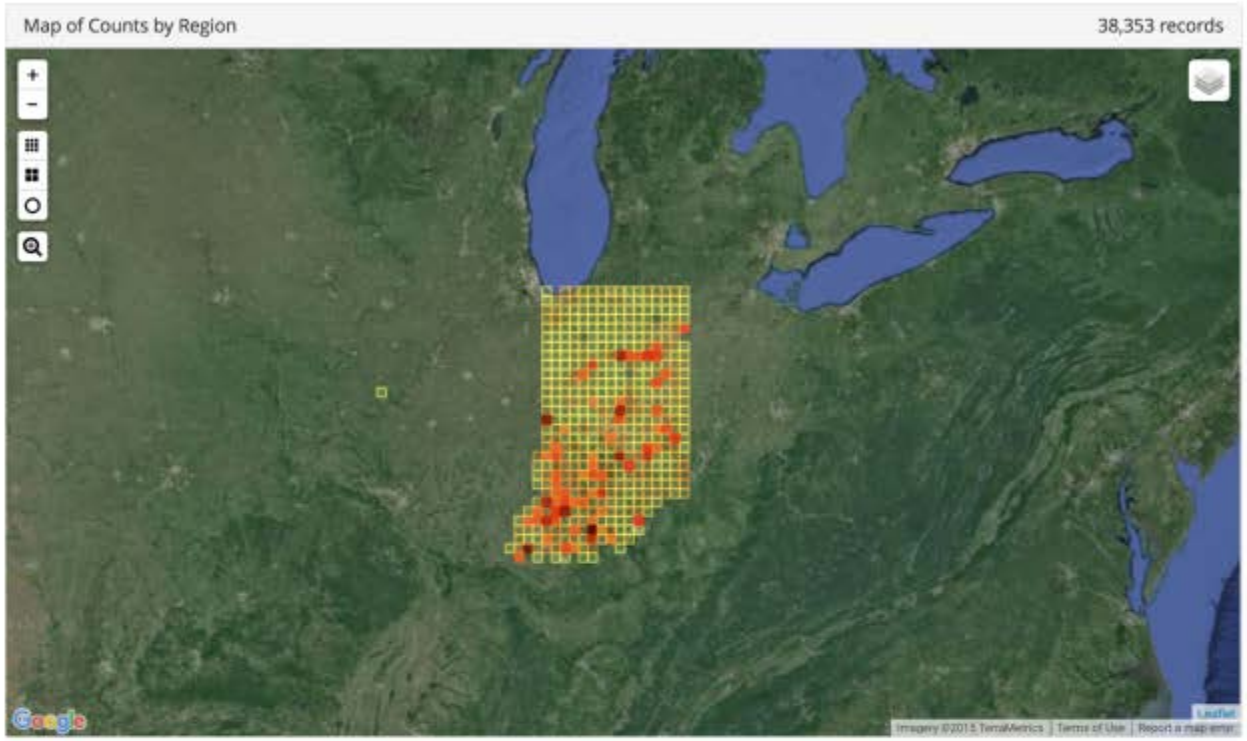
Blogroll

- Documentation
- Plugins
- Suggest Ideas
- Support Forum
- Themes
- WordPress Blog
- WordPress Planet

Tags

archaeology architecture book collaboration common needs common tools communication community needs cultural heritage cuneiform Cyberinfrastructure database data publication data publishing data sharing digital humanities DINAA dream tools Egypt Greco-Roman Greece Israel/Palestine Linked Open Data maps Medieval Mesoamerican Mesopotamia museum open access open context Open Data policy primary content repository scan Scholarly Communications site review survey user communities user experience user needs web tools working groups workshop zooarchaeology

DINAA Search for Indiana Archaeological Sites



Applied Filters

General Keyword Search
Search Term: 'Indiana'

Description
Has period

Project
Digital Index of North American Archaeology (DINAA)

General Keyword Search

Indiana

Filtering Options

- Has period
  - Period 38,353
- Context
- Digital Index of North American Archaeology (DINAA)
- Descriptions (Common Standards)
- Descriptions (Project Defined)
- Numeric Descriptions (Project Defined)
- Date Descriptions (Project Defined)

Paging

Search results 1 to 20 of 38,353

Search Results

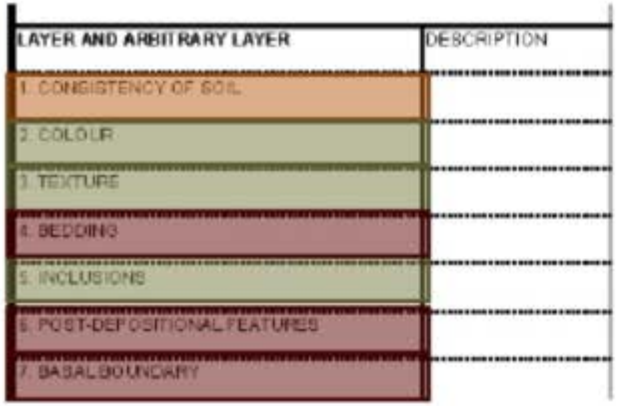
Item Label	Project	Context	Category	Early (BCE/CE)	Late (BCE/CE)	Thumbnail

Linked Open Data

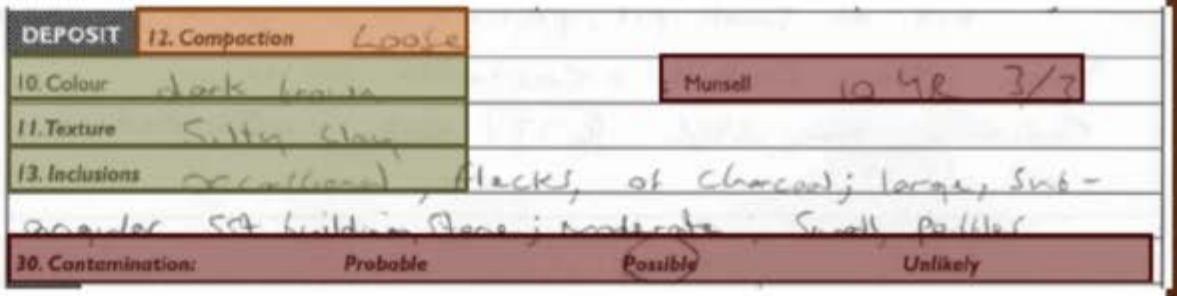
Is this a solution for the problems of managing and sharing archaeological data?

Documentation

- Different excavation methods bring differing documentation
- Comparison of different documentation sheets



LAYER AND ARBITRARY LAYER	DESCRIPTION
1. CONSISTENCY OF SOIL	
2. COLOUR	
3. TEXTURE	
4. BEDDING	
5. INCLUSIONS	
6. POST-DEPOSITIONAL FEATURES	
7. BASAL BOUNDARY	



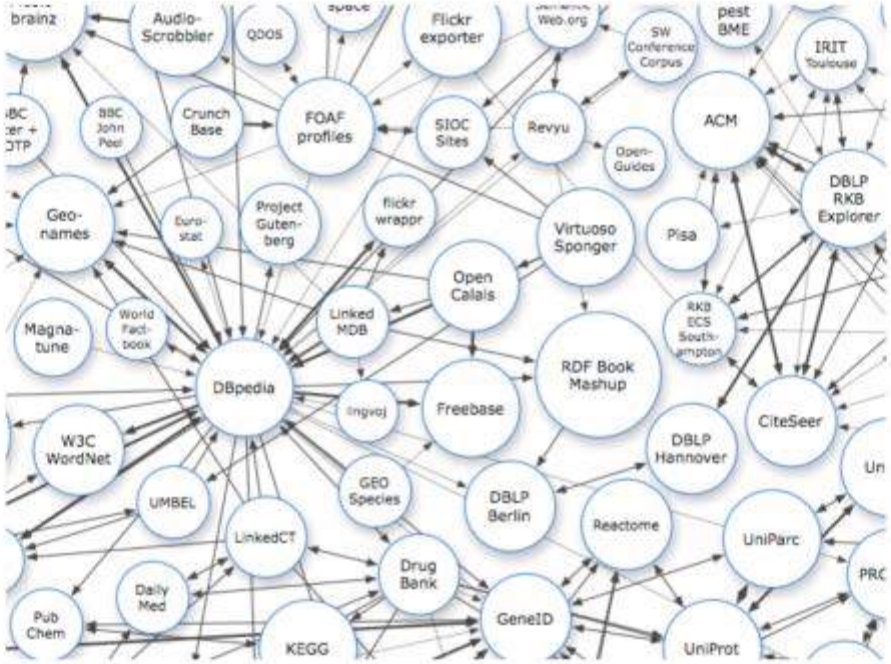
Linked Data - Connect Distributed Data across the Web

- Home
- Guides and Tutorials
- Frequently Asked Questions
- Glossary
- Images and Posters
- Presentations
- Data Sets
- Tools
- Events
- Calls for Papers
- Research
- News and Blogs
- Domains
- See Also

Linked Data

Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. More specifically, Wikipedia defines Linked Data as "a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF."

This site exists to provide a home for, or pointers to, resources from across the Linked Data community.



Part of the Linking Open (LOD) Data Project Cloud Diagram, click for full and historical versions...

Login to edit (authorised users only at present)

Username: *

Password: *

Log in

- Log in using OpenID
- Request new password

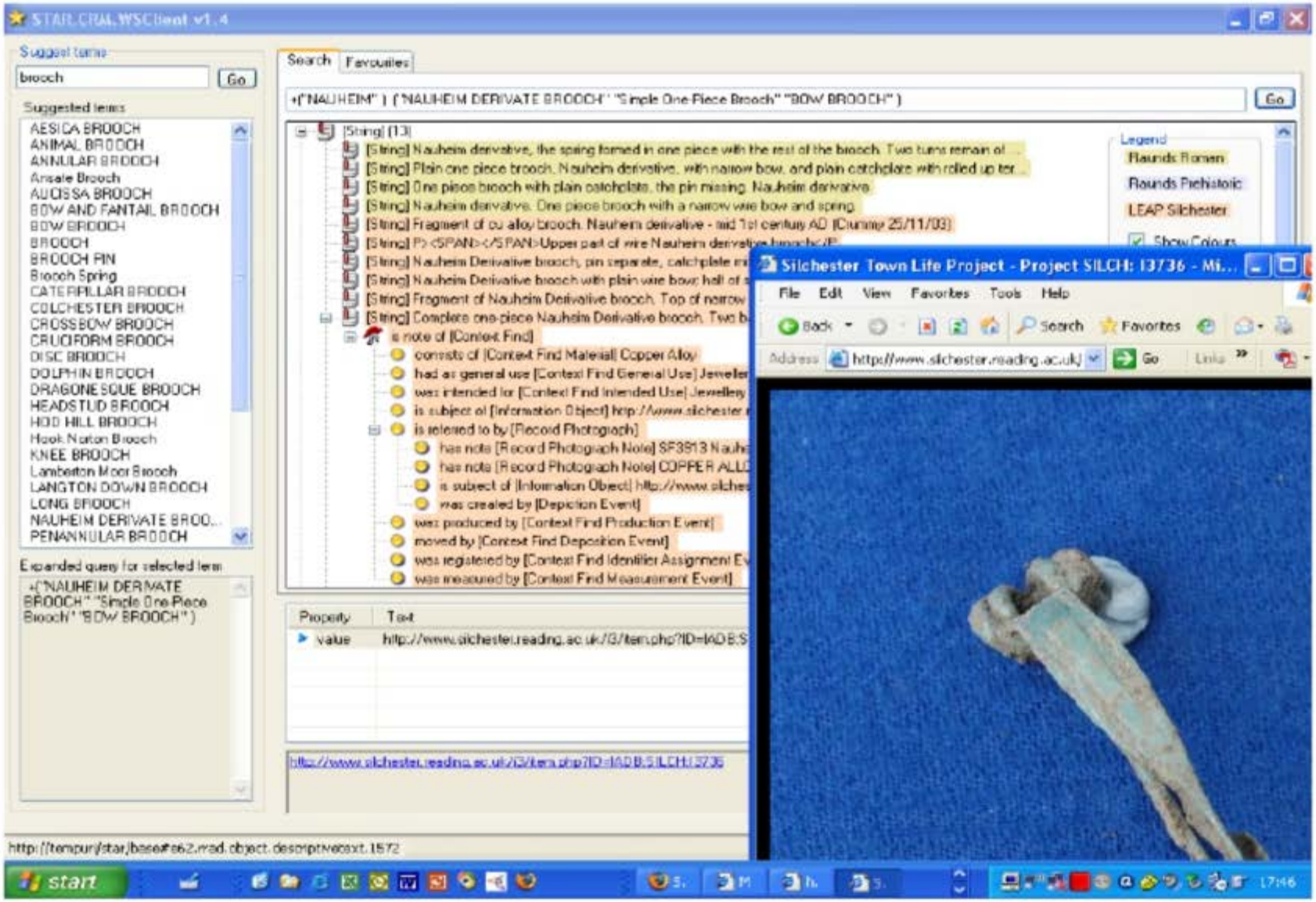
Navigation

- Feed aggregator

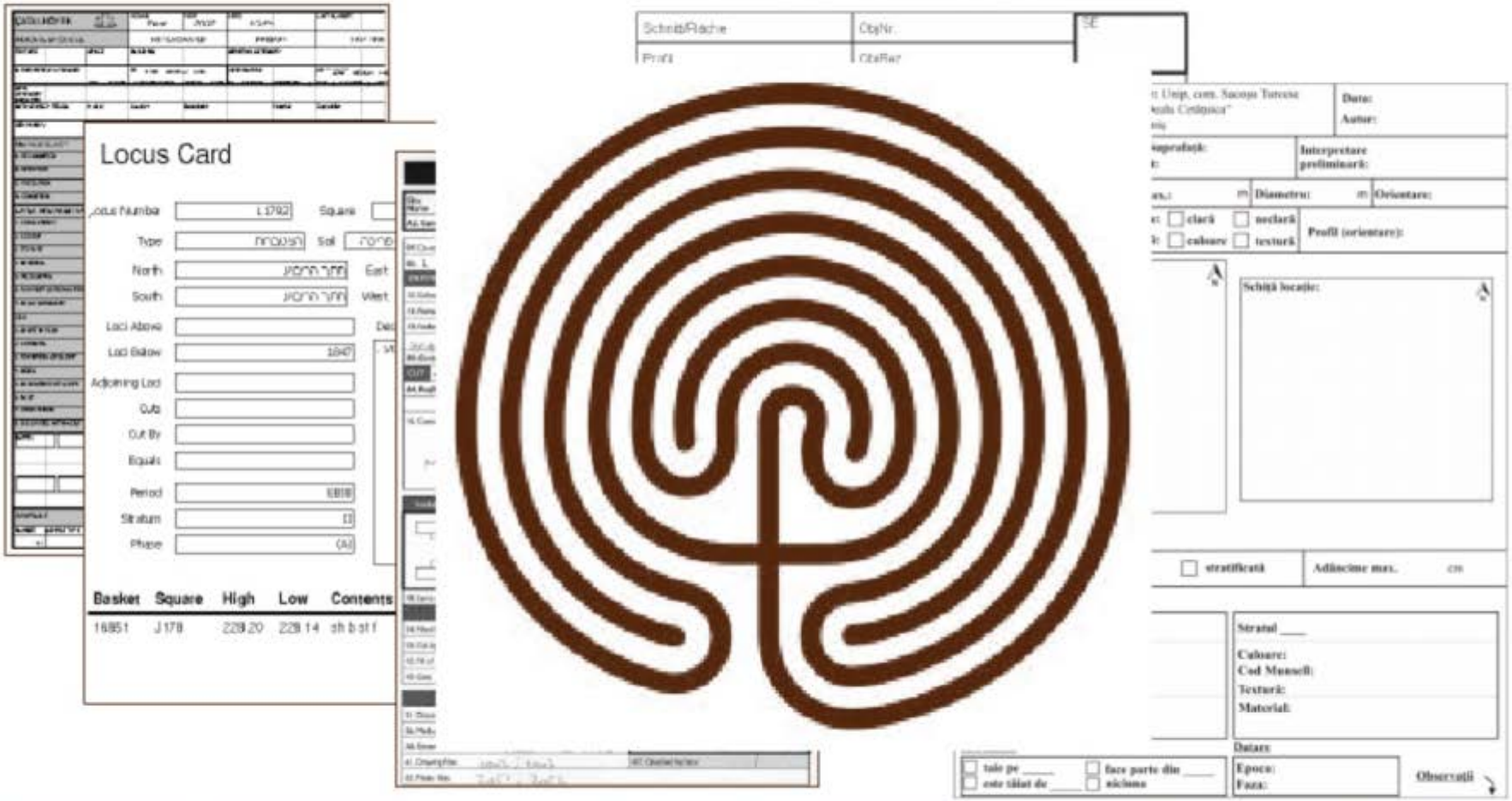
Syndicate

About this Site

Prototype Controlled Vocabulary searching



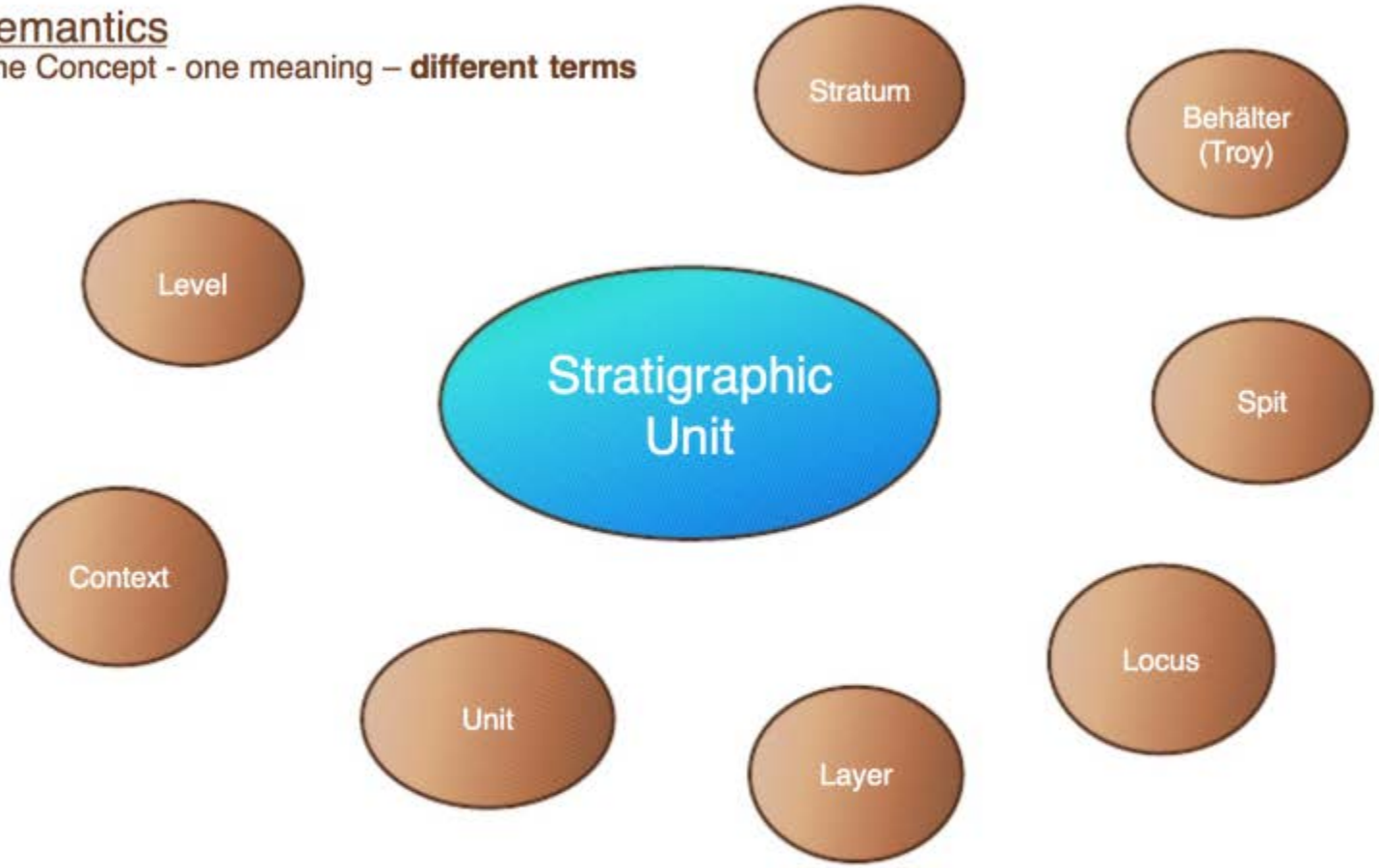
What about comparing records across different countries?



ARIADNE

Semantics

One Concept - one meaning – **different terms**



ARIADNE

Barrier: Semi-controlled vocabularies…

Deposit Colour			Deposit Texture	Deposit Compaction
(Reddy) Brown	Dark orange/brown	Orangy brown, very light	Firm	Plastic
9Reddy) brown				Sticky
Brown				y (wet)
Brown red				y/firm
Brown/redd				s
Dark brown				
Dark brown				
Dark grey b				
Dark orange				
Dark orange				
darker patche				
Dark orange loam		Yellow brown		
		Yellow/orange brown		

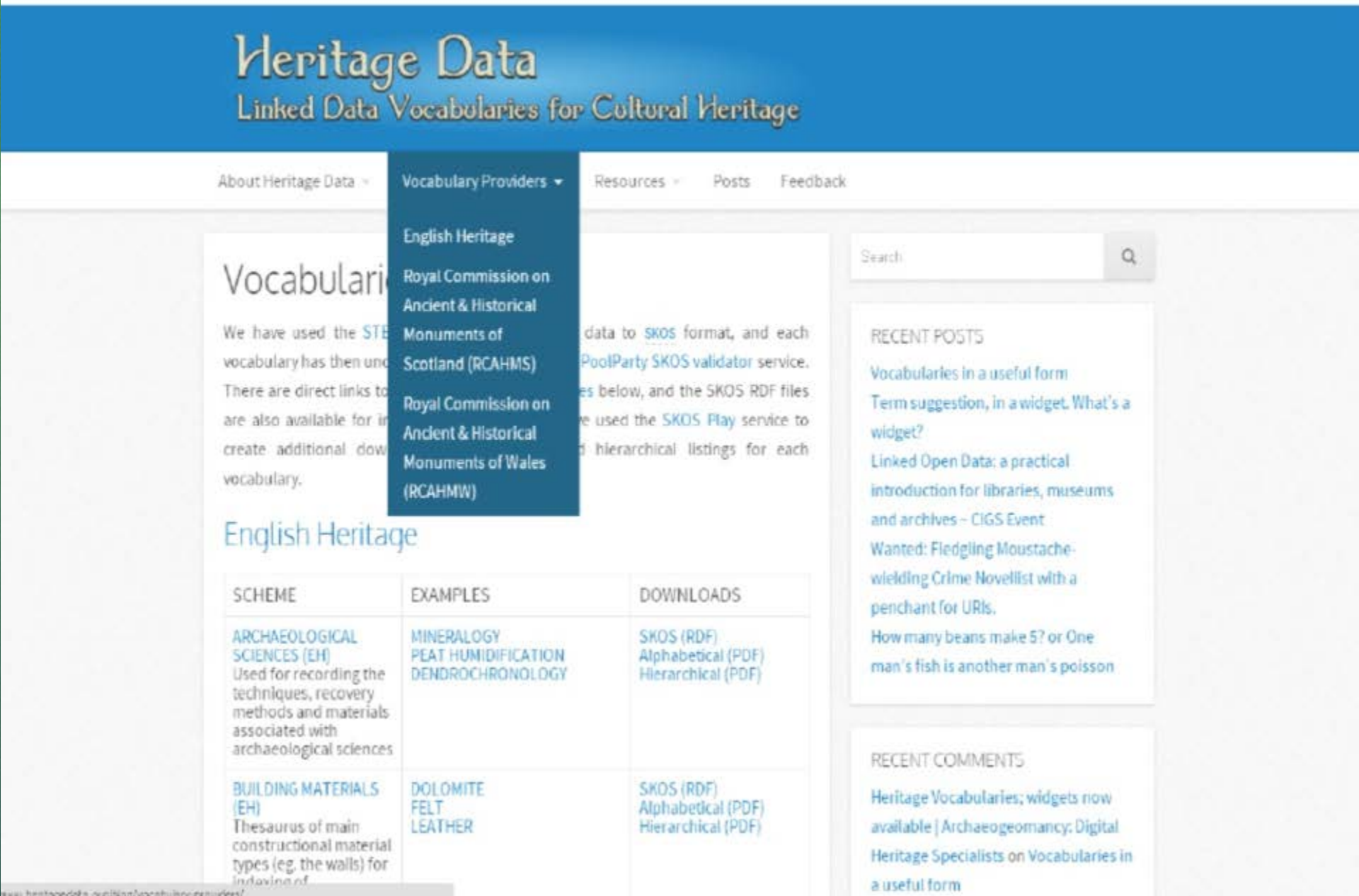
"…another of my examples has something about some flint that is 'snuff coloured' & I don't know if I've ever seen snuff, let alone know what colour it is, or might have been over 150 years ago, and I would think it would make sense to take some kind of integrated approach from the outset,…." [G. Carver]

For data entry: Semi-controlled vocabularies represent a useful compromise somewhere between descriptive & controlled vocabularies, ***the best of both worlds!***

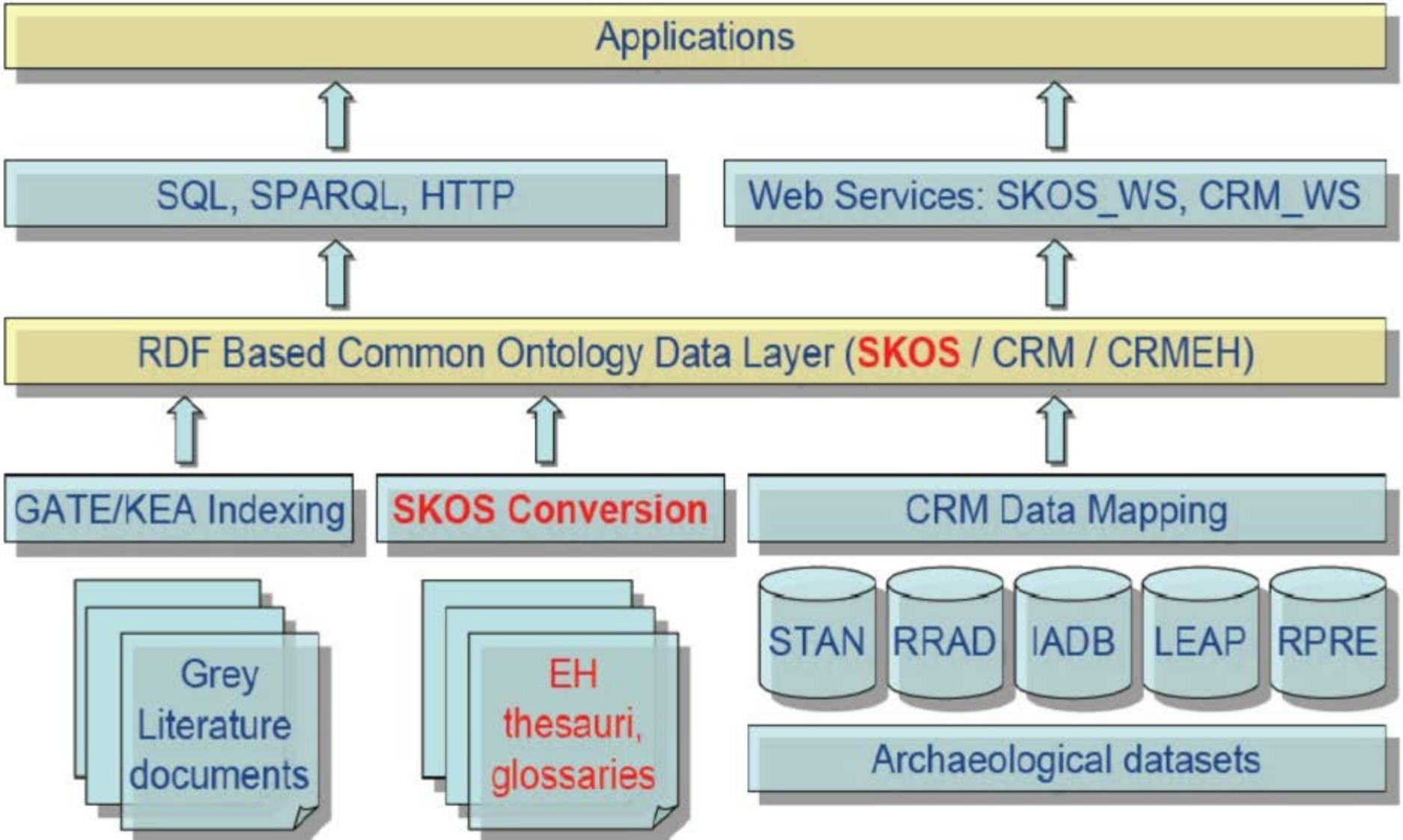
For data retrieval: ***The worst of all worlds*** *(Re. find all the iron age post holes)*

This problem arises from trying to do two different things within a single input field. Should do both, but separately – ***1) describe using free text description fields,*** *and* ***2) index using controlled index fields***

LOD Heritage Vocabularies: http://www.heritagedata.org



STAR - General Architecture



Example: Aztalan

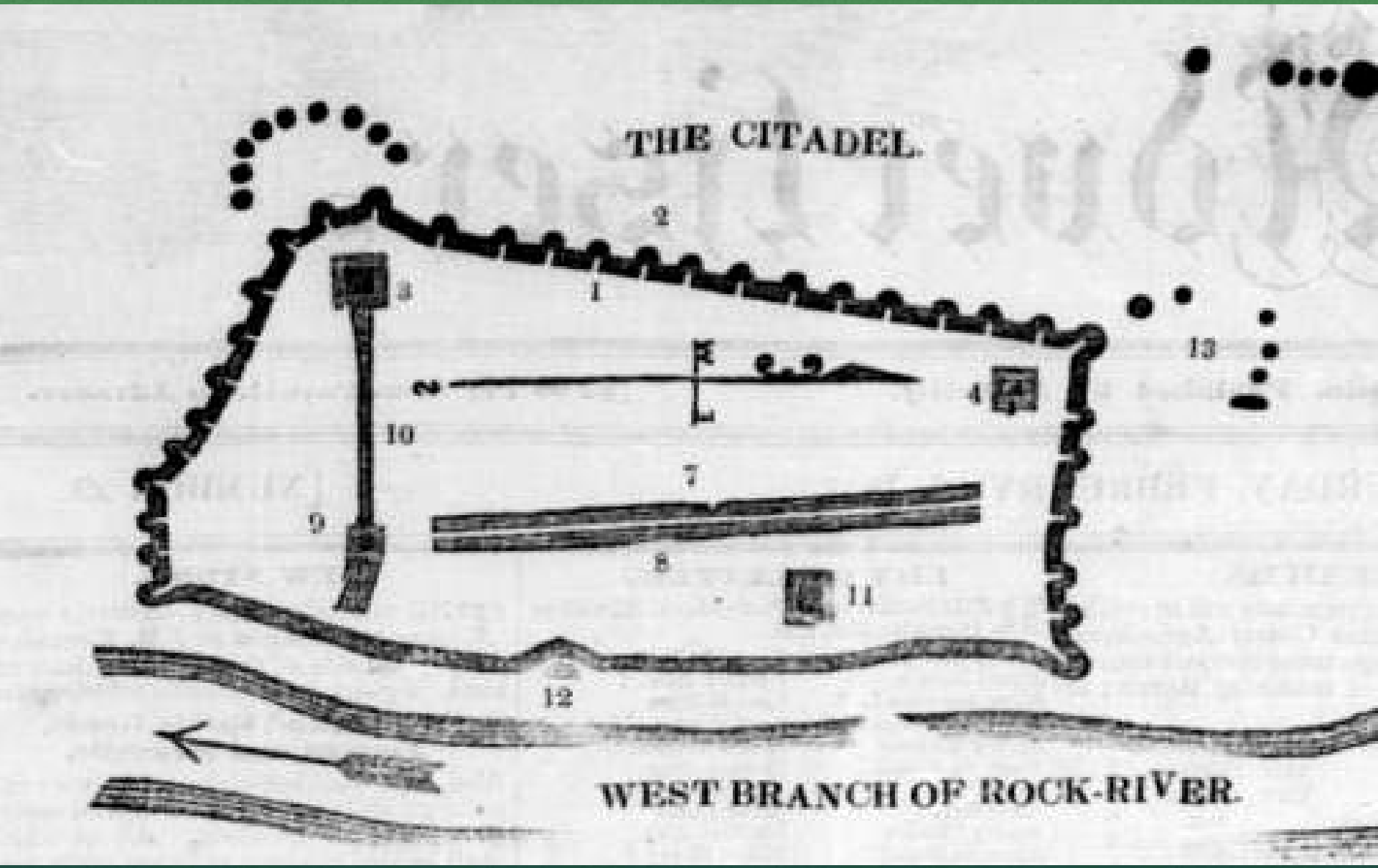
- Aztalan – a prehistoric archaeological site in Southern Wisconsin dating AD 800-1200.
- Has been examined and excavated by both amateur and professional archaeologists for over 175 years

The Aztalan Site [47 JE001] Jefferson County, Wisconsin

Location of Aztalan State Park

Aztalan - Looking west

Aztalan – looking south


THE CITADEL.
2
3
1
13
4
10
7
9
8
11
12
WEST BRANCH OF ROCK-RIVER.

WATERTOWN CHRONICLE.

VOL. IV.—NO. 32. WATERTOWN, WIS.—WEDNESDAY, JANUARY 22, 1851. WHOLE NO. 188.

WATERTOWN CHRONICLE.

BY J. A. HADLEY.

Terms of Subscription and Advertising.

List of Letters.

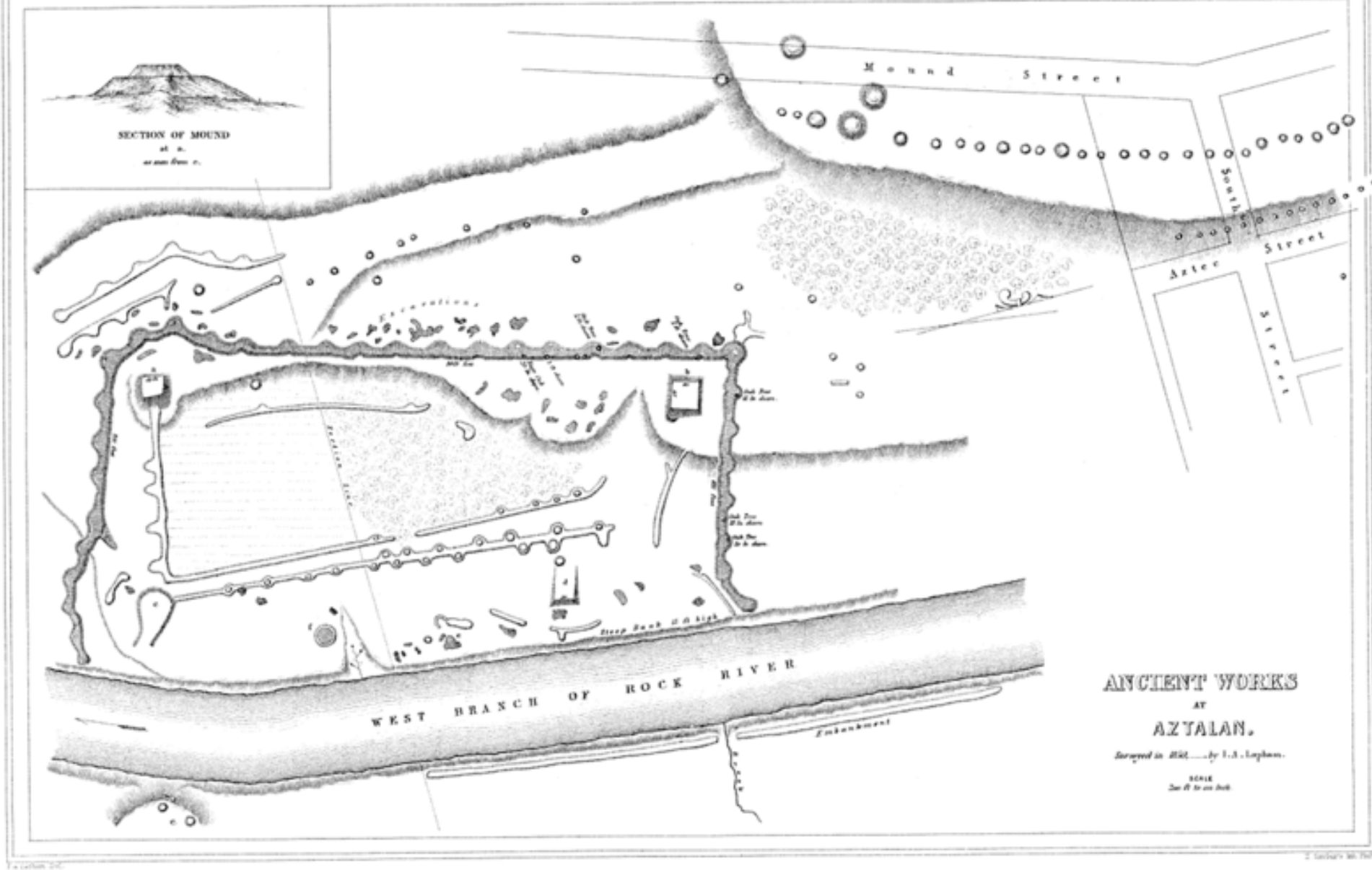
THE CHRONICLE.

Second Annual Pioneer Festival

Timothy Johnson's Narrative.

Luther A. Cole's Narrative.

Remarks of David Noggle, Esq.


SECTION OF MOUND
Mound Street
Aztec Street
WEST BRANCH OF ROCK RIVER
ANCIENT WORKS
AT
AZTALAN.

Madison July 6 – 1850

Dear Sir

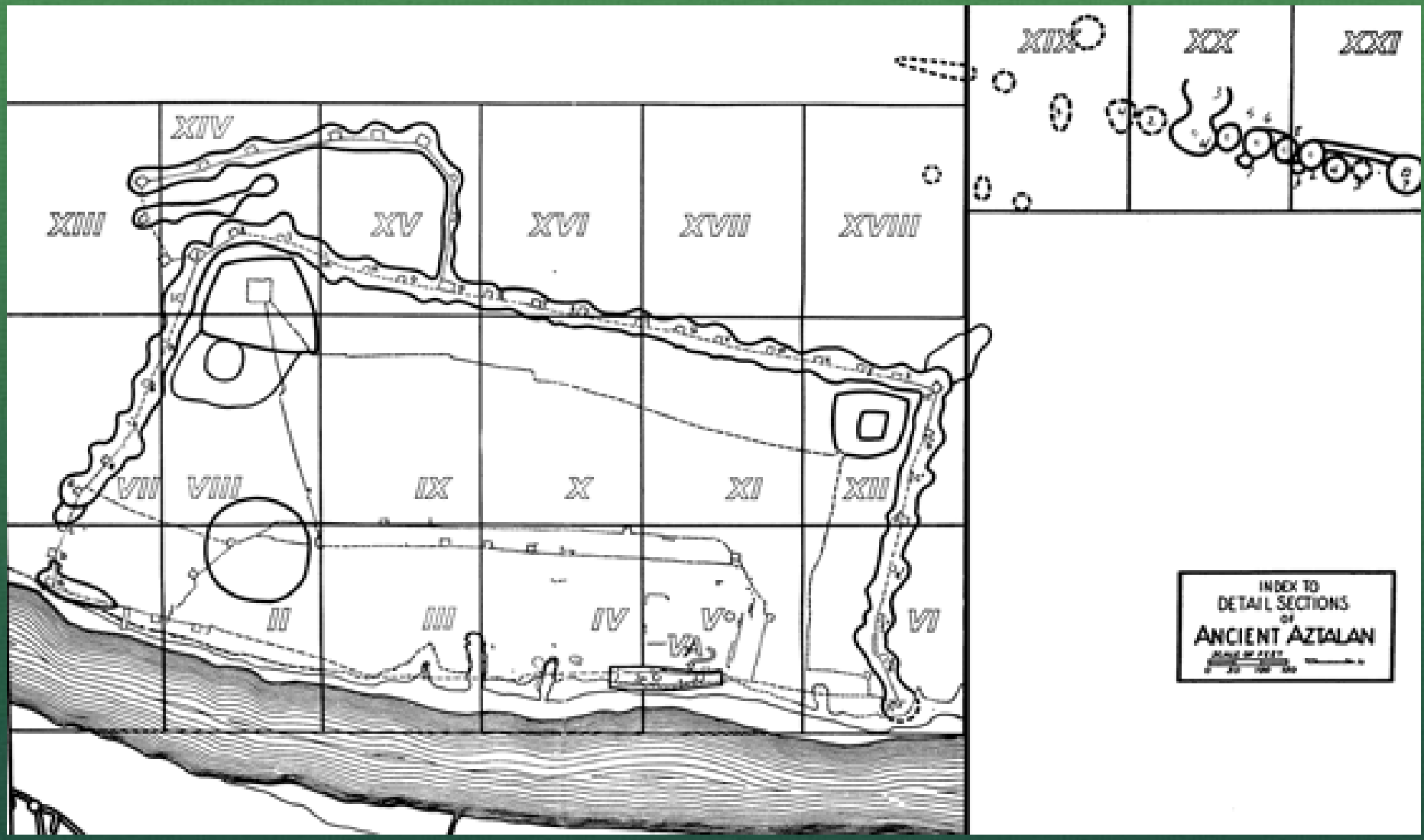
When at Aztalan a few days ago examining & surveying the ancient works, I was informed that you had made some investigations there in former times; and wishing to ascertain all the facts respecting these interesting remains I take the liberty of asking you to give me an account of the results of your examinations and reflections.

Did you find the remains of cloth? Are the works in your opinion works of defense – and for what reasons? Generally I would be happy to receive any full information you may communicate on the subject

I find Judge Hyer's original map requires some corrections and additions. I ~~shall present to the public~~ have made a thorough and correct one. —

Yours truly
I. A. Lapham

Geo. Hyer.


XIIV
XIII
XV
XVI
XVII
XVIII
XIX
XX
XXII
VII
VIII
IIX
X
XI
XIII
II
III
IV
V
VA
VI
INDEX TO
DETAIL SECTIONS
OF
ANCIENT AZTALAN

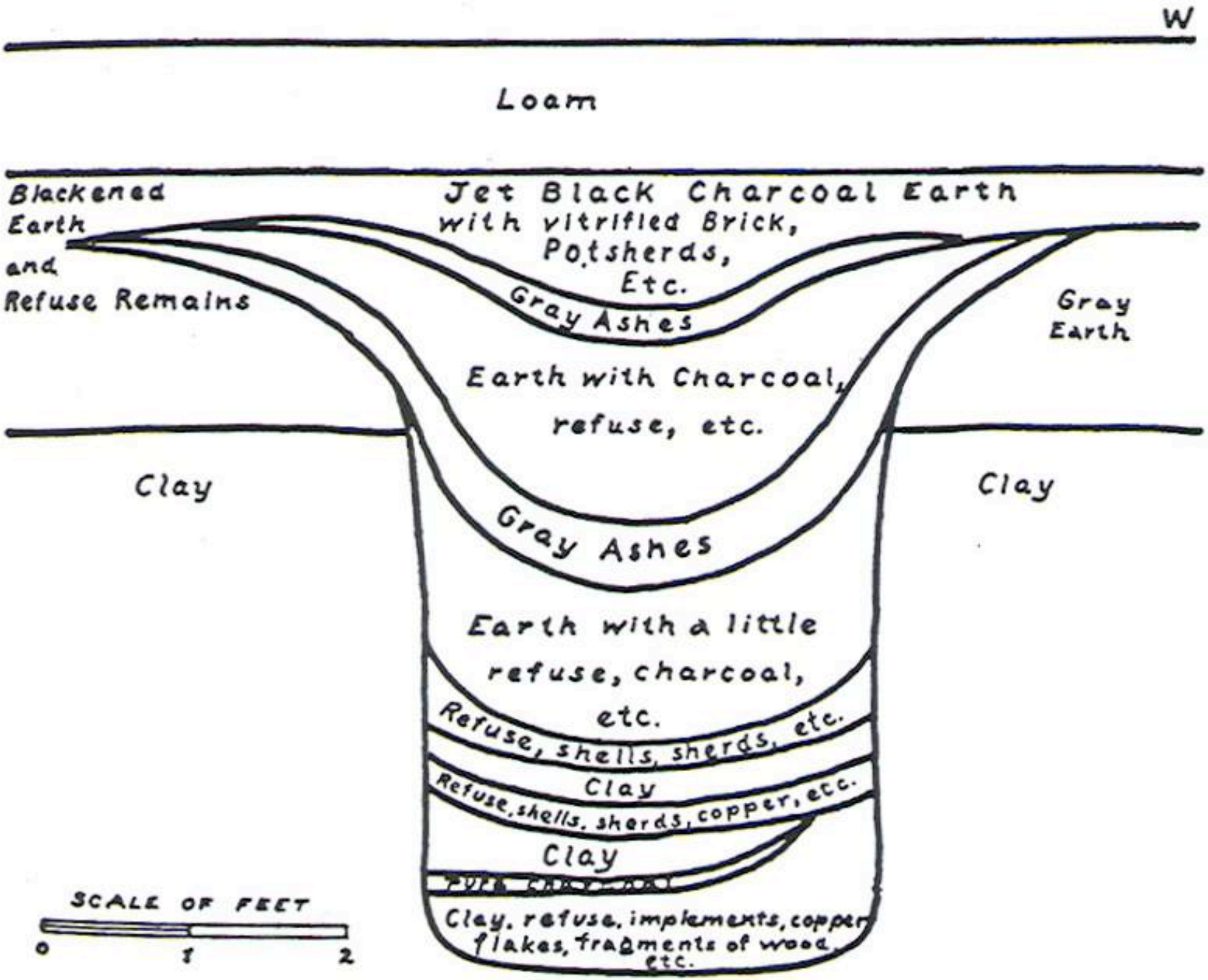


Fig. 34.—Section of refuse pit V-81.

Some examples of Aztalan ceramics

Chipped stone hoes

Projectile points or “arrowheads”

Ground stone axes

Discoidals or gaming stones

Ground stone pendant

Sandstone abraders

Animal bone

Animal bone made into artifacts

Shell beads; shell from Florida coast

Mussel shell hoes

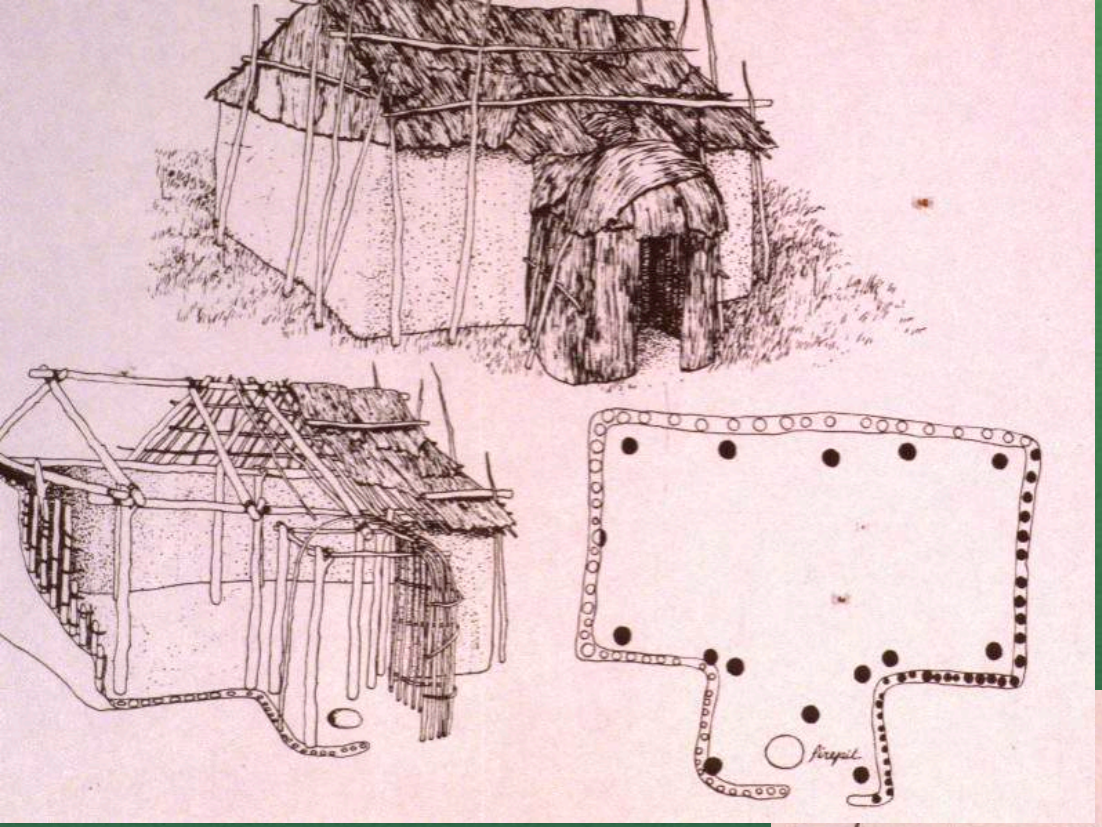
Copper tools; Lake Superior copper

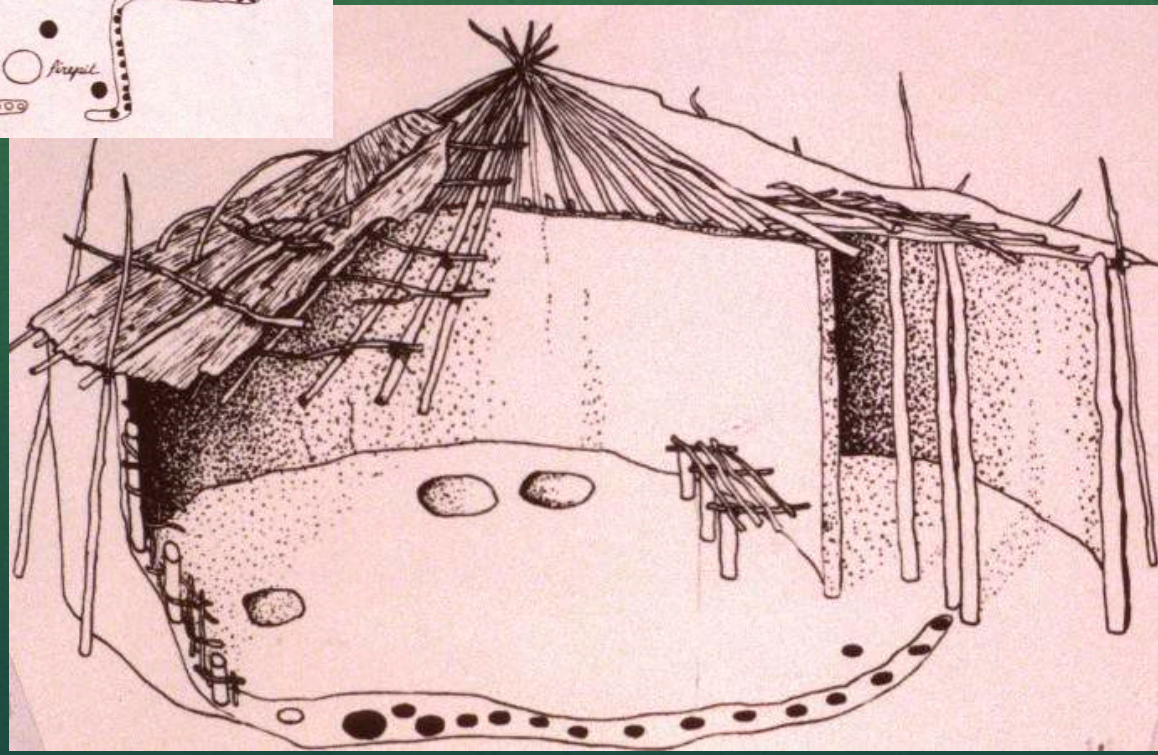
Copper maskettes

“Aztalan brick”

Charred corn kernels

Aztalan houses





Assiniboine Hole Shirt

" " " " end. 16 in long

" " ~~tibia~~ Fibula " " " 14 " "

" " " one end 12 " "

" " Femur a small (bird) shoulder bone length 5/2 inches

" burnt broken toe bone, black several pieces.

(RESTORED) Potsherds large sections of very large pot plain, simple rim, reddish; shell temper

" section of rim and side; plain " " "

" " " " " " "

" " " " " " " painted on rim " " glued many pieces

" " " " corrugated piece; incised on one edge, red, small

" " " " stone tempered. " " shell tempered.

" " " shell " " " lip. glued

" " " designed l 9 7/8 in " ends outside

" " " " three pieces. broken good specimen glued

" " " " brown

bone, good point; polished; broken at end. l. 2 1/4 in.

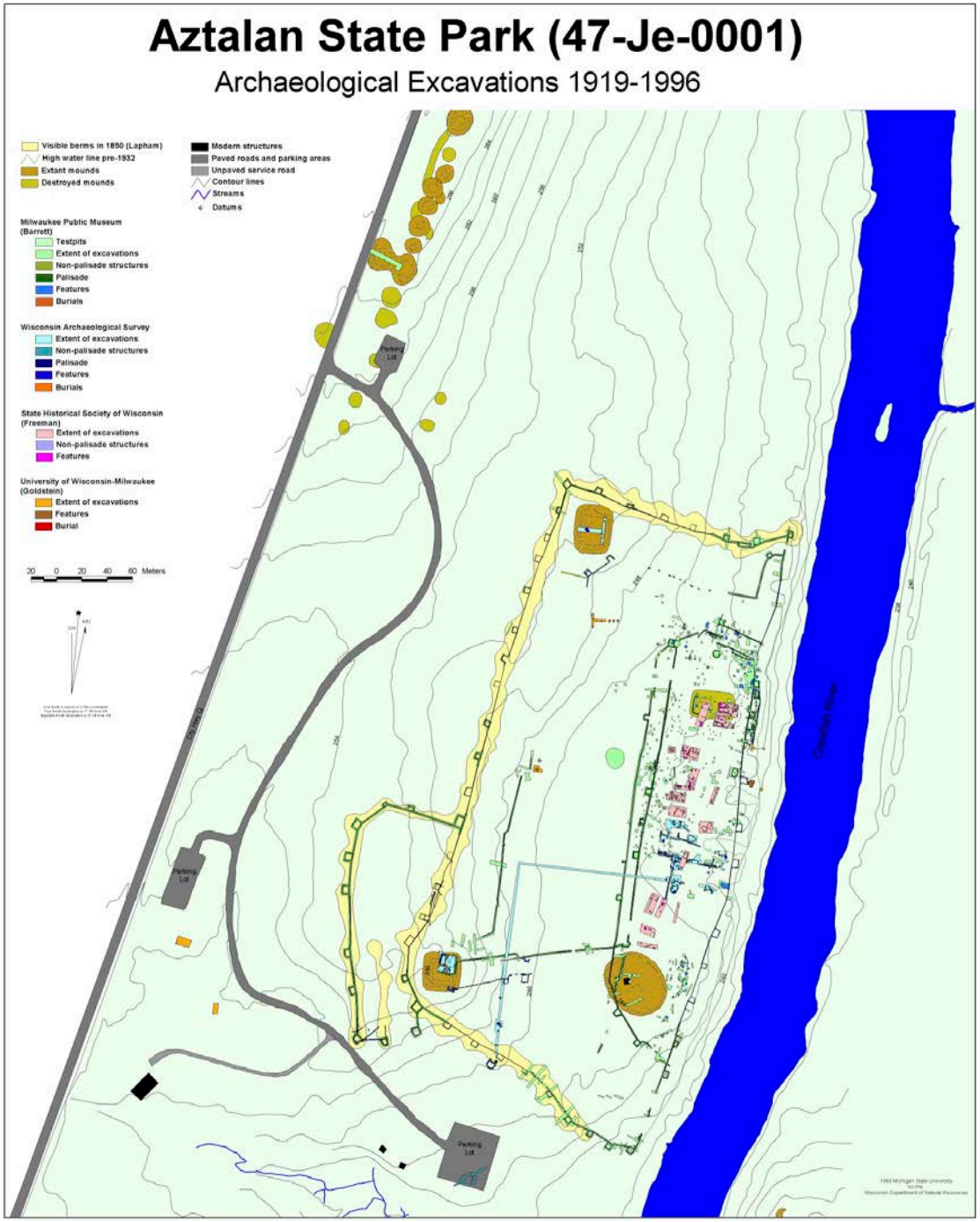
" " broken at base l. 3 inches.

piece; rolled shell coffee l. 1 1/4 in

[illegible]on of rim; simple, brown shell tempered.

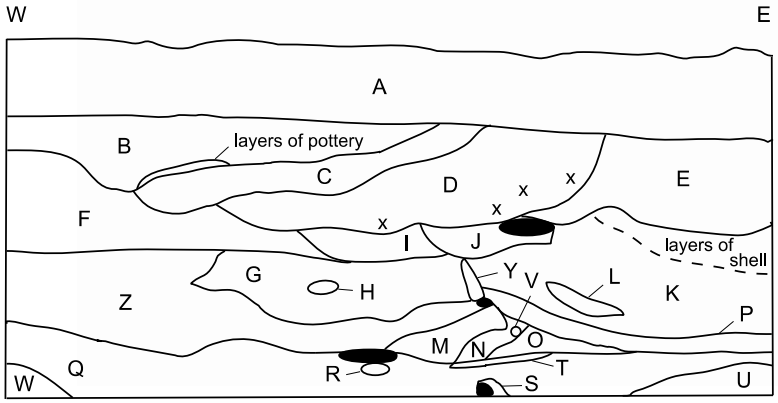
[illegible] of ladle; [illegible], shell tempered

[illegible] perforated; broken across hole. small



Aztalan GIS showing all excavations through 1996

Test Pit 3 North Profile

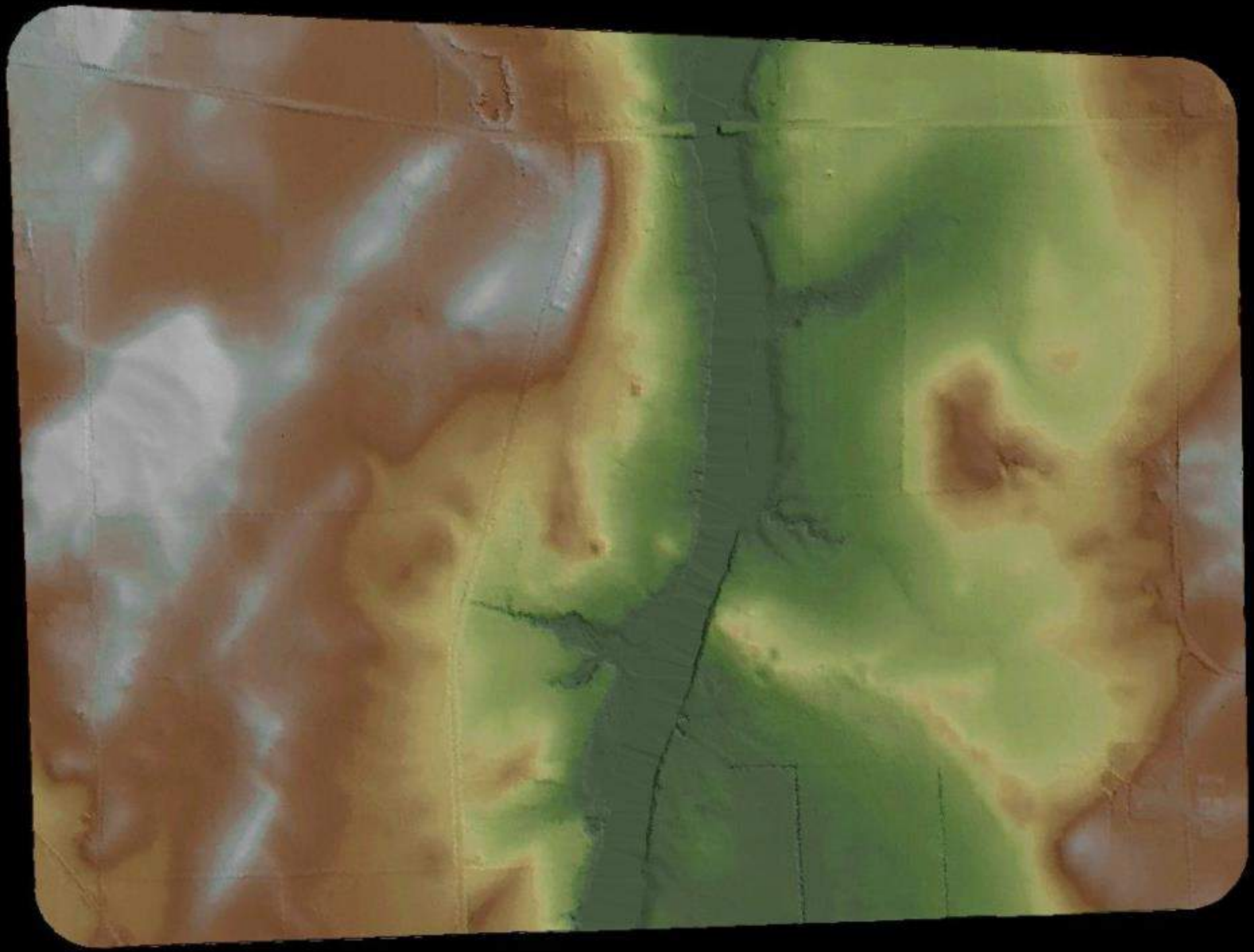


centimeters
0 50

= rock

A. 10YR3/2 Clay Loam
B. 10YR3/2 Clay Loam with flecks of charcoal
C. 10YR4/3 Sand with fine gravel
D. 10YR3/2 Clay Loam with 10YR3/4 sand mottles (x)
E. 10YR2/2 Loam
F. 10YR3/4 Silt Loam with large mottles of 10YR6/8 clay; mussel shell scattered throughout
G. 10YR3/1 Silty Clay Loam with thin layers of other soils
H. 10YR3/4 Sand
I. 10YR3/3 Clay Loam
J. 10YR5/3 Clay with shell
K. 10YR3/3 Silty Loam
L. 10YR5/4 Clay Loam
M. 10YR3/2 Clay
N. 10YR4/3 Sand
O. 10YR3/3 Silt Loam
P. 10YR2/1 Silty Clay Loam with thin layers of soils of different colors
Q. 10YR4/3 Silty Clay Loam
R. 10YR3/4 Sand
S. 10YR2/1 Silty Clay Loam
T. very mottled layer with sand that is 10YR5/4, Clay that is 2.5YR2.5/3, Clay that is 10YR6/3, and Clay Loam that is 10YR4/2
U. 10YR2/1 Silty Clay Loam
V. 2.5YR6/2 Silty Clay Loam
W. 10YR33 Sand with some charcoal
Y. 10YR4/4 Sand
Z. 10YR6/3 Clay with lots of mussel shell

Crawfish River



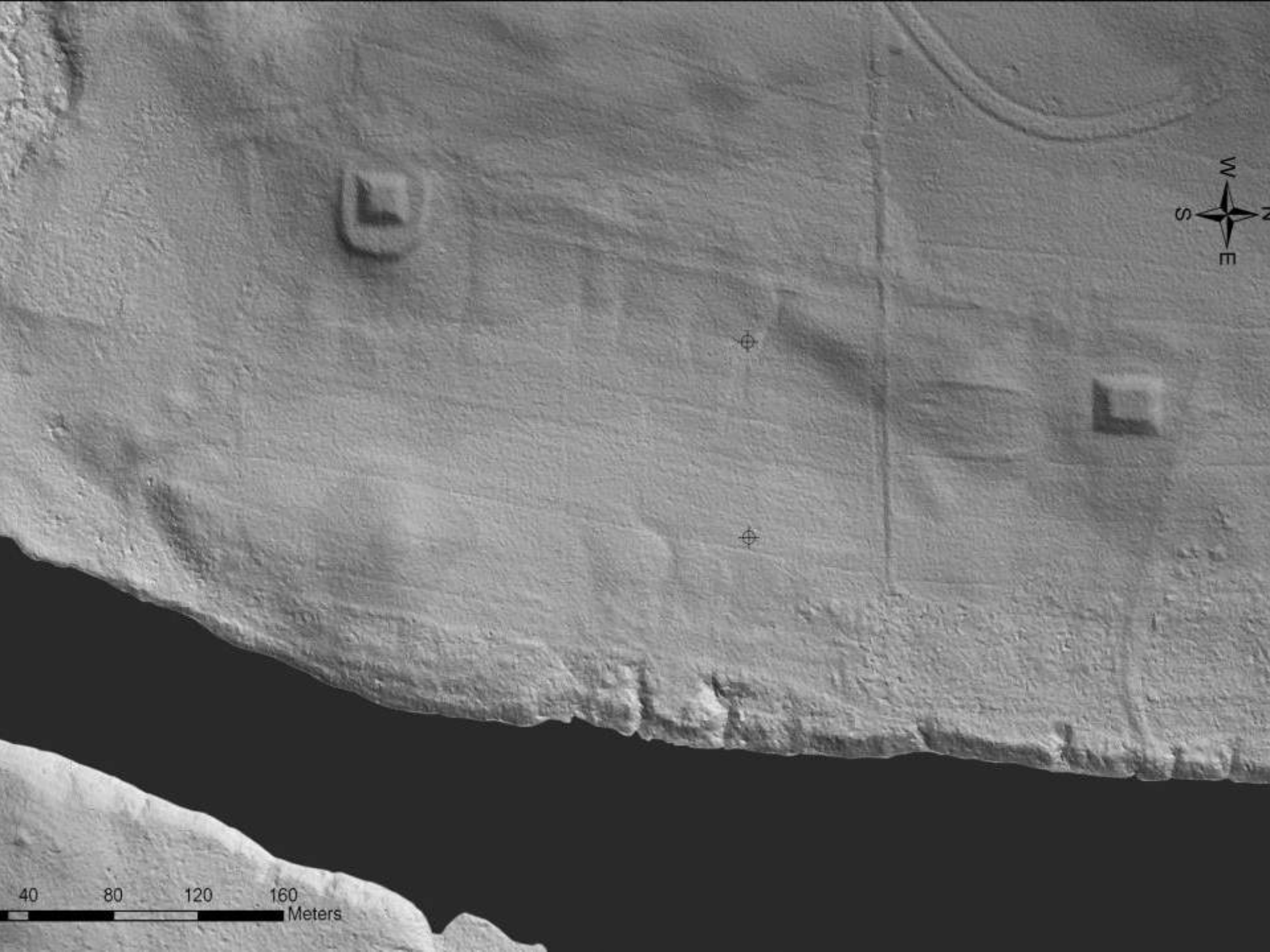

W
S
N
E
40
80
120
160
Meters

TABLE 4. Aztalan 2013 Gravel Knoll and West-of-Palisade Radiocarbon Samples

Sample Name	**Measured C14 Years (Years RCYBP)**	Mean Calibrated C14 Years (BP)	Error +/- (Years)	Upper Limit (Cal BP)	Lower Limit (Cal BP)	**Mean Calibrated Year AD**	Lower Age Limit (AD) - 2 sigma	Upper Age Limit (AD) – 2 sigma	Material dated Context
GK1 Beta-374816	860	742	26	795	690	1210	1155	1260	Charcoal – AMS Unit 4, Fea 4a
GK2 Beta-374817	190	150	75	300	0	1800	1650	1950	Carbonized wood – AMS – Unit 3, PM2
GK3 Beta-374818	120	140	70	280	0	1810	1670	1950	Carbonized wood – AMS – Unit 3, PM2
GK4 Beta-374819	960	855	35	925	785	1095	1025	1165	Wood: Radiometric Unit 2, Fea 2
WP1 Beta-374821	960	860	35	930	790	1090	1020	1160	Charcoal – AMS Unit 25, Fea 101
WP2 Beta-374822	910	820	45	910	730	1130	1040	1220	Charcoal – AMS Unit 25, Fea 101
WP3 Beta-375567	960	860	35	930	790	1090	1020	1160	Charcoal – AMS Unit 24, Fea 103

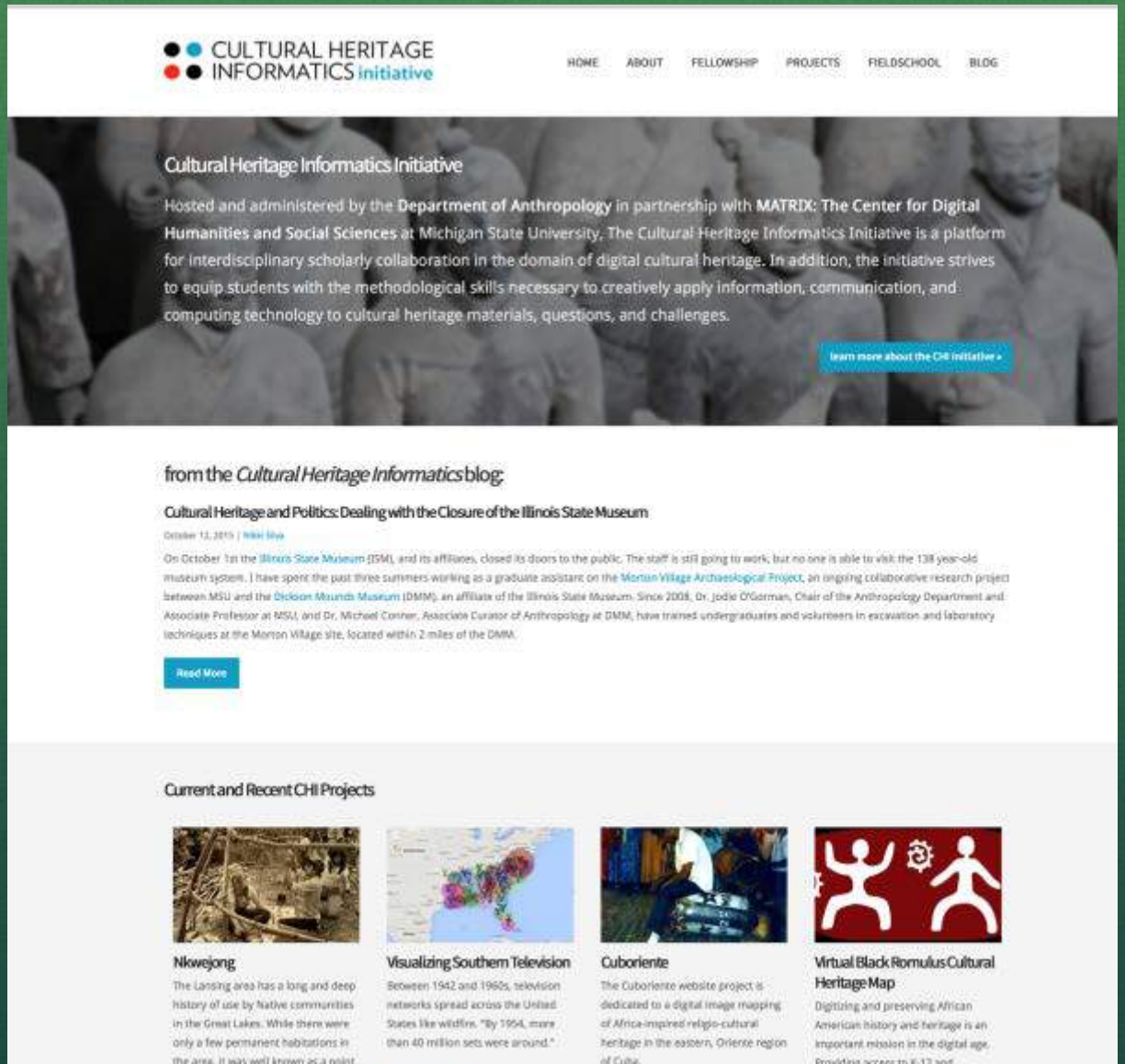
*Dates reported as radiocarbon years before present consider the year 1950 to be the 'present' date.

*Upper and lower age limits are calculated as the two-sigma standard deviation of the calibrated radiocarbon age.

The Future

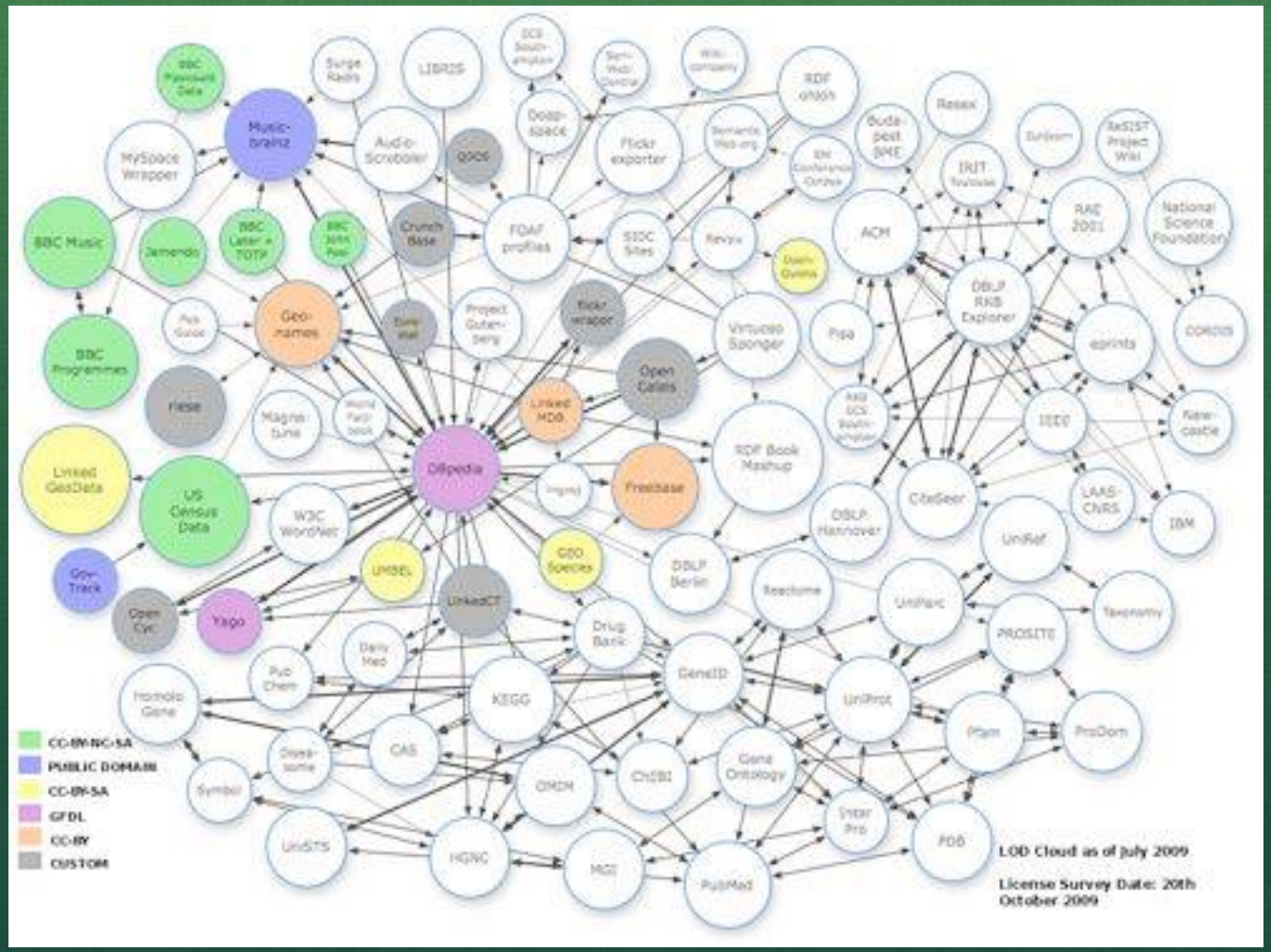
Teach to the Next Generation

CHI: Cultural Heritage Informatics

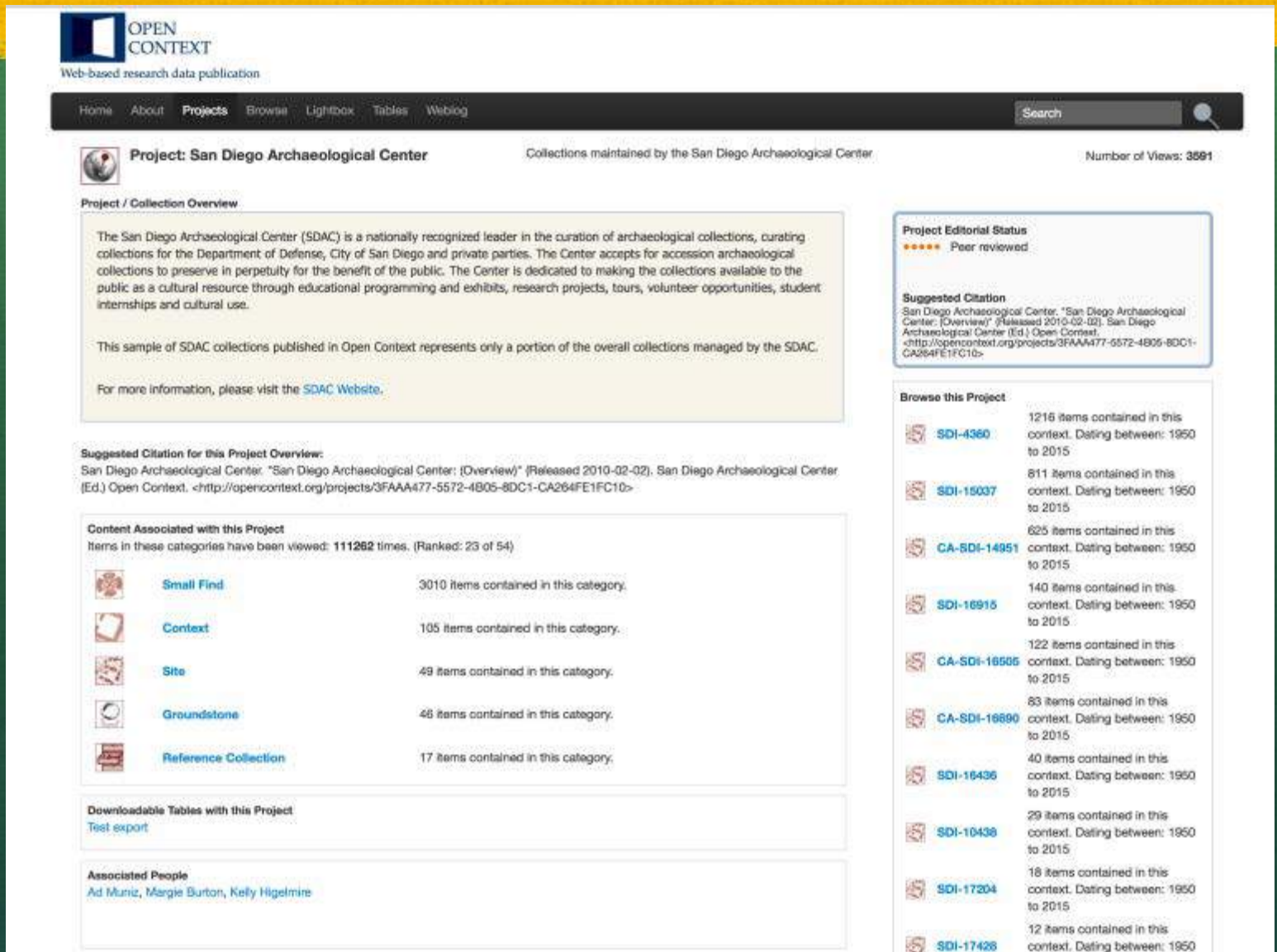


MSU Campus Archaeology Program

Think About Data Differently



Credit for Publishing Data



OPEN CONTEXT
Web-based research data publication

Home | About | Projects | Browse | Lightbox | Tables | Weblog | Search

Project: San Diego Archaeological Center

Collections maintained by the San Diego Archaeological Center

Number of Views: **3591**

Project / Collection Overview

The San Diego Archaeological Center (SDAC) is a nationally recognized leader in the curation of archaeological collections, curating collections for the Department of Defense, City of San Diego and private parties. The Center accepts for accession archaeological collections to preserve in perpetuity for the benefit of the public. The Center is dedicated to making the collections available to the public as a cultural resource through educational programming and exhibits, research projects, tours, volunteer opportunities, student internships and cultural use.

This sample of SDAC collections published in Open Context represents only a portion of the overall collections managed by the SDAC.

For more information, please visit the SDAC Website.

Project Editorial Status
★★★★★ Peer reviewed

Suggested Citation
San Diego Archaeological Center. "San Diego Archaeological Center: (Overview)" (Released 2010-02-02). San Diego Archaeological Center (Ed.) Open Context. <http://opencontext.org/projects/3FAAA477-5572-4B05-8DC1-CA264FE1FC10>

Suggested Citation for this Project Overview:
San Diego Archaeological Center. "San Diego Archaeological Center: (Overview)" (Released 2010-02-02). San Diego Archaeological Center (Ed.) Open Context. <http://opencontext.org/projects/3FAAA477-5572-4B05-8DC1-CA264FE1FC10>

Content Associated with this Project
Items in these categories have been viewed: **111262** times. (Ranked: 23 of 54)

Small Find	3010 items contained in this category.
Context	105 items contained in this category.
Site	49 items contained in this category.
Groundstone	46 items contained in this category.
Reference Collection	17 items contained in this category.

Browse this Project

SDI-4360	1216 items contained in this context. Dating between: 1950 to 2015
SDI-15037	811 items contained in this context. Dating between: 1950 to 2015
CA-SDI-14951	625 items contained in this context. Dating between: 1950 to 2015
SDI-16915	140 items contained in this context. Dating between: 1950 to 2015
CA-SDI-16505	122 items contained in this context. Dating between: 1950 to 2015
CA-SDI-16890	83 items contained in this context. Dating between: 1950 to 2015
SDI-16436	40 items contained in this context. Dating between: 1950 to 2015
SDI-10438	29 items contained in this context. Dating between: 1950 to 2015
SDI-17204	18 items contained in this context. Dating between: 1950 to 2015
SDI-17428	12 items contained in this context. Dating between: 1950

Downloadable Tables with this Project
Test export

Associated People
Ad Muniz, Margie Burton, Kelly Higelmire