Principles of IR

Hacettepe University
Department of Information Management
DOK 324: Principles of IR

Geographic IR

Some Slides taken from: Ray Larson

Overview

- What is Geographic Information Retrieval?
- Geographic and Spatial Querying and Browsing.
- Geographic and Spatial Indexing.
- Examples of GIR Systems and Geographically Indexed Information.

Introduction

- What is Geographic Information Retrieval?
  - GIR is concerned with providing access to georeferenced information sources. It includes all of the areas of traditional IR research with the addition of spatially and geographically oriented indexing and retrieval.
  - It combines aspects of DBMS research, User Interface Research, GIS research, and Information Retrieval research.

Introduction

- The need for Geographic and Spatial Information Retrieval.
  - Digital Libraries
    - Sequoia 2000
    - UC Berkeley NSF/NASA/ARPA Digital Library Project
    - UC Santa Barbara Alexandria Project
    - NSDI - National Spatial Data Infrastructure

Geographic and Spatial Querying

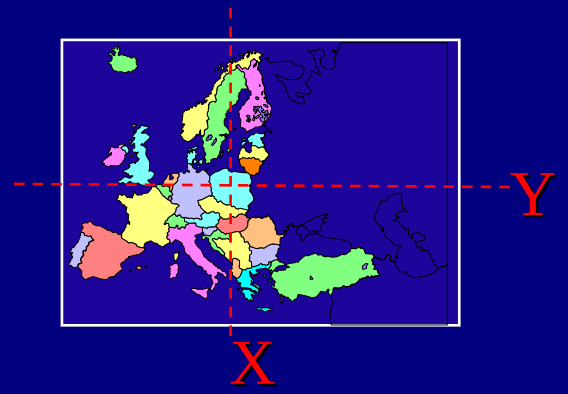
- Both imply querying on relationships within a particular coordinate system
- Spatial querying is the more general term
- Can be defined as queries about the spatial relationships (intersection, containment, boundary, adjacency, proximity) of entities geometrically defined and located in space

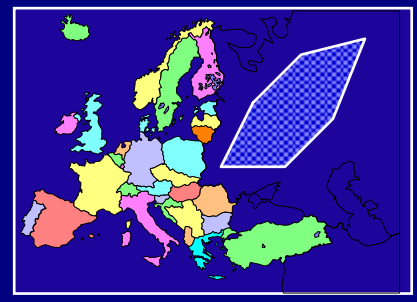
Geographic and Spatial Querying

- Geographical coordinates are geometric relationships (distance and direction can be measured on a continuous scale)
  - E.g. "5.21 miles north of Champaign"
- Spatial relations may be both geometric and topological (spatially related but without measureable distance or absolute direction)
  - E.g.: "inside the city limits"
  - "left side of Beckman Institute"

Geographic and Spatial Querying

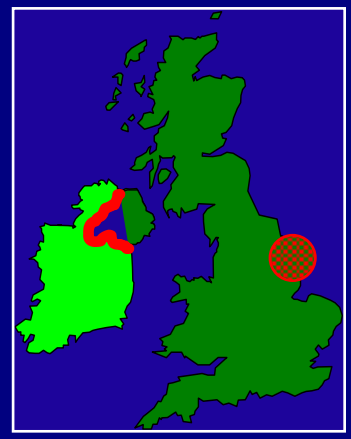
- Types of spatial queries
  - Point – in - polygon : "What do we have at this X,Y point?"
  - Region Queries : "What do we have in this region?"
    - Which point encoded items lie within the region
    - What lines (borders, etc.) lie within or the cross the region
    - What areas overlap the region area

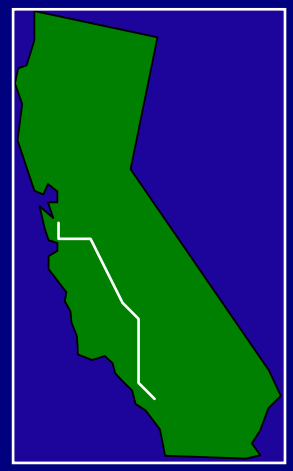




Geographic and Spatial Querying

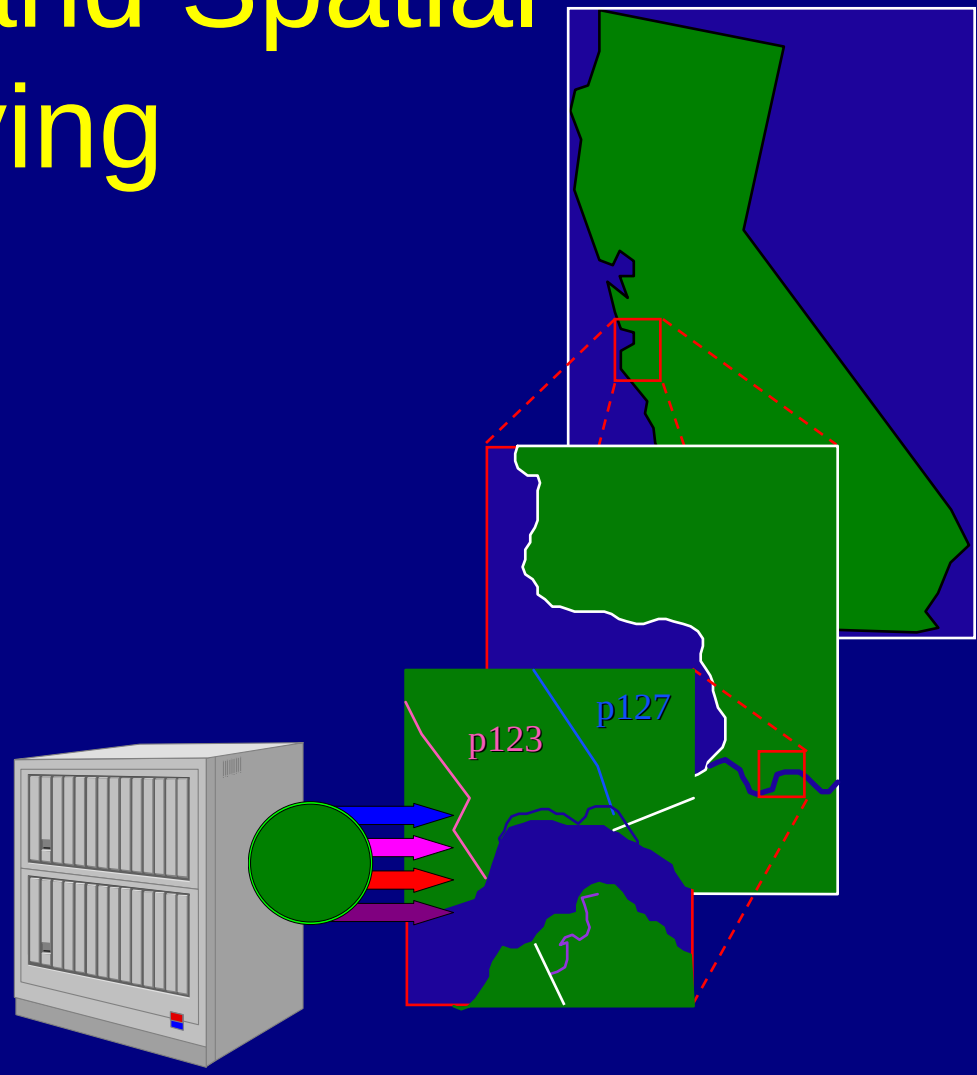
- Types of spatial queries, cont.
  - Distance and Buffer Zone Queries
    - What cities lie within 40 miles of the border of Northern and Southern Ireland?
    - What wetlands lie within 50 miles of London?
  - Path Queries
    - What is the shortest route from San Francisco to Los Angeles?





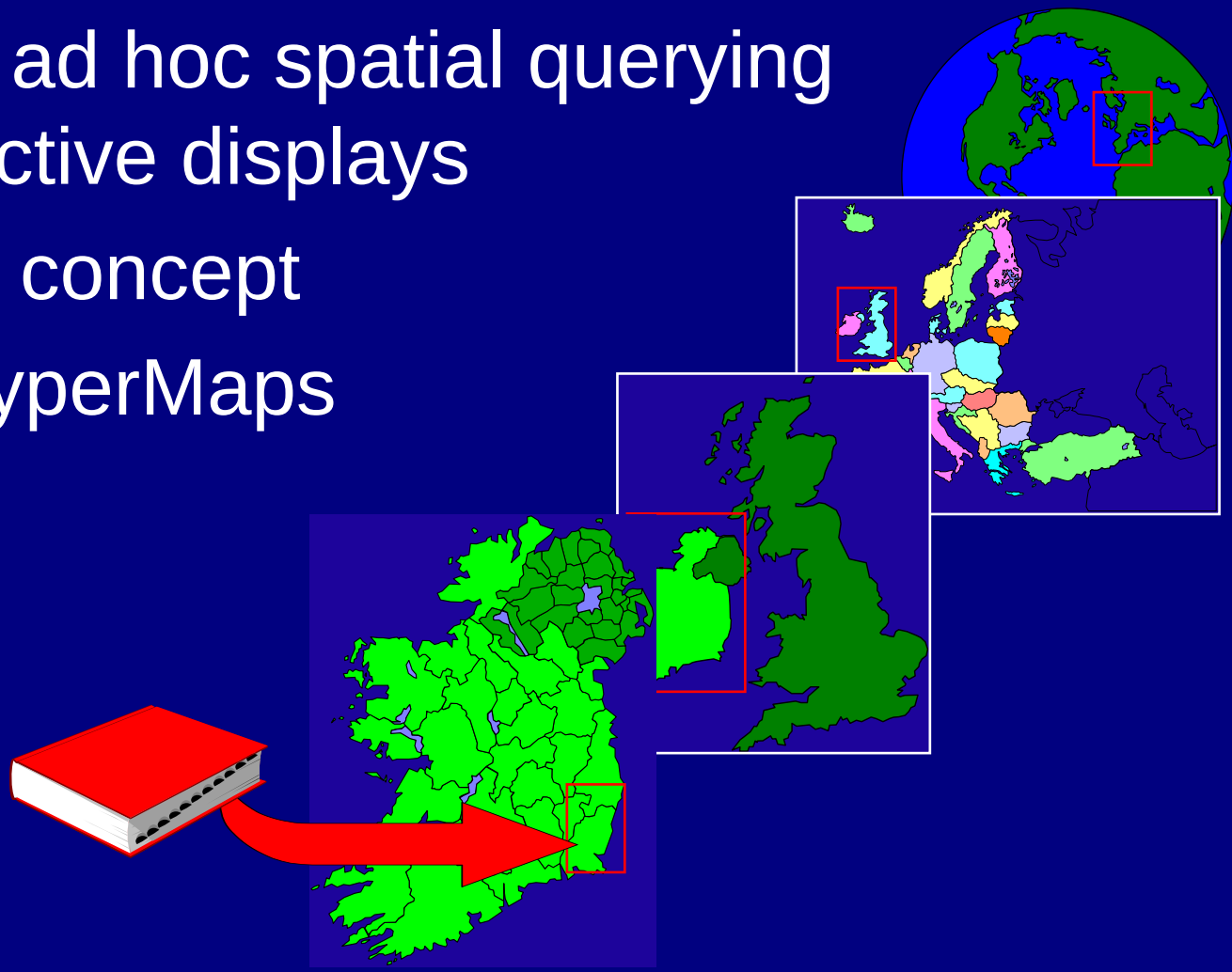
Geographic and Spatial Querying

- Types of spatial queries, cont.
  - Multimedia Queries : Use non - map georeferenced information.
    - What are the names of farmers affected by flooding in Monterey and Santa Cruz Counties?



Spatial Browsing

- Combines ad hoc spatial querying with interactive displays
- HyperMap concept
- Pseudo-HyperMaps



Spatial Browsing

- Advantages:
  - May not need the accuracy of a full GIS
  - Comprehensible searching metaphor for many materials
- Problems:
  - Clutter and differing scales.
  - Requires good (and preferably accurate) geographical indexing
  - Assumes that the user knows some geography

Geographic and Spatial Indexing

- Traditional geographic indexing involves using place names from LCSH and name authorities. These have some problems:
  - Names are not unique
  - The places referred to change size, shape and names over time
  - Spelling variations
  - Some places are temporary conventions (study areas, etc.)

Digital Gazetteers

- Geographic names are and will remain the primary *Entry Vocabulary* for DL spatial queries
  - The gazetteer must support as many variant forms of the name as possible
    - Including temporal ranges for particular names
  - querying must support spatial reasoning based on gazetteer and other geographic and temporal information in the system or accessible by network access

ADL Gazetter Content Standard

Geographic Feature ID: 1200
Geographic Name: Barbuda
Name Authority: U.S. Board on Geographic Names (BGN)
Variant Name: Barbuda Island
Type of Geographic Feature
Scheme: ADL Feature Type Thesaurus
Type: islands
Scheme: NIMA Feature Designation
Type: ISL (island)
Geographic Feature Code
Scheme: NIMA
Code: AC01
Related Feature
Relationship type: isPartof
Name: Antigua and Barbuda; Leeward Islands, West Indies
Spatial Location
Longitude:- 61.8
Latitude: 17.6

Geographic and Spatial Indexing

- Geographic coordinates have some advantages over names:
  - They are persistent regardless of name, political boundary or other changes
  - They can be simply connected to spatial browsing interfaces and GIS data.
  - They provide a consistent framework for GIR applications and spatial queries.
- However, the geographic extents and boundaries of entities also change over time
  - This may be the primary interest of historical scholarship

Geographic and Spatial Indexing

- GIPSY: Automatic georeferencing of texts (Geographic Info Processing System)
  - The work of Allison Woodruff and Christian Plaunt - Later DBMS-based version by Jolly Chen -- New version planned
  - Designed to operate on the full text of documents
  - Extracts geographic terms and attempts to identify the coordinates of the places discussed in the text using a combination of evidence

Geographic and Spatial Indexing

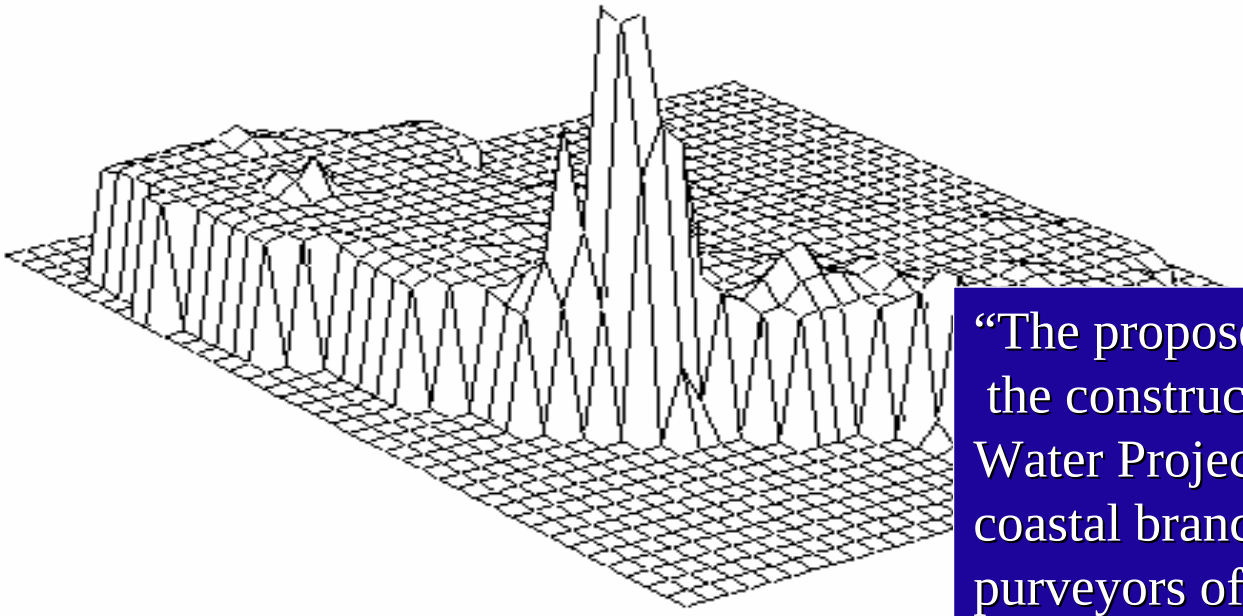
- GIPSY cont.
  - Used the USGS Geographic Names Information System (GNIS) and Geographic Information Retrieval and Analysis System (GIRAS) to associate names with coordinates of named places, geographic features and land use characteristics.

Geographic and Spatial Indexing

- GIPSY cont.
  - Identified places are added as “elevations” with each place adding a weight based on its frequency in the text and database characteristics
  - The resulting map is analysed to identify the most likely locations, and coordinates for those locations are extracted

Geographic and Spatial Indexing

- GIPSY Map Overlay



“The proposed project is the construction of a new State Water Project facility, the coastal branch... by water purveyors of northern Santa Barbara County... delivering water to San Luis Obispo ... “

Geographic and Spatial Indexing

- To be useful for the range of cultural and humanities materials being collected in digital libraries, the GIPSY gazetteer must
  - Support many different time ranges, location and boundary changes
  - Support synonymous and variant names with differing locations for the same entity
  - Support names in multiple languages, scripts and usages

ECAI

- The Electronic Cultural Atlas Initiative is a collaboration between IT professionals and humanities scholars
- ECAI is developing a globally distributed spatio-temporal library of cultural and historical resources with a centralized metadata catalogue and a GIS viewer
- Currently the ECAI consortium includes over 250 projects

ECAI

- Projects range from small works by individual scholars to large nationally and internationally funded efforts. E.g.:
  - geography of Greco-Roman culture (Perseus project)
  - toponym locations for over 300,000 images of Buddhist art and architecture
  - Seals of the Sassanian Empire
  - historical trade routes of Eurasia
  - the map of Hideyoshi's invasion of Korea
  - historical GIS projects for China, Great Britain, the United States, the Black Sea and Tibet

ECAI

"DOK 324 PIR"

Electronic Cultural Atlas Initiative (ECAI) - Microsoft Internet Explorer

Dosya Düzen Görünüm Sık Kullanılanlar Araçlar Yardım

Geri Ara Sık Kullanılanlar

Adres http://www.ecai.org/ Git

Google Search Web 74 blocked AutoFill Options Bağlantılar

ECAI

Electronic Cultural Atlas Initiative

ECAI Clearinghouse
Directory
Members Area
Search | Home

Activities Community Projects Technology Participating

Dynamic Map Catalog

Create digital maps that display a wide range of cultural material by using place and time as a common element.

Community

Participate in the ECAI community through twice yearly conferences and international working groups.

Technical Infrastructure

ECAI technical infrastructure illustrates the vision of sharing distributed data and using time enabled mapping tools.

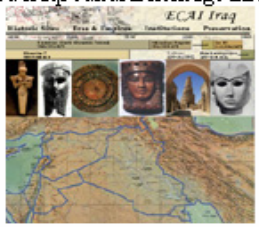
ECAI uses time and space to enhance understanding and preservation of human culture.

Highlights

ECAI Meeting, May 9 - 13, 2005
Fudan University, Shanghai
Registration ends April 18th

ECAI Iraq Cultural Atlas

ECAI Iraq provides integrated access to widely distributed Iraqi cultural heritage information.



See: http://ecai.org/iraq

New **David Rumsey Map Collection - Over 5000 georegistered historical**

18 Mayıs 2005 Çarşamba

Başlat H E. TR 08:51

Y.T.

ECAI Clearinghouse Search - Microsoft Internet Explorer

Dosya Düzen Görünüm Sık Kullanılanlar Araçlar Yardım

Geri Ara Sık Kullanılanlar

Adres http://ecaimaps.berkeley.edu/clearinghouse/ Git

Google Search Web 74 blocked AutoFill Options Bağlantıla

Developed by University of Sydney
TimeMap™ Project

TimeMap™

Electronic Cultural Atlas Initiative

Clearinghouse Search

Find area of interest with zoom and drag tools

Click MapSpace bounding boxes for interactive map

Search Extent

Longitude	26	to
	40	
Latitude	36	to
	50	
Time	-10000	to
	2005	

Reset

Search Settings

Search collections:
- ECAI clearinghouse
- David Rumsey map collection

3. Text search: turkey — Blank = broadest search

Sort by: Spatial relevance

Başlat H E K. E. TR

Y.T.

Electronic Cultural Atlas Initiative

Developed by University of Sydney
*Time*Map™ Project

TimeMap™

ECAI Clearinghouse Search Results

Refine search

Longitude 26 to 40 Latitude 50 to 36 Time -10000 to 2005

Find turkey Search Search within results

Results 1 to 2 shown (2 results in total)

= MapSpace | = password required | = time-enabled | = link-enabled | = ECAI publication
= Framework - global | = Framework - regional | = Framework - local
= Key cultural dataset | = ECAI rated dataset | = Non-geographic dataset | = WorkInProgress
[tms] = Show MapSpace file | [wms] = View with Web Map Service | [specs] = View WMS capabilities | [swf] = View as Macromedia Flash file

Make Map

www 162: **A Retrospective Look at Earthquakes in the Ýzmit Gulf Area** (-1440 CE)
Contributors: Department of Geography, Fatih University
Responsibility: Novakowski, Nicholas <nicknova2000@hotmail.com>
Created by Novakowski, Nicholas
Turkey, Izmit, Not yet rated

1511: **Rivers - Iraq Region** (2003 CE)
Contributors: ECAI
Responsibility: Jeanette Zerneke <jlz@pdlink4.berkeley.edu>

Results 1 to 2 shown (2 results in total)

Map this dataset

Connection information

Dataset title A Retrospective Look at Earthquakes in the Ýzmit Gulf Area
Dataset type Website
Server type WWW web page/site (not accessible as dataset)
Creation date 2000-11-16 23:08:29
Last updated 2002-02-03 13:14:15

Descriptive metadata

dc.title	**A Retrospective Look at Earthquakes in the Izmit Gulf Area**
dc.subject.specific	earthquakes
dc.subject.specific	land use
dc.subject.specific	Izmit
dc.subject.specific	August 17, 1999
dc.subject.specific	Turkey
dc.description	A retrospective look at seismic activity in the Izmit Gulf area, including a focus on the catastrophic August 17, 1999 earthquake. In particular, normative land use strategies for ameliorating the impacts of seismic events are explored.
dc.publisher	Department of Geography, Fatih University
dc.publisher.address	Karaagac Yolu, Buyukcekmece-Istanbul, Turkey, 34900
dc.publisher.address	http://www.geography.fatih.edu.tr
dc.creator.Person.Name	Novakowski, Nicholas
dc.creator.Person.Affiliation	nicknova@fatih.edu.tr
dc.creator.Person.Affiliation	Department of Geography, Fatih University, Buyukcekmece-Istanbul, Turkey, 34900

Electronic Cultural Atlas Initiative

ECAI Technology

IT Infrastructure Components

- The ECAI Metadata Clearinghouse
- *Time*Map™ Software
- ECAI Servers
- ECAI User Support System and Knowledgebase

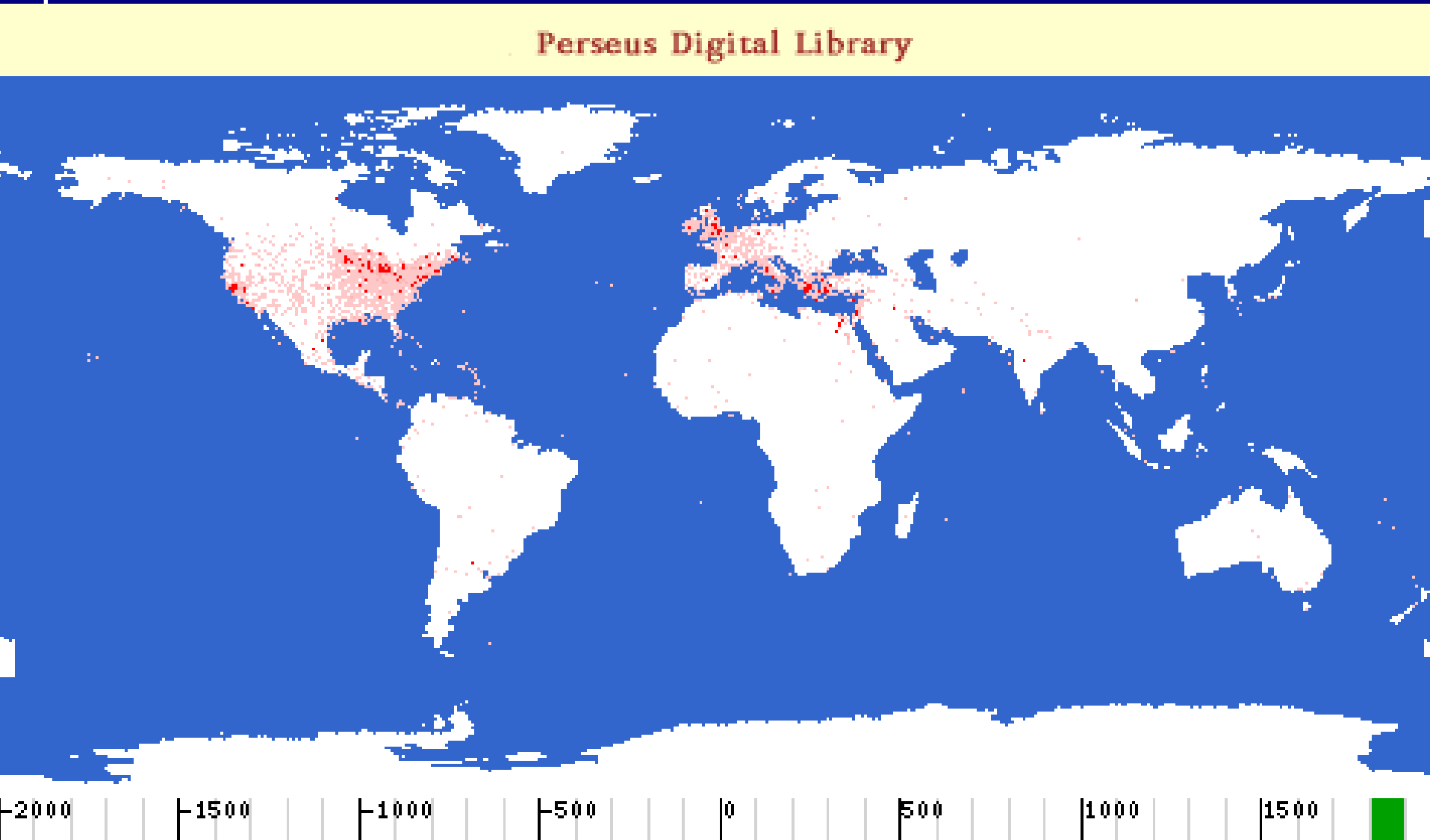
IT Architecture Goals

- Use distributed institutional resources
- Follow best practices for distributed internet computing
- Encourage interoperability
- Create custom solutions to link and enhance technologies when needed
- Create services and tools to support users
- Use a team structure to organize contributions of ECAI affiliates

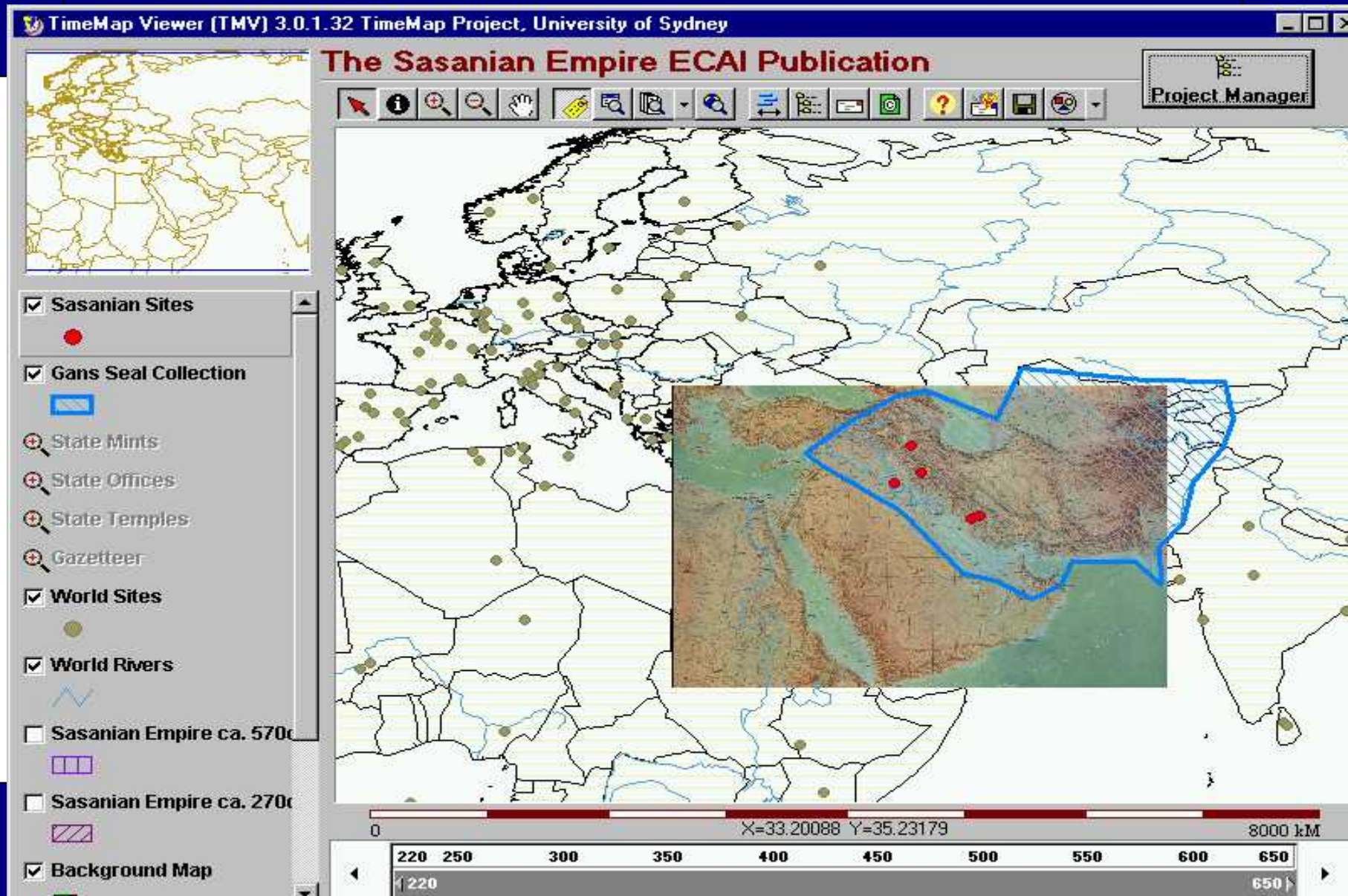
Technical Methodology

- The technical architecture and technology design is developed collaboratively by participants

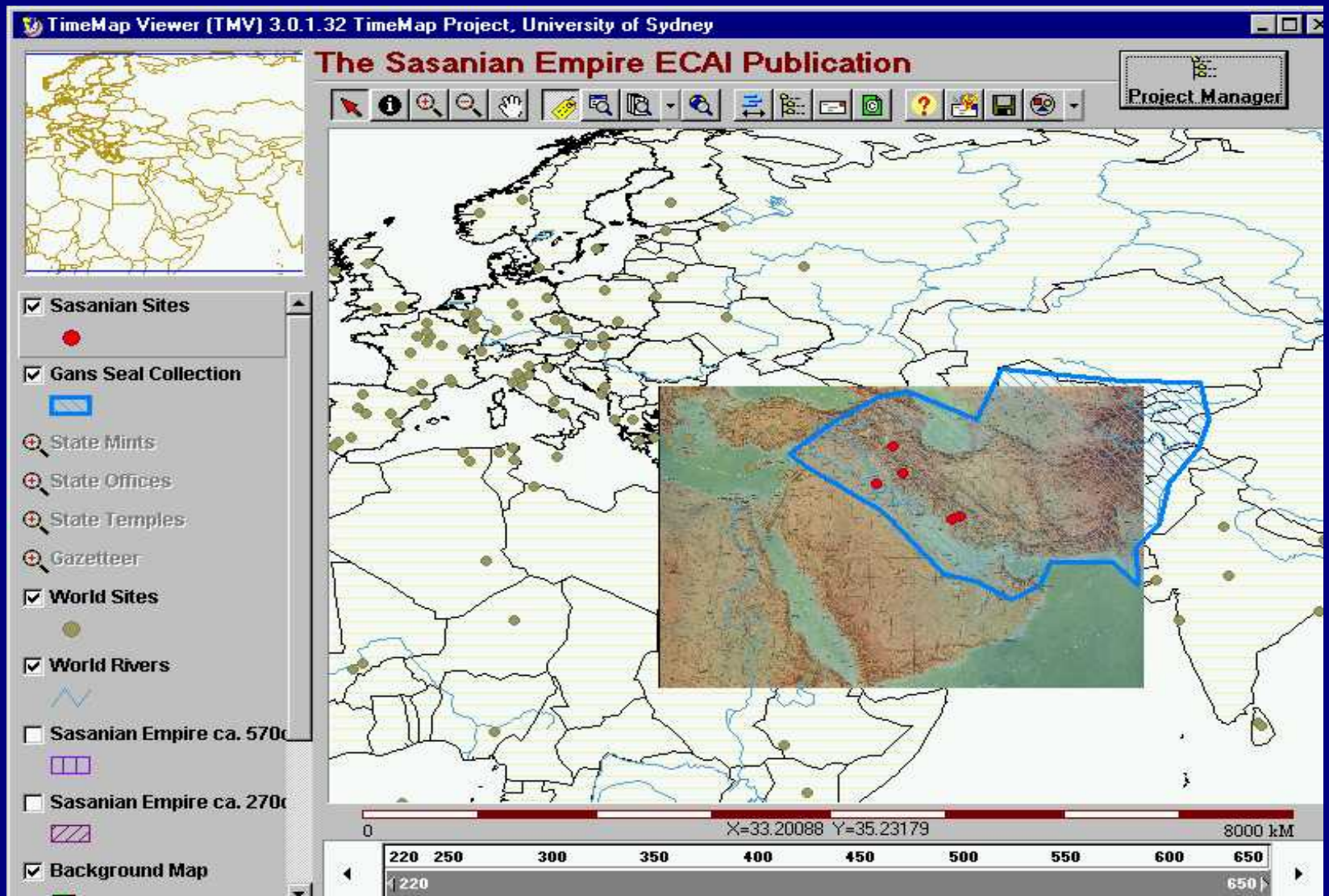
Perseus



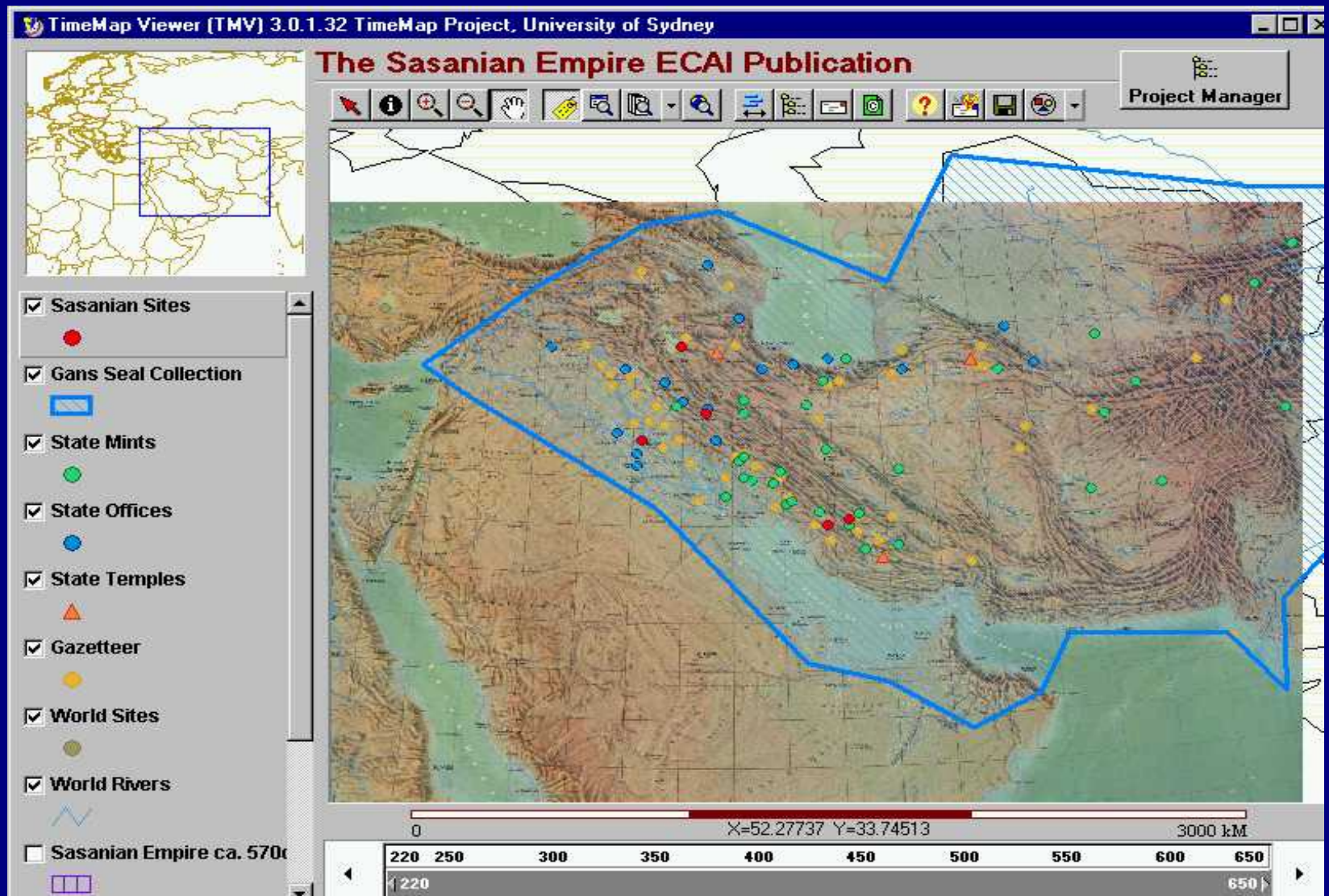
The Sasanian Empire



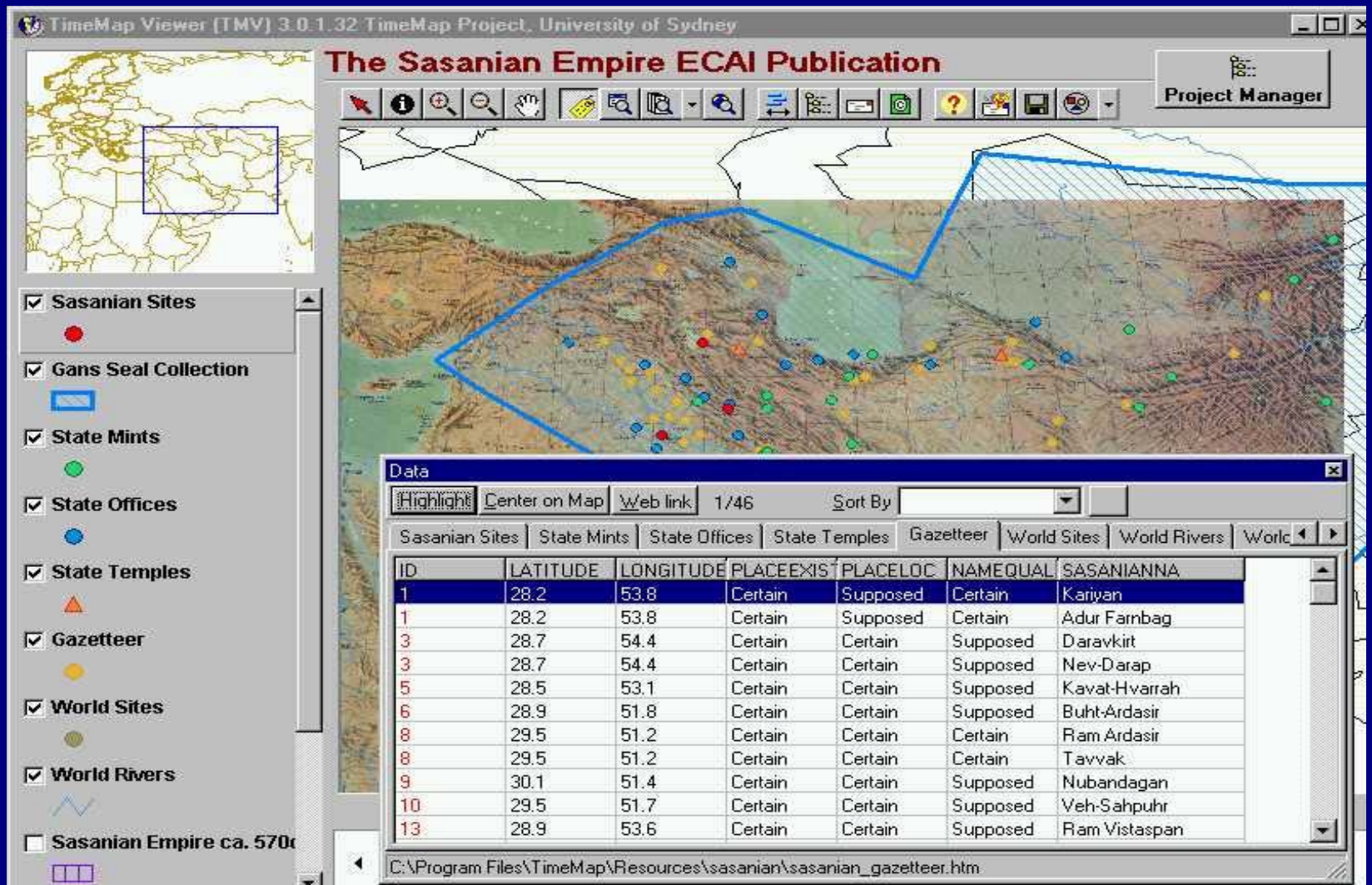
Opening shot of the Sasanian Empire ECAI project, showing a map with diverse resources, a timeline, and a menu of available map layers.



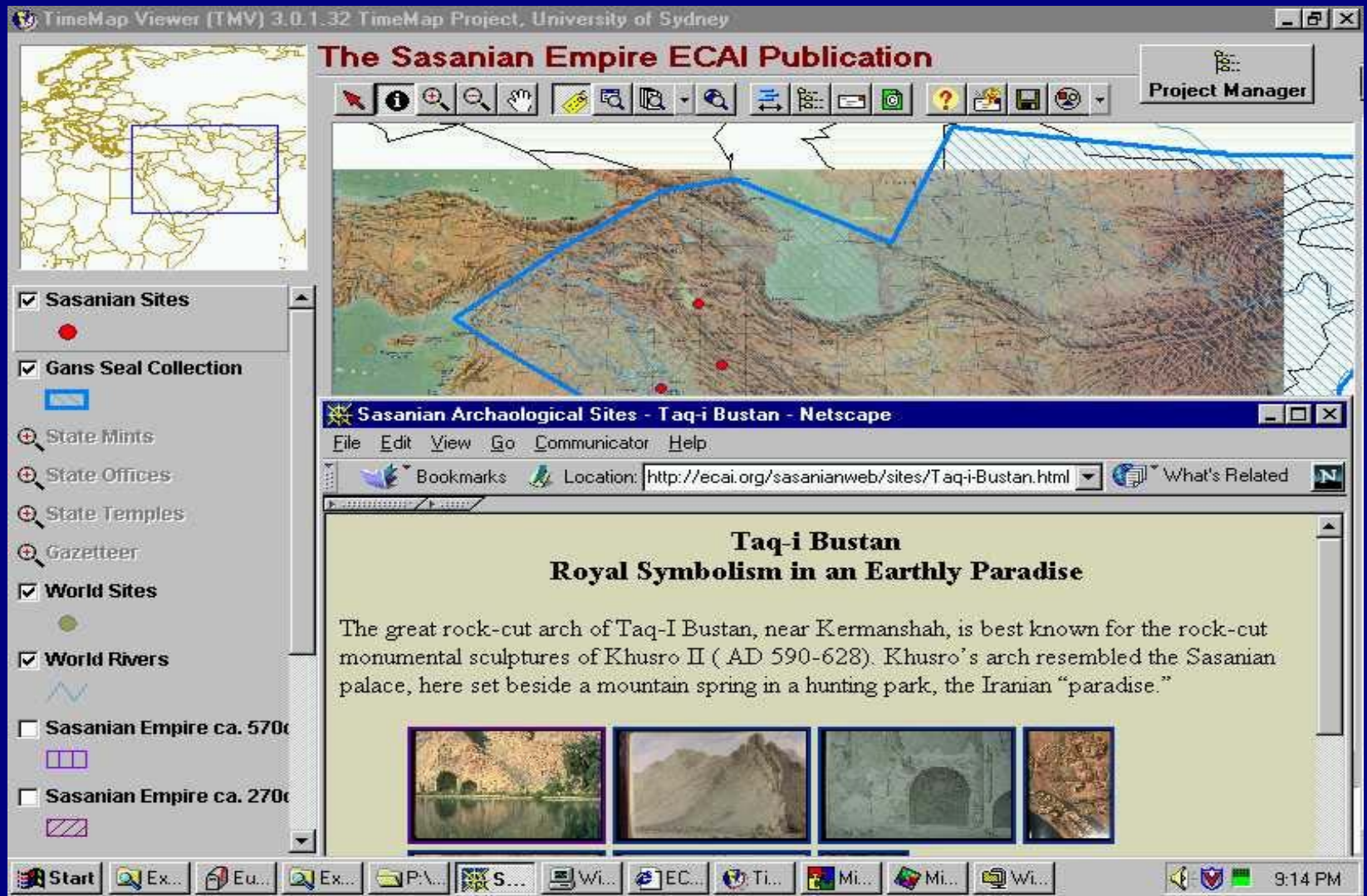
Users may zoom in to see resources that are only visible at a higher level of detail.



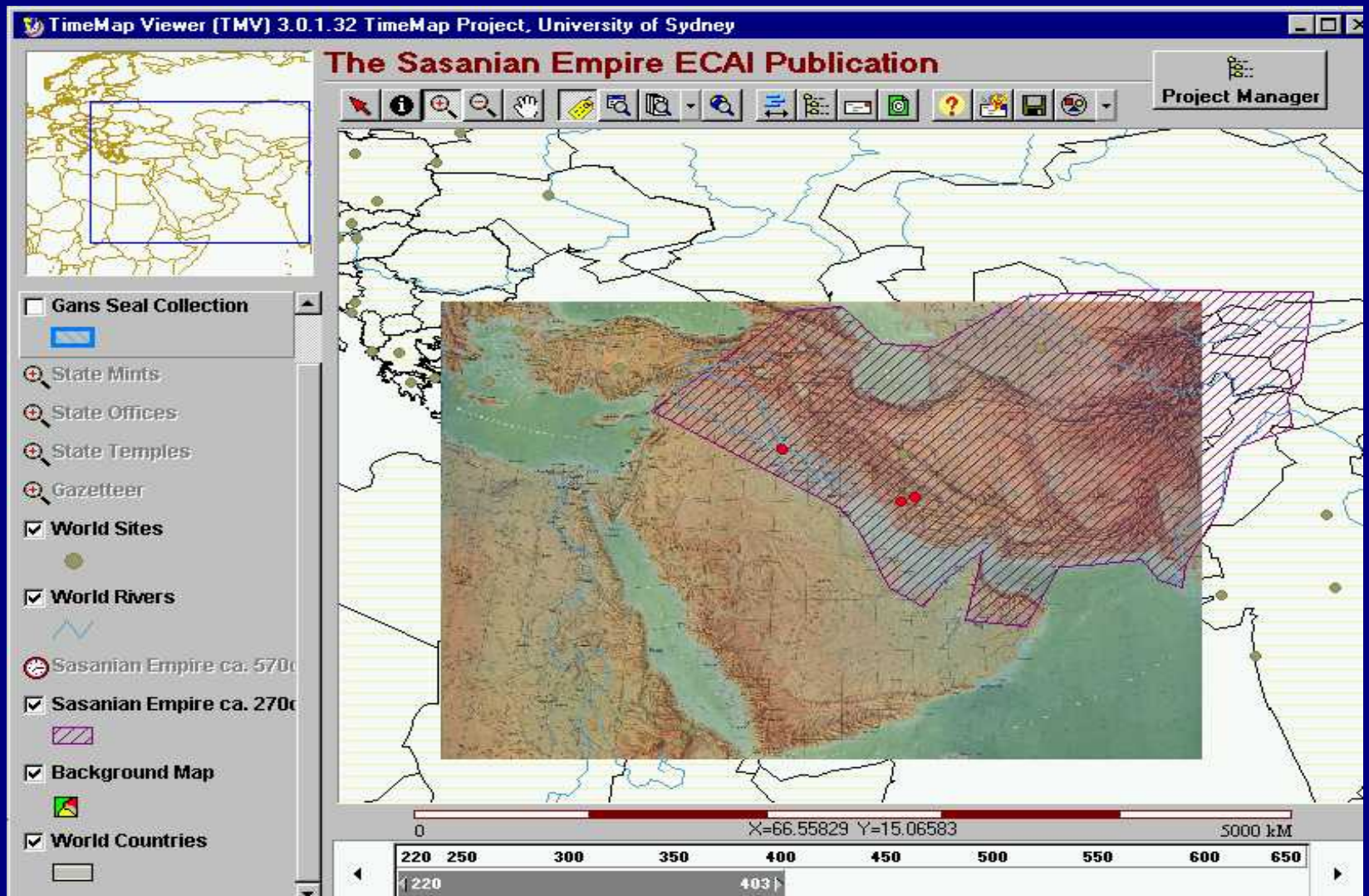
Spatial objects on the map are linked to a table of attributes, which may include any information about the objects. Note that this is a scholarly tool. By creating a “name quality” field, the author has noted that there is disagreement about the locations and names of places in the Sasanian Empire.



Sites on the map may be linked to resources elsewhere on the internet. In this case, important archaeological sites on the map are linked to web-based tours.

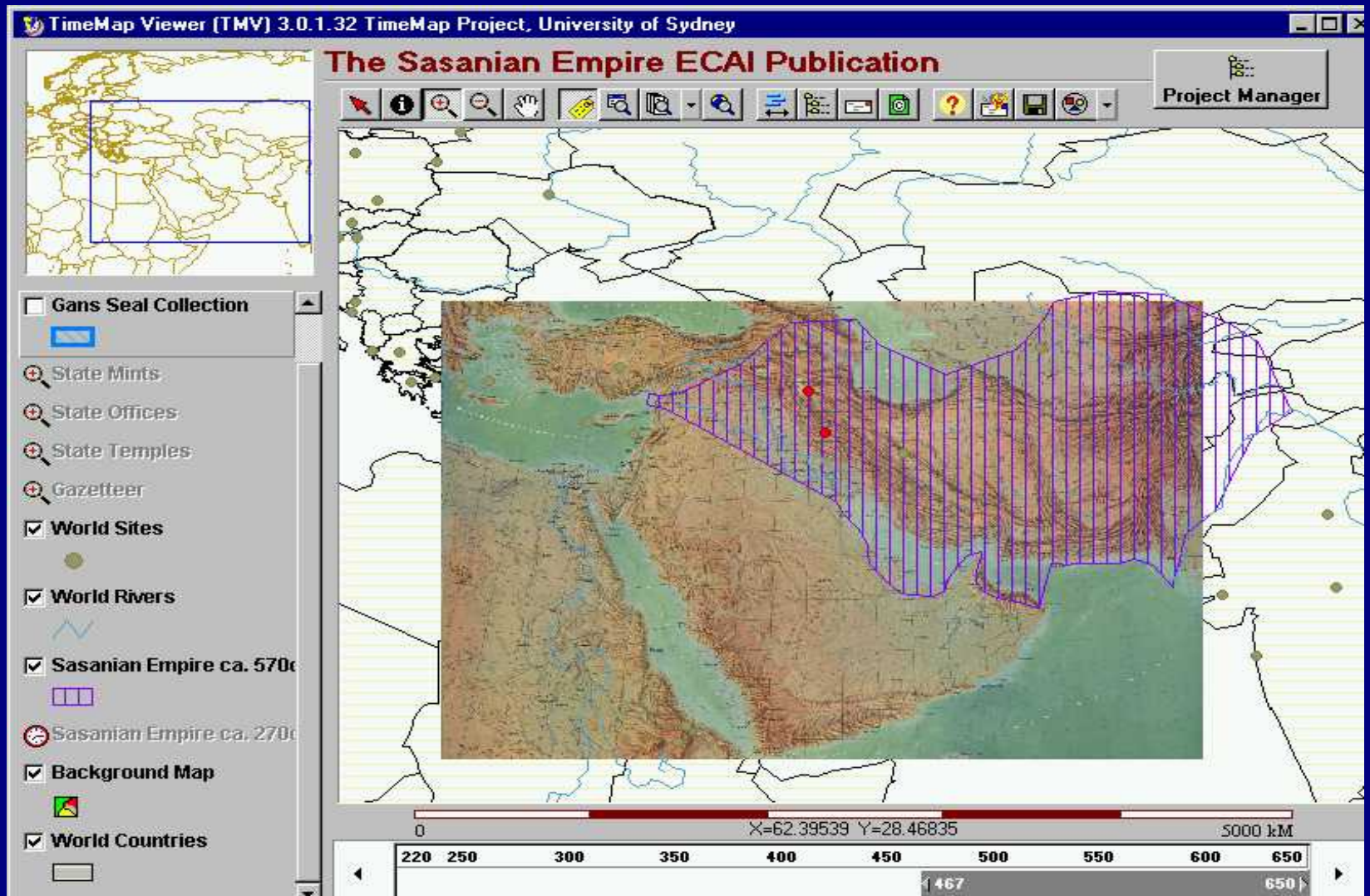


The map interface may be used to show change over time. The “Sasanian Empire ca. 270s” resource is highlighted, and the “Sasanian Empire ca. 570s” is greyed out. If a user slides the timeline bar, the new boundary of the empire will appear.



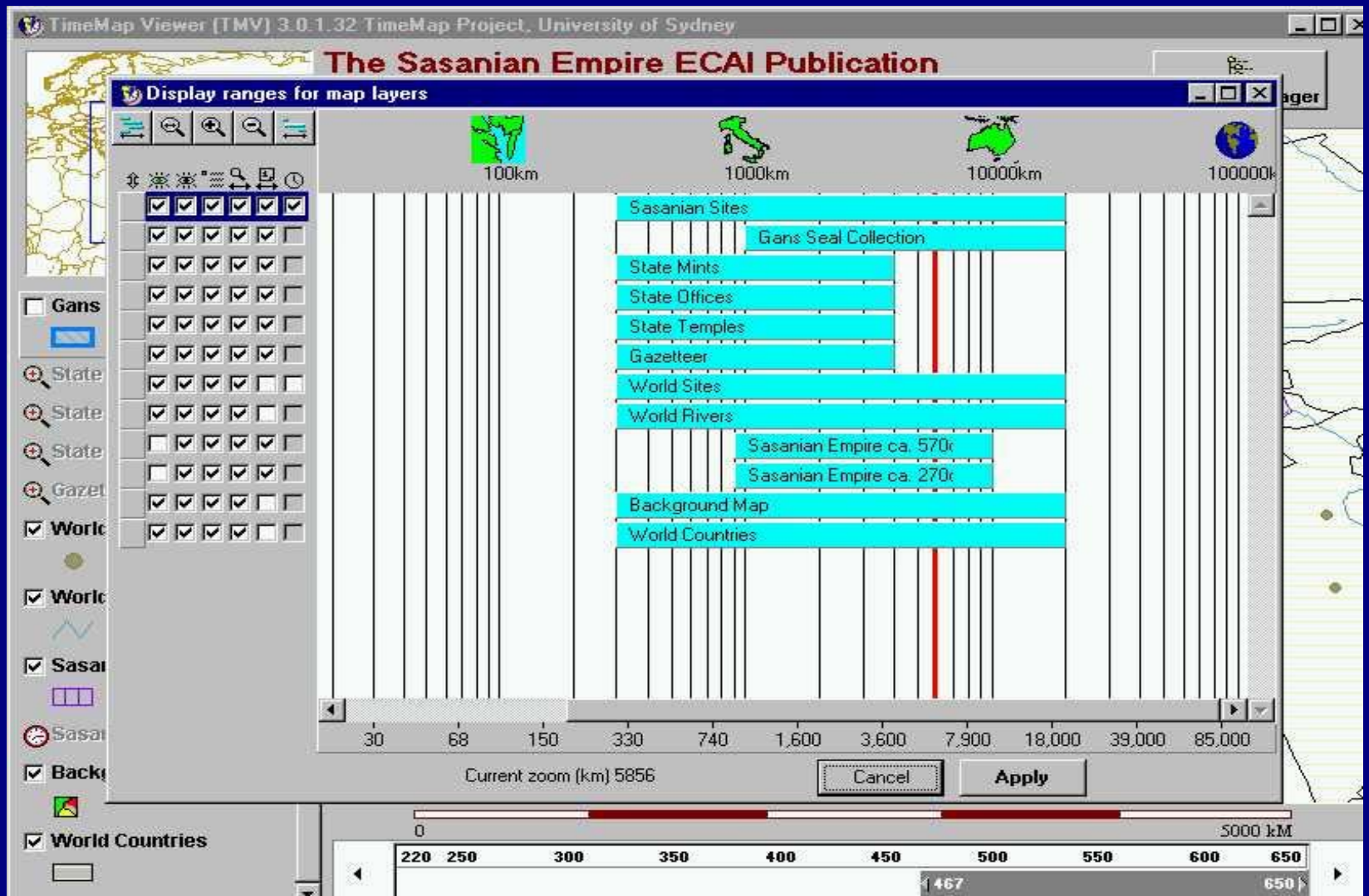
Y.T.

In a different time range, not only do the boundaries of the empire appear different, but the sites that were active during the earlier era (the red dots) have moved as well.



Y.T.

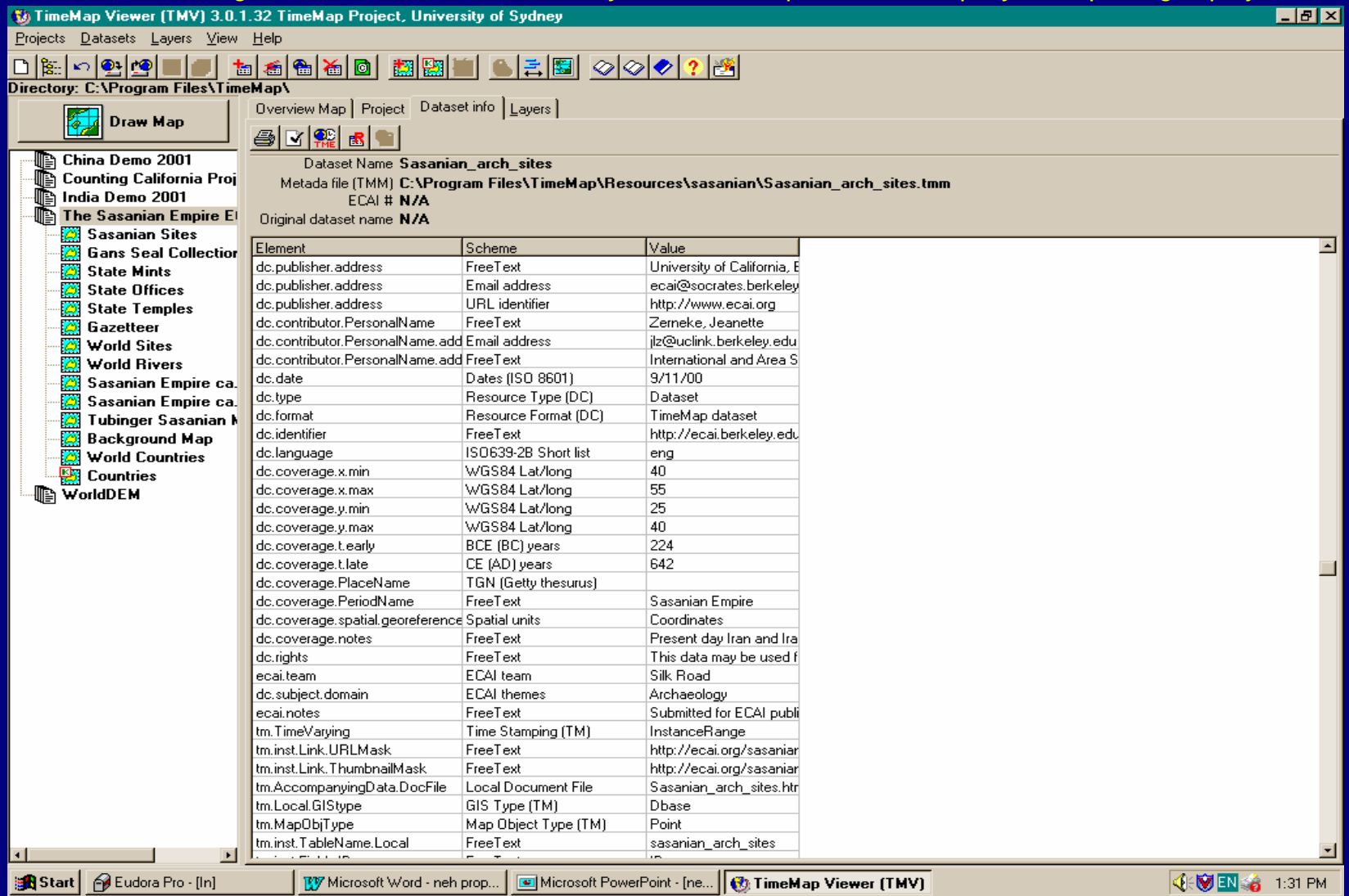
TimeMap is a user authoring tool, not merely a viewer. Users can control the look of the icons, the map layers that comprise a project, and, as shown here, the map scale at which different layers will become visible.



Y.T.

“DOK 324 PIR”

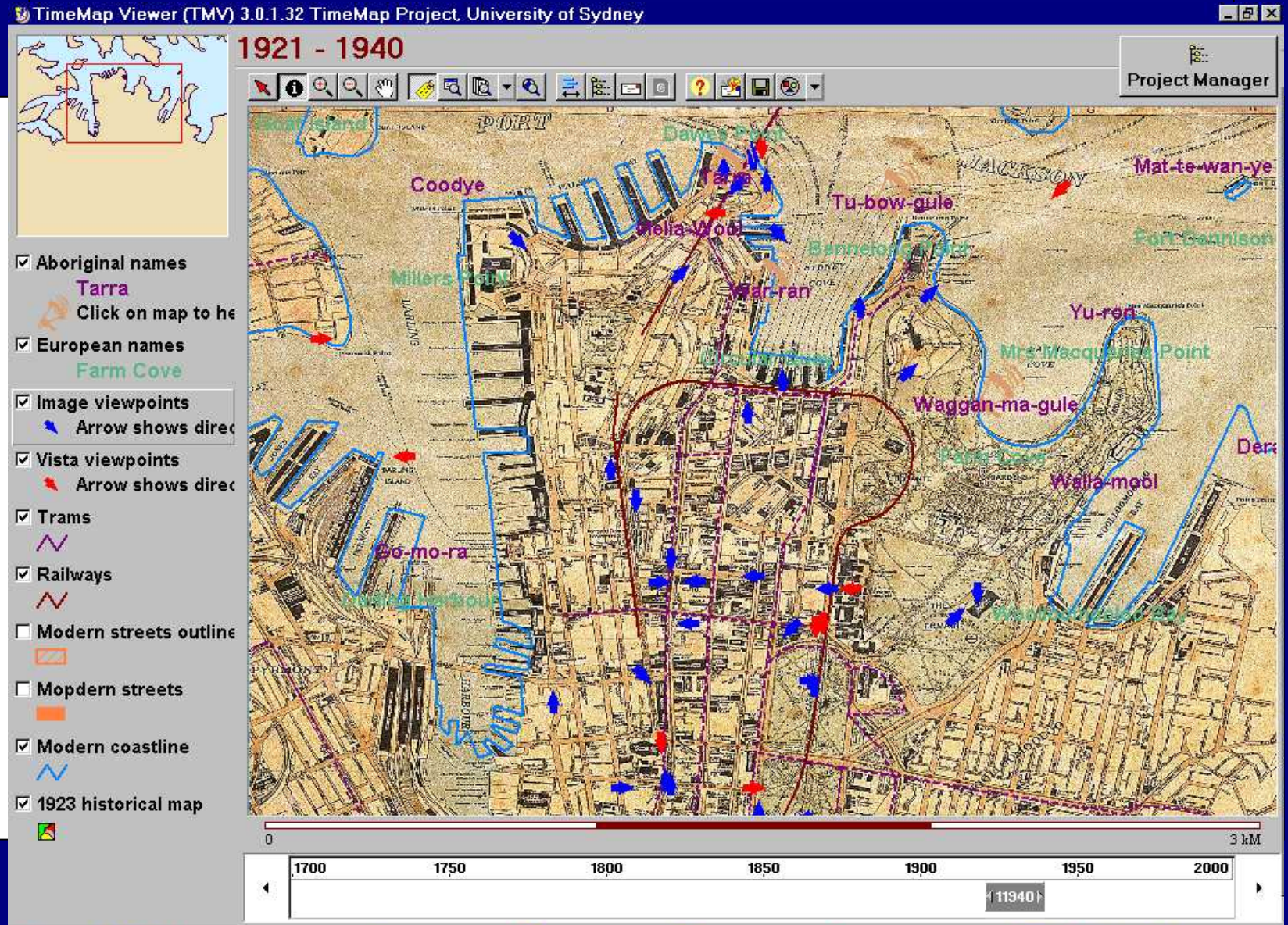
This screen displays the metadata for the a part of the Sasanian Empire project. The metadata includes functional (tm.) metadata to enable connection to the map interface in addition to cataloguing (dc. and ecai.) metadata. Using the menu on the left, users may choose to map individual map layers or packaged projects.



Y.T.

Historic Sydney

PIR”



Y.T.

The Mongol Empire