

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

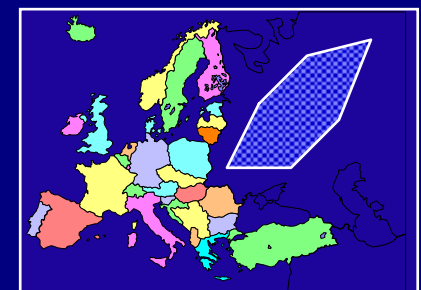
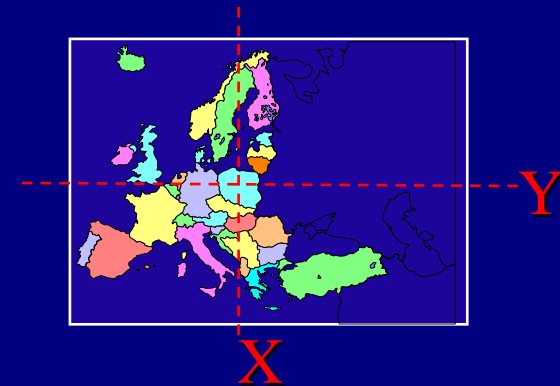
- ➔ Geographical coordinates are **geometric** relationships (distance and direction can be measured on a continuous scale)
 - E.g. “5.21 miles north of Champaign”
- ➔ Spatial relations may be both geometric and **topological** (spatially related but without measurable distance or absolute direction)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



Geographic and Spatial Querying

☞ Types of spatial queries

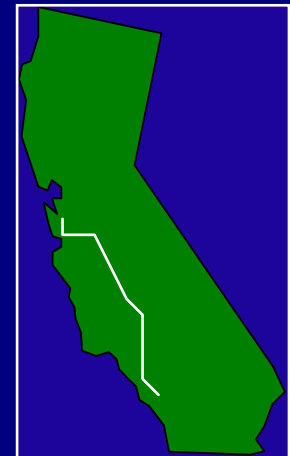
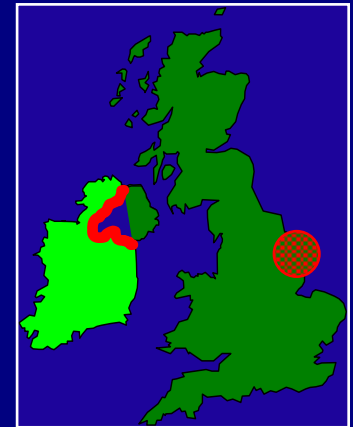
- Point – in - polygon : “What do we have at this X,Y point?”
- Region Queries : “What do we have in this region?”
 - ◆ Which point encoded items lie within the region
 - ◆ What lines (borders, etc.) lie within or the cross the region
 - ◆ What areas overlap the region area





Geographic and Spatial Querying

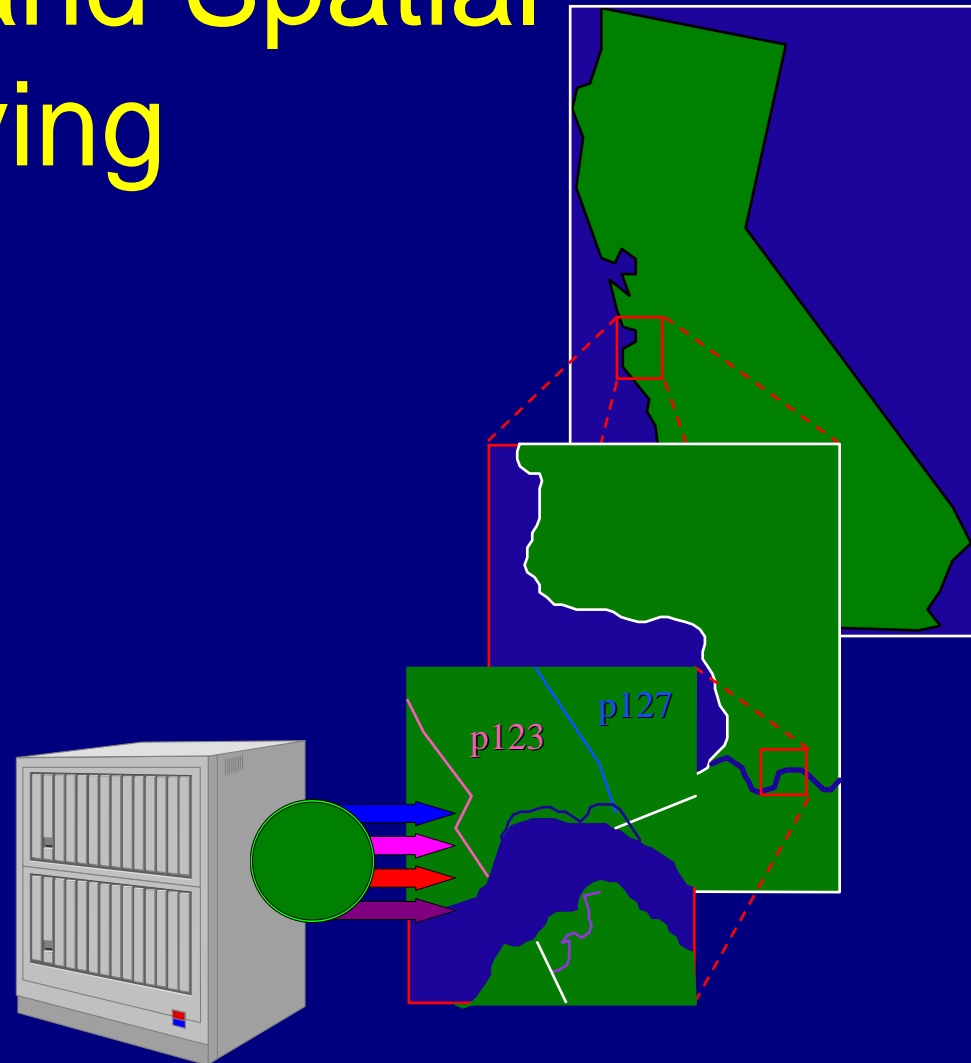
- ☞ Types of spatial queries, cont.
- Distance and Buffer Zone Queries
 - ◆ What cities lie within 40 miles of the border of Northern and Southern Ireland?
 - ◆ What wetlands lie within 50 miles of London?
 - Path Queries
 - ◆ What is the shortest route from San Francisco to Los Angeles?





Geographic and Spatial Querying

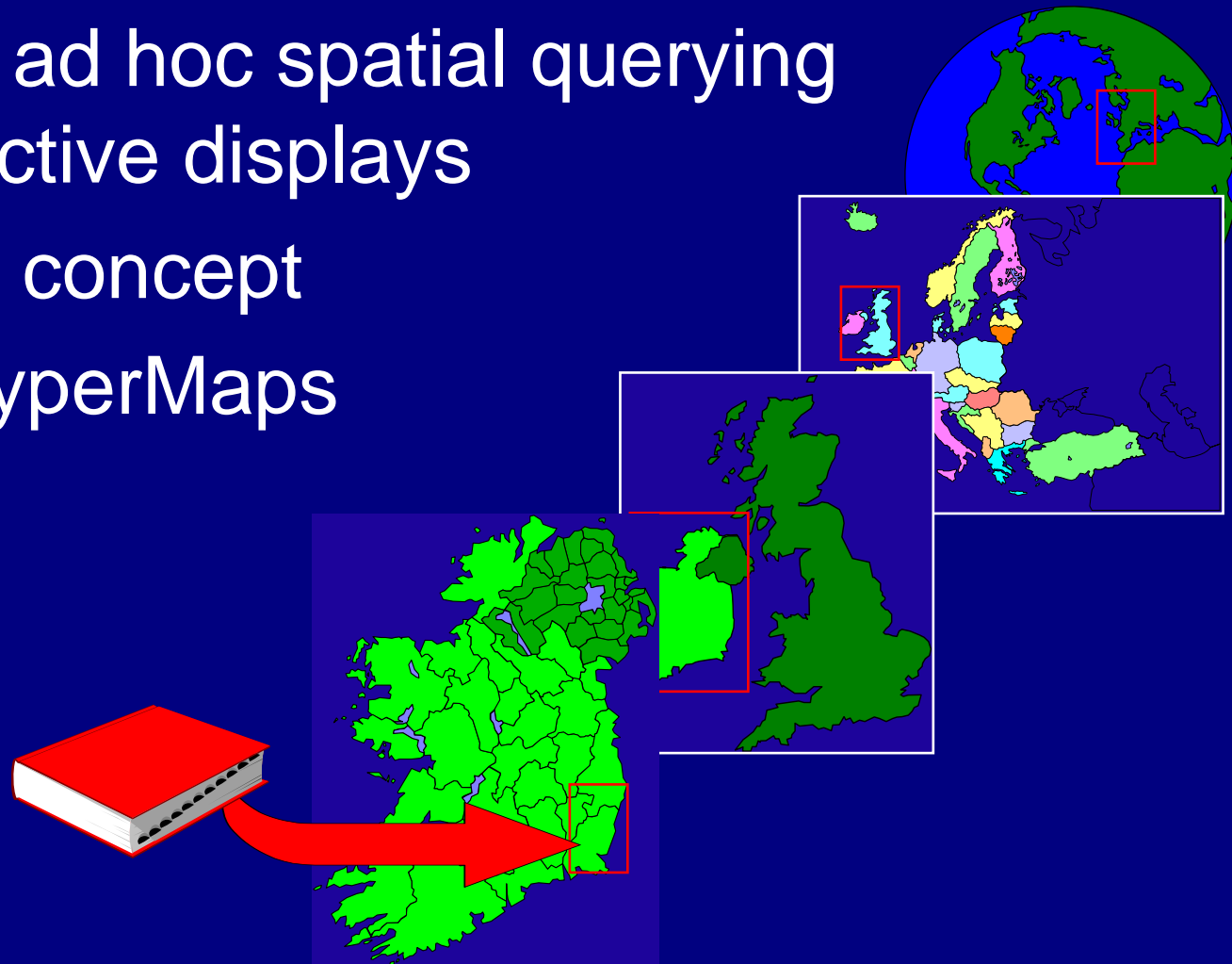
- ➔ Types of spatial queries, cont.
 - Multimedia Queries : Use non - map georeferenced information.
 - ◆ What are the names of farmers affected by flooding in Monterey and Santa Cruz Counties?





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

Geographic Feature ID: 1200

ADL Gazetter Content Standard

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

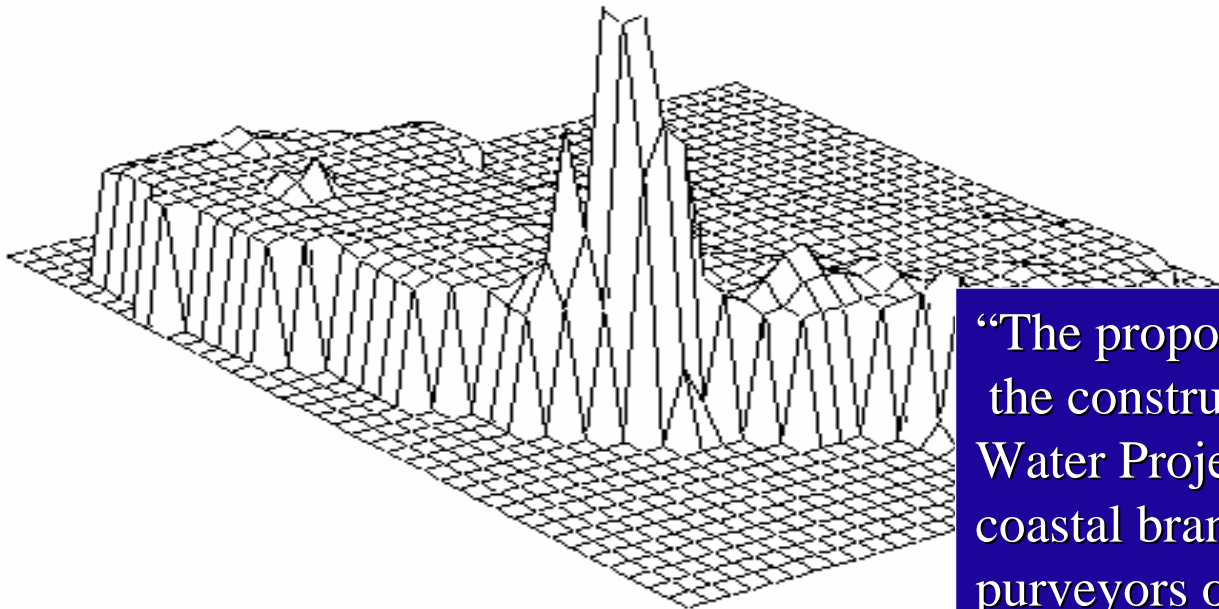
☞ GIPSY cont.

- 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 Clearinghouse](#)
[Directory](#)
[Members Area](#)
[Search](#) | [Home](#)

Electronic Cultural Atlas Initiative

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>



David Rumsey Map Collection -
Over 5000 georegistered historical

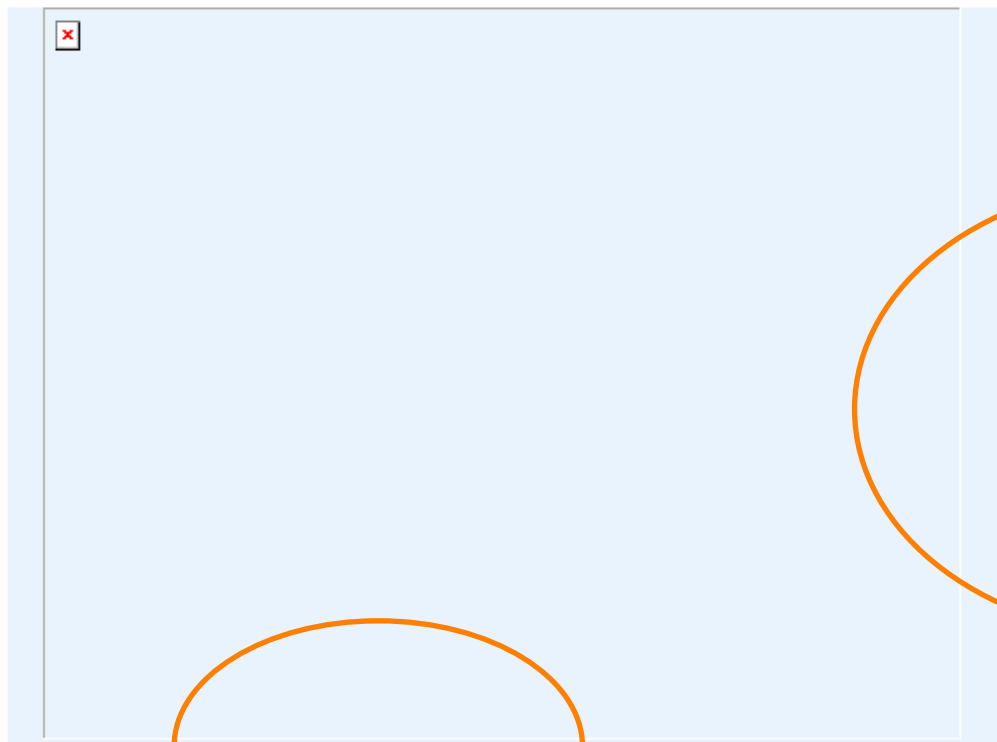


Electronic Cultural Atlas Initiative

Developed by University of Sydney
TimeMap™ Project



Clearinghouse Search



- Find area of interest with zoom and drag tools
- Click MapSpace bounding boxes for interactive map

Search Extent

Longitude to
Latitude to
Time to

Search Settings

Search collections:

Sort by:

3. Text search Blank = broadest search



Developed by University of Sydney
TimeMap™ Project



Electronic Cultural Atlas Initiative

ECAI Clearinghouse Search Results

Refine search

Longitude to Latitude to Time to

Find

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

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

1511: **[Rivers - Iraq Region](#)** (2003 CE)
Contributors: ECAI
Responsibility: Jeanette Zerneck <jz@uclink4.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 İzmit 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 İzmit 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 İzmit 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 Geograpy, Fatih University, Buyukcekmece-Istanbul, Turkey, 34900



[Clearinghouse](#)
[Directory](#)
[Members Area](#)
[Search](#) | [Home](#)

Electronic Cultural Atlas Initiative

[Activities](#) [Community](#) [Projects](#) [Technology](#) [Participating](#)

ECAI Technology

IT Infrastructure Components

- [The ECAI Metadata Clearinghouse](#)
- [TimeMap™ 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

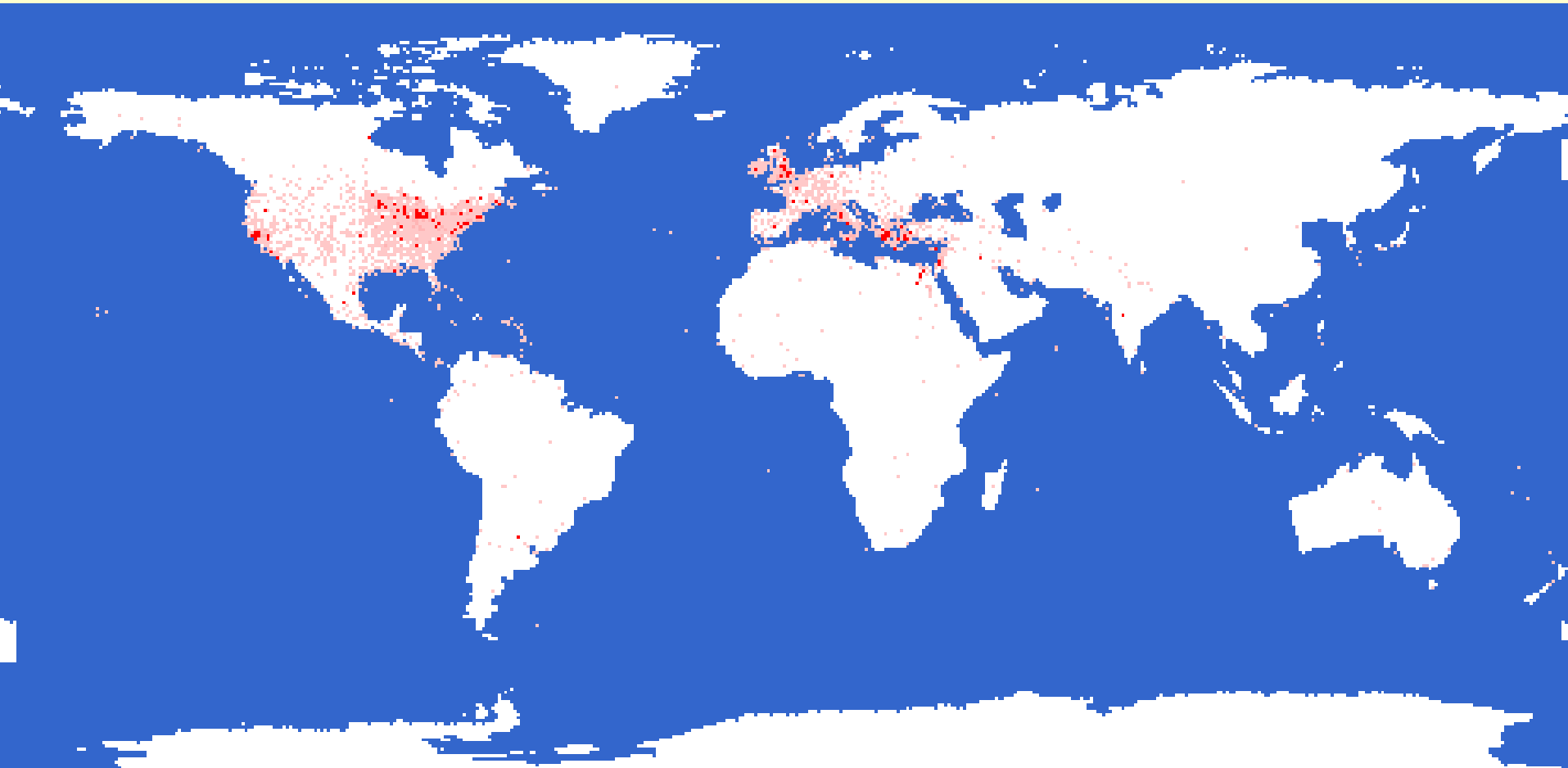
- The technical architecture and technology design is developed collaboratively by participants



=

Perseus

Perseus Digital Library





The Sasanian Empire

PIR”

TimeMap Viewer (TMV) 3.0.1.32 TimeMap Project, University of Sydney

The Sasanian Empire ECAI Publication

Project Manager

- Sasanian Sites
- Gans Seal Collection
- State Mints
- State Offices
- State Temples
- Gazetteer
- World Sites
- World Rivers
- Sasanian Empire ca. 570c
- Sasanian Empire ca. 270c
- Background Map

X=33.20088 Y=35.23179

0 8000 km

220 250 300 350 400 450 500 550 600 650

220 650



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

TimeMap Viewer (TMV) 3.0.1.32 TimeMap Project, University of Sydney

The Sasanian Empire ECAI Publication

Project Manager

- Sasanian Sites
- Gans Seal Collection
- State Mints
- State Offices
- State Temples
- Gazetteer
- World Sites
- World Rivers
- Sasanian Empire ca. 570c
- Sasanian Empire ca. 270c
- Background Map

0 X=33.20088 Y=35.23179 8000 km

220 250 300 350 400 450 500 550 600 650

220 650



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

The screenshot displays the TimeMap Viewer (TMV) 3.0.1.32 interface. The title bar reads "TimeMap Viewer (TMV) 3.0.1.32 TimeMap Project, University of Sydney". The main window title is "The Sasanian Empire ECAI Publication". The interface includes a toolbar with various navigation and tool icons, a "Project Manager" button, and a legend on the left side. The legend lists several categories with corresponding symbols:

- Sasanian Sites (Red circle)
- Gans Seal Collection (Blue square)
- State Mints (Green circle)
- State Offices (Blue circle)
- State Temples (Red triangle)
- Gazetteer (Yellow circle)
- World Sites (Brown circle)
- World Rivers (Blue line)
- Sasanian Empire ca. 570 (Pink square)

The main map area shows a topographic map of the Sasanian Empire region, outlined in blue. A scale bar at the bottom indicates 0 to 3000 km. The current coordinates are X=52.27737 Y=33.74513. The bottom status bar shows a zoomed-in view of the X-axis from 220 to 650.



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.

The screenshot shows the TimeMap Viewer (TMV) 3.0.1.32 interface. The main window is titled "The Sasanian Empire ECAI Publication". On the left, there is a legend with several checked items: Sasanian Sites (red dot), Gans Seal Collection (blue envelope), State Mints (green circle), State Offices (blue circle), State Temples (orange triangle), Gazetteer (yellow circle), World Sites (green circle), and World Rivers (blue line). The main map area shows a geographical map of the Sasanian Empire with various colored markers corresponding to the legend. A blue polygon highlights a specific region on the map. In the foreground, a "Data" window is open, displaying a table of attributes for the selected region. The table has columns for ID, LATITUDE, LONGITUDE, PLACEEXIS, PLACELOC, NAMEQUAL, and SASANIANNA. The data rows are as follows:

ID	LATITUDE	LONGITUDE	PLACEEXIS	PLACELOC	NAMEQUAL	SASANIANNA
1	28.2	53.8	Certain	Supposed	Certain	Kariyan
3	28.2	53.8	Certain	Supposed	Certain	Adur Farnbag
3	28.7	54.4	Certain	Certain	Supposed	Daravkirt
3	28.7	54.4	Certain	Certain	Supposed	Nev-Darap
5	28.5	53.1	Certain	Certain	Supposed	Kavat-Hvarrah
6	28.9	51.8	Certain	Certain	Supposed	Buht-Ardasir
8	29.5	51.2	Certain	Certain	Certain	Ram Ardasir
8	29.5	51.2	Certain	Certain	Certain	Tavvak
9	30.1	51.4	Certain	Certain	Supposed	Nubandagan
10	29.5	51.7	Certain	Certain	Supposed	Veh-Sahpuhr
13	28.9	53.6	Certain	Certain	Supposed	Ram Vistaspan

The bottom of the Data window shows the file path: C:\Program Files\TimeMap\Resources\sasanian\sasanian_gazetteer.htm.



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 screenshot displays a web browser window with the following elements:

- Browser Title Bar:** "TimeMap Viewer (TMV) 3.0.1_32 TimeMap Project, University of Sydney"
- Map Interface:** A main map window titled "The Sasanian Empire ECAI Publication" with a toolbar and a "Project Manager" button. A smaller inset map is visible in the top-left corner.
- Left Panel:** A list of site categories with checkboxes and corresponding icons:
 - Sasanian Sites (red dot icon)
 - Gans Seal Collection (blue square icon)
 - State Mints (magnifying glass icon)
 - State Offices (magnifying glass icon)
 - State Temples (magnifying glass icon)
 - Gazetteer (magnifying glass icon)
 - World Sites (green dot icon)
 - World Rivers (blue line icon)
 - Sasanian Empire ca. 570c (purple square icon)
 - Sasanian Empire ca. 270c (pink square icon)
- Browser Window:** A Netscape browser window titled "Sasanian Archaeological Sites - Taq-i Bustan - Netscape" is open. The address bar shows the URL: <http://ecai.org/sasanianweb/sites/Taq-i-Bustan.html>. The page content includes:
 - Section Header:** "Taq-i Bustan Royal Symbolism in an Earthly Paradise"
 - Text:** "The great rock-cut arch of Taq-I Bustan, near Kermanshah, is best known for the rock-cut monumental sculptures of Khusro II (AD 590-628). Khusro's arch resembled the Sasanian palace, here set beside a mountain spring in a hunting park, the Iranian "paradise.""
 - Images:** Four small thumbnail images showing different views of the Taq-i Bustan site, including the rock-cut arch and surrounding landscape.
- Taskbar:** The Windows taskbar at the bottom shows the Start button and several open application icons, including Explorer, Eu..., P:\..., S..., Wi..., EC..., Ti..., Mi..., and Wi... The system clock shows 9:14 PM.



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.

The screenshot displays the TimeMap Viewer (TMV) 3.0.1.32 interface. The title bar reads "TimeMap Viewer (TMV) 3.0.1.32 TimeMap Project, University of Sydney". The main window title is "The Sasanian Empire ECAI Publication". The interface includes a toolbar with various navigation and tool icons, a "Project Manager" button, and a "Gans Seal Collection" checkbox. The map shows the Sasanian Empire boundaries, with the "Sasanian Empire ca. 270s" highlighted in red hatching and the "Sasanian Empire ca. 570s" greyed out. The timeline bar at the bottom shows a range from 220 to 650, with a current position at 403. The map also displays a background map and world countries.

Legend:

- Gans Seal Collection
- State Mints
- State Offices
- State Temples
- Gazetteer
- World Sites
- World Rivers
- Sasanian Empire ca. 570s
- Sasanian Empire ca. 270s
- Background Map
- World Countries

Timeline: 220 250 300 350 400 450 500 550 600 650

Coordinates: X=66.55829 Y=15.06583

Scale: 5000 km



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.

The screenshot shows the TimeMap Viewer (TMV) 3.0.1.32 interface. The main window title is "The Sasanian Empire ECAI Publication". The map displays the Sasanian Empire boundaries in purple hatching. A red hatched area is overlaid on the map, and two red dots are visible within it. The left sidebar contains a legend with the following items:

- Gans Seal Collection
- State Mints
- State Offices
- State Temples
- Gazetteer
- World Sites
- World Rivers
- Sasanian Empire ca. 570c
- Sasanian Empire ca. 270c
- Background Map
- World Countries

The bottom of the window shows a scale bar from 0 to 5000 km and a coordinate display: X=62.39539 Y=28.46835. A coordinate grid at the bottom shows values from 220 to 650.



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.

TimeMap Viewer [TMV] 3.0.1.32 TimeMap Project, University of Sydney

The Sasanian Empire ECAI Publication

Display ranges for map layers

Layer	100km	1000km	10000km	100000km
Sasanian Sites	✓	✓	✓	✓
Gans Seal Collection	✓	✓	✓	✓
State Mints	✓	✓	✓	✓
State Offices	✓	✓	✓	✓
State Temples	✓	✓	✓	✓
Gazetteer	✓	✓	✓	✓
World Sites	✓	✓	✓	✓
World Rivers	✓	✓	✓	✓
Sasanian Empire ca. 570c	✓	✓	✓	✓
Sasanian Empire ca. 270c	✓	✓	✓	✓
Background Map	✓	✓	✓	✓
World Countries	✓	✓	✓	✓

Current zoom (km) 5856

0 5000 km

220 250 300 350 400 450 500 550 600 650

467 650



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.

TimeMap Viewer (TMV) 3.0.1.32 TimeMap Project, University of Sydney

Projects Datasets Layers View Help

Directory: C:\Program Files\TimeMap\

Draw Map

Overview Map | Project | Dataset info | Layers

Dataset Name **Sasanian_arch_sites**
Metadata file (TMM) C:\Program Files\TimeMap\Resources\sasanian\Sasanian_arch_sites.tmm
ECAI # **N/A**
Original dataset name **N/A**

Element	Scheme	Value
dc.publisher.address	FreeText	University of California, E
dc.publisher.address	Email address	ecai@socrates.berkeley
dc.publisher.address	URL identifier	http://www.ecai.org
dc.contributor.PersonalName	FreeText	Zerneke, Jeanette
dc.contributor.PersonalName.add	Email address	jlz@uclink.berkeley.edu
dc.contributor.PersonalName.add	FreeText	International and Area S
dc.date	Dates (ISO 8601)	9/11/00
dc.type	Resource Type (DC)	Dataset
dc.format	Resource Format (DC)	TimeMap dataset
dc.identifier	FreeText	http://ecai.berkeley.edu
dc.language	ISO639-2B Short list	eng
dc.coverage.x.min	WGS84 Lat/long	40
dc.coverage.x.max	WGS84 Lat/long	55
dc.coverage.y.min	WGS84 Lat/long	25
dc.coverage.y.max	WGS84 Lat/long	40
dc.coverage.t.early	BCE (BC) years	224
dc.coverage.t.late	CE (AD) years	642
dc.coverage.PlaceName	TGN (Getty thesurus)	
dc.coverage.PeriodName	FreeText	Sasanian Empire
dc.coverage.spatial.georeference	Spatial units	Coordinates
dc.coverage.notes	FreeText	Present day Iran and Ira
dc.rights	FreeText	This data may be used f
ecai.team	ECAI team	Silk Road
dc.subject.domain	ECAI themes	Archaeology
ecai.notes	FreeText	Submitted for ECAI publi
tm.TimeVarying	Time Stamping (TM)	InstanceRange
tm.inst.Link.URLMask	FreeText	http://ecai.org/sasanian
tm.inst.Link.ThumbnailMask	FreeText	http://ecai.org/sasanian
tm.AccompanyingData.DocFile	Local Document File	Sasanian_arch_sites.htr
tm.Local.GISType	GIS Type (TM)	Dbase
tm.MapObjType	Map Object Type (TM)	Point
tm.inst.TableName.Local	FreeText	sasanian_arch_sites

Start | Eudora Pro - [In] | Microsoft Word - neh prop... | Microsoft PowerPoint - [ne... | TimeMap Viewer (TMV) | EN | 1:31 PM



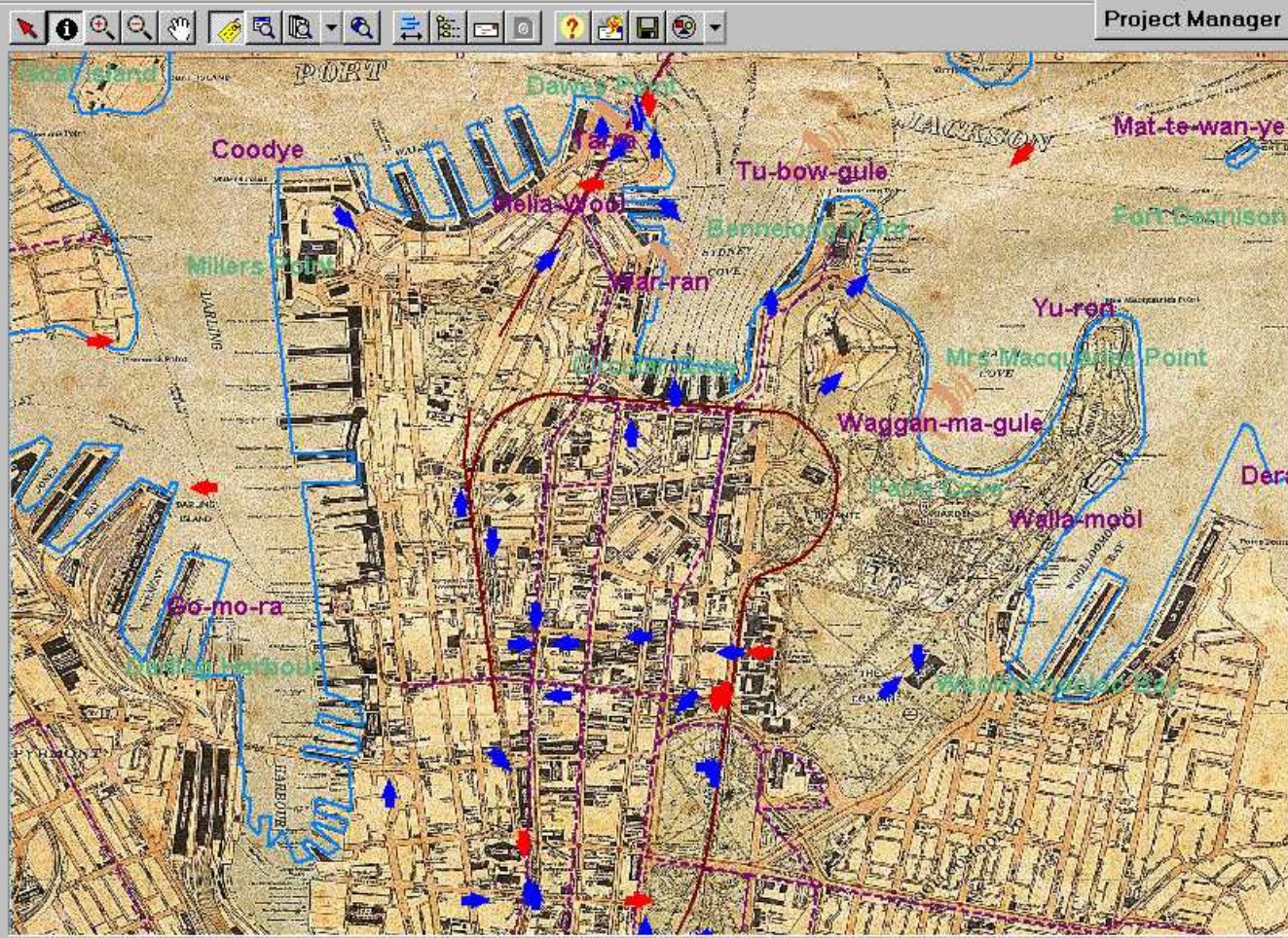
Historic Sydney

PIR”

TimeMap Viewer (TMV) 3.0.1.32 TimeMap Project. University of Sydney



1921 - 1940



Y.T.



The Mongol Empire

PIR''



A.D. 1100