Alexandria Digital Library

Greg Janée
gjanee@alexandria.ucsb.edu

http://www.alexandria.ucsb.edu/

Outline

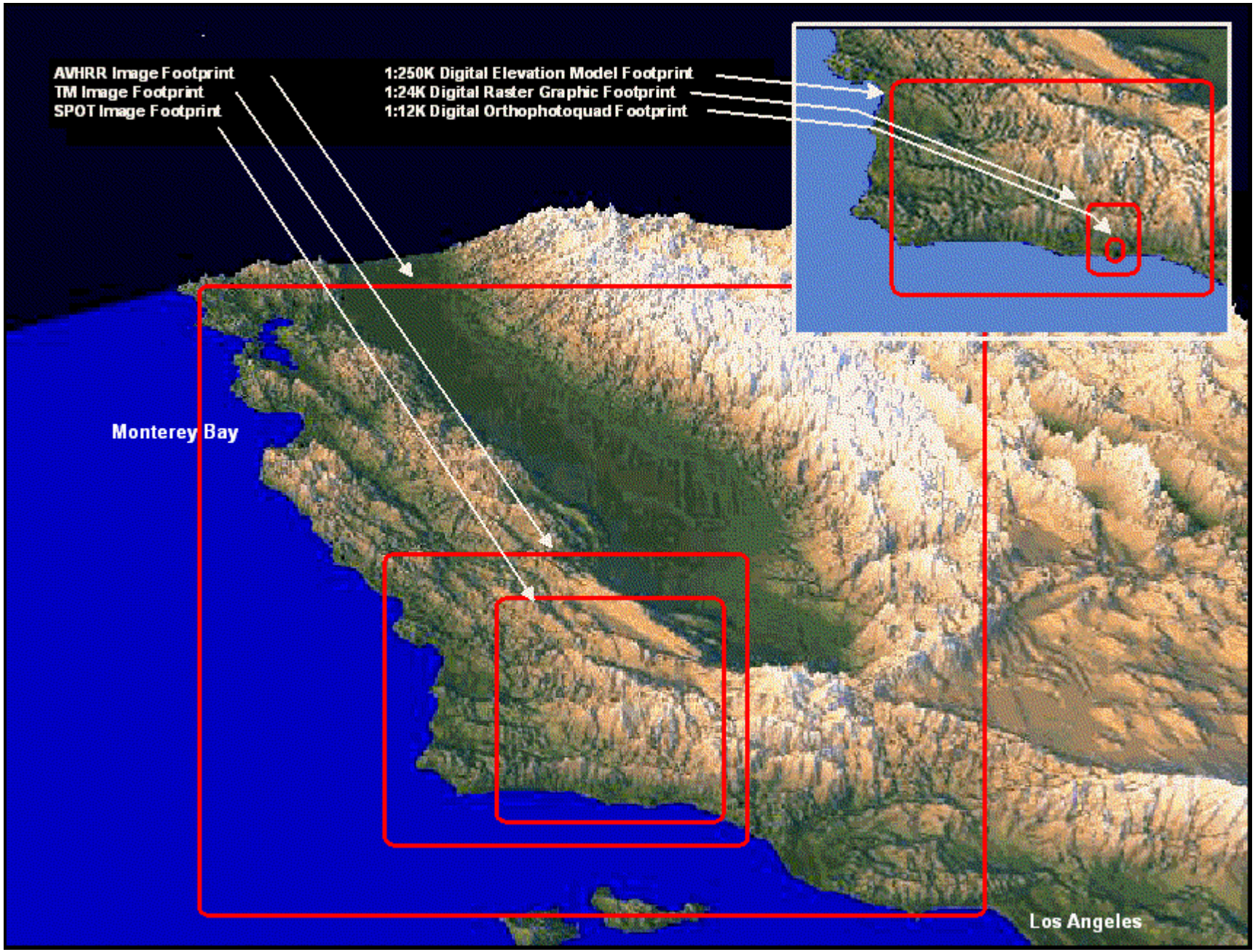
- Overview
- Data
- User Interface
- Architecture
- Development in progress
- Summary

Overview

- ADL is one of the 6 DLI/1 projects
- Focus is on georeferenced information
- Holistic approach: address all aspects of the problem of digital library creation
- Evolution of prototypes:
  - GIS-based
  - Web-, HTML-based
  - Latest: Java client with standardized middleware interfaces
- Slated to become part of the California Digital Library

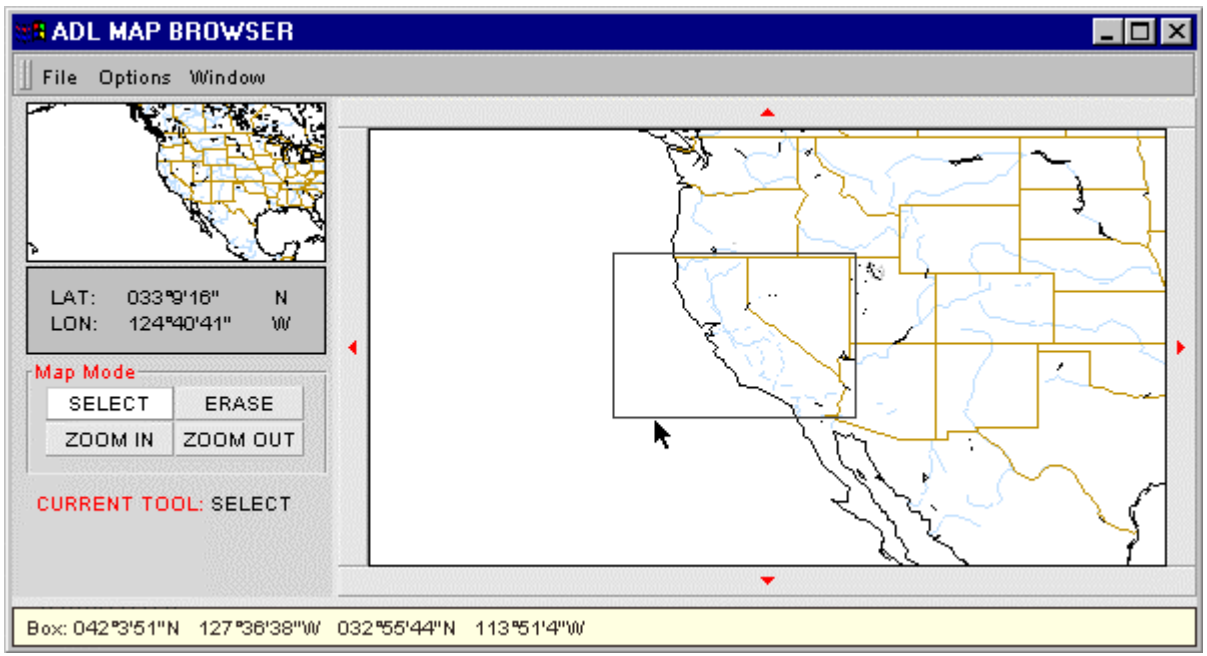
Data

- Geospatial and georeferenced materials
- Approaching 1M catalog holdings, 6M gazetteer entries, 0.5TB storage
- Emphasis on data heterogeneity
  - wide variation in content type: aerial photographs, U.S.G.S. products, scientific datasets, gazetteer data including seismic and volcanic activity, bibliographic references, etc.
  - wide variation in metadata content: FGDC, USMARC, numerous *ad hoc* content standards
  - wide variation in geographic extent and resolution

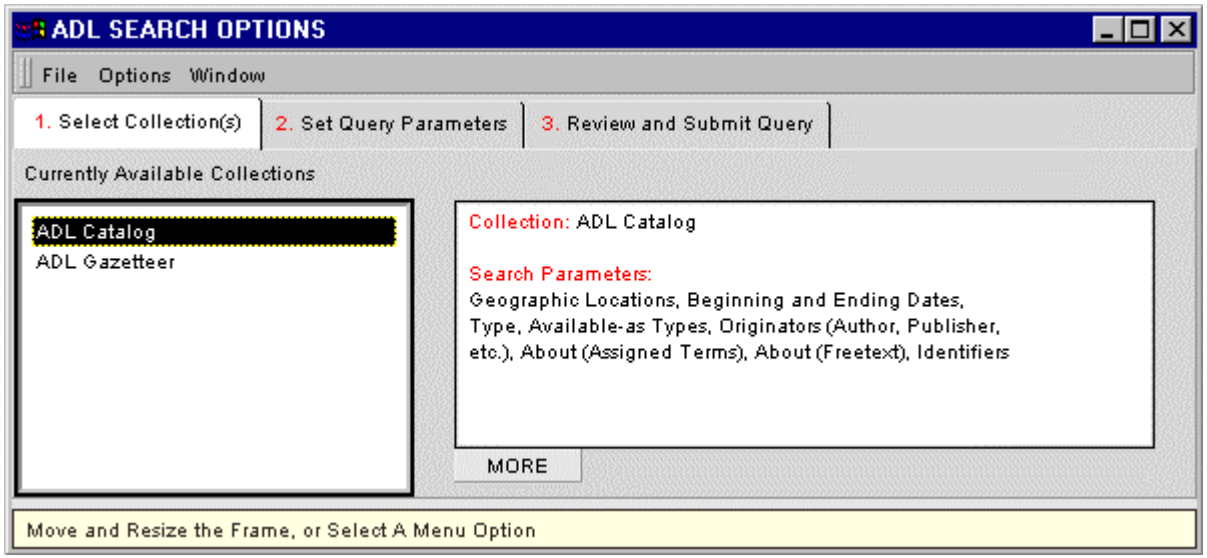

AVHRR Image Footprint
TM Image Footprint
SPOT Image Footprint
1:250K Digital Elevation Model Footprint
1:24K Digital Raster Graphic Footprint
1:12K Digital Orthophotoquad Footprint
Monterey Bay
Los Angeles

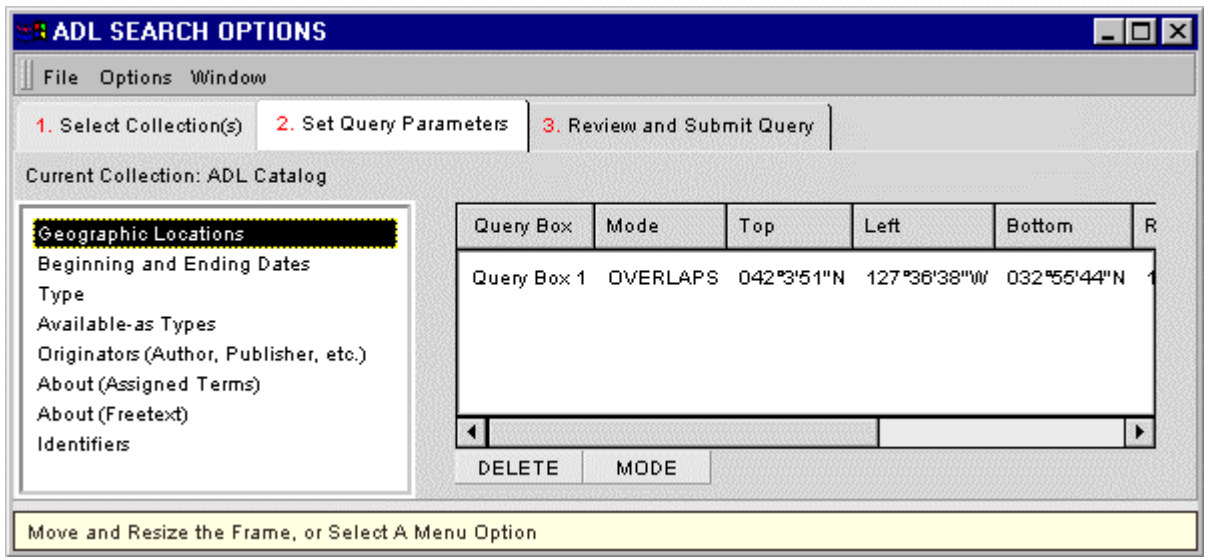
User interface

- Java client applet/application
- Maintains session and workspace
- Offers both spatial and textual paths into the system



User interface

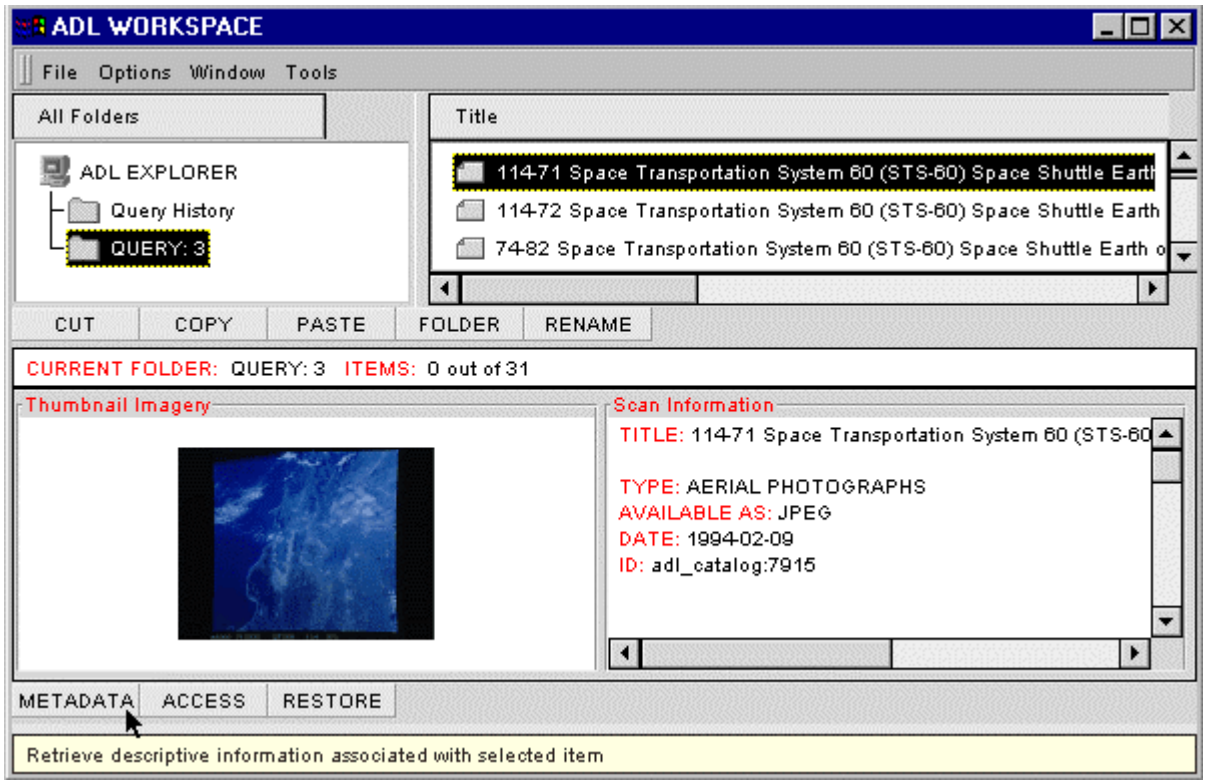




User interface

Query history and organization of results

Current result set

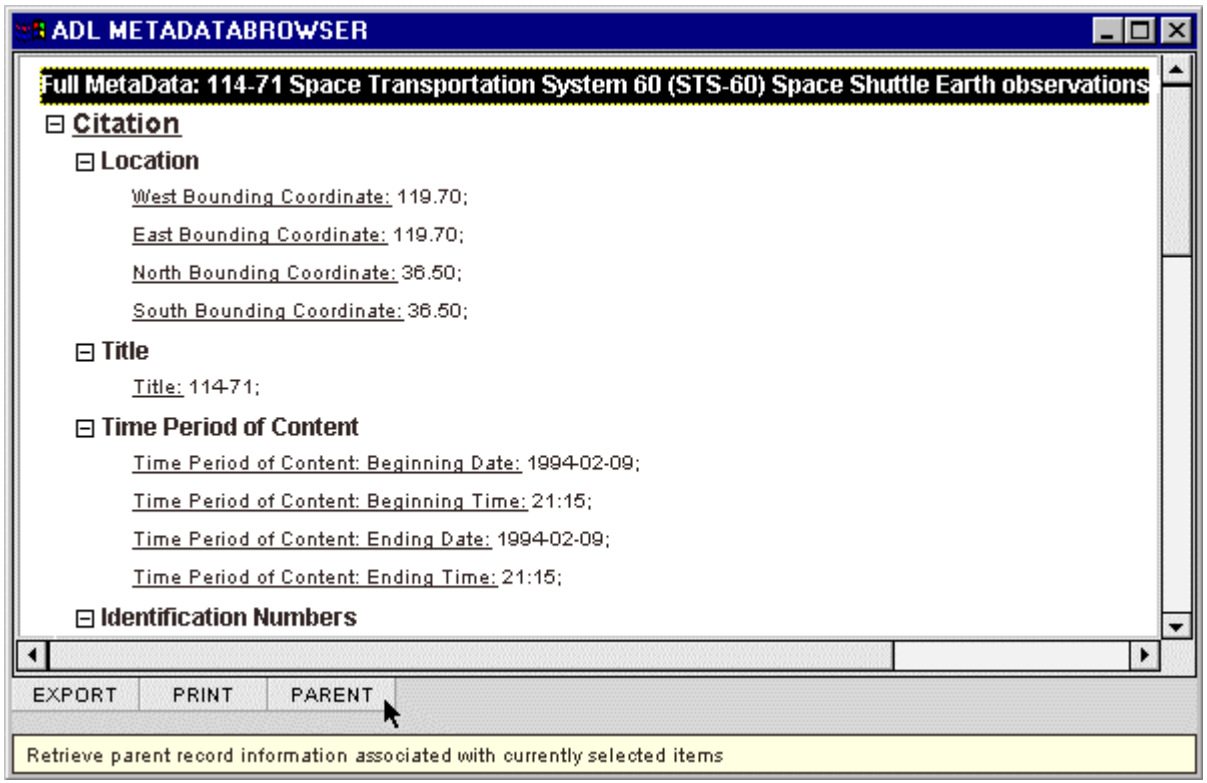


Thumbnail image

Metadata synopsis

User interface

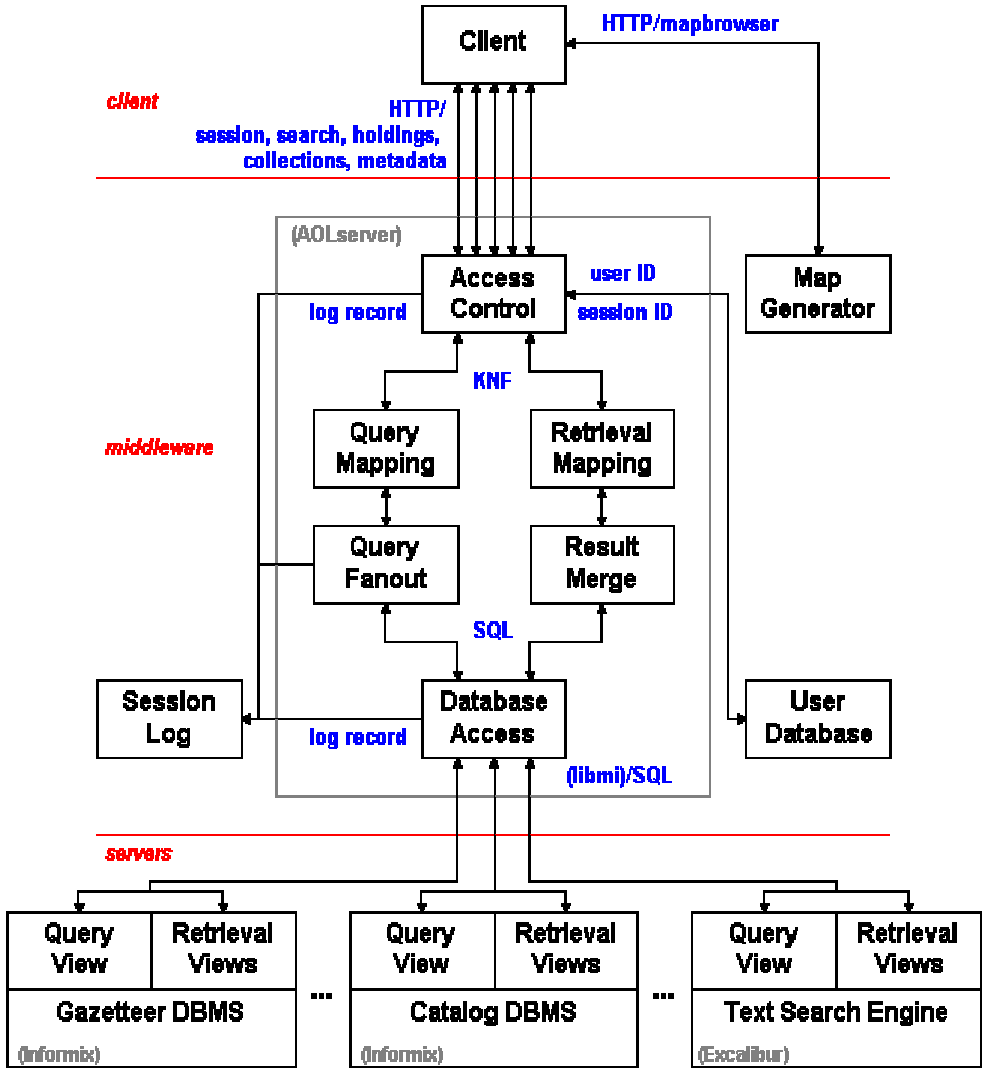
Metadata report:



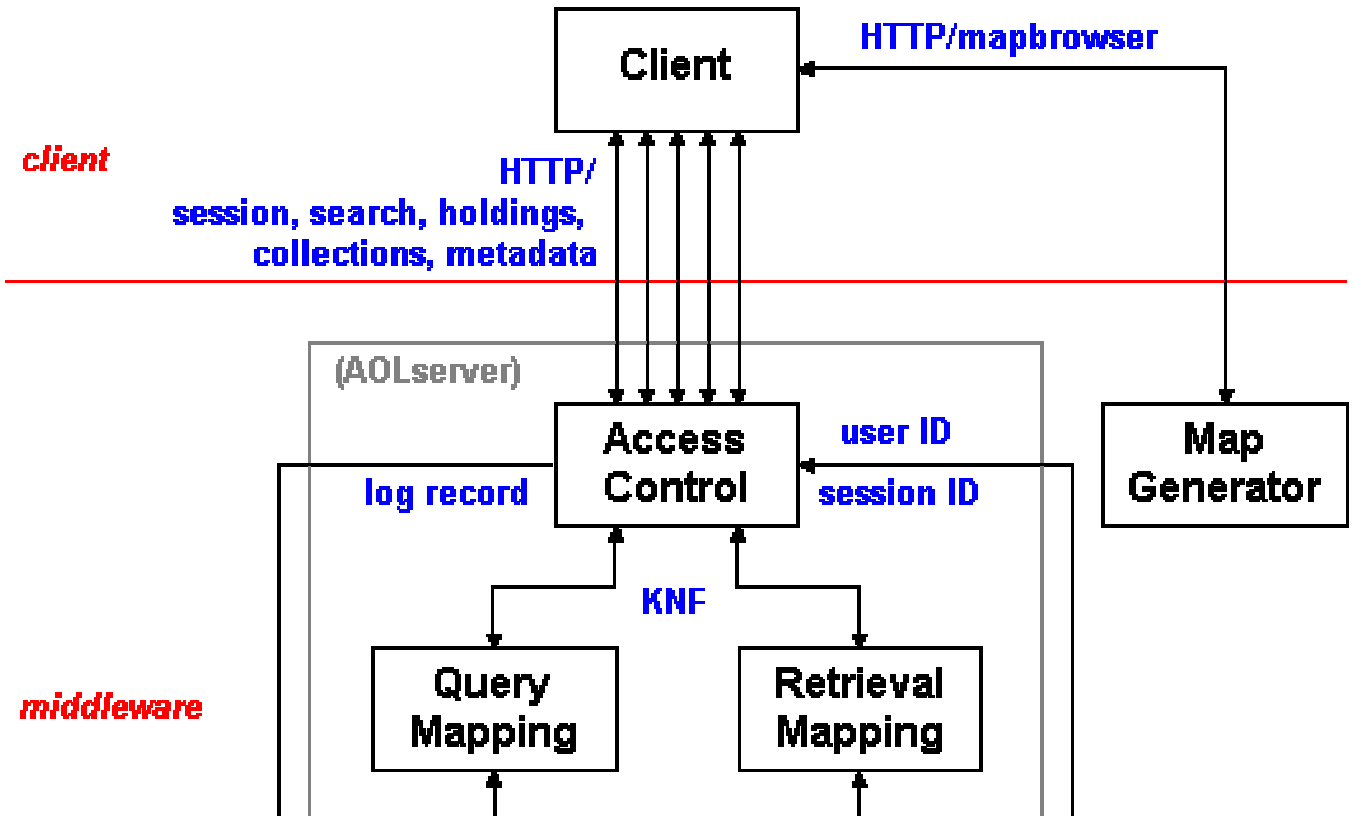
Access report is similar; includes “browse graphic,” access and use constraints, and hypertext link to actual file

Architecture

- 3-tier, client-server architecture
- Admits multiple clients, multiple collection servers
- Driving feature is the middleware server, which presents standard, collection-independent services to clients



Architecture



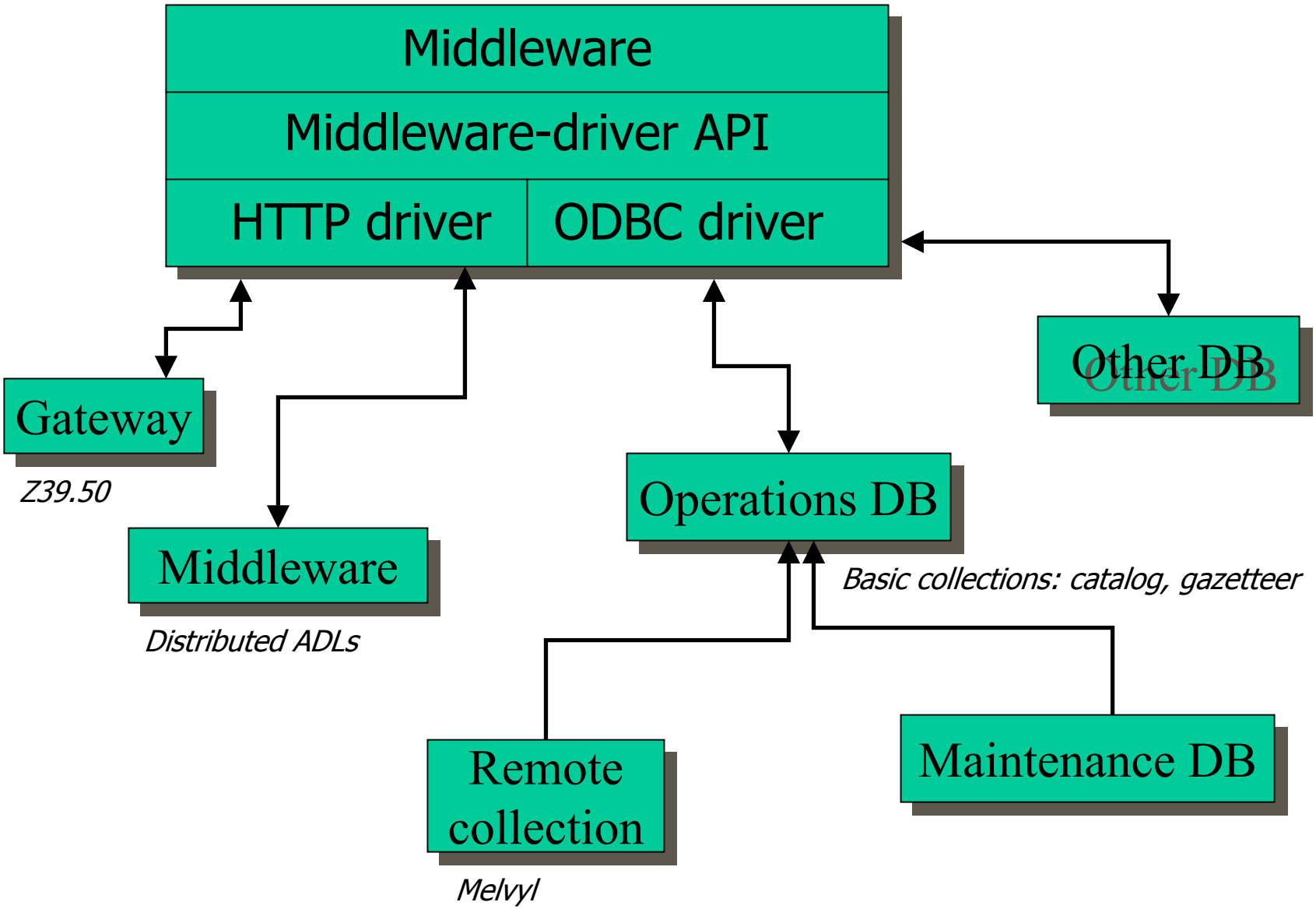
Architecture

- Search buckets:
  - Abstract, searchable indexes
  - Similar to GILS and Dublin Core, but buckets define the allowable content and search semantics, and they are optimized for geospatial searching
  - Designed to be easy for collections to populate
- Location, Time, Type, Format, Originator, Assigned terms, Free terms, Identifier

Development in progress

- Additional clients, e.g., Z39.50
- Java servlet-based middleware
- Better delineation of middleware-server interface
- Additional types of collections:
  - Gateways (e.g., Z39.50)
  - Remote collections (e.g., existing, traditional library catalogs)

Development in progress

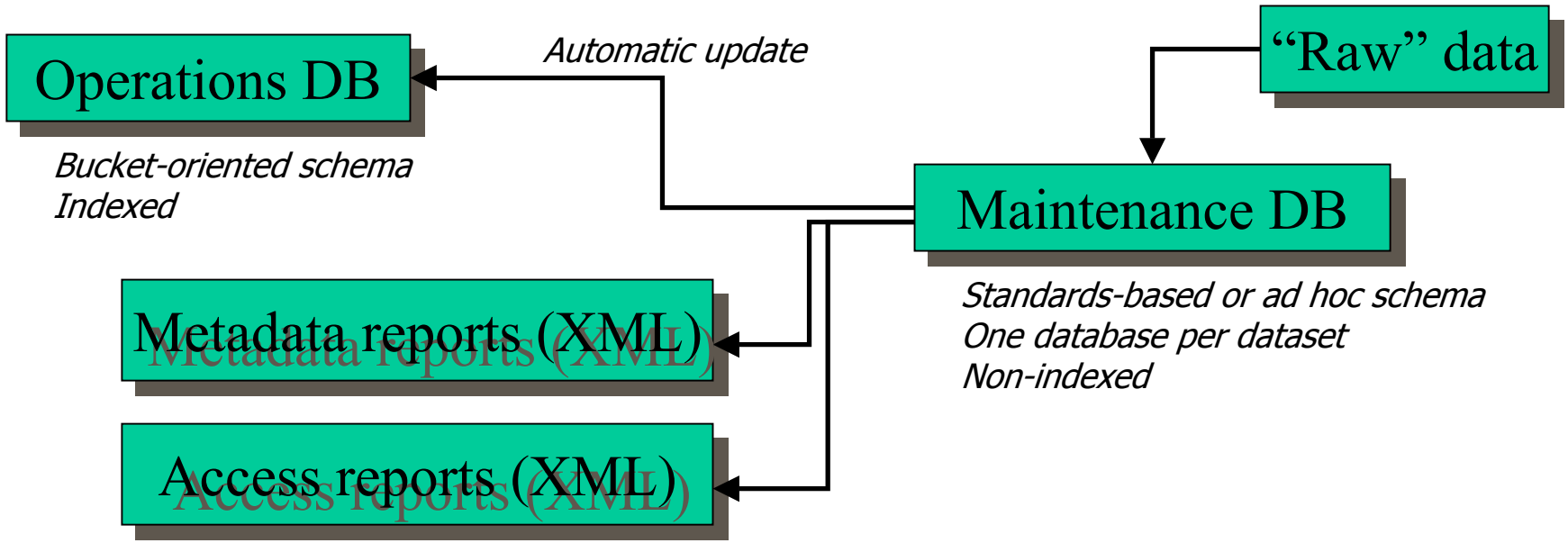


Development in progress

- Problem: incoming metadata must be mapped to FGDC and/or USMARC metadata standards
  - Current catalog schema explicitly implements FGDC and USMARC; has ~80 tables
- Ingest is inordinately difficult
- Report creation is slow

Development in progress

- Solution: use minimal, bucket-oriented schema that is independent of metadata content standards
- Solution: pre-compute the reports



Summary

- ADL: a digital library for georeferenced information with user and programmatic interfaces
- Offers uniform access to heterogeneous collections
- Transitioning from research prototype to operational system
- Actively ingesting new material
- More portable system is under development