



Architectural Directions for Distributed Geolibraries

Greg Janée
gjanee@alexandria.ucsb.edu



Outline

- Vision & goals
- Architecture: foundation
 - CRADDL
 - FEDORA
- Architecture: additions
 - Standard thesaurus interface
 - Two standard metadata models
 - Core index/search service
 - Tile-based browse/aggregation service
 - Collection discovery service
 - Gazetteer service
- Goals, revisited



Vision

- Fundamental organization of information
 - Self-contained, georeferenced digital objects...
 - ...aggregated into collections...
 - ...networked into libraries
 - **T**he **L**ibrary constitutes a “Digital Earth”
- Static and dynamic content
- Personal, customizable collections
- Collaborative use of distributed resources
- Component-based approach
 - Specify interfaces and protocols
 - Build representative services



Goals

□ Find

- Find appropriate collections
- Find items within collections
 - Using simplified, uniform methods
 - Using more refined, perhaps collection-specific methods

□ Assemble, structure, publish

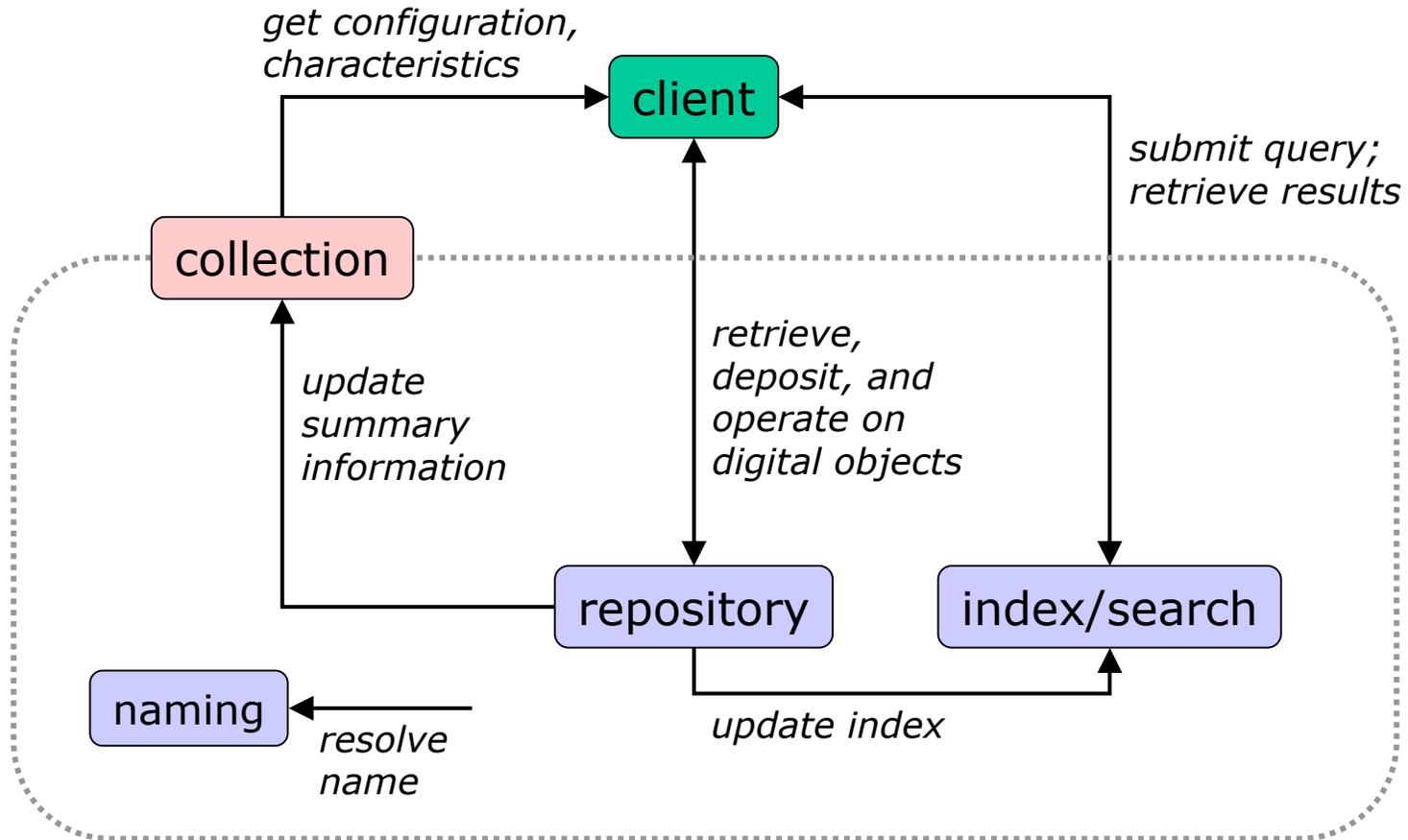
- Create and populate new collections
- Structure collections using domain-specific thesauri
- Make available to others

□ Use

- Invoke operations on items
- Integrate library into user application environment



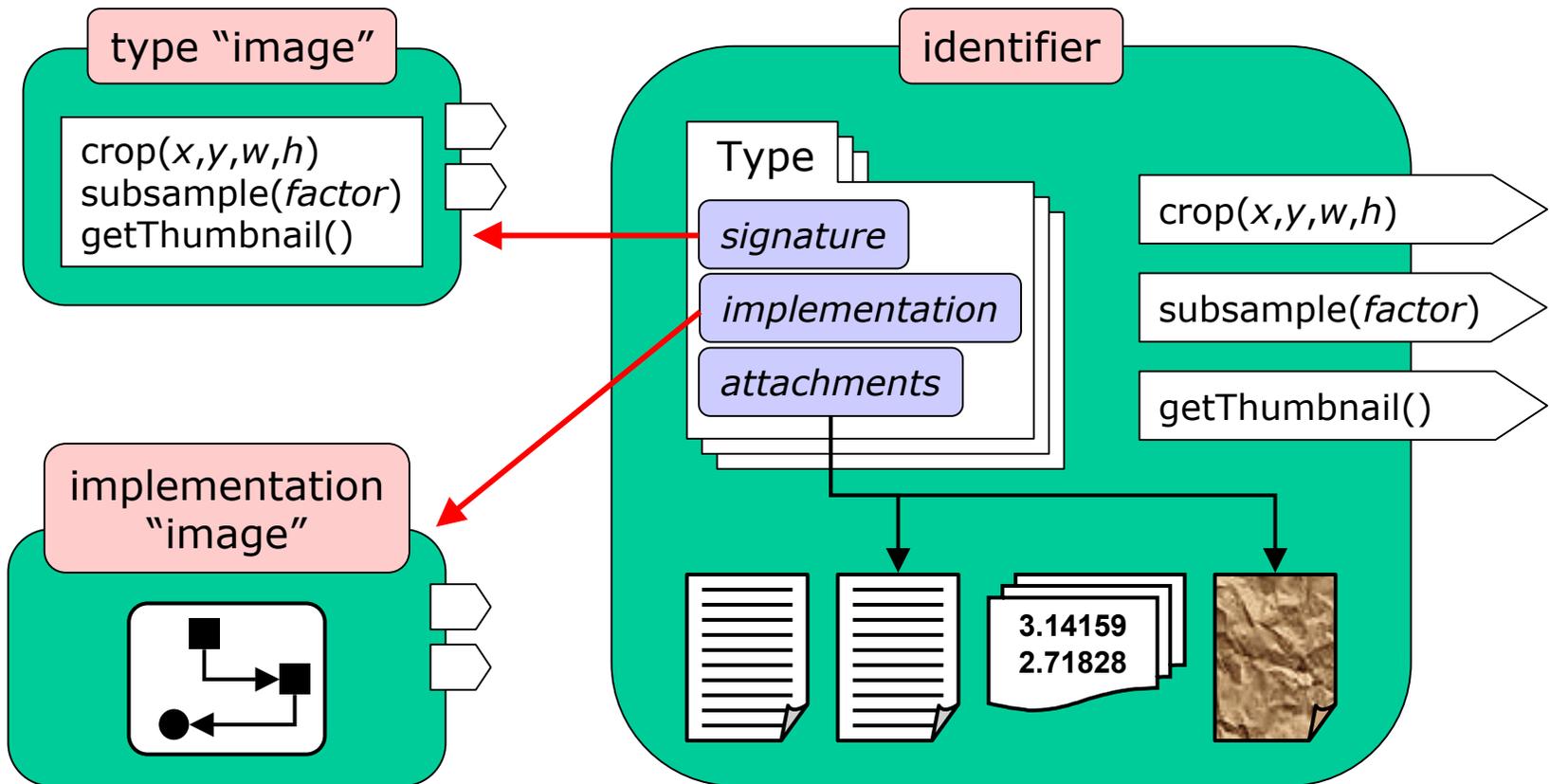
CRADDL



D-Lib Magazine, Nov. 1998



FEDORA



apologies to Christophe Blanchi

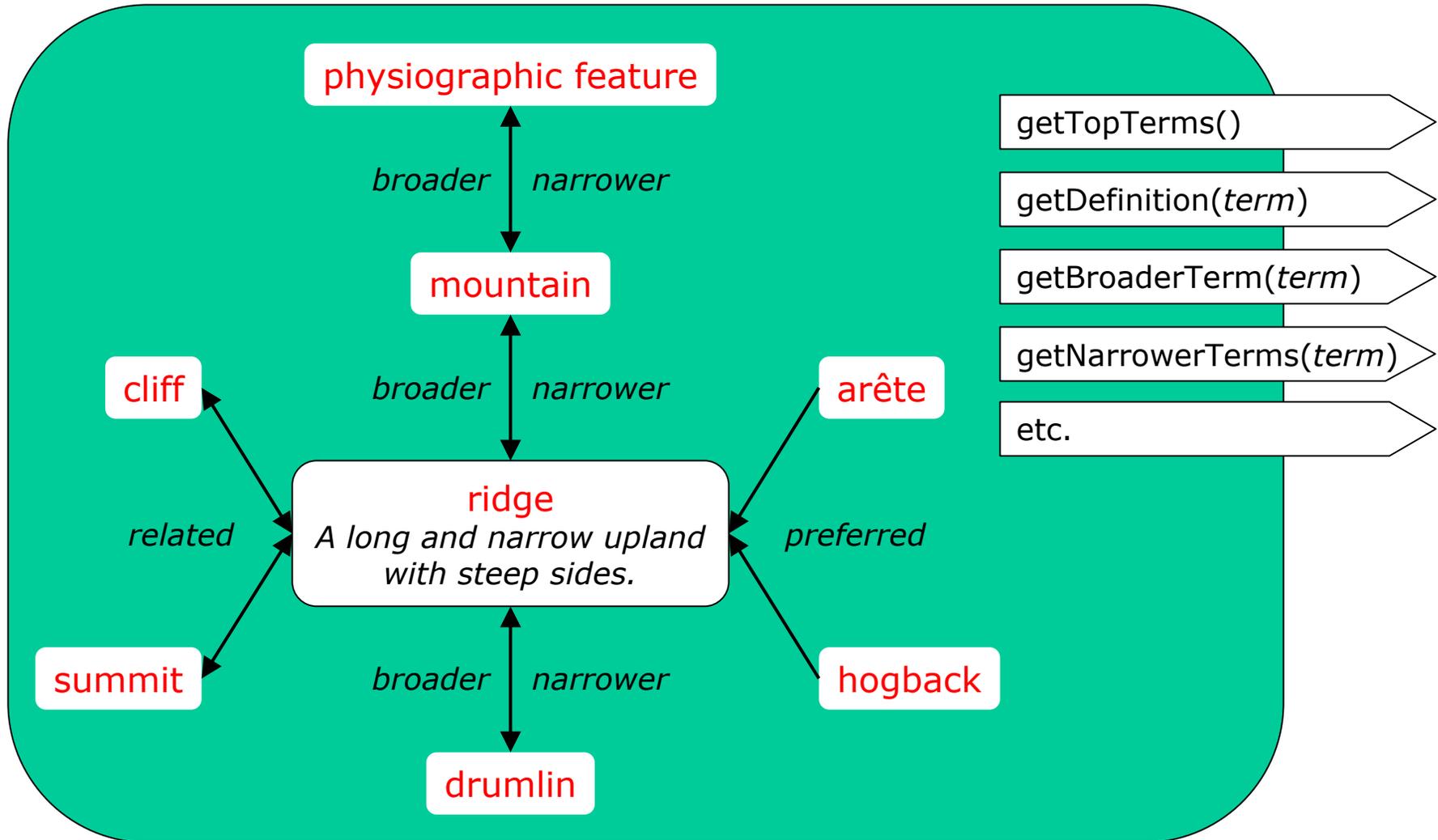


Architecture: additions

- Standard thesaurus interface
- Two standard metadata models
 - *ADL-Basic*: supports uniform description & search
 - *ADL-Full*: supports XML-based querying on entire metadata
- Standard index/search services
 - Core index/search service
 - Tile-based browse/aggregation service
- Collection metadata
 - Characterizes collection
 - Supports collection discovery
- Gazetteer service



Standard thesaurus interface



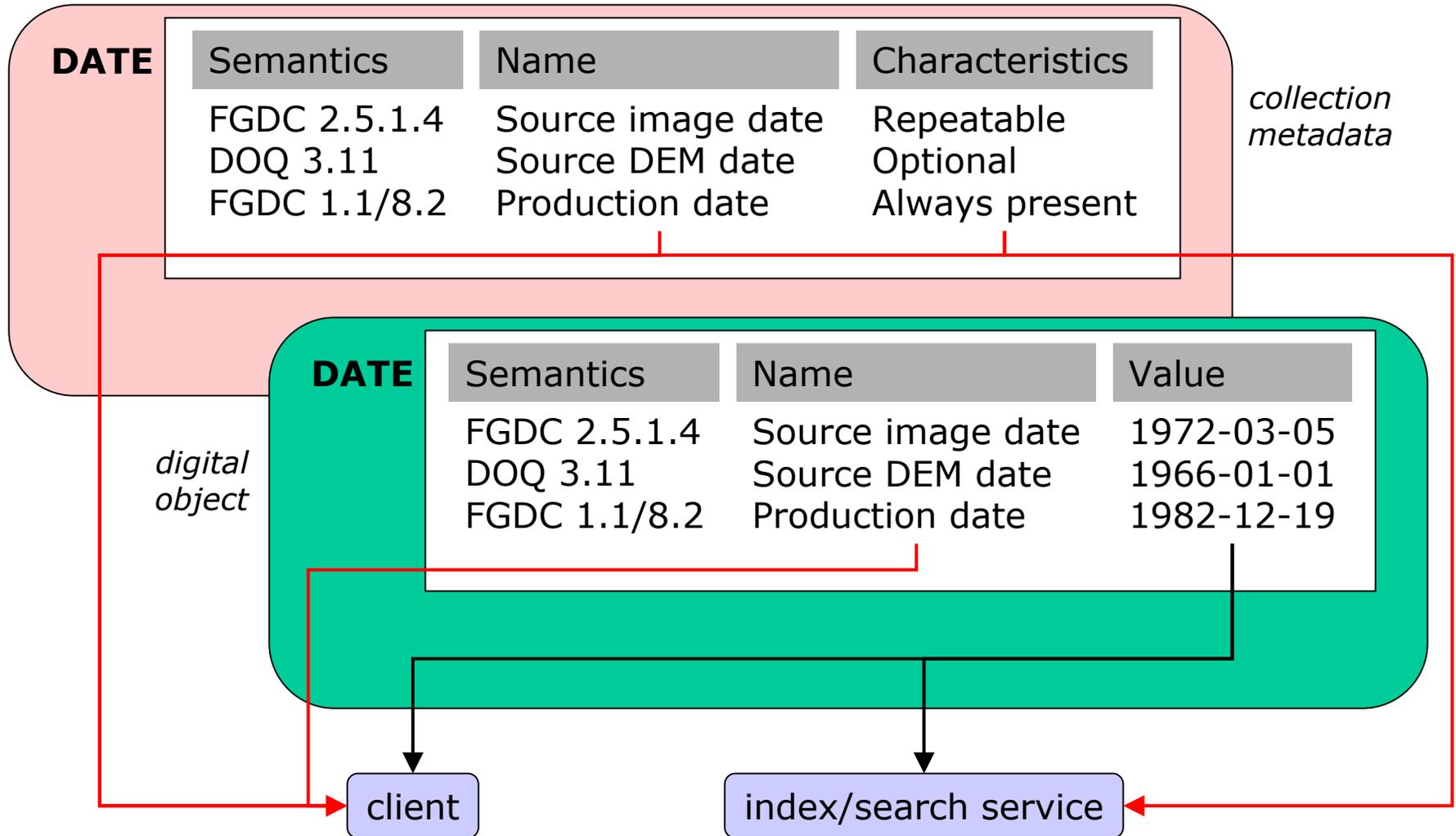


ADL-Basic fields & subfields

- Originator
- Subject-related text
 - Title
 - Assigned terms
- Type
- Format
- Spatial domain
- Date
 - Time period of content
- Identifier



ADL-Basic general structure





ADL-Basic field definition

Type

- **Semantics** identifies the nature, genre, meaning, or intellectual content of the item
- **Content** zero or more terms drawn from identified thesauri
- **XML representation** `<!ELEMENT ...>`
- **Collection metadata implications** collection metadata lists all referenced thesauri
- **Query value** single term from an identified thesaurus
- **Query operator** "is a"

```
{("Object Types", "aerial photograph"), ("Geology Concepts", "erosion")}
```

```
Type is a ("Object Types", "image")
```



ADL-Full

- Encodes the full, native metadata in a standard syntactic representation
 - RDF
 - ADL's generic encoding

```
<source prefix="MARC">http://lcweb.loc.gov/marc</source>
...
<group name="Data quality">
  <field name="Accuracy" source="MARC:514g">
    <value type="number" unit="meters">25</value>
  </field>
  <field name="Contour interval" source="MIL-B:06200a">
    <value type="number" unit="meters">10</value>
  </field>
</group>
```

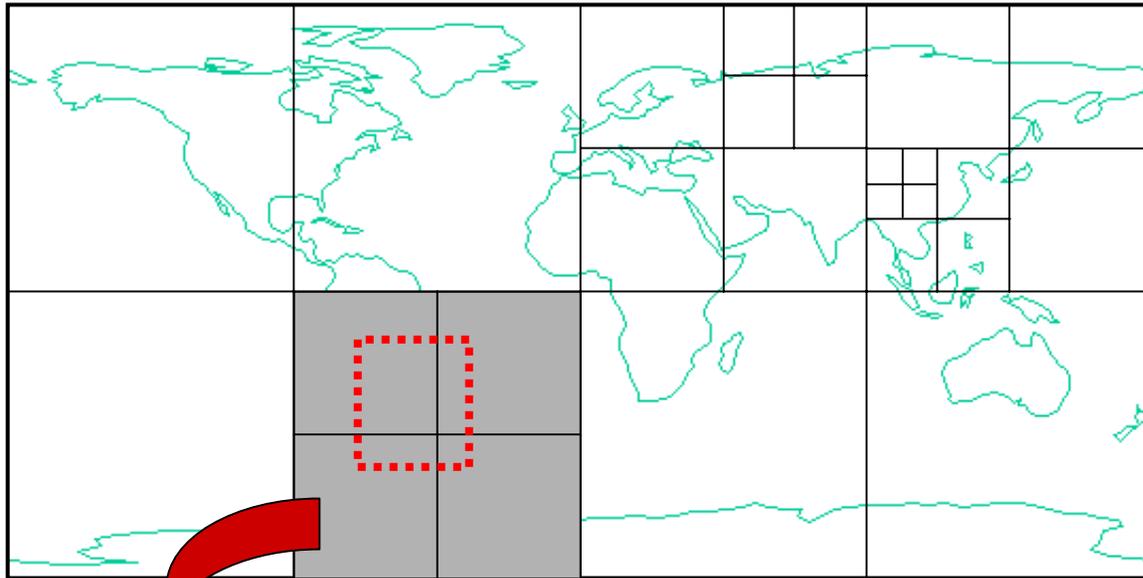


Index/search services

- Core index/search service
 - Based on *ADL-Basic*
 - Boolean combinations of constraints on *ADL-Basic* fields and subfields (only)
- Support for other, more refined search services
 - Utilizing *ADL-Basic* metadata mappings
 - Based on *ADL-Full*



Browse/aggregation service



- ❑ Aggregate statistics
- ❑ Digital objects (scale-dependent)
- ❑ By type, format, and date



Collection metadata

- Static
 - Scope and purpose, maintaining agency, etc.
- Derived
 - Referenced thesauri
 - Referenced metadata standards
 - Native → *ADL-Basic* metadata mappings
- Statistical summarization via browse/aggregation service



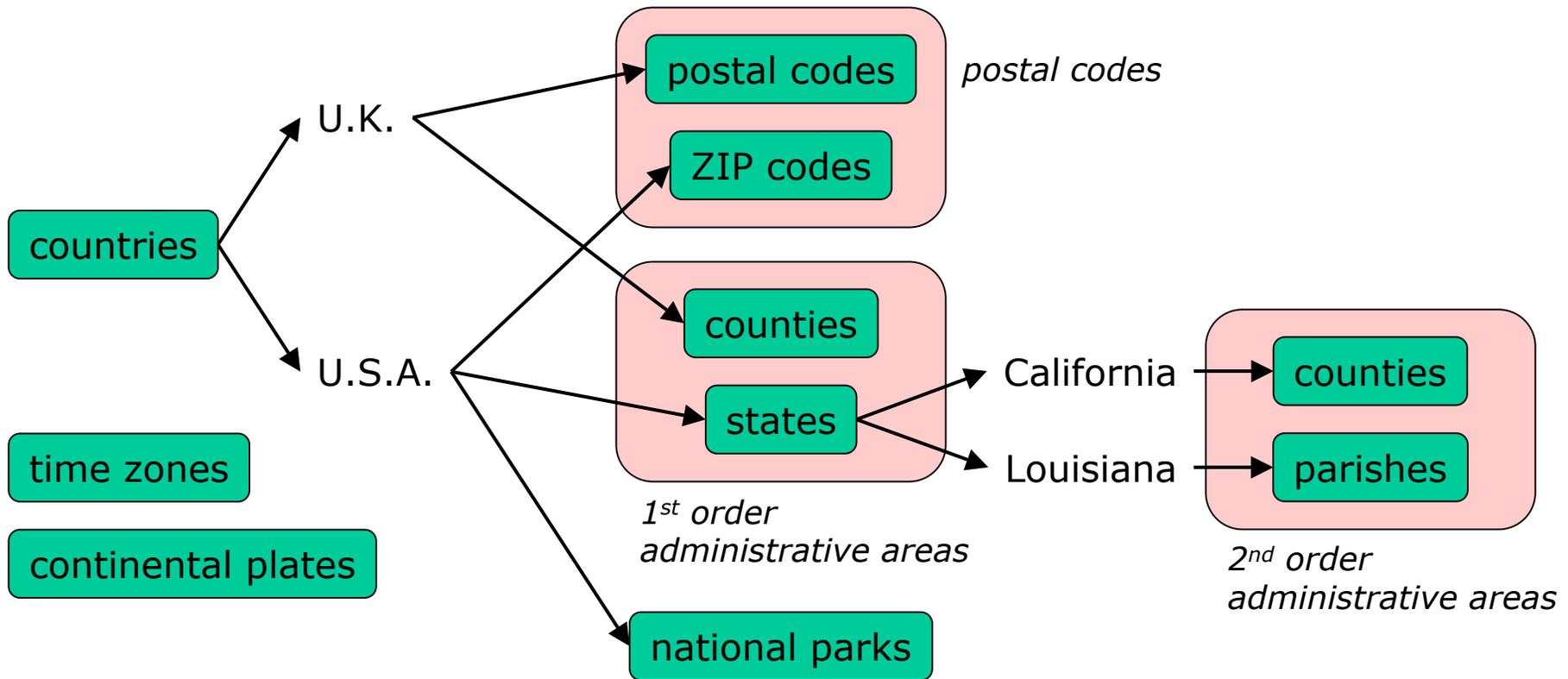
Gazetteer service (1/3)

- Stratford-upon-Avon
 - **Variant names** Stratford upon Avon; Stratford
 - **Location** N 52° 11', W 1° 42'
 - **Feature type** populated place (NIMA); town (local)
 - **Time period** 1196–present
- 805
 - **Feature type** U.S. telephone area code
- 37T
 - **Feature type** UTM zone
- Mississippi
 - **Feature type** drainage basin



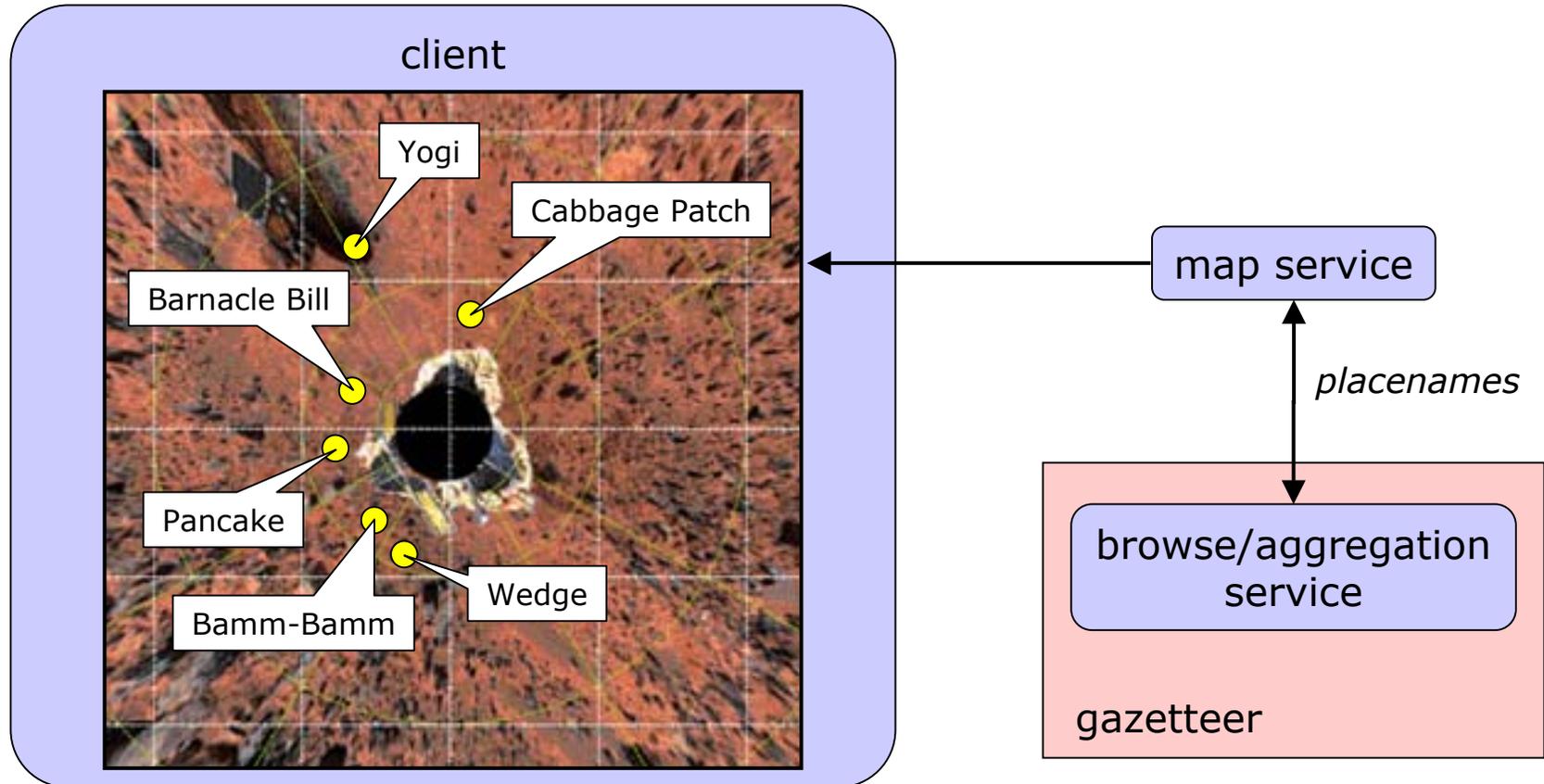
Gazetteer service (2/3)

- Geographic namespace: spatial partition of a region into uniquely named subregions





Gazetteer service (3/3)





Goals, revisited

□ Find

- Discovery service based on rich collection metadata
- Uniform searching based on *ADL-Basic*
- More refined searching based on *ADL-Basic* and *ADL-Full*
- Browse/aggregation service

□ Assemble, structure, publish

- Uniform use of collections
- Thesauri and inheritance of collections and digital objects support customization & structure
- Collections designed to span gamut big → small

□ Use

- FEDORA-like extensible digital object model