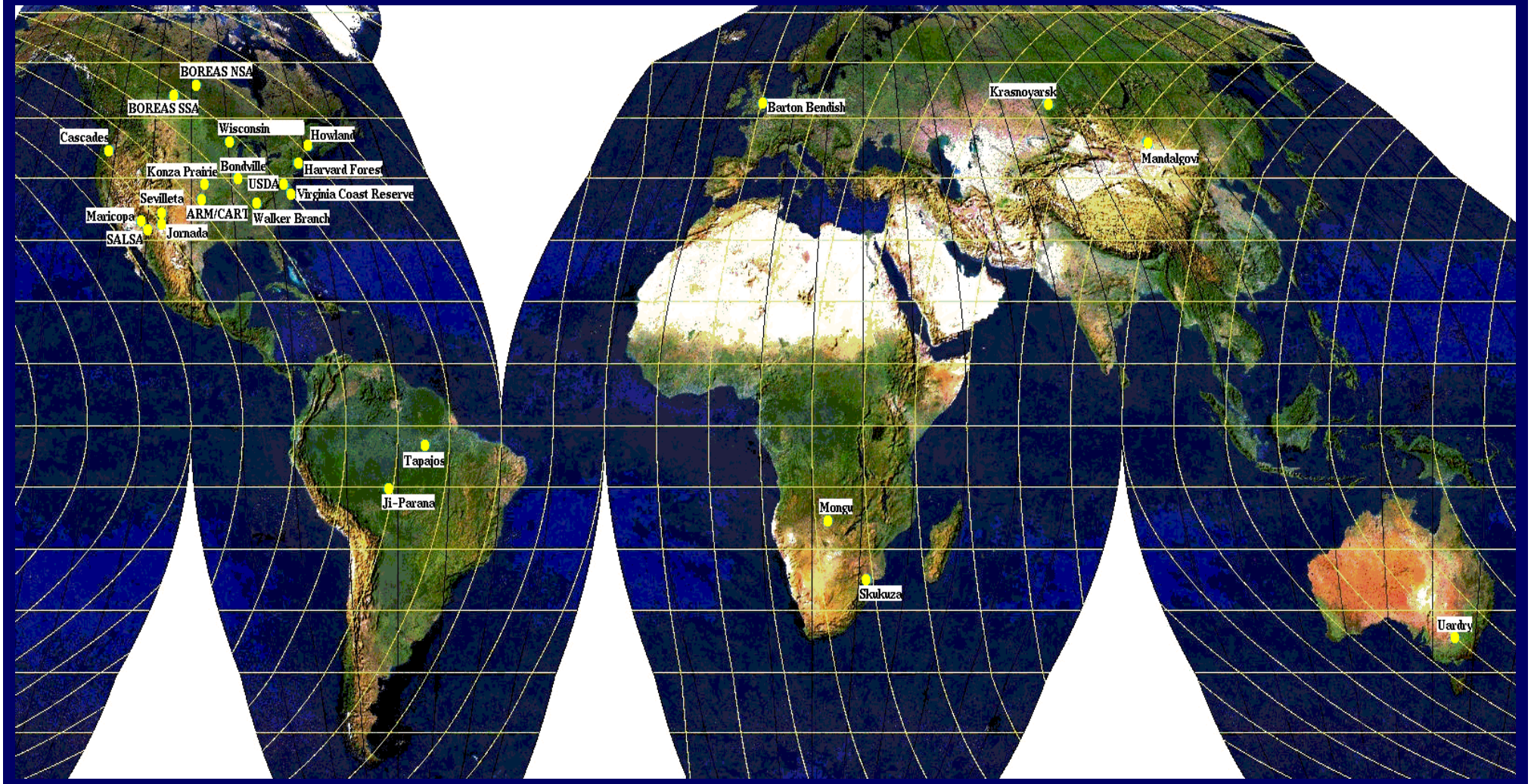
Data Validation User Interface (DVUI) on ECHO

Beth Weinstein

Bob Harberts

June 5, 2002

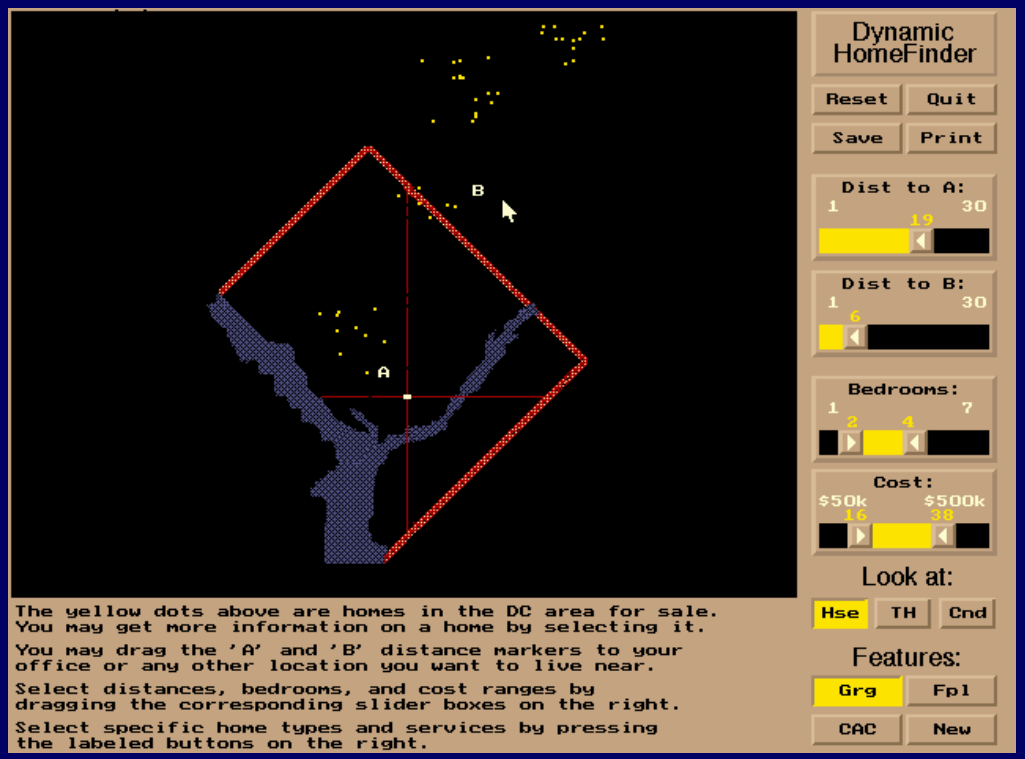
MODIS Land Validation

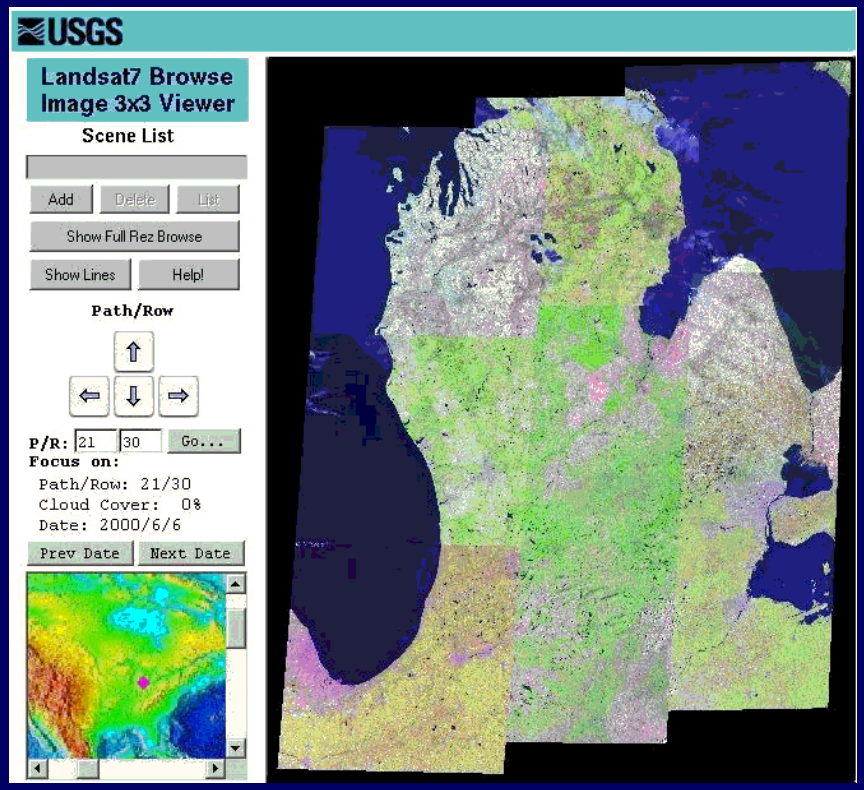


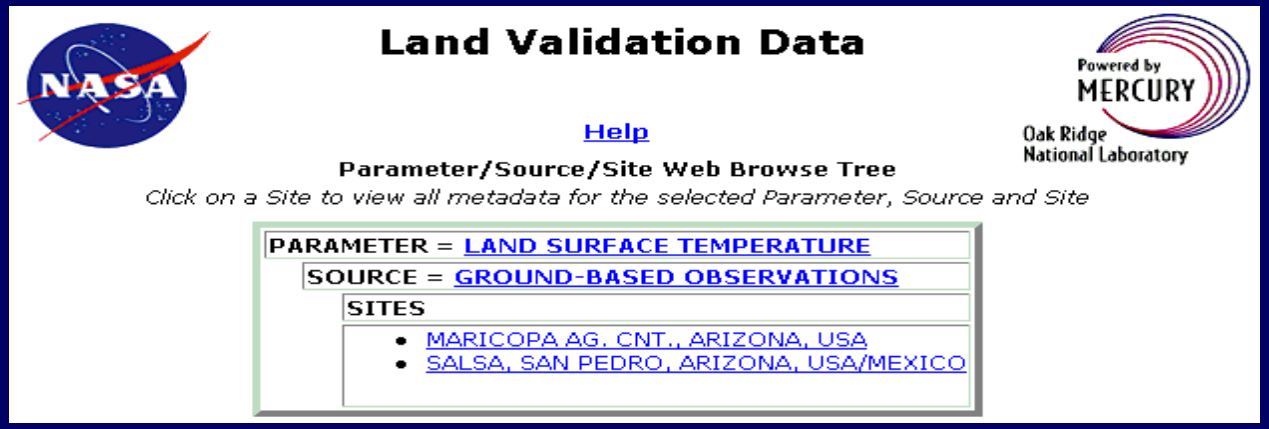
http://modarch.gsfc.nasa.gov/MODIS/LAND/VAL/

Search and Retrieval Data Access Paradigm

Navigation/Discovery Data Access Paradigm







http://journals.ecs.soton.ac.uk/~lac/ht97/pdfs/cunliffe.pdf, http://flp.cs.tu-berlin.de/~okp/sketch.html

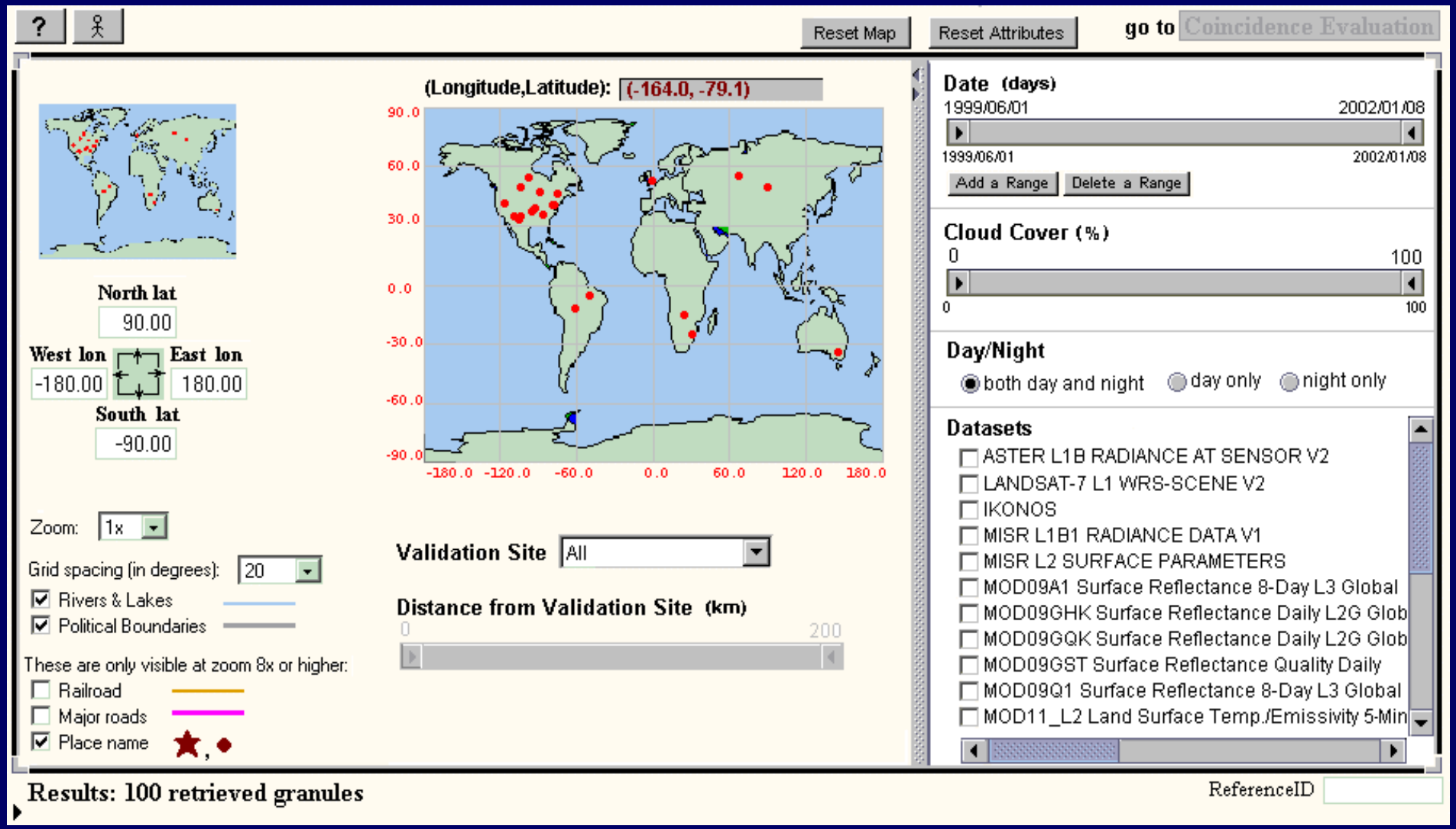
Data Validation UI (DVUI)

- Navigation/Discovery Data Access Paradigm
- Dynamic Query User Interface (The query and results are represented visually, and users have dynamic, continuous, reversible control of the query.)

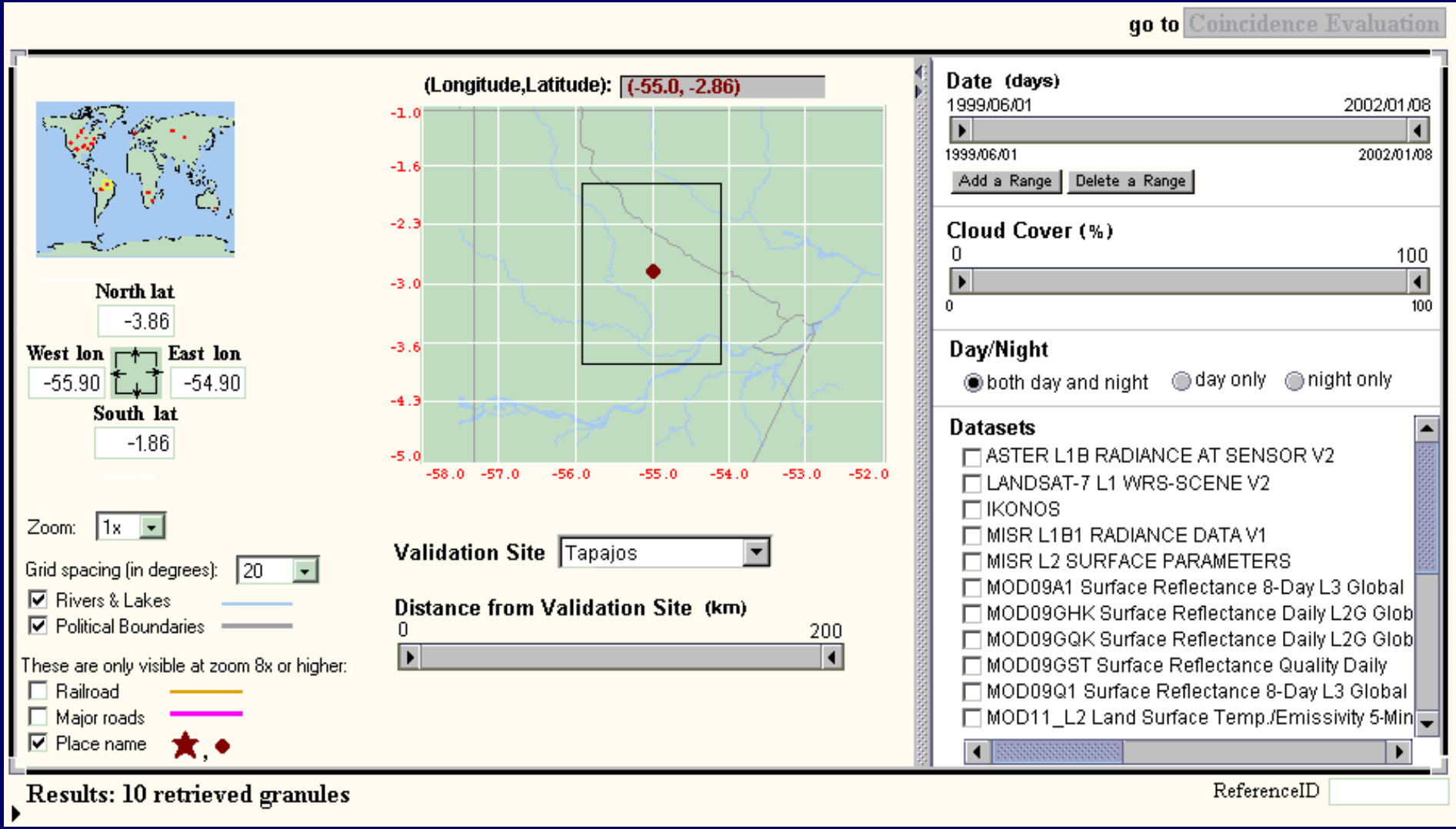
	System Approaches			Earth Science Issues		
	Preprocess	Hybrid	Personalized Information Space	Data Relationships	Resolution	Spatial Accuracy
Previous Earth Science UIs	X	X	X			
DVUI	X		X	X		X

ftp://ftp.cs.umd.edu/pub/hcil/Reports-Abstracts-Bibliography/97-02.html

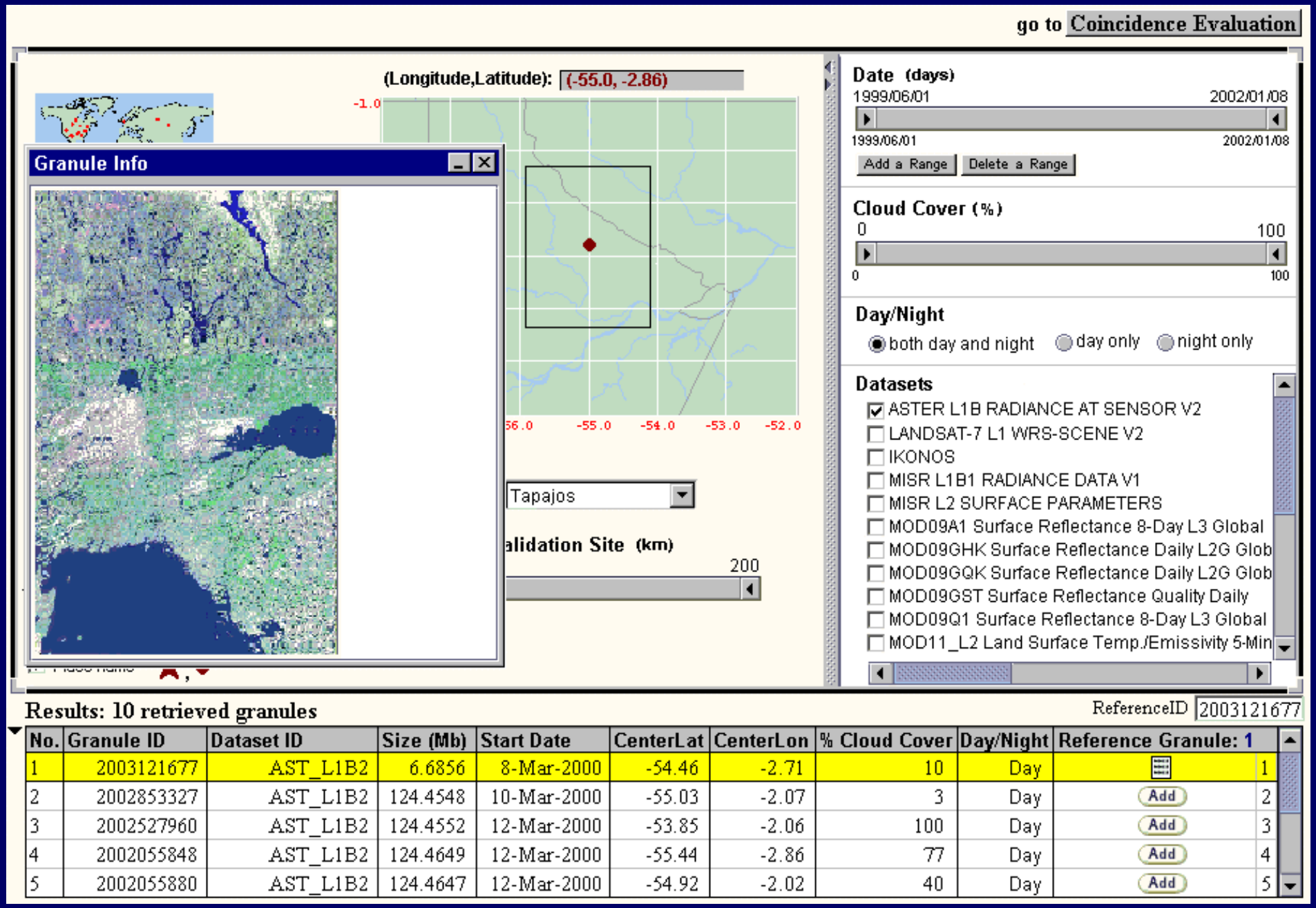
Reference Granule Search



Choose a Validation Site

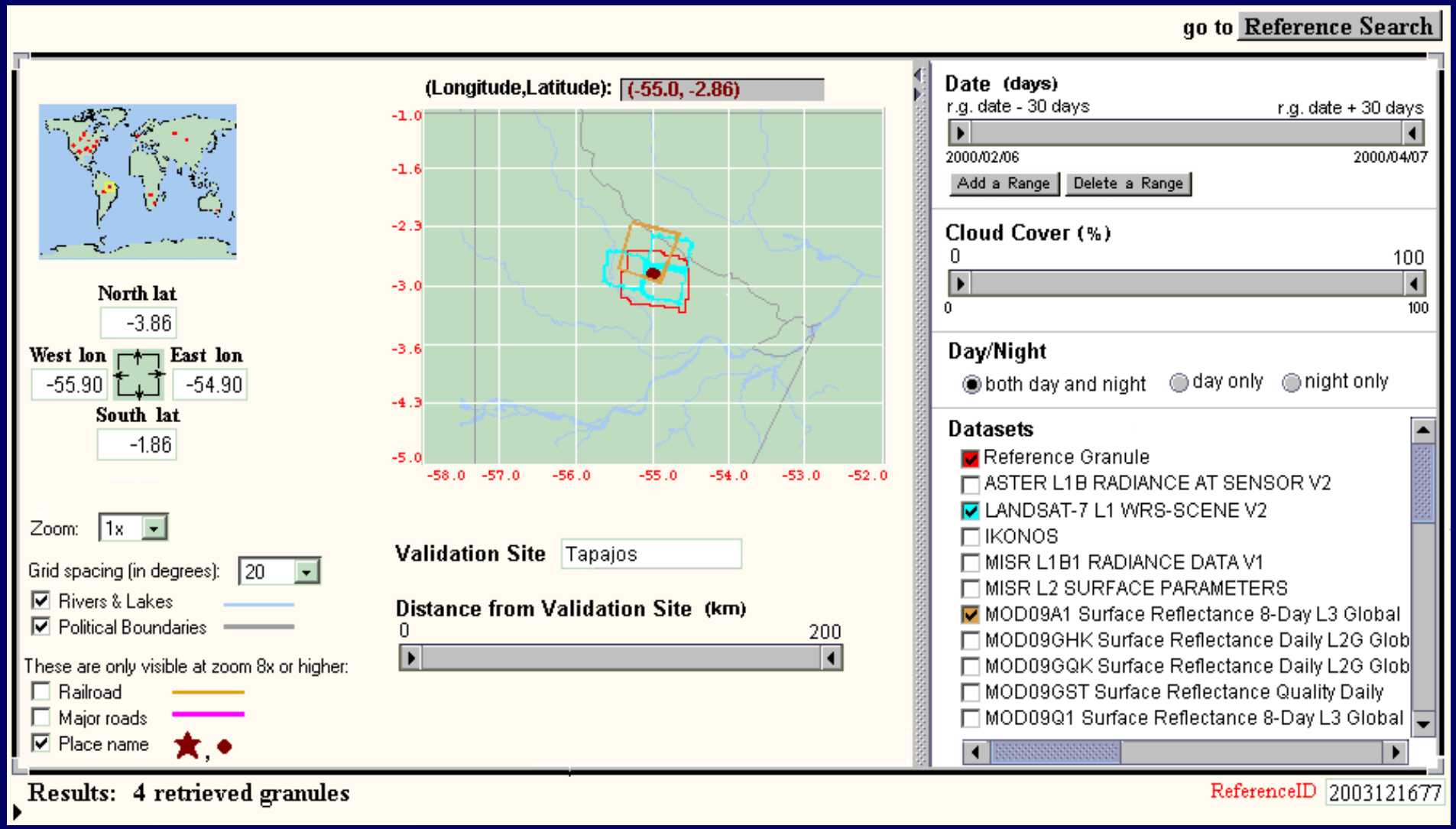


Display Results and Browse

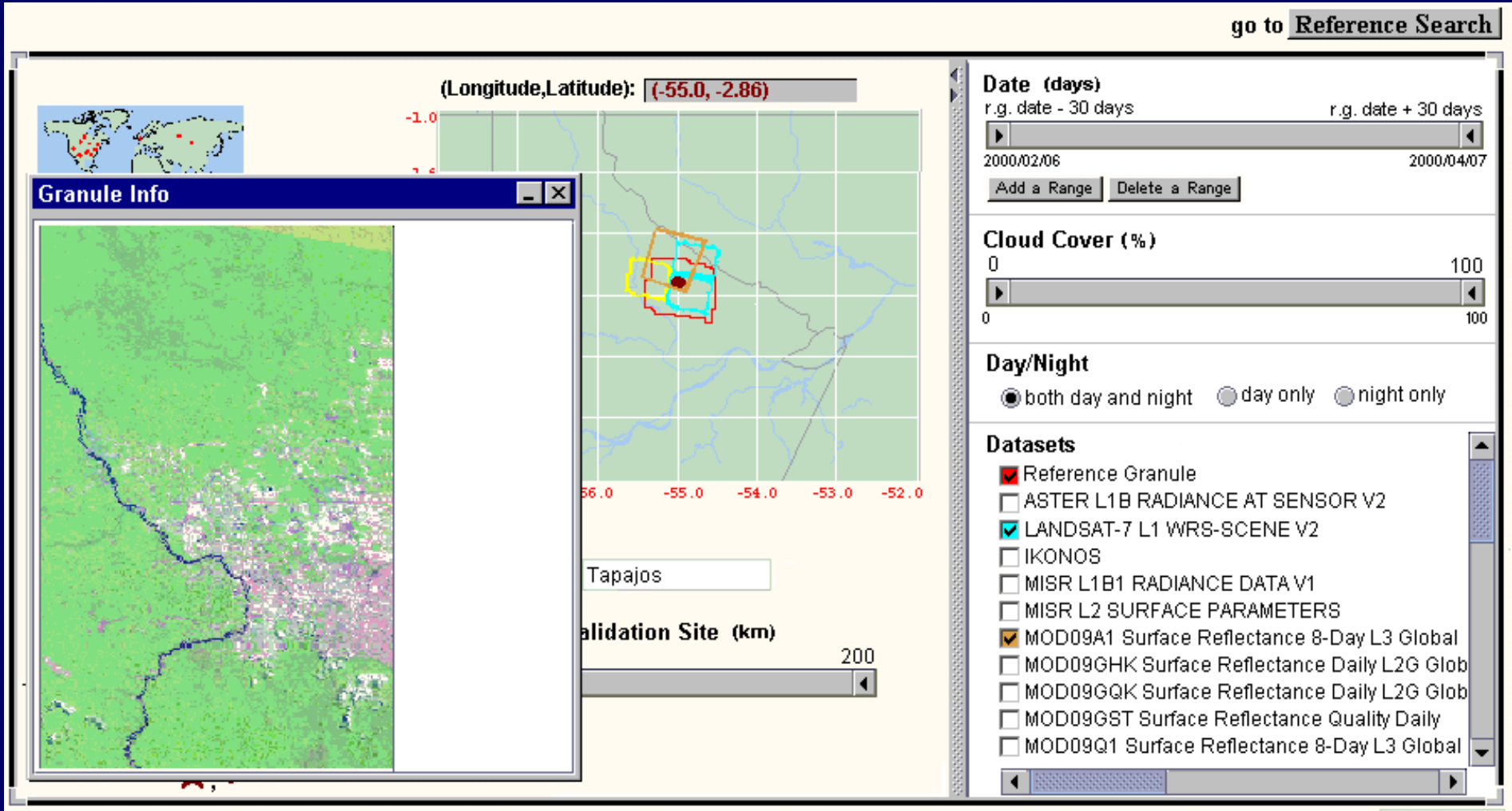


No.	Granule ID	Dataset ID	Size (Mb)	Start Date	CenterLat	CenterLon	% Cloud Cover	Day/Night	Reference Granule: 1	
1	2003121677	AST_L1B2	6.6856	8-Mar-2000	-54.46	-2.71	10	Day		1
2	2002853327	AST_L1B2	124.4548	10-Mar-2000	-55.03	-2.07	3	Day	Add	2
3	2002527960	AST_L1B2	124.4552	12-Mar-2000	-53.85	-2.06	100	Day	Add	3
4	2002055848	AST_L1B2	124.4649	12-Mar-2000	-55.44	-2.86	77	Day	Add	4
5	2002055880	AST_L1B2	124.4647	12-Mar-2000	-54.92	-2.02	40	Day	Add	5

Coincident Evaluation



Coincident Results



Results: 4 retrieved granules

ReferenceID 2003121677

No.	Granule ID	Dataset ID	Size (Mb)	Start Date	CenterLat	CenterLon	% Cloud Cover	Day/Night	Shopping Cart: 2	
1	2003121677	AST_L1B2	6.6856	8-Mar-2000	-54.46	-2.71	10	Day	x 2 Add	1
2	2003121533	L7_L12	124.6856	8-Mar-2000	-55.16	-2.76	3	Day	Add	2
3	2002853327	L7_L12	124.4548	10-Mar-2000	-55.03	3.07	100	Day	Add	3
4	2002527960	L7_L12	124.4552	12-Mar-2000	-55.25	-2.76	77	Day	Add	4
5	2002055848	MODIS_L3	124.4649	12-Mar-2000	-55.00	-2.86	40	Day	Add	5

Datasets

ASDC DAAC

- MISR Level 1B1 Radiance Data V001
- MISR Level 2 Surface Parameters

Stennis (Data Buy data)

- IKONOS Multispectral 4m
- IKONOS Panchromatic 1m
- IKONOS Stereo
- IKONOS DEM

LP DAAC

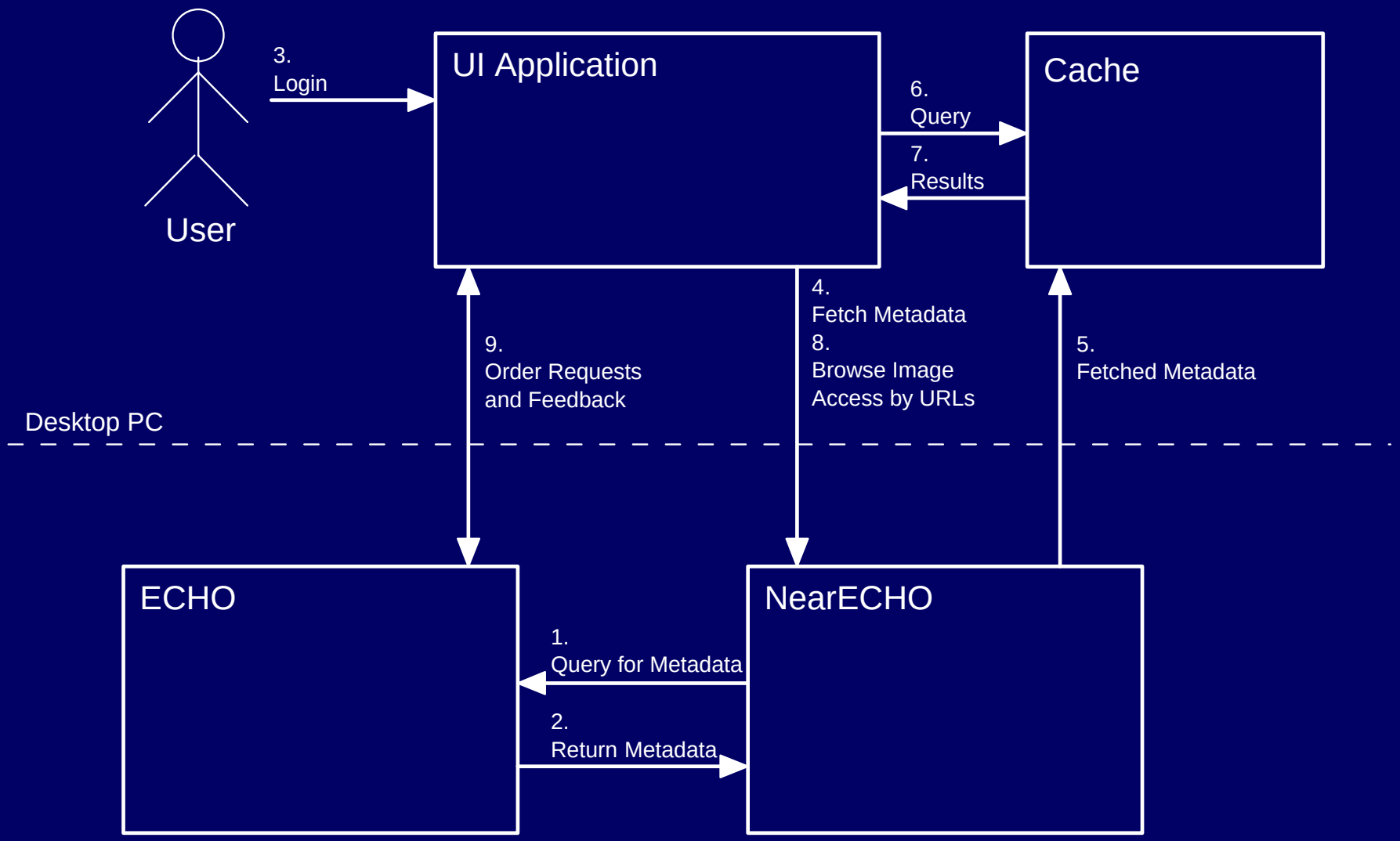
- ASTER Level 1B Data Set Registered Radiance at the Sensor V2
- LANDSAT-7 Level-1 WRS-SCENE V2
- MOD09A1 Surface Reflectance 8-Day L3 Global 500m Grid V3
- MOD09GHK Sur. Reflect. Daily L2G Global 500m Grid V3
- MOD09GQK Sur. Reflect. Daily L2G Global 250m Grid V3
- MOD09GST Sur. Reflect. Quality Daily L2G Global 1km Grid V3
- MOD09Q1 Sur. Reflect. 8-Day L3 Global 250m Grid V3
- MOD11_L2 Land Surface Temperature/Emissivity 5-Min L2 Swath 1km V3
- MOD11A1 Land Sur. Temp./Emiss. Daily L3 Global 1km Grid V3
- MOD11A2 Land Sur. Temp./Emiss. 8-Day L3 Global 1km Grid V3
- MOD11B1 Land Sur. Temp./Emiss. Daily L3 Global 5km Grid V3
- MOD12Q1 Land Cover Type 96-Day L3 Global 1km Grid V3
- MOD13A1 Vegetation Indices 16-Day L3 Global 500m Grid V3
- MOD13A2 Vegetation Indices 16-Day L3 Global 1km Grid V3
- MOD13Q1 Vegetation Indices 16-Day L3 Global 250m Grid V3
- MOD14 Thermal Anomalies/Fire 5-Min L2 Swath 1km V3
- MOD14A1 Th. Anom./Fire Daily L3 Global 1km Grid V3
- MOD14A2 Th. Anom./Fire 8-Day L3 Global 1km Grid V3
- MOD14GD Th. Anom./Fire Daily L2G Global 1km Grid Day V3
- MOD14GN Th. Anom./Fire Daily L2G Global 1km Grid Night V3
- MOD15A2 Leaf Area Index/FPAR 8-Day L4 Global 1km Grid V3
- MOD17A2 Net Photosynthesis 8-Day L4 Global 1km Grid V3
- MOD43B1 BRDF/Albedo Model-1 16-Day L3 Global 1km Grid V3
- MOD43B3 Albedo 16-Day L3 Global 1km Grid V3
- MOD43B4 Nadir BRDF-Adjusted Reflectance 16-Day L3 Global 1km Grid V3
- MODPT1KD Observation Pointers Daily L2G Global 1km Grid Day V3
- MODPT1KN Obs. Pointers Daily L2G Global 1km Grid Night V3
- MODPTHKM Obs. Pointers Daily L2G Global 500m Grid V3
- MODPTQKM Obs. Pointers Daily L2G Global 250m Grid V3
- MODMGGAD Geolocation Angles Daily L2G Global 1km Grid Day V3
- MODMGGAN Geo. Angles Daily L2G Global 1km Grid Night V3

Size of Information Space

- 31 Validation Sites
- 37 Datasets
- 1,000: average # of granules per dataset per site
- 37,000: average # of granules at a site
- 1,147,000: total number of granules
- 200 bytes: size of one granule's metadata (25 attributes)
- 7 MB: amount of metadata per site
- 220 MB: total amount of metadata

All numbers are approximated.

Context Diagram



Near-ECHO Study

- Query Preview was a study to support preprocessed metadata to tie spatial and temporal metadata into dependent valids. QP was deemed not feasible due to the distributed architecture
- ECHO offers an architecture that will support efficient preprocessing of metadata
- Near-ECHO, an extension of the Query Preview, is a study to determine feasibility and requirements for supporting preprocessed metadata
- Experience gained supporting DVUI benefits both ECHO development and science application developers

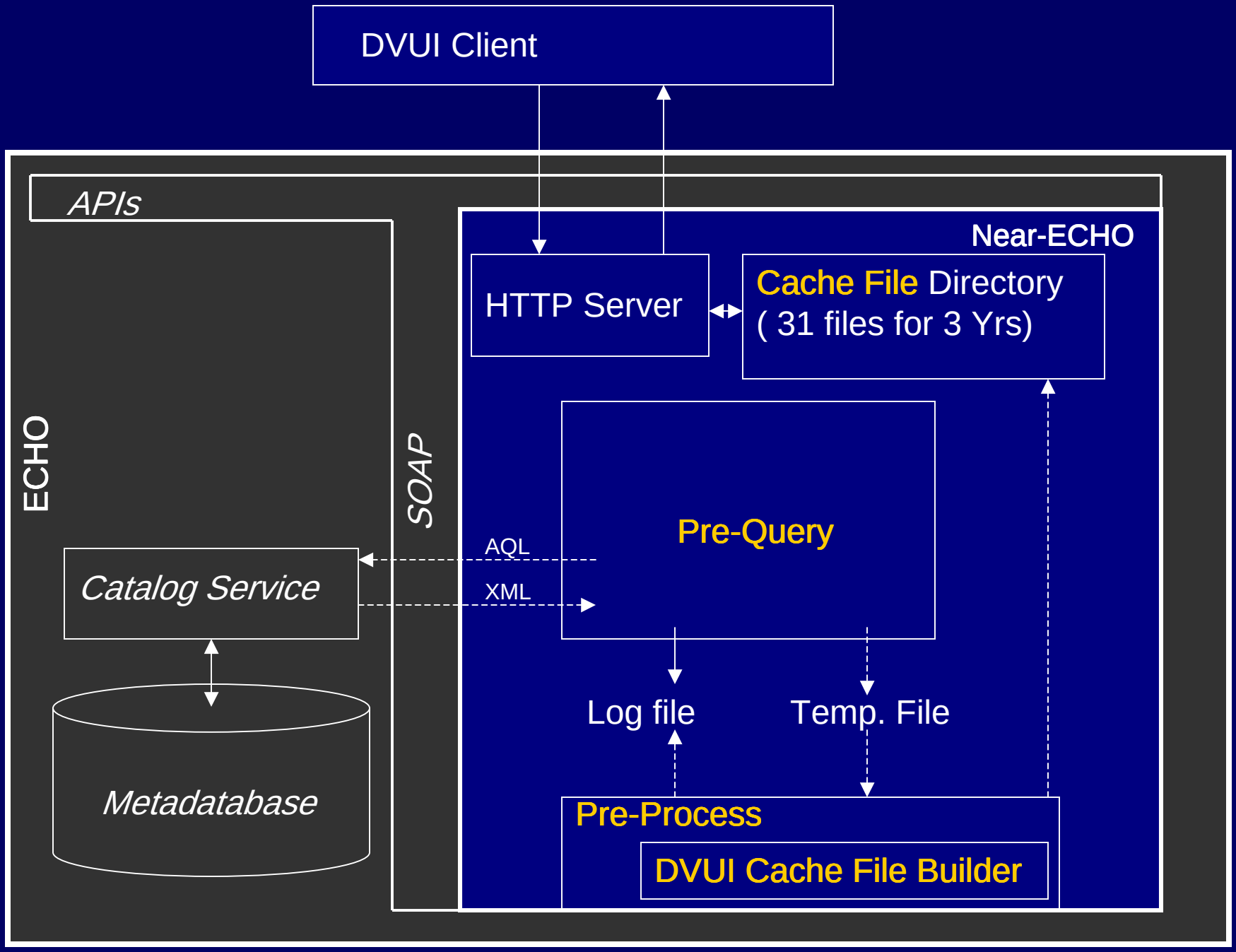
Required DVUI Collection Metadata Attributes

- Dataset ID
- Dataset short name
- Dataset long name
- Package Info
  - Data format
  - Media type/format

Required DVUI Granule Metadata Attributes

- Granule ID (=GranuleUR)
- Granule Size
- Beginning/Ending date and time
- Geographic extent – (spatial foot-print geometry coordinates)
- *Center lat/lon** (derived)
- Browse URL (in JPEG format)
- Cloud Cover (handling convention)
- *Day-Night flag* (derived for ASTER)
- *Distance between granule and validation point.** (derived)

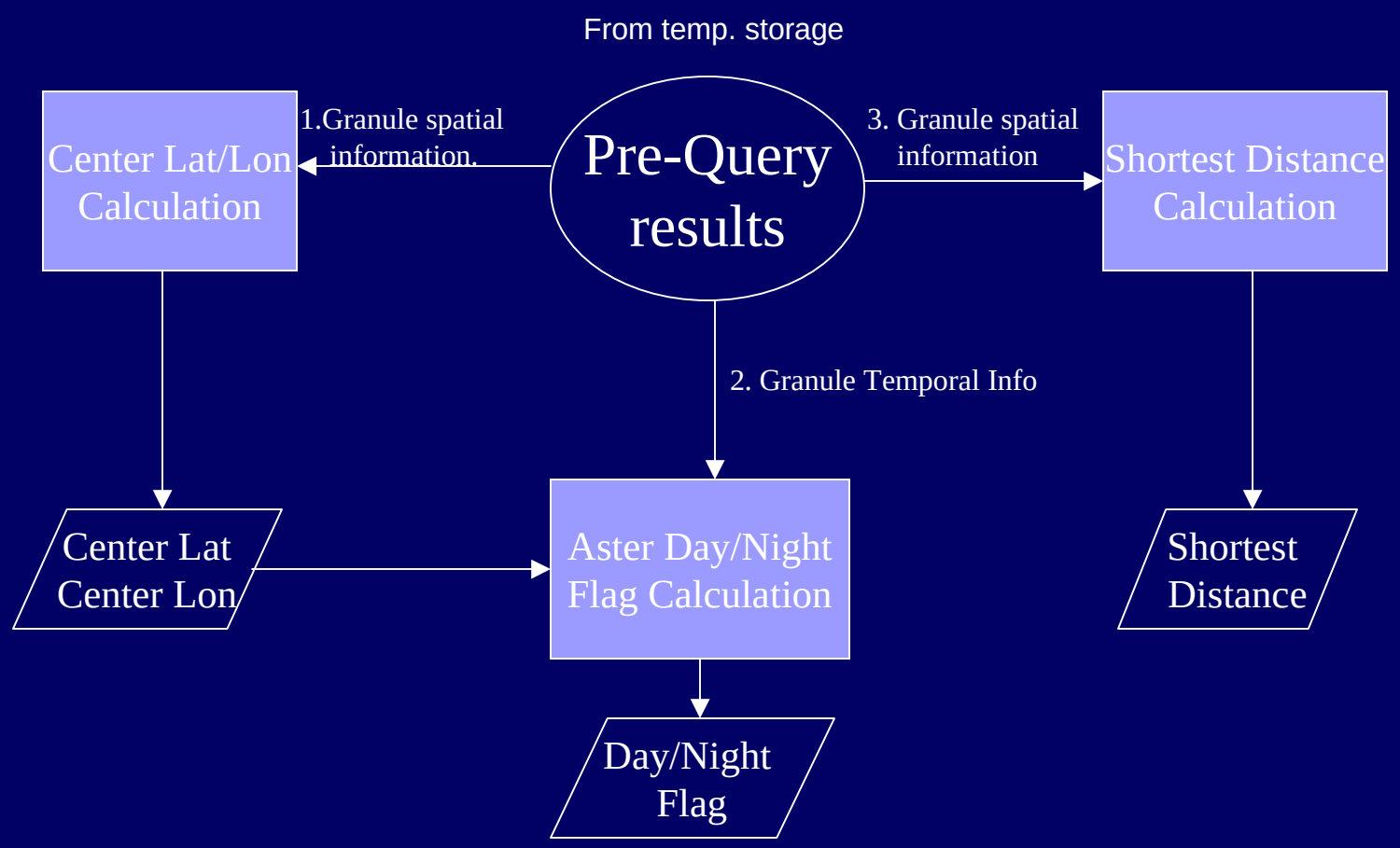
* center lat/lon and distance are not currently available in ECHO, N-ECHO pre-process will calculate them


DVUI Client
APIs
ECHO
SOAP
Near-ECHO
HTTP Server
Cache File Directory
(31 files for 3 Yrs)
Pre-Query
AQL
XML
Catalog Service
Metadatabase
Log file
Temp. File
Pre-Process
DVUI Cache File Builder

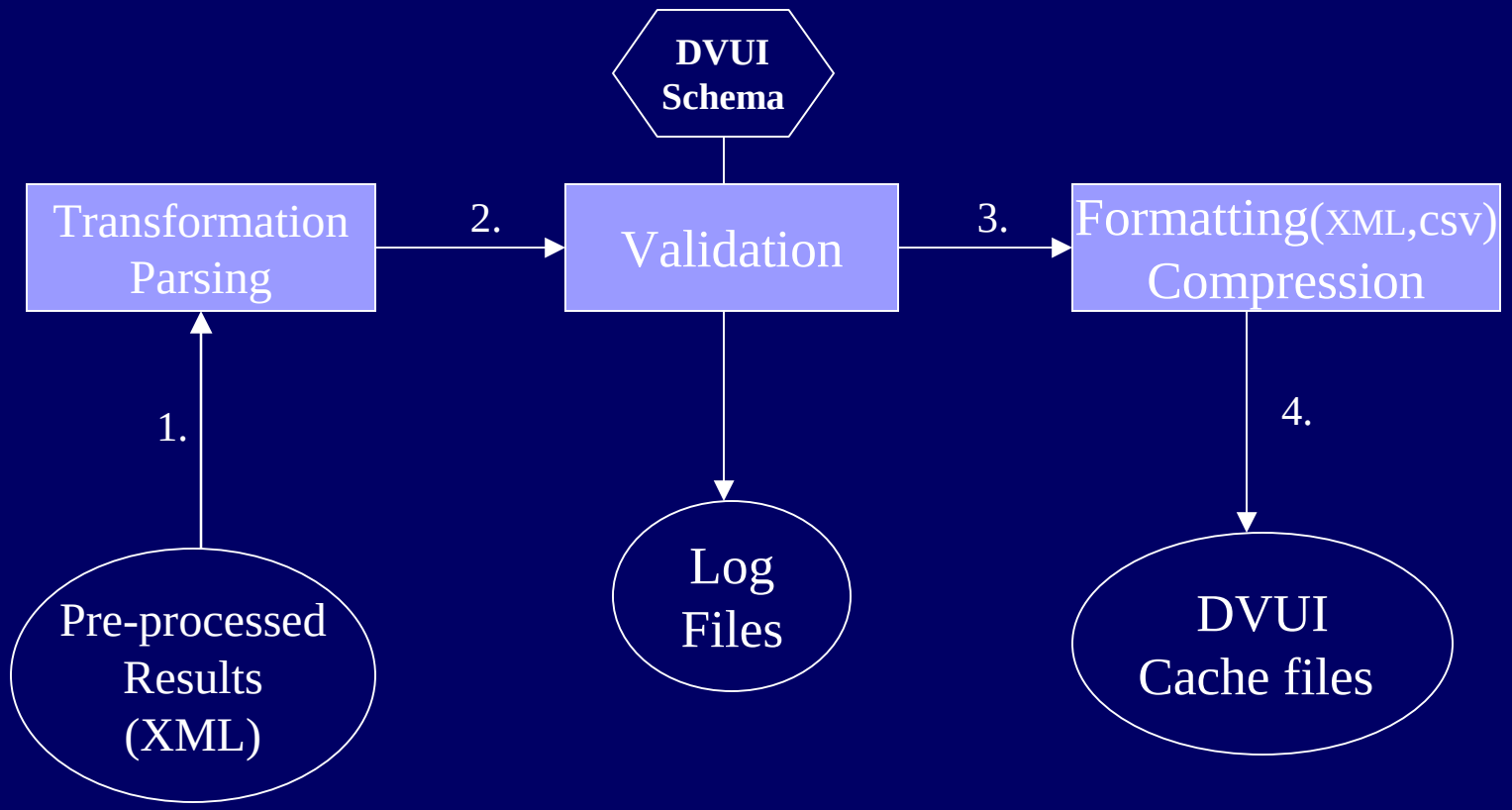
Pre-Query Activity (ECHO-based)

- Pre-Query to ECHO involves three sets of queries:
  - Collection Level Metadata: one query per validation site
  - Collection level Packaging info: one query per dataset using the ECHO's order entry process
  - Granule Level Metadata: one query per validation site

DVUI Pre-Processing Context Diagram



DVUI Cache Builder Context Diagram



Operational Mechanisms

- Two ways to get metadata for the DVUI:
  - ECHO (XML) – preferable but isn’t currently populated with the required metadata
  - EDG (ODL) and Direct Archive – only feasible for development and testing

EDG Operational Experience

- Metadata for 31 sites and 37 datasets for 3 years could potentially result in 1 million granules
- EDG Client and DAAC servers have constraints on executable queries (i.e. 1 hour timeout, 1000 granule result limit)
- Queries must be broken down to fit EDG and DAAC constraints for automated processing

EDG Operational Experience

- For operations, all metadata needs to be refreshed daily
- However, querying EDG for 2 validation sites will take three working days
- Therefore, it would take 45 days to populate Near-ECHO for all 31 sites
- Breaking down queries creates extra processing to integrate and aggregate query results including cache file building (includes ODL-to-XML conversion and parsing steps)

Assessment

- An ideal daily update of a comprehensive DVUI cache file is not possible because EDG does not permit “update only” queries (ECHO does)
- An operational pre-query mechanism with ECHO significantly reduces complexities of the EDG/Direct Archive approach
- Although the ECHO approach seems more feasible for the DVUI, it will not be successful unless there is metadata in ECHO

Contact Info

Beth Weinstein

beth.weinstein@gsfc.nasa.gov

Bob Harberts

harberts@gst.com