THREDDS

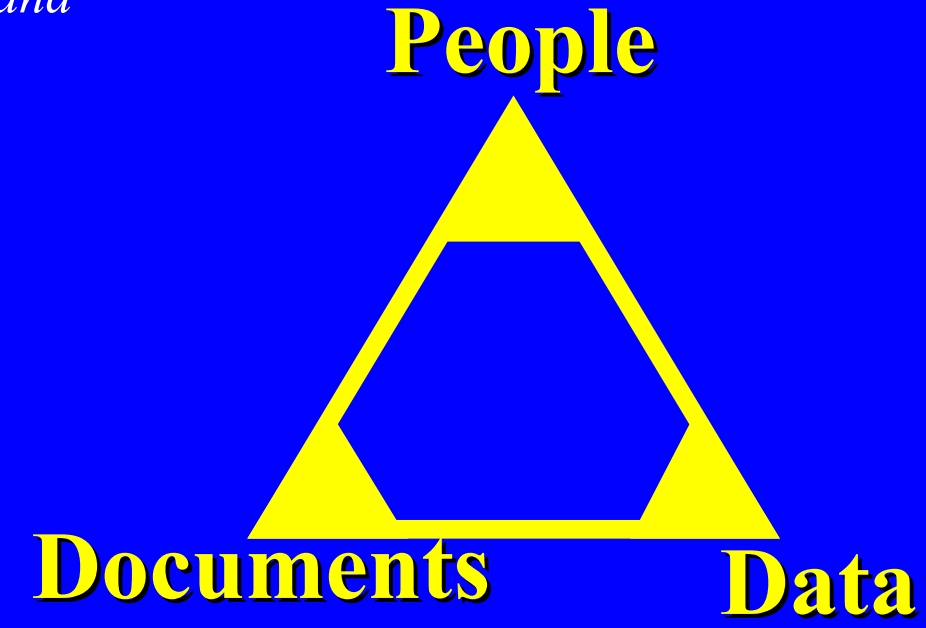
THematic Real-time Environmental Distributed Data Services

Connecting people, documents and data

Ben Domenico, John Caron, Ethan Davis, Robb Kambic, Stefano Nativi

Unidata Program Center and
University of Florence
March 2003

Sponsored by NSF



THREDDS Overview

- National Science Digital Library (NSDL) "collections" project
- Integrating real-time environmental data into
  - Online educational materials
  - Digital libraries (DLESE, NSDL)
- Two-year grant from NSF Department of Undergraduate Education (DUE)
- Led by Unidata Program Center (UPC)

THREDDS Partners: Data Providers

- **University of Alabama Huntsville** (Sara Graves, Rahul Ramachandran, Steve Tanner, Ken Keiser)
- **ARM** (Atmospheric Radiation Measurement, Chris Klaus)
- **CDC**, the Climate Diagnostic Center (Roland Schweitzer)
- **COLA**, Center for Oceans Land Atmosphere (Joe Wielgosz)
- **University of Florence** (Stefano Nativi)
- **GMU**, George Mason University (Menas Kafatos and Ruixin Yang)
- **IRI/LDEO**, International Research Institute/Lamont Doherty Earth Observatory (Benno Blumenthal)
- **ESG**, the Earth System GRID (Luca Cinquini, NCAR/SCD)
- **IRIS DMC**, Incorporated Research Institutes for Seismology Data Management Center (Rob Casey)
- **NCAR**, the National Center for Atmospheric Research (Don Middleton)
- **NCDC**, the National Climatic Data Center (Ben Watkins)
- **NGDC**, National Geophysical Data Center (Ted Habermann)
- **NOMADS,**NOAA Operational Model Archive and Distribution System, (Glenn Rutledge, NCDC)
- **University of Oklahoma** (Kelvin Droegemeier)
- **PMEL**, the Pacific Marine Environment Laboratory (Steve Hankin)
- **FNMOC**, Fleet Numerical Meteorological and Oceanographic Center (Phil Sharfstein)
- **SSEC**, the Space Science and Engineering Center., U. of Wisconsin-Madison (Steve Ackerman, Tom Whittaker)
- **Unidata Community ADDE servers** (Tom Yoksas, Unidata Program Center)

THREDDS Partners: Analysis/Display Tool Builders

- **Data Discovery Toolkit and Foundry** based on **EDMI** (Earth Data Multimedia Instrument, **New Media Studio**, Bruce Caron).
- **GDS**, GrADS/DODS Server **(COLA**, Center for Oceans Land Atmosphere, Joe Wielgosz)
- **IDV**, Integrated Data Viewer **(Unidata Program Center**, Don Murray)
- **INGRID** (**IRI/LDEO**, International Research Institute/Lamont Doherty Earth Observatory, Benno Blumenthal)
- **LAS**, Live Access Server **(PMEL**, the Pacific Marine Environment Laboratory, Steve Hankin)
- **VGEE**, Virtual Geophysical Exploration Environment (**NCAR**, DLESE, U. of Illinois, Unidata, many collaborators)
- **WXWISE Applets (SSEC**, the Space Science and Engineering Center., U. of Wisconsin-Madison, Tom Whittaker)

THREDDS Partners: Interoperability

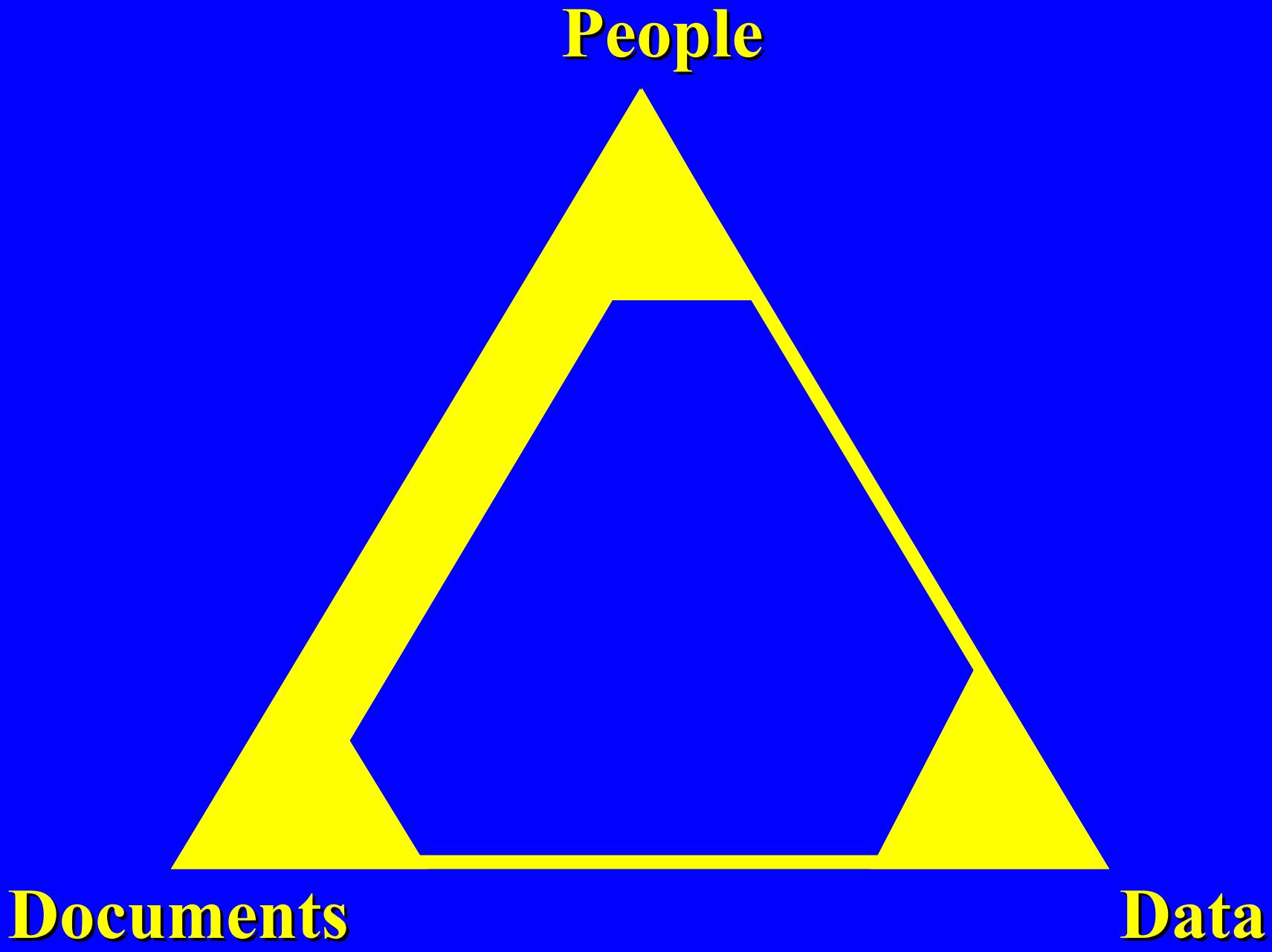
- **ADDE,** Abstract Data Distribution Environment (University of Wisconsin – Madison, Tom Yoksas)
- **DIMES**, DIstributed MEtadata System (George Mason University, Ruixin Yang)
- **DODS/OPeNDAP/Aggregation Server**, Distributed Oceanographic Data System/Open source Project for a Network Data Access Protocol (University of Rhode Island, Unidata, Ethan Davis)
- **DLESE**, Digital Library for Earth System Education (Rajul Pandya)
- **ESML**, Earth System Markup Language (University of Alabama-Huntsville, Rahul Ramachandran)
- **NCML**, netCDF Markup Language (Earth System Grid/Unidata)
- **ESRI**, Environmental Science Research Institute (various)
- **GCMD**, Global Change Master Directory (Gene Major)
- **OGC and ISO Standards** (University of Florence, Stefano Nativi)

Unidata's Contributions

- A large, (inter)national, active, cooperative user community
- Coordination of many disparate contributors (universities, government agencies, digital libraries, commercial vendors, standards bodies…)
- Reliable, automated, real-time data systems
- Platform-independent 5D visualization with HTML document integration
- Basic inventory catalog generator and server software
- Client-side catalog access modules

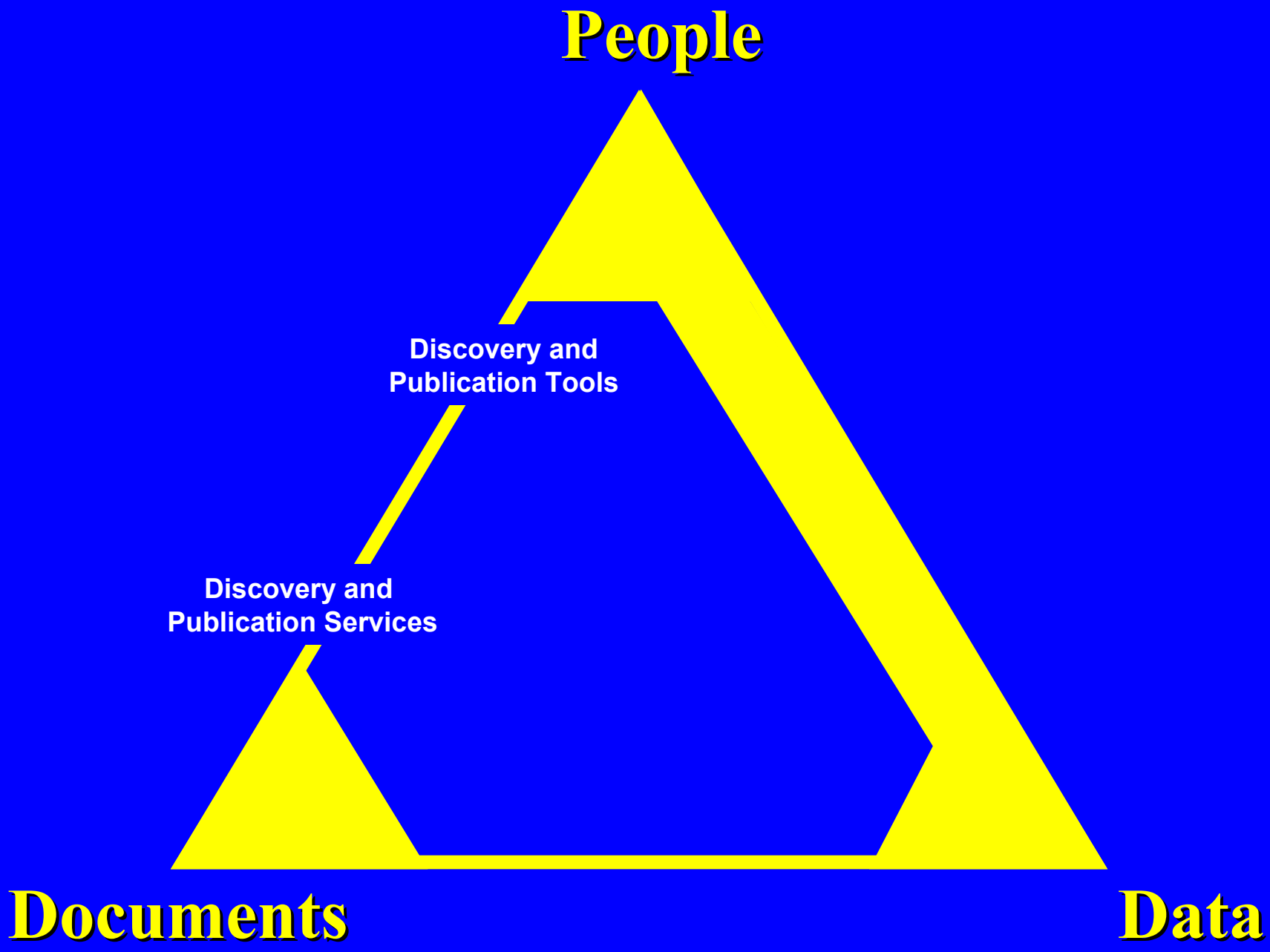
Funding Sources

- Unidata 2003/2008 (NSF Atmospheric Science Division)
- THREDDS NSDL Collections Grant (NSF Department of Undergraduate Education)
- DODS/OPeNDAP (University of Rhode Island subcontract on Naval Ocean Partnership Program Grant)
- NWS/COMET Case Studies (NOAA NWS)


People
Documents
Data

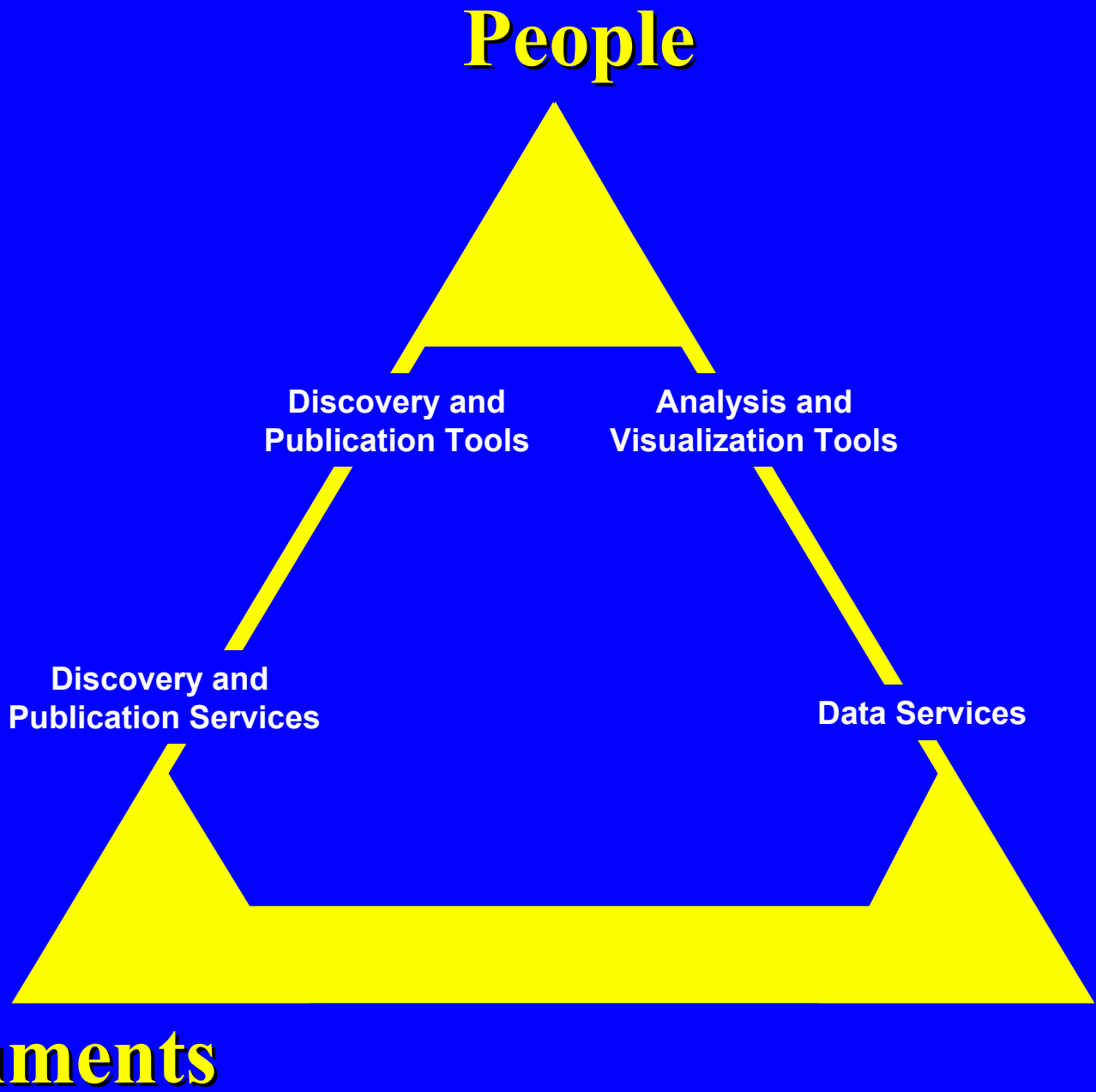
People – Documents: The Web

- Well-developed connections
  - Document references
  - Embedded multimedia
  - Embedded interactive applets
- Powerful tools
  - Google
  - Dreamweaver
  - Web-site management tools
  - Web services


People
Discovery and
Publication Tools
Discovery and
Publication Services
Documents
Data

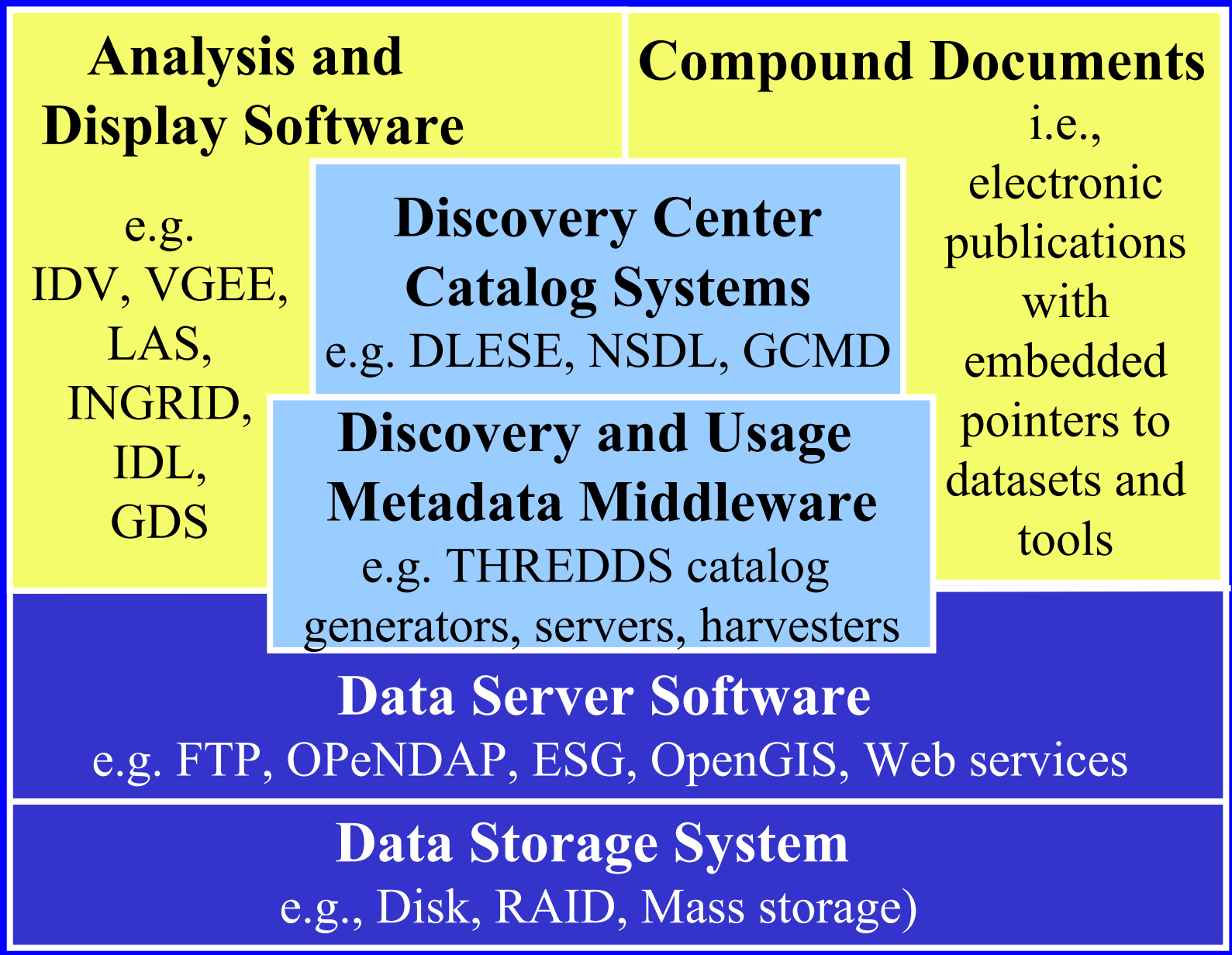
People – Data
Ad Hoc Tools/Services

- Traditional Unidata approach
  - IDD moves data to local network
  - McIDAS, GEMPAK, IDV (thick clients)
  - Most analysis work done on local client
- Web-based data interactions
  - Simple (passive) gif images
  - LAS, INGRID, GDS (thin clients)
  - Most analysis work done on remote server
- Combinations
  - Web browse/catalogs with FTP delivery/local analysis
  - Client/server (DODS/OPeNDAP, ADDE…)
  - Embedded data access applets (WXWISE)
- All lack sophisticated, text-based Web search/discovery tools and coherent integration


People
Discovery and Publication Tools
Analysis and Visualization Tools
Discovery and Publication Services
Data Services
Documents
Data

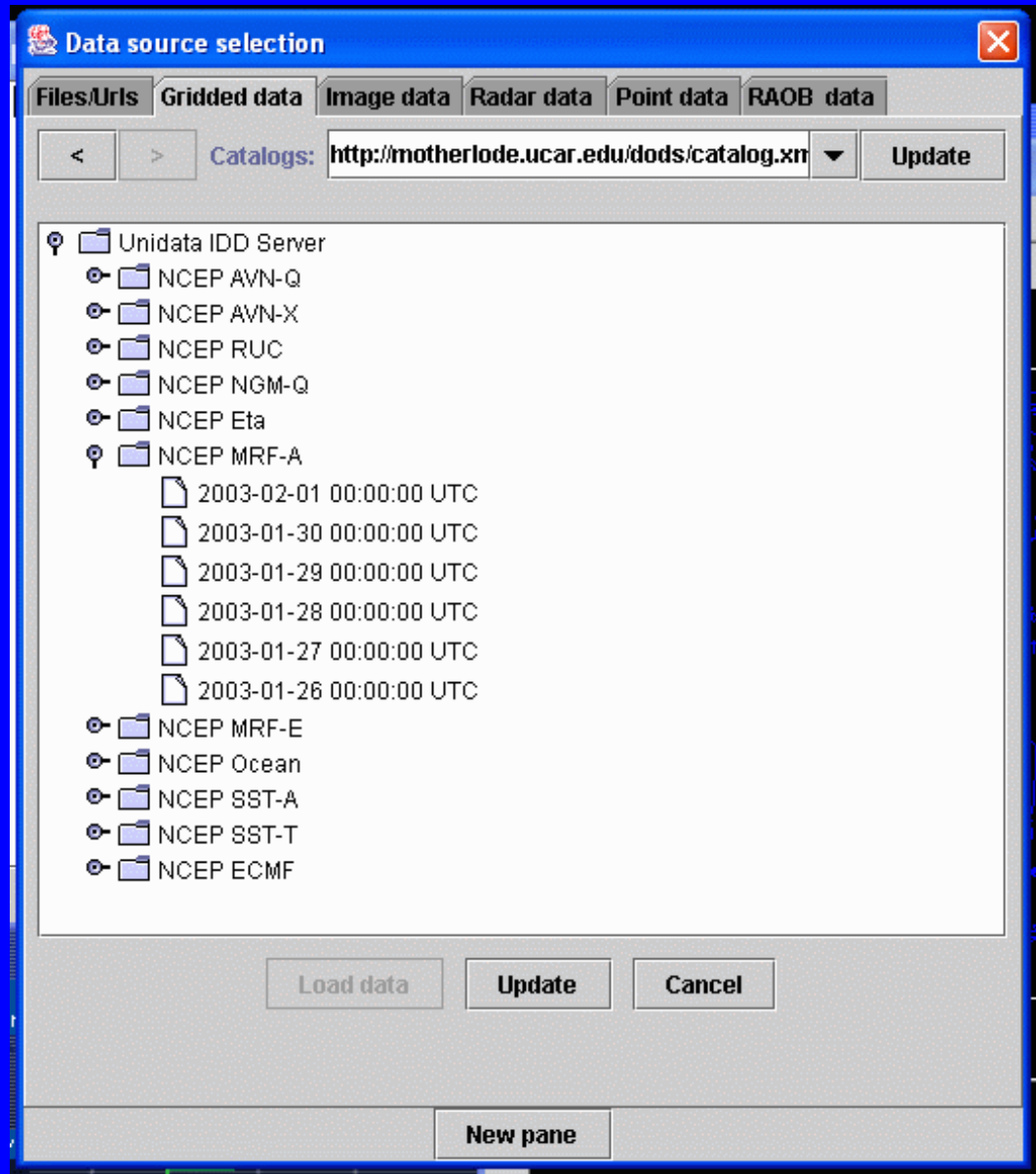
Documents – Data Connect Words and Datasets

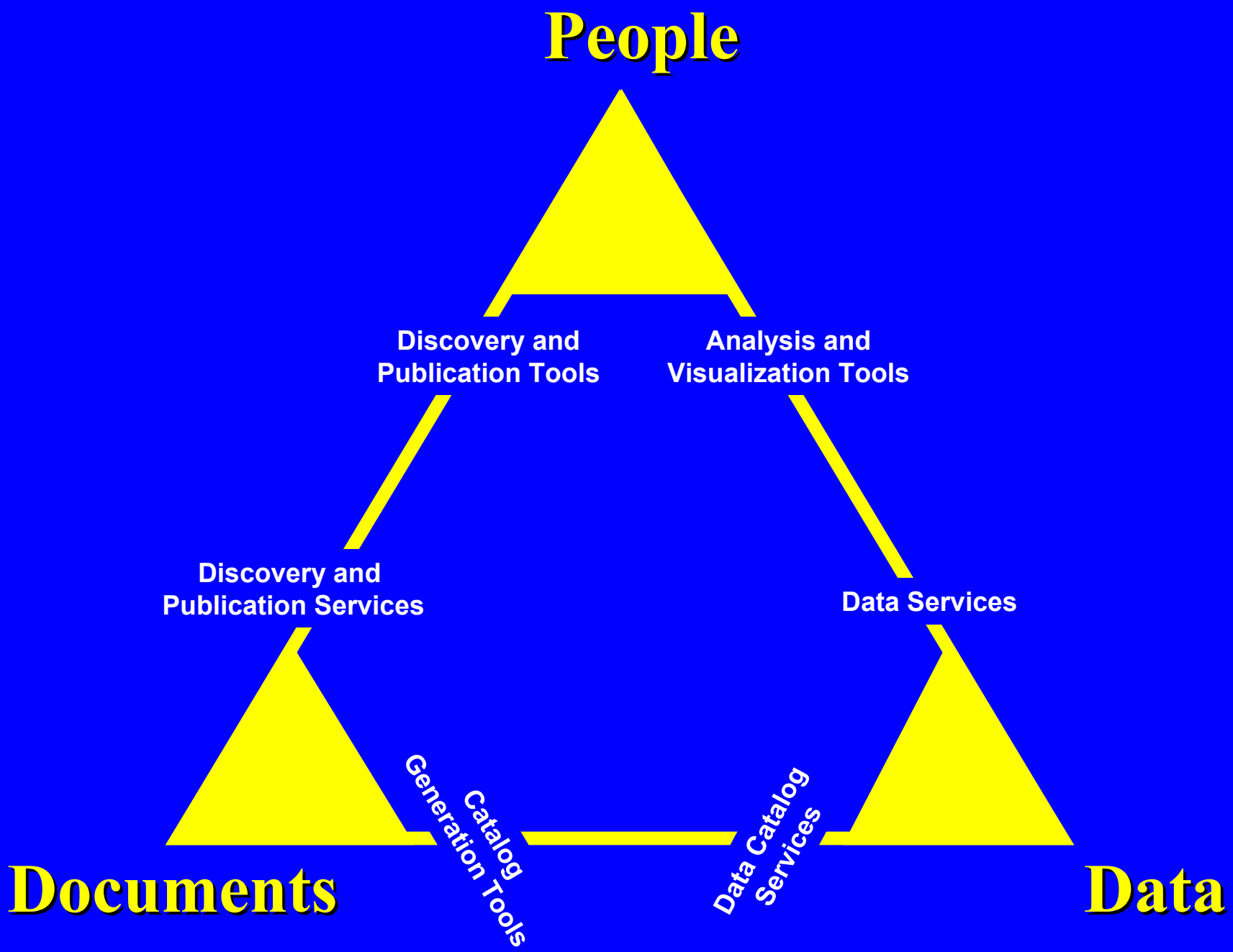
- THREDDS primary focus
  - Associate words of the science with available datasets
  - Create “compound” documents pointing to datasets
  - Connect analysis tools to documents and datasets
- Wide range of compound documents
  - Lists of datasets available on server with brief description of dataset classes
  - Online publications pointing to datasets illustrating concepts
- Massive arsenal of Web and Digital Library search/discovery tools can be applied to compound documents

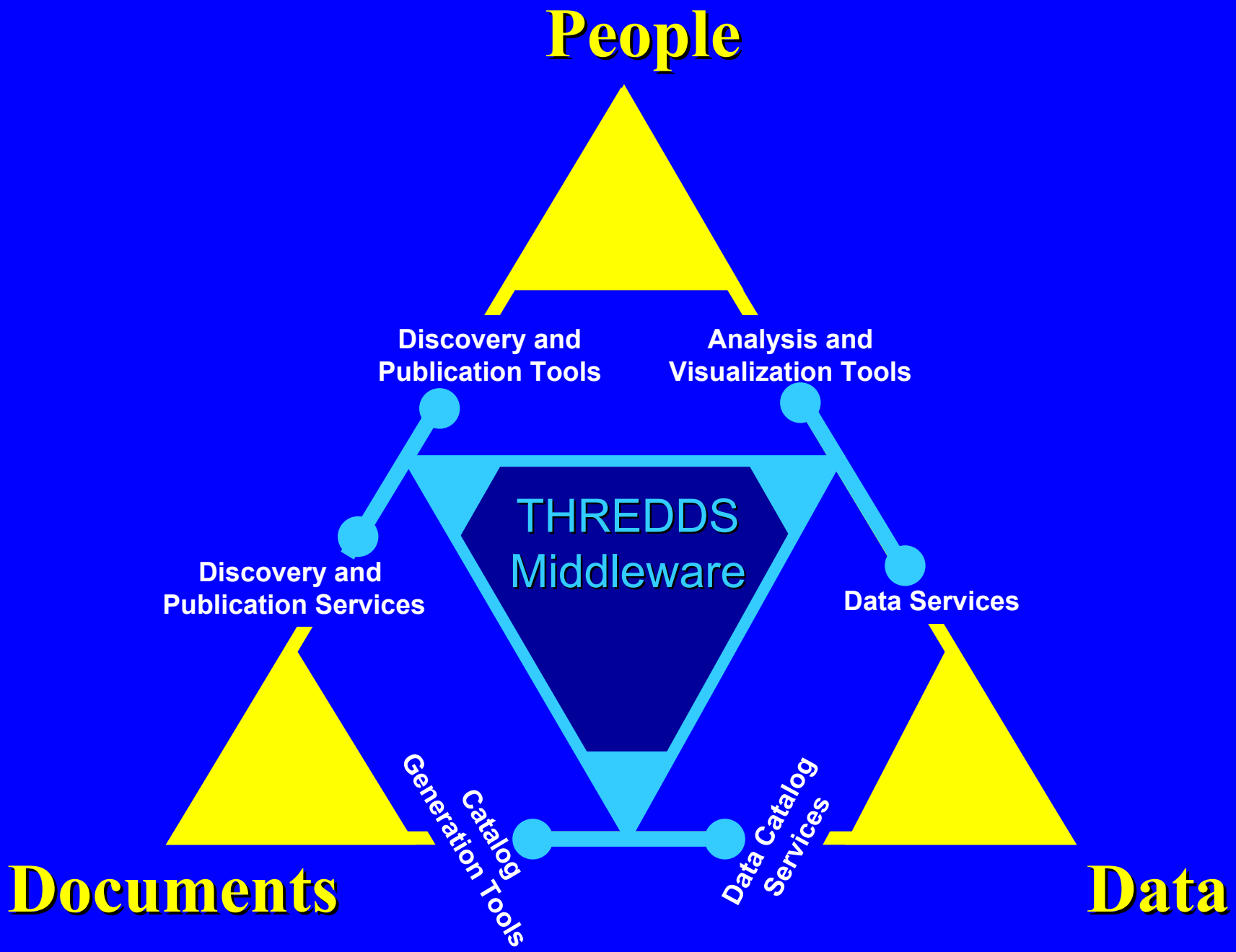

Analysis and Display Software
e.g. IDV, VGEE, LAS, INGRID, IDL, GDS
Compound Documents
i.e., electronic publications with embedded pointers to datasets and tools
Discovery Center Catalog Systems
e.g. DLESE, NSDL, GCMD
Discovery and Usage Metadata Middleware
e.g. THREDDS catalog generators, servers, harvesters
Data Server Software
e.g. FTP, OPeNDAP, ESG, OpenGIS, Web services
Data Storage System
e.g., Disk, RAID, Mass storage)

Basic THREDDS InventoryCatalog

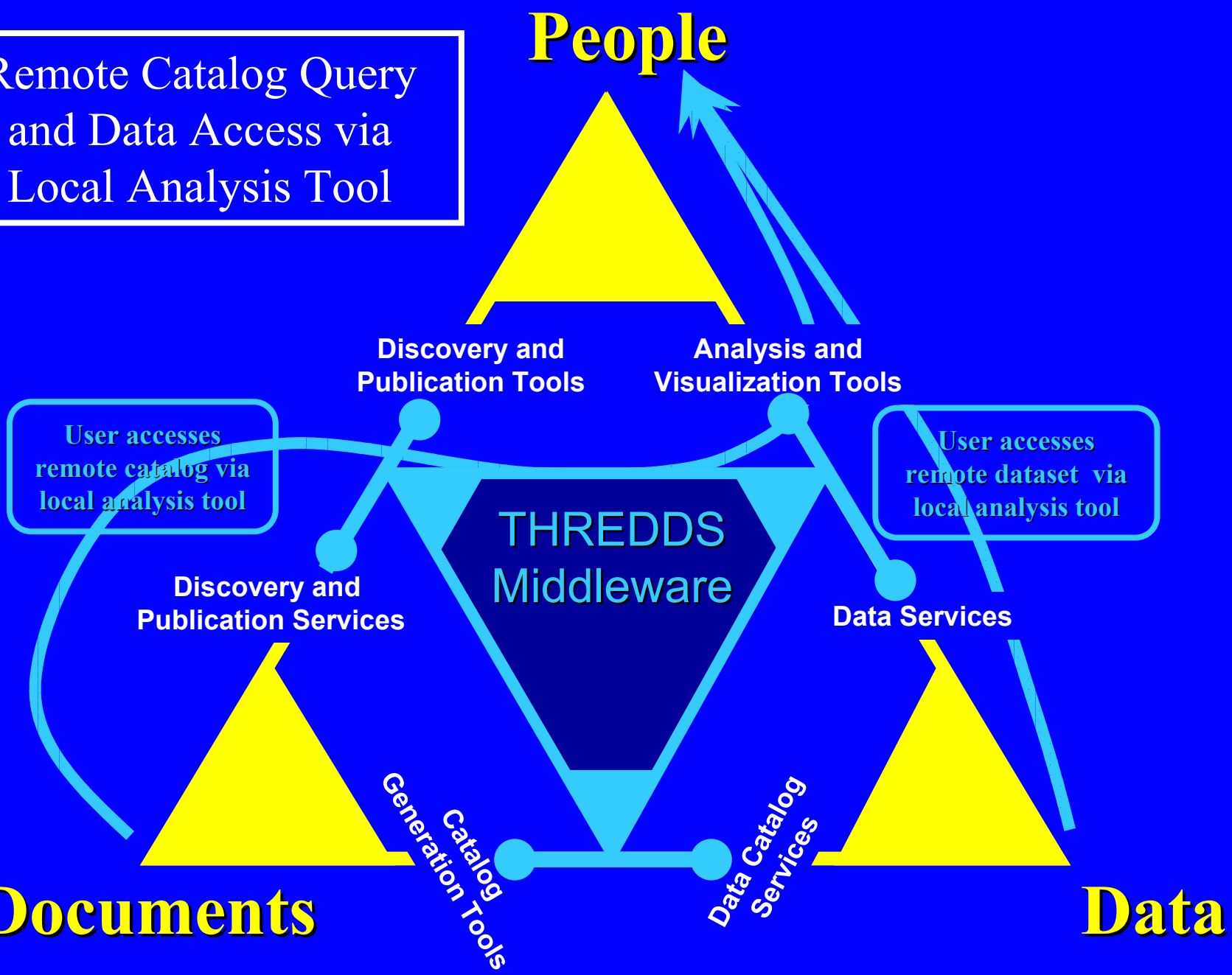
- Inventory list of datasets on server
- Generated automatically with minimal human input
- Viewed from within analysis and display application
- Can be harvested for inclusion in GCMD, DLESE, NSDL




People
Discovery and Publication Tools
Analysis and Visualization Tools
Discovery and Publication Services
Data Services
Catalog Generation Tools
Data Catalog Services
Documents
Data

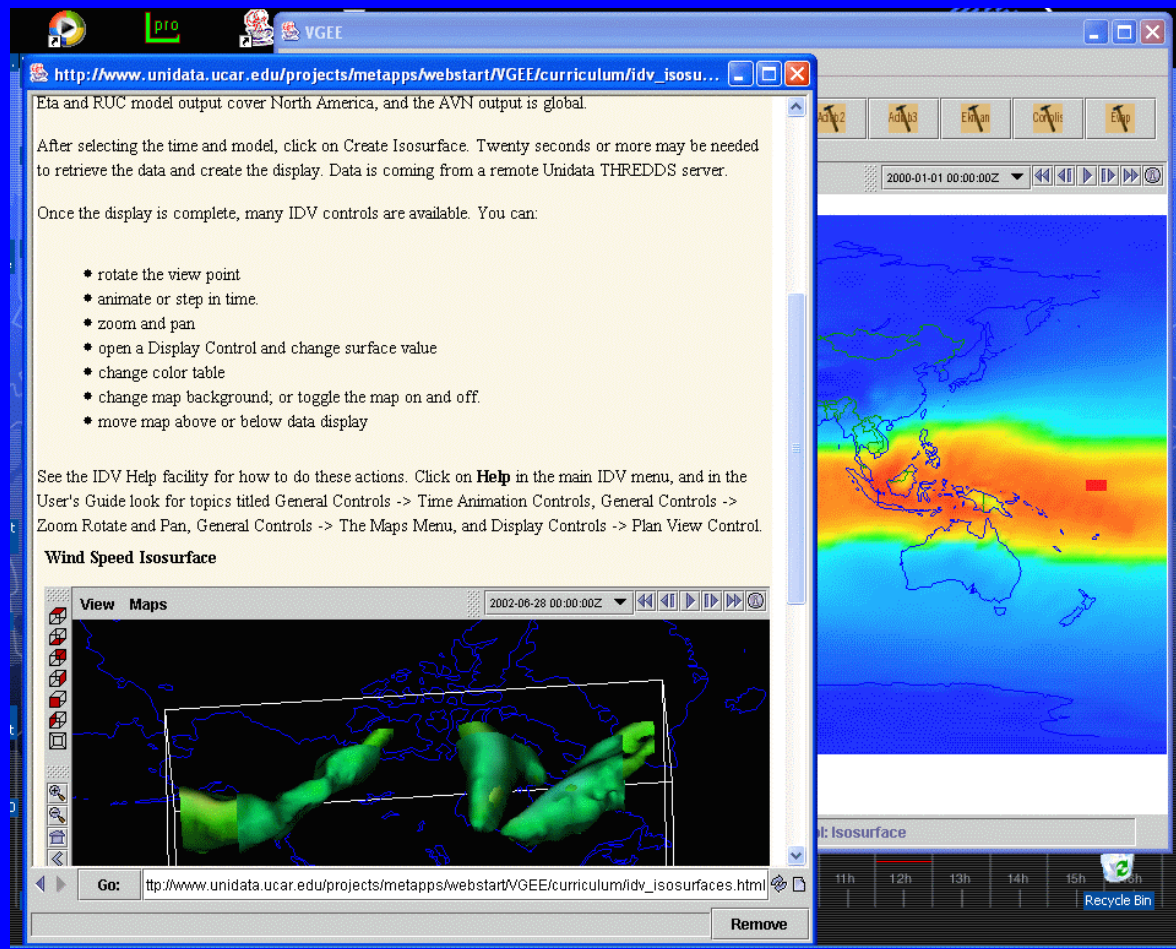

People
Discovery and Publication Tools
Analysis and Visualization Tools
THREDDS Middleware
Discovery and Publication Services
Data Services
Catalog Generation Tools
Data Catalog Services
Documents
Data

Remote Catalog Query and Data Access via Local Analysis Tool



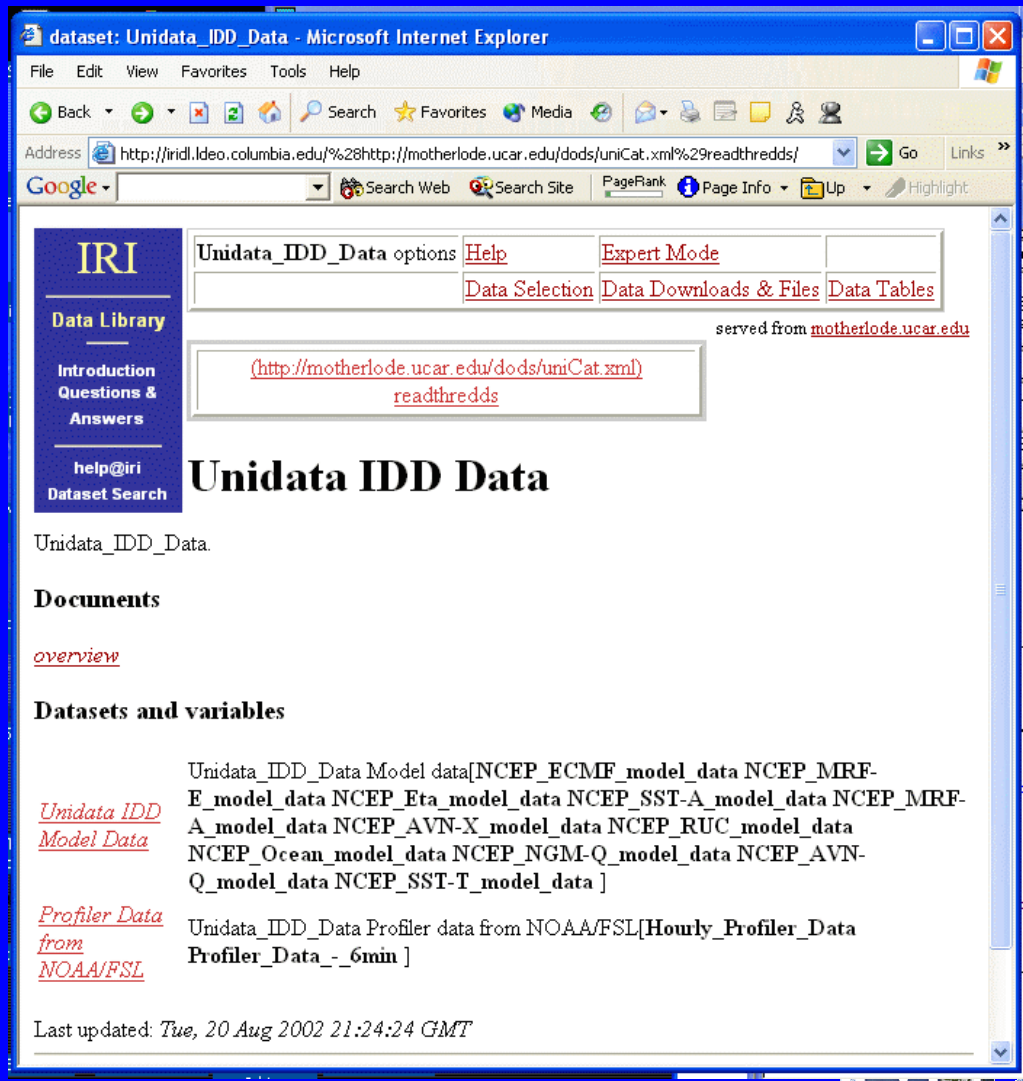
Online Publication Viewed from within Analysis Tool

- Discovery at DLESE,
- module at DPC, VGEE
- tool at Unidata,
- datasets at NCAR
- Lends itself well to Web discovery tools, DL integration
- Can be:
  - education module
  - online scientific publication

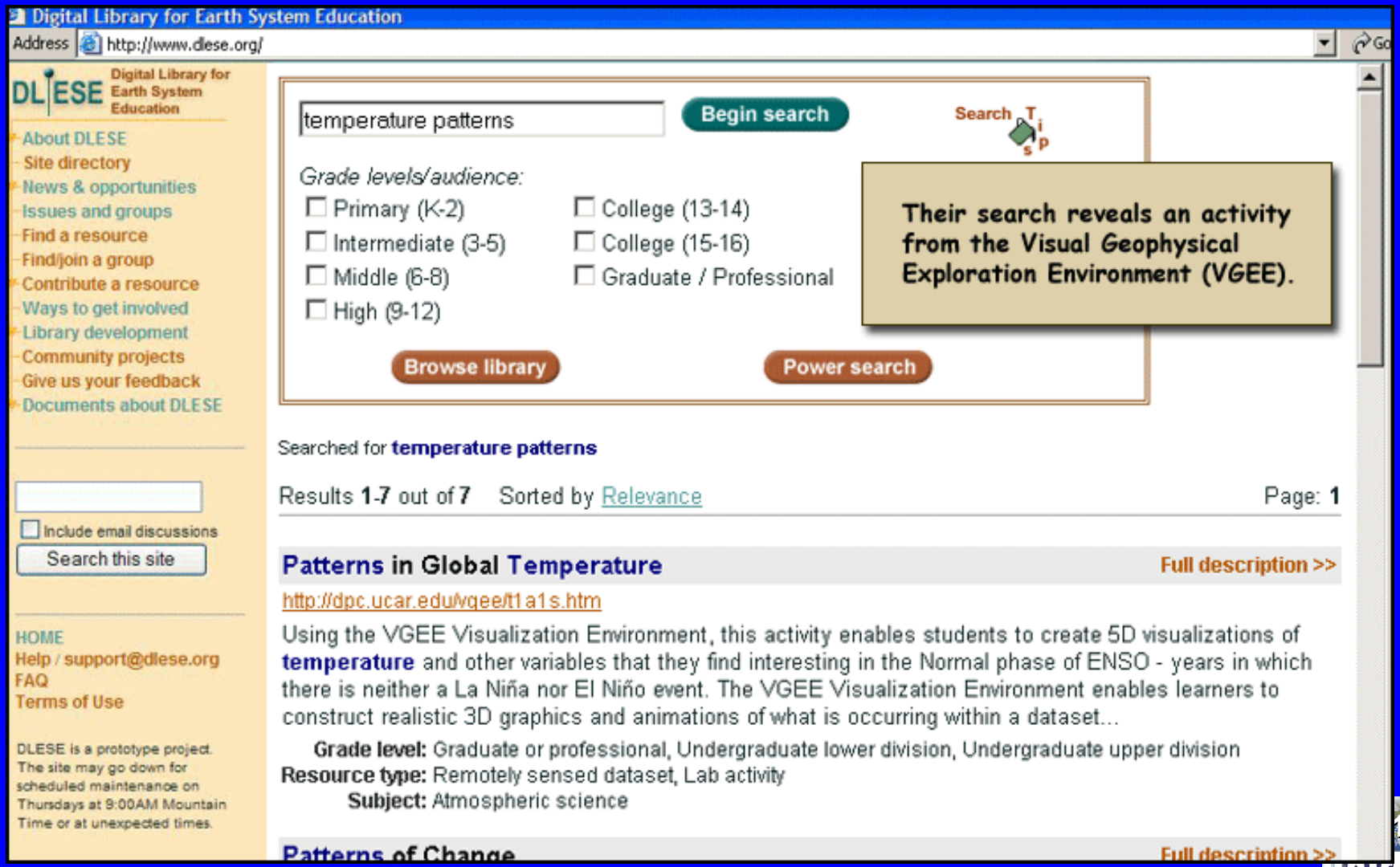


Browser-base Thin Client Access

- LDEO/IRI web site publishes catalog of datasets available on server at UCAR
- Catalog resides and is updated at UCAR
- Browsing of datasets on UCAR server from LDEO server
- Also enables analysis and display of datasets on UCAR server using tools on LDEO server



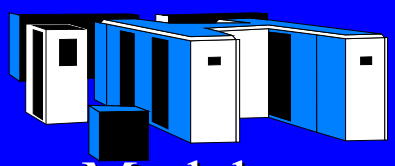
Discovery in DLESE



Stepwise creation of third-party enhanced catalogs/case studies

- Begin with basic inventory catalog
- Crawler traverses datasets listed in basic catalog and adds location “bounding box” to location-enhanced catalog
- Gazetteer service examines location-enhanced catalogs to create a catalog of datasets associated with named region on Earth
- Evolve to “event” gazetteer with 5-dimensional bounding box (e.g., model output datasets related to “Storm of the Century” with vorticity above a threshold – a distributed case study)

En Digital Library Catalog System atalogs

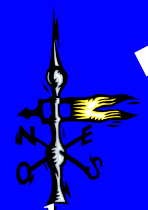


Model output

IDD

IDD

IDD



Weather Obs

Data Catalog Harvesting

NcML or ESML Generator

Enhanced Catalogs

Inventory Catalog Generator

Enhanced Inventory Catalog and Data Server

Data Mining Engine

Case Study Catalog

Event Gazetteer

Third Party THREDDS Catalog Server

Catalog Server

Catalog Server

Enhanced (via ESG) Metadata Catalog

metadata

dc		
	title	Clear Air Turbulence
	dcmitype	Collection
	identifier	ucar:ucar.scd.vets.vg.cat
	language [encoding=RFC1766]	EN
	subject	Turbulence
	description	In December of 1992, a DC-8 cargo plane westbound out of Denver encountered severe turbulence. Despite losing an engine (far right side) and 6 meters of wing the crew man the plane safely. This event - and others like it - are a strong reminder that we don't y understand clear air turbulence nearly well enough. In this particular case, observationa revealed pronounced Horizontally-aligned Vortex Tubes (HVT). NCAR and NOAA researcl together to study the incident and using NOAA weather data, developed a very high re numerical model to see what a simulation might reveal.
	coverage	temporal: December, 2002 spatial: Rocky Mountains, Boulder, Colorado, USA
	creator	Terry Clark, Larry Radke, Bill Hall, Bob Kerr, Don Middleton
	publisher	National Center for Atmospheric Research (NCAR)
	rights	Copyright 2002, University Corporation for Atmospheric Research (UCAR)
	date	created: null available: 2002-08-05

ISCCP Collection Metadata

DUBLIN CORE METADATA

title	ISCCP: International Satellite Clould Climatology Project
creator	CGD (Climate and Global Dynamics division), part of NCAR (National Center for Atmospheric Research
subject	Atmospheric Science, ISCCP, Satellite data,
description	

abstract	ISCCP 2.5 degree grid data were linearly interpolated, to a T42 grid using NCL fu linint2. see: /fs/cgd/home0/shea/ncld/ISCCP/gridInterp.ncl
tableOfContents	Mean Ozone abundance Mean Precipitable water for 680-310 mb Mean Precipita for 1000-680 mb Mean Stratosphere temperature at 50 mb Mean Tropopause te Mean Tropopause pressure TOVS: Mean Temperature at 375 mb TOVS: Mean Temperature at 500 mb TOVS: Mean Temperature at 740 mb TOVS: Mean Near- air temperature TOVS: Mean Surface pressure Mean snow/ice amount Mean RS sky composite Standard deviation over space Mean TS from clear sky composite for cloud type 15 = Deep convective Mean TAU for cloud type 15 = Deep conve Mean TC for cloud type 15 = Deep convective Mean PC for cloud type 15 = Dee convective Mean CA for cloud type 15 = Deep convective Mean WP for cloud ty Cirrostratus Mean TAU for cloud type 14 = Cirrostratus Mean TC for cloud type Cirrostratus Mean PC for cloud type 14 = Cirrostratus Mean CA for cloud type 1 Cirrostratus Mean WP for cloud type 13 = Cirrus Mean TAU for cloud type 13 = Mean TC for cloud type 13 = Cirrus Mean PC for cloud type 13 = Cirrus Mean CA type 13 = Cirrus Mean WP for cloud type 12 = Nimbostratus,ice Mean TAU for cl 12 = Nimbostratus,ice Mean TC for cloud type 12 = Nimbostratus,ice Mean PC fo type 12 = Nimbostratus,ice Mean CA for cloud type 12 = Nimbostratus,ice Mean cloud type 11 = Altostratus,ice Mean TAU for cloud type 11 = Altostratus,ice M cloud type 11 = Altostratus,ice Mean PC for cloud type 11 = Altostratus,ice Mea cloud type 11 = Altostratus,ice Mean WP for cloud type 10 = Altocumulus,ice M for cloud type 10 = Altocumulus,ice Mean TC for cloud type 10 = Altocumulus,ic for cloud type 10 = Altocumulus,ice Mean CA for cloud type 10 = Altocumulus,ic WP for cloud type 9 = Nimbostratus liquid Mean TAU for cloud type 9 = Nimbost

More ISCCP Metadata

publisher	Climate Analysis Section, Climate and Global Dynamic Division, National Center for Atmospheric Research
contributor	Hongjun Zhang (zhangho@ucar.edu)
date	**created** [encoding=W3CDTF]: 2002-09-23
type [encoding=DCMI]	collection
format	netcdf
identifier	isccp/D2/
identifier [encoding=URI]	http://www.cgd.ucar.edu/cas/catalog/satellite/isccp/D2/
source	ftp://isccp.giss.nasa.gov
language	en
relation	**isVersionOf**: ISCCP **references** [encoding=URI]: http://isccp.giss.nasa.gov/products/dataview.html
coverage	**spatial** **box** [name=the whole earth] [projection=geographic, height relative to mean-sea-level] **northlimit** [units=signed decimal degrees]: 90 **eastlimit**: 0

Next Steps

- Bring catalog servers online at remaining data provider sites
- Complete work on client side catalog access modules
- Incorporate THREDDS modules into additional clients
- Add enhanced catalog capabilities
- Complete work to make catalogs compatible with DLESE catalog entries
- Add server components to enable DL harvesting of catalogs
- Work with DLESE to create sample educational modules with embedded data and tool pointers

Next Phase

- Participate in DLESE Data Services workshops:
  - Technical sessions
  - Educational module developers
  - End users: professors, teachers, students
- Work with education partners to develop more educational materials with embedded data and tools
- Work with current THREDDS collaborators on tools for creating enhanced catalogs/virtual case studies
- Incorporate GIS (Geographic Information System) datasets into collections
- Develop new partnerships for integrating GIS data and tools into educational modules
- Work with OGC (Open GIS Consortium) protocols with focus on Web Coverage Service (WCS)
- Incorporate MyWorld GIS client into THREDDS
- Integrate THREDDS with evolving GRID technology

Funding Opportunities

- Unidata 2008 proposal to NSF/ATM: Distributed, organized collections of digital material
- DLESE Data Services
- NASA REASON CAN (2) include OPeNDAP support resources
- LEAD Large ITR proposal includes resources for IDV, THREDDS, LDM development
- NSDL Collections: follow on proposal due April 21

More Information

- http://www.unidata.ucar.edu/projects/THREDDS/

- ben@unidata.ucar.edu