LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY

University Corporation for Atmospheric Research

Colorado State University

Howard University

Indiana University
Quality Education. Lifetime Opportunities.®

The University of Oklahoma

UNC Chapel Hill

UAH The University of Alabama in Huntsville

University of Illinois at Urbana-Champaign

Millersville University

NSF

Sponsored by the National Science Foundation

What is LEAD?

- A 5-year NSF Large Information Technology Research (ITR) Grant
- Began 1 October 2003
- $10.25 million/5 years
- Nine institutions (including Unidata), 105 people
- Target audience: Undergraduates and graduate student researchers & faculty
- A research project with a long-term view as a deployed capability for the community
- A “model” of collaboration among disciplines that has been studied and profiled in the literature by organizational behavior experts

The LEAD Vision

Revolutionize the ability of scientists, students, and operational practitioners to observe, analyze, predict, understand, and respond to intense local weather by interacting with it dynamically and adaptively in real time

Motivation #1 for LEAD

- Weather technologies are very hard to use in sophisticated ways because they're so complicated and linked together using cumbersome scripts - creating a HUGE DIVIDE BETWEEN THE HAVES AND THE HAVE NOTS

Motivation #2 for LEAD

- High impact local weather is VERY DYNAMIC while our meteorological tools, cyber environments and learning modalities are almost entirely STATIC

Two Principal Goals of LEAD

- **Goal #1**: Lowering the barrier for using complex end-to-end weather technologies
  - Democratize the availability of advanced weather technologies for research and education
  - Empower application in a grid context
  - Facilitate rapid understanding, experiment design and execution
- **Goal #2**: Dynamic Adaptation to Weather
  - Models and hazardous weather detection systems responding to observations and their own output
  - Models and hazardous weather detection systems driving the collection of observations
  - IT infrastructures providing on-demand, fault tolerant services

The Potential Payoff

- An improved understanding of mesoscale weather by studying it in ways that are consistent with its behavior
- Advanced meteorological capabilities available to a much broader community of users
- Much shallower learning curve -- design and execute experiments in minutes rather than months.
- Application of weather technologies in a grid context
- An environment for ongoing basic research in computer science (e.g., data, workflow, monitoring, QoS)
- An environment where meteorologists can learn about new computing concepts

Important Ingredient: The Notion of a Web Service

LEAD Uses a Service-Oriented Architecture

Service A
(Analysis)

Service B
(Model)

Service C
(Radar Stream)

Service D
(Work Space)

Service E
(VO Catalog)

Service F
(Viz Engine)

Service G
(Monitoring)

Service H
(Scheduling)

Service I
(Decoder)

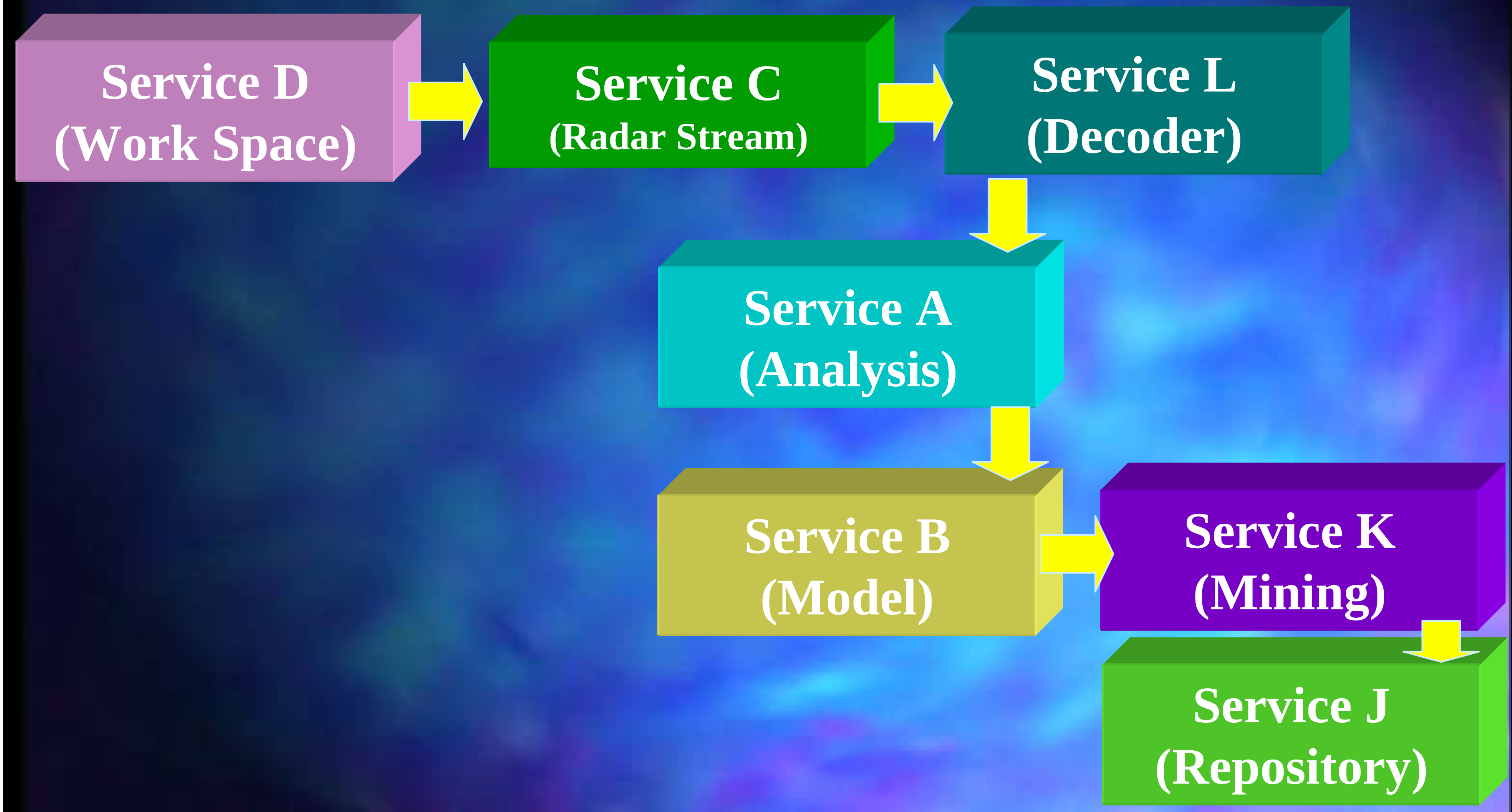
Service J
(Repository)

Service K
(Mining)

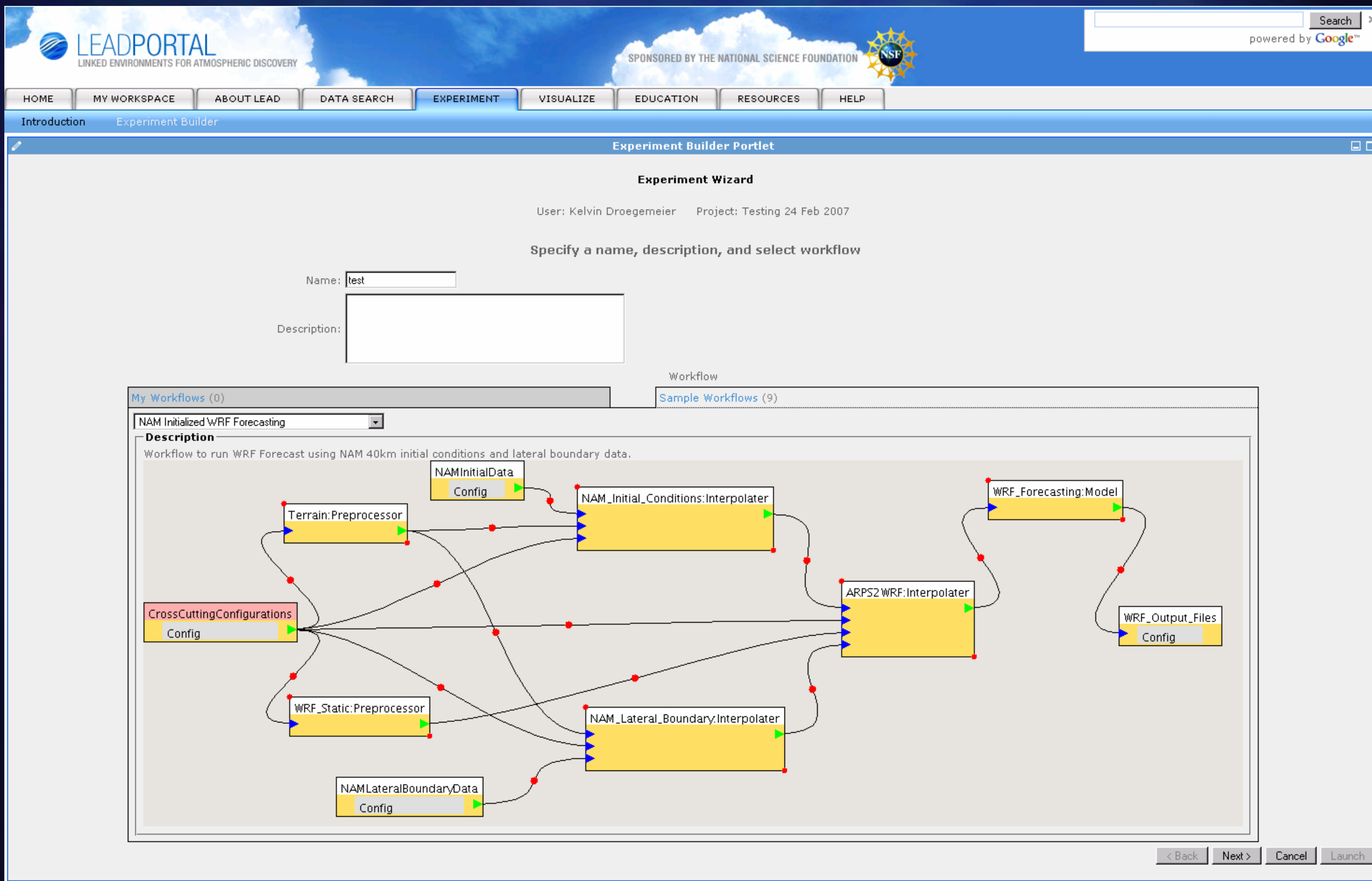
Service L
(Decoder)

Many others…

Can Solve Broad Classes of Problems by Linking Services Together in Workflows

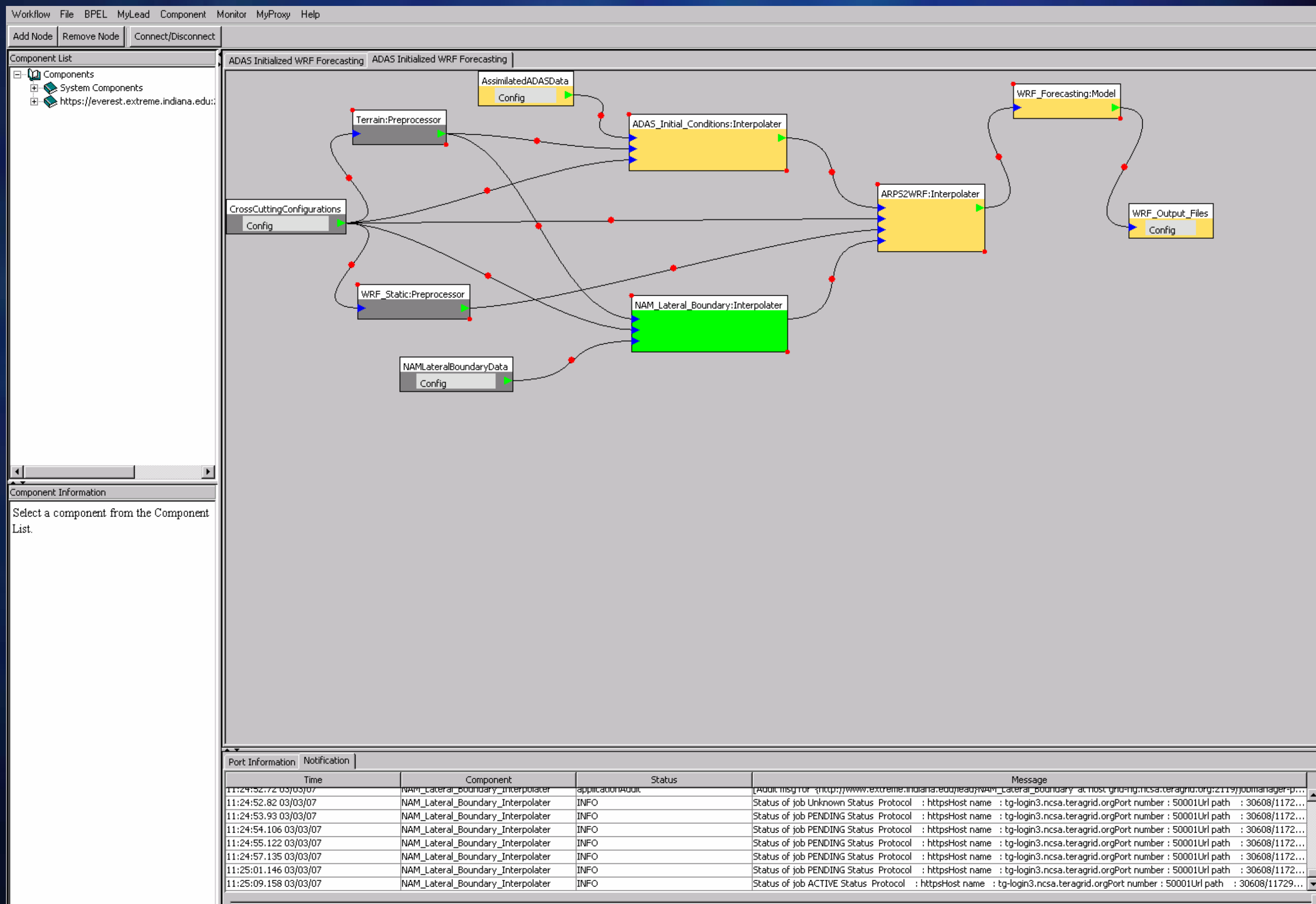


Experiment Builder



http://portal.leadproject.org

Monitoring & Resource Prediction Tool



http://portal.leadproject.org

The LEAD Framework/Environments

Fundamental Capabilities

Query & Acquire | Simulate | Assimilate | Analyze/Mine | Predict | Vi e

Foundational User Tools

Portal | ADAS | MyLEAD | WRF | ADaM

Enabling Functions

Define | Query | Store | Publish | Allocate | Monitor | Edit | Authorize | Build | ge

Acquire | Execute | Compile | Steer | Catalog | Configure | Move | Trigger | ode

Middleware Services

- Authorization
- Authentication
- Notification
- Monitoring
- Workflow
- Security
- ESML

- VO Catalog
- THREDDS Catalog
- MyLEAD
- Control
- Query
- Stream
- Transcoder

- Ontology
- Host Environment
- GPIR
- Application Host
- Execution Descriptio
- Application Descript

Outcomes

Data Sets | Model Output | Gridded Analyses | Animations | Static Images | Relation

Meteorology

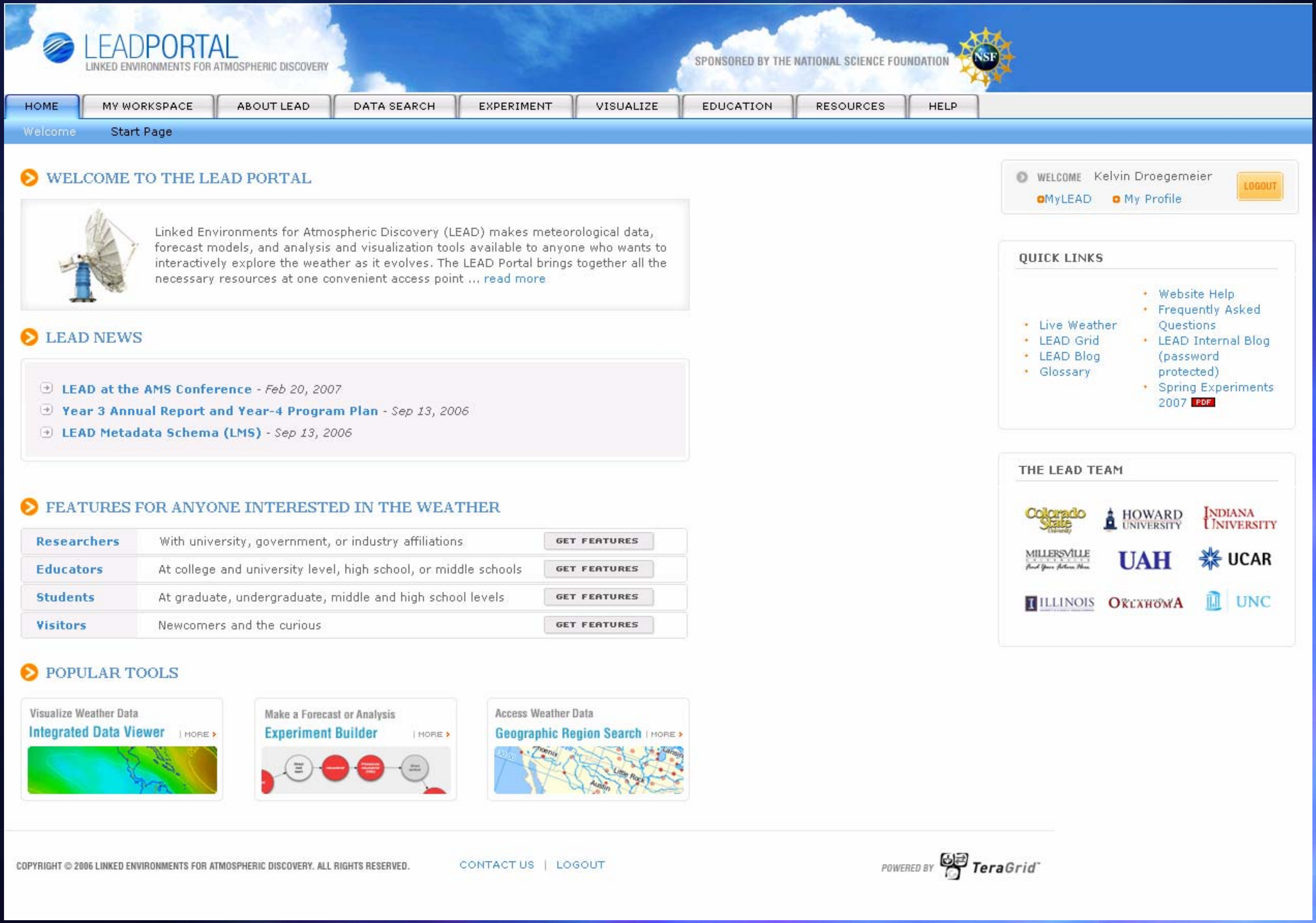
CS

CS

Meteorology

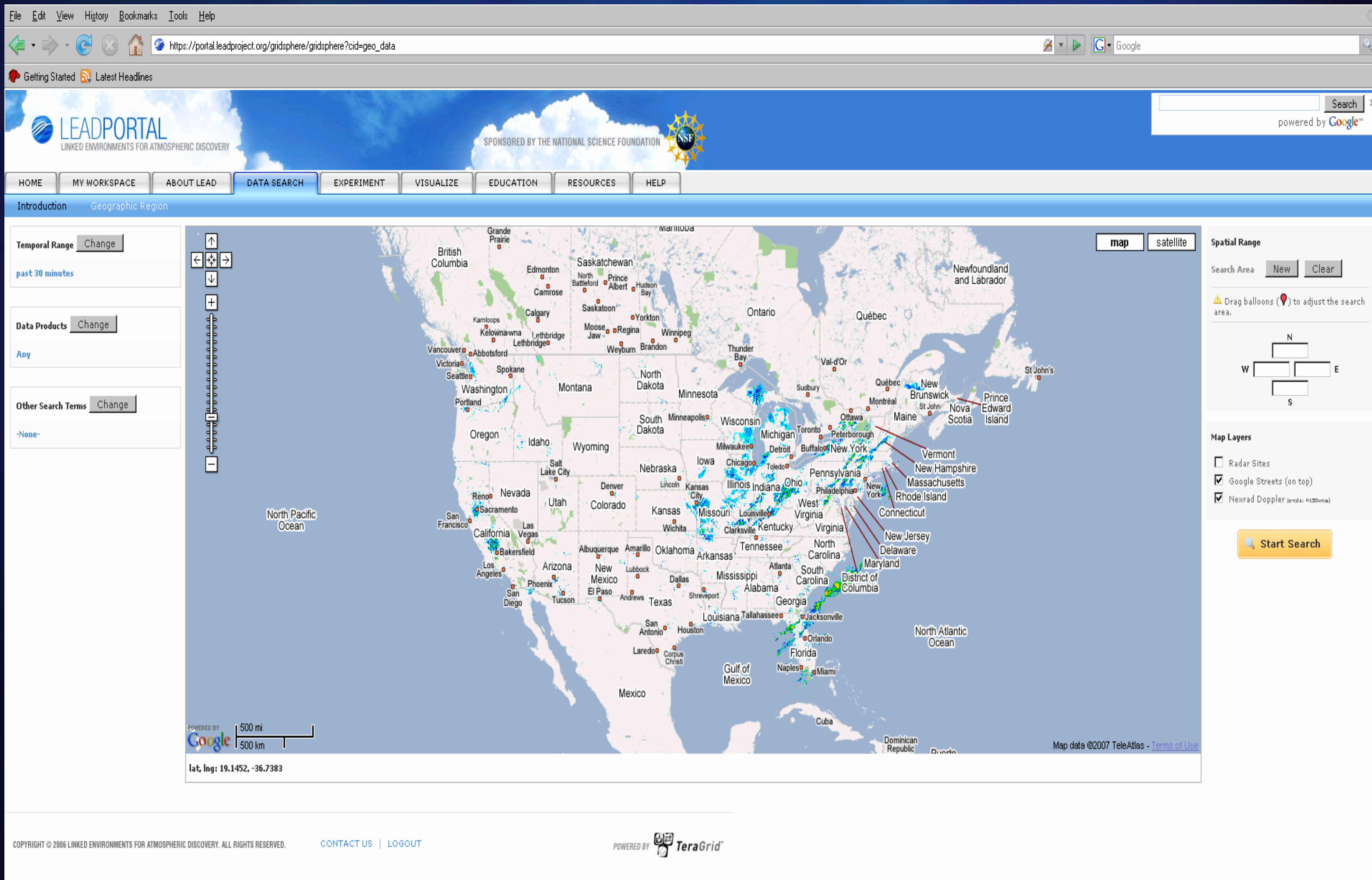
New Knowledge, Understanding, Ideas

LEAD Portal Home Page



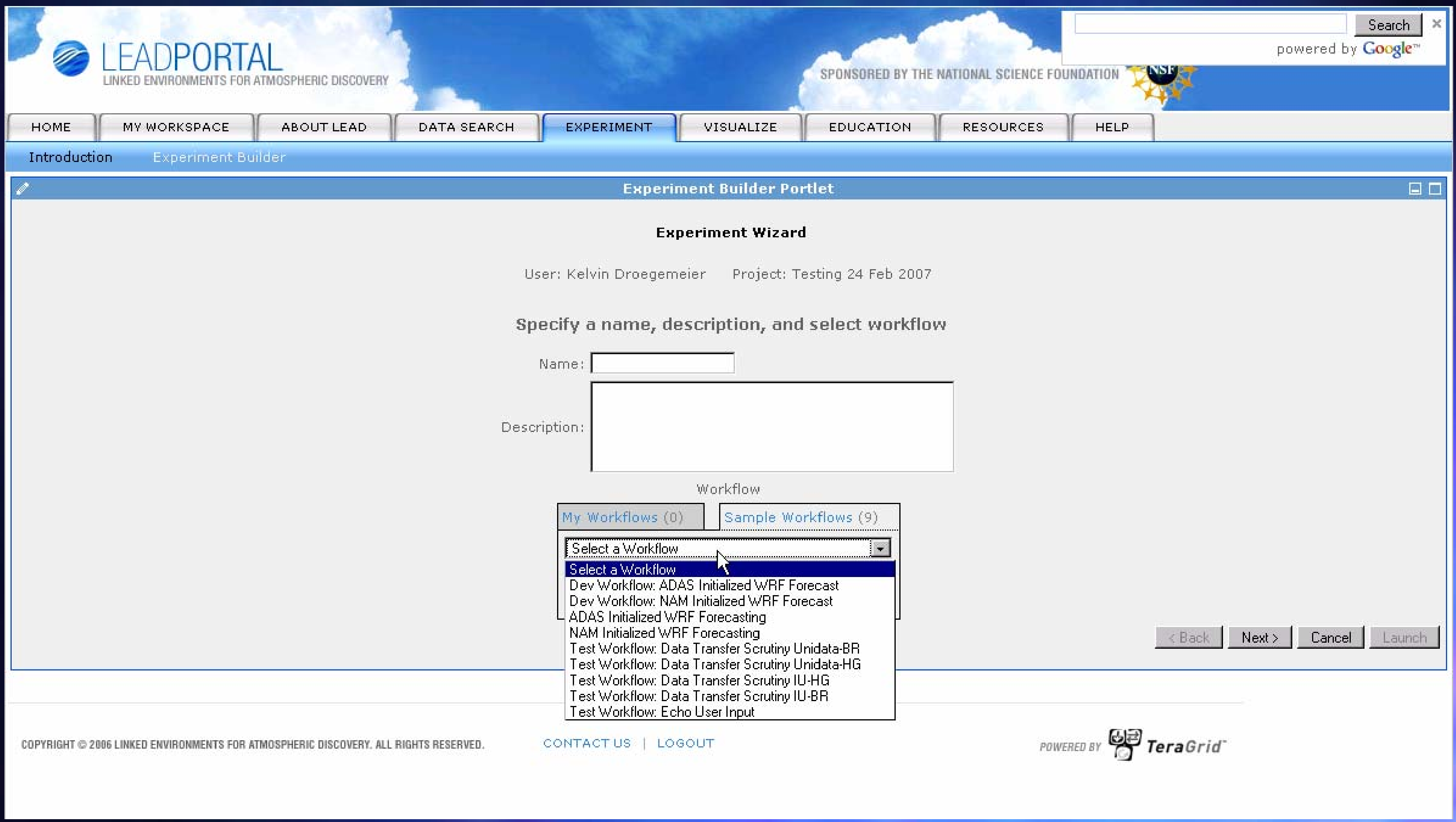
http://portal.leadproject.org

Data Query System



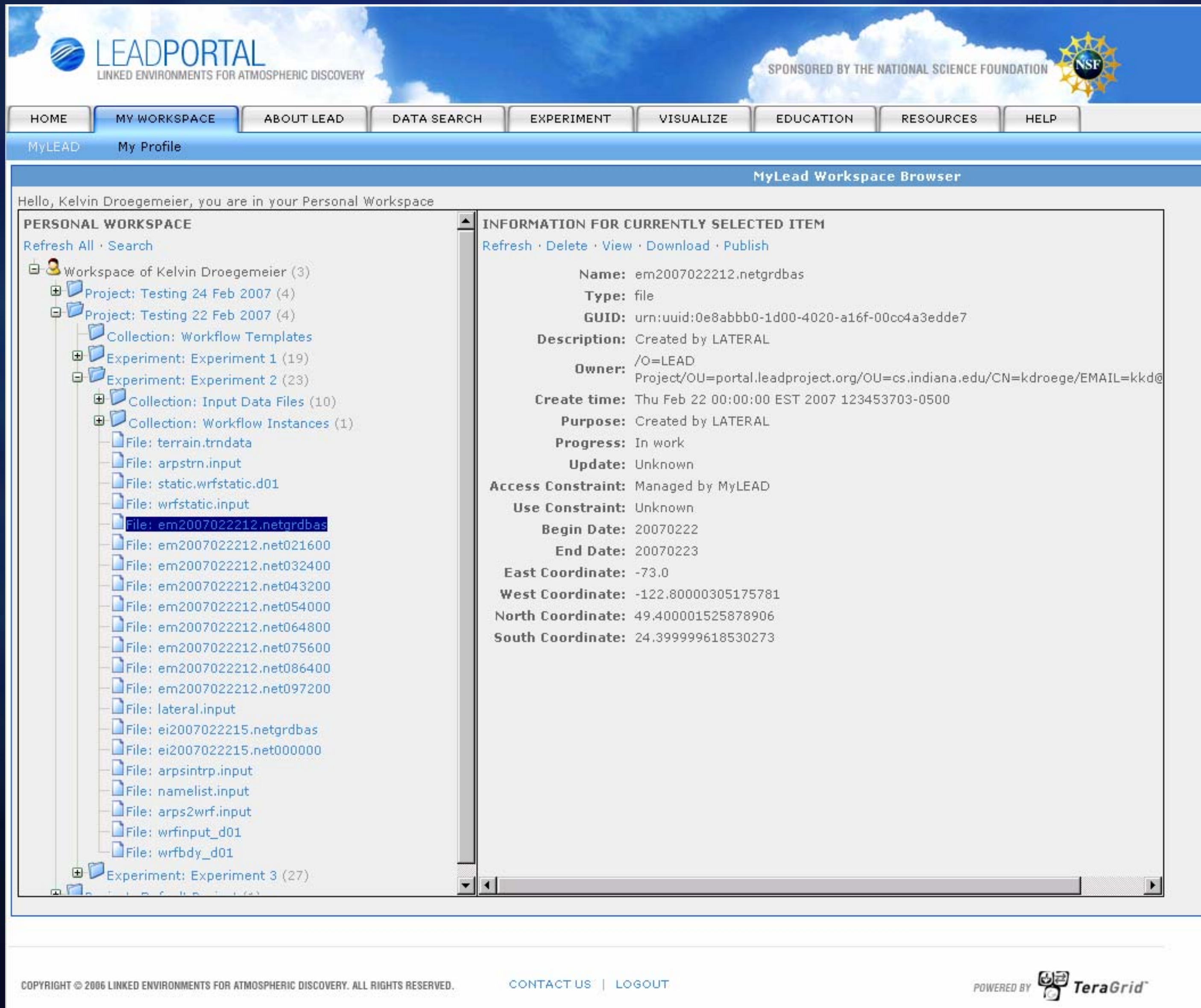
http://portal.leadproject.org

Experiment Builder

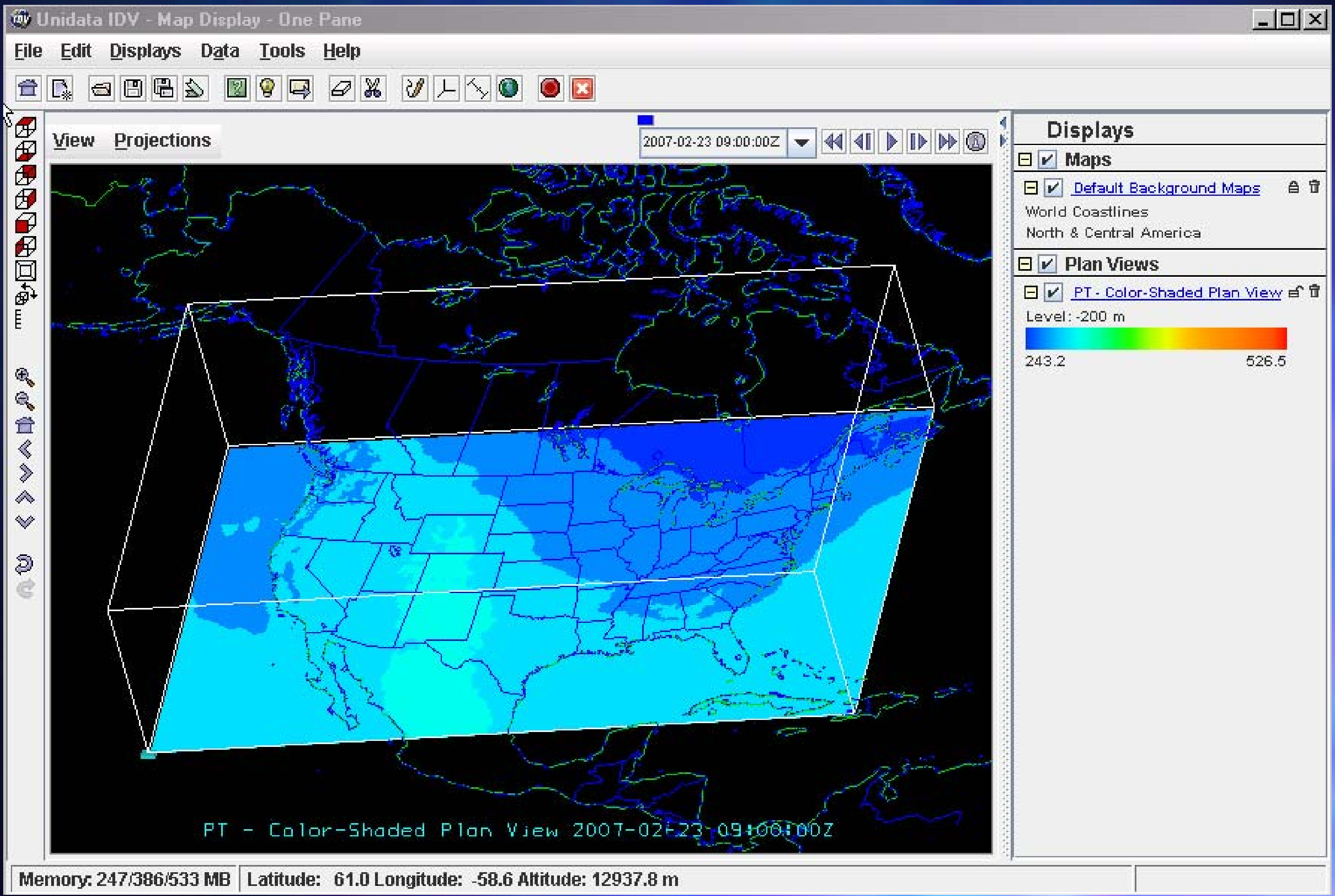


http://portal.leadproject.org

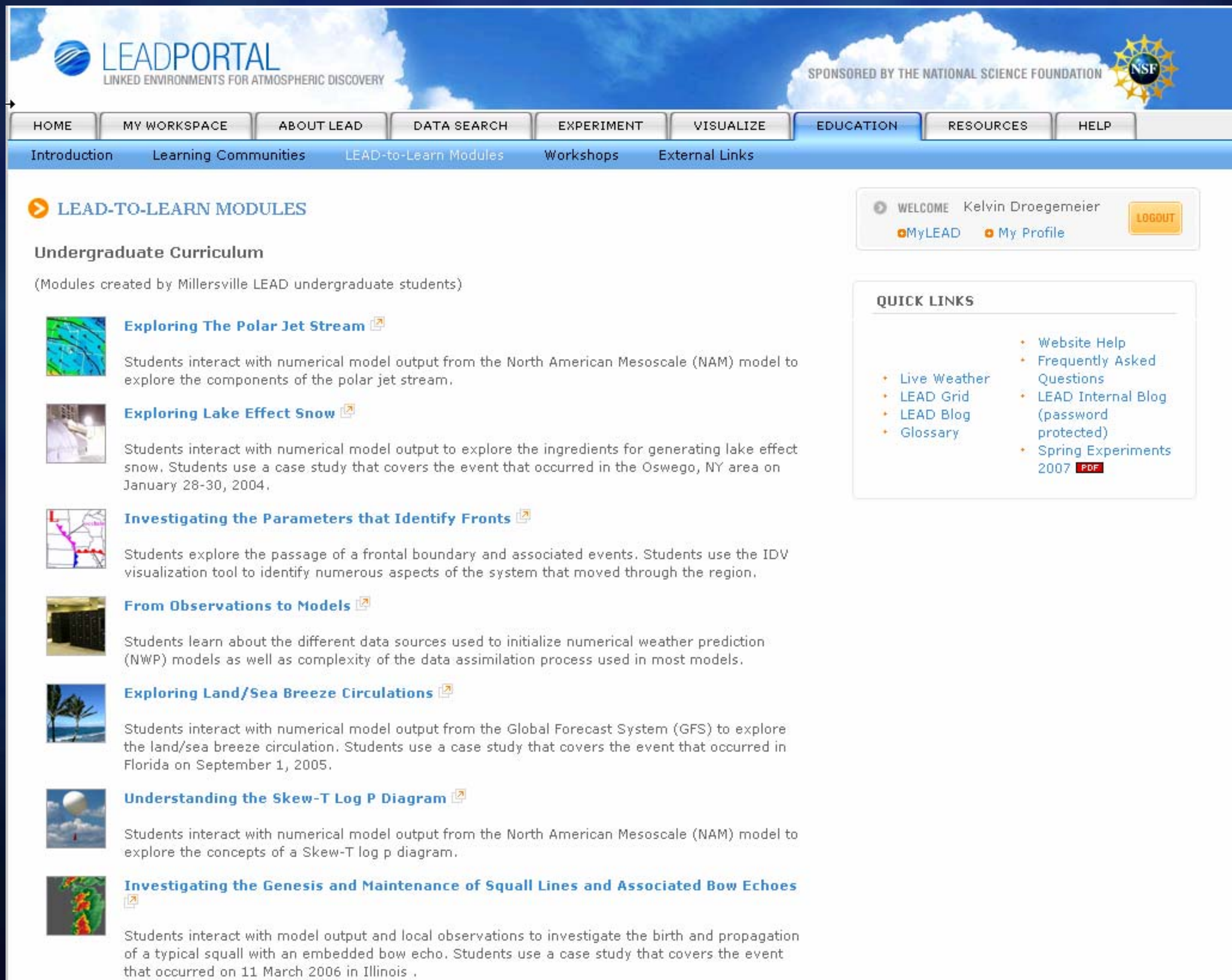
MyLEAD Personal Workspace



Visualization and Analysis



LEAD Education Programs



LEADPORTAL
LINKED ENVIRONMENTS FOR ATMOSPHERIC DISCOVERY
SPONSORED BY THE NATIONAL SCIENCE FOUNDATION

HOME | MY WORKSPACE | ABOUT LEAD | DATA SEARCH | EXPERIMENT | VISUALIZE | EDUCATION | RESOURCES | HELP

Introduction | Learning Communities | LEAD-to-Learn Modules | Workshops | External Links

LEAD-TO-LEARN MODULES

Undergraduate Curriculum

(Modules created by Millersville LEAD undergraduate students)

Exploring The Polar Jet Stream

Students interact with numerical model output from the North American Mesoscale (NAM) model to explore the components of the polar jet stream.

Exploring Lake Effect Snow

Students interact with numerical model output to explore the ingredients for generating lake effect snow. Students use a case study that covers the event that occurred in the Oswego, NY area on January 28-30, 2004.

Investigating the Parameters that Identify Fronts

Students explore the passage of a frontal boundary and associated events. Students use the IDV visualization tool to identify numerous aspects of the system that moved through the region.

From Observations to Models

Students learn about the different data sources used to initialize numerical weather prediction (NWP) models as well as complexity of the data assimilation process used in most models.

Exploring Land/Sea Breeze Circulations

Students interact with numerical model output from the Global Forecast System (GFS) to explore the land/sea breeze circulation. Students use a case study that covers the event that occurred in Florida on September 1, 2005.

Understanding the Skew-T Log P Diagram

Students interact with numerical model output from the North American Mesoscale (NAM) model to explore the concepts of a Skew-T log p diagram.

Investigating the Genesis and Maintenance of Squall Lines and Associated Bow Echoes

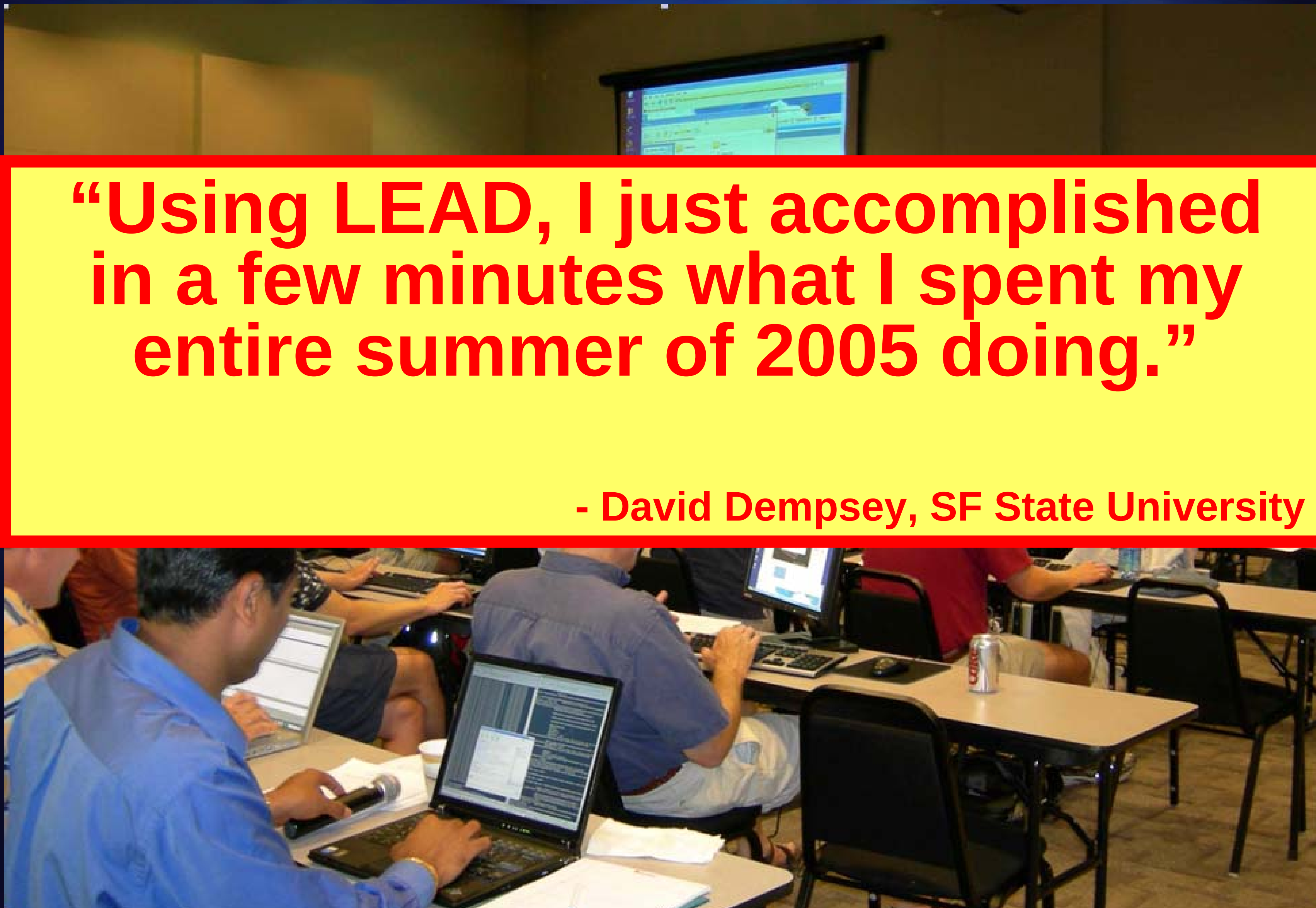
Students interact with model output and local observations to investigate the birth and propagation of a typical squall with an embedded bow echo. Students use a case study that covers the event that occurred on 11 March 2006 in Illinois .

WELCOME Kelvin Droegemeier LOGOUT
MyLEAD | My Profile

QUICK LINKS

- Live Weather
- LEAD Grid
- LEAD Blog
- Glossary
- Website Help
- Frequently Asked Questions
- LEAD Internal Blog (password protected)
- Spring Experiments 2007 PDF

Unidata Users Workshop



“Using LEAD, I just accomplished in a few minutes what I spent my entire summer of 2005 doing.”

- David Dempsey, SF State University

2007 Real Time Operational Tests: The SPC/NSSL Spring Program

- Each day, LEAD will be used to create...
  - 10-member ensemble WRF forecast at 4 km grid spacing
  - 2 km forecasts launched automatically over tornado watches
  - 2 km forecasts launched by forecasters on demand
- Meteorological impact plus cognition research on decision making
- Simultaneous testing of prediction with CASA radar data

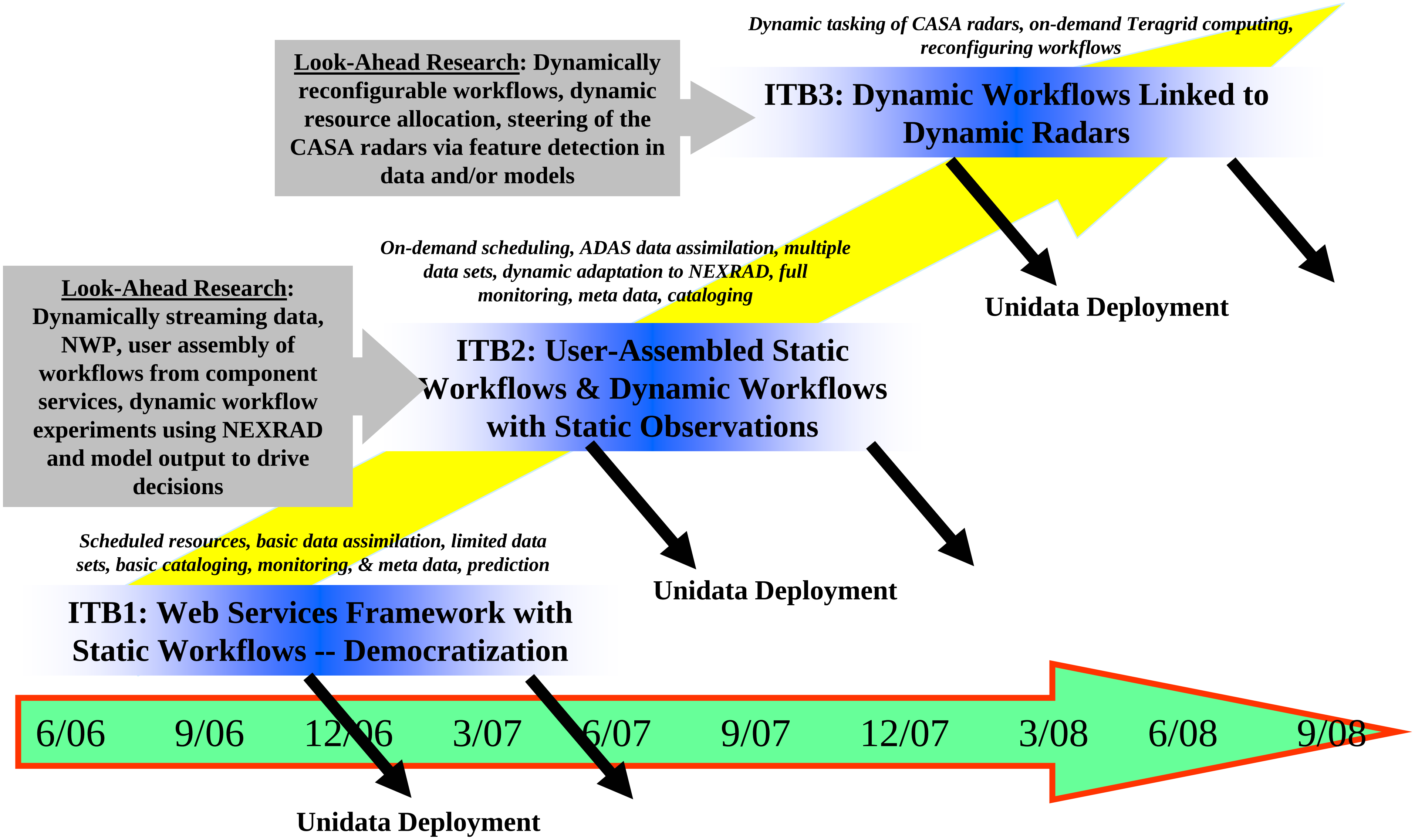
Broadening Exposure of LEAD

- National Collegiate Forecast Contest (now called the Weather Challenge)
  - Impact of “personal” WRF runs on individual forecasts
  - Decision science and human factors research
  - Informal testing in fall 2006/spring 2007
  - Prepare 3-year NSF proposal for funding beginning in fall 2007
  - Utilize TeraGrid

Recent Student User Comment

- Rich Clark and Don Murray gave a talk on LEAD at last week’s NE Storms Conference
- To Prof. Todd Sakora, Millersville University from Eric Meyers
  - “Attached is a display of 2-m T and wind depicting the WRF's interpretation of the coastal front on 14 February 2007. It's interesting that I found an example using IDV that parallels our discussion of mesoscale boundaries in class. It illustrates very nicely the transition to a coastal low and the strong baroclinic zone with a location very similar to Markowski's depiction. **I created this image in IDV after running a 5-km WRF run (initialized with NAM output) via the LEAD Portal.** This simple 1-level plot is just a precursor of the many capabilities IDV will eventually offer to visualize high-res WRF output. Enjoy!”

LEAD Strategic Timeline



Role of Unidata in LEAD (1)

- Unidata is a core partner in LEAD chosen strategically for its
  - Expertise and track record in developing and deploying robust software to a huge community of users
  - Expertise in data, servers, catalogs
  - Linkages to the academic and research community
- Quoting from the original LEAD proposal:
  - "The impacts of LEAD will be broad and sustained. As a virtual extension of the user's desktop, and via deployment through Unidata... LEAD will enable researchers, users, educators, and students to use atmospheric models and other tools in more realistic, real time settings than is now possible."

Role of Unidata in LEAD (2)

- Mohan Ramamurthy was the intellectual driver behind LEAD's adoption of the web services concept
- Anne Wilson is co-leader of the Data Thrust
- Unidata was the architect of software engineering principles, repositories and bug tracking used in LEAD
- Tom Baltzer has developed several components of LEAD and is heading a rigorous testing and deployment program
- LEAD was featured prominently as part of the Users Workshop last summer and the community reception was highly positive.
- LEAD is one of the most important use cases for many Unidata technologies, including IDV, THREDDS, TDS catalogs, TDR, and metadata crosswalks.
- IDV is the visualization and analysis engine and THREDDS is the remote data access technology for LEAD
- LEAD is heavily leveraging the THREDDS system and has stimulated many new ideas and funded new activities in Unidata (e.g., THREDDS Data Repository, advances in TDS and IDV w.r.t. staggered grids and sub-setting services, establishment of data repository for 3-6 months of 7 key data streams, developments in storage and cluster technologies, next generation case studies, and provision of high resolution model output for GIS communities including CUAHSI)
- Deployment of Unidata technologies by LEAD testbed sites and collaboration with LEAD partners has been most valuable and is benefiting Unidata and the larger Unidata community in a significant way.

The Question

- Should Unidata become the home for LEAD following its formal tenure as an ITR Grant?
- Three-part role
  - Deploy LEAD to the community
  - Continue a smaller program of research and technology infusion with some of the existing LEAD partners
  - Home for open-source LEAD IP

Considerations

- Timing is consistent with the next 5-year Unidata proposal/cooperative agreement
  - Inclusion of LEAD would be done apart from the renewal proposal
- NSF seems to be encouraging this direction, which is consistent with the original LEAD proposal
- Unidata remains the ideal home for LEAD
  - No other organizations have the requisite software engineering & deployment/support capabilities, or linkages to the broad community
- No other discussion has occurred -- the Policy Committee is the first formal vetting of the idea

Opportunities (1)

- LEAD would bring a completely new array of capabilities into the Unidata framework while leveraging a number of existing assets
- Would help move Unidata in the direction of connecting its technologies to grid computing in a bold new way
- Potential for significant contribution to atmospheric science research and education, especially modeling
- Could provide another mechanism to link Unidata with other geoscience disciplines and communities
- Would be a perfect application area for NCAR's new computing facilities, especially if they become the Geo-Collaboratory; NCAR is already a TeraGrid node.

Opportunities (2)

- Could allow Unidata to become an on-demand prediction facility for HIAPER, field campaigns, and the community
- Unique framework for K-12 education, especially applied to traditionally underrepresented and technologically disadvantaged groups
- Significant interest by Microsoft with possibility for collaboration in both technology and education
- Would provide a mechanism to continue UOP's established role of linking research with community resources
- Good opportunity for UCAR to interact with the basic CS and meteorology research enterprise
- Links with the DTC could provide opportunity to move technology into operations

Challenges and Risks (1)

- LEAD is a notably different type of application than traditionally deployed by Unidata, consisting of multiple components with overall less stability than single applications such as IDV, GEMPAK and LDM. This poses risks to Unidata's strong established reputation
- Many of the technologies in LEAD have originated from other institutions, and not all of those technologies have been completely hardened yet.
- Currently, no grid computing expertise resides within Unidata; New staff with appropriate skills will be needed
- Given the scope of the project, LEAD might become the 800 pound gorilla within Unidata. Internal dynamic will need to be managed carefully.
- Unknown amount of software re-engineering will be required to convert research system to a deployed facility

Challenges and Risks (2)

- Broad adoption by the community is not assured owing to limited exposure at the present time
- Could become a drain on Unidata resources even with sufficient new funding
- NSF support for facilities is planned to decrease, and Unidata would have to pick among its assets, with LEAD perhaps a low priority given its newness
- The change in project leadership would have to be carefully managed to maintain the extant positive chemistry of those continuing to be involved and those who won’t
- A dual research-operations framework involving external partners could be challenging

Other Considerations (1)

- LEAD is an NSF TeraGrid Gateway Project and through the TG, NSF has funded 2 LEAD FTEs at Indiana University
- NSF holds LEAD up as a notable success story, partly because LEAD encompasses almost the entire spectrum of cyberinfrastructure -- from instruments in the field (via its collaboration with CASA) to end users and education
- LEAD governance and relationship to Unidata's current governance
- Currently, LEAD funds 2.5 FTEs in Unidata; Additional technical and administrative staff will be needed to successfully manage the project.

Other Considerations (2)

- Operations and maintenance cost for the LEAD facility will need to estimated carefully. Unidata cannot do this out of hide.
- Has potential to connect Unidata to many NCAR divisions and their communities
- In the long term, LEAD technology can be leveraged for many other applications and communities (e.g. Climate Modeling, Coupled Hydrologic Prediction, Trajectory modeling)

Input from Rick Anthes, UCAR President

Mohan,

Thanks for sending me the slides on LEAD--what an exciting novel project!

You asked me what I thought about Unidata possibly taking on the management and leadership, working with the community, of LEAD after the current 5-year funding ends (a LEAD-2). I understand there are technical risks with doing this, but I don't think we should shy away from this possibility because of these risks. As I mentioned to you, we at UCAR/NCAR should be on the "bleeding edge" as Kelvin has called it in the past!

I do think in taking on LEAD-2 we must be sure additional adequate additional funding is provided. I agree completely with you that Unidata cannot take on this challenging program "out of hide." We can't let this new program put the core programs of Unidata at risk. With this caveat, you can be sure I support the concept and look forward to hearing about the Policy Committee discussion and its outcome.

Best wishes,
Rick