



G · A · B · B · S
geospatial data analysis building blocks

Project Overview & Opportunities

Carol Song, PI (NSF DIBBs)

Rosen Center for Advanced Computing

Purdue University

Brown Dog Workshop, July 23, 2014

Overarching goals

Making it easy for scientists to share geospatial data and tools

Reach broad user community

- Anyone can create an online app and share
- Anyone can share geospatial data

Building for self service (DIY) – Leverage successful software – Develop building blocks



Project goals

- Data and Tool are one
- Support geospatial data processing, analysis and visualization
 - Data services interface
 - Rapid tool creation APIs
 - Map and image renderers for online tools
 - Enabling geospatial data driven workflows
- All of these integrated with HUBzero core
 - Open source release
 - Hosting



Funding

- A National Science Foundation grant
- Data Infrastructure Building Blocks (DIBBs) program
- GABBs: 1 of 4 implementation awards in 2013
- \$4.5M, 10/2013 – 9/2017
- Collaboration with other awards



Team (11+)

Carol Song, PI

Larry Biehl (remote sensing, GIS)

Venkatesh Merwade (hydrology, Civil Eng)

Nelson Villoria (global geospatial data, Ag Econ)

Ed Lee (project manager)

Michael McLennan (HUBzero architect)

Rob Campbell (sr developer, tool development)

Leif Delgass (sr developer, visualization)

George Howlett (sr developer, RAPPTURE Toolkit)

Lan Zhao (research scientist, geospatial applications, data management)

Rajesh Kalyanam (GIS data processing, management)



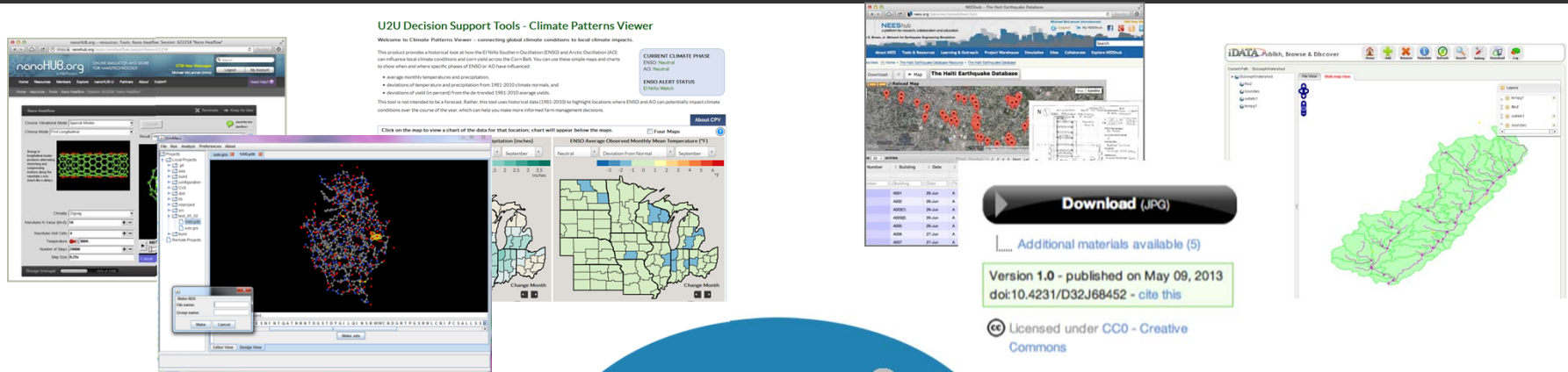
Geospatial hub projects

- Efforts in developing integrated geospatial data/modeling capabilities using HUBzero
 - Drinet hub (<http://drinet.hubzero.org>)
 - Geoshare hub (<http://geoshareproject.org>)
 - Water hub (<http://water-hub.org>)
 - Useful to Useable (u2u)
<http://drinet.hubzero.org/groups/u2u>

- Many hubs can make use of GABBs, such as NEES (network of earthquake engineering simulation), GENI (k-12 education), PURR (Purdue Research Repository), etc.

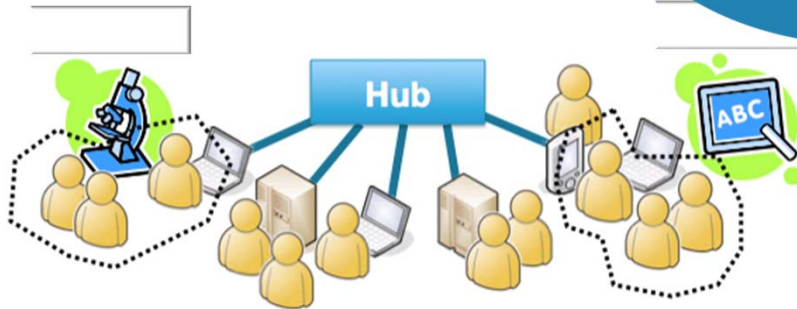


Platform for Scientific Collaboration



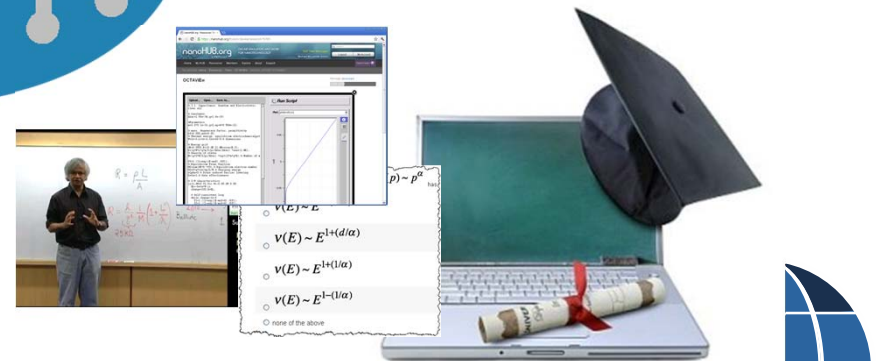
Computational Tools

Databases / Publications



Group/Project Collaboration

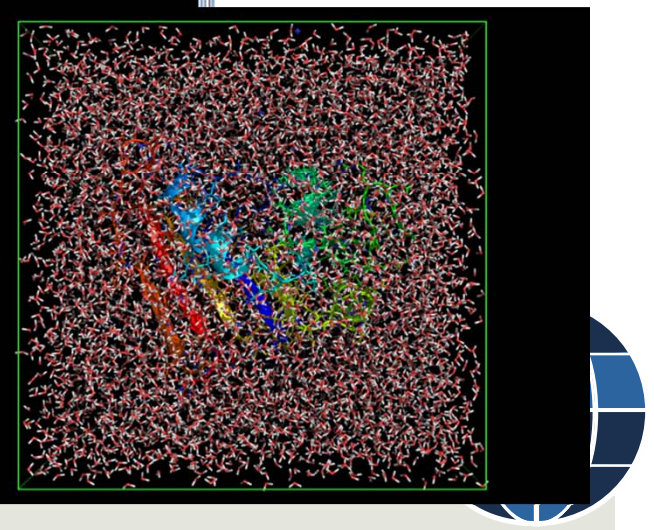
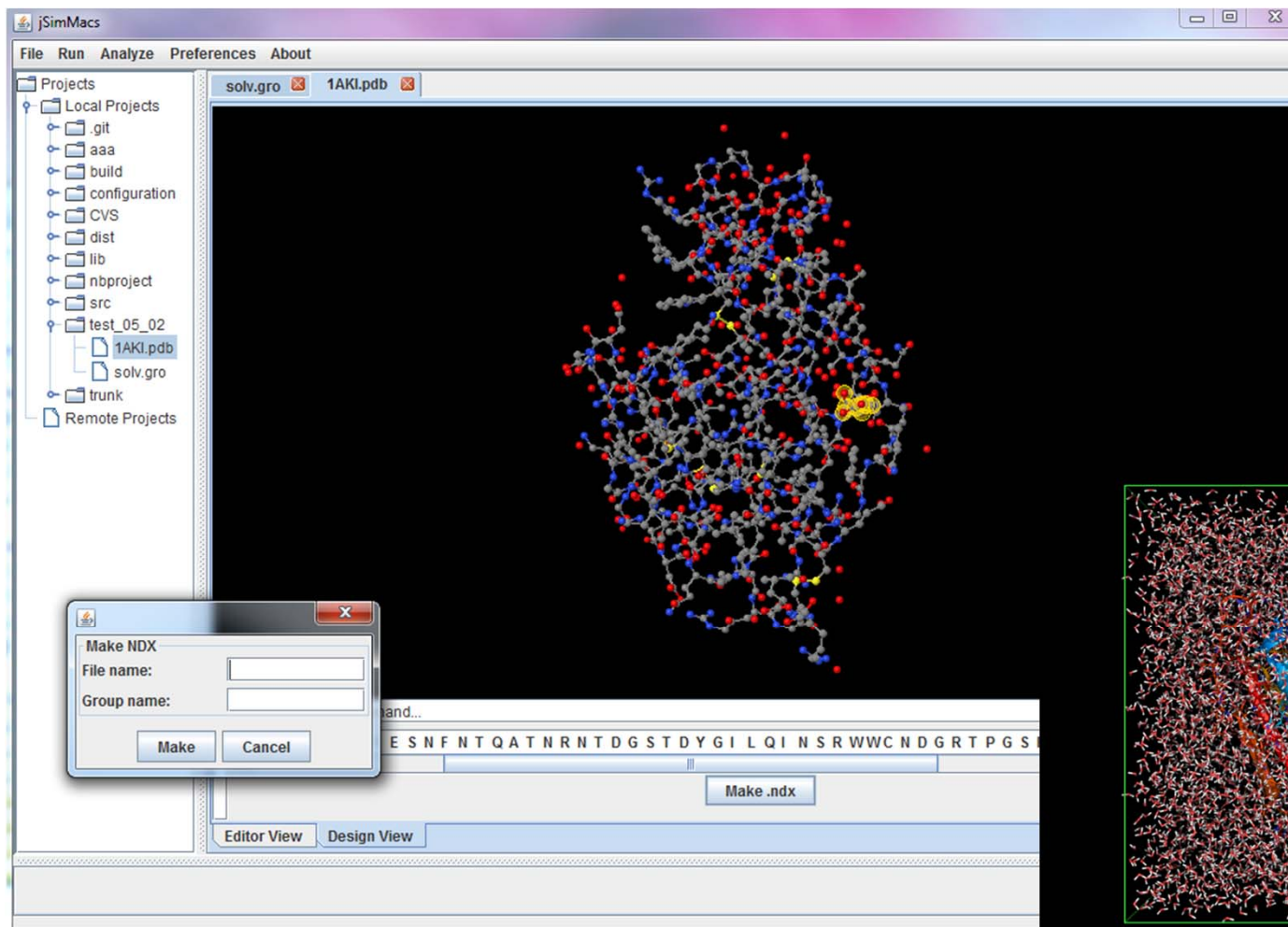
Courtesy of M. McLennan, Purdue University



Learning Management



Online tools



HUBzero for Scientific Collaboration

- An open source collaborative cyberinfrastructure for research and education
- Provides out-of-the-box support for developing web portals with content contribution, tool development workflows, user groups, wikis, ticketing systems, etc.
- Transparent access to large scale computation resources from online tools (no download, installation)
- Adopted by more than 50 science gateways cross many disciplines
- Serves more than 1,000,000 unique visitors in past 12 months
- Currently no integrated geospatial capabilities



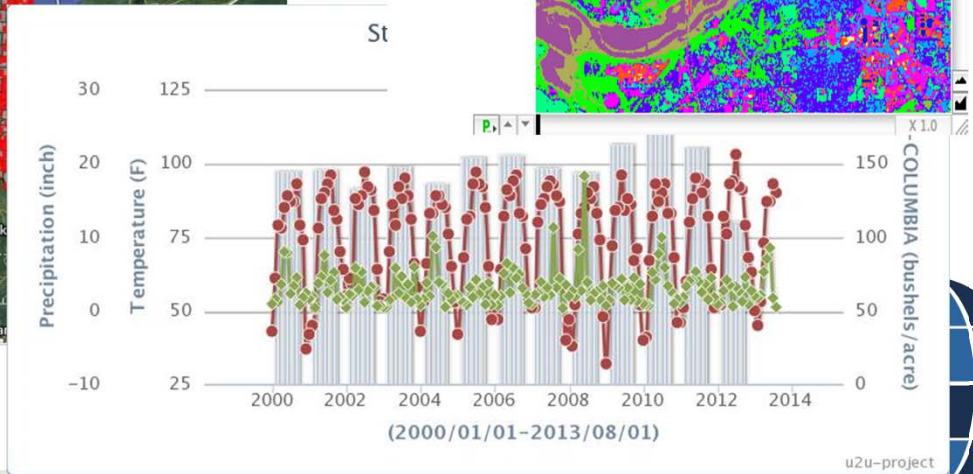
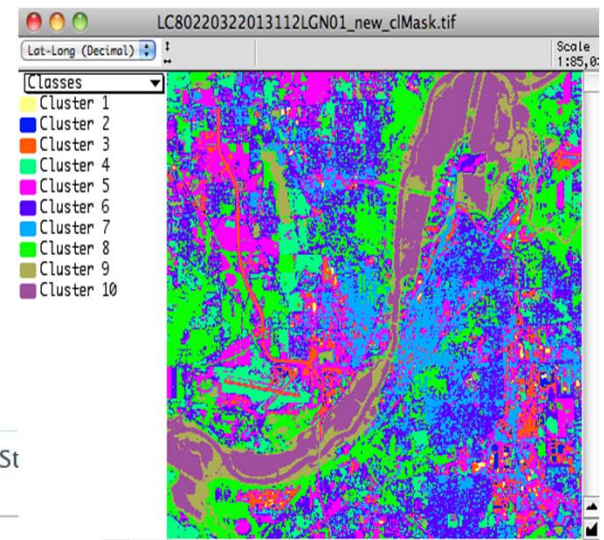
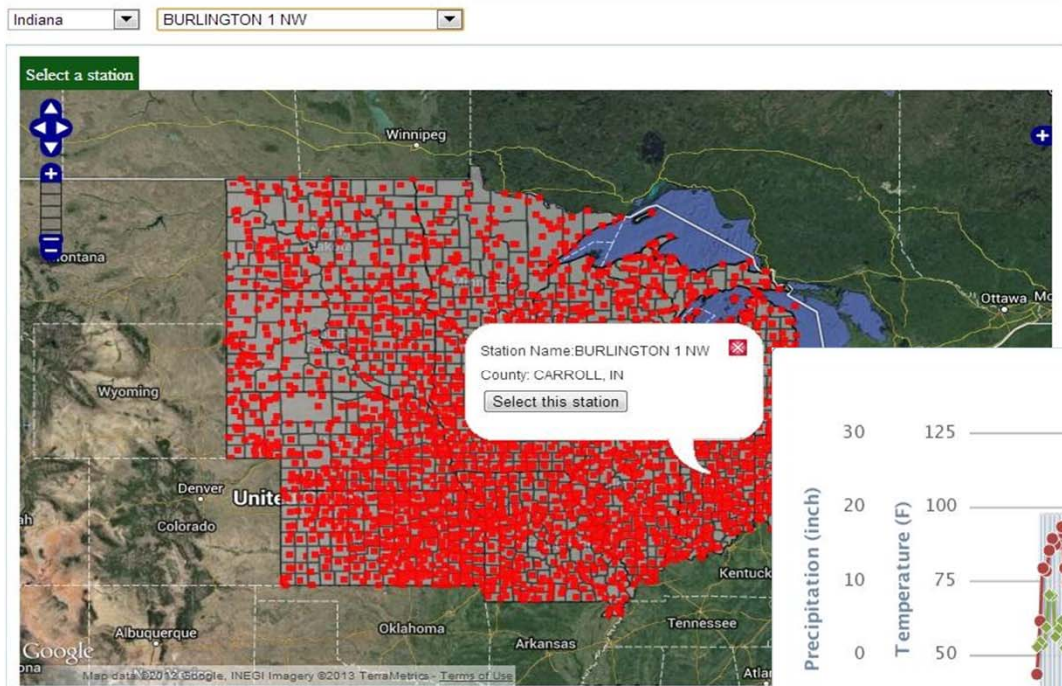
Driving Examples



Driving examples

- Easy deployment of geospatial tools

Please select a station from the drop down list or directly on the map.



Driving example

The screenshot displays the iDATA web application interface. At the top, the logo "iDATA" is followed by the text "Manage ge" and "Publish, Browse & Discover". A navigation bar contains icons for Home, Add, Remove, Metadata, Refresh, Search, Setting, Download, and Log. The main interface is divided into several sections:

- My Collections / Shared Collections:** A list of collections with columns for Name and Description. The "My" collection is expanded, showing a sub-collection "pegasus1" with a "File View" and "Raster Map View" option. A small map thumbnail is visible next to "pegasus1".
- Current Path:** "stJosephRiv2".
- File View:** A tree view showing the following items:
 - stJoseph
 - Riv2 (selected)
 - boundary
 - county83
 - outlets1
 - tempg1
- Map View:** A map showing a watershed area with a network of rivers and outlets. The map is overlaid with a grid. A "Layers" panel on the right lists the following layers:
 - tempg1
 - outlets1
 - boundary
 - Riv2
- Data View:** A panel for viewing data, currently empty.

At the bottom of the map view, there are two buttons: "Save map" and "Add a layer". The coordinates "-84.91148, 41.53824" are displayed at the bottom left of the map area.

Driving example

- Multi-scale and multi-disciplinary data and modeling for addressing hydrologic and ag economic issues

The screenshot displays the Pegasus Tool [1.0] interface. The main window shows a map of Iowa with a model selection window open. The map includes a toolbar with 'View', 'Upload', 'Edit', 'Run', and 'Visualization' options. The model selection window lists several models with their 'isShared' status:

Model name	isShared
Upperlowa_ca	Yes
Upperlowa_Se	Yes
Upper_lowa_R	Yes
Upper_lowa_R	No

The 'Upper_lowa_R' model is selected. A 'Property Value' window is also open, showing the following details:

Property	Value
User ID	tomotoso
Model Name	Upper_lowa_River_nea
Model Type	normal
Version	SWAT2009
HUC ID	07060002
Country	United States
State	IA

The interface also features a 'My Models' sidebar on the right, listing various models such as 'NorthRaccoon_Sensitivity', 'North_Raccoon_River_I', 'North_Raccoon_River_C', 'Sangamon_Calibrated', 'Sangamon_River_at_Fi', 'Sangamon_Sensitivity', 'SouthBranch_calibration', 'SouthForkSangamon_C', 'SouthFork_sensitivity', 'SOUTH_FORK_SANGAM', 'SpoonRiver_Sensitivity', 'SPOON_RIVER_AT_LON', 'SPOON_RIVER_AT_LON', 'SugarRiver_Sensitivity', 'SUGAR_RIVER_calibrati', 'SUGAR_RIVER_NEAR_BI', and 'testflatriver'. The bottom of the interface includes a 'Download File' button and a 'Download Archive' button, with a note that 'all files can also be downloaded.'

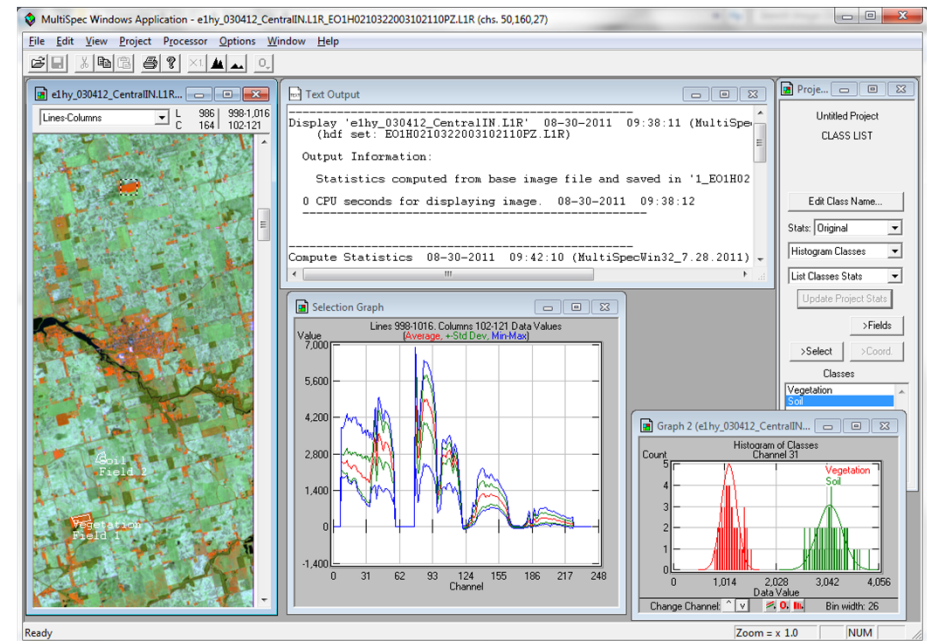
Starting point

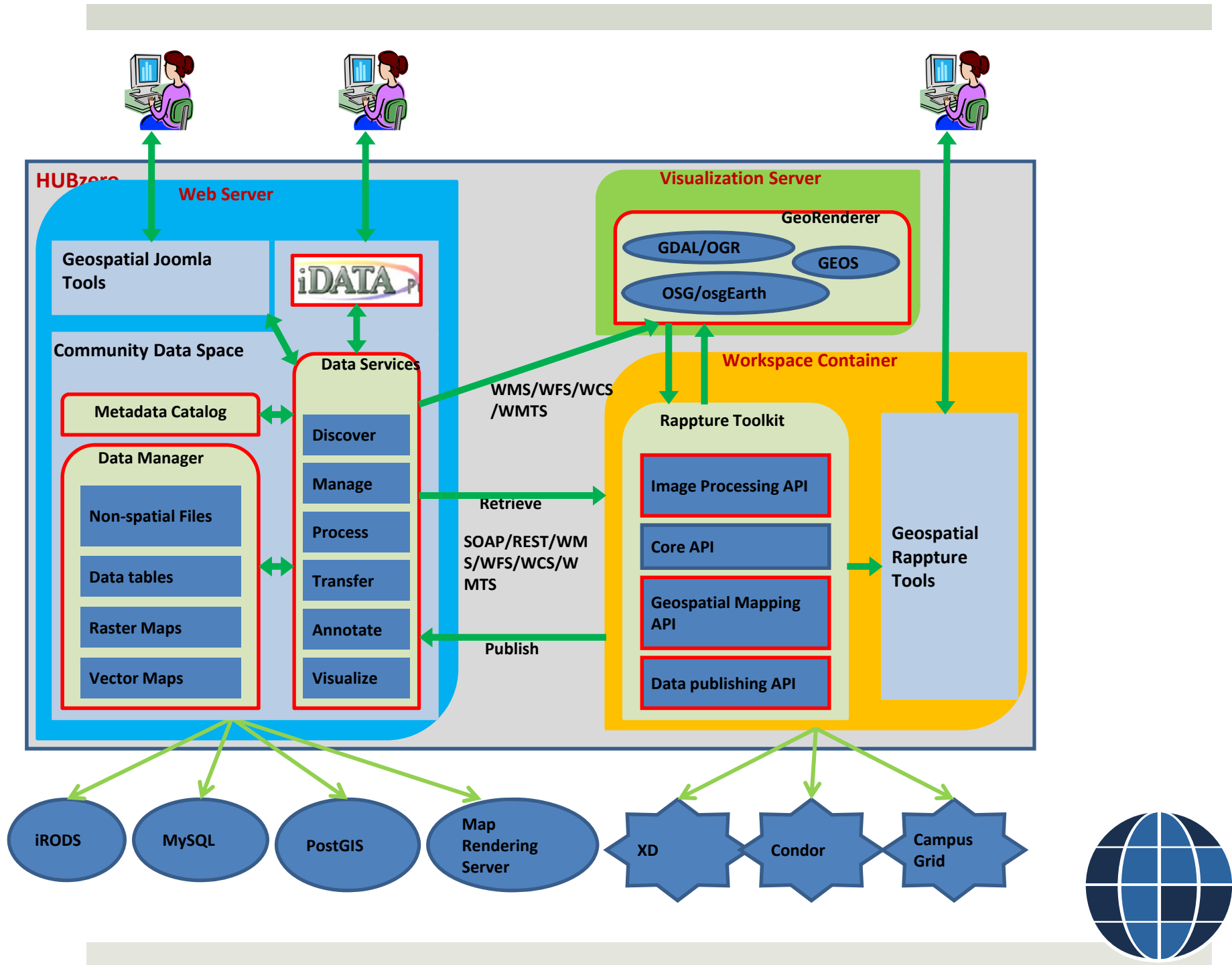
- **HUBzero** (Rappture, graphics rendering, collaborative web tools)
- **iData** (tool for self service data publishing and management)
- **Multispec** (tool for analyzing multispectral/hyperspectral image data)
- **Geospatial hub projects** (DRINET, Geoshare, WaterHUB, U2U etc)
- Leveraging software developed elsewhere
 - iRODS
 - Globus data transfer



Multispec

- Image data analysis software application
- Developed for interactive analysis of Earth observational multispectral images
- Processors for image displays, unsupervised & supervised classification and other feature extraction techniques
- Open source
- Used by research, higher education, k-12 learning and government institutions





Challenges

- ❑ Dealing with large data sets
- ❑ Seamless data/tool integration
- ❑ Map rendering in hub VM workspace
- ❑ Service interfaces
- ❑ Interfacing with other systems (Google drive, Dropbox, GIS servers)



Collaborations

- iRODS
- Globus
- DataOne
- Other DIBBs



More to come

- New hub: <http://MyGeoHub.org>
- HUBbub 2014 – annual HUBzero conference

