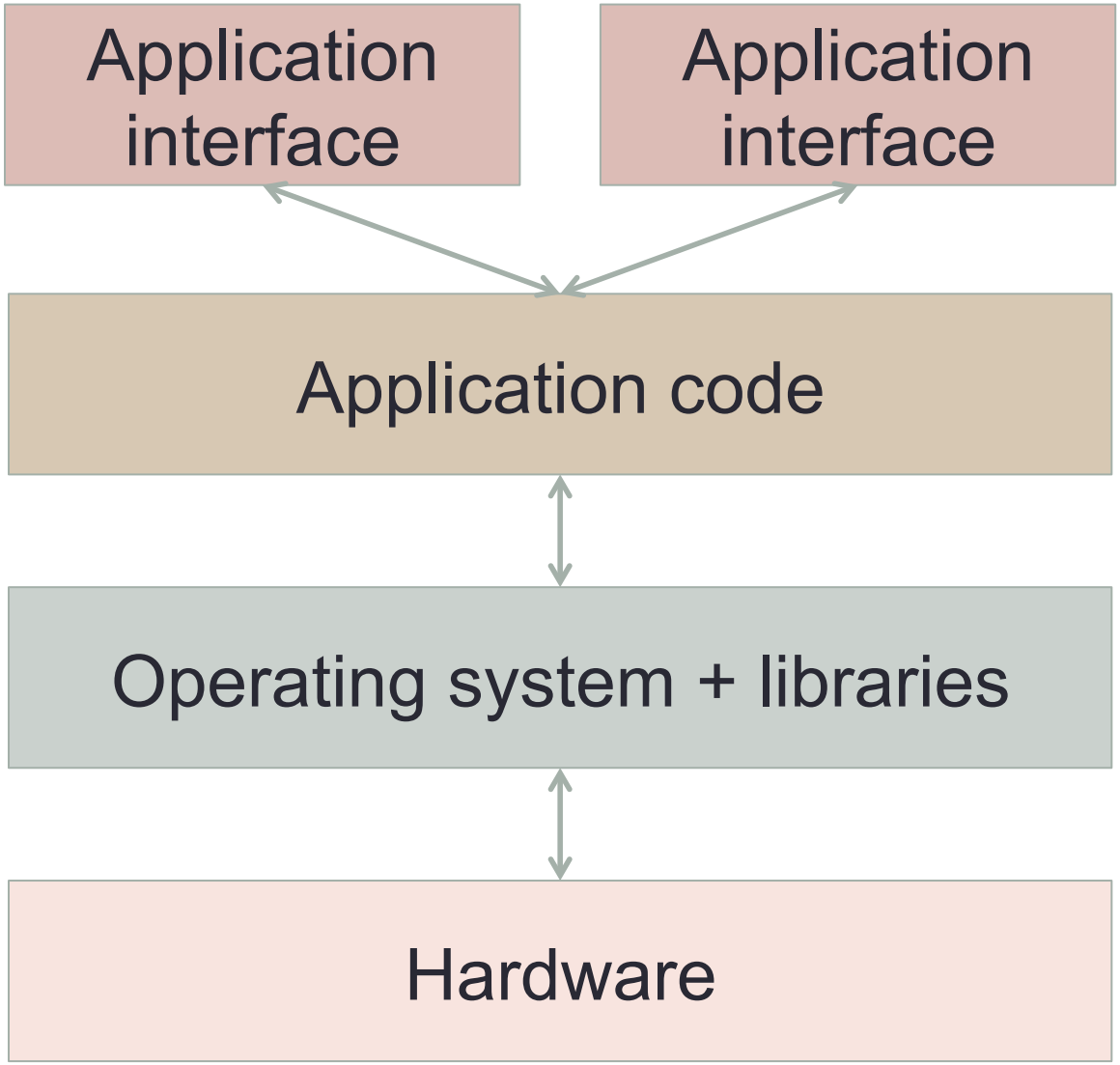
GEOL 5303 TOPIC: DATA COLLECTION

Dr. Deana Pennington, Associate Professor
Dept. of Geological Sciences
University of Texas at El Paso
Spring semester, 2017

Agenda

- Review of command line and GitHub (why are we doing this?)
- Schedule
- Real time data & robotics
- Sensor networks
- Internet of things (IoT)
- Hands on

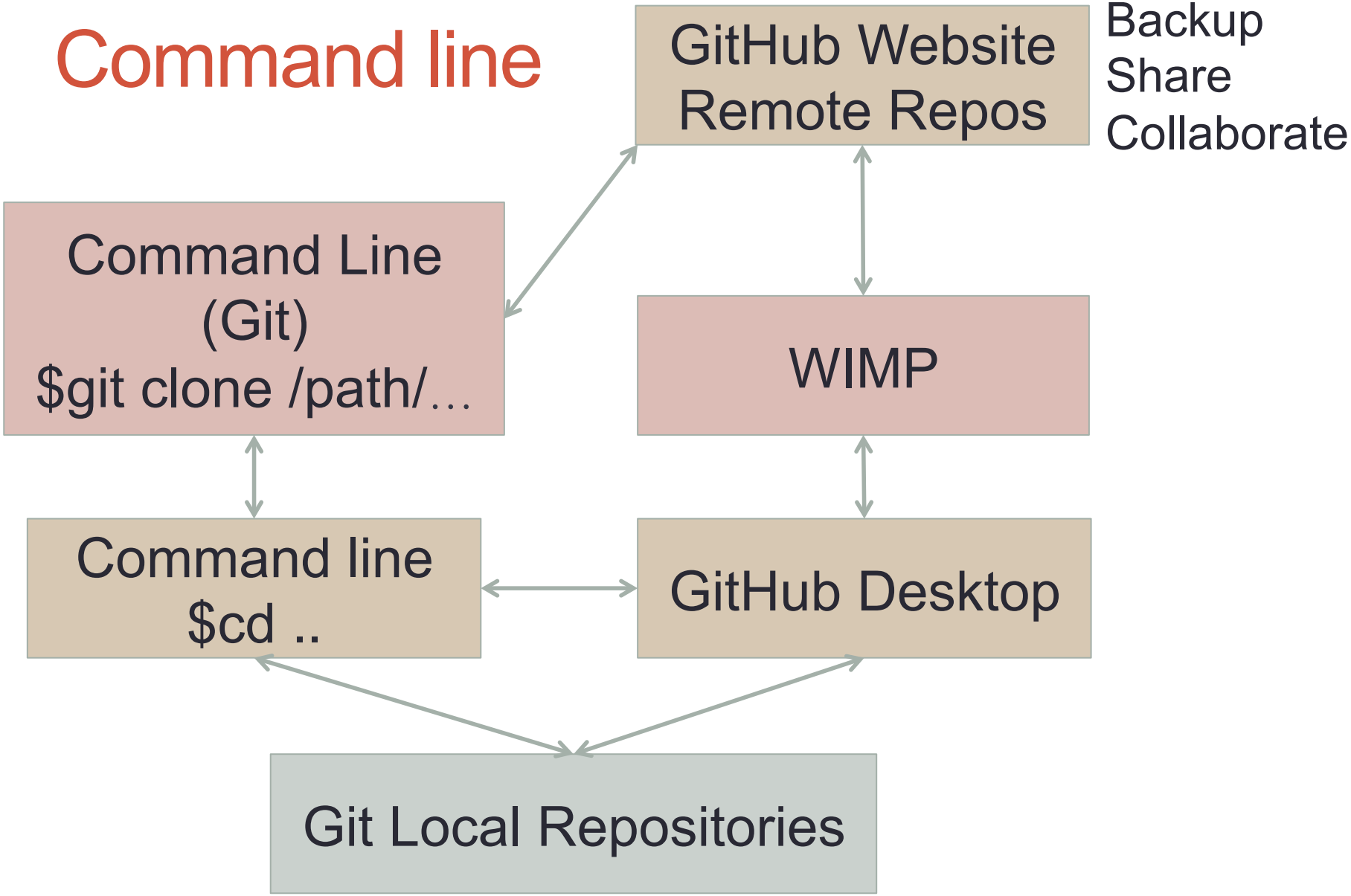
Computer applications



Types of interfaces

- **Command line interface**: Displays a prompt, the user types commands on the keyboard, the computer executes the command and provides textual output
- **Menu driving interface:** The user is given a list of items to choose from and selects from the list by typing a number into the keyboard or arrowing to the selection and entering the selection
- **Graphical user interface:** Uses windows, icons, menus and pointers (WIMP) that are manipulated by a mouse and/or keyboard
- **Natural language interface**: Instructions given through voice-activated command
- **Brain-machine interface:** Direct communication pathway between brain activity and machine

Command line



Why use command line instead of WIMP?

- POWER: Often provides more/different functionality
- SPEED: For repetitive tasks, much faster
- SCIENCE APPLICATIONS: Many widely used applications written by scientists are only available with command line

You may or may not ever need this, but it is a core skill for anyone who intends to work with emerging technologies

What will we use GitHub for?

Share information about new data collection and management techniques relevant to earth and environmental science, and/or ongoing earth and environmental science research using these techniques

- Who is involved?
- What are they working on?
- Where is it being applied? (institution and/or field site)
- When did their project start/end.
- Why is their work relevant? What problem are they addressing?
- How does it work technically, if you can figure this out and articulate it in a simple way.
- Glossary development

Collaboration in teams to compile this info for final presentations

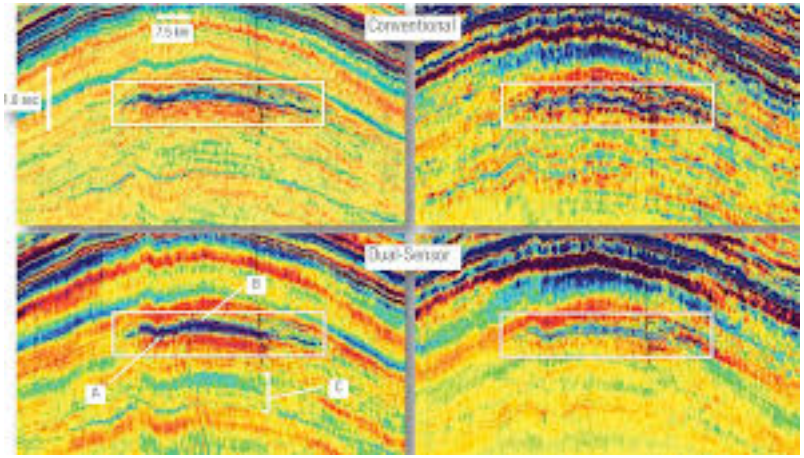
Schedule

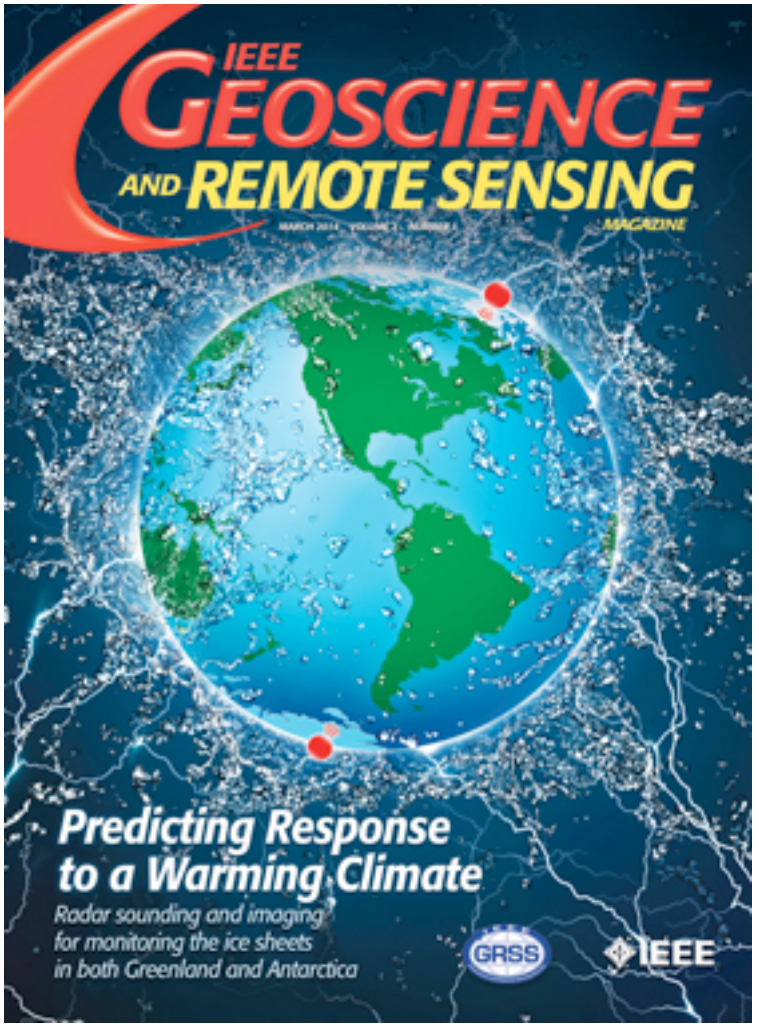
- Today: data collection
- Next Friday: data management
- Next Sunday midnight: first technical report due covering data collection and management
- Assignment is in Blackboard with instructions and a grading rubric (rules for grading)
- Goal: Survey of new data collection and management techniques relevant to earth and environmental science, and/or ongoing earth and environmental science research using these techniques

Agenda

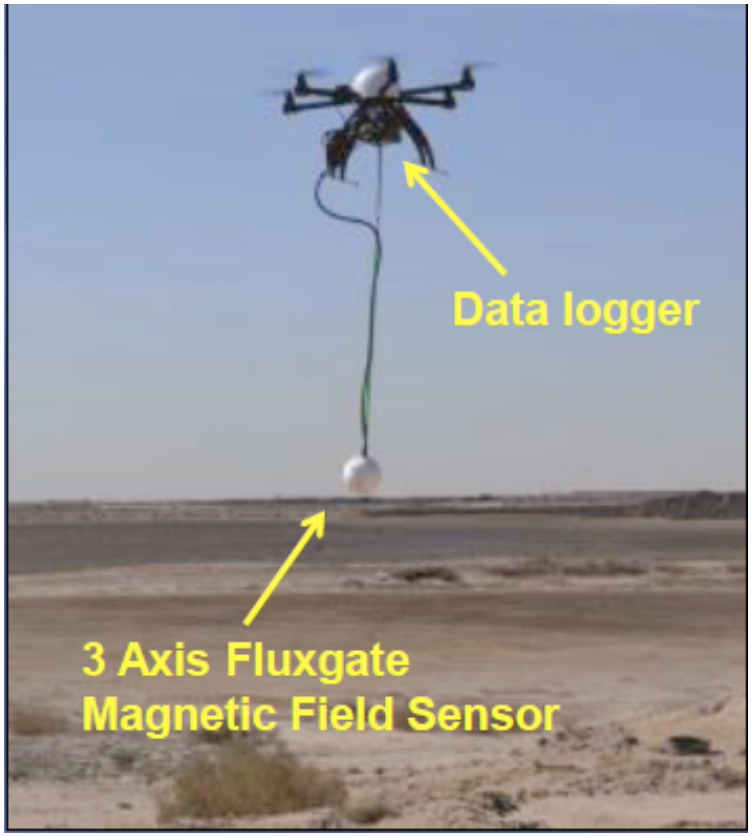
- Review of command line and GitHub (why are we doing this?)
- Schedule
- Real time data & robotics
- Sensor networks
- Internet of things (IoT)
- Hands on

Sensors, sensors, everywhere!

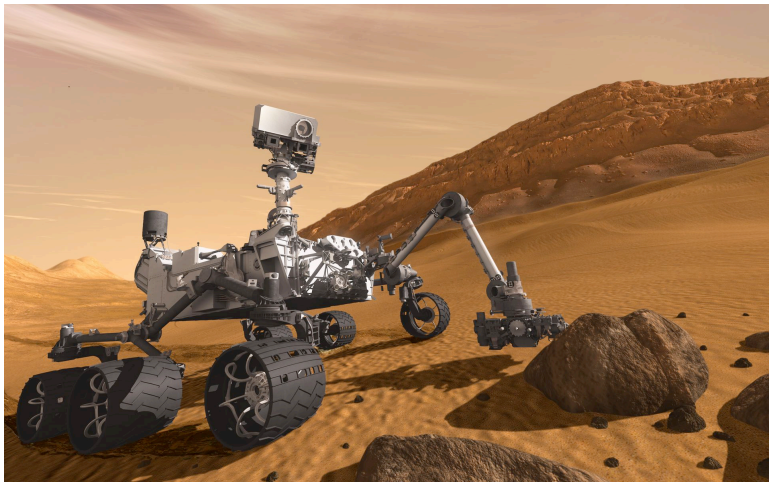




Drones



Other robotics



Collecting data in difficult places

EarthCube RCN: Intelligent systems in geosciences

Model-driven sensing

- Devices that “know” about the scientific context

CHORDS

- Cloud Hosted Real time Data Services
- Tools for creating a CHORDS portal
- Stream data into the portal
- Access data in other CHORDS portals

http://portal.chordsrt.com/

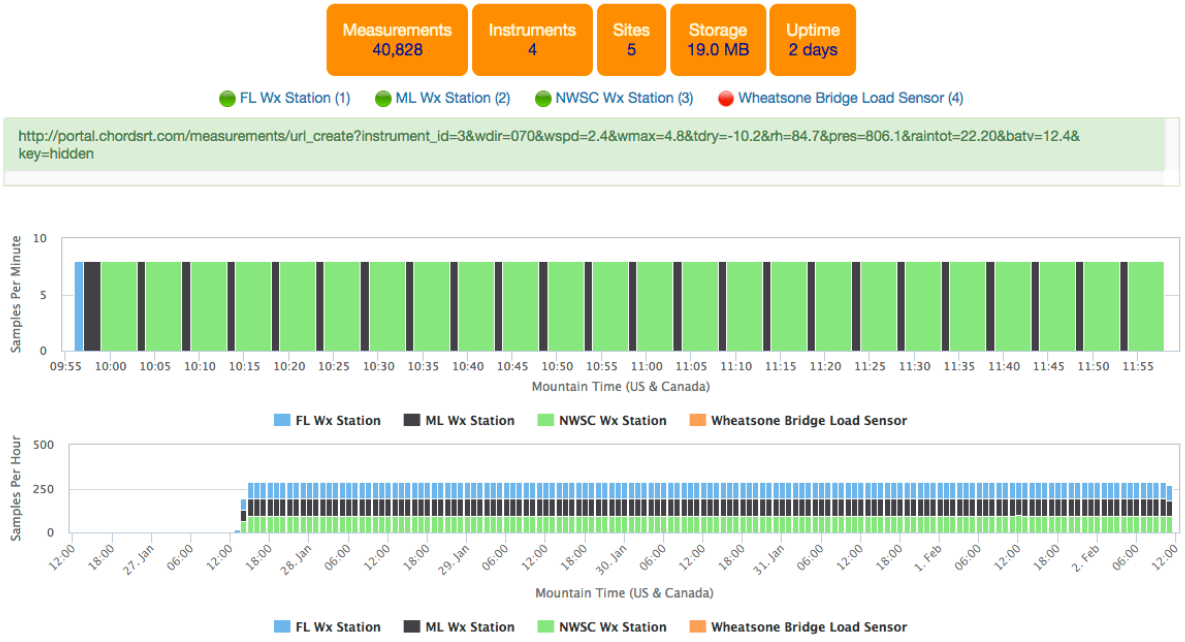
Data Functions

Download Data By Day

Start Date (start of day)

Instrument Name	Site
FL Wx Station	NCAR Foothills Lab
ML Wx Station	NCAR Mesa Lab
NWSC Wx Station	NCAR-Wyoming Supercomputing Center
Wheatsone Bridge Load Sensor	U.E.M.R.A.D.L

Dashboard



Sensor networks

- Multiple sensors communicate and act as one distributed instrument
- Events at one sensor trigger other sensor to respond (e.g. collect data more frequently)
- Event: Preset thresholds of measurements
- Event: Outlier identification
- Real time sensor steering based on real time data

ESIP EnviroSensing Cluster

- Earth Science Information Partnership

http://wiki.esipfed.org/index.php/EnviroSensing_Cluster

Document

- Introduction (Contact: Corinna Gries)
- Sensor Site and Platform Selection (Contact: Scotty Strachan and Adam Kennedy)
- Sensor Data Acquisition (Contact: Renee F. Brown)
- Sensor Management Tracking and Documentation (Contact: Corinna Gries)
- Sensor Data Management Middleware (Contact: Christine Laney and Branko Zdravkovic)
- Sensor Data Quality (Contact: Don Henshaw and Mary Martin)
- Sensor Data Archiving (Contact: Christopher Jones)
- Download complete document as of 1 Dec 2014 File:Sensor best practices 1Dec2014.pdf

Other Resources

- Cross-Domain Observational Metadata for Environmental Sensing - X-DOMES
- EnviroSensing Work Plan 2016
- Sensor Software Collaborations
- Workshops and Training Material
- Outside Links
- EnviroSensing Monthly telecons
- Monthly PDF archive
- Publications about Quality Control
- Publications about Sensors and Field Techniques

Contributors

Each chapter has a lead editor who is responsible for periodically compiling comments and contributions into stable versions of this document which will be archived as PDF versions and can be found here. If you contribute to this document by editing or adding text, images or comments you agree to the use of that material in the regularly published PDF versions of this document. Please add your name to the list of contributors if you feel you made a significant contribution.

- Renee F. Brown, Sevilleta LTER & UNM Sevilleta Field Station, rfbrown at sevilleta.unm.edu
- Richard Cary, Coweeta LTER, rcary1 at uga.edu
- Jason Downing, Bonanza Creek LTER, jpdowning at alaska.edu
- Corinna Gries, North Temperate Lakes LTER, cgries at wisc.edu
- Don Henshaw, Andrews LTER, don.henshaw at oregonstate.edu
- Christopher Jones, National Center for Ecological Analysis and Synthesis, cjones at nceas.ucsb.edu
- Adam Kennedy, Andrews Forest LTER, adam.kennedy at oregonstate.edu
- Christine Laney, University of Texas at El Paso, Jornada Basin LTER, cmlaney at utep.edu
- Mary Martin, Hubbard Brook LTER, mary.martin at unh.edu
- Jennifer Morse, Niwot Ridge LTER, jennifer.f.morse at colorado.edu
- John Porter, Virginia Coast Reserve LTER, jporter at virginia.edu
- Jordan Read, US Geological Survey, jread at usgs.gov
- Andrew Rettig, University of Cincinnati, andrewrettig at gmail.com
- Wade Sheldon, Georgia Coastal Ecosystems LTER, sheldon at uga.edu
- Scotty Strachan, University of Nevada, Reno, scotty at dayhike.net
- Branko Zdravkovic, University of Saskatchewan, branko.zdravkovic at usask.ca
- Fox Peterson, Andrews LTER, fox.peterson at tinybike.net

Citizen science

- People as “sensors”
- Anytime/anywhere data collection
- Mobile apps

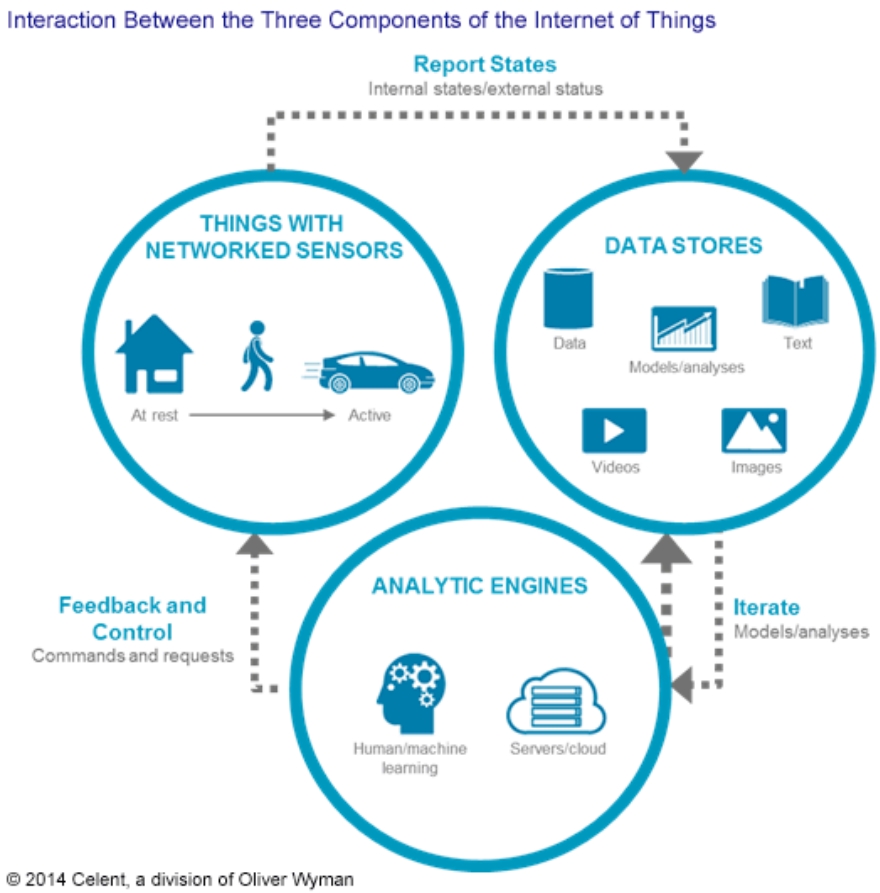
Example: Tweedie lab

- Stephen Escarzaga slides

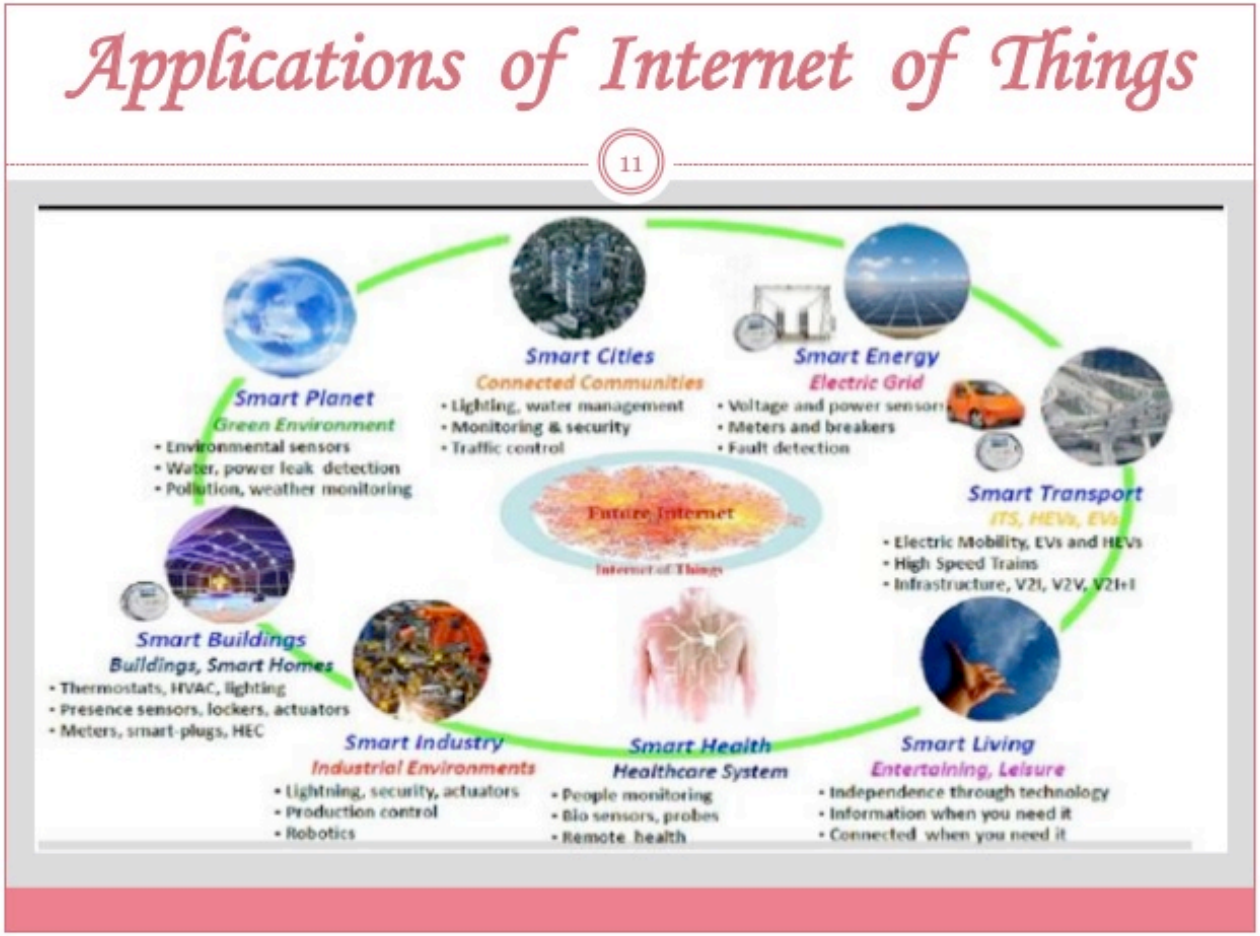
X-DOMES

- Cross Domain Observational Metadata Environmental Sensing Network
- Automatically generate metadata (data about data) associated with sensor data in real time
- Fully described data, including “provenance”
- Provenance: source, processing and QA/QC
- Metadata + provenance => “trust”
- Working with sensor manufacturers

Internet of Things (IoT)



Applications: “Smart” X

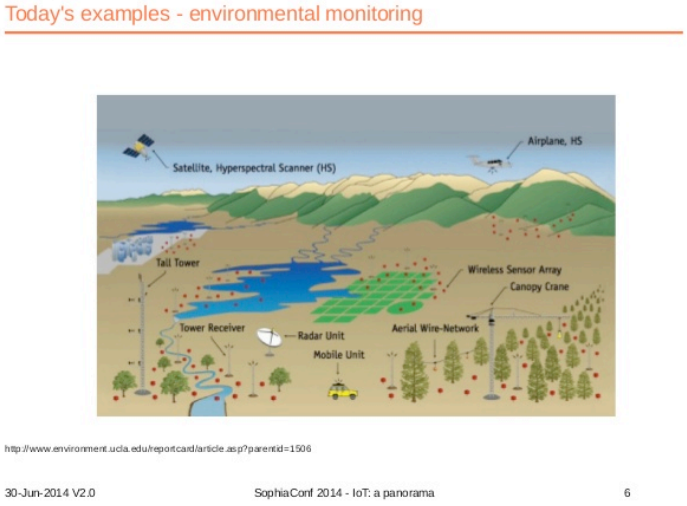


Examples

SmartLab
http://www.labvolution.de/en/conferences-events/themenschwerpunkte/smartlab/

- Monitor instruments
- Stream data to electronic lab notebooks
- Track device usage and workload
- Control experimental workflows

TetraScience
http://tetrascience.com/



iSamples

- Connect physical samples and sample collections across the Earth Sciences with digital data infrastructures
- Improve discovery, access, sharing, analysis, and curation of physical samples
- Facilitate the management of samples, sample collections, and sample-based data in the field, in the lab, in repositories, in data systems, and in scientific publications

Sample registry

http://www.geosamples.org/

get your igsn
Register your samples with SESAR to obtain IGSNs for unique sample identification.

search the catalog
Search the SESAR catalog to find registered samples and their current location.

sample curation
Learn about the DESC initiative to build a Digital Environment for Sample Curation.

interoperability
Access IGSN metadata profiles and register samples via web services.

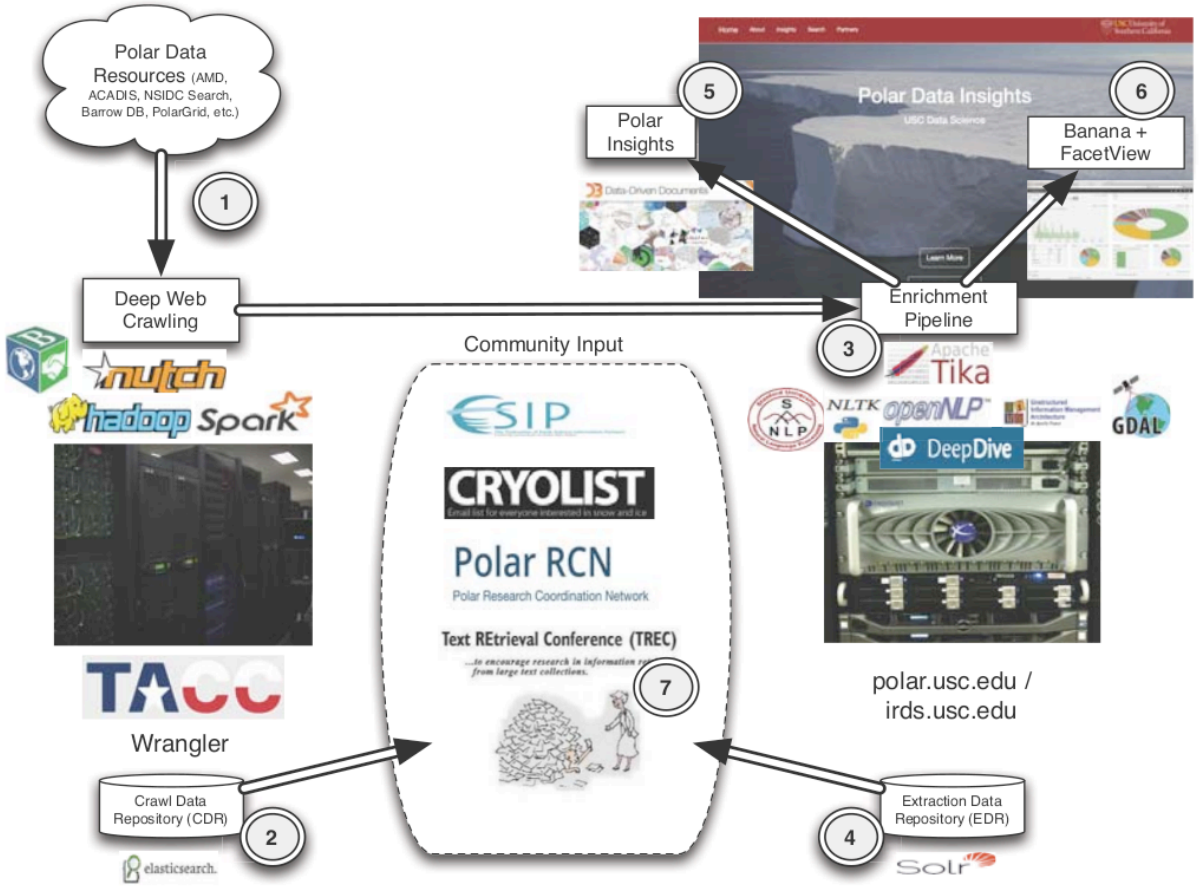
new user?
Get a MySESAR account to register your samples.

- Alphanumeric code - Unique identifier
- Unique URI (web location) => sample is a “thing” on the internet
- Rocks, water, cores, sites, stratigraphic sections

IGSN: IECUR0002 (Registered object: Metamorphic rock powder from Australia, registered by Curtin University)

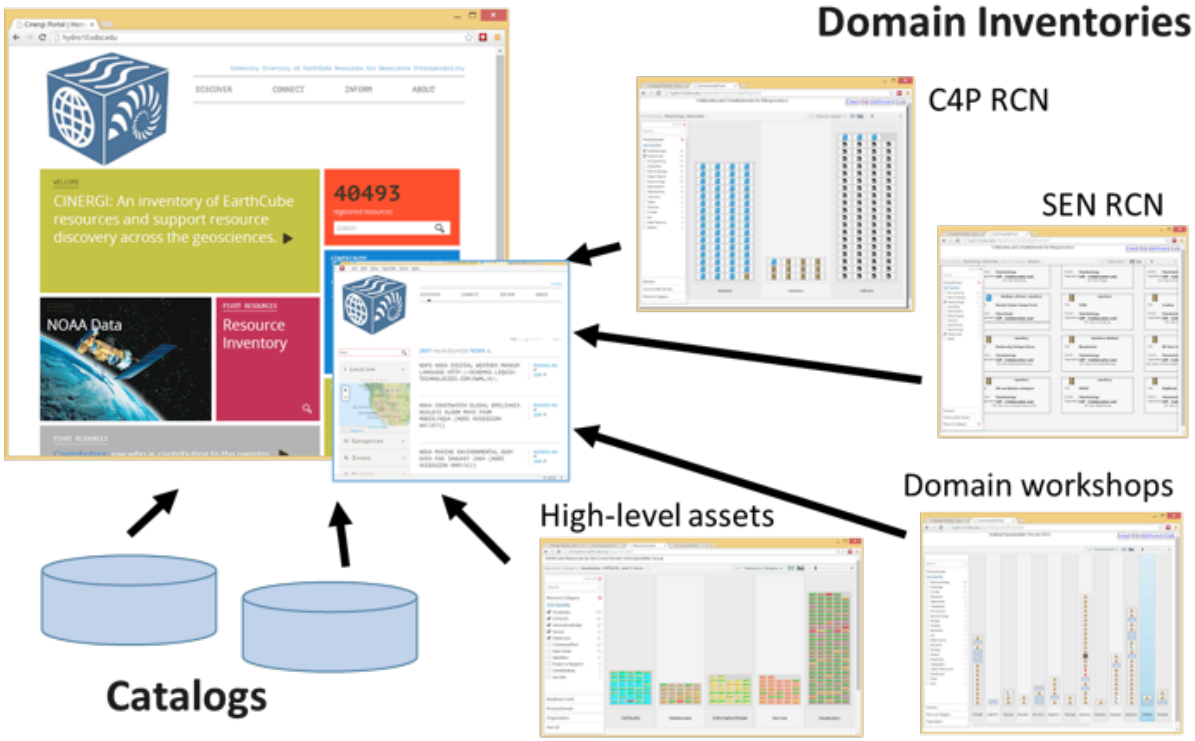
Polar Data: Deep Scientific Web

- Collect, analyze, and make interactive
- Text, data, journals, multimedia, web pages



CINERGI

- Community Inventory of EarthCube Resources for Geosciences Interoperability
- Linking geoscience resources, users, publications, usage information



CINERGI Resource Viewer

- http://pivots.azurewebsites.net/ecogeo.html

https://www.earthcube.org/workspace/cinergi/cinergi-resource-viewer-overview

Step 1: Create an inventory using the CINERGI template
Step 2: CINERGI will set up a viewer
Step 3: Community editing and curation

CINERGI Resource Inventory

Required

Resource Name
Short name of the resource, but spell out acronyms
e.g. NCALM - National Center for Airborne Laser Mapping

CINERGI Resource Type
The Resource Type from the CINERGI list.
See list in a google spreadsheet: http://goo.gl/e99KRj
Common resource types are: Organization, Webpage, Colle

Classification (Theme)
The classification or theme is a term that the community wo
e.g. topography, climate, hydrology, laboratory

Resource Description

Few sentences that informs the user what is the resource ar

Resource URL
URL to the resource. In the case that the URL does not exist
e.g. http://calm.geo.berkeley.edu/ncalm/ddc.html

Optional

- Resource steward person, organization, email, phone
- Geographic keywords
- Link Lable and Link function
- Any other ISO Metadata fields

Spreadsheet in GitHub
Fork for each team
Merge into master

Remaining time today

- Create teams. Choose a name for your team
- Create a project board for your team
- Assign known resources to team members
- One team member fork CINERGI template
- Add team members to the team fork
- Each team member branch the team fork
- Start adding resources to the inventory & to the wiki
- Each time you stop working, commit and submit pull request
- Team lead: approve pull requests
- Team lead: Next Thursday commit to master and pull request