

CZOData Cyber-Seminar #3: 2014-03-20

Protocols for sharing CZOData: Part 2: Registering data via CriticalZone.org

David Lubinski, UC-Boulder.

Tom Whitenack, SDSC.

Anthony K. Aufdenkampe, Stroud Water Research Center

Kerstin Lehnert, IEDA/Columbia.

Ilya Zaslavsky, SDSC.

Jeff Horsburgh, USU.

Emilio Mayorga, UW-APL

David Tarboton, USU



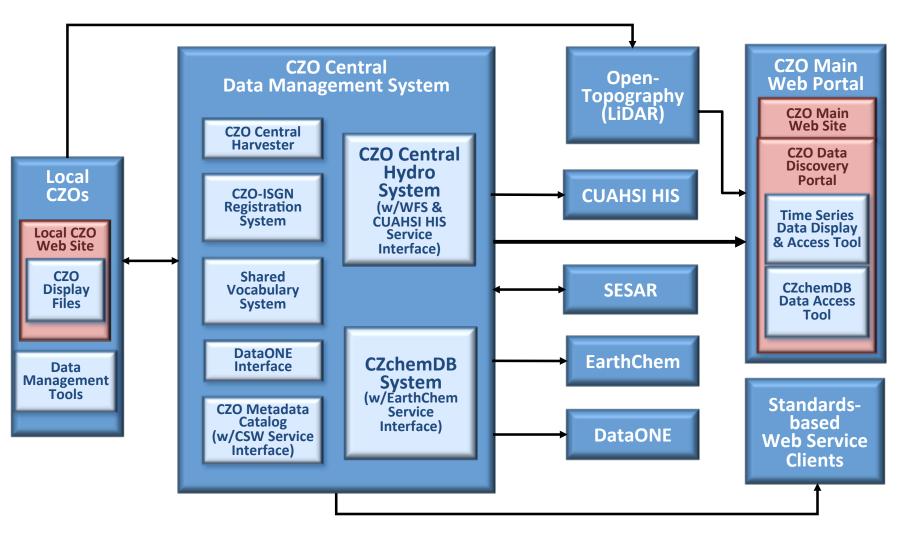








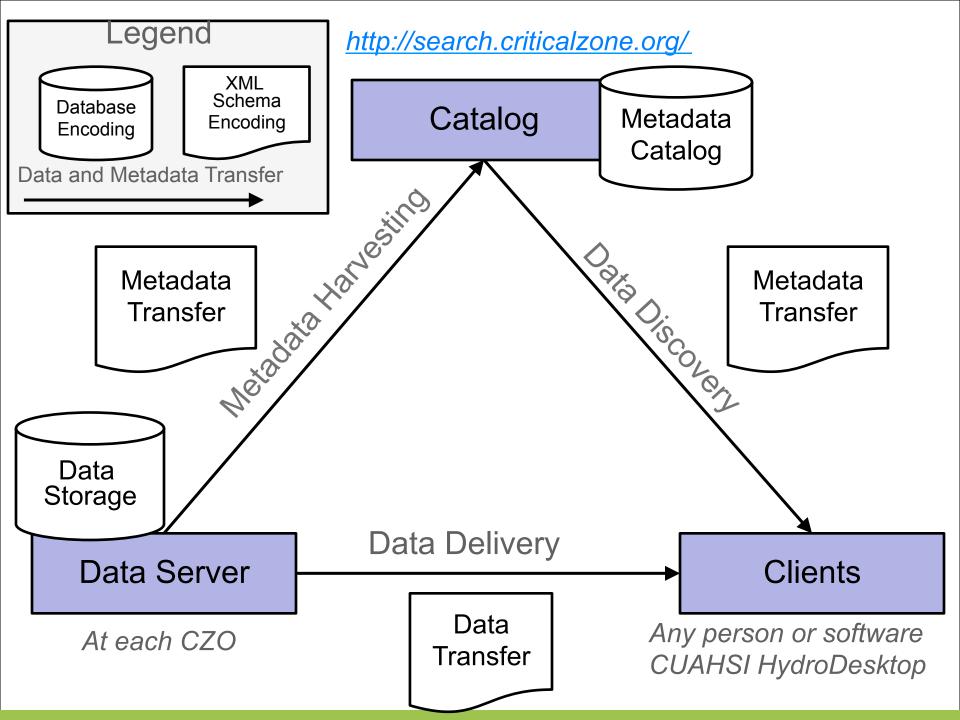
CZOData II Architecture



Local CZOs

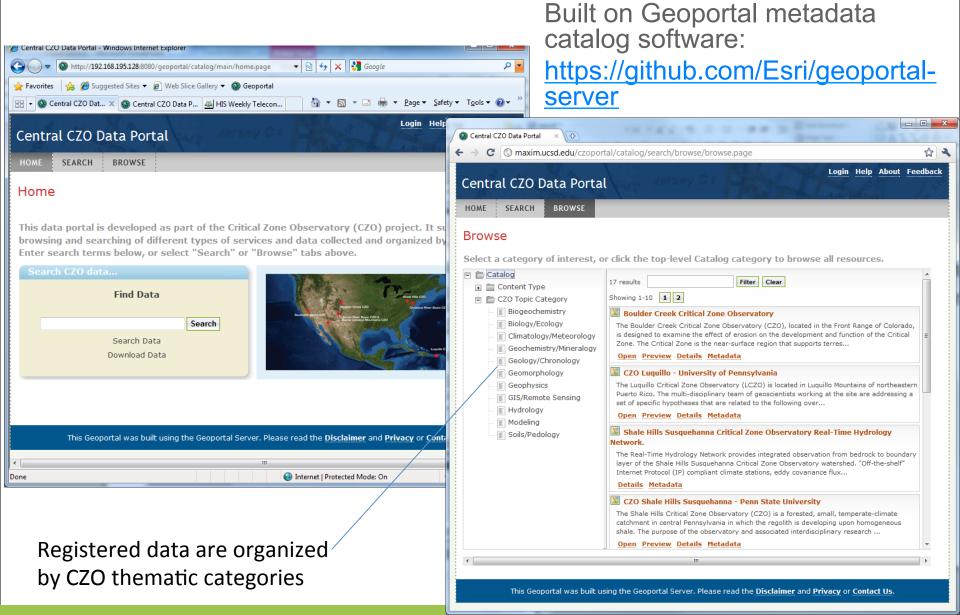
CZO Central Coordination Functions CZO Central Data Repositories Non-CZO
Integrated Data
& Discovery Sites

Clients

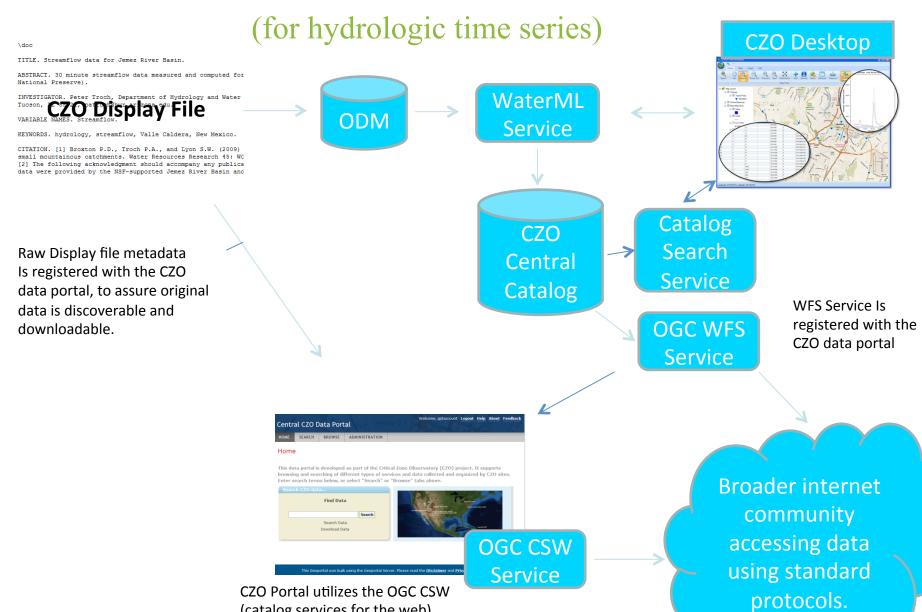


CZO Data Portal

At http://search.criticalzone.org/



Data Publication Process



(catalog services for the web)

TSS concentrations and elemental/isotopic composition (starting 2005) of baseflow and stormflow.

Variables: Solids_total suspended, carbon to nitrogen molar ratio, carbon_ particulate organic, nitrogen_ particulate organic, nitrogen-15 stable isotope ratio delta

Date Range: (1993-2012)

Dataset Creators/Authors: Aufdenkampe, A.K.; Newbold, J.D.; Anderson, B. A.; Richardson, D.; Damiano, S.G.

Contact: Sara Geleskie Damiano, Stroud Water Research Center, 970 Spencer Road, Avondale, PA 19311,

sgeleskie@stroudcenter.org

Field Area: Boulton Run | Christina River Basin | Forest Endmember: Spring Brook | White Clay Creek @ SWRC | Construction Endmember: White Clay Creek below landfill | Lower White Clay Creek | Agricultural

Endmember: South Branch Doe Run

● Water Chemistry

■ Geomorphology

□ Biogeochemistry

➤ Hydrology

■ Geochemistry / Mineralogy

Christina

Description Keywords Citation Publications

Acknowledgements

Description

Total suspended solids (TSS) and Volatile Suspended Solids (VSS) from White Clay Creek near the Stroud Water Research Center, Avondale, PA, USA. The purpose is to quantify export of inorganic and organic particulate matter from the 725-hectare watershed. Samples consist of those taken at monthly intervals, normally the first Wednesday of each month regardless of weather or flow conditions and those taken after precipitation events. The monthly samples are manual grab samples collected in 5-L polyethylene "space saver" bottles from a few centimeters below the surface and without disturbance to the stream bed. The event samples were collected in response to precipitation of 20 mm or more using an ISCO automated sampler which collected 1-L samples s in polyethylene bottles at hourly intervals through an intake approximately 20 cm above the bed. Each of approximately four events per year are represented by approximately 10 samples selected from the hourly series to characterize the rise, peak, and falling limb of the hydrograph. Additional events are represented by the three samples nearest peak flow.

Data

Christina River Basin - Stream Suspended Sediment 1993

_

CZO Field Areas







Christina River Basin



Forest Endmember: Spring Brook





Entering Dataset Metadata into the CMS

(CriticalZone.org)

&

Automatic registration of CMS Dataset Metadata into Geoportal

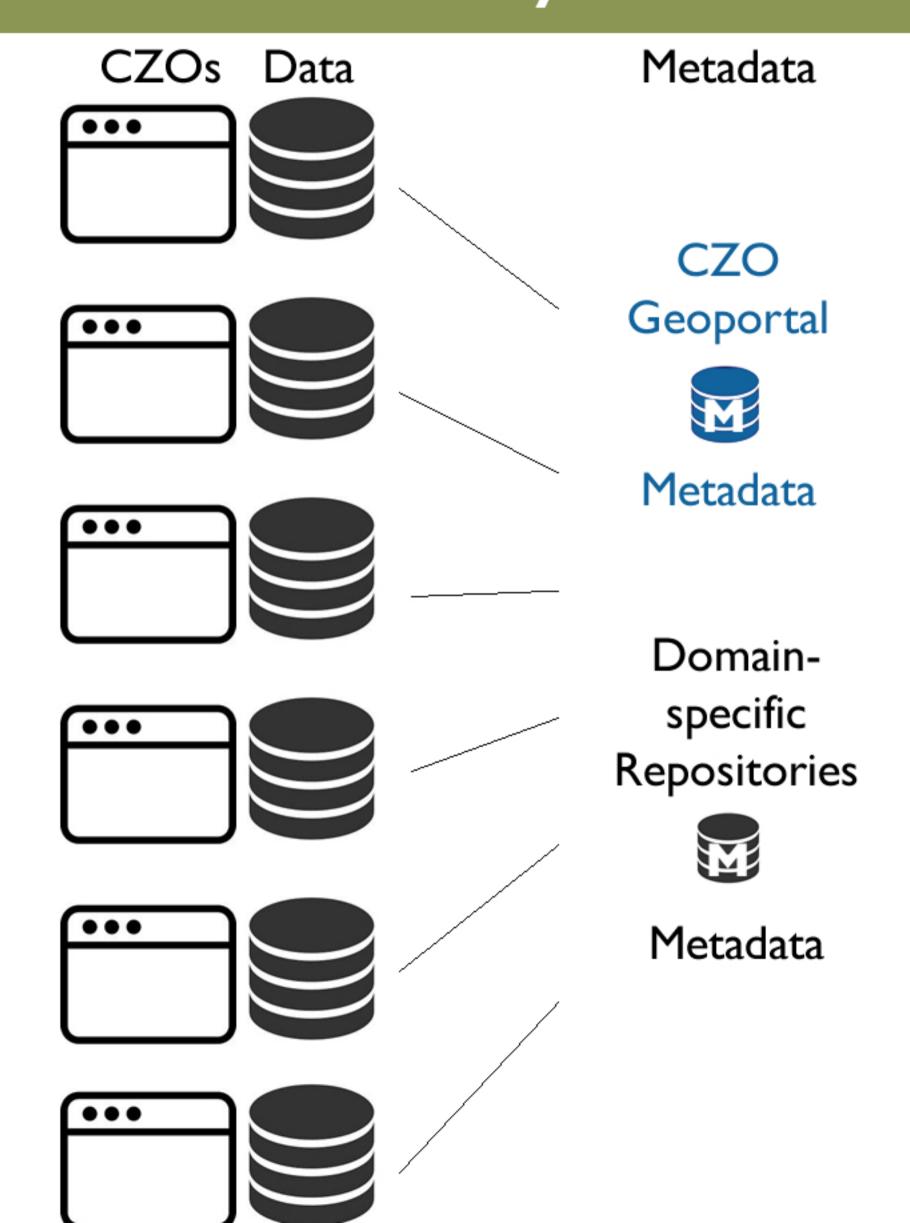
(Search.CriticalZone.org)



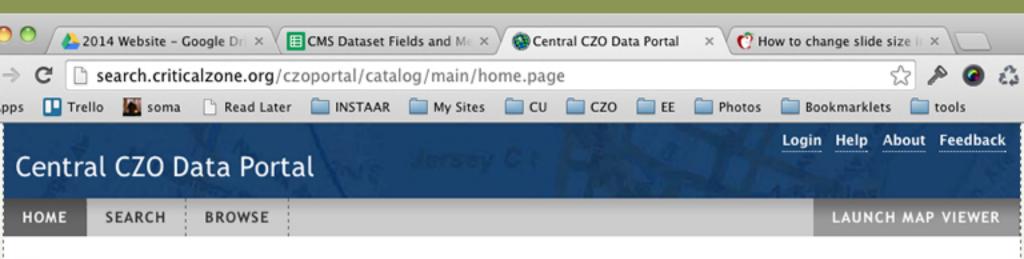
David Lubinski CZO "Webmaster"

INSTAAR CU-Boulder

A year ago... Only a small fraction of metadata widely available



CZO Geoportal



Home

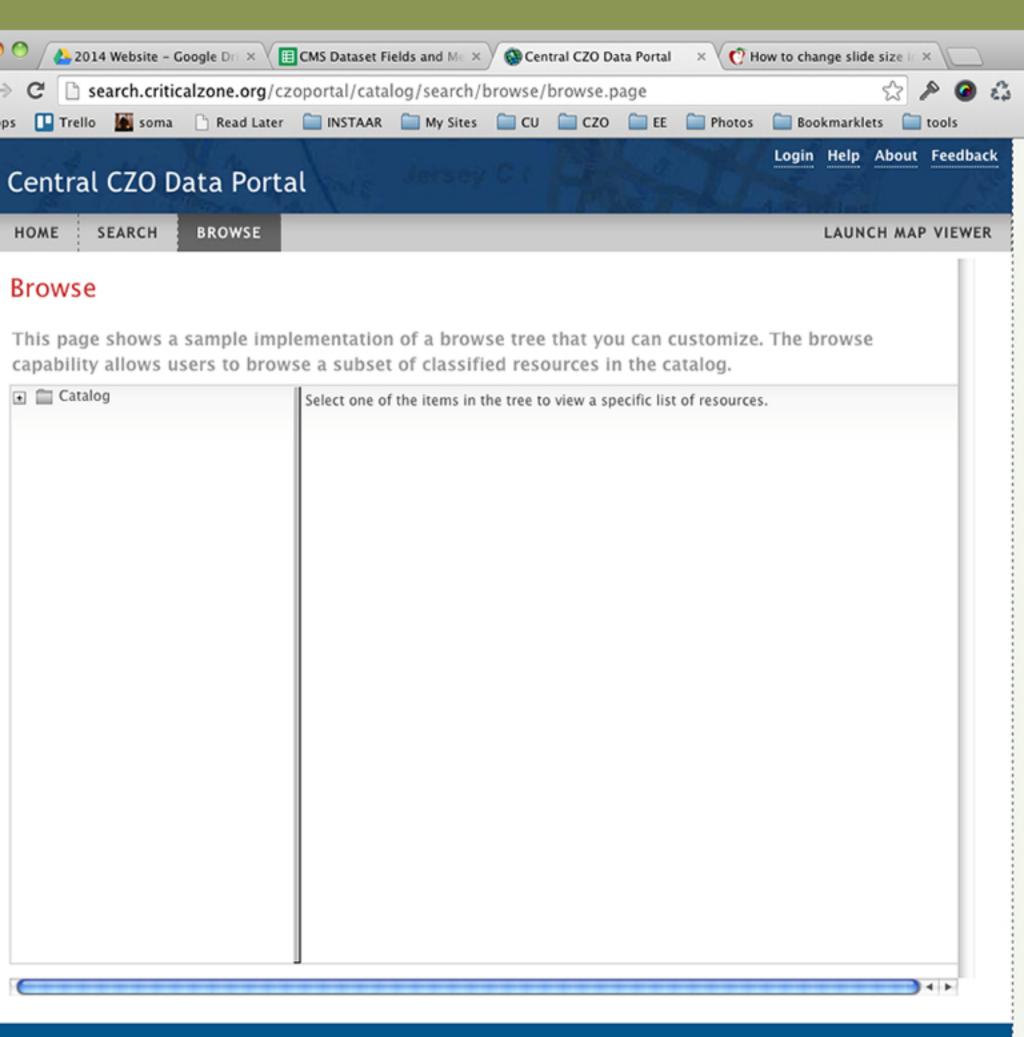
This data portal is developed as part of the Critical Zone Observatory (CZO) project. It supports browsing and searching of different types of services and data collected and organized by CZO sites. Enter search terms below, or select "Search" or "Browse" tabs above.





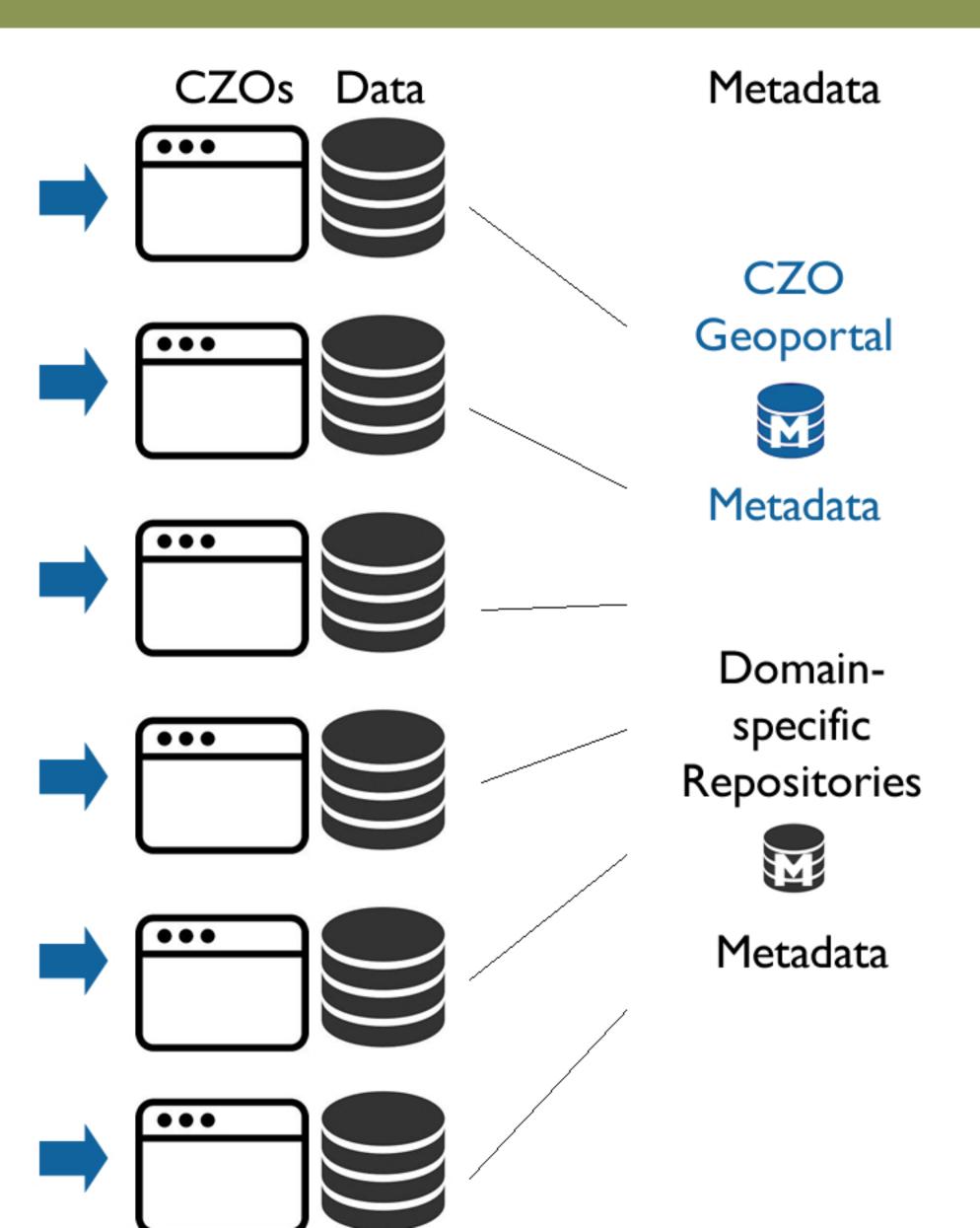
This Geoportal was built using the Geoportal Server. Please read the Disclaimer and Privacy or Contact Us.

CZO Geoportal



This Geoportal was built using the Geoportal Server. Please read the Disclaimer and Privacy or Contact Us.

Even more data...



Triple whammy of Difficult Data Discovery

Workflow/Training/Capabilities

Default Geoportal UI

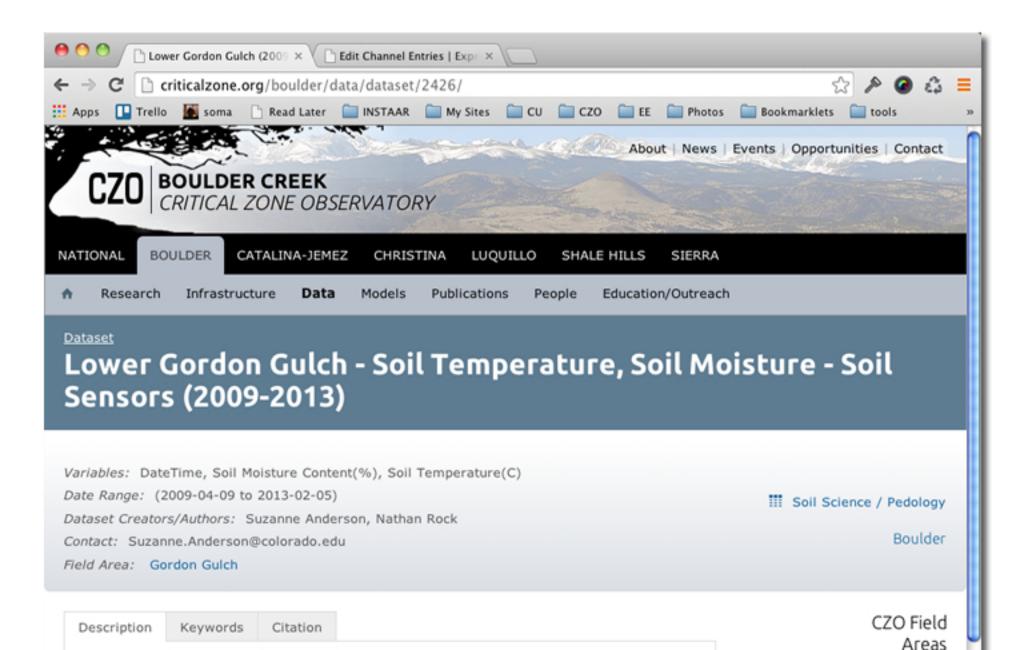
More data coming in...



Solution: Use the CMS

Dec 2012 Meeting at Stroud

Winter-Spring 2013
Collaborative Design
Development



179 Datasets entered

Each dataset consists of I-30 components that point to individual online data resources.

1018 dataset component resources are listed for the 179 datasets.

These resources are all online, mostly on individual CZO web servers.

Data Demez River Basin - Flux Tower data (.jsp) Data Level 1 Demez River Basin - Flux Tower data 2007 (.csv) Data Level 1, Metadata Demez River Basin - Flux Tower data gap filled 2007 (.csv) Data Level 2, Metadata Demez River Basin - Flux Tower data 2008 (.csv) Data Level 1, Metadata Demez River Basin - Flux Tower data gap filled 2008 (.csv) Data Level 2, Metadata

CZOs Data









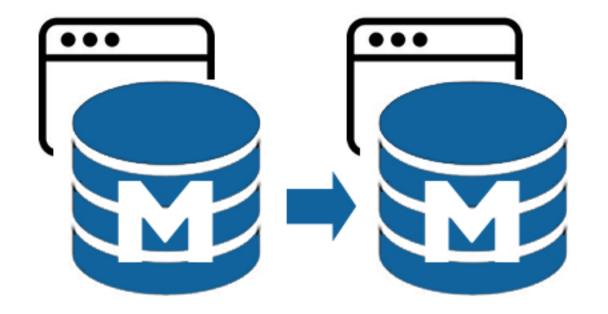




One-way metadata sync

CMS metadata

CZO Geoportal metadata



CZOs Data







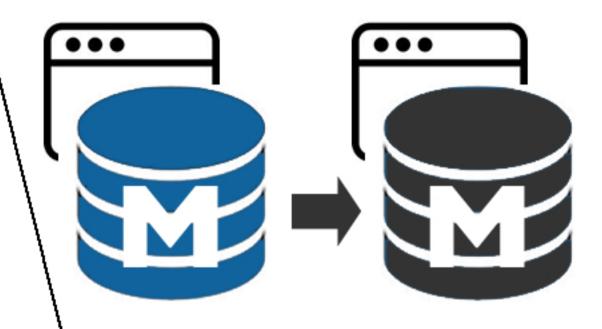






One-way metadata sync

CMS metadata CZO Geoportal metadata

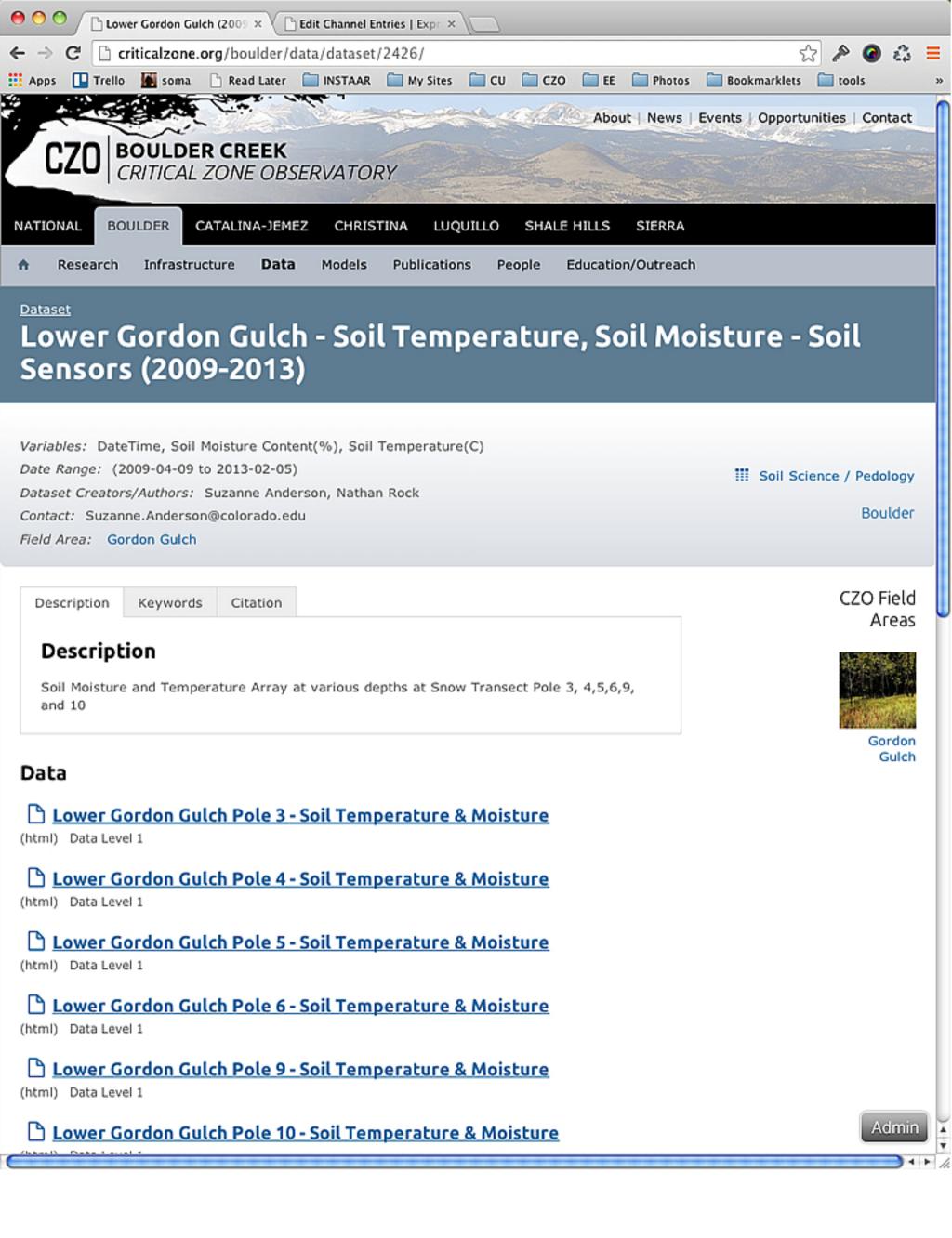


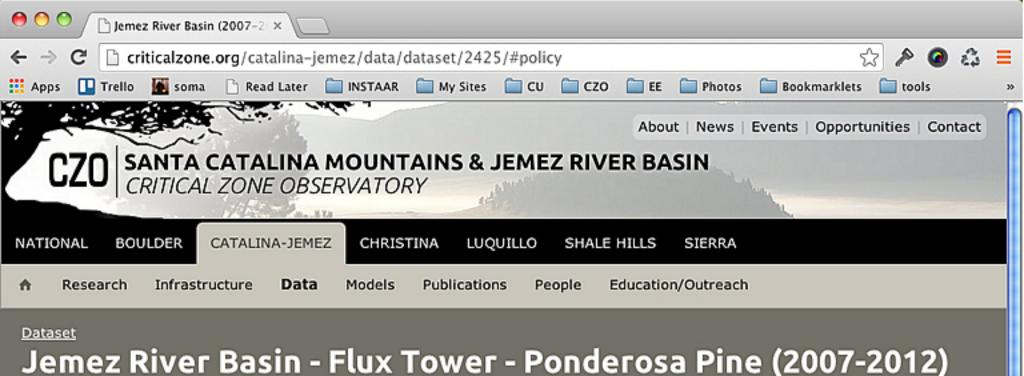


CMS-triggered
Data Value
Harvesting

CZO Central







Valles Caldera Ponderosa Pine Flux Tower

Variables: Friction velocity, Temperature, Wind direction, Wind speed, Carbon dioxide flux, Sensible heat flux, Latent heat flux, Precipitation, Relative humidity, Barometric pressure, Carbon dioxide, Vapor pressure deficit, Volumetric water content, Radiation net, Radiation incoming PAR, Radiation incoming shortwave, Radiation outgoing shortwave, Radiation incoming longwave, Radiation outgoing longwave, Water vapor concentration, Ecosystem respiration, Gross primary productivity.

Date Range: (2007-2012)

Dataset Creators/Authors: Marcy Litvak, Paul Brooks

Contact: Marcy Litvak, Biology Department, University of New Mexico, Biology MSC03 2020, 167

Castetter Hall, Albuquerque, NM 87131, mlitvak@unm.edu.

Field Area: Jemez River Basin

Biology / Ecology
Climatology /
Meteorology

Catalina-Jemez

Description

Keywords

Citation

Description

The flux tower is located in the Jemez River basin of the Jemez Mountains in north-central New Mexico at the southern margin of the Rocky Mountain ecoregion in the Valles Caldera National Preserve. The climate can be characterized as semi-arid, montane. Vegetation at this site is composed of a Pinus ponderosa overstory with Gambel oak scrubland (Quercus gambelii) understory. Tower height is 25 m. CZO Field Areas



Jemez River Basin

Data

Jemez River Basin - Flux Tower data

(.isp) Data Level 3

Jemez River Basin - Flux Tower data 2007

(.csv) Data Level 1, Metadata

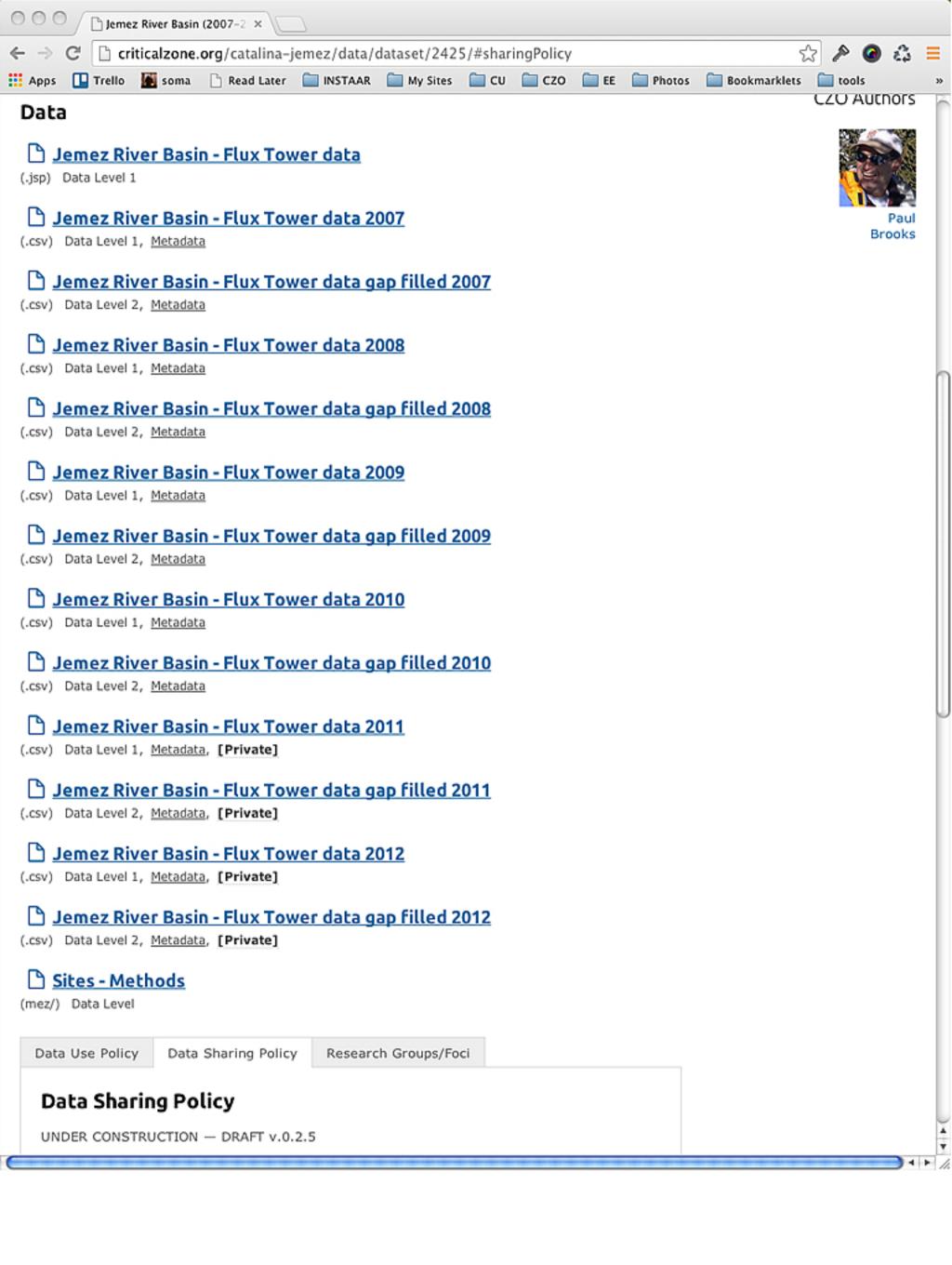
Jemez River Basin - Flux Tower data gap filled 2007

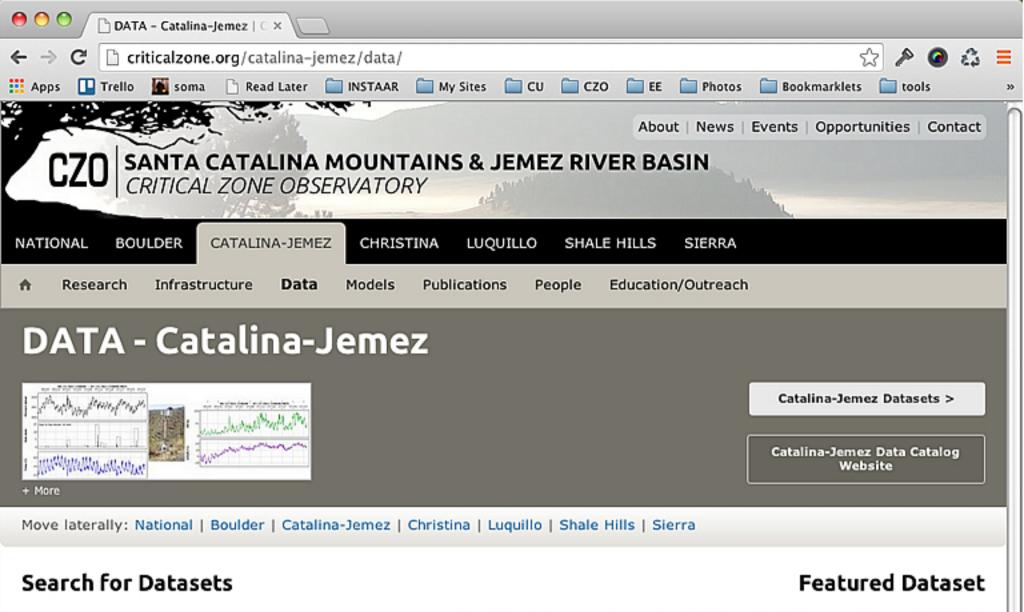
(.csv) Data Level 2, Metadata

CZO Authors



Paul Brooks





Search Tips]

Shale Hills, Boulder, Luquillo, JRB-SCM - Soil Geochemistry (2001-2013)

All Catalina-Jemez Datasets >

Integrated Data System

We are working with other CZOs to develop cross-CZO capabilities for sharing, integrating, analyzing, and preserving CZO data.

More Info >

CZO datasets are a powerful resource for developing and testing Earth system models. CZOs produce a large volume of data from a wide range of physical, chemical, and biological disciplines as well as spatial and temporal scales. This "flood" of disparate observational data partly reflects a new generation of integrated measurements. For example, most CZOs are using LIDAR imagery, copious isotopic samples, and low-cost sensors; others are also using eddy correlation and embedded sensor networks.

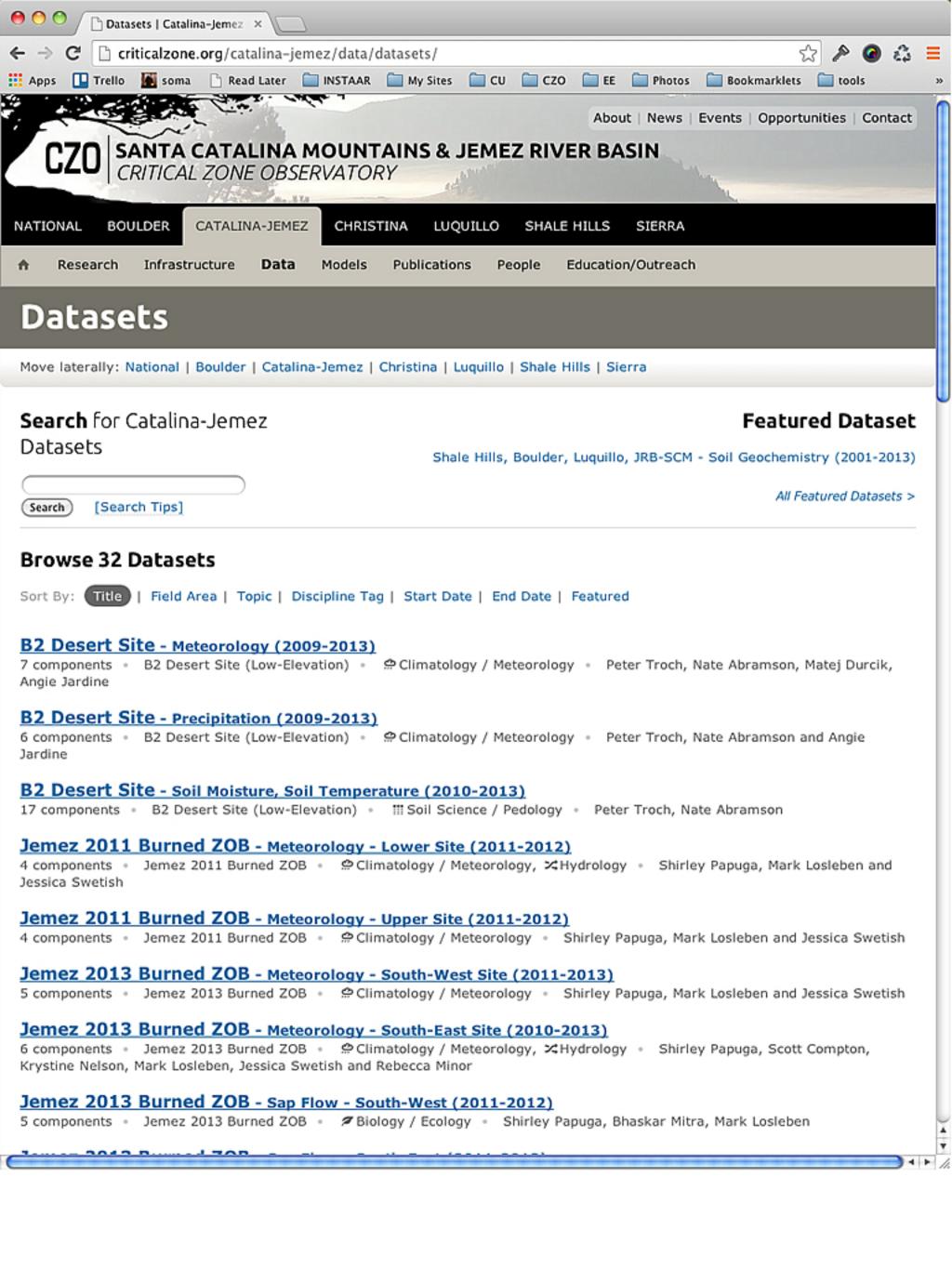
The Santa Catalina Mountains and Jemez River Basin CZO is collecting data from two field areas:

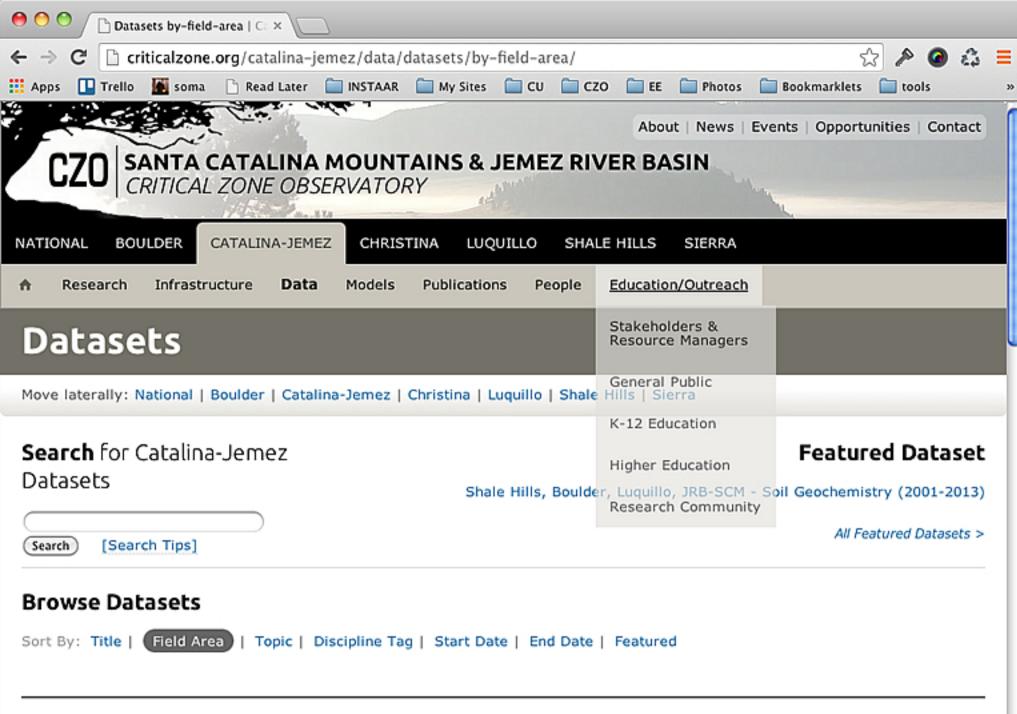
- · Santa Catalina Mountains located north of Tucson in southern Arizona three sites
- Jemez River Basin located in north-central New Mexico in the Valles Caldera National Preserve two zero order basins.

Our efforts thus far have concentrated on:

Land and Atmosphere

a) Airborne LiDAR: Digital Surface Model (DSM), Digital Elevation Model (DEM), Canopy height, intensity.





SANTA CATALINA MOUNTAINS



805-2790 m elevation, 11-18 °C, 420-940 mm/yr

The Santa Catalina Mountains (Catalina) are located northeast of Tucson in southern Arizona. This location includes three sites: one in the Sonoran desert on the mountain foothill; the second in the mid elevation and the third - high elevation in the Marshall Gulch Creek catchment.

Santa Catalina Mountains - Soil Water Chemistry (2006-2010)

2 components → Santa Catalina Mountains → ♦ Water Chemistry,

Geochemistry / Mineralogy → Jon Chorover, Peter Troch,
Julia Perdrial, Ingo Heidbuechel, Nate Abramson

Santa Catalina Mountains - Stream Water Chemistry, Rainfall Chemistry (2006-2010)

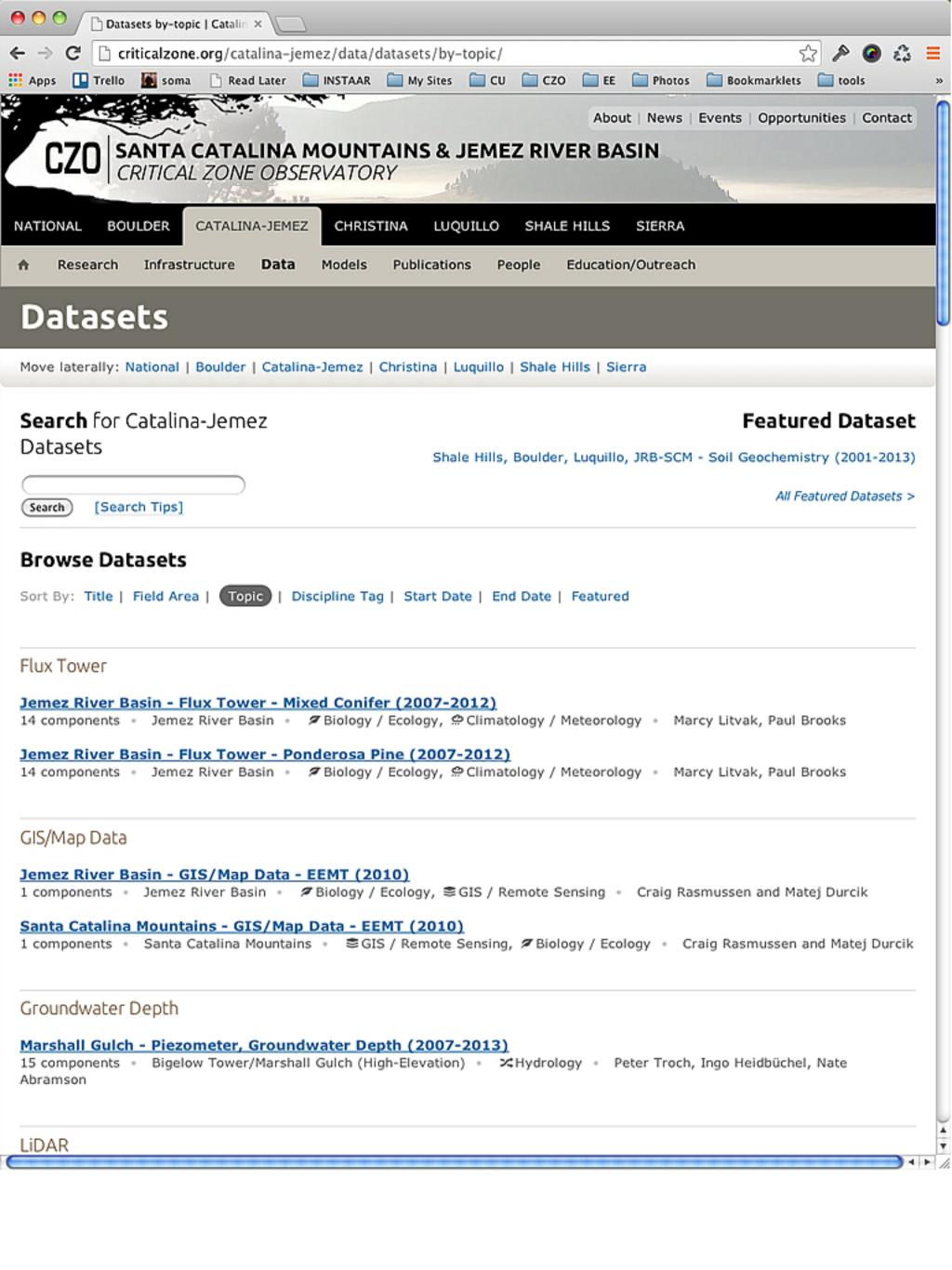
3 components • Santa Catalina Mountains • • Water Chemistry • Jennifer McIntosh, Paul Brooks, Peter Troch, Ingo Heidbüchel, Nate Abramson

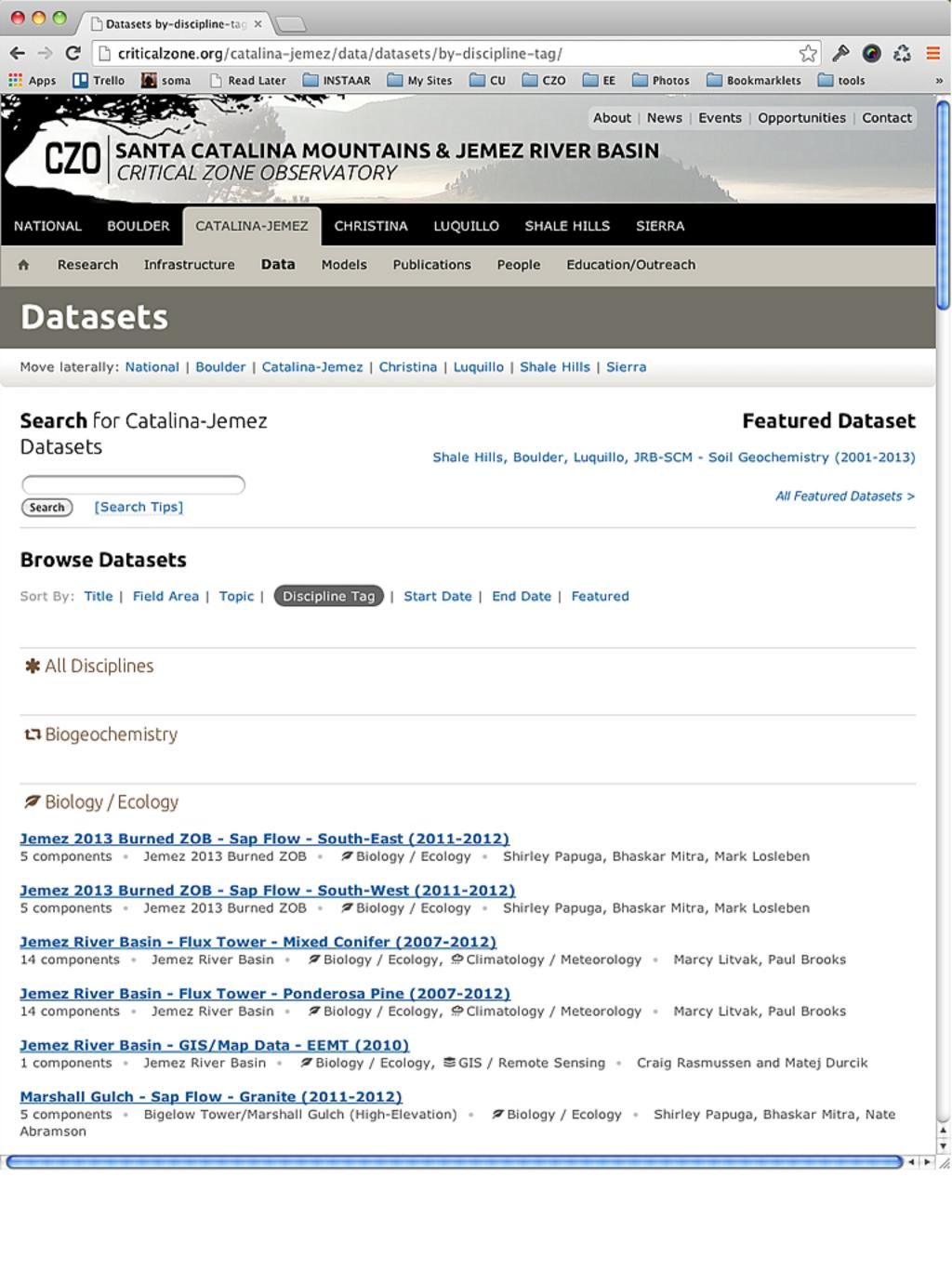
Santa Catalina Mountains - GIS/Map Data - EEMT (2010)

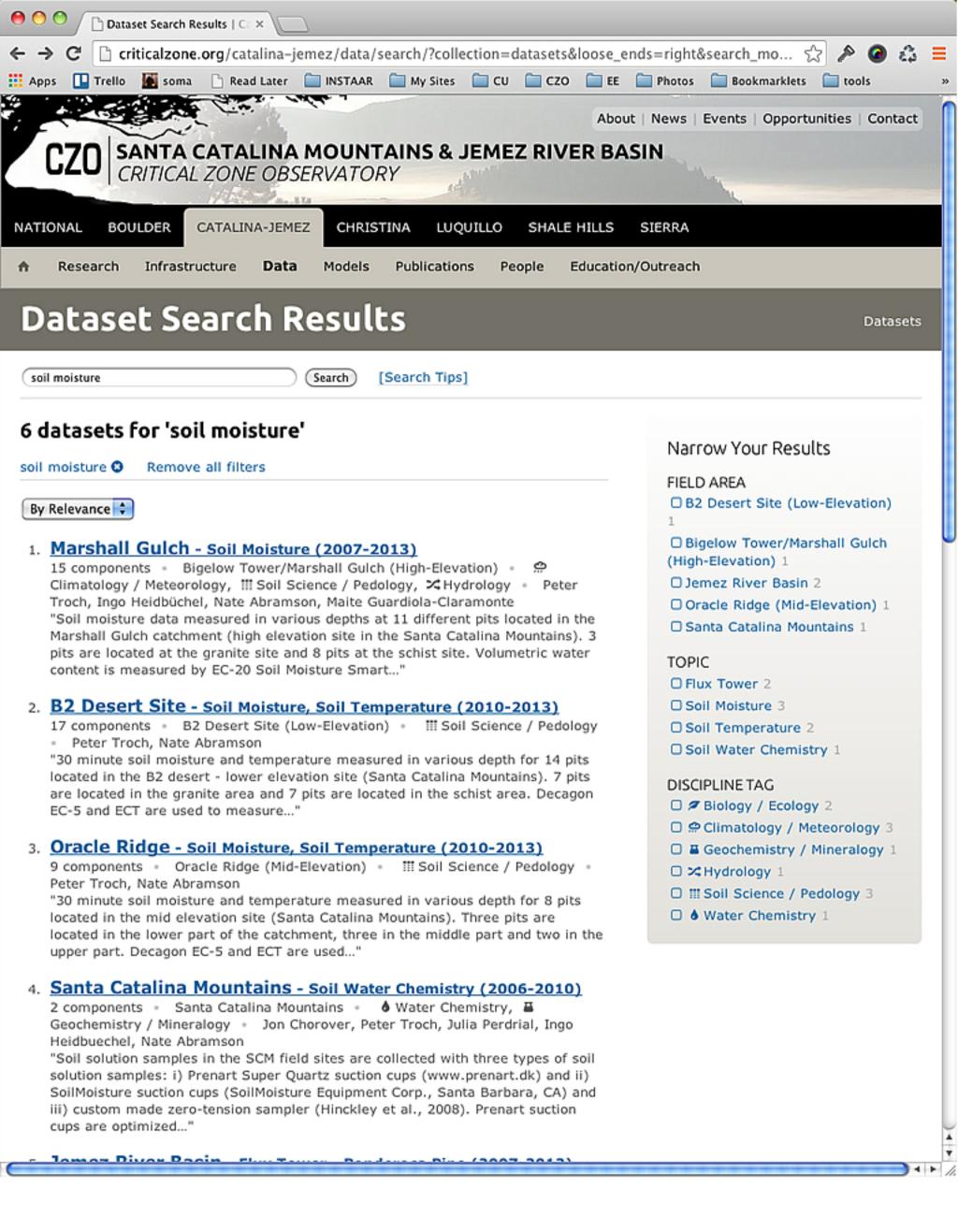
1 components . Santa Catalina Mountains . SGIS / Remote Sensing, Biology / Ecology . Craig Rasmussen and Matej Durcik

Shale Hills, Boulder, Luquillo, JRB-SCM - Soil Geochemistry (2001-2013)

1 components • Susquehanna Shale Hills Critical Zone Observatory, Bisley, Boulder Creek Watershed, Northeastern Puerto Rico and the Luquillo Mountains, Santa Catalina Mountains • **E** Geochemistry / Mineralogy, Diegeochemistry • Niu, Xianzeng; Williams, Jennifer; Brantley, Susan; Miller, Doug; Bills, Brian







Critical Zone Observatories

CP Home

Joe Tester

Log-out

reveal sidebar



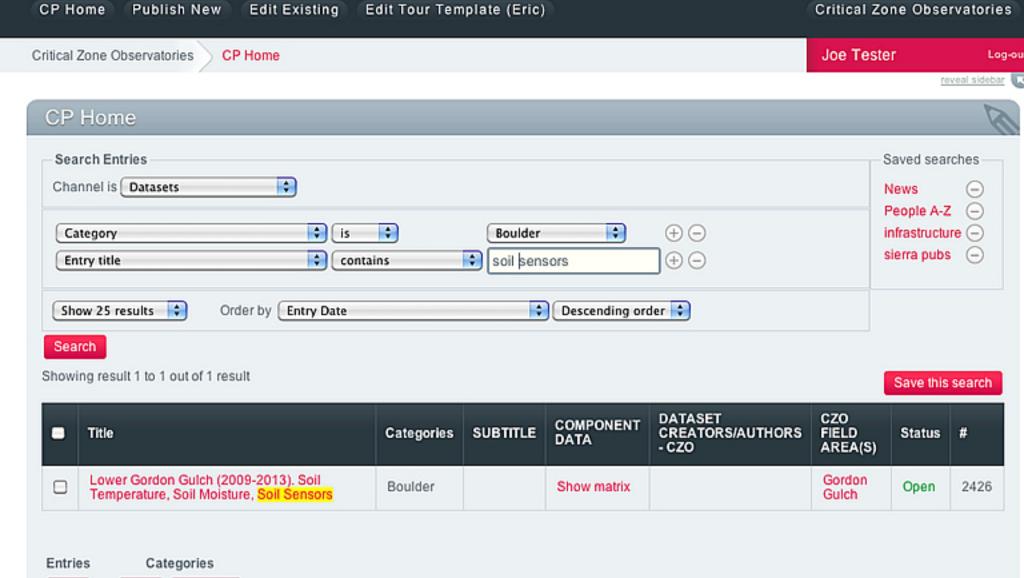
Search Entries Channel is Datasets \$\displant{\text{Datasets}}\$	Saved searches News People A-Z
Category	People A-Z — infrastructure — sierra pubs —
Show 25 results Order by Entry Date Descending order Descending order	siella pubs

Search

Showing results 1 to 25 out of 46 results

Save this search

1 2 🖸								
-	Title	Categories	SUBTITLE	COMPONENT DATA	DATASET CREATORS/AUTHORS - CZO	CZO FIELD AREA(S)	Status	#
0	Gordon Gulch (2008- 2011). Stream Water Chemistry,	Boulder	Gordon Gulch Surface Water Chemistry	Show matrix	Anderson, Suzanne	Gordon Gulch	Open	3065
0	Betasso (2008-2011). Stream Water Chemistry,	Boulder	Stream water chemistry samples collected in Betasso	Show matrix	Anderson, Suzanne	Betasso	Open	3064
	Boulder Creek (2010). LiDAR, Point Cloud Data at OpenTopography.com (Snow off)	Boulder		Show matrix		Boulder Creek Watershed	Open	2921
	Upper Gordon Gulch (2009-2011). Streamflow / Discharge,	Boulder	at gauge GGU_SW_0	Show matrix		Gordon Gulch	Open	2919
	Lower Gordon Gulch (2010-2012). Streamflow / Discharge,	Boulder		Show matrix		Gordon Gulch	Open	2918
	Boulder Creek (2010). LiDAR, Derived DEM from LIDAR	Boulder		Show matrix		Boulder Creek Watershed	Open	2915
	Lower Gordon Gulch (2012-2013). Air Temperature, Meteorology, South- Facing Meteorological Tower	Boulder		Show matrix		Gordon Gulch	Open	2889
0	Lower Gordon Gulch (2012-2013). Air Temperature, Meteorology, North- Facing Met Tower	Boulder, Luquillo		Show matrix		Gordon Gulch	Open	2888
	Boulder Creek (2010- 2013). GIS/Map Data, Google Earth with BCCZO locations and LIDAR	Boulder		Show matrix		Boulder Creek Watershed	Featured	2866
	Green Lakes Valley (1985-2011). Stream Water Chemistry, (data from Niwot Ridge I TER)	Boulder		Show matrix		Green Lakes Valley	Open	2865



ExpressionEngine v2.4.0 - Build: date 20120123 - © Copyright 2003 - 2014 EllisLab, Inc. Script executed in 2.4104 seconds - 63 SQL queries used

Edit

Add

Remove

Air Temperature

->

4

- Chlorophyll
- Climate
- Cosmogenic Radionuclides
- Diatoms
- Digital Elevation Model (DEM)
- Electrical Conductivity
- Flux Tower
- Geophysics

Instructions: Don't want the full Day, Month, and Year to appear on the website? If so, switch to "NO" and only the year will be displayed

YES, Show full dates

▼ DATE RANGE COMMENTS

Instructions: OPTIONAL. Additional info about the date range such as periodicity or consistency (e.g., weekly, irregular).

▼ SUBTITLE

Ciorro

Instructions: OPTIONAL. Additional important information that doesn't fit within the auto-generated composite title format of Location - Topic, Subtopi StartYear-EndYear. Try to keep concise. (e.g. Sampled at depths of 15 cm, 40 cm and 70 cm.)

Categories Status Author National . Chi Yang 0 Open Boulder Calhoun Catalina-Jemez Christina □ Eel ☐ IML Luquillo Reynolds Shale Hills

▼ Categories	▼ Status	▼ Author
National	Open 💠	Chi Yang 🗘
✓ Boulder		
Calhoun		
 Catalina-Jemez 		
Christina		
☐ Eel		
□ IML		
☐ Luquillo		
Reynolds		
Shale Hills		
Sierra		

™ ■ COMPONENT DATA

Instructions: Enter information about the individual data components of the dataset.

LOCATION Add additional info as needed, such as offset.

TOPIC Add additional info as needed.

URL Each component must have a URL, which points to a data file or a web page where you can access the data.

PRIVATE Website will show a label saying "Private". Restricting access is the responsibility of the individual CZO.

DATA LEVEL. For more info on the five Data Levels, see http://sv.criticalzone.org/edit.aspx?id=517576882&tbl=DataLevel&.

DOI (Optional). Digital Object Identifier. Just give the identifier, not a link, A URL link will be automatically generated for you.

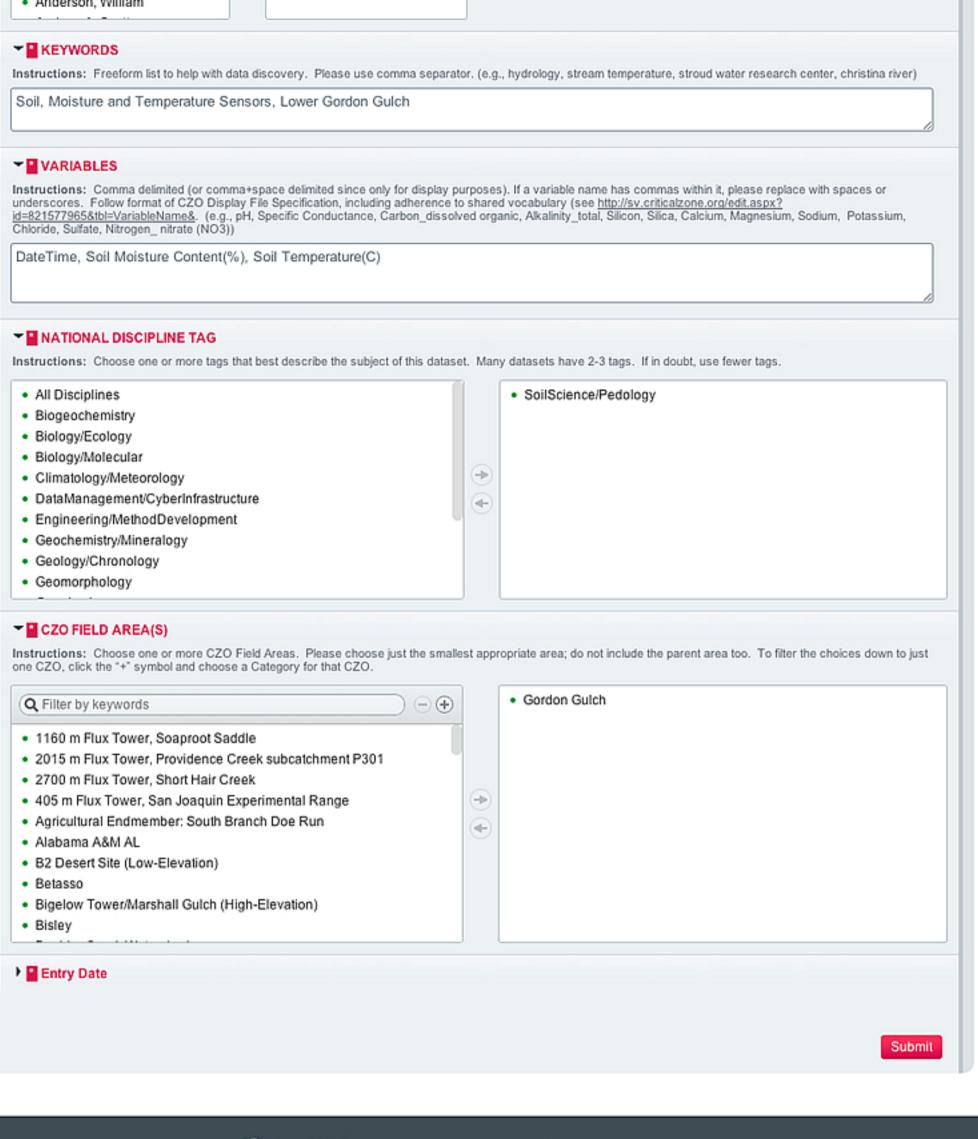
HARVEST UNIQUE METADATA? (Optional). If you want unique metadata to be harvested to SDSC for just this one component row, you must paste in a unique URL from your CZO's webserver. If left blank, then atuomatic harvesting will be based on a combination of the other columns in this component row (ie Location, Topic) and the overall dataset field values (ie Description/Abstract, Citation for this Dataset, Contact Person & Info, Subject/Keywords, Variables, Comments).

50 character limit. (e.g.,				LEVEL		HARVEST UNIQUE METADATA?
Gordon Gulch)	100 character limit. (e.g., Soil Temperature & Moisture 2011)	e.g., http://myczo.edu/datafile.html	Is the dataset currently private?	See above instructions	Digital Object Identifier (DOI) (e.g., 10.1029/2007GL031765)	See above Instructions.
Lower Gordon Gulch Pole 3	Soil Temperature & Moisture	http://czo.colorado.edu/q	NO YES	1 0		http://
Lower Gordon Gulch Pole 4	Soil Temperature & Moisture	http://czo.colorado.edu/q	NO YES	1 0		http://
Lower Gordon Gulch Pole 5	Soil Temperature & Moisture	http://czo.colorado.edu/q	NO YES	1 0		http://
Lower Gordon Gulch Pole 6	Soil Temperature & Moisture	http://czo.colorado.edu/q	NO YES	1 0		http://
Lower Gordon Gulch Pole 9	Soil Temperature & Moisture	http://czo.colorado.edu/q	NO YES	1 0		http://
Lower Gordon Gulch Pole 10	Soil Temperature & Moisture	http://czo.colorado.edu/q	NO YES	1 0		http://
	Lower Gordon Gulch Pole 4 Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 6 Lower Gordon Gulch Pole 9	Lower Gordon Gulch Pole 4 Soil Temperature & Moisture Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 6 Soil Temperature & Moisture Soil Temperature & Moisture Soil Temperature & Moisture Lower Gordon Gulch Pole 9 Lower Gordon Gulch Pole 9 Soil Temperature & Moisture	Lower Gordon Gulch Pole 3 Lower Gordon Gulch Pole 4 Soil Temperature & http://czo.colorado.edu/1 Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 6 Soil Temperature & http://czo.colorado.edu/1 http://czo.colorado.edu/1 http://czo.colorado.edu/1 Lower Gordon Gulch Pole 9 Lower Gordon Gulch Pole 9 Soil Temperature & http://czo.colorado.edu/1 http://czo.colorado.edu/1	Lower Gordon Gulch Pole 4 Soil Temperature & Moisture Moisture http://czo.colorado.ede 1 NO YES Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 6 Soil Temperature & http://czo.colorado.ede 1 NO YES Lower Gordon Gulch Moisture http://czo.colorado.ede 1 NO YES Lower Gordon Gulch Pole 9 Lower Gordon Gulch Soil Temperature & http://czo.colorado.ede 1 NO YES Lower Gordon Gulch Moisture http://czo.colorado.ede 1 NO YES Lower Gordon Gulch Moisture http://czo.colorado.ede 1 NO YES Lower Gordon Gulch Moisture	Lower Gordon Gulch Pole 3 Lower Gordon Gulch Pole 4 Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 6 Lower Gordon Gulch Pole 6 Lower Gordon Gulch Pole 6 Lower Gordon Gulch Pole 9 Lower Gordon Gulch Pole 10 Lower Gordon Gulch Pole 10 Lower Gordon Gulch Pole 10 Lower Gordon Gulch Pole 10	Lower Gordon Gulch Pole 4 Lower Gordon Gulch Pole 5 Lower Gordon Gulch Pole 6 Lower Gordon Gulch Pole 9 Lower Gordon Gulch Pole 10 Lower Gordon Gulch Pole 10

▼ ■ DESCRIPTION/ABSTRACT

Instructions: Who?, What?, When? Where?, Why? and How? Please make your text helpful to people unfamiliar with this dataset and with CZO. Length may be from a few words to multiple paragraphs.

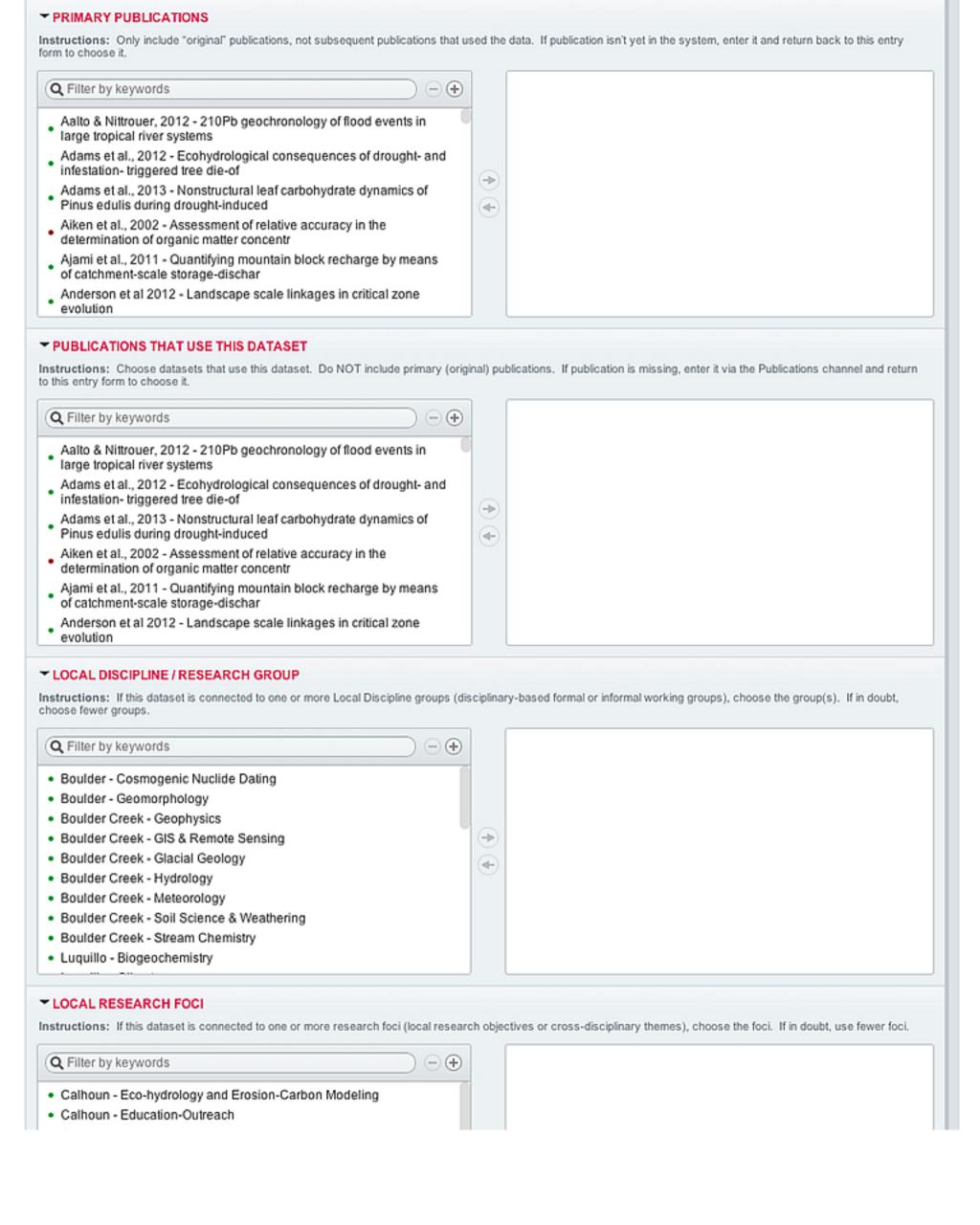
words to multiple paragraphs.	
Soil Moisture and Temperature Array at various depths at Snow Transec	ct Pole 3, 4,5,6,9, and 10
▼■ DATASET CREATORS/AUTHORS	▼ CITATION FOR THIS DATASET
Instructions: Data requires authors for the same reasons journal articles require authors. Please list specific people (e.g., Brantley, Susan L.; Ketchum, Blake; White, Tim; Sullivan, Pamela L.). If individuals not known, enter the organization/entity name.	Instructions: Data requires citations for the same reasons journal articles require citations. Please provide the following: author, title, year of publication, publisher (for data this is often the archive where it is housed), edition or version, and access information (a URL or other persistent identifier). For further guidance refer to
Suzanne Anderson, Nathan Rock	http://libquides.lib.msu.edu/citedata none
▼ DATASET CREATORS/AUTHORS - CZO	▼ CONTACT PERSON & INFO
Instructions: OPTIONAL. Are some of the dataset creators associated with	Instructions: Name and contact info. (e.g., Denis Newbold, Stroud Water
CZO? If so, choose them.	Research Center, 970 Spencer Road, Avondale, PA 19311. newbold@stroudcenter.org)
Q Filter by keywords — +	Suzanne.Anderson@colorado.edu
Aalto, Rolf Abramago, Note	
Abramson, Nate Adams, Hallie	CONTACT PERSON IS FROM CZO?
Almaraz, Maya	Instructions: Is the contact person associated with CZO? If "Yes", then the top
Anarde, Katherine	person on the CZO Contributors list will be the main contact.
Anders, Alison M	NO YES, CZO PERSON
Anderson, Ray	
Anderson, Robert S.	
Anderson, Suzanne	
Anderson, William	
▼ ■ KEYWORDS	
Instructions: Freeform list to help with data discovery. Please use comma separator	. (e.g., hydrology, stream temperature, stroud water research center, christina river)
Soil, Moisture and Temperature Sensors, Lower Gordon Gulch	
■ VARIABLES Instructions: Comma delimited (or comma+space delimited since only for display pur underscores. Follow format of CZO Display File Specification, including adherence to sid=821577965&tbl=VariableName&, (e.g., pH, Specific Conductance, Carbon_dissolve Chloride, Sulfate, Nitrogen_nitrate (NO3))	shared vocabulary (see http://sv.criticalzone.org/edit.aspx?
DateTime, Soil Moisture Content(%), Soil Temperature(C)	
▼ ■ NATIONAL DISCIPLINE TAG	
Instructions: Choose one or more tags that best describe the subject of this dataset.	Many datasets have 2-3 tags. If in doubt, use fewer tags.
All Disciplines	SoilScience/Pedology
Biogeochemistry	
Biology/Ecology	
Biology/Molecular	
Climatology/Meteorology	→
DataManagement/CyberInfrastructure	(4)

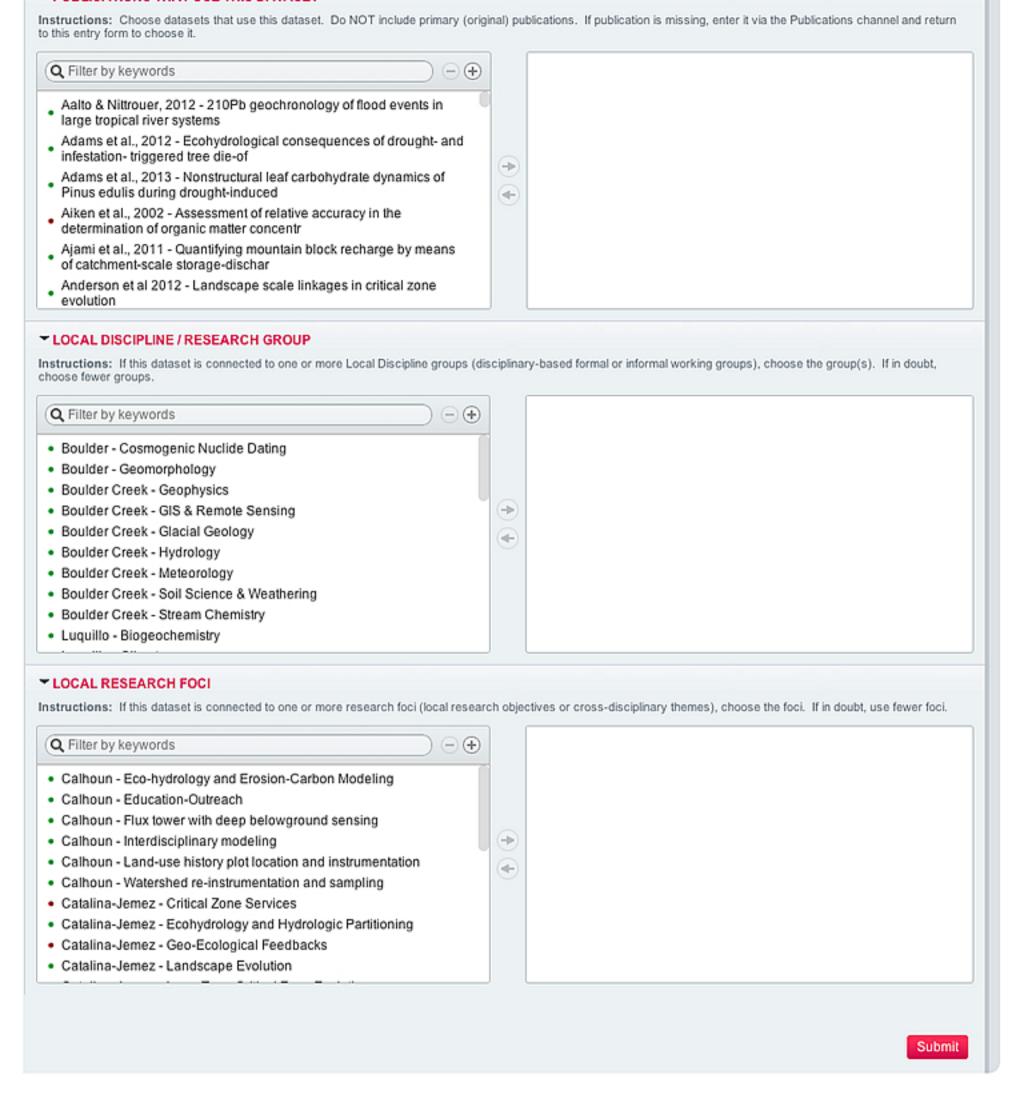


Betasso (2009-2014). Air Temperature, Climate, Meteorology, Live
 Meteorological from Betasso Meteoro

▼ PRIMARY PUBLICATIONS

Instructions: Only include "original" publications, not subsequent publications that used the data. If publication isn't yet in the system, enter it and return back to this entry form to choose it.





Critical Zone Observatories

CP Home P

Publish Edit Entry

Joe Tester

Log-out

reveal sidebar

Edit Entry

Publish Other Fields

Maps

MAP INSTRUCTIONS

▼ GENERAL MAP UPLOADS

Instructions: OPTIONAL. Upload general maps as images (e.g., png, jpg).

*TITLE *MAP FILE CAPTION *MAP SIZE
Enter a map title Upload or choose existing image file. (OPTIONAL) Description of the map Pick a size for the website

(+)

■ GENERAL MAP LAYER UPLOADS

Instructions: OPTIONAL. Upload general map layers as images (e.g., png, jpg). Each map layer image should be the exact same pixel dimensions.

*TITLE
Enter a very short title for the layer (e.g., Satellite)

*MAP FILE
Upload or choose existing image file.

CAPTION
(OPTIONAL) Description of the map layer

4

Instructions: Longitude of the approximate center of the area, in decimal degrees (e.g., -105.469). It's needed for Google Maps.

CENTROID LATITUDE (dd)

Instructions: Latitude of the approximate center of the area, in decimal degrees (e.g., 40.0128). It's needed for Google Maps.

▼ POINT MARKERS

Instructions: Optional. Enter information for up to 15 point markers that will appear in the Google Map. If you need more than 15 markers, please create and upload kml files.

*MARKER TITLE	*LATITUDE (dd)	*LONGITUDE (dd)	DESCRIPTION	COLOR
Brief title of the marker (e.g., Flux Tower A1B200).	Marker latitude in decimal degrees (e.g., 40.0128)	Marker longitude in decimal degrees (e.g., -105.469)	Brief description of the marker (100 characters max)	Pick a color

(+)

▼ KML/KMZ FILES

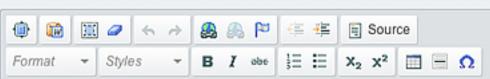
Instructions: OPTIONAL. Upload one or more KML or KMZ files in a Google Maps compatible format to plot points, polygons, polylines, and imagery. Consider uploading at least a single kml with a polygon border around your area.

CONTROL MAP EXENT? *TITLE *KML/KMZ FILE DESCRIPTION Do you want the locations within this Brief title of the file Upload a kml or kmz file in Google Maps Description of the file (e.g., Locations of soil pits dug in (e.g., soil pit kml/kmz file to control map extent? compatible format or choose an existing 2011 and 2012, includes popup info windows with site CHOOSE ONLY ONE FILE. Multiple locations) id) controls may create undesired results.

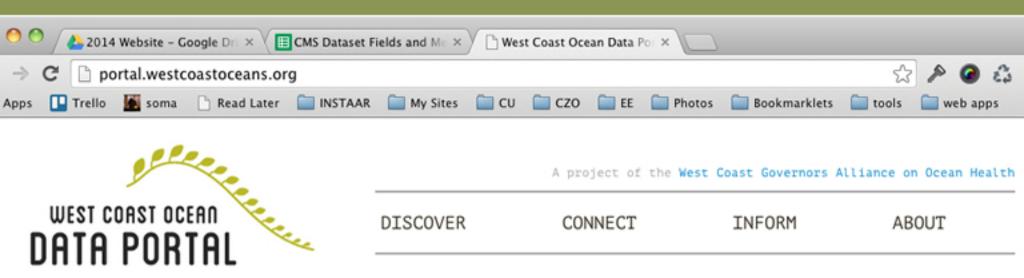


▼ MAP FREEFORM TEXT & CODE

Instructions: Have some additional geospatial info or metadata that doesn't fit elsewhere and needs to be discussed? Want to include links to external resources? Need to embed a dynamic map that's is housed elsewhere? Include those items or anything else here.

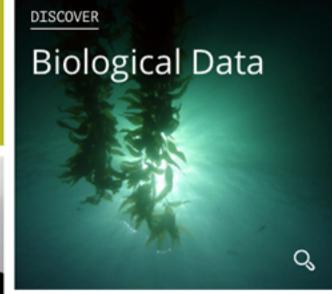


portal.westcoastoceans.org



Discover & contribute valuable data resources to inform ocean planning, policy development and resource management on the West Coast.









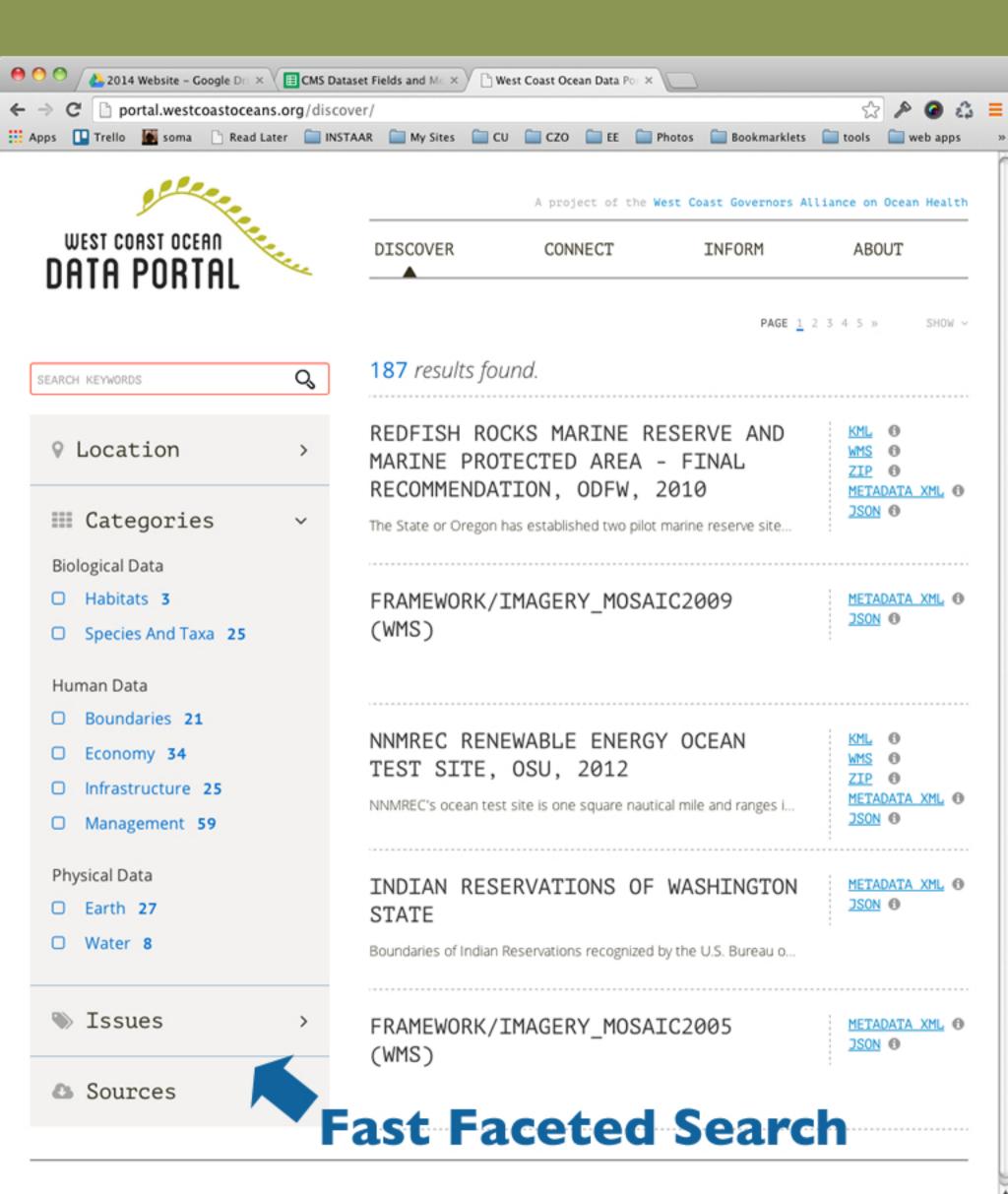
Physical Data

EXPLORE

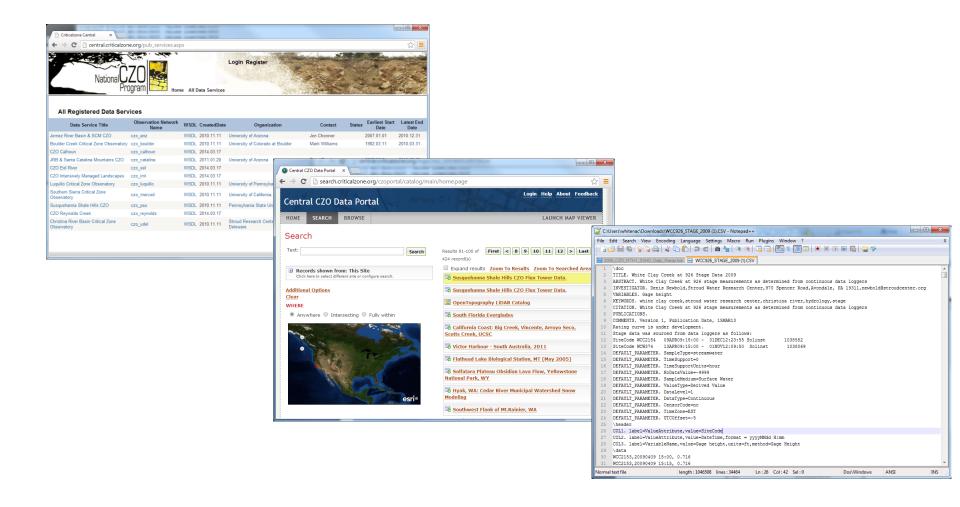
WELCOME

Contributors: see who is contributing to the registry.

portal.westcoastoceans.org



CZO Central, Geoportal Search interface, and Display Files



Outline

- How registering data at central.criticalzone.org works with ODM/hydroservers, and how those hydroservers are visible from cuahsi clients
- Display files formats.
- Explain the relationship with Central.critcialzone.org, and www.criticalzone.org, and search.criticalzone.org
- Was is Geoportal? Our plans for the criticalzone search interface.



Integrated CZO data system

Synthesizing information management experience and software from CZO partners and neighboring earth science projects into a standards-based system for publishing environmental data to emphasize the "critical zone" nature of our shared data sets







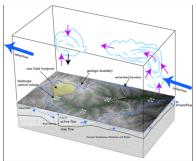








ADVANCED DATA MANAGEMENT IN SOLID EARTH GEOCHEMISTRY



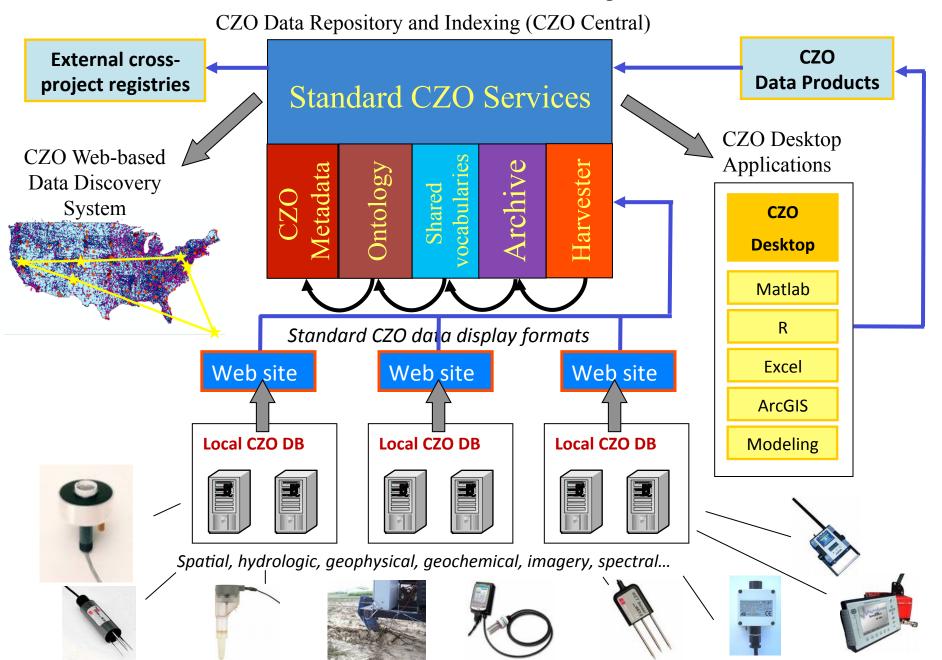




Ť	Туре	Name 🖶 🖶	Size	Last Modified
0		Bull Creek Lower Met		2008-03-25 09:20
0		Bull Creek Upper Met		2008-03-25 09:20
0		Critical Zone Tree		2009-05-29 10:57
0		Documents		2009-12-11 11:36
0		Maps		2010-03-22 13:35
0		P301 Flux Tower		2010-07-22 14:15
0		P301 Water Balance Transect		2009-05-29 11:50

Southern Sierra CZO - KREW

CZO Data Publication System

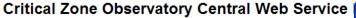




CZO Central web service registry

Logistical support and/or data were provided by the

NSF-supported Boulder Creek Critical Zone Internet



POINT OBSERVATION DATA SERVICES

This website supports the sharing of hydrologic data published using WaterOn publication services support the publication of data by the research community Data Model (ODM). For more informatin on publishing hydrologic data, visit: U

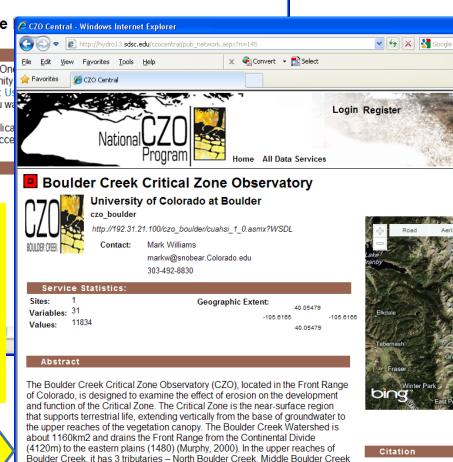
- Server Set-up. Establish a relational database server to host the data you was a pata Loading Load the data into CDM.
- Data Loading, Load the data into ODM.
- Network Set-up. Establish web services to make the data from ODM publica
- 4. Data Indexing. Register the web services with CUAHSI so that they are acce

WEB SERVICE REGISTRATION SYSTEM

- Listing of registered public data services (47 registered services)
- Hydrologic Concept Ontology (requires java).

CZO display file is automatically ingested in CZO data repository, a service is updated, making new data available

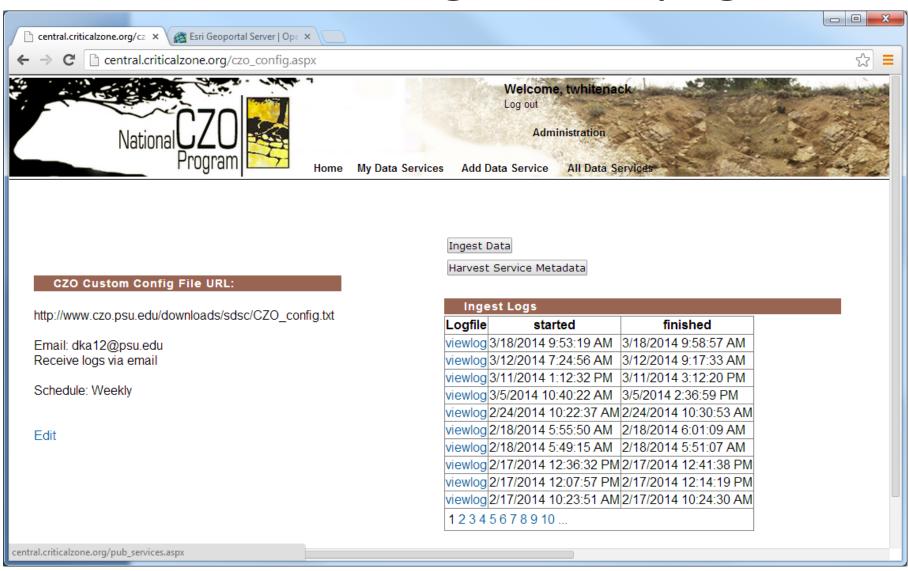
Boulder Creek CZO web service



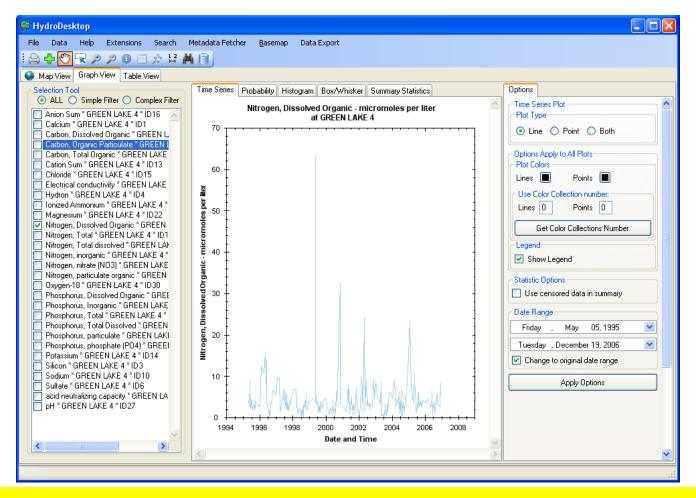
and South Boulder Creek. The confluence of North Boulder Creek and Middle

Boulder Creek is at ?m. South Boulder Creek joins Boulder Creek near the City

Central Configuration page



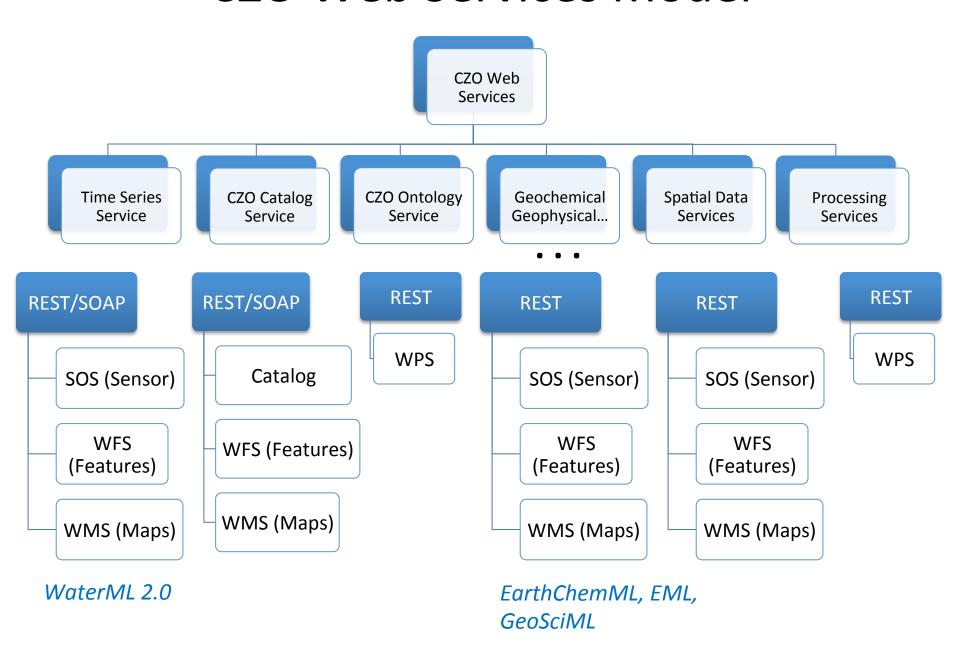
Working with CZO Time Series Data



Once CZO web service is updated and registered in CZO Central, it can be discovered in HydroDesktop an open source application with rich mapping and time series analysis capabilities

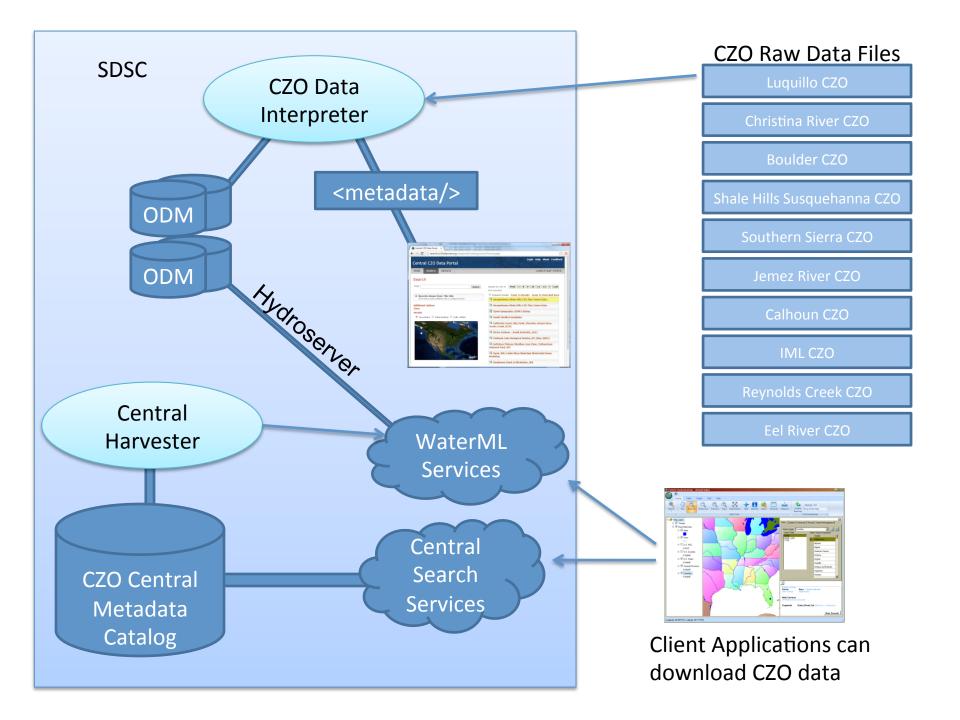
HydroDesktop, showing one of 31 newly ingested time series

CZO Web Services Model



CZO Data Publication Model

- Relies on individual CZO data management systems to generate display files
 - Display file is modeled on LTER data file, and allows adding series-level and data valuelevel attributes as defined in CUAHSI Observations Data Model
- When additional display files are generated and referenced in the config file, they are picked up and ingested in a CZO repository / hydroserver at SDSC
- The time series in the files are then automatically exposed as water data services (WaterML-compliant web services used by CUAHSI HIS)
- These services are available for data discovery and analysis by a variety of applications: CZO Desktop (a version of HydroDesktop), Google Earth, etc.
- A non-intrusive system: no change in how one would normally publish data on CZO web sites; no additional software/hardware needed.
- Can be a good model for the community wishing to publish their data in an easy and inexpensive way
 - note the NSF requirement for data management plans with every proposal from October 2010



Config file

```
\Root URL
http://www.czo.psu.edu/downloads/sdsc/
\Category Hydrological Time Series
\Sites
sites.csv
\Methods
methods.csv
\Headers
ts precip/daily/2006 CZO RTH1 SSHO Daily Precip.hdr
ts precip/daily/2007 CZO RTH1 SSHO Daily Precip.hdr
ts precip/daily/2008 CZO RTH1 SSHO Daily Precip.hdr
ts precip/daily/2009 CZO RTH1 SSHO Daily Precip.hdr
ts precip/daily/2010 CZO RTH1 SSHO Daily Precip.hdr
ts precip/daily/2011 CZO RTH1 SSHO Daily Precip.hdr
ts_precip/daily/2012_CZO_RTH1_SSHO_Daily_Precip.hdr
ts precip/hourly/2006 CZO RTH1 SSHO Hourly Precip.hdr
ts precip/hourly/2007 CZO RTH1 SSHO Hourly Precip.hdr
ts precip/hourly/2008 CZO RTH1 SSHO Hourly Precip.hdr
ts precip/hourly/2009 CZO RTH1 SSHO Hourly Precip.hdr
ts precip/hourly/2010 CZO RTH1 SSHO Hourly Precip.hdr
ts precip/hourly/2011 CZO RTH1 SSHO Hourly Precip.hdr
ts precip/hourly/2012 CZO RTH1 SSHO Hourly Precip.hdr
```

Sites file

- SiteCode, SiteName, Latitude, Longitude, LatLong Datum, Elevation, Vertical Datum, Local X, Local Y, Local Projection, Pos Accuracy, Comments
- CZO RTH1,CZO RTH 1,40.6658173,-77.9040146,WGS84,302.7,NAVD88,,,,,weather station
- CZO_RTH2_1,CZO RTH 2 SM1,40.66522,-77.90315,NAD 1983,279.875398,NAVD88,587053.5885,147909.778,PA South,x-y < 10 cm; z < 5 cm,soil moisture 1
- CZO_RTH2_2,CZO RTH 2 SM2,40.66522,-77.90315,NAD 1983,279.761841,NAVD88,587041.7909,147907.6064,PA South,x-y < 10 cm; z < 5 cm,soil moisture 2
- CZO_RTH2_3,CZO RTH 2 SM3,40.66522,-77.90315,NAD 1983,277.155926,NAVD88,587049.6105,147898.5509,PA South,x-y < 10 cm; z < 5 cm,soil moisture 3
- CZO_RTH3_1 ,CZO RTH 3 SM1 ,40.6645164,-77.90552,NAD 1983,266.057135,NAVD88,586855.3065,147828.4402,PA South,x-y < 10 cm; z < 5 cm,soil moisture 1 and groundwater well 1

Methods file

Methodcode, method Description, method Link

- 253MP, Water (matric) potential Campbell Scientific 253 probes calib using mfr sugg eqn fr raw mV measure, http://www.campbellsci.com
- AirRH_HMP45C,Relative humidity measured w Campbell Scientific HMP45C Temperature and Rel Humidity Probe,http://www.campbellsci.com
- AirTemp_CR1K,measured with internal sensors in the CR1000 data loggers,http://www.campbellsci.com
- AirTemp_CSat3, Air temperature measured with a Campbell Scientific CSAT3 Three Dimensional Sonic Anemometer, http://www.campbellsci.com
- AirTemp_HMP45C,Air temperature measured with a Campbell Scientific HMP45C Temperature and Relative Humidity Probe,http://www.campbellsci.com

Header File / Display File

\doc

TITLE. White Clay Creek at 926 Stage Data 2009

ABSTRACT. White Clay Creek at 926 stage measurements as determined from continuous data loggers

INVESTIGATOR. Denis Newbold, Stroud Water Research Center, 970 Spencer Road, Avondale, PA

19311,newbold@stroudcenter.org

VARIABLES. Gage height

KEYWORDS. white clay creek, stroud water research center, christina river, hydrology, stage

CITATION. White Clay Creek at 926 stage measurements as determined from continuous data loggers PUBLICATIONS.

COMMENTS. Version 1, Publication Date, 15MAR13

Rating curve is under development.

Stage data was sourced from data loggers as follows:

SiteCode WCC2154 09APR09:15:00 - 31DEC12:23:55 Solinst 1038582

SiteCode WCW374 13APR09:15:00 - 01NOV12:09:50 Solinst 1038069

DEFAULT PARAMETER. SampleType=streamwater

DEFAULT_PARAMETER. TimeSupport=0

DEFAULT_PARAMETER. TimeSupportUnits=hour

DEFAULT PARAMETER. NoDataValue=-9999

DEFAULT PARAMETER. SampleMedium=Surface Water

DEFAULT PARAMETER. ValueType=Derived Value

DEFAULT PARAMETER. DataLevel=1

DEFAULT_PARAMETER. DataType=Continuous

DEFAULT PARAMETER. CensorCode=nc

DEFAULT PARAMETER. TimeZone=EST

DEFAULT PARAMETER. UTCOffset=-5

\header

COL1. label=ValueAttribute,value=SiteCode

COL2. label=ValueAttribute,value=DateTime,format = yyyyMMdd H:mm

COL3. label=VariableName,value=Gage height,units=ft,method=Gage Height

\data

WCC2153,20090409 15:00, 0.716

WCC2153,20090409 15:15, 0.716

How www.criticalzone.org and search.criticalzone.org will work together

- Metadata will be created from Datafiles registered at www.criticalzone.org, and synced with geoportal*
- Data files that comply with the "displayfile" specification, will also be ingested into the hydroserver.*



^{*}not yet functionally complete

What are the roles of each?

• www.criticalzone.org: The main, public facing site.

Search.criticalzone.org: Search

 Central.criticalzone.org: not intended to be a public site. Only used by data managers to control ingest of display files into the hosted hydroserver.

Geoportal

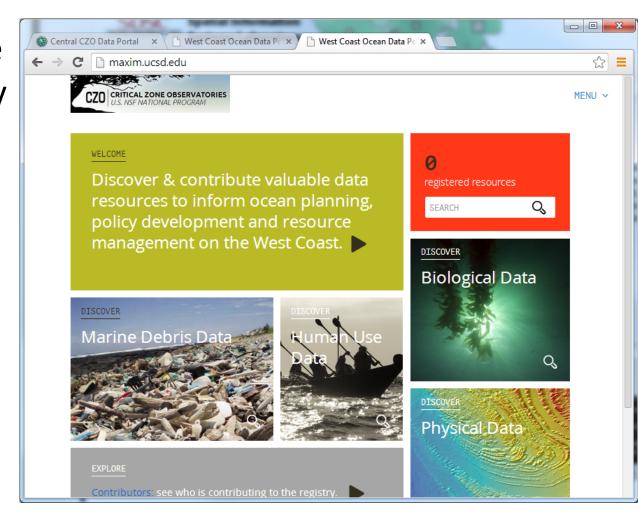
- OGC compliant services for CSW (catalog services for the web).
- Open source, but developed and supported by ESRI.
- Extendable.



Geoportal Facets and Custom Interfaces

- Utilizes Apache Solr technology for fast searching.
- Custom

 interface from
 ecotrust.org



Questions? Comments?

CZO Display File v1

Requirements:

- Human readable
- Machine parsable
- Conform to ODM 1.1 & CZO Shared Vocabulary

Solution:

- ASCII text file, with comma-separated data
- Header with detailed, structured metadata
- Separate Sites and Methods files
- http://criticalzone.org/national/publications/pub/whitenack-et-al-2011czo-display-file-specification/

CZO Display File v1

Issues:

- Many parsing errors,
- Limited applicability to many disciplinary data
 - Designed around CUAHSI HIS
 - Can't handle soil intervals, specimen genealogies, etc.
 - Only accepts terms from CZO Shared Vocabulary
- Relatively "flat" information model.
 - Not rich enough for modern data systems (incl. ODM2)
- Does not meet archival requirements!
 - Separate "header" file for metadata
 - No explicit in-file references to Sites or Methods file/info
- Do not open nicely in Excel

CZO Display File v2

Requirements:

- Human readable & Machine parsable
- Conform to ODM2 & new Shared Vocabulary

Solution:

- JSON text file, with comma-separated data array
- Header with detailed, structured metadata
 - Structural validation likely via http://json-schema.org/
- Separate Sites and Methods files
 - but with clear links in each file
- Still working on specification

Thank You