Building the LTAR Data Portal

Susan McCarthy

National Agricultural Library

January 7, 2016

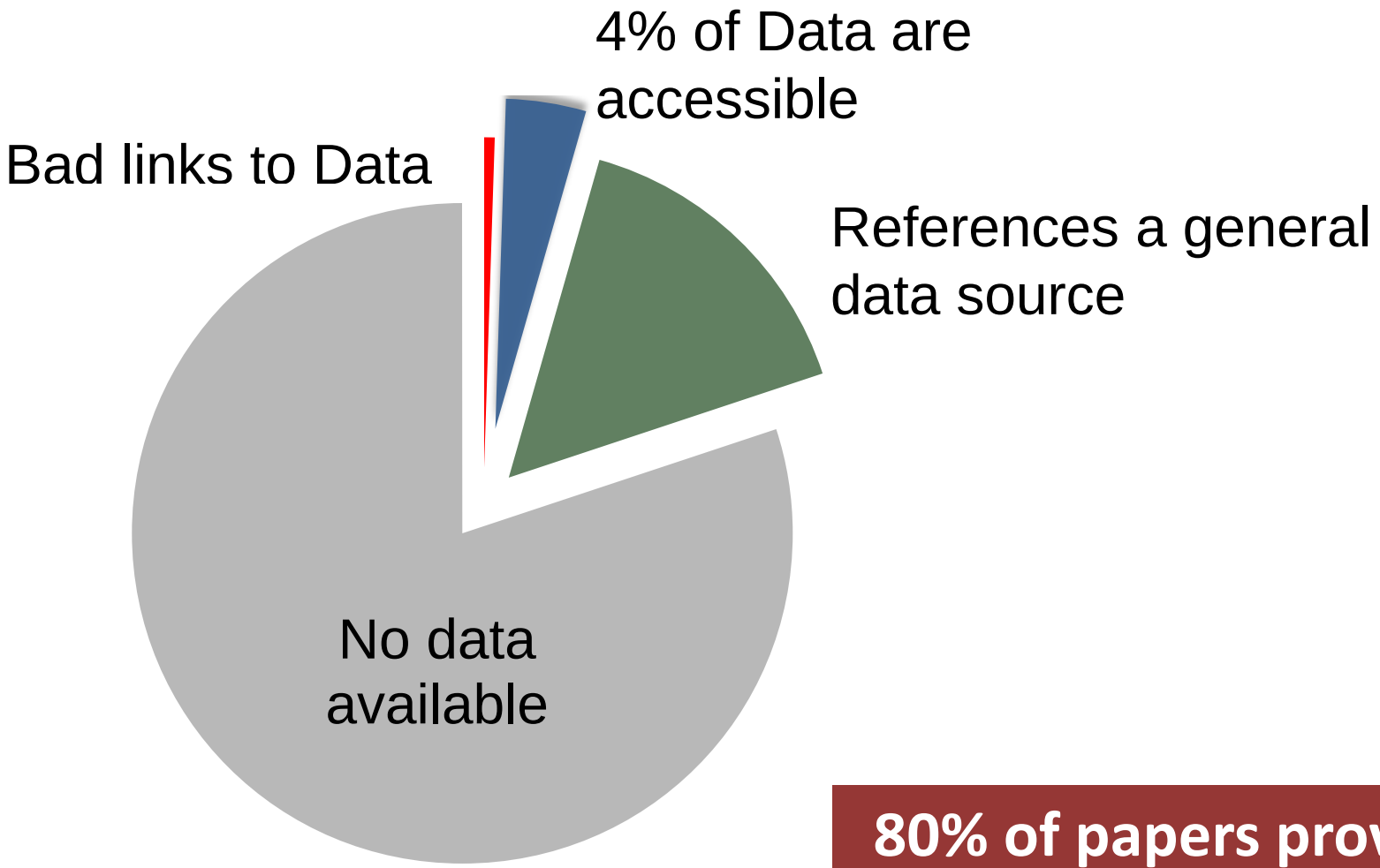
Goals and Motivation *for* LTAR Data Portal

1. Federal Directives / mandates for open and machine readable access to Federally funded research results
2. For the science!

USDA

LTAR Publication-Related Data Loss

N=194 of ~500 citations in 2011 LTAR site proposals



80% of papers provide no way to obtain data

Putting the “Long-Term” in LTAR Data

In order to have data of known quality for re-use in a new application…

50 years from now, someone would need to:

- Discover that the data exists;
- Find the data;
- Obtain and read the data;
- Determine the data’s suitability for re-use

Long-Term Data Requirements

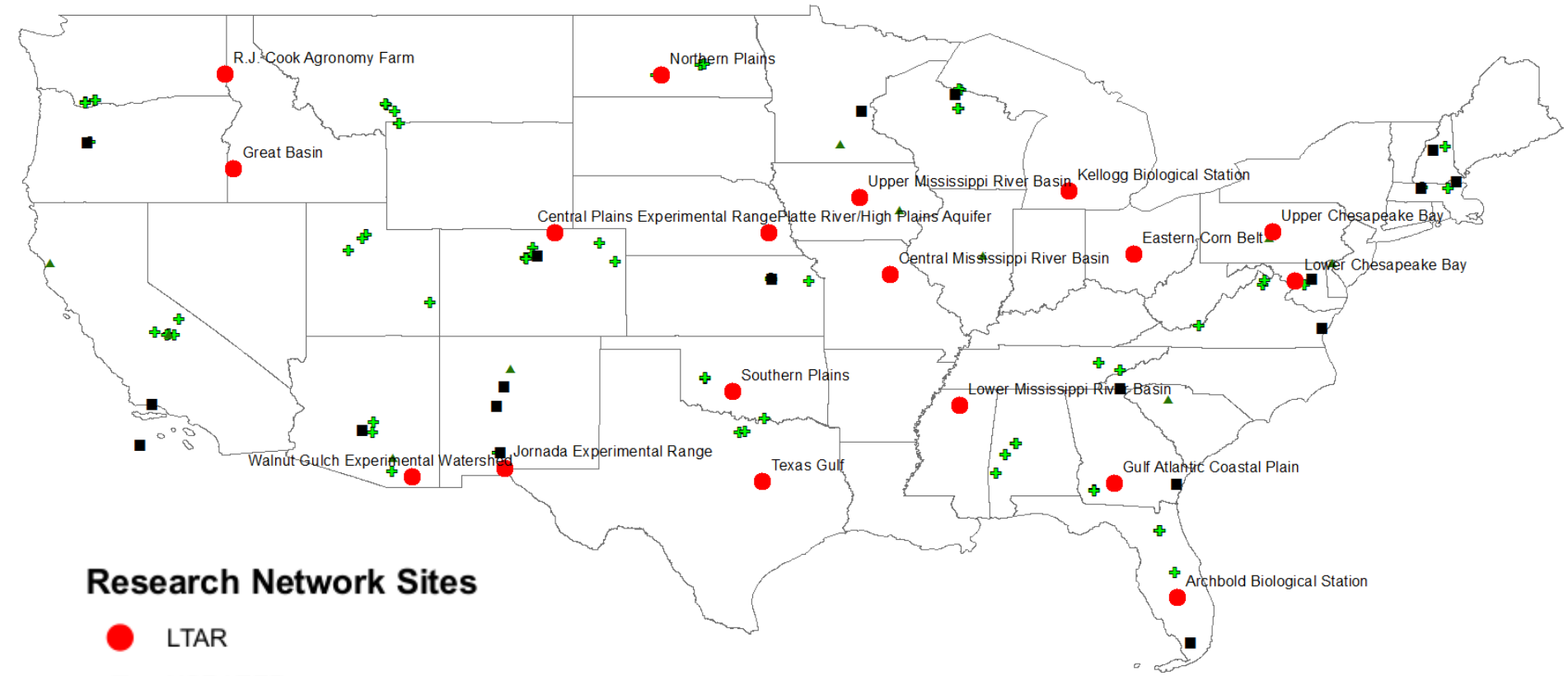
- **Preserved** – data needs to be secured for the long-term
- **Discovered** – consistent application of standards-based metadata description
- **Accessed** -- available through standards-compliant Web services
- **Re-used** – multiple data output formats; adequate metadata including descriptions of intended use, limitations of use, and data quality

Data Curation for the Long-Term

Archiving is not a back-up!

- Initiating a data curation program at NAL:
  - In collaboration with the University of Maryland iSchool faculty and postdoc (R. Punzalan, A. Kreisberg)
  - Currently conducting a self-assessment
  - Developing requirements and recommendations

Long Term Agro-ecosystem Research (LTAR)



USDA

LTAR Research

Common Observatory Data (CORe)

Immediate and Near Term

- Meteorology
- Phenocam (camera)
- Hydrology
- Eddy flux: CO_2 and Non-CO_2 gasses
- Soil

Long Term

- Biological

Research Approach

Common Experiment

- Management Studies: Business as usual compared with Aspirational
- Crop and livestock

Common Experiment Data

- Management practices
- Results

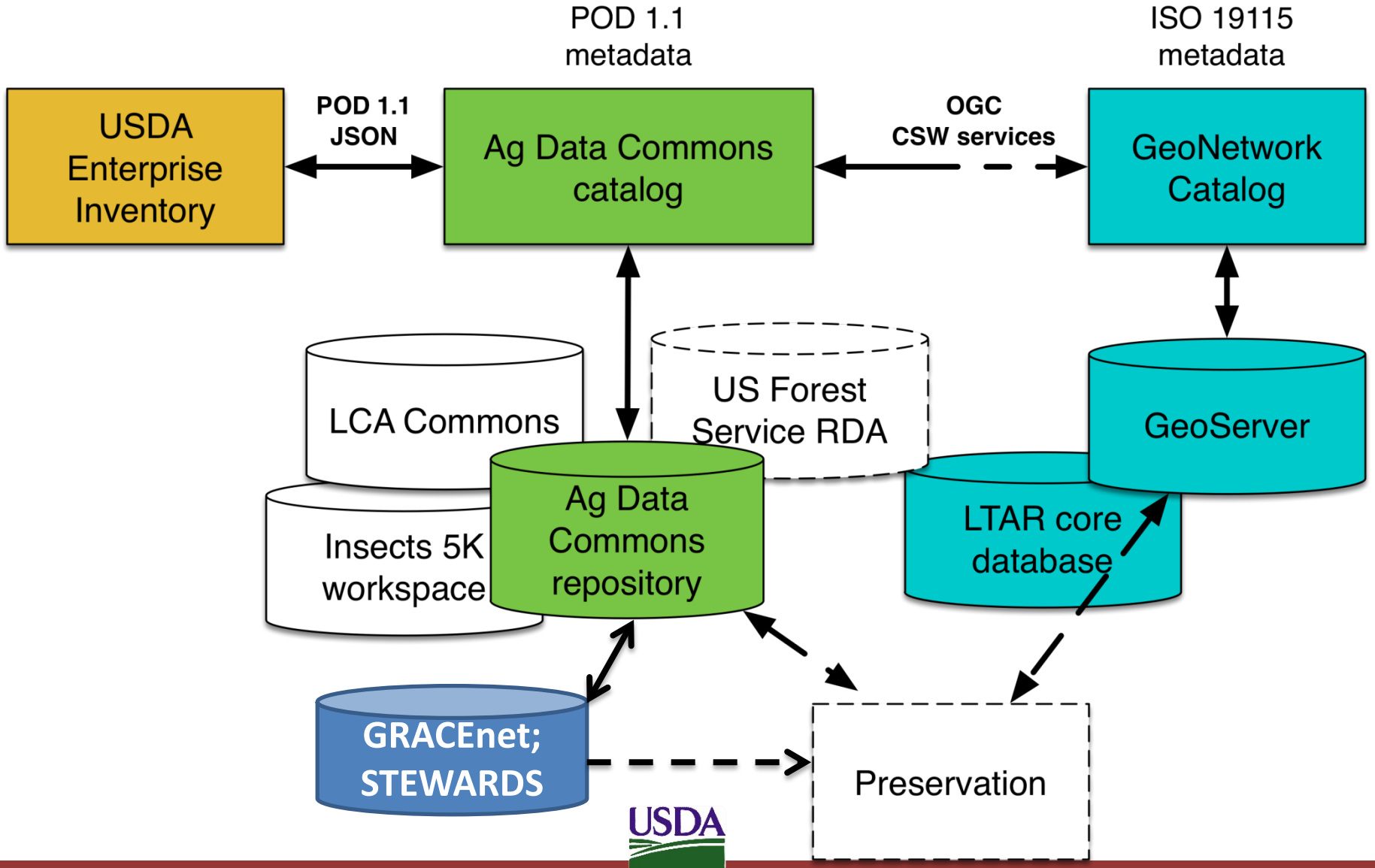
Approach for LTAR Data Management

- Provide a central data catalog (registry) linking to wherever the data resides
- Describe the data for long-term re-use and discovery
- Leave existing data in specialized systems (STEWARDS, GRACENet, PhenoCam, etc.)
- Common Observatory data stored at NAL
- Preserve (curate) the data for the long-term

LTAR Data Under Management

1) Near real-time CORe
2) Publication-related datasets
3) Aerial Imagery

LTAR Data IT Infrastructure @ NAL



Two Interconnected Systems

Ag Data Commons

System for Agricultural data

- Central registry for USDA data
  - Link to data residing outside of the Ag Data Commons
- Provide a repository for data without logical “home”
- Provide citable DOI’s
- Feed USDA Enterprise Inventory >>> Data.gov

Geospatial Data System

System LTAR CORe and GIS-based data

- Central registry – GIS data
- Robust descriptive information
  - Enhanced discovery and display for GIS data
  - Automatic feed to Ag Data Commons >> … >>> Data.gov
  - Cross-site linkage
  - Citable DOI’s

https://ltar.nal.usda.gov/

USDA Long-Term Agroecosystem Research

Data Home | About LTAR | Real-Time Meteorology | Real-Time Cameras | Published Data and Articles | Alpha Pilot Version

Long-Term Agroecosystem Research Data Overview

The USDA Agricultural Research Service (ARS) Long-Term Agroecosystm Research network consists of 18 Federal and university agricultural research sites with an average of over 50 years of history. The goal of this research network is to ensure sustained crop and livestock production and ecosystem services from agroecosystems, and to forecast and verify the effects of environmental trends, public policies, and emerging technologies.

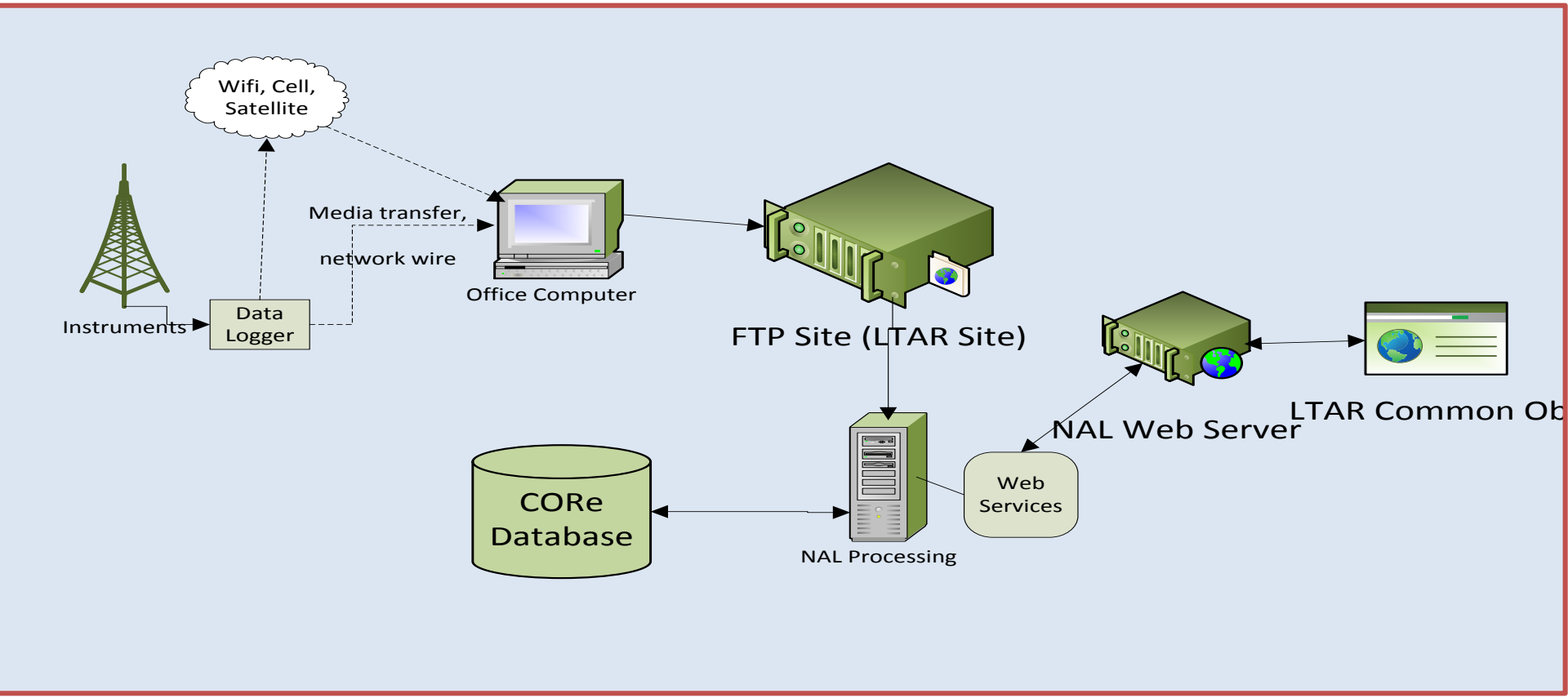
Display Options

- LTAR
  - LTAR sites
  - Soil Climate Data
  - Atmospheric Gas D
  - Camera Images
- Other Research Networ
- Boundaries
- Ecological Context
  - Watersheds
  - Seasonal Drought C
  - Ecosystems
  - Bioclimates
  - Land Surface Form
  - Surface Lithology
  - Topographic Positic
- Base Maps

Click on the LTAR sites on the map to see more information about each site. Other national ecosystem research networks are available on the map to show how LTAR contributes to nation-wide coverate. The map also links for LTAR data available through the Ameriflux, Soil Climate Analysis Network, and PhenoCam research data networks.

USDA

CORe Meteorological Data Flow



LTAR CORe Data Status

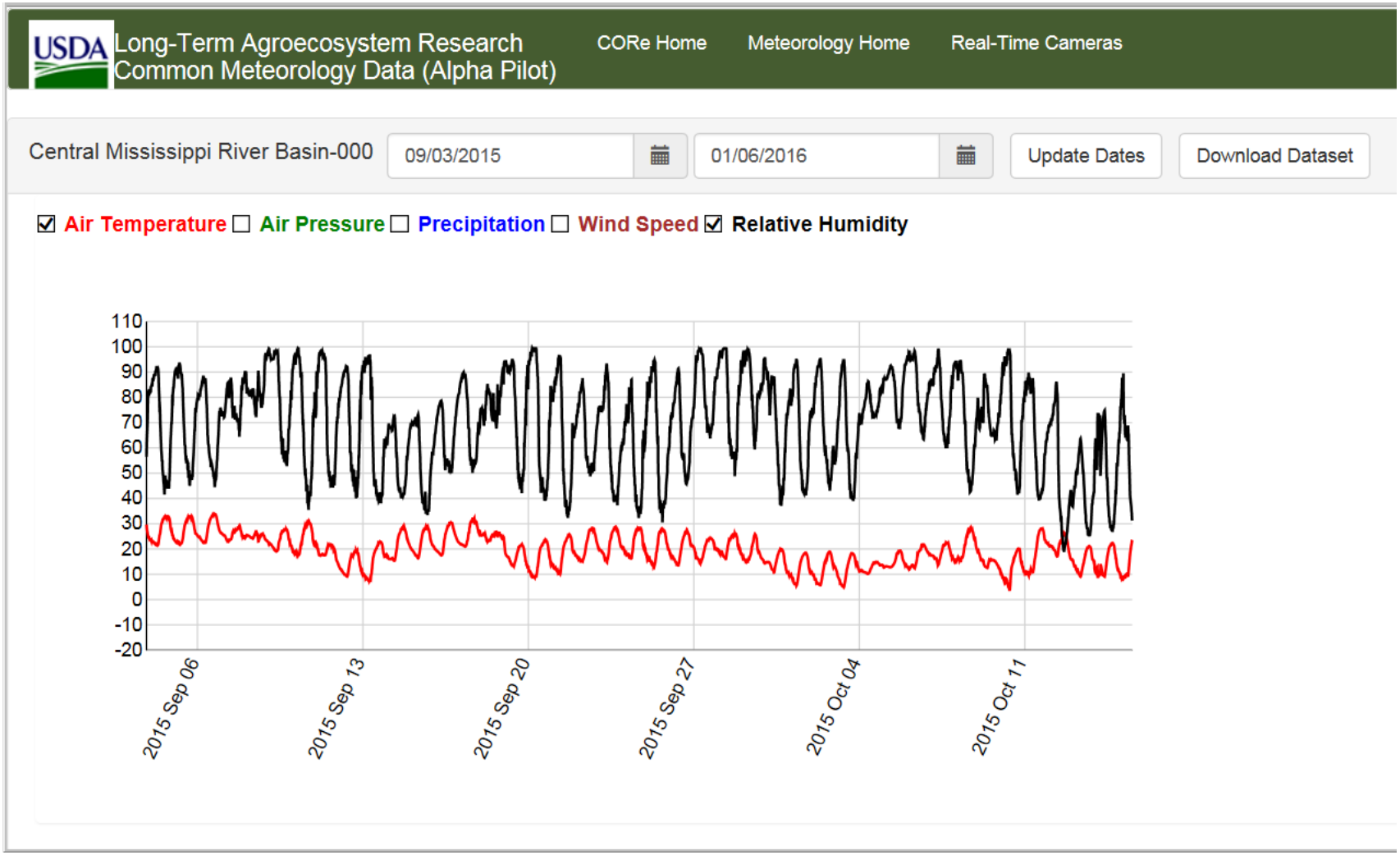
Current Status

- Meteorological Data:
- Near Real-Time Data for 11 stations at 7 LTAR locations
- Test batch data from 4 stations at 2 LTAR locations
- Phenocam images for 18 cameras at 7 LTAR sites
  - Phenocam network performs image analysis to measure seasonal changes (spectrum)

Future Value-Enhancements

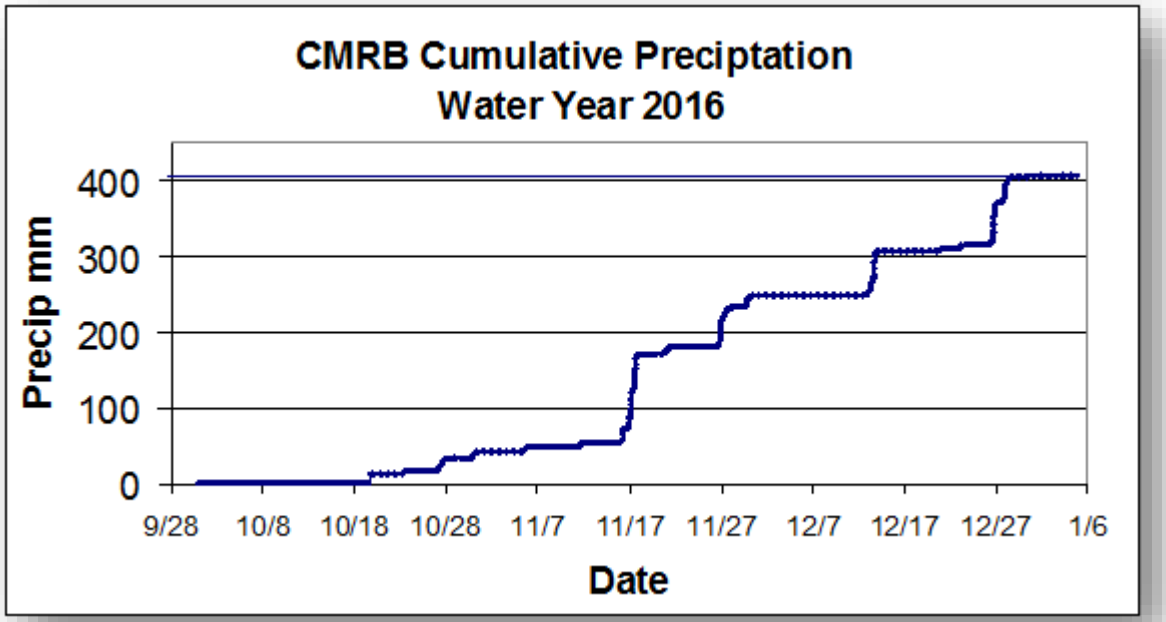
- Respond to environmental events (e.g. flooding)
  - Tie-in LTAR data (graphic display of cumulative precipitation)
  - Create a news feed
- Build out API Web services
  - ability to automatically serve downloads
- ***And….***

Near Real-Time Meteorological Data



Future Enhancements

- Greatly enhanced information visualization and user interaction with data
- Possible timely news features



Cumulative precipitation at Central Mississippi River Basin LTAR site (Columbia, MO) ties into the current flooding in Missouri (and downstream)

https://data.nal.usda.gov/

USDA Ag Data Commons Beta
ARS National Agricultural Library

Datasets About News Login

Highlighted Program: Long Term Agroecosystem Research

A network of 18 research sites around the U.S. are collecting long-term data on agriculture, climate change, ecosystems and natural resource conservation. These data are vital to developing safe and sustainable agricultural production systems.

Search Ag Data Commons

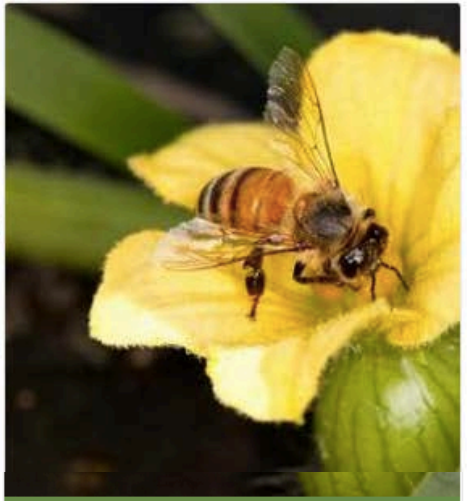
Topics

- Agricultural Products
- Agroecosystems & Environment
- Animals & Livestock
- Bioenergy
- Food & Nutrition
- Genomics & Genetics
- Maps & Multimedia
- Plants & Crops

Highlighted Datasets

Scientists are using the Blue Berry Genomics Database (BBGD) to investigate which genes are involved in acclimation to cold in blueberries.

The Baylor College of Medicine recently sequenced and annotated the *Oncopeltus fasciatus* genome as part of the i5k pilot project.



What is the role of pathogens in Honey Bee Colony Collapse Disorder? Deep RNA sequencing was used to study microbial diversity and seek answers.

LTAR Data @ Ag Data Commons

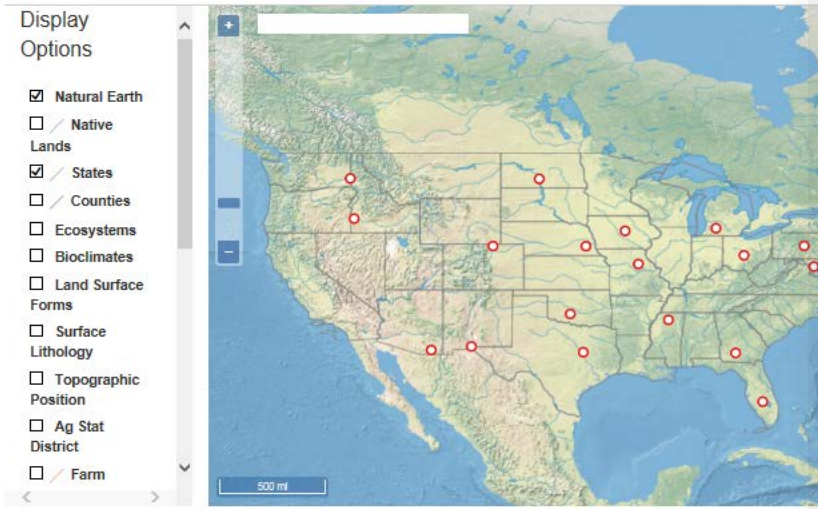
Home / Long Term Agroecosystem Research Landing

Long Term Agroecosystem Research Overview

Agriculture faces tremendous challenges in meeting multiple, diverse societal goals, including a safe and plentiful food supply, supplying sources of bioenergy, improving water/air/soil quality, and maintaining biodiversity. The LTAR network enables long-resource regions to address these challenges. The goal of this research network is to ensure sustained crop and livestock prod agroecosystems, and to forecast and verify the effects of environmental trends, public policies, and emerging technologies. Ult array of clients, partners, and stakeholders with four basic outcomes:

1. Agroecosystem productivity is sustainably enhanced by the development and application of new technologies
2. Mitigation and adaptation of agroecosystems to climate change is improved by more accurate predictions of resource re
3. Stronger linkages to other long-term research networks improves conservation and environmental quality in agricultura
4. The socio-economic viability of, and opportunities for, rural communities are enhanced through educational outreach by

For the fully interactive version of this LTAR data explorer visit https://ltar.nal.usda.gov.



Click on the LTAR sites on the map to see more information about each site. Other national ecosystem research networks are contributes to nation-wide coverate. The map also links for LTAR data available through the Ameriflux, Soil Climate Analysis N networks.

LTAR Research Sites

Data from the following LTAR sites are presented. They are related to topics such as agricultural sustainability, climate change, ecosystem services, and natural resource conservation at the watershed or landscape scale.

- Archbold Biological Station (ABS), Venus, FL
- Cook Agronomy Farm (CAF), Pullman, WA
- Central Mississippi River Basin (CMRB), Columbia, MO
- Central Plains Experimental Range (CPER), Cheyenne, WY; Fort Collins, CO; Nunn, CO
- Eastern Corn Belt (ECB), Columbus, OH; West Lafayette, IN
- Gulf Atlantic Coastal Range (GACP), Tifton, GA
- Great Basin (GB), Boise, ID; Burns, OR
- Jornada Experimental Range (JER), Las Cruces, NM
- Kellogg Biological Station (KBS), southwestern MI
- Lower Mississippi River Basin (LMRB), Oxford, MS; AR; LA; MO
- Northern Plains (NP), Mandan, ND
- Platte River-High Plains Aquifer (PRHP), Lincoln, NE
- Southern Plains (SP), El Reno, OK
- Texas Gulf (TG), Riesel, TX
- Upper Chesapeake Bay (UCB), University Park, PA
- Upper Mississippi River Basin (UMRB), Ames, IA; MN; WI
- Walnut Gulch Experimental Watershed (WGEW), Tucson, AZ; Tombstone, AZ
- Lower Chesapeake Bay (LCB), Beltsville, MD

Datasets

107 datasets

Search... | Date changed | Desc | Apply | Reset

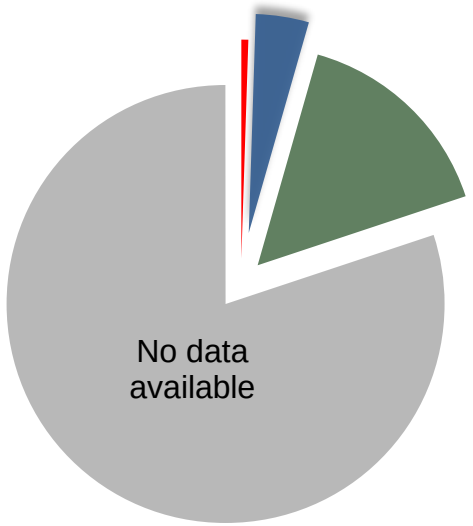
Upper Washita River Experimental Watersheds: Nutrient Water Quality Data

https://data.nal.usda.gov/long-term-agroecosystem-research

Publication-related LTAR Data

Publications Related

- LTAR site proposal references were reviewed for related data
  - 194 of more than 500 citations were reviewed
  - These citations are cross referenced to full-text papers in PubAg



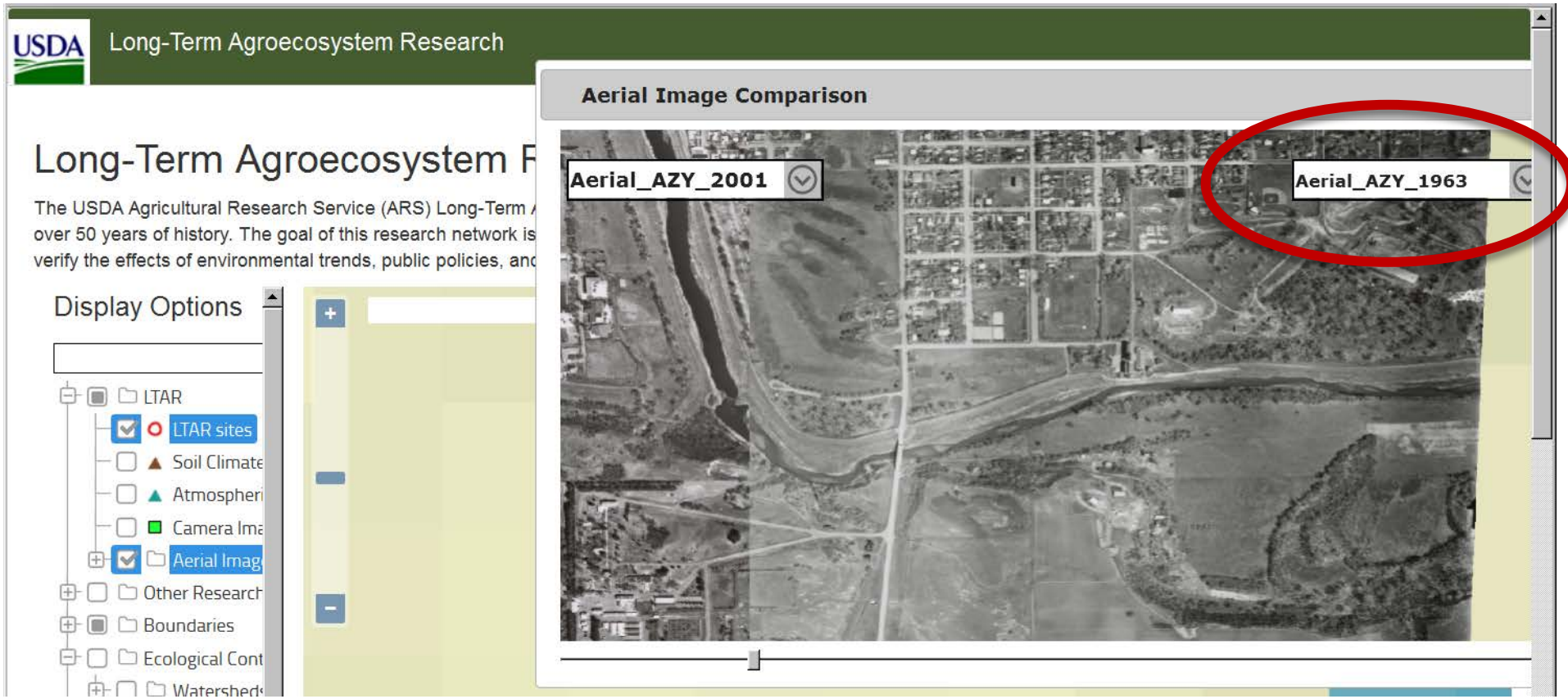
4% of Data are accessible

107 LTAR-Data Records in Ag Data Commons (including records for other network data, e.g. SCAN)

Aerial Imagery

- More than 600 images from Mandan, ND and Cheyenne, WY; NAL:
  - Digitized the images
  - Created metadata records
- Under development: image comparison tools
- Expected public release this fiscal year

Aerial Imagery: Sneak Peek



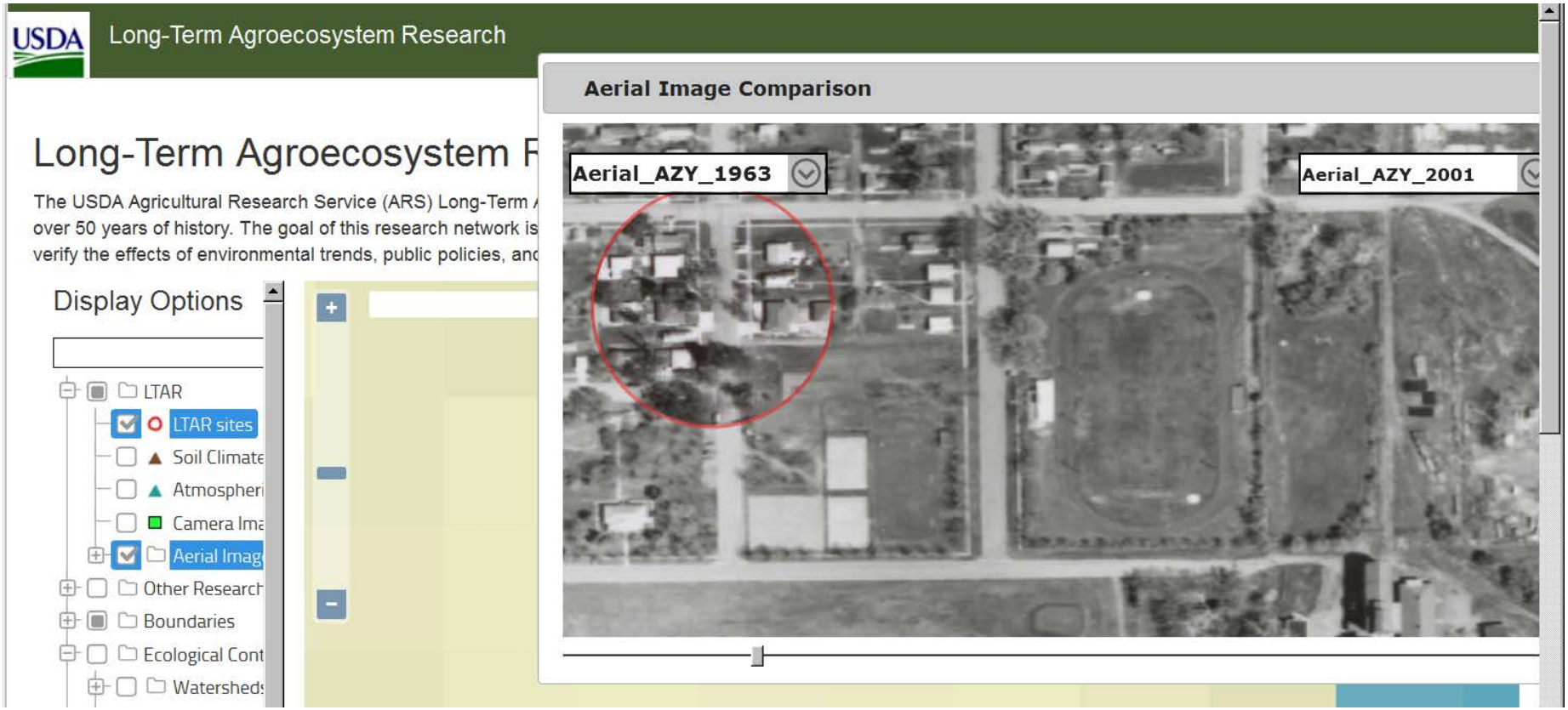
Land occupation 1963

Aerial Imagery: Sneak Peek

Land occupation 2001

Aerial Imagery: Before & After

Spyglass shows 2001 image over the 1963 base



Immediate Next Steps for LTAR Data

- Continue to add new LTAR sites with near real-time meteorological data
- Coordinate QA/QC processes for LTAR-wide comparability (ARS statisticians)
- Establish a means to collect “station event” metadata (sensor calibration; wildlife interference; major weather events, etc.)
- Publish the Mandan and Cheyenne aerial images

LTAR Project Acknowledgements

Jeffrey Campbell	Cynthia Parr	Ursula Pieper
Susan McCarthy	Bruce Vandenberg	NuCivic Team
Mark Walbridge	Jorge Delgado	Qing Qu
Charlie Walthall	Simon Liu	Gary Moore
Sally Schneider	Jim Ascough	Jiabin Heng

Jeffrey Campbell

Jeffrey.campbell@ars.usda.gov

301-504-5767

USDA

Ag Data Commons Acknowledgements

Susan McCarthy, NAL – KSD
Ursula Pieper, NAL – ISD
Qing Qu, NAL – KSD contractor
Jeff Campbell – NAL – KSD
Jaylen Nathwani, NAL – student intern
NüCivic, Angry Cactus Team
Jocelyn McNamara -- NAL – KSD contractor
Kerry Huller – UMD graduate fellow
Erin Antognoli – UMD graduate fellow

Cynthia.Parr@ars.usda.gov

USDA