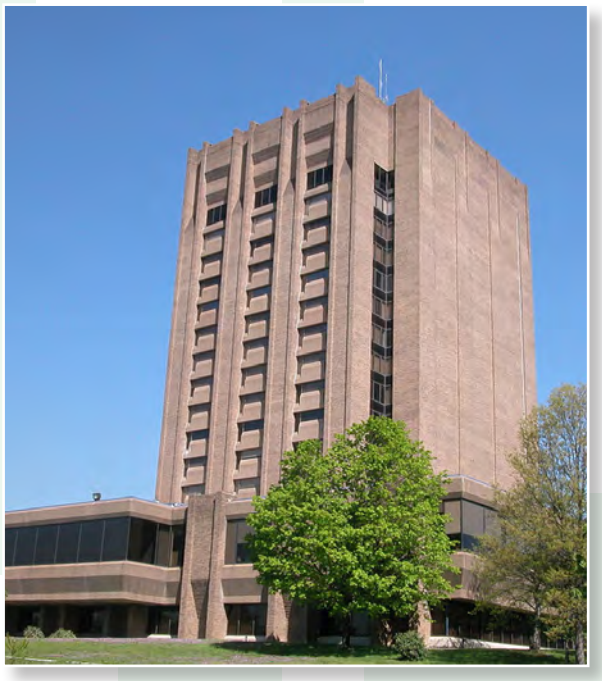
AgMIP

The Agricultural
Model Intercomparison
and Improvement Project

REPORT FROM THE AGMIP-USDA DATA HARMONIZATION WORKSHOP

MAY 11-15, 2015
NATIONAL AGRICULTURAL LIBRARY
BELTSVILLE, MARYLAND



Report from the AgMIP-USDA Data Harmonization Workshop
May 11-15, 2015
National Agricultural Library
Beltsville, MD

Table of Contents

Report from the AgMIP-USDA Data Harmonization Workshop
May 11-15, 2015
National Agricultural Library
Beltsville, MD

Introduction

A workshop was organized by AgMIP and USDA to determine how to reduce the wide gap between the goal of open, accessible, and usable agricultural data vs. the current reality regarding data collected by researchers nationwide in projects funded by USDA-ARS and NIFA. Vastly greater value could be obtained if data were combined across locations, time, and management conditions so that researchers could develop or evaluate models that help inform decision support systems, assess the benefits of new technologies or management, analyze impacts of changes in climate, and evaluate tradeoffs between productivity gains and environmental risks. There are various initiatives aiming to improve data access and use that need to be harmonized so that we do not end up with multiple, different data management solutions that are difficult to interconnect. The workshop was hosted by the National Agricultural Library in Beltsville; there were 55 invited attendees.

The goals of the workshop were to understand how to harmonize agricultural data collected from sites across the USA, demonstrate how a National Agricultural Data Network might work, develop ideas for a roadmap on how to create such a network, and make recommendations to the USDA (ARS and NIFA) for developing an operational data network. Specific objectives were to:

1) Develop a prototype system to harmonize databases from representative NIFA and ARS projects that will make data accessible and usable for multiple crop models and for statistical analyses;
2) Expand the AgMIP IT tools used to operate multiple crop models to include nitrogen and phosphorus inputs and outputs and to complete translators for additional US-based cropping system models;
3) Select and document metadata and minimum variables that should be included in harmonizing data in other USDA research areas (e.g., dairy, beef, Life Cycle Assessment, and biofuels);
4) Create recommendations for USDA and a draft roadmap that leads to broader harmonization of data with capabilities for on-line publication of harmonized, discoverable, accessible, and usable datasets.

These efforts will help researchers and USDA staff comply with the 2013 Federal Open Data Policy.

Workshop Program and Breakout Sessions

The full agenda for the workshop is given in the Appendix. The workshop consisted of a series of plenary sessions in which presentations were made to expose attendees to current efforts by the USDA National Agricultural Library (e.g., the Ag Data Commons, https://data.nal.usda.gov/), by AgMIP (e.g., https://data.agmip.org/cropsitedb), and others to make data and associated publications available and to datasets provided by participants for testing the data harmonization national web access approaches.

The workshop broke into four sessions for carrying out the work needed to reach the objectives. Breakout Session 1 included corn and wheat researchers who provided example datasets in addition to IT and database managers from corn and wheat projects and AgMIP. In these breakouts, participants worked on harmonizing data and uploading datasets to an AgMIP data node that had been implemented on the NAL server by NAL and UF/AgMIP staff prior to the workshop. They also worked on the web interface to the data to allow datasets to be discovered and downloaded. This set of Breakout Sessions (#1) continued throughout the workshop, working with data from seven locations.

The second set of Breakout Sessions consisted of crop modelers and IT experts who are working on different crop modeling teams and on the soil model intercomparison effort in AgMIP. This set of Breakout Sessions worked on translators to convert the AgMIP harmonized data into input files for running their models, reviewed the metadata and ICASA data dictionary, and expanded the variables to include a minimum set of needed variables for including N and P soil and crop model analyses and for simulation of long-term crop rotations. This Breakout session spanned the duration of the workshop.

A third set of Breakout Sessions identified metadata and minimum sets of variables that are typically collected and needed for dairy, beef, and life cycle analysis models. A major output of these breakouts was an initial set of metadata for each type of system identified by the dairy, beef, and LCA projects. In addition, these breakouts developed initial lists of the minimum set of variables that should be harmonized for each type of system, including variable names, descriptions, and units. These Breakouts also continued throughout the workshop.

A fourth set of Breakout Sessions was held on Days 1 and 2 of the workshop for the purpose of developing recommendations for USDA and AgMIP for future work and to develop ideas for a roadmap that would lead to a National Agricultural Data Network. This breakout included leaders in USDA (ARS and NIFA), and the NAL and AgMIP who discussed strategies for expanding this effort to ultimately create a distributed system of harmonized databases. These efforts will also assist researchers and USDA staff in complying with the 2013 Federal Open Data Policy.

Progress in all of the teams was presented during the last morning of the workshop in a plenary session that was videoed and streamed to others who were unable to attend on the last day. These video presentations are being posted to the AgMIP web site.

Toward a Shared Vision and Goal

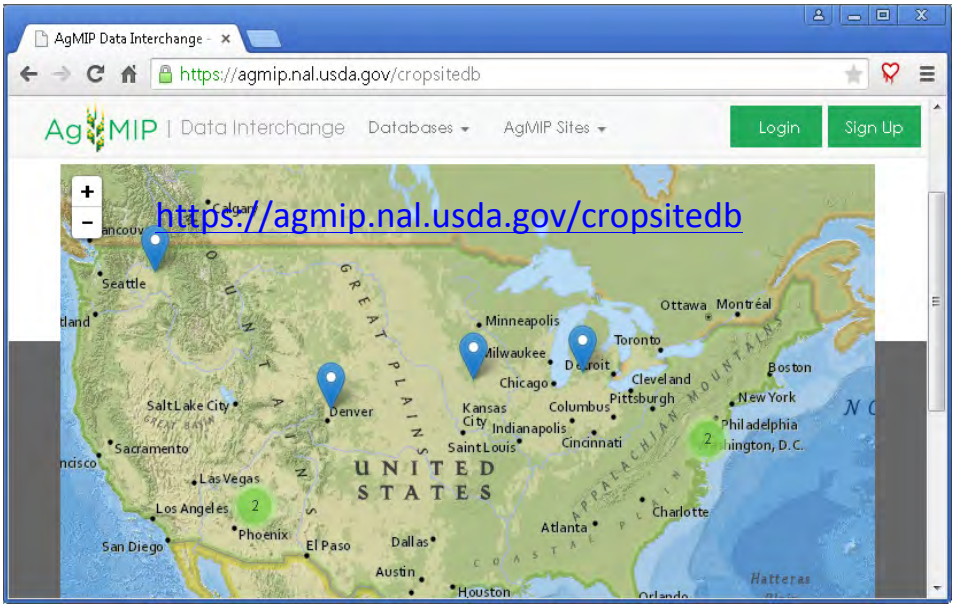
New statements were drafted for a shared vision and goal for a national effort on data harmonization for access and use. The vision statement drafted was that we should strive to develop a "*distributed network of linked, compatible agricultural databases into which researchers provide data that are easily shared among users with maximum impact of contributions and harmonized for easy discovery, open access and use in models and statistical analyses*". The complexity of agricultural challenges facing the nation and the world are such that agricultural data stewardship and advanced tools are needed to enable sustainable production that can meet future national and international food, fiber, and bioenergy needs. A National Agricultural Data Network (NADN) will accelerate progress towards sustainability and resilience to a changing climate by greatly enhancing the efficiency with which data from USDA-supported research are applied to research on agricultural systems analysis and modeling.

Workshop Progress

<u>Objective/Breakout Session 1</u>
Progress on this objective exceeded our plans. The participants in this breakout successfully harmonized seven datasets from ARS and NIFA locations across the USA:

- USDA ARS – Corn data from ARS site in Ames, Iowa (Jerry Hatfield)
- USDA ARS - Wheat FACE data from ARS site in Arizona, Maricopa, AZ (Jeff White)
- USDA ARS - China Wheat data collected by ARS, Lubbock TX (Jeff White)
- REACCH NIFA CAP - Wilke 2013 Spring Wheat, Davenport, WA (University of Idaho)
- USDA ARS - LIRF Corn data (means), Greeley, Colorado (Pat Bartling and Laj Ahuja, ARS)
- MSU - LTAR Kellogg Biological Station (Brian Baer and Bruno Basso, Michigan State)
- USDA ARS – BARC Corn dataset from ARS sites in MD & DE (Dennis Timlin, ARS)

All data were uploaded into the NAL AgMIP database site. The figure below shows the locations of the sites, where the circle with numbers indicates the number of datasets represented by that dot. More details about the datasets and progress made during the workshop are given in an attached file.



Considerable progress was made by participants from the Corn Cap on aligning the data codes from the CAP with the ICASA Data Dictionary used to harmonize the databases for the NAL site. Work will be continued to complete this work so that Corn CAP data can be loaded to the NAL site when investigators in that project are ready to release their data. The overall path forward includes finalizing data started in this workshop, adding more LTAR sites, and mapping the ICASA data dictionary to variables from GraceNET and REAP sites.

Objective/Breakout Session 2.

Good progress was made on the objectives to include N and P data and to add translators to allow other models to use the harmonized data. In particular, N and P data are now available for the input and output translators and are usable for DSSAT, CROPSYST and APSIM models. Translators for the SALUS and WOFOST models are still under development. Harmonization of USDA data during the workshop now allows ensemble model use of the data, thus facilitating model intercomparison and improvement and re-use of data beyond original scope of the research in which data were collected. See attached for more details on accomplishments from this set of breakout sessions.

Objective/Breakout Session 3

This set of breakout sessions worked on 1) a review of metadata used in the NAL's Ag Data Commons, 2) developing a shared understanding of the roles of data dictionaries, vocabularies, and ontologies in harmonizing, searching, accessing, and using data that are collected across sites and research initiatives, and 3) developing data dictionary terms and metadata for dairy, livestock, and life cycle analysis databases. A very good start was made on describing metadata for the Dairy CAP and on variables for data needed for a data dictionary. For the livestock CAP, existing data were reviewed and initial ideas were developed for continuing work to harmonize data for the livestock CAP. Also, the existing ICASA Data Dictionary used for the NAL AgMIP data node was reviewed relative to variables that need to be

included in order for the harmonized data to be used for Life Cycle Analysis. Additional work is needed for each of these three types of data, with input from domain experts as well as IT specialists. This breakout also identified additional domains where work is needed to develop metadata and data dictionaries for harmonizing data useful in those domains (including data on phenomics and genomics, plant pathology, and insect pest. Additional details are in the attached reports presented (as power point presentations) during the workshop.

Objective/Breakout Session 4

Participants in this set of breakout sessions during days 1 and 2 discussed what is needed to move forward toward a National Agricultural Data Network (NADN) with harmonized data that are easily discovered (e.g., through a NAL portal), accessible, and usable for cross-space and -time analyses in models and other tools. We drafted three components for use as recommendations to USDA (ARS, NIFA, and NAL) administrators and staff and to AgMIP contributors, all aimed at achieving the goal and vision summarized above in this report. First, we listed important guiding principles that should be strived for in developing a NADN. Then, we outlined a Roadmap to ensure that the technical design and community endorsement and use of the NADN are successful. Finally, we summarized recommendations for consideration by NIFA, ARS, the NAL and AgMIP to invest in financial, human, and institutional resources to create a NADN that could very likely become a defacto standard way for agricultural open data to be harmonized, shared, and used. Each component is presented below.

Guiding Principles

1. Fair attribution to data providers (joint authorships, citations, etc.)
2. Compatibility with international efforts
3. Community-driven approaches
4. Open access
5. Inclusion
6. Quality
7. Standards and protocols
8. Shared vision
9. Transparency
10. Support for the spectrum of agricultural production systems
11. Sustainability in environmental, social, and economic realms

Roadmap for a National Agricultural Data Network (NADN): Data-Model Integration Initiative

1. Conduct inventories of existing systems and datasets, both national and international (e.g., from EarthCube, the CGIAR and AgMIP to benefit from lessons learned in those initiatives). This includes evaluating and reconciling existing standards for data and metadata, data dictionaries, ontologies and prioritizing datasets to be harmonized for distribution. Determine relationships among the Ag Data Commons, AgMIP, REEPort, NSF DataOne, and other systems and what coordination is needed.
2. Develop a funding model for supporting the design and development of a national agricultural data network, with support to contributing members and clear accountability.
3. Design infrastructure including data dictionaries, ontologies, and use cases that are compatible with international efforts. Establish a USDA NAL-AgMIP-Ag Professional Societies team to jointly extend the data dictionary or dictionaries and metadata that are needed to serve as the

backbone of data harmonization (to involve the broader agricultural modeling community). Start with a workshop focusing on LTAR data and evolving a common data dictionary that will join communities around data management and harmonization, modeling, and tool development (including LTAR leaders/scientists, NAL, and AgMIP).

4. Develop priority use cases, implementation and harmonization tools, meta data, pilot databases, and user interfaces for data provision and access and QAQC procedures (e.g., LTAR, GxExM, climate change impacts, …).
5. Address legal issues including data policies, licensing, attribution, and liability related to data use.
6. Engage stakeholders in this process, including researchers in government and universities as well as those from the private sector who may use the data system created. The Tri-society leadership can help identify partners in crops, soils, agronomy, and environment. This group can help advocate for this NADN effort with key leaders in government.
7. Build capacity and culture for open data sharing and model integration. Education and outreach should include:
    a. Symposia, workshops, special meetings, newsletters, and journal papers to (re)educate students and existing professionals about data science
    b. Tools to assist researchers in recording and managing data in harmonized formats.
    c. Tools for authors, reviewers, and editorial boards to ensure published results/data can be accessed and easily reused
    d. Formal on-campus courses to educate graduate students about data. Include content in a course on data that is analogous to the ethics training currently required of graduate students.
8. Conduct a series of workshops to continually engage the community in various activities associated with items 1-7.
9. Assess use and impact: identify suitable metrics, Alt Metrics, etc.
10. Sustainably curate and maintain the data and tools

Recommendations to USDA Administrators

1. *Strengthen and clarify draft ARS data policy*. This should involve sending draft to selected groups and individuals for them to review and make recommendations for strengthening and clarifying the draft. For example, require that terms and conditions must have a data management section describing how and when they will provide metadata and full datasets from funded projects. There should be guidelines in the RFPs providing concrete options with clear requirements and consequences of noncompliance. Contracts would then require submission of metadata and data by certain times. Encourage publication in one of the new data journals. Publish and archive data and associated papers. Data provision compliance would be required in addition to a data management plan. Policy must be relevant going forward with global networking.
2. *Invest in a data repository at NAL* plus a portal for connecting to other locations/nodes in a National Agricultural Data Network (NADN). The NAL would host the retrievable metadata with pathways for access to the full set of data via APIs that harmonize a minimum set of priority data that can also be retrieved for analyses. The data portal would allow access and have the legal terms of data use downloaded/displayed to inform users the terms of their use of accessed data (e.g., citations). Downloads and other metrics would be automatically collected. Data would be identified by the DOI for citations. Legal statement about data usage expectations and crediting can be attached through the portal; these must be consistent with "open data" stipulations.
3. *Invest in the broader community for IT tool development and data provision* to make it easy for researchers/data providers to upload data and for modelers and others to access and use these harmonized data. Some of this investment should be allocated to the community via grants and other mechanisms to ensure that high quality data and effective/efficient processes for provision, access, and usability.

4. *Implement an initial three- to five-year initiative* with funding to build the NADN. The modeling and data and metadata standards need to be defined by the research communities, some of which are just getting started (e.g., Dairy, Beef, Grazing, LCA, etc.). The agricultural modeling community should help identify metadata and minimum set of observations that need to be harmonized and be engaged to help design and develop tools. NIFA CAPs, ARS Labs, and AgMIP should be engaged in the initiative to work with the NAL in the 3 to 5-year initiative for designing and developing the network.
5. *Develop a sustainable and adaptive business model.* Work with funding agencies to fairly and robustly include true costs of data sharing in overall research budgets and institutional infrastructure support on a continuing basis.
6. *USDA leaders work with Experiment Station Directors and Deans to promote the NADN* and to have them get buy-in from Land Grant institutions and faculty. There is a need for the Directors and Deans to incentivize faculty and to understand the value of data metrics for institutions as well as individual faculty members. There is a need to have a process to engage these administrators and faculty across the nation, being careful not to be too top-down or there could be push back.
7. *Support additional workshops, building on success of the recent one at the NAL.* Several workshops were suggested, each targeting specific objectives for designing and developing the NADN. These workshops will join communities around data management and harmonization, modeling, and tool development. Suggested examples were:
    - LTAR leadership/scientists, NAL, and AgMIP with a focus on harmonizing long term experiment data from the past, tools, and plan for future improvements (also connect with LTERs)
    - Crop modelers and genomics/plant breeder group, for connecting genotype and phenotype data to enable G x E x M analyses and gene-based modeling of traits and processes
    - Climate Change Management Adaptation Strategies – Identify the minimum data sets and harmonization required to support models. One could focus specifically on NIFA CAP goals.
    - Training workshops for researchers aimed at helping to help them understand and comply with the policy and to help develop a data-friendly culture among ag researchers
    - Pests and Diseases research and modeler community for defining metadata and minimum data needed to model pest and disease dynamics and impacts on production
    - Bioenergy CAPs x AgMIP (Bioenergy MIP), also links to the Global Research Alliance, LCA analysis is part of this and other themes
    - Model-data integration, developing use cases for a range of agricultural systems, using data to improve models and vice versa.
8. *Develop incentives as well as consequences for non-compliance to data policy.* "Carrots" will generally be preferred because "sticks" may result in poorly described, marginally useful data being provided in order to "check the box".

AgMIP The Agricultural Model Intercomparison and Improvement Project

AgMIP-USDA Data Harmonization Workshop

May 11-15, 2015

National Agricultural Library

Beltsville, MD

James W. Jones
University of Florida

USDA NIFA

United States Department of Agriculture
National Institute of Food and Agriculture




AgMIP The Agricultural Model Intercomparison and Improvement Project

Toward a Shared Vision:

Distributed network of linked, compatible agricultural databases into which researchers provide data that are easily shared among users with maximum impact of contributions and harmonized for easy discovery, open access and use in models and statistical analyses

USDA ARS

USDA NIFA

UF IFAS UNIVERSITY of FLORIDA




AgMIP The Agricultural Model Intercomparison and Improvement Project

Goal:

The complexity of agricultural challenges facing the nation and the world are such that agricultural data stewardship and advanced tools are needed to enable sustainable production. A NADN will accelerate progress towards sustainability and resilience to changing climate by greatly enhancing the efficiency with which data from USDA research programs are applied to research on agricultural systems analysis and modeling.

USDA ARS

USDA NIFA

UF IFAS UNIVERSITY of FLORIDA




AgMIP The Agricultural Model Intercomparison and Improvement Project

Objectives

1. Implement a prototype approach to harmonize databases from several USDA and NIFA projects that will make data accessible and usable for multiple crop models and for statistical analyses (Breakout 1),
2. Expand the AgMIP IT tools to include nitrogen and phosphorus inputs and outputs and to complete translators for at least two additional US-based cropping system models (e.g., among CROPSYST, EPIC, SALUS, WOFOST, and RZWQM) (Breakout 2),
3. Develop and document metadata and minimum variables that should be included in harmonizing data in other USDA research areas (e.g., dairy, beef, Life Cycle Assessments, and biofuels) (Breakout 3), and
4. Develop recommendations for USDA and others responsible for Agricultural R&D that will lead to broader harmonization of data with capabilities for on-line data publication, discovery, access, and use. (Breakout 4).

AgMIP The Agricultural Model Intercomparison and Improvement Project

Toward a Shared Vision:

Distributed network of linked, compatible agricultural databases into which researchers provide data that are easily shared among users with maximum impact of contributions and harmonized for easy discovery, open access and use in models and statistical analyses

USDA ARS

USDA NIFA

UF IFAS UNIVERSITY of FLORIDA




AgMIP The Agricultural Model Intercomparison and Improvement Project

Some Guiding Questions?

- Should AgMIP, USDA (NIFA and ARS), USAID, BMGF or some combination convene a series of workshops over the next few years to continue what we start in this workshop?
- Should funding agencies/donors require project to attend such workshops to continue development and implementation of a national data initiative for harmonizing data and models?
- Should donors require all projects to enter meta data w/in the first year and a minimum set of data after 2 years?
- Should USDA and other donors invest in tool and database development to make this a successful endeavor and ensure credit?
- Should funding agencies require modeling efforts that they fund to be connectable with the harmonized data?
- Should NAL and the AgMIP community co-develop the tools? Are there funds to do this, with engagement of the broader community?
- What to call the initiative? ADMI – the AgMIP Data-Model Initiative)?

Possible Recommendations

- USDA (NIFA and ARS/NAL) and AgMIP convene series of workshops over the next 3-5 years to implement tools and a process for developing and maintaining a national database network system..
- NIFA RFPs that provide opportunities for researchers to develop tools and databases for the National Database System.
- Develop language in NIFA RFPs that require project investigators to attend such workshops for training on use of the national database system for harmonizing data and models?
- USDA (NIFA and ARS) should require all projects to enter meta data w/in the first year and a minimum set of data after 2 years.
- USDA ARS (NAL) should invest in tool and database development to make this a successful endeavor, working with AgMIP to help ensure community buy-in and use.
- NIFA should require modeling efforts that they fund to be connectable with the harmonized data.

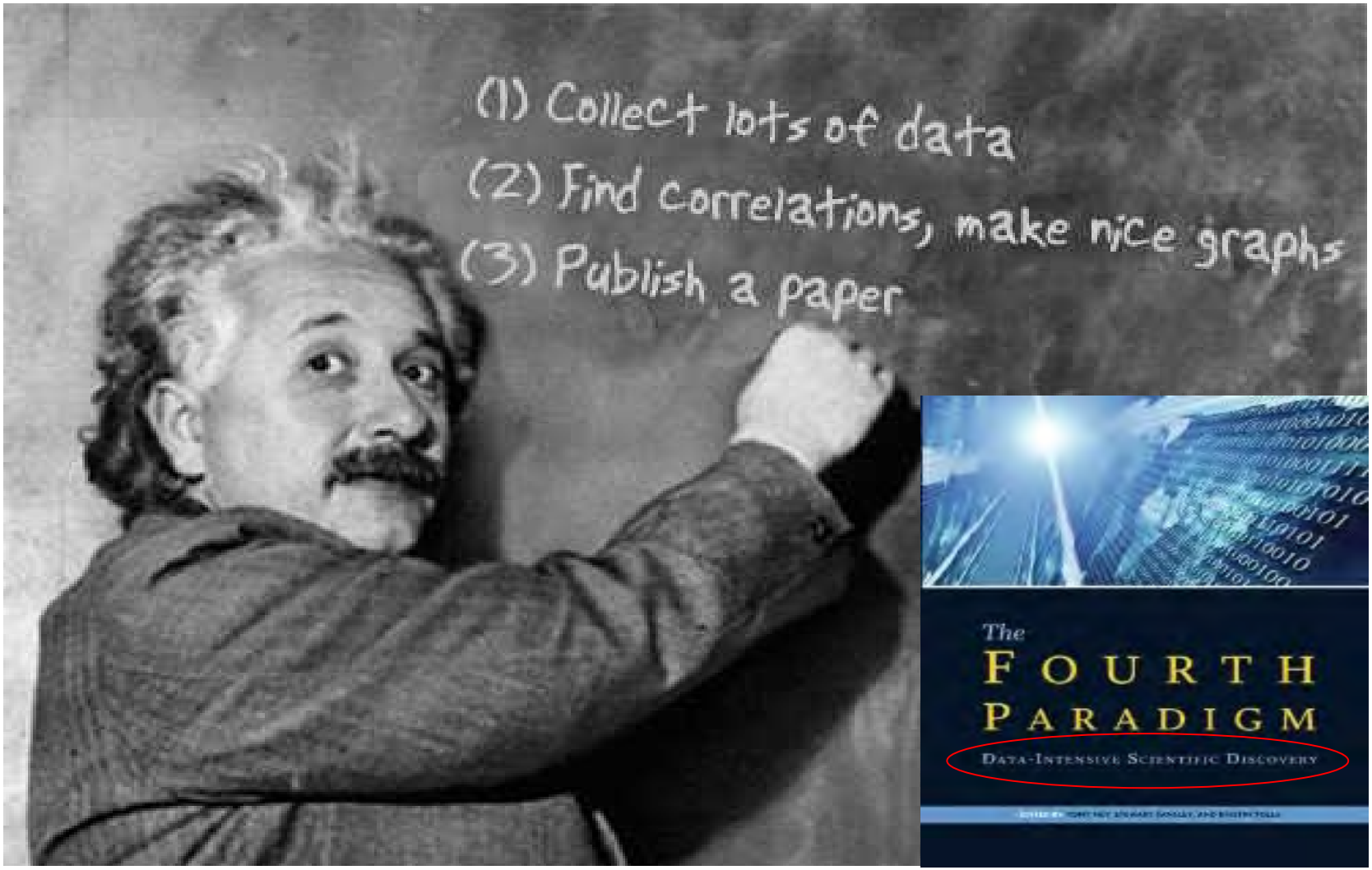
Toward a National Agricultural Data Network ~ Potential Roles of NAL

Simon Liu, Ph.D.
Associate Administrator
Agricultural Research Service
May 12, 2015

Agenda

- Introduction
- Drivers & Needs
- Draft Policies
- Technical Approaches
- A Logical Journey
- Q&A

Introduction



Drivers & Needs

- Mandates
  - America COMPETES Reauthorization Act (12/2010)
  - Office of Science & Technology Policy (OSTP) Public Access Memo (02/2013)
  - Executive Order – Making Open and Machine Readable the New Default for Government Information (05/2013)
- Research support
  - Data-intensive research
  - Computational-intensive research
  - Multidisciplinary research
  - Open science

Scholarly Publications (Draft Policy)

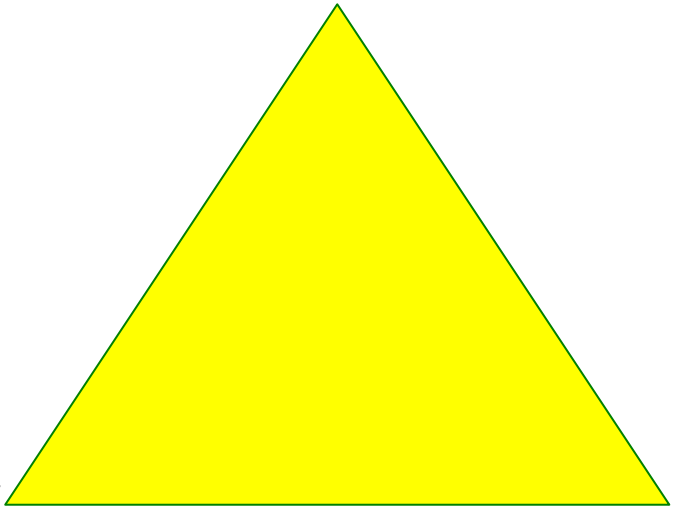
- Who is responsible?
  - Investigators/organizations funded by the USDA
- What is submitted?
  - Final peer-reviewed manuscripts or published papers
- Where to submit?
  - USDA public access archive system
- When to submit?
  - Upon acceptance for publication
- When to be made publically available?
  - No later than *12* months after the official date of publication

Scientific Digital Data (Draft Policy)

- Who is responsible?
  - Investigators/organizations funded by the USDA
- What are submitted?
  - High value datasets
  - Metadata associated with datasets
- Where to submit?
  - Recognized & public accessible databases
  - USDA scientific data catalog
- When to submit?
  - No later than *36* months of completion of data collection
- When to be made publically available?
  - No later than *36* months of completion of data collection

Pillars of National Agricultural Data Network

Scholarly Publications
(PubAg – Centralized)



Scientific
Digital Data
(Ag Data Commons – Distributed)

Models &
Tools
(AgMIP, …)

PubAg (I)

Data Management

APIs

Search & Discovery

Repository (Citations + Full Text)

Indexing

NAL Thesaurus (NALT)

Compliance Reporting

Manuscript Submission System

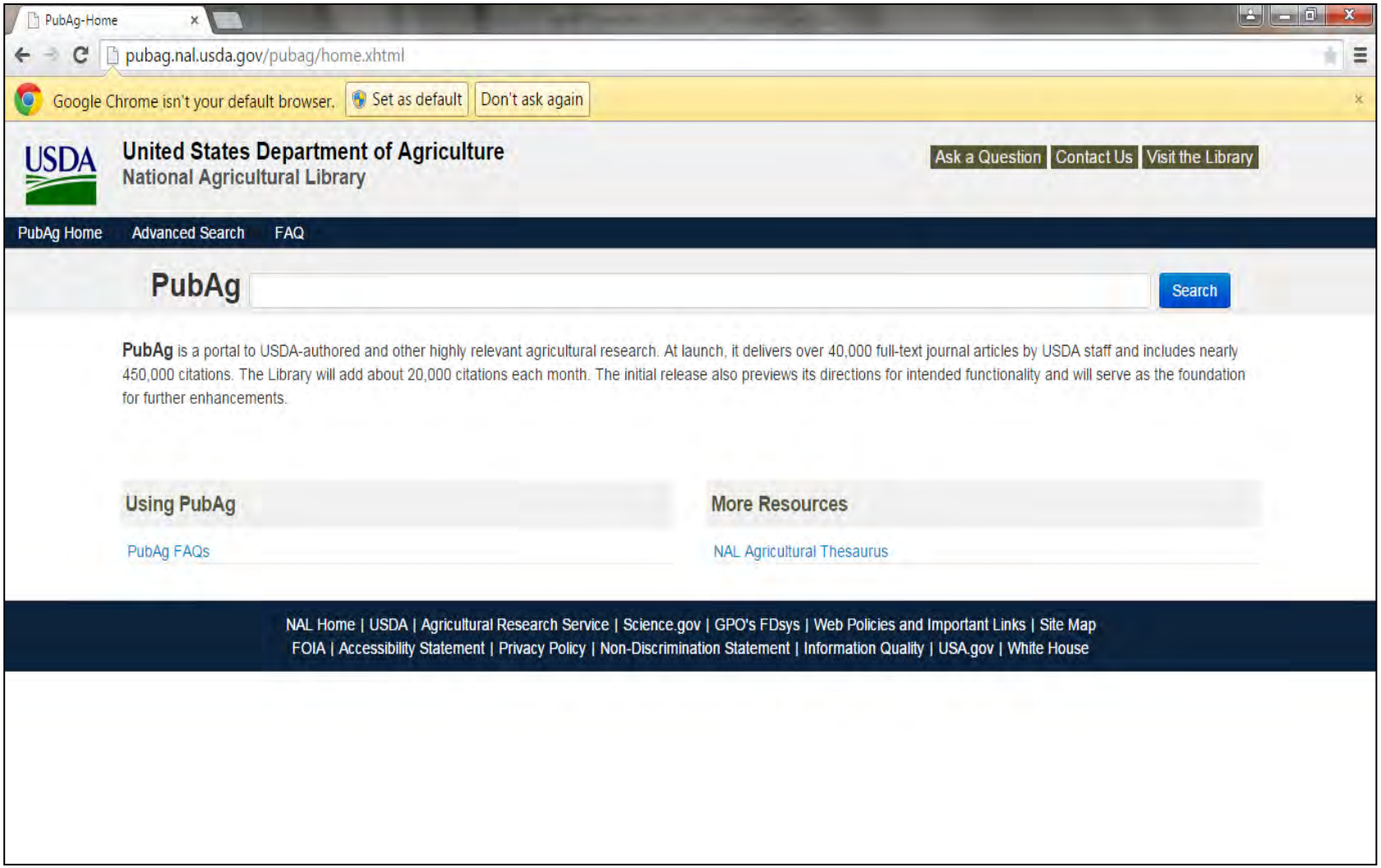
Authors

ELSEVIER

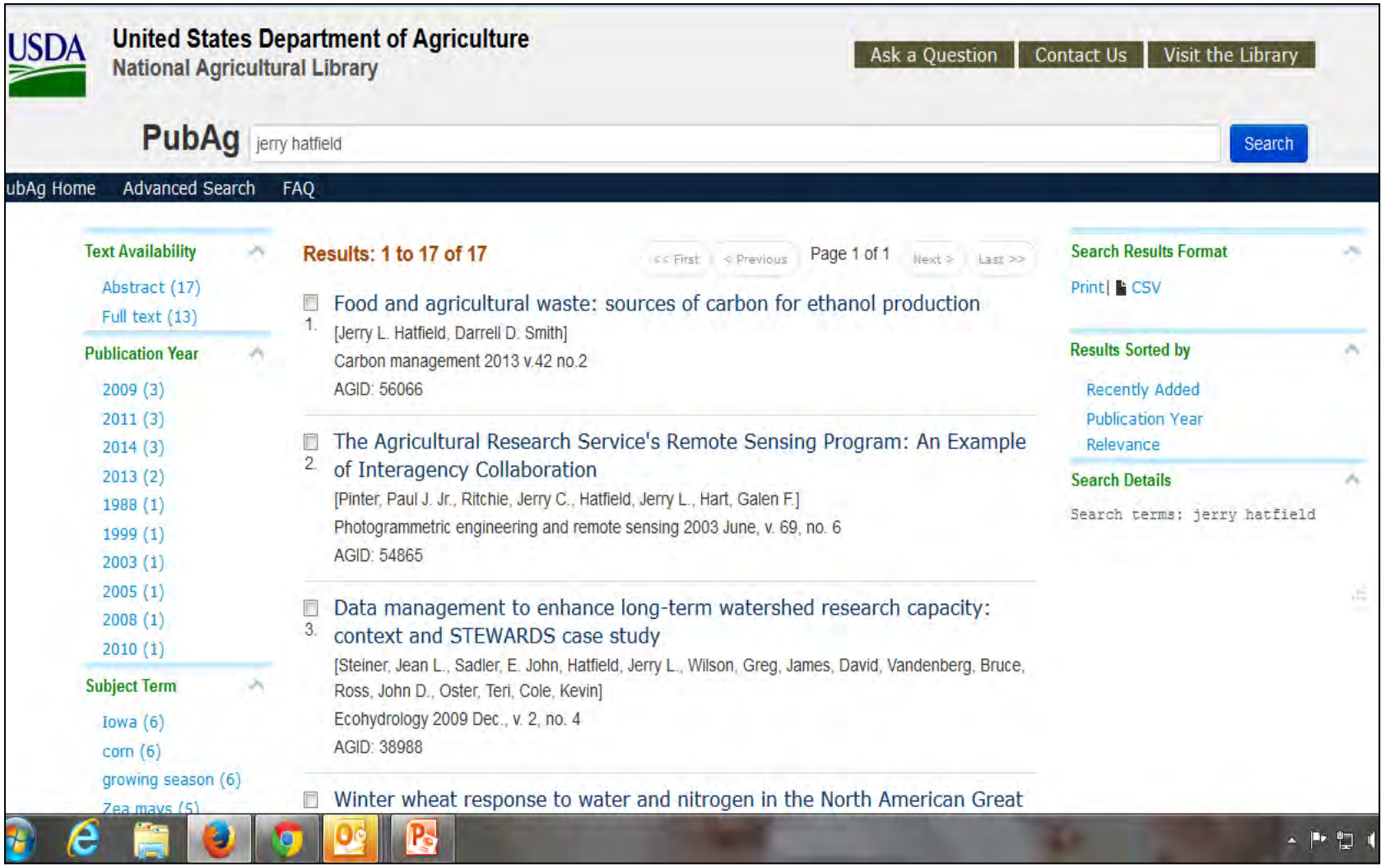
CHORUS

Publishers

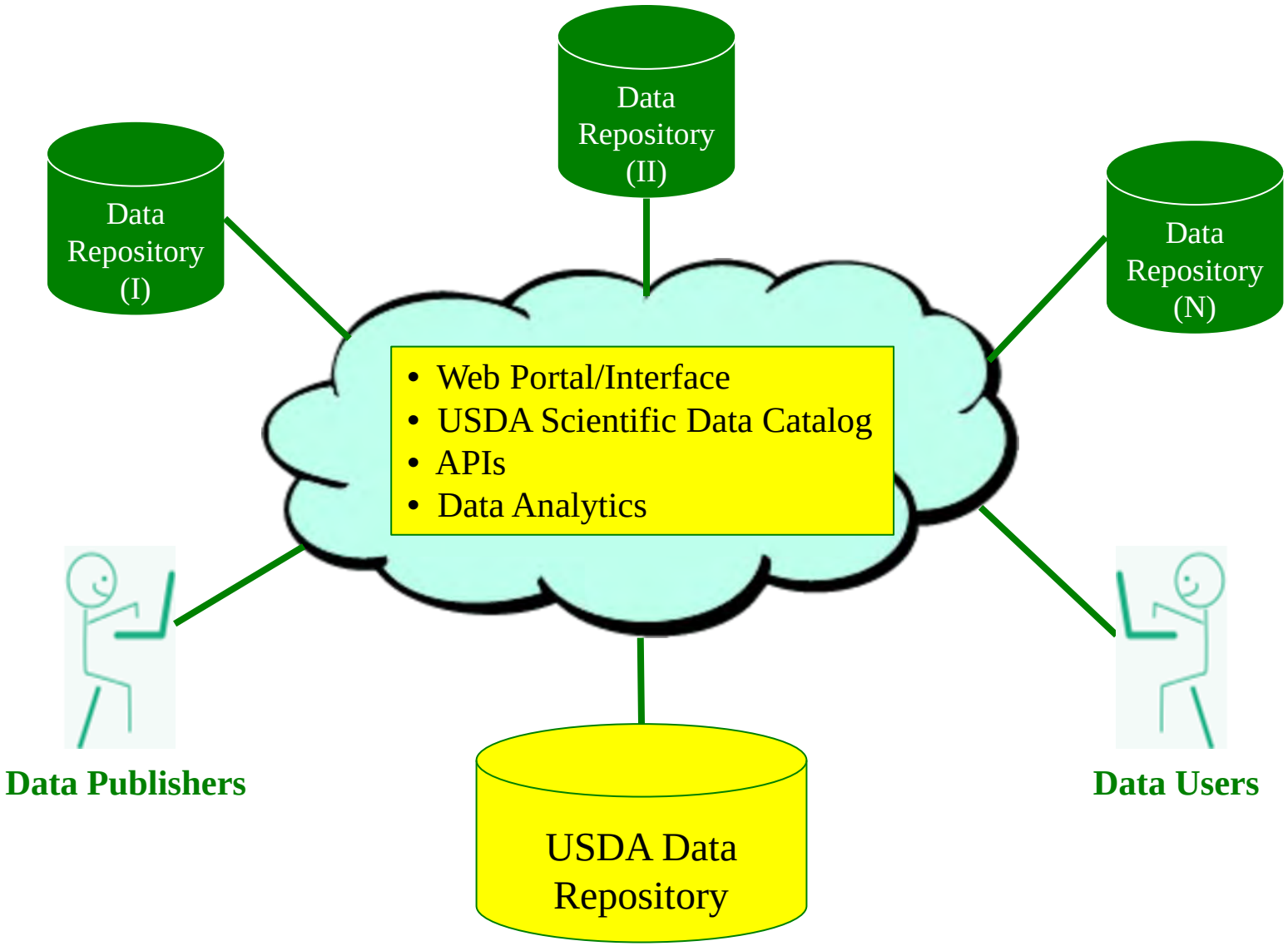
PubAg (II)


PubAg-Home
pubag.nal.usda.gov/pubag/home.xhtml
Google Chrome isn't your default browser.
Set as default
Don't ask again
United States Department of Agriculture
National Agricultural Library
Ask a Question
Contact Us
Visit the Library
PubAg Home
Advanced Search
FAQ
PubAg
Search
PubAg is a portal to USDA-authored and other highly relevant agricultural research. At launch, it delivers over 40,000 full-text journal articles by USDA staff and includes nearly 450,000 citations. The Library will add about 20,000 citations each month. The initial release also previews its directions for intended functionality and will serve as the foundation for further enhancements.
Using PubAg
PubAg FAQs
More Resources
NAL Agricultural Thesaurus
NAL Home | USDA | Agricultural Research Service | Science.gov | GPO's FDsys | Web Policies and Important Links | Site Map
FOIA | Accessibility Statement | Privacy Policy | Non-Discrimination Statement | Information Quality | USA.gov | White House

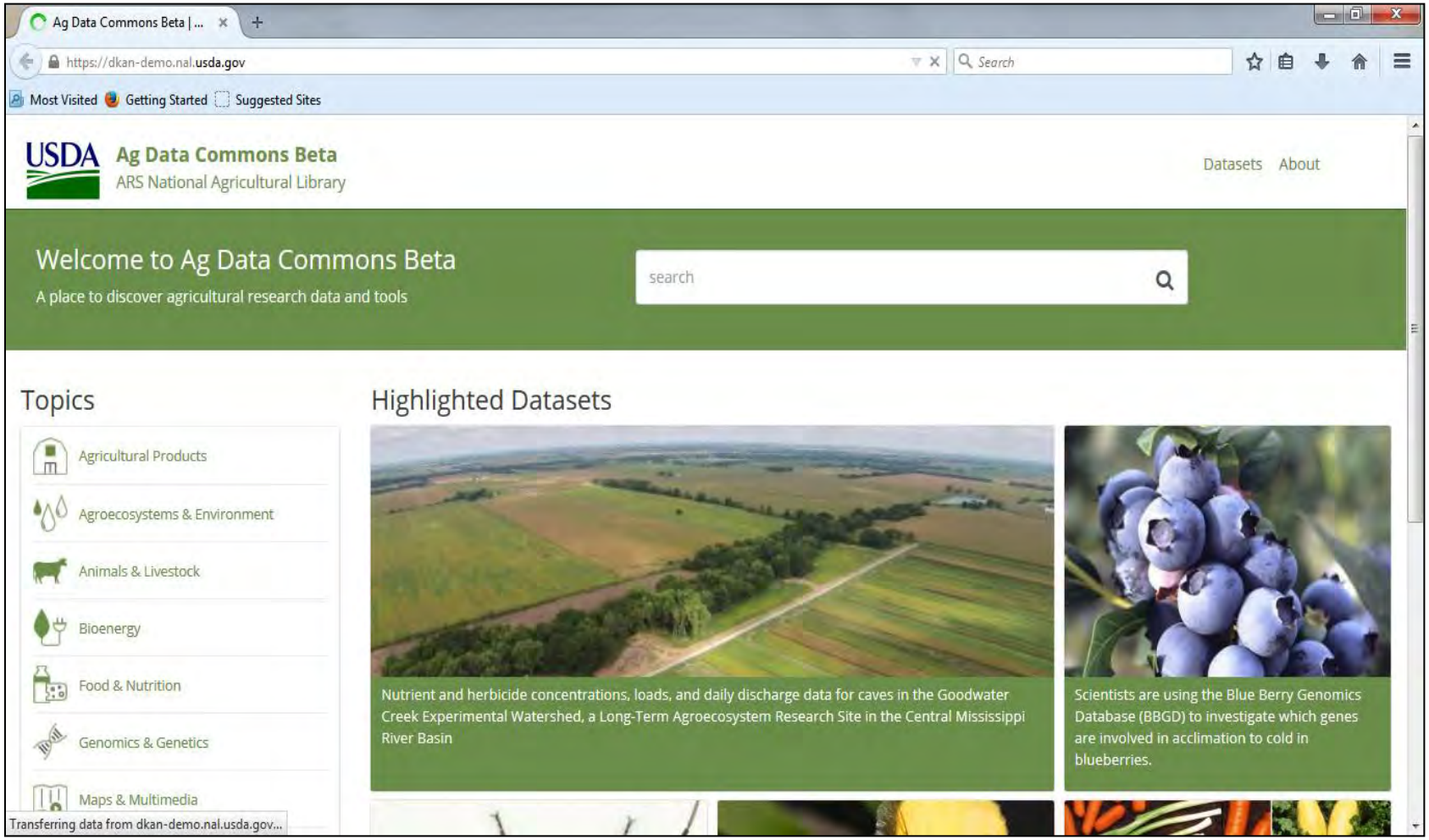
PubAg (III)


United States Department of Agriculture
National Agricultural Library
Ask a Question
Contact Us
Visit the Library
PubAg
jerry hatfield
Search
ubAg Home
Advanced Search
FAQ
Text Availability
Abstract (17)
Full text (13)
Publication Year
2009 (3)
2011 (3)
2014 (3)
2013 (2)
1988 (1)
1999 (1)
2003 (1)
2005 (1)
2008 (1)
2010 (1)
Subject Term
Iowa (6)
corn (6)
growing season (6)
Zea mays (5)
Results: 1 to 17 of 17
<< First
< Previous
Page 1 of 1
Next >
Last >>
1. Food and agricultural waste: sources of carbon for ethanol production
[Jerry L. Hatfield, Darrell D. Smith]
Carbon management 2013 v.42 no.2
AGID: 56066
2. The Agricultural Research Service's Remote Sensing Program: An Example of Interagency Collaboration
[Pinter, Paul J. Jr., Ritchie, Jerry C., Hatfield, Jerry L., Hart, Galen F.]
Photogrammetric engineering and remote sensing 2003 June, v. 69, no. 6
AGID: 54865
3. Data management to enhance long-term watershed research capacity: context and STEWARDS case study
[Steiner, Jean L., Sadler, E. John, Hatfield, Jerry L., Wilson, Greg, James, David, Vandenberg, Bruce, Ross, John D., Oster, Teri, Cole, Kevin]
Ecohydrology 2009 Dec., v. 2, no. 4
AGID: 38988
Winter wheat response to water and nitrogen in the North American Great
Search Results Format
Print| CSV
Results Sorted by
Recently Added
Publication Year
Relevance
Search Details
Search terms: jerry hatfield

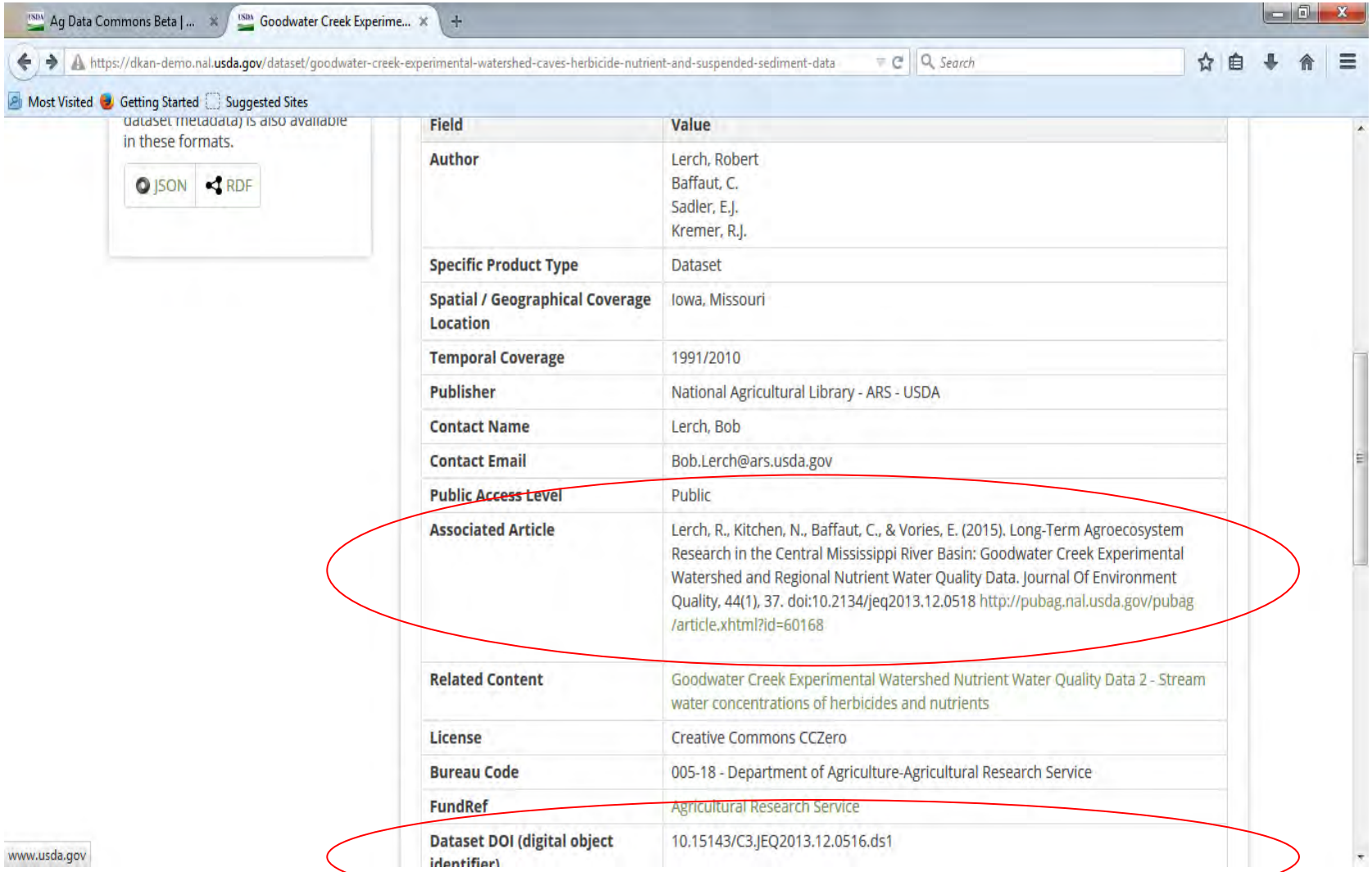
Ag Data Commons (I)



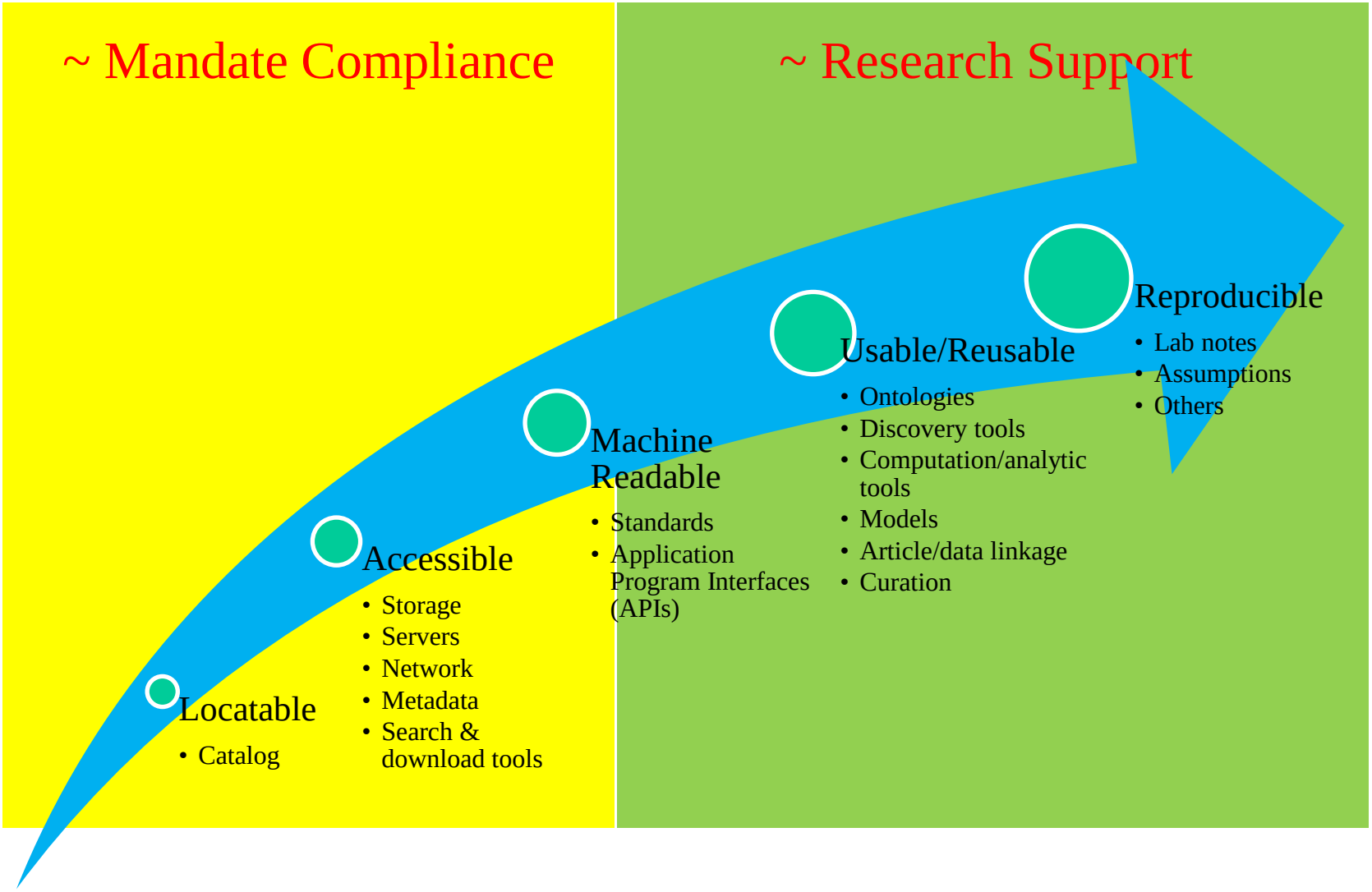
Ag Data Commons (II)



Ag Data Commons (III)



A Logical Journey




29
Q&A
Simon Liu
simon.liu@ars.usda.gov
202-720-3597

Harmonizing Agricultural Data for Open Access and Improving Cropping Systems Models

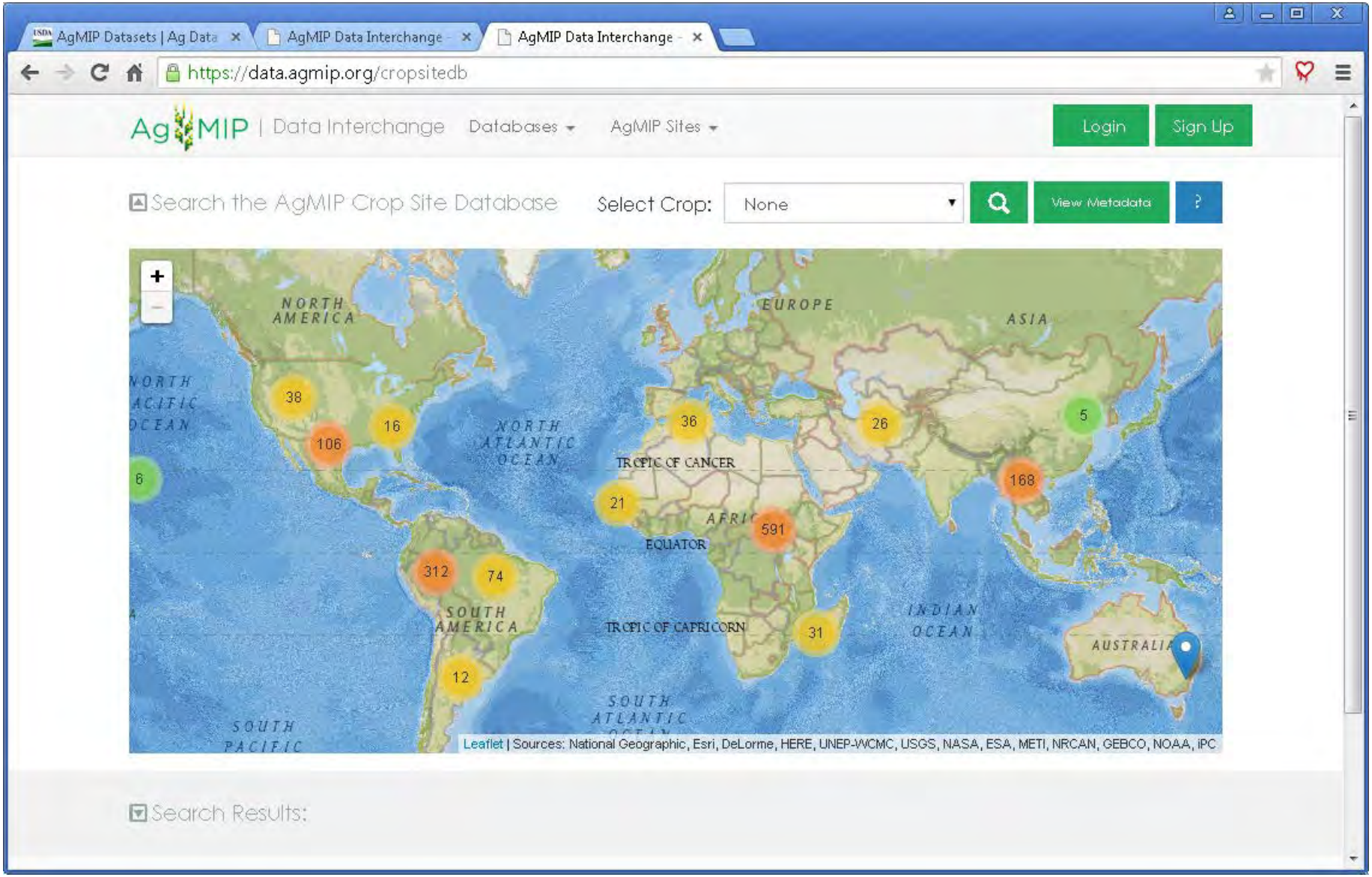
Joint AgMIP-USDA Workshop
National Agricultural Library
Beltsville, MD
Agenda
May 11-15, 2015

Breakout 1 Report

Breakout 1

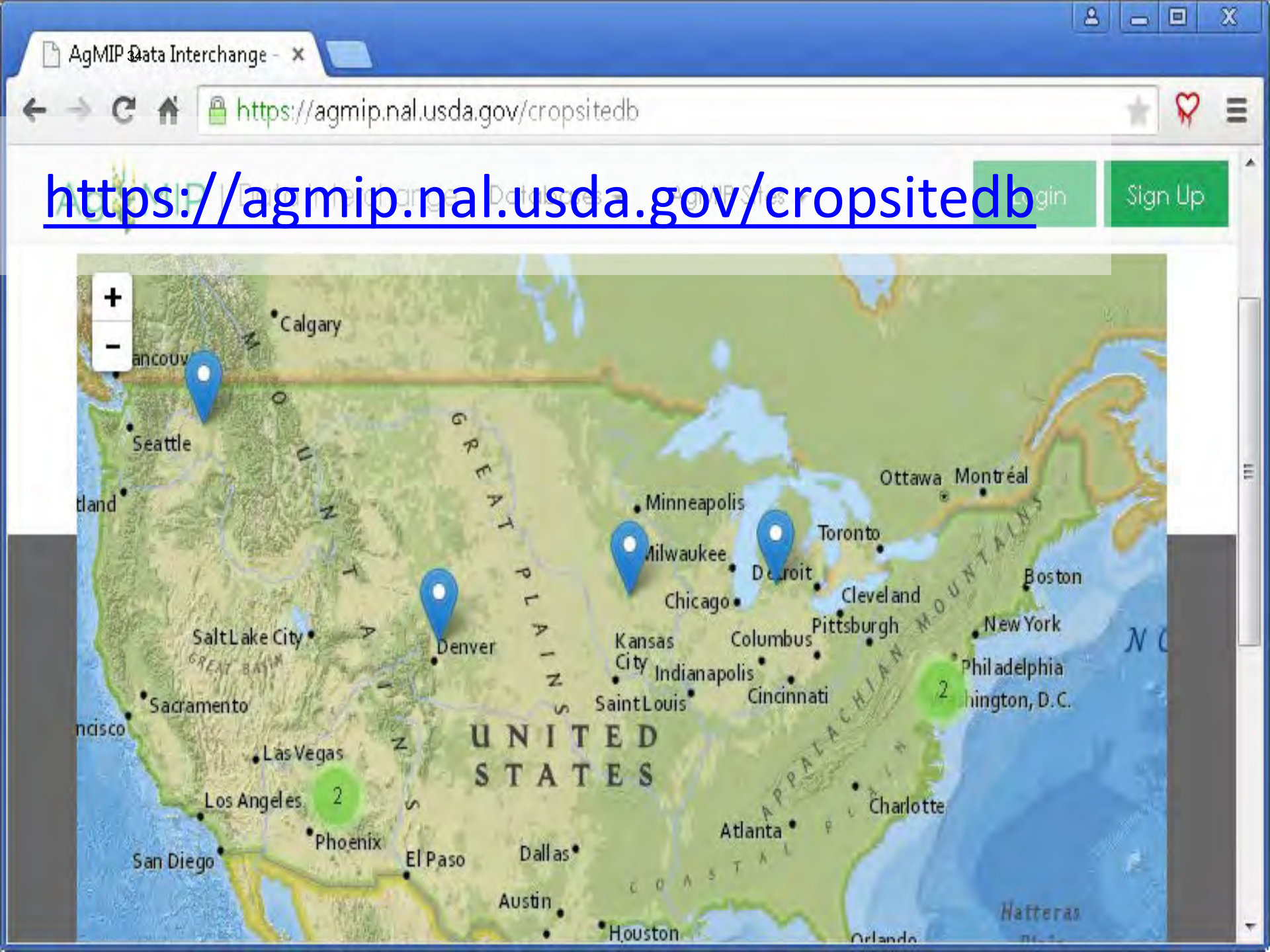
- Review implementation of metadata in corn and wheat AFRI and ARS projects as well as status of the databases, data dictionaries, and APIs for discovering and accessing the data via the internet.
- Develop specific goals for work to be completed during workshop along with an initial plan and timetable for what will be done after the workshop..

https://data.agmip.org/

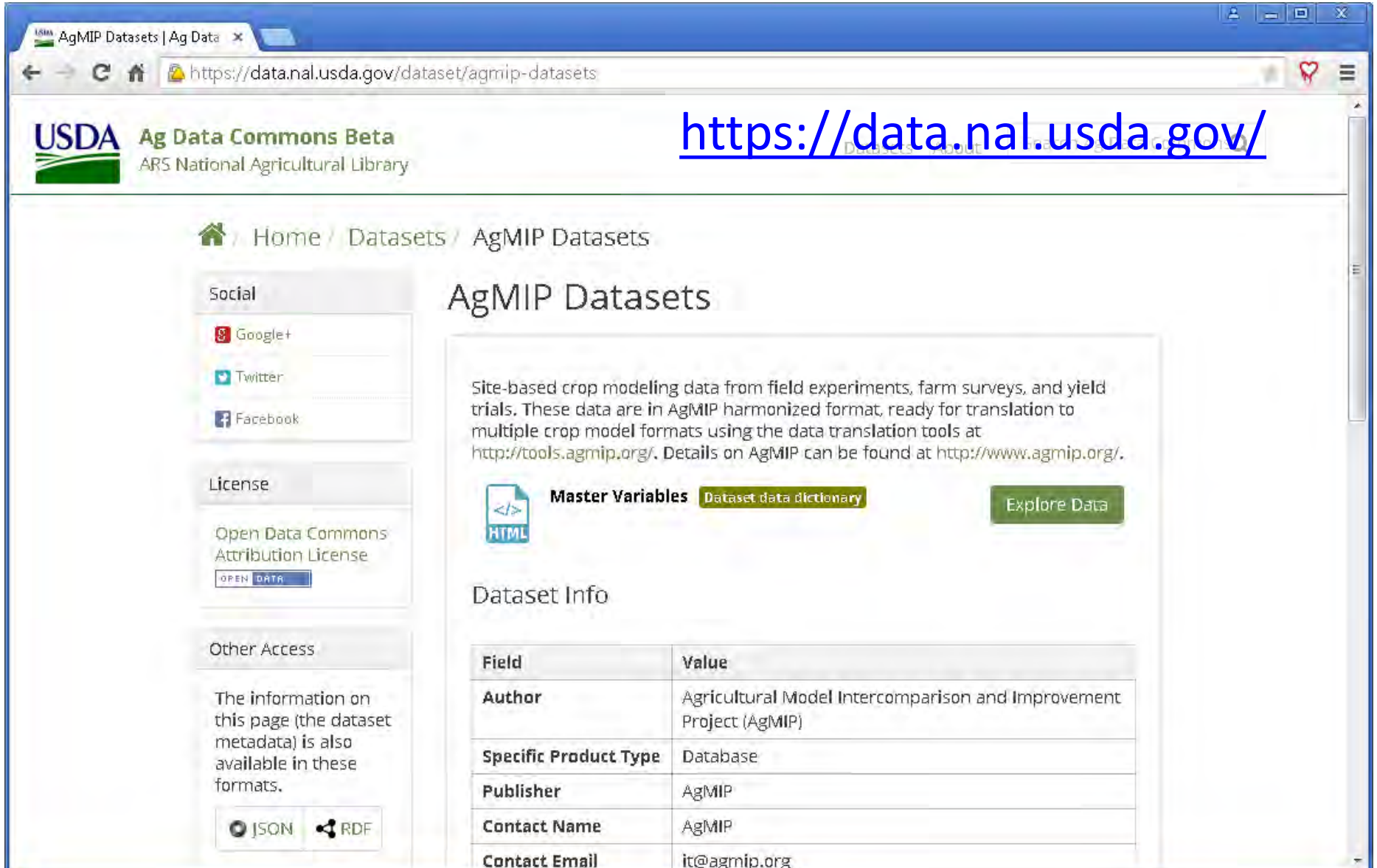


Datasets successfully uploaded to NAL

- USDA ARS - Corn, Ames, Iowa (Jerry Hatfield)
- USDA ARS - Wheat FACE, Maricopa, AZ (Jeff White)
- USDA ARS - China Wheat, Lubbock TX (Jeff White)
- REACCH PNA - Wilke 2013 Spring Wheat, Davenport, WA (University of Idaho)
- USDA ARS - LIRF Means Corn, Greeley, Colorado (Pat Bartling)
- MSU - LTAR Kellogg Biological Station (Brian Baer)
- USDA ARS – BARC Corn dataset, MD & DE (Dennis Timlin)


AgMIP Data Interchange
https://agmip.nal.usda.gov/cropsitedb
https://agmip.nal.usda.gov/cropsitedb
Sign Up
Calgary
Seattle
Minneapolis
Milwaukee
Toronto
Ottawa
Montréal
Boston
Chicago
Cleveland
Pittsburgh
New York
Salt Lake City
Denver
Kansas City
Columbus
Indianapolis
Philadelphia
Sacramento
Saint Louis
Cincinnati
Las Vegas
UNITED STATES
Los Angeles
Charlotte
Atlanta
Phoenix
San Diego
El Paso
Dallas
Austin
Houston
Hatteras
GREAT PLAINS
APPALACHIAN MOUNTAINS

AgMIP on Ag Data Commons



Agricultural and Forest Meteorology, 44 (1988) 105–116
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

WINTER WHEAT RESPONSE TO WATER AND NITROGEN IN THE NORTH AMERICAN GREAT PLAINS*

ROBERT J. REGINATO

USDA-ARS, U.S. Water Conservation Laboratory, 4331 East Broodway, Phoenix, AZ 85040 (U.S.A.)

JERRY L. HATFIELD

USDA-ARS, Cropping Systems Research Laboratory, Lubbock, TX 79401 (U.S.A.)

ARMAND BAUER

USDA-ARS, Northern Great Plains Research Laboratory, Mandan, ND 58554 (U.S.A.)

KENNETH G. HUBBARD, BLAINE L. BLAD, SHASHI B. VERMA

Center for Agricultural Meteorology and Climatology, University of Nebraska, Lincoln, NE 68583 (U.S.A.)

EDWARD T. KANEMASU

Evapotranspiration Laboratory, Kansas State University, Manhattan, KS 66506 (U.S.A.)

DAVID J. MAJOR

Agriculture Canada, Lethbridge, Alberta, Canada TIJ 4B1

(Received October 2, 1987; revision accepted May 2, 1988)

ABSTRACT

Reginato, R.J., Hatfield, J.L., Bauer, A., Hubbard, K.G., Blad, B.L., Verma, S.B., Kanemasu, E.T.,

USDA ARS Water Management Unit and Agricultural Systems Research Unit Collaboration
Limited Irrigation Research Farm, Greeley, Colorado

Producer Objective : The most crop for the drop in limited water resource environment in the west
ARS Objective : Sustainable production with limiting resources

Experiment:
Treatments: Vary the amount of irrigation water applied in terms of %ET between 40 and 100% ET. Applying water savings in the reproductive phase.

Data Set : Corn 2008-2011
Irrigation Levels of %ET: 40,55, 70, 70+reprod, 85, 100
Current Data : Observed Treatment Means, Biomass, Yield, Soil Water, LAI, plant height, ETcalc
Model Compatibility: RZWQM2 and DSSAT
Use in Crop ET Initiative for improving Crop ET simulation in Models.

USDA ARS – BARC Corn dataset

Modeling Temperature Responses of Leaf Growth, Development, and Biomass in Maize with MAIZSIM

Soo-Hyung Kim,* Yang Yang, Dennis J. Timlin, David H. Fleisher, Annette Dathe, Vangimalla R. Reddy, and Kenneth Staver

ABSTRACT

Mechanistic crop models capable of representing realistic temperature responses of key physiological processes are necessary for enhancing our ability to forecast crop yields and develop adaptive cropping solutions for achieving food security in a changing climate. Leaf growth and phenology are critical components of crop growth and yield that are sensitive to climate impacts. We developed a novel modeling approach that incorporates a set of nonlinear functions to augment traditional thermal time methods (e.g., growing degree days) for simulating temperature responses of leaf expansion and phenology in maize or corn (*Zea mays* L.). The resulting leaf expansion and phenology models have been implemented into a new crop model, MAIZSIM, that simulates crop growth based on key physiological and physical processes including C_4 photosynthesis, canopy radiative transfer, C partitioning, water relations, and N dynamics for a maize plant. Coupled with a two-dimensional soil process model, 2DSOIL, MAIZSIM was applied to simulate leaf growth, phenology, biomass partitioning, and overall growth of maize plants planted at two field sites on the Eastern Shore of Maryland and in Delaware for 3 yr of data. The model parameters were estimated using data from outdoor sunlit growth chambers and the literature. No calibration was performed using the field data. The MAIZSIM model simulated leaf area, leaf addition rate, leaf numbers, biomass partitioning and accumulation with reasonable accuracy. Our study provides a feasible method for integrating nonlinear temperature relationships into crop models that use traditional thermal time approaches without sacrificing their current structure for predicting the climate change impacts on crops.

Michigan State Univerity Kellogg Biological Station/LTER Main Cropping System Experiment

- Seven Treatments with six replicate blocks
- Started in 1989 – continuing
- Key personnel to collect/manage data set
  - Phil Robertson (Lead PI/Director)
  - Joe Simmons (Agronomy Manager)
  - Sven Bohm (Information Manager)

KBS LTER MCSE Treatments

- T1 Std input corn/soybean conv till
- T2 Std input corn/soybean no till
- T3 Low input wheat/corn/soybean w/cover crop
- T4 Zero input wheat/corn/soybean w/cover crop
- T5 Perennial biomass -- *Populus* trees
- T6 Perennial biomass -- Alfalfa
- T7 Successional community -- Historically tilled

REACCHPNA - Wheat CAP
Data Harmonization Progress 05.15.2015
www.reacchpna.org

Initial workshop goals

- Evaluate REACCHPNA data with regards to AGMIP needs
- Examine research station data for a select location
- Prepare discovery and agricultural metadata
- Test basic metadata upload
- Test access to data via REST compliant web services (optimal)

REACCHPNA - Wheat CAP
Data Harmonization Progress 05.15.2015
www.reacchpna.org

Dataset overview

- Wilke research station central WA
- 2013 spring wheat rotation
- Select fields: Biomass, harvest dates, yields, NxKxP soil testing, moisture testing

REACCHPNA - Wheat CAP
Data Harmonization Progress 05.15.2015
www.reacchpna.org

Results

- Test metadata records parsed and successful
- Test web service was implemented with JSON output (Wilke research station 2013 spring wheat)
- Initial parsing successful!!

REACCHPNA - Wheat CAP
Data Harmonization Progress 05.15.2015
www.reacchpna.org

Conclusion and next steps

- Web service model for data access is a potential solution for long term AGMIP data access
- ICASA variable matchups still need to occur, but existing variables for REACCHPNA appear to be close approximations
- Opportunity for REACCH cropping systems team to incorporate ICASA templates into workflow

Corn CAP - Accomplishments

- Umbrella site on Ag Data Commons has been created
  - Initially private
  - Will contain multiple datasets
  - Subset of data aligned with ICASA will include AgMIP tag
- AgMIP, NAL, and CAP understanding of mutual goals & synergies
- Alignment of Corn CAP and ICASA variables and metadata – 50% complete

Corn CAP – Next Steps

- Finish remaining 50% of work related to aligning ICASA codes with CAP data dictionary.
- Communicate with AgMIP on ICASA variables that need reconciling/clarification as well as determining how to code for corn CAP data that is not in ICASA. Use ISO standards or others?
- Determine how to integrate 55 corn CAP treatments into codes that are understandable by AgMIP, NAL, etc. Possibly use corn CAP codes as starting point if treatment codes are still being worked out.
- Programmatically switch codes in corn CAP database and generate exported test case dynamically from the database for upload testing by AgMIP.
- Discuss with AgMIP what metadata is considered minimum set for harmonization and of greatest use by modeling teams.
- Determine language, grouping of data, authorship, and assign DOIs for corn CAP data hubs on the NAL Ag Data Commons site. Expect to have edit capabilities this fall; make changes directly at that time.

The Path Forward

- Additional datasets
  - Finalize datasets started in this workshop
  - Soybean
  - More LTAR sites
- ICASA variables mapping
  - GraceNET / REAP
  - Stewards – watershed datasets
- New harmonized data dictionaries
  - Dairy, Grazing and Livestock
  - Pest and Disease
  - Life Cycle Analysis
- Long Term: Overlay ICASA data dictionary with ontology(ies) to facilitate linkage with other systems

Harmonizing Agricultural Data for Open Access and Improving Cropping Systems Models

Joint AgMIP-USDA Workshop
National Agricultural Library
Beltsville, MD
Agenda
May 11-15, 2015

Breakout 2 Report

Crop Models represented

- CropSyst – Roger Nelson, Claudio Stockle
- DSSAT – Meng Zhang, Cheryl Porter
- PSIMS modeling platform – David Kelly
- RZWQM2 – Pat Bartling
- Salus – Brian Baer, Bruno Basso
- WOFOST – Rob Knapen

- MaizeSIM - Dennis Timlin
- SpudSIM – David Fleisher

Nitrogen and Phosphorus modeling

- Soil variables
  - Inorganic N – NO3, NH4
  - Inorganic P – labile, active, stable
  - Organic – SOM pools, fresh organic matter pools
- Management variables
  - Fertilizer (inorganic) applications
    - NO3, NH4, urea
    - Labile P
  - Organic matter applications – N & P concentrations
- Plant variables
  - Observed plant concentrations and masses
  - In season, at maturity
  - Leaf, stem, whole plant, etc.

Status of Translators

- DSSAT
  - N & P data are used in ACE input and ACE output translators
- CropSyst & APSIM
  - N data are used in ACE output translator
- WOFOST
  - Under development
- SALUS
  - Under development

Accomplished this week

- CropSyst ACE input translator
  - Soils, weather & initial conditions complete
- PSIMS ACE input translator
  - Allows additional crop models to be run using regional and global gridded data with PSIMS platform
- SALUS – weather ACE output translators
- WOFOST – weather ACE output translator

Focused discussions

- Environmental (and other) Modifications
  - Protocol defined
  - To be implemented in June AgMIP regional workshop, Victoria Falls, Zimbabwe
- Sub-daily time steps
  - Strategy outlined
- Crop Rotations
  - pending

Summary – Breakouts 1 & 2

- Harmonization of USDA data to standard formats
  - Allow ensemble modeling exercises
  - Facilitate meta-analyses of projects and data
  - Facilitate improved science through
    - Model intercomparison
    - Model improvement
    - Greater understanding of interactions of genetics, environments, and management (GxExM)
  - Allow extensive re-use of data beyond scope of the original experiment

Breakout Session 3. Developing Metadata and Data Dictionaries for Other Agricultural System Components.

Co-led by Cheryl Porter (UF) and Jeff White (USDA ARS)

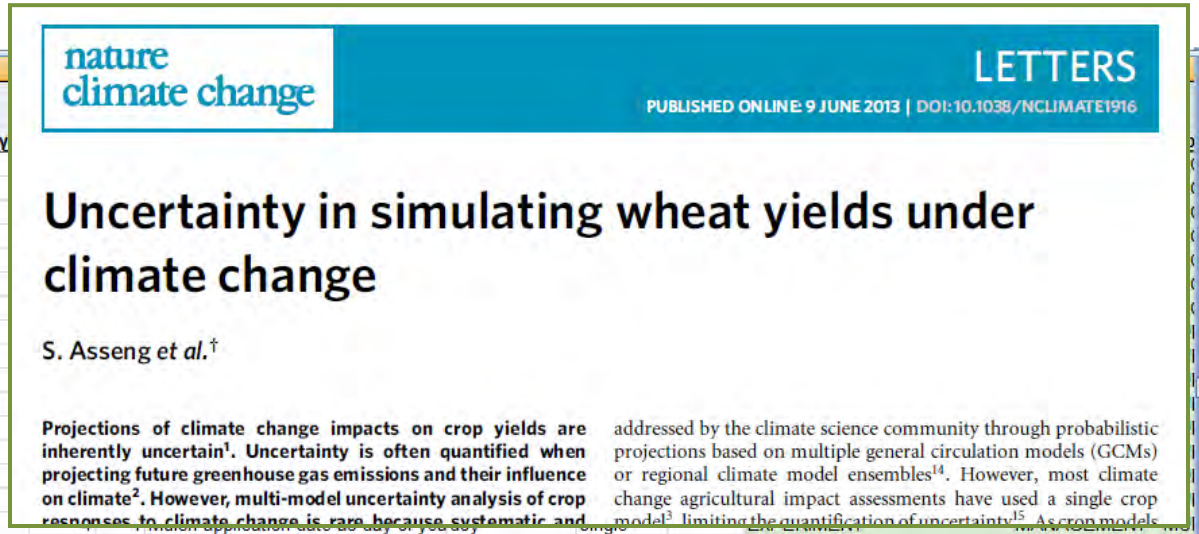
- Develop metadata for the following types of systems: dairy, beef, forestry, economics, and Life Cycle Analysis.
- Document these metadata and identify USDA databases (in NIFA, ARS, FRS, and ERS) that should be connected for use in model-based analyses and statistical analyses.
- Develop list of minimum set of variables that need to be harmonized.

Friendly reminder: ICASA/AgMIP Data Dictionary

- Former “Master Variable List”
- Framework for AgMIP data harmonization
- Best reached at: tinyurl.com/icasa-mvl

	A	B	C	D	E								
1	Variable_Name	Name_ Length	Code_Display	Code length	Code_Query								
308	org_matter_moisture_conc	24	OMH2O	5	OMH2O								
309	org_matter_carbon_conc	22	OMC%	4	OMCPC								
310	organic_material_N_conc	23	OMN%	4	OMNPC								
311	organic_material_P_conc	23	OMP%	4	OMPPC								
312	organic_material_K_conc	23	OMK%	4	OMKPC								
313	org_material_lignin_conc	24	OMLI%	5	OMLPC								
314	org_material_c_to_n	19	OMC2N	5	OMC2N								
315	id		id		id								
316	mulch_level	11	ML	2	ML								
317	mulch_level_name	16	ML_NAME	7	ML_NAME								
318	mulch_level_notes	17	ML_NOTES	8	ML_NOTES								
319	mulch_applic_date	17	MLADAT	6	MLADAT								
320	id		id		id								
321	mulch_level	11	ML	2	ML								
322	mulch_applic_year	17	MLAYR	5	MLAYR								
323	mulch_applic_day	16	MLADY	5	MLADY								
324	mulch_type	10	MLTP	4	MLTP	7	Mulch type by codes for soild polyth	code	text	6	EXPERIMENT	MANAGEMENT	MUI
325	mulch_thickness	15	MLTHK	5	MLTHK	8	Mulch thickness	mm	single		EXPERIMENT	MANAGEMENT	MUI
326	mulch_color	11	MLCOL	5	MLCOL	9	Mulch color	code	text	6	EXPERIMENT	MANAGEMENT	MUI
327	mulch_removal_date	18	MLRDAT	6	MLRDAT	-98	Mulch removal date	date	single		EXPERIMENT	MANAGEMENT	MUI
328	mulch_level	11	ML	2	ML	1	Mulch treatment level	number	single		EXPERIMENT	MANAGEMENT	MUI
329	mulch_removal_year	18	MLRYR	5	MLRYR	5	Mulch removal year	year	single		EXPERIMENT	MANAGEMENT	MUI
330	mulch_removal_day	17	MLRDAY	6	MLRDAY	6	Mulch removal date as day of year	doy	single		EXPERIMENT	MANAGEMENT	MUI
331	mulch_type_remove	17	MLTPR	5	MLTPR	7	Mulch type removed by codes for soi	code	text	6	EXPERIMENT	MANAGEMENT	MUI
332	id		id		id	0	1 id		single		EXPERIMENT	MANAGEMENT	CHE
333	chemical_applic_level	20	CH	2	CH	1	Chemicals application level		number		single	EXPERIMENT	MANAGEMENT CHE

Management_info | Measured_data | Explanatory_notes | Metadata_codes | Crop_codes | Management_codes



nature climate change

LETTERS

PUBLISHED ONLINE: 9 JUNE 2013 | DOI: 10.1038/NCLIMATE1916

Uncertainty in simulating wheat yields under climate change

S. Asseng *et al.*†

Projections of climate change impacts on crop yields are inherently uncertain[1]. Uncertainty is often quantified when projecting future greenhouse gas emissions and their influence on climate[2]. However, multi-model uncertainty analysis of crop

addressed by the climate science community through probabilistic projections based on multiple general circulation models (GCMs) or regional climate model ensembles[14]. However, most climate change agricultural impact assessments have used a single crop model[3], limiting the quantification of uncertainty[15]. As crop models

AgMIP The Agricultural Model Intercomparison and Improvement Project

Harmonizing Agricultural Data for Open Access and Improving Cropping Systems Models: Joint AgMIP-USDA Workshop

USDA

Data Dictionary – Just the basics

	A	B	C	D	E	F	G
1	Variable_Name	Description	Unit_or	Dataset	Subset	Group	Sub-Group
142	soil_identifier	Soil identifier	code	EXPERIMENT	MANAGEMENT	FIELDS	
143	soil_file	File containing relevant soil profile description	text	EXPERIMENT	MANAGEMENT	FIELDS	
144	water_table_depth	Water table depth	cm	EXPERIMENT	MANAGEMENT	FIELDS	
145	field_notes	Field, short notes	text	EXPERIMENT	MANAGEMENT	FIELDS	
146	field_soil_texture	Field observed soil texture	code	EXPERIMENT	MANAGEMENT	FIELDS	
147	soil_analysis_date	Soil analysis date	date	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
148	id	id		EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
149	soil_analysis_level	Soil analysis, treatment level identifier	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
150	soil_analysis_name	Soil analysis name	text	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
151	soil_family_NRCS_system	Soil family, NRCS system	text	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
152	soil_analysis_year	Soil analysis year	year	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
153	soil_analysis_day	Soil analysis day of year or days after planting	day	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
154	soil_analys_meth_pH_buff	Soil anlysis method for pH in buffer solution	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
155	soil_anal_meth_P_extract	Soil analysis method for P extraction	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
156	soil_anal_meth_exch_K	Soil analysis method for exchangeable	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
157	drainage_rate_per_day	Drainage rate as fraction per day	1/d	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
158	runoff_curve_no_SCS	Runoff curve no. (Soil Conservation Service)	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
159	id	id		EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
160	soil_analysis_level	Soil analysis, treatment level identifier linkage	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
161	soil_layer_top_depth	Soil layer depth to upper boundary of sample (top)	cm	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
162	soil_layer_base_depth	Soil layer depth to lower boundary of sample (bottom)	cm	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
163	soil_bulk_dens_moist	Soil bulk density, moist, determined on field sample	g/cm3	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
164	organic_carbon_conc	Organic carbon	%	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
165	soil_N_conc_total	Nitrogen, total soil, total	%	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
166	pH_in_water	pH of soil in water, from in-field sample	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
167	pH_in_buffer	pH of soil measured in buffer solution, from in-field sar	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
168	phosphorus_extractable	Phosphorus, extractable	mg/kg	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
169	potassium_exchangeable	Potassium, exchangeable	cmol/kg	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
170	soil_compac_fact_by_lyr	Soil compaction factor by layer (0 to 1 scale)	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS

Climate_codes | Other_codes | Pest_codes | Form Responses 1 | Form Responses 2 | Sheet1

What we really did …

- No researchers from forestry or economics
  - Scattered expertise relating to economics
- Focused on :
  - Reviewing metadata for discovery in Ag Data Commons
  - Understanding how the I/A Data Dictionary relates to schemes for organizing knowledge:
    - Vocabularies
    - Dictionaries
    - Ontologies
  - Examining how the current I/A Data Dictionary relates to domains represented in the workshop
    - Dairy CAP
    - Livestock CAP
    - Life Cycle Analysis database

Metadata: improvements to facilitate discovery

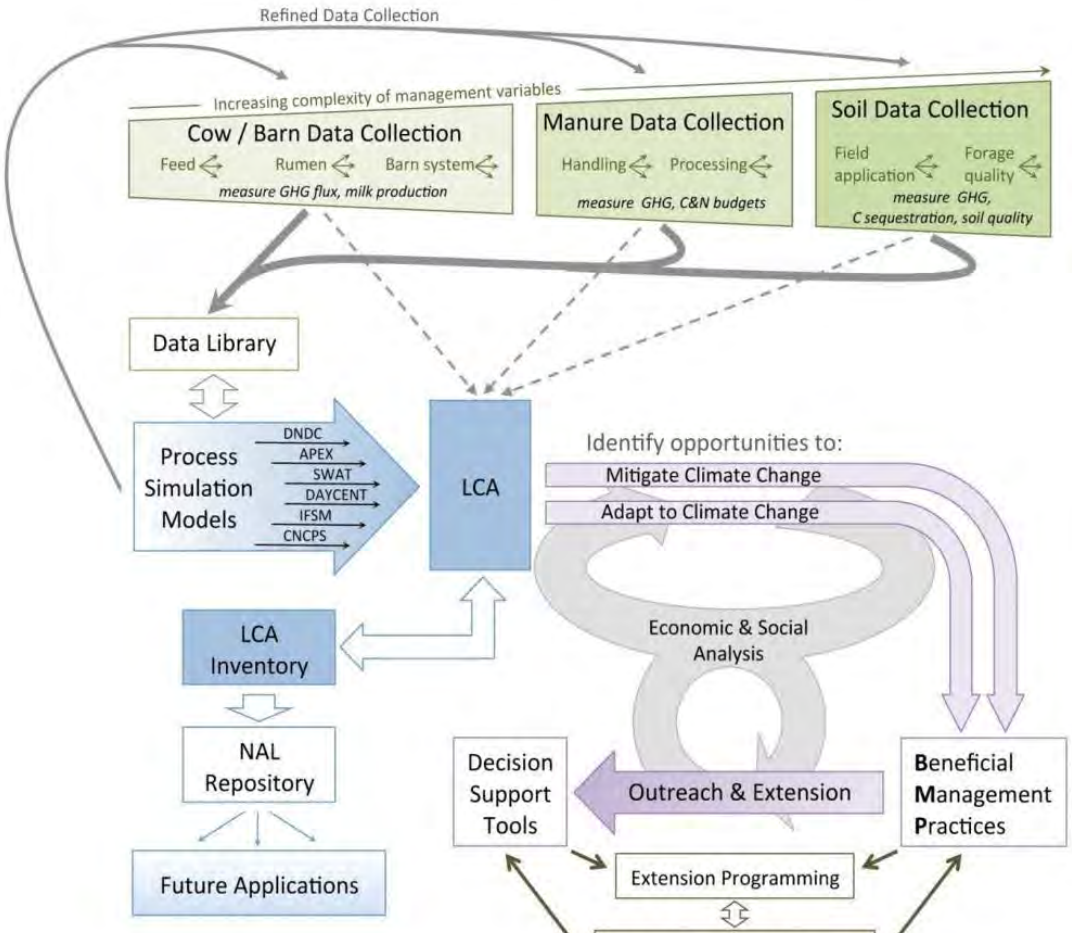
- Cropping/production systems:
  - I/A Data Dictionary has an unintentional mono-crop bias
  - Need to review existing descriptors
    - GRACEnet
    - NRCS
- Tillage practices
- “Organic” management
- Crop type categories: grain, bioenergy, fiber, etc.
- Less cryptic description of treatment factors
- Appropriate citations
  - Digital object identifiers (DOI) for published data
  - Publications
- Funding sources

I/A Data Dictionary

- Soil water retention
  - Van Genuchten and Brooks & Corey parameters
  - Used in RZWQM, MaizeSim
- Tillage descriptors
  - Link to RUSLE descriptors
- Curation

Dairy CAP

SUSTAINABLE DAIRY
SCIENCE FOR SUSTAINABLE PRODUCTION



Data acquisition activities:

- Cow/barn
- Manure
- Soil data in farm fields

Dairy CAP

SUSTAINABLE DAIRY
SCIENCE FOR SUSTAINABLE PRODUCTION

Two hands-on sessions working from Carol Barford's initial interpretation of I/A Data Dictionary

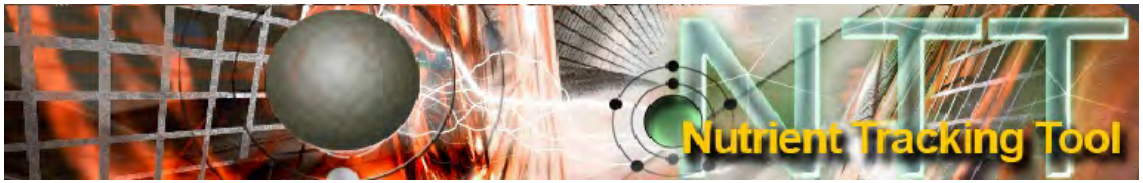
	A	B	C	D	E	F
1	**Measurement / characteristic**	**Proposed Units**	**Topic**		**Notes**	**Jeff White comment**
27	**Cow Status**					
28	**Cow ID / name**		cow		unique identifier , stable	
29	Breed		cow		dropdown list	
30	Stage of production		cow	time series	e.g. lactating, growing, dry	
31	Body weight	kg	cow	time series		
32	Body condition score		cow	time series	integer 1-5	
33	Lactation number		cow		integer	Also time series?
34	Days in milk (DIM)	days	cow	times series		
35	Milking frequency	times per day	cow	time series		
36	Rumen pH		cow	time series		
37	Rumen ammonia-N	mmol/L	cow	time series	time series data	
38	Rumen total volatile fatty acids (VFA)	mmol/L	cow	time series		
39	Rumen acetate	mol % of VFA	cow	time series		
40	Rumen propionate	mol % of VFA	cow	time series		
41	Rumen butyrate	mol % of VFA	cow	time series		
42	Rumen valerate	mol % of VFA	cow	time series		
43	Rumen branched-chain VFAs	mol % of VFA	cow	time series	2-methylbutyrate, isobutyrate, isovalerate	
44	Housing type		housing	?? Separate group		
45	Manure removal type		housing			
46	Manure removal frequency	times per week	housing		?	
47	Animal time in barn	hours/day	housing		vs. outdoors	
48	Air temperature	degrees Celsius	housing	time series	need max, min per day, indoors (two variables)	
49	Ventilation		housing		barn air volume exchanges per time	
50	Stall type		housing		tie stall, free stall	

Dairy CAP: next steps

SUSTAINABLE DAIRY
SCIENCE FOR SUSTAINABLE PRODUCTION

- Cows and barns
  - Carol Barford to discuss draft variables with researchers
  - Iterative revisions with Jeff White
- Manure descriptions
  - Easy to expand characteristics in I/A Data Dictionary
- Farm soils and greenhouse gas emissions
  - Based on GRACEnet template
  - Need to map I/A Data Dictionary to GRACEnet terminology – Jeff White
- Prepare a prototype data dictionary
  - Not discussed: format of dictionary, managing revisions, future harmonization (e.g., with GRACEnet and AgMIP)

Livestock CAP:



- Reviewed NTT
  - Weather, soil and management databases
  - Very impressive framework for model applications
  - Relation to experiment description unclear
- Data types not in I/A Data Dictionary
  - Erosion
    - Control structures
    - Soil loss and deposition
  - Livestock
    - Types
    - Stocking rates
    - Production
    - Grazing behavior
    - Grazing impacts
    - Manure and urine deposition
- Need input from the discipline scientists in the CAP
- Data types seem tractable

LCA

- Initial list of LCA terms or concepts in relation to I/A Data Dictionary
  - Missing in I/A Data Dictionary
  - Poorly or incompletely defined in I/A Data Dictionary from LCA perspective
- Tillage
  - Tractor characteristics: power, age, fuel amount and type …
- Irrigation
  - Water sources: well (depth), canal, pond …
- Farm or field vs. “point” data
  - I/A Data Dictionary is per unit land area (usually hectare)
  - LCA requires a known field or farm area

I/A DD – Curation & promotion

	A	B	C	D	E	F	G
1	Variable_Name	Description	Unit_or	Dataset	Subset	Group	Sub-Group
142	soil_identifier	Soil identifier	code	EXPERIMENT	MANAGEMENT	FIELDS	
143	soil_file	File containing relevant soil profile description	text	EXPERIMENT	MANAGEMENT	FIELDS	
144	water_table_depth	Water table depth	cm	EXPERIMENT	MANAGEMENT	FIELDS	
145	field_notes	Field, short notes	text	EXPERIMENT	MANAGEMENT	FIELDS	
146	field_soil_texture	Field observed soil texture	code	EXPERIMENT	MANAGEMENT	FIELDS	
147	soil_analysis_date	Soil analysis date	date	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
148	id	id		EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
149	soil_analysis_level	Soil analysis, treatment level identifier	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
150	soil_analysis_name	Soil analysis name	text	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
151	soil_family_NRCS_system	Soil family, NRCS system	text	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
152	soil_analysis_year	Soil analysis year	year	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
153	soil_analysis_day	Soil analysis day of year or days after planting	day	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
154	soil_analys_meth_pH_buff	Soil anlysis method for pH in buffer solution	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
155	soil_anal_meth_P_extract	Soil analysis method for P extraction	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
156	soil_anal_meth_exch_K	Soil analysis method for exchangeable	code	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
157	drainage_rate_per_day	Drainage rate as fraction per day	1/d	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
158	runoff_curve_no_SCS	Runoff curve no. (Soil Conservation Service)	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	
159	id	id		EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
160	soil_analysis_level	Soil analysis, treatment level identifier linkage	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
161	soil_layer_top_depth	Soil layer depth to upper boundary of sample (top)	cm	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
162	soil_layer_base_depth	Soil layer depth to lower boundary of sample (bottom)	cm	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
163	soil_bulk_dens_moist	Soil bulk density, moist, determined on field sample	g/cm3	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
164	organic_carbon_conc	Organic carbon	%	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
165	soil_N_conc_total	Nitrogen, total soil, total	%	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
166	pH_in_water	pH of soil in water, from in-field sample	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
167	pH_in_buffer	pH of soil measured in buffer solution, from in-field sar	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
168	phosphorus_extractable	Phosphorus, extractable	mg/kg	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
169	potassium_exchangeable	Potassium, exchangeable	cmol/kg	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS
170	soil_compac_fact_by_lyr	Soil compaction factor by layer (0 to 1 scale)	number	EXPERIMENT	MANAGEMENT	SOIL_ANALYSES	SOIL_ANALYSES_LAYERS

Climate_codes | Other_codes | Pest_codes | Form Responses 1 | Form Responses 2 | Sheet1

I/A Data Dictionary – From an IT or knowledge management perspective, what is the "MVL"?

Candidates

- Ontology
- Data dictionary
- Thesaurus
- Glossary
- Controlled vocabulary

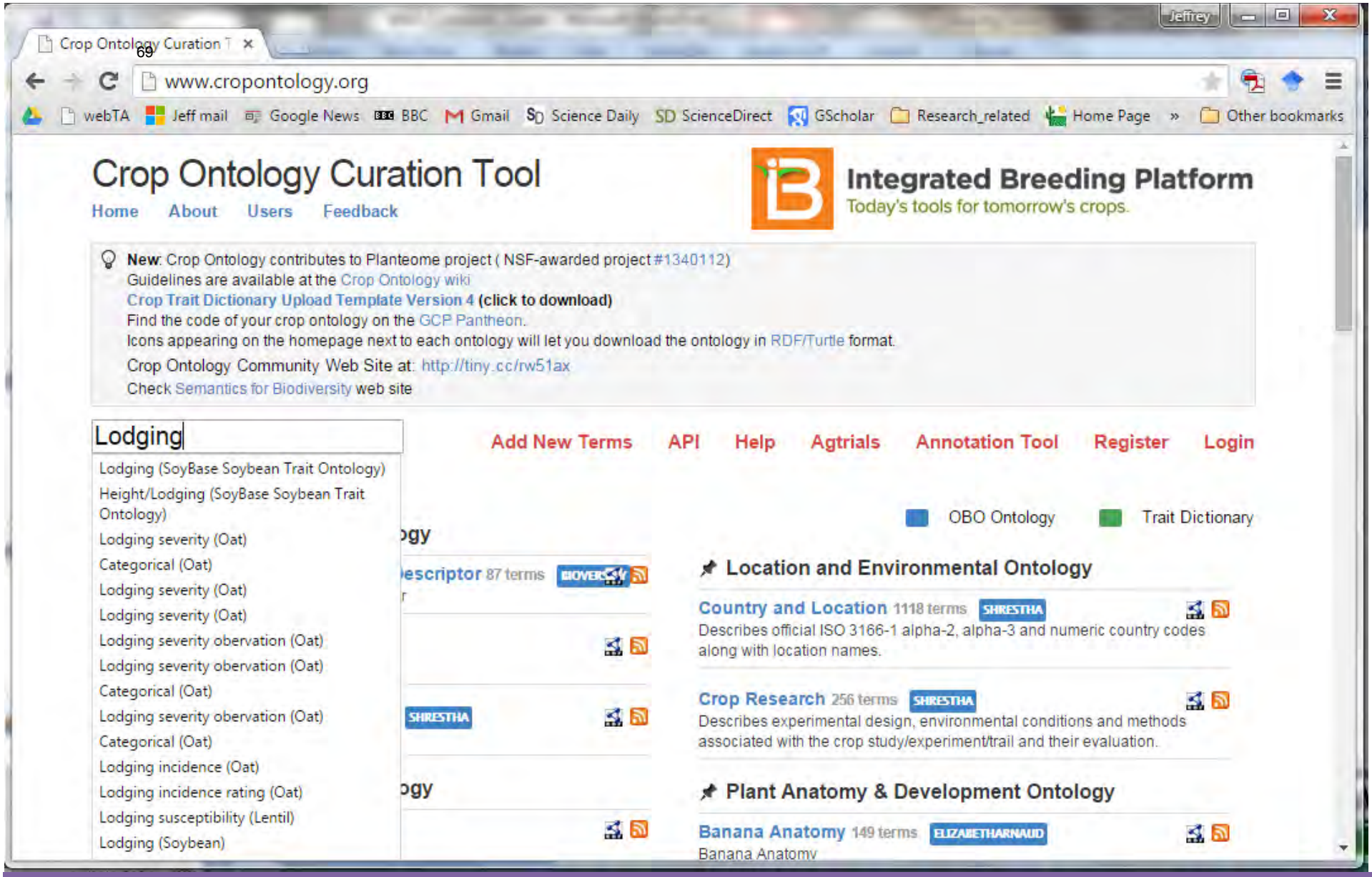
Content

- Name & synonyms
- Meaning
- Hierarchy of relations
- Units
- Data type
- Validation criteria

Wikipedia: A **data dictionary**, as defined in the *IBM Dictionary of Computing*, is a "centralized repository of information about data such as meaning, relationships to other data, origin, usage, and format."

I/A Data Dictionary Curation & promotion

- Need to provide the I/A Data Dictionary with an intelligent interface:
  - Searchable (not “Ctrl-F”)
  - Provision for user questions and suggestions
  - Hide the “geek-speak”
  - Provide links to:
    - Protocols/assays
    - More complete information on variables
- Solution through tools such as Protogé or CropOntology.org interface


Crop Ontology Curation T
www.cropontology.org
webTA
Jeff mail
Google News
BBC
Gmail
Science Daily
ScienceDirect
GScholar
Research_related
Home Page
Other bookmarks
Crop Ontology Curation Tool
Home
About
Users
Feedback
Integrated Breeding Platform
Today's tools for tomorrow's crops.
New: Crop Ontology contributes to Planteome project (NSF-awarded project #1340112)
Guidelines are available at the Crop Ontology wiki
Crop Trait Dictionary Upload Template Version 4 (click to download)
Find the code of your crop ontology on the GCP Pantheon.
Icons appearing on the homepage next to each ontology will let you download the ontology in RDF/Turtle format.
Crop Ontology Community Web Site at: http://tiny.cc/rw51ax
Check Semantics for Biodiversity web site
Lodging
Lodging (SoyBase Soybean Trait Ontology)
Height/Lodging (SoyBase Soybean Trait Ontology)
Lodging severity (Oat)
Categorical (Oat)
Lodging severity (Oat)
Lodging severity (Oat)
Lodging severity obervation (Oat)
Lodging severity obervation (Oat)
Categorical (Oat)
Lodging severity obervation (Oat)
Categorical (Oat)
Lodging incidence (Oat)
Lodging incidence rating (Oat)
Lodging susceptibility (Lentil)
Lodging (Soybean)
Add New Terms
API
Help
Agtrials
Annotation Tool
Register
Login
OBO Ontology
Trait Dictionary
Location and Environmental Ontology
Country and Location 1118 terms SHRESTHA
Describes official ISO 3166-1 alpha-2, alpha-3 and numeric country codes along with location names.
Crop Research 256 terms SHRESTHA
Describes experimental design, environmental conditions and methods associated with the crop study/experiment/trail and their evaluation.
Plant Anatomy & Development Ontology
Banana Anatomy 149 terms ELIZABETHARNAUD
Banana Anatomy

Conclusions

- *Special thanks*: Cynthia Parr, Jeff Campbell, Ezra Kahn, Simon Liu, Carol Barford, Ali Saleh and end-users
- Metadata: new descriptors, refine existing descriptors
- ICASA/AgMIP approach is extensible to other agricultural systems
  - Case 1: Dairy CAP
  - Case 2: Livestock CAP
  - Case 3: Life Cycle Analysis
  - [Case 4: Phenomics – with K-State]
  - [Case 5: Plant pathology – with U of Arizona]
- At workshops, domain experts should be paired with data management specialists
- Current I/A Data Dictionary is usable but can be improved
  - Improve user-interface
  - Data entry tools
  - Sub-daily measurements: “time stamp(s)” + suffix & prefix
- Role for NAL Knowledge Services Division to work with USDA locations and projects, AgMIP community, professional societies, and other stakeholders

Harmonizing Agricultural Data for Open Access and Improving Cropping Systems Models
Joint AgMIP-USDA Workshop
National Agricultural Library
Beltsville, MD
Agenda
May 11-15, 2015

Goal:

Through harmonization of data formats and tools, the workshop seeks to greatly enhance the efficiency with which data from USDA research programs are applied to research on agricultural systems analysis and modeling.

Objectives:

The objectives of the workshop are to:

1. Develop a prototype system to harmonize databases from several existing USDA and NIFA projects that will make data accessible and usable for multiple crop models and for statistical analyses (Breakout 1),
2. Expand the AgMIP IT tools used to operate multiple crop models to include nitrogen and phosphorus inputs and outputs and to complete translators for at least two additional US-based cropping system models (e.g., among CROPSYST, EPIC, SALUS, and RZWQM) (Breakout 2),
3. Develop and document metadata and minimum variables that should be included in harmonizing data in other USDA research areas (e.g., dairy, beef, Life Cycle Assessments, and biofuels) (Breakout 3), and
4. Develop recommendations for USDA that will lead to broader harmonization of data with capabilities for on-line publication of datasets. This includes development of plans for publication(s) from the workshop (Breakout 4).

Summary of Breakout Sessions:

Breakout Session 1. Corn and Wheat Database Harmonization. (Co-led by J. White and B. Basso). Review implementation of metadata in corn and wheat AFRI and ARS projects as well as status of the databases, data dictionaries, and APIs for discovering and accessing the data via the internet. Develop specific goals for work to be completed during workshop along with an initial plan and timetable for what will be done after the workshop. We will likely need to break this one into separate sub-groups.

Breakout Session 2. Harmonizing Inputs for Additional Cropping System Models, Incorporating N and P Variables. (Co-led by Meng Zhang and C. Porter). Demonstrate existing AgMIP tools for using the same data (APSIM and DSSAT). Review status of translators for CROPSYST, EPIC, SALUS, and RZWQM crop models. Develop work plan for completing these translators and a timetable that includes completion of at least one or two of these models during the workshop.

Breakout Session 3. Developing Metadata and Data Dictionaries for Other Agricultural System Components. (Co-led by C. Porter and J. White). Develop metadata for the following types of systems: dairy, beef, forestry, economics, and Life Cycle Analysis. Document these metadata and identify USDA databases (in NIFA, ARS, FRS, and ERS) that should be connected for use in model-based analyses and statistical analyses. Develop list of minimum set of variables that need to be harmonized.

Breakout Session 4. Strategic Planning with a Roadmap for a National Agricultural Data Network. (Co-Led by J. Jones, S. Liu) During these breakouts over a 2-day period, selected participants will summarize data initiatives and how they might fit into a national or global distributed agricultural data network. This will include short (~ 5 minute) summaries from **Ruth Bastow** of the Global Plant Council, **Pankaj Jaiswal** from Oregon State U, **Rob Knapen** from Alterra, **Moffatt Ngugi** from USAID, **Stan Wood** from BMGF, and others. This is intended to help identify mechanisms for collaboration on the major task ahead regarding creation of a national or global approach for publishing, storing, and accessing agricultural site data (e.g., experiments, breeding trials, on-farm field data, etc.). The breakout will outline a report and recommendations for use by USDA, AgMIP, and other institutions that are contributing to improving agricultural data archiving, publication, and retrieval for analyses.

Monday, May 11, 2015		
8:00 am	Registration	
	Plenary Session (Moderated by J. Jones)	
8:30	Welcomes	C. Woteki, S. Liu, C. Rosenzweig
9:00	Goals	J. Hatfield, J. Jones
9:15	Introductions	All
10:00	Break	
10:30	AgMIP Overview	C. Rosenzweig
10:50	USDA Goals and Strategies	A. Bartuska
11:10	Data Harmonization: AgMIP Approach/Progress	C. Porter
11:30	Status of Selected USDA databases	B. Basso, S. Eigenbrode, J. Hatfield, J. White
12:15 pm	Discussion	All
12:30	Lunch	
1:30	Charge to Breakout Sessions	J. Jones, J. Hatfield
	Breakout Sessions	
1:45	Breakout Sessions 1, 2, 3, 4	
3:00	Break	
3:15	Continue Work in Breakout Sessions 1, 2, 3, 4	
	Plenary Session	
4:30	Plenary Session: Status Report, Each Breakout	Moderator: C. Rosenzweig
5:30 pm	Adjourn	

Tuesday, May 12, 2015		
	Plenary Session	
8:30 am	Feedback, Goals for the Day	J. Hatfield, J. Jones
8:45	Potential Roles of NAL in Harmonizing USDA Data	S. Liu
9:15	Tri-Societies Data Initiatives Overview	J. Volenec
9:45	Summary Status of Dairy, Beef, LCA, Other Data	C. Barford, A. Saleh, E. Kahn
10:15	Discussion	
10:30	Break	
11:00	Charge to Breakout Sessions	J. Jones, J. Hatfield
	Breakout Sessions	
11:15	Breakout Sessions 1, 2, 3, 4	
12:30 pm	Lunch	
1:30	Continue Work in Breakout Sessions 1, 2, 3, 4	
3:00	Break	
3:30	Continue Work in Breakout Sessions 1, 2, 3, 4	
	Plenary Session	
4:30	Plenary Session: Status Report, Each Breakout	Moderator: S. Liu
5:30 pm	Adjourn	

Wednesday, May 13, 2015		
	Plenary Session	
8:30 am	Feedback, Goals for the Day	J. Jones
9:00	Charge to Breakout Sessions 1, 2 and 3 only	J. Jones
	Breakout Sessions	
9:15	Breakout Sessions	
10:30	Break	
11:00	Breakout Sessions	
12:30 pm	Lunch	
1:30	Continue Breakout Sessions	
3:00	Break	
3:30	Continue Breakout Sessions 1, 2, and 3	
	Plenary Session	
4:30	Plenary Session: Status Report, Each Breakout	Moderator: J. Hatfield
5:30	Adjourn	

Thursday, May 14, 2015		
	Plenary Session	
8:30 am	Feedback, Goals for the Day	J. Hatfield, J. Jones
8:45	Report from Breakout Sessions 1, 2, 3	
9:45	Overview of Other Potential Collaborative Initiatives	J. Jones, J. Hatfield
10:15	Break	
	Breakout Sessions	
11:00	Continue with Breakout Sessions 1, 2, 3	
12:30 pm	Lunch	
1:30	Continue Breakout Sessions 1, 2, 3	
3:00	Break	
3:30	Continue Breakout Sessions 1, 2, 3	
	Plenary Session	
4:30	Plenary Session: Status Report, Each Breakout	Moderator: J. White
5:00 pm	Adjourn	

Friday, May 15, 2015		
	Plenary Session	
8:30 am	Goals for the Day	J. Hatfield, J. Jones
	Breakout Sessions	
8:45	Continue Breakout Sessions 1, 2, 3 Reconvene Breakout Session 4 (with remote connections)	
10:30	Break	
	Plenary Session	
11:00	Plenary Session: Status Report, Each Breakout	Moderator: J. Jones
12:15 pm	Concluding Remarks	J. Hatfield, J. Jones, S. Liu
12:30 pm	Adjourn, End of Workshop	

Participants Attending the AgMIP-USDA Data Harmonization Workshop May 11-15, 2015

Abendroth, Lori	labend@iastate.edu	Iowa State
Ahuja, Laj	Laj.Ahuja@ARS.USDA.GOV	USDA-ARS
Baer, Brian	baerb@msu.edu	Michigan State University
Bartling, Pat	pat.bartling@ars.usda.gov	USDA-ARS
Bartuska, Ann	Ann.Bartuska@osec.usda.gov	USDA, OSEC
Basso, Bruno	basso@msu.edu	Michigan State University
Bastow, Ruth	ruth@globalplantcouncil.org	Global Plant Council
Birkett, Clay	claybirkett@gmail.com	
Boote, Kenneth	kjboote@ufl.edu	University of Florida
Campbell, Jeffrey	Jeffrey.Campbell@ARS.USDA.GOV	USDA/NAL
Collier, Sarah	smcollier@wisc.edu	University of Wisconsin
Devare, Medha	m.devare@cgiar.org	CGIAR Consortium
Eigenbrode, Sanford	sanforde@uidaho.edu	University of Idaho
Fleischer, David	David.Fleisher@ARS.USDA.GOV	USDA-ARS
Gerik, Tom	tgerik@brc.tamus.edu	Texas A & M
Gessler, Paul	paulg@uidaho.edu	University of Idaho
Gustafson, Dave	dgustafson@ilsi.org	CIMSANS/ILSI
Hatfield, Jerry	jerry.hatfield@ars.usda.gov	USDA-ARS
Herzmann, Daryl	akrherz@iastate.edu	Iowa State
Hyman, Glenn	G.HYMAN@CGIAR.ORG	CGIAR
Izaurralde, Cesar	cizaurra@umd.edu	University of Maryland
Jacobs-Young, Chavonda	C.JacobsYoung@ARS.USDA.GOV	USDA, ARS Administrator
Jahn, Molly	mjahn@cals.wisc.edu	University of Wisconsin
Jaiswal, Pankaj	jaiswalp@science.oregonstate.edu	Oregon State University
James, Rosalind	Rosalind.James@ARS.USDA.GOV	USDA-ARS
Jannink, Jean-Luc	jeanluc.work@gmail.com	Cornell University
Jones, Jim	jimj@ufl.edu	University of Florida
Jones, Kristal	kjones@sesync.org	NSF-Synthesis Center
Kahn, Ezra	Ezra.Kahn@ars.usda.gov	USDA/NAL
Kelly, David	davidkelly@uchicago.edu	University of Chicago
Kim, Soonho	SOONHO.KIM@cgiar.org	IFPRI
Knapen, Rob	rob.knapen@wur.nl	Wageningen University
Knighton, Ray	Rknighton@nifa.usda.gov	UDSA-NIFA
Liu, Simon	Simon.Liu@ARS.USDA.GOV	USDA-ARS
Martin, Clyde	clyde.f.martin@ttu.edu	Texas Tech
Moore, Gary	Gary.Moore@ARS.USDA.GOV	USDA-ARS
Nelson, Roger	rnelson@wsu.edu	Washington State University
Okamuro, Jack	Jack.Okamuro@ARS.USDA.GOV	USDA-ARS
Parr, Cynthia	Cynthia.Parr@ARS.USDA.GOV	USDA-ARS, NAL
Porter, Cheryl	cporter@ufl.edu	University of Florida

Reddy, V.R.	vr.reddy@ars.usda.gov	USDA-ARS
Rosenzweig, Cynthia	crr2@columbia.edu	NASA, Columbia University
Saleh, Ali	asaleh@tiaer.tarleton.edu	Tarleton State University
Seamon, Erich	'erichs@uidaho.edu'	University of Idaho
Shafer, Steve	Steven.Shafer@ARS.USDA.GOV	USDA-ARS
Simmons, Kay	Kay.Simmons@ARS.USDA.GOV	USDA-ARS
Stockle, Claudio	stockle@wsu.edu	Washington State University
Timlin, Dennis	Dennis.Timlin@ARS.USDA.GOV	USDA-ARS
Villalobos, Chris	cvillalobos@ufl.edu	University of Florida
Volenec, Jeff	jvolenec@purdue.edu	Purdue University
Walthall, Charlie	Charlie.Walthall@ARS.USDA.GOV	USDA-ARS
Wayne, Rick	fewayne@wisc.edu	University of Wisconsin
White, Jeff	Jeffrey.White@ARS.USDA.GOV	USDA-ARS
Wood, Stan	Stanley.Wood@gatesfoundation.org	Gates Foundation
Woteki, Cathy	Catherine.Woteki@osec.usda.gov	USDA, OSEC
Zhang, Meng	meng.zhang15@ufl.edu	University of Florida