Building an Agriculture Data Infrastructure for Arizona

Arthur "Barney" Maccabe
Executive Director
Institute for Computation and Data-Enabled Insight
The University of Arizona

Presentation for Multicore World 2025
Christchurch, NZ
February 19, 2025

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight

We respectfully acknowledge the University of Arizona is on the land and territories of Indigenous peoples. Today, Arizona is home to 22 federally recognized tribes, with Tucson being home to the O’odham and the Yaqui. Committed to diversity and inclusion, the University strives to build sustainable relationships with sovereign Native Nations and Indigenous communities through education offerings, partnerships, and community service.

THE UNIVERSITY OF ARIZONA
Institute for Computation & Data-Enabled Insight

A Little Bit About Barney

- 11 years: Kansas, Virginia, Ohio
- 11 years: Tucson, Arizona

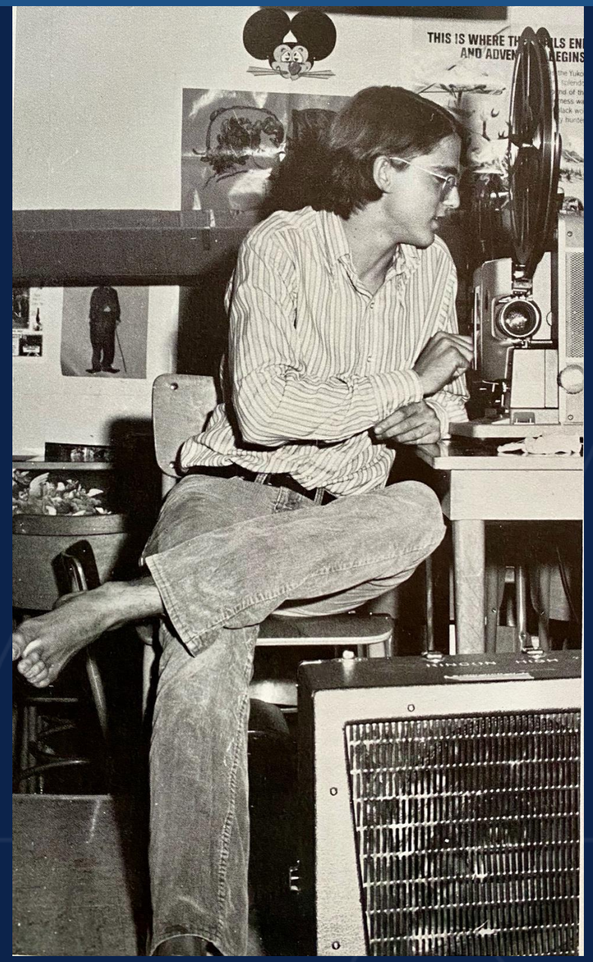
  Undergraduate degree: Math
- 5 years: Atlanta, Georgia

  Graduate degrees: Info and Computer Science
- 27 years: Albuquerque, New Mexico

  System software for massively parallel systems
- 13 years: Oak Ridge, Tennessee

  Computer Science and Mathematics
- 3 years: Tucson, Arizona

  Institute for Computation and Data-Enabled Insight



ICDI:
Innovation and opportunity at the intersection of culture and technology

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight

Theme 1

It’s the Data, stupid

Data assets have become *the* commodity

It’s unlikely that the academic community will have significant assets in energy and computing (the other critical assets)

We probably don’t even have much data

Focus on the **public good**

THE UNIVERSITY OF ARIZONA
Institute for Computation & Data-Enabled Insight

Theme 2

Collecting data is expensive; Reusing data is dangerous

We have made significant advances in AI reusing data (e.g., Google translate).

However, data has context and the reuse context is never an exact fit for the original context.

Need to be intentional about data collection.

Goal: Describe the agriculture related data collection ecosystem we're building in Arizona

THE UNIVERSITY OF ARIZONA
Institute for Computation & Data-Enabled Insight

Arizona’s Economy: 5 C’s

- **Copper**
- **Cattle**
- **Cotton** (Pima)
- **Citrus**
- **Climate**

Arizona’s Economy: 5 C’s

Caterpillar, a manufacturer of construction and mining equipment, has built a prototype for an all electric truck. The 240-ton large mining truck was demonstrated in November 2022. The truck is on display at Caterpillar’s Tucson Proving Ground in Green Valley.

Mamta Popat, Arizona Daily Star

- **Copper**
- **Cattle** → Cows (dairy)
- **Cotton** (Pima) → Computing?
- **Citrus** → Crops (leafy greens)
- **Climate**:

“We have the climate of tomorrow, today.”

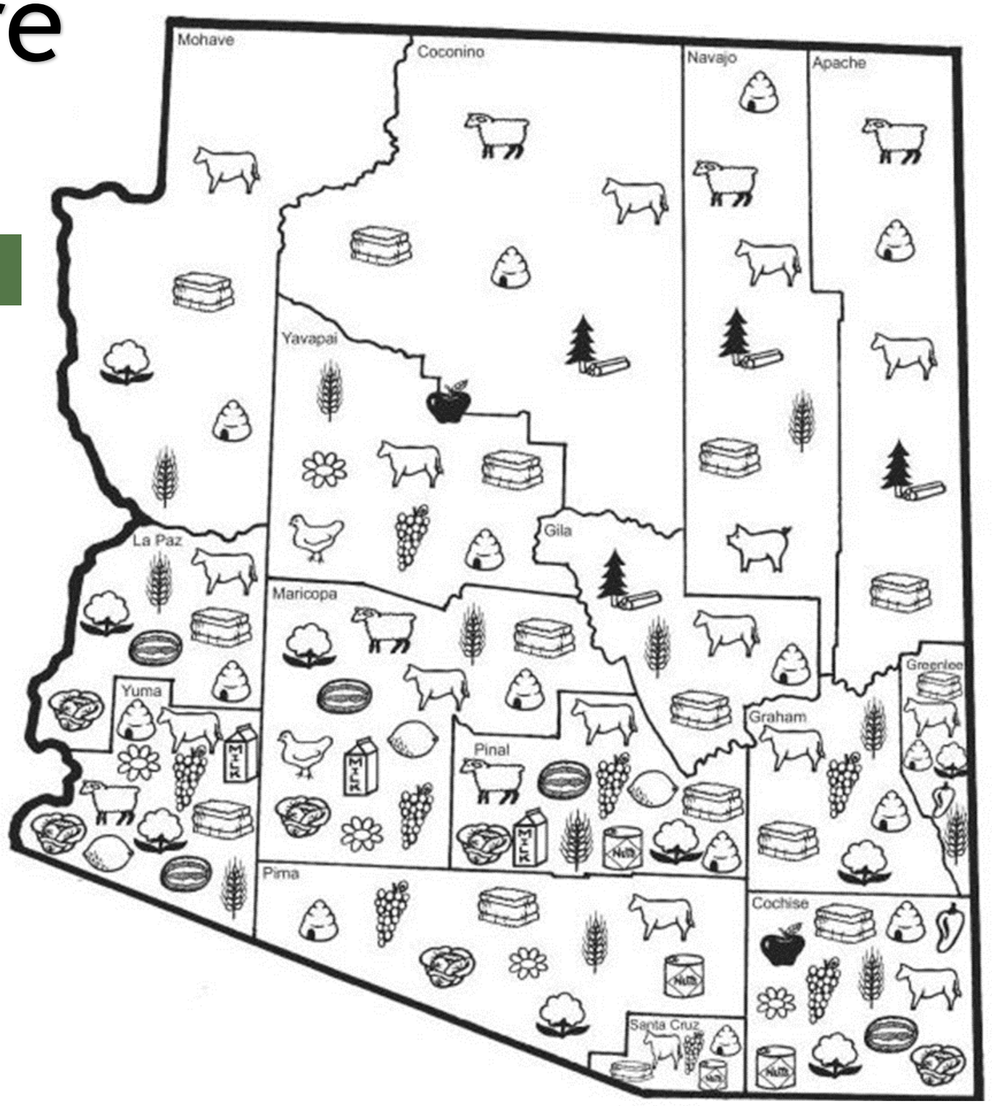
–Duke Pauli

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight

Agriculture in AZ

- **$23.3B Economic Impact**
- **Supports 138,000+ jobs**
- **20,000 farms/ranches (26M acres)**
- **Exports to over 70 countries**

3 of Arizona's "Five C's"



Source: nass.usda.gov/az/

"Arizona is known as one of the most productive and efficient agricultural regions in the world and one of the most diverse agricultural production states in the nation."

AZ Farm Bureau

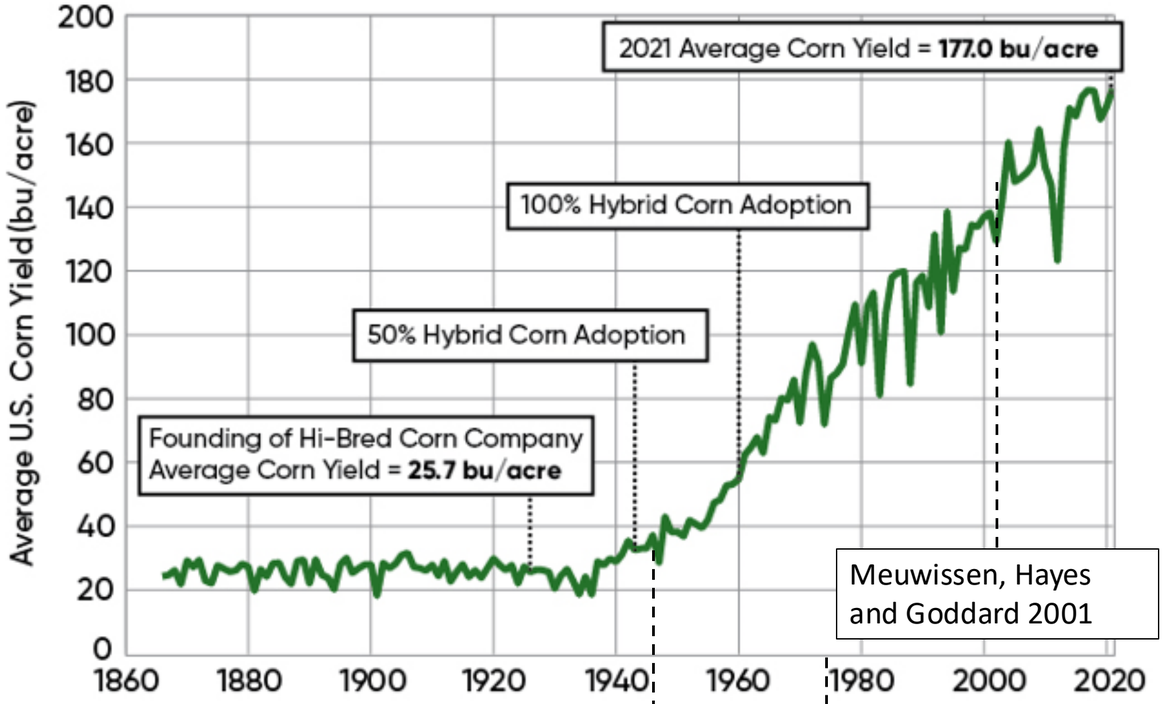
Slide courtesy of Connor Osgood

Building Models in Agriculture

Genetic Models
Crop Thinning and Weeding
Irrigation and Fertilizer Management

Improvements over the last century due to genetic and computational advancements

Slide courtesy of Mary-Francis LaPorte



United States average corn yield, 1866-2021 (USDA NASS.) via a Pioneer report

Meuwissen, Hayes and Goddard 2001

Applications of BLUPs

Malécot 1948: Mathematical Theory of heredity

- Genetic advancements
  - Hybrid corn and Rice
- Computational Methods aid in breeding
  - identifying lines that contain traits of interest, identifying markers for traits of interest
  - predicting which lines will contain traits of interest using the genetics

GENETICS

Information for Authors Editorial Board Submit a Manuscript

Genetics. 2013 Jul;194(3):597–607. doi: 10.1534/genetics.113.152207

Genomic BLUP Decoded: A Look into the Black Box of Genomic Prediction

David Habier [*,†,1], Rohan L Fernando [*], Dorian J Garrick [*]

Ex. 1: Robotic crop thinning & weeding

- Lettuce seeds are planted with 2” spacing, but then crop is thinned to 12” spacing
  - Helps to guarantee ‘a plant in every spot’
- Stand count and plant uniformity for is critical for profitability in lettuce
- Thinning is done manually by contract labor crews (typically H2A)
- While crews work quickly, there are issues with consistency
- Robotic lettuce thinning & cultivation
  - Image recognition to target plants for removal
  - Precision spot spray of herbicide or lasers to kill extra plants
  - One company claims 7 acres/hour

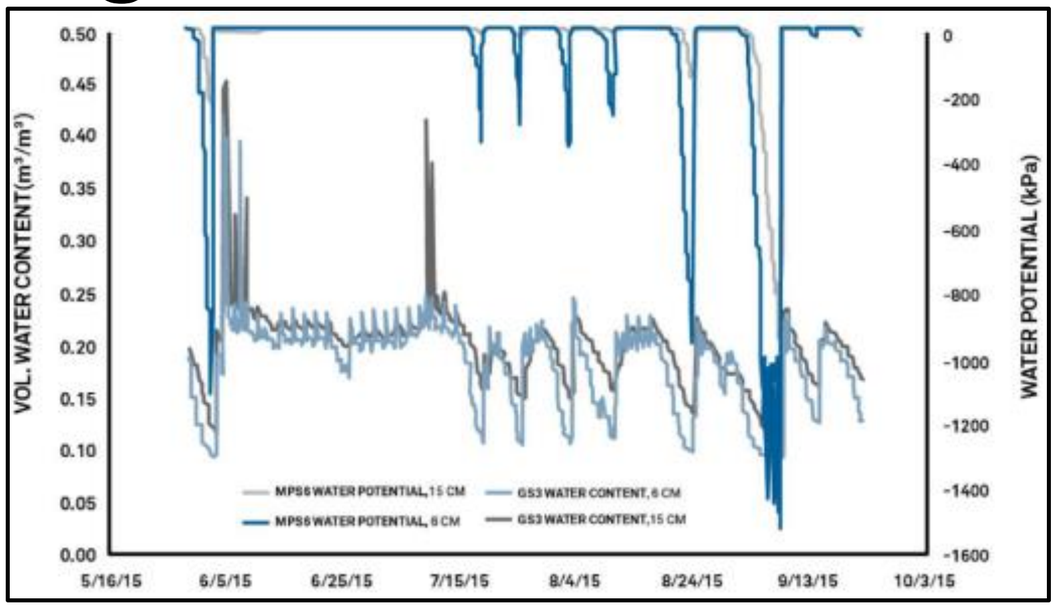
Vision Robotics | RDO website

University of Arizona | FarmProgress

Slide courtesy of Connor Osgood

Ex. 2: Irrigation and fertilizer management

- Delivering water to the crop based on what is needed, not on time schedule
- Advanced sensors
- Remote/proximal crop sensing
- Crop water stress index (CWSI)
- Nutrient-specific sensors (some day…)
- Variable-rate fertilizer based on local soil maps
- Validation studies needed to ensure accuracy of models



Water content vs. water potential (METER Group)


Irrometer

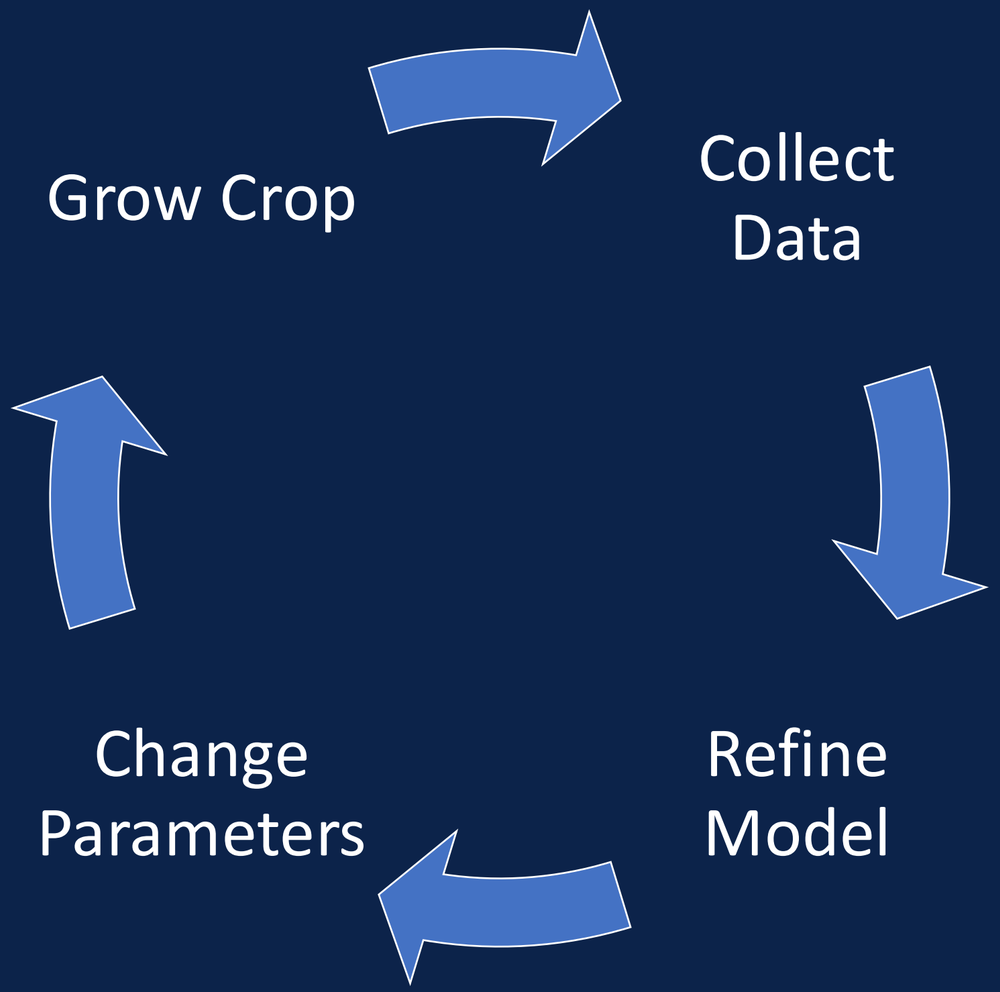
TEROS 21 | METER Group


RS130 Telemetry Node +

Slide courtesy of Connor Osgood

The Virtuous Cycle

Collect data, refine model



- Growing crops is time consuming and expensive
- Potential to reduce field productivity
- Need incremental data collection capabilities

THE UNIVERSITY OF ARIZONA
Institute for Computation & Data-Enabled Insight

Incremental Data Collection

Controlled Environments
Biosphere 2
The Gantry

Controlled Environment Agriculture Center

Controlled Environments

- University of Arizona Controlled Environment Agriculture Center (CEAC) supports research related to plant growth in alternate environments

- Upper right image shows the Prototype Lunar Greenhouse

- Lower right shows agrovoltaics where part of the incident solar radiation not required for crop growth is harvested for generating power

Biosphere 2: An Analog Twin

Space Analog for the Moon and Mars

- The University of Arizona's Biosphere 2 is the world's largest controlled environment dedicated to understanding the implication, mitigation and adaptation solutions for resilience of our planet (Biosphere 1) due to the global climate crisis.
- Biosphere 2 is a meso-scale Earth science facility encompassing 3.14 acres and houses five synthetic ecosystems encased in a glass and metal spaceframe. Operational for 30 years and counting, these meso-scale ecosystems include the world's largest controlled systems of tropical rain forest, desert, savanna, mangrove, and ocean.

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight

Big Data Problem: Phenotyping 30K Individual Plants

- University of Arizona is home to the world's largest, outdoor phenotyping platform – the Field Scanner.
- Provides unique opportunity to study how plants grow and respond to extreme environmental conditions, like those predicted in the future climate scenarios.
- Collecting and providing high resolution data that has never been available to the research community.

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight

Yuma

Leafy Greens
Cyber Experiment Station
Yuma Center for Excellence in Desert Agriculture

Agriculture in Yuma

- High value winter specialty crop industry ($2B)
- Supplies ~80% of the winter leafy green vegetables for USA and Canada (Nov – Mar)
- Top 0.1% of U.S. Counties in vegetable sales, Top 0.5% for all agricultural products
- One of the top worldwide producing areas of medjool dates (7,000+ acres)
- Year-round growing season, 150+ crops
  - Rotation crops: Wheat, Melons, Cotton, Hay
  - Also: Citrus and Seed Crops






- 230,000 acres of fertile soil
- Available workforce
- Senior water rights
- Great community support
- Progressive farmers exploring/utilizing technology

Slide courtesy of Connor Osgood

THE UNIVERSITY OF ARIZONA
Yuma Center of Excellence for Desert Agriculture

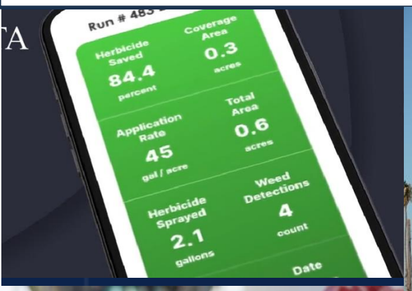
THE UNIVERSITY OF ARIZONA
Arizona
Experiment Station

Cyber Experiment Station

- Team of data scientists, programmers/developers, cloud-solution experts, videographers, and telepresence engineers.
- Provides the digital-communications vehicle to deliver our “story”. ½ Communications + ½ Cyber Technologies
- Unique to the University of Arizona

THE UNIVERSITY OF ARIZONA
Yuma Center of Excellence for Desert Agriculture

YCEDA



- YCEDA is an innovative public-private partnership that connects top scientists to the desert agricultural industry.
- Together we put science to work developing solutions to the challenges of arid-land crop production.
- Our work focuses on high-priority issues identified by industry stakeholders, including but not limited to increasing production efficiencies through disease and water management, crop yield maximization, food safety, and technology utilization.

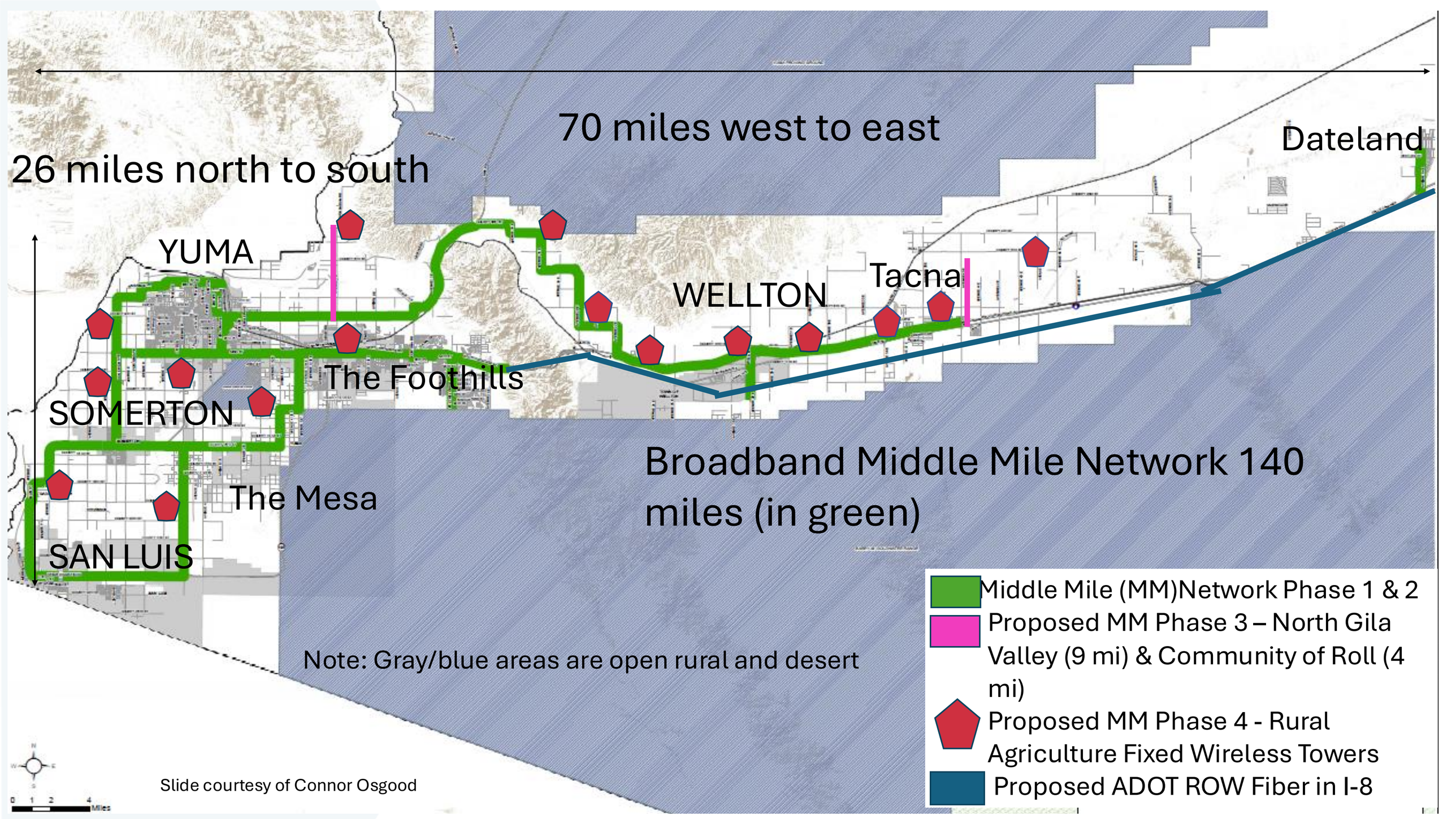
Connecting Fields

AZ Broadband Strategy
Yuma County Last-Mile for Agriculture

AZ Broadband Strategic Plan: Build the Middle Mile

- 80% of unserved and underserved households can be served by the middle-mile network on interstate and select state highways
- The middle-mile network creates the open access network that last mile initiatives can tap into to cover the remaining 21% of underserved

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight


70 miles west to east
26 miles north to south
Dateland
YUMA
WELLTON
Tacna
The Foothills
SOMERTON
The Mesa
SAN LUIS
Broadband Middle Mile Network 140 miles (in green)
Note: Gray/blue areas are open rural and desert
Middle Mile (MM)Network Phase 1 & 2
Proposed MM Phase 3 – North Gila Valley (9 mi) & Community of Roll (4 mi)
Proposed MM Phase 4 - Rural Agriculture Fixed Wireless Towers
Proposed ADOT ROW Fiber in I-8
Slide courtesy of Connor Osgood

Connecting fields in Yuma County

Image courtesy of Connor Osgood

- NSF CC* grant
  - Provides connectivity to Experiment Station
  - Local education facilities
- Longer term Yuma Board of Supervisors
- 27 CBRS towers
  - 10GB upload and download
  - ~180,000 acres
  - ~450 farms
- Three tiers of service: Wi-Fi, Broadband cellular (CBRS), and LoRaWAN for battery and solar devices

RESEARCH, INNOVATION & IMPACT
Institute for Computation & Data-Enabled Insight

Summary of Data Collection

	Size	Controlled	Measurement	Cost
Greenhouse	Tiny	Completely	Excellent	Very Low
Biosphere 2	Small	Mostly	Good	Low
Gantry	Medium	Minimally	Good	Moderate
Fields in Yuma	Large	Minimally	Fair	Expensive

Governance

Data Sharing/Discovery Infrastructure
Data Protection
Data Sovereignty

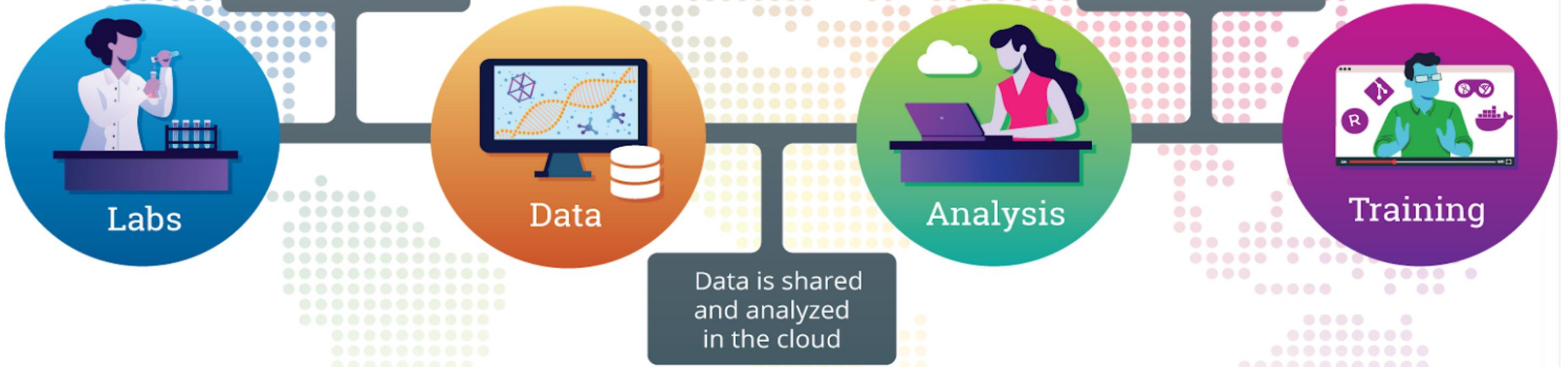
CYVERSE® Public

100,000 users

1500 publications

11 PBs of data

40,000 trained



CYVERSE® Professional

- Federation with local and commercial cloud and high-performance computing
- Integration with local user identity management systems
- Security compliance

CYVERSE® Health

- Secure Perimeter (VPN)
- HIPAA Compliant Platforms inside of Perimeter
- XNAT
- Training

CYVERSE® Defense

- ITAR Compliance
- Receive data from multiple sensors
- Policy based data visibility and sharing
- Support for multiple teams and data partitioning

Data Protection

Walled Garden with Function-based Access Control

Shared Models

- Multiple contributors.
- Share models, not data (directly).
- Federated or trusted third party.

Walled Garden

- Controlled admission.
- Admission requires shared responsibility/accountability
- Don’t trust, “the data protects itself”

Access Control

- Role based access control.
- Each role has limited access through a set of defined functions.

Questions?

maccabe@arizona.edu

THE UNIVERSITY OF ARIZONA
Institute for Computation
& Data-Enabled Insight