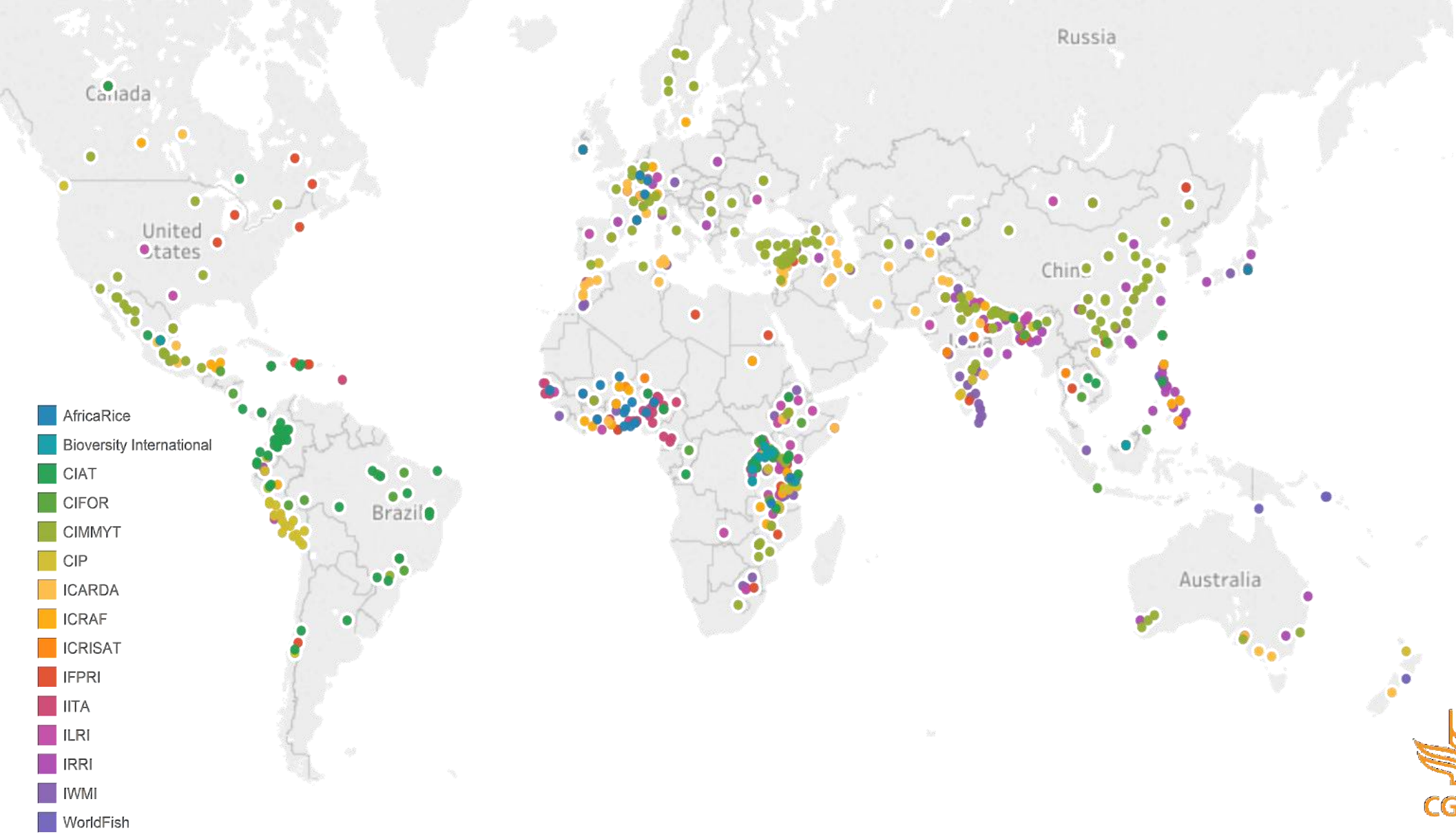


Dots show where CGIAR's research took place in 2000-2016

Koo, Jawoo, Glenn Hyman, Silvia-Elena Castaño, and Grant McKenzie. "CGIAR Scientometric Trends." International Food Policy Research Institute, 16 Mar. 2017. http://scientometrics.ifpri.org.

CGIAR Platform for Big Data in Agriculture

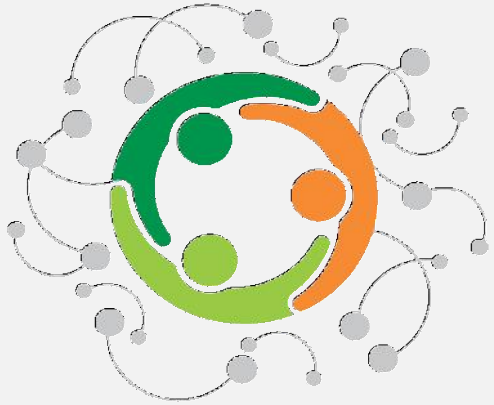
JAWOO KOO

Opportunities to Collaboration with BIG DATA | December 14, 2018

Modules = Collaboration Entry Points

Organize

Support data generation, open access, and management



Convene

Collaborate and convene to support incorporating data science at **Centers**

Inspire

Inspire how data science can enhance science and deliver impacts on **CRP-selected** special topics

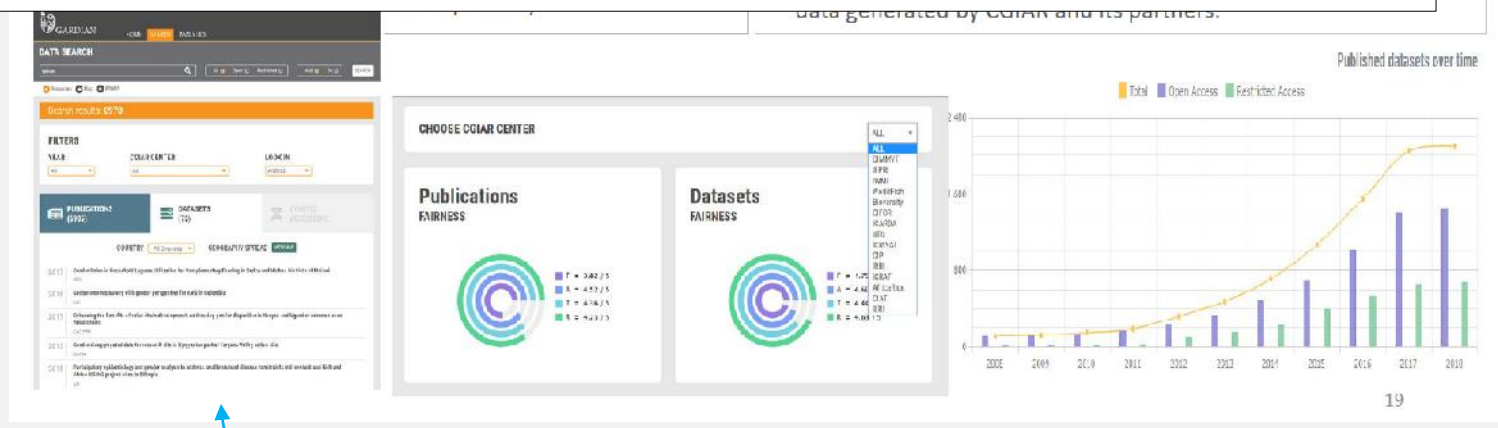
Organize

Open Access, Open Data, Data Management/Analysis

- **Centrally** establishing an online system and workflow to make CGIAR data more visible.
- **De-centrally** building the capacity of Centers to make data assets F.A.I.R.
- ***Plus***, developing common data analytics environment.

Yellow – **Potential for second or later business cycles**

- Projects in program have credible documentation of objectives and assumptions, and clear explanation of how they are aligned with program objectives (1)
- Capacity development is appropriately designed and delivered in support of the Program Theory of Change. (10)
- Program has made adequate progress towards open and FAIR (Findable, Accessible, Interoperable, Reusable) data. (15)
- Program produces high quality evidence of its claims for outcomes and impacts (16)
- Program progress reporting to CGIAR (annual reports, common reporting indicators) is substantially complete and adequately evidenced (16)
- Program effectively plans and manages budgets (17)



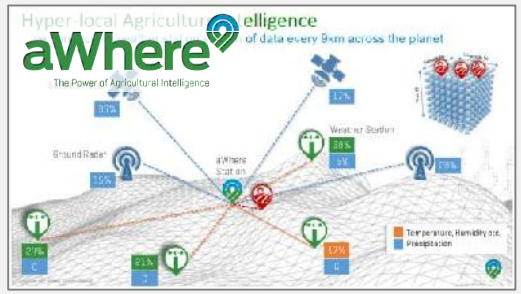
gardian.bigdata.cgiar.org

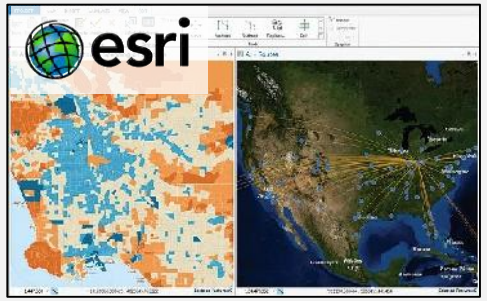
Organize

Open Access, Open Data, and Data Management/Analysis

- GARDIAN | Agroknow
- **Geospatial data cataloging/visualization** | Critigen
- **Data analysis pilots** | AgMIP/U. of Florida
- **Secured data storage/encryption** | U. of Minnesota
- **Strengthening CGIAR's data science** | UC Davis
- **Ontologies** | Bioversity International
- **Metadata standards** | IITA
- **Breeding and trial data standardization** | CIP
- **Privacy and ethics** | One Planet Solution
- *… and more!*

Agroknow
accenture
UCDAVIS
CRITIGEN
Earlham Institute
TechChange
the institute for technology and social change
UNB
EST. 1785
UNIVERSITY OF NEW BRUNSWICK
UF IFAS
UNIVERSITY of FLORIDA
Ontocale
WAGENINGEN
UNIVERSITY & RESEARCH
UNIVERSITY OF MINNESOTA
Driven to Discover℠
KDDart
Knowledge Discovery System
Altmetric
ONE PLANET SOLUTION
ORCID



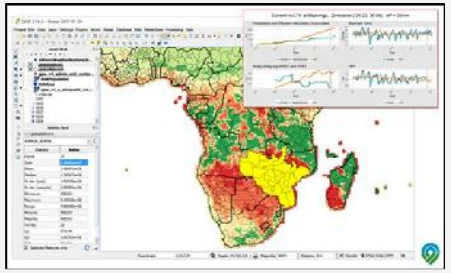


Agroknow
accenture
UCDAVIS
CRITIGEN
Datawheel
viamo
improving lives via mobile
We Robotics
Earlham Institute
aws
IBM Research
UNB
EST. 1785
UNIVERSITY OF NEW BRUNSWICK
Ontocale
UF IFAS
UNIVERSITY of FLORIDA
WAGENINGEN
UNIVERSITY & RESEARCH
ONE PLANET SOLUTION
dial Digital Impact Alliance
UNIVERSITY OF MINNESOTA
Driven to Discover℠
KDDart
Knowledge Discovery System
TechChange
the institute for technology and social change

Convene

Collaborate and convene around data and agricultural R4D

- Developing new **Technical Partnerships**
- Provision of **Shared Services (data and tools)**
- Provision of **Technical Training**
- Supporting six **Community of Practice**
- Mini-Grants for **Key Datasets**



aws

Community of Practices

Data-Driven Agronomy | CIAT

Crop Modeling | CIMMYT

Geospatial Analysis | IFPRI

Livestock Data | Univ. Edinburgh

Ontologies | Bioversity Int'l

Socio-Economic Data | CIMMYT

Inspire

Innovation process to implement data science research in CRPs

- **Competition for pilots** (100K) and **scaling-up** grants (250K)
- **Topics**
  - Revealing Food System Flows
  - Monitoring Pests & Diseases
  - Disrupting Impact Assessment
  - Empowering Data-Driven Farming
- **Criteria**
  - Data use
  - Scale
  - Impact
  - Sustainability
  - Innovation

2017

Real Time Diagnostics for Wheat Rust | **WHEAT** | CIMMYT, EIAR, John Innes Centre
IVR Marketing Service | **MAIZE** | CIMMYT & VOTO Mobile
Livestock Disease Detection using Social Media | **LIVESTOCK** | ILRI & Farm.ink
Pest and disease monitoring by using AI | **RTB** | CIAT, CIP, Bioversity International, Google, PSU
Picture-based Insurance and Extension | **PIM** | IFPRI & CABI

2018

CubicA: Agriculture Advisory App | Bioversity, Dalberg Data Insights & Viamo
Revealing Informal Food Flows through Free Wifi | CIAT & GSO (Vietnam)
An Integrated Data Pipeline for Small Fisheries | WorldFish & PDS
Smart Seed Selection | CIMMYT & BioSense Institute
Use CML to Estimate Rainfalls for Agriculture | IFPRI & Cornell University

Geospatial research in CGIAR

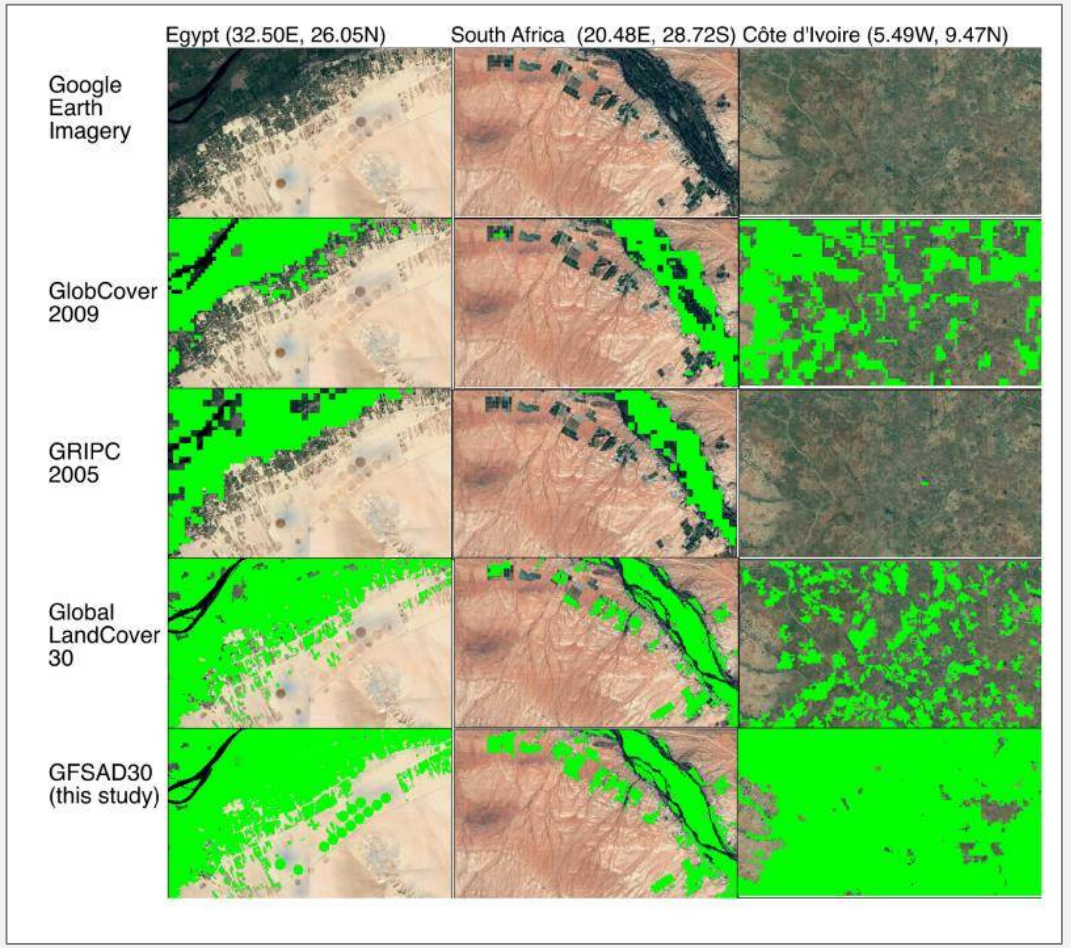
Mapping Baseline/Projections

- Landcover and land-use
- Crop/livestock distribution
- Historic climate
- Soils
- Population
- Market access
- Health
- Technology adoption
- Poverty
- Suitability
- Crop modeling
- Trade modeling
- Climate change impacts
- Socioeconomic projections
- *…and many more!*

Special Spatial Challenges… and ways to address

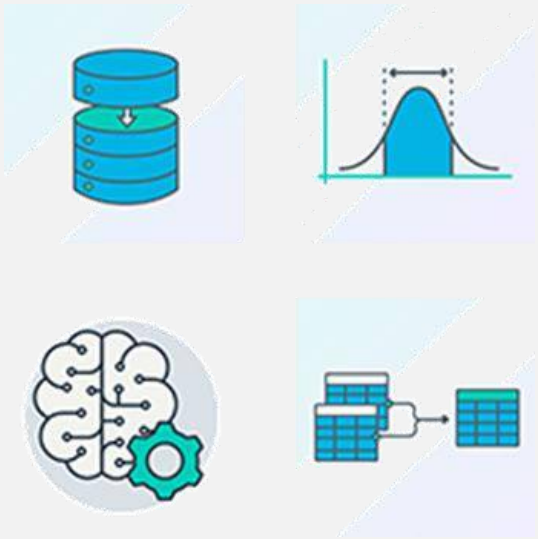
1. **Managing large files**
   → Cloud
2. **Developing/Managing portals**
   → Invest more on data
3. **Need insiders' knowledge**
   → CoP is here to help!
4. **Need programming skills**
   → Invest to (re)train staffs.
5. **Proprietary or FOSS4G**
   → Don't sweat; whichever works.

Seeing is (not always) believing!

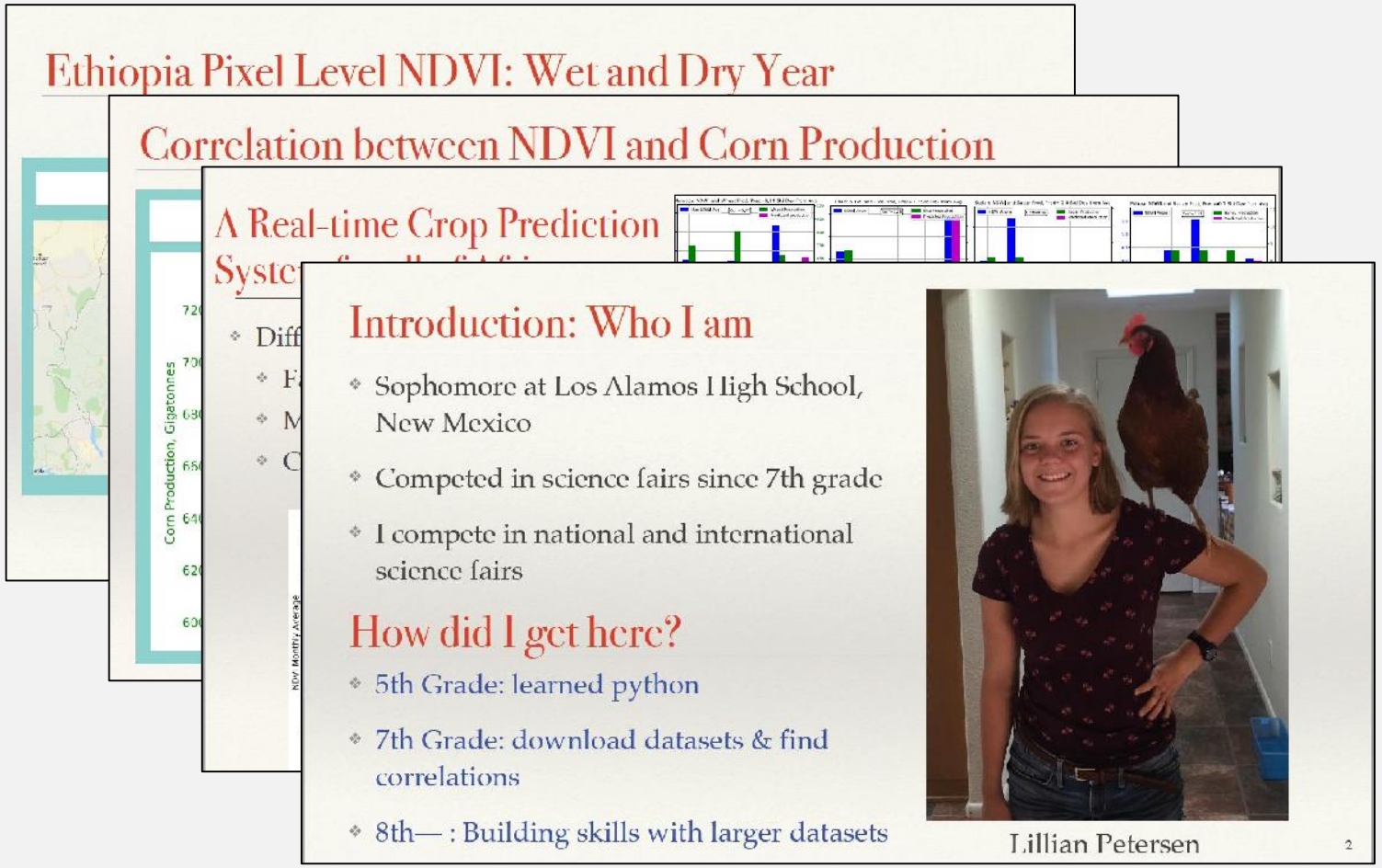


A visual comparison of all crop extent products, shown in green, overlaid on Google Earth Imagery
Xiong, J., et al. (2017). "Nominal 30-m Cropland Extent Map of Continental Africa by Integrating Pixel-Based and Object-Based Algorithms Using Sentinel-2 and Landsat-8 Data on Google Earth Engine." Remote Sensing 9(10): 1065.

Data Scientists…?



Programming Skills
Statistics
Machine Learning
Data Wrangling



Summary / Key Takeaways

We, Platform for Big Data in Agriculture,

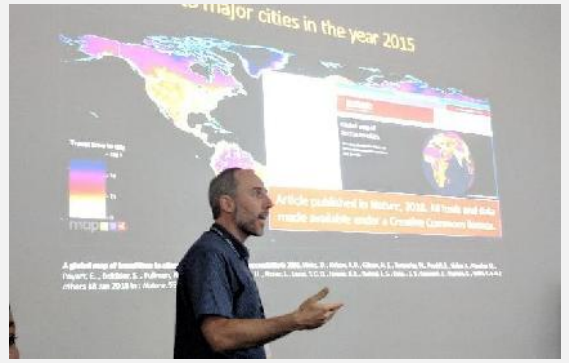
- Work with Centers to make CGIAR's data assets F.A.I.R. (Findable, Accessible, Interoperable, and Reusable).
- Support CRPs/Centers to enhance (big) data management and improve analytical capacity through trainings, partnerships, shared services, and innovation processes.
- Facilitate the coordination between our scientists and (external) technical capacities to bring rapidly advancing data science into CGIAR.
- Provide innovation space for CRPs to pilot novel research ideas using data science with new (external) technical partners.

Call for Action

1. **Engage with us!**
   Everything we do is about *Collaboration*. We don't do research ourselves. We support CRPs/Centers to incorporate (big) data science in your research through the collaborative partnerships, facilitated both internally (CoPs) and externally (*Convention* and *Inspire Challenge*).

2. **Leverage us!**
   We welcome to learn your data science needs and challenges. We're here to strengthen CGIAR (you)!

3. **Let us all be data scientists!**
   We need more CGIAR scientists with data science skills.



CGIAR at the 2018 FOSS4G

3 training workshops and 26 academic presentations