

# #CoP\_SED

Gideon Kruseman, CIMMYT

Coordinator community of practice on socio-economic data

1<sup>st</sup> annual scientific conference and capacity development workshop of CGIAR Collaborative Platform for Gender Research December 5-8 2017 at KIT Royal Tropical Institute in Amsterdam.



Platform for  
Big Data  
in Agriculture



Community of practice  
on socio economic data  
#CoP\_SED



# Overview

- Quick recap of the CGIAR Platform for Big data in agriculture
- Introduction to #CoP\_SED
- Flavor of what the platform holds in store
  - Blockchain
  - Data integration
  - Nifty new analytical tools
- Getting involved

# CGIAR Platform for Big data in Agriculture



Platform for  
Big Data  
in Agriculture

[bigdata.cgiar.org](http://bigdata.cgiar.org)

#BigDataInAg

# CGIAR Platform for Big Data in Agriculture

*Harness the capabilities of big data to accelerate and enhance the impact of international agricultural research*



Bringing Big Data to Agriculture,  
and Agriculture to Big Data



Platform for  
Big Data  
in Agriculture

CIMMYT

Community of practice  
on socio economic data  
#CoP\_SED



## A Place, A Movement, A Mission...



Platform for  
Big Data  
in Agriculture



Platform for  
Big Data  
in Agriculture

to solve agricultural  
development problems faster,  
better and at greater scale

# #CoP\_SED



Platform for  
Big Data  
in Agriculture

CIMMYT

Community of practice  
on socio economic data  
#CoP\_SED



# Our challenge

CGIAR collects lots of socio-economic data.

This data is not always made available for further use, other than the primary purpose of collection.

If it is archived for further use, it is not interoperable with other data sets.

Besides CGIAR data there are other data sets that are relevant: *e.g.* LSMS-ISA

Our challenge is to start making the data available and interoperable.



Platform for  
Big Data  
in Agriculture



Community of practice  
on socio economic data  
#CoP\_SED





# Our challenge

Socio-economic data is by far the most complicated type of data:

- Structured, semi-structured, unstructured
- No standards for kind of data collected
- No standards for values collected

We do not seek to enforce standards that will be obsolete before they are developed!

Where useful, we will use standards

But, we seek interoperability through different means



Platform for  
Big Data  
in Agriculture

CIMMYT

Community of practice  
on socio economic data  
#CoP\_SED



# Challenge

Ensure that data sets have relevant metadata attached.

1. Technical metadata: format, location
2. Descriptive metadata: CGCore, DublinCore
3. Structural metadata: machine readable information that allows harvesting data from wherever it is deployed into workflows and analytical tools



The screenshot shows the CIMMYT Research Data Dataverse interface. At the top, there are logos for CIMMYT (International Maize and Wheat Improvement Center) and the Research Data Repository. Below the navigation bar, the page title is 'CIMMYT Research Data Dataverse'. A search bar and 'Advanced Search' link are visible. The main content area displays a list of studies under the 'Socio-Economics' category. The first study listed is 'Zero-Tillage Service Provision as a Business Opportunity' by Rab Allen, Michalski, Joel, with ID 1152910841. The second study is 'Zero-Tillage Adoption and its Welfare Impacts at the Farm Household Level' by Rab Allen, Michalski, Joel, with ID 1152910840. The third study is 'Pathways to Sustainable Intensification in Eastern and Southern Africa - Ethiopia 2010' by Mameya Fisseha, Kassie Haneke, Yige, Chib, Maricho, Geoffrey Alemu, Solomon, with ID 1152910746. The fourth study is 'Pathways to Sustainable Intensification in Eastern and Southern Africa - Tanzania 2010' by Mameya Fisseha, Kassie Haneke, Hani, Fulgence, Maricho, Geoffrey Alemu, Solomon, with ID 1152910754.



Platform for  
Big Data  
in Agriculture



Community of practice  
on socio economic data  
#CoP\_SED



# Working groups

- Survey of surveys
  - Define common questions
- SociO!
  - Concepts and classifications (ontology) needed to understand the data content
- Structural metadata schema
  - Make the data interoperable in a flexible and extensible way



Platform for  
Big Data  
in Agriculture



Community of practice  
on socio economic data  
#CoP\_SED



# New working groups

- Blockchain  
Solving the insuperable problems in value chains with data – bridging the trust gap and democratizing VCs
- Ethics, privacy and cybersecurity  
Moving goal posts regarding human subjects data – how to not do harm with really sensitive data
- CAPI  
Making the most of Computer assisted personal interviews
- *Gender*  
*Making gender data FAIR (Findable, Accessible, Interoperable, Reusable)*



Platform for  
Big Data  
in Agriculture



Community of practice  
on socio economic data  
#CoP\_SED

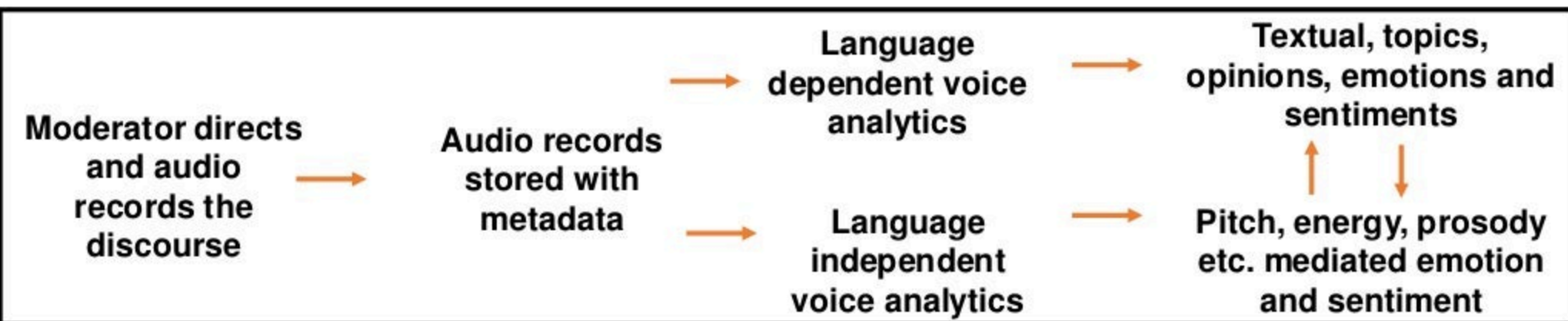


# Flavor of what the platform holds in store



Platform for  
Big Data  
in Agriculture





# Example of data integration

## Rural transformation and tipping points

- Combining quantitative and qualitative data
- Combining data sources to understand the point at which a series of small changes or incidents becomes significant enough to cause a larger, more important change.



# New tools example

## Blockchain technology overcome value chain challenges

- Cutting edge technology under development
- Using CGIAR Platform for Big Data in Agriculture as an organizational ecosystem
- Make nutritional quality traceable





# Getting engaged

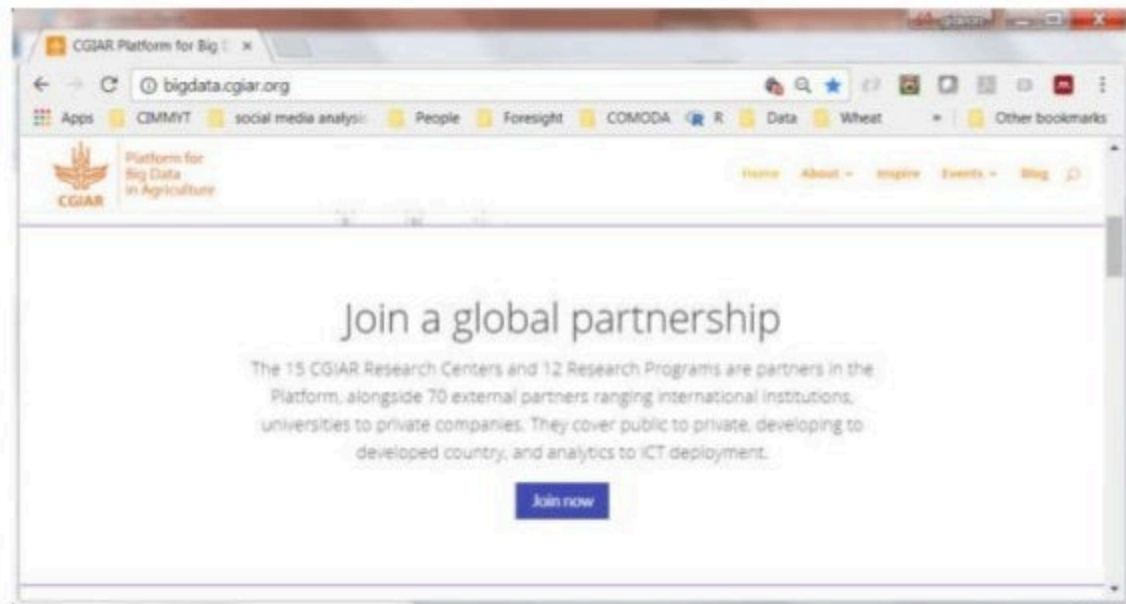


Platform for  
Big Data  
in Agriculture



# What can you do

1. Sign-up on [bigdata.cgiar.org](http://bigdata.cgiar.org)



Platform for  
Big Data  
in Agriculture

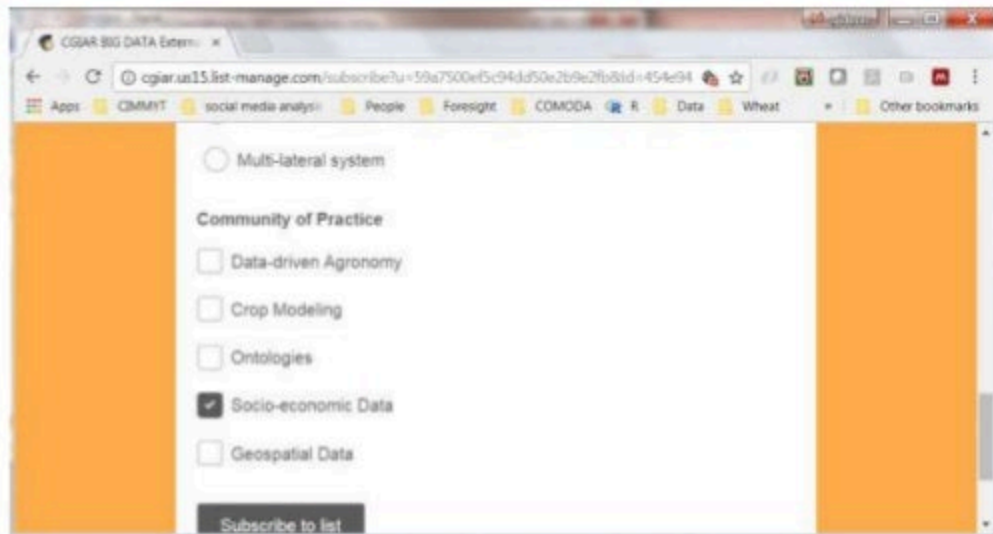
CIMMYT

Community of practice  
on socio economic data  
#CoP\_SED



# What can you do

1. Sign-up on [bigdata.cgiar.org](http://bigdata.cgiar.org)
2. In the form check the box that you are interested in the community of practice on socio-economic data



Platform for  
Big Data  
in Agriculture

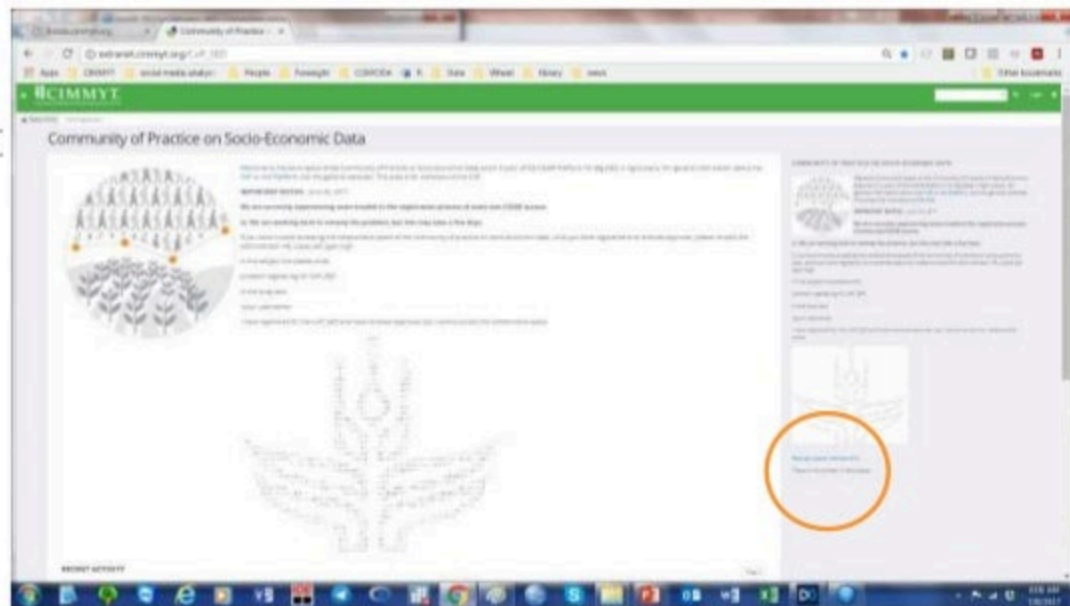


Community of practice  
on socio economic data  
#CoP\_SED



# What can you do

1. Sign-up on [bigdata.cgiar.org](http://bigdata.cgiar.org)
2. In the form check the box that you are interested in the community of practice on socio-economic data
3. Register in the collaborative space of the CoP



Platform for  
Big Data  
in Agriculture

CIMMYT

Community of practice  
on socio economic data  
#CoP\_SED



# What can you do

1. Sign-up on [bigdata.cgiar.org](http://bigdata.cgiar.org)
2. In the form check the box that you are interested in the community of practice on socio-economic data
3. Register in the collaborative space of the CoP
4. Sign-up for active participation in working groups

extranet.cimmyt.org/working-groups

## Working groups

This section contains links to various working groups within the community of practice. This is where the real work takes place, with the ultimate goal of harnessing the capabilities of socio-economic big data to enhance the impact of our research for development, through making data interoperable, doing new analyses using interoperable data or enhancing decision-making by farmers, policy makers and other stakeholders.

Using the drop down menu at the top of the page you can access the spaces of the different working groups, both active and inactive, we currently have the following active groups:

- [Survey of surveys](#)
- [Concepts and classifications: SociOI](#)
- [Structural metadata schema](#)

You can sign-up for the working groups using the following [form](#).

If you want to start a new working group, please feel free to [contact the coordinator](#)

[Survey of surveys group](#)

This page is Public Challenge:



Platform for  
Big Data  
in Agriculture

CIMMYT

Community of practice  
on socio economic data  
#CoP\_SED





Platform for  
Big Data  
in Agriculture

[bigdata.cgiar.org](http://bigdata.cgiar.org)

#BigDataInAg