

Community of practice on *socio-economic data* within the CGIAR Platform for Big Data in Agriculture

Plan of work and budget

Introduction

After 18 months of planning The CGIAR Platform for Big Data in Agriculture is now operational, led jointly by CIAT and IFPRI, with all 15 centers involved in implementation. We hope to serve both the CGIAR Research programs (CRPs) and the Centers in a number of ways, and help move CGIAR as a whole into an area of comparative advantage in terms of digital agriculture. We are convinced that the success of this platform will hinge on us listening to the needs of the CRPs and the Centers, and being agile and cutting-edge in providing solutions for those needs. This is a very exciting domain where we think CGIAR can build reputation, and be a source of growth over the coming years.

Here's a quick overview of how the CGIAR Platform for Big Data in Agriculture is organized and what we're hoping to achieve:

- The platform is co-led by CIAT and IFPRI, and will have a Steering Committee made up of participating centers + external partners. An International Advisory Board will be set up, consisting of representatives from other related global and regional initiatives to ensure that the Platform is relevant, adding value, and partnering appropriately. More on governance to follow in next 2-3 weeks.
- The platform has 3 modules:
 - **Organise:** this is about ensuring the CGIAR data is generated, managed and made available to the highest standards. Led by IFPRI, with Medha Devare at the helm (sat in the SMO), the immediate goal is to support centers and CRPs to reach compliance on [the CGIAR Open Access Data Management \(OADM\) policy](#) and enable CGIAR resources to be seamlessly discoverable and interoperable. This module is an enabler – if we are serious about big data, we need to ensure our own data is appropriately managed and FAIR (Findable, Accessible, Interoperable, Reusable).
 - **Convene:** this is about putting CGIAR in the center of the world of big data related to developing agriculture. The module consists of Communities of Practice around key themes to ensure collective action across the CGIAR on topics of strategic importance. We expect to solidify new partnerships for CGIAR with big data powerhouses globally and regionally through an annual convention, focused on idea development, capacity enhancement, and partnership building.
 - **Inspire:** this is about hitting the ground with big data related solutions to CGIAR problems. We will identify key topics in the CRP portfolio where a big data approach might solve development problems faster, better and at greater scale. Calls will then be made for concept notes outlining novel means of solving the problems, and funds assigned to teams (CGIAR + partners) to go about and innovate.

- **Our partners have the opportunity to engage on a number of fronts**, especially through participation in the Communities of Practices (CoPs) to raise the profile of your work, access new partnerships, and help fill critical gaps, and also to help develop novel Inspire ideas and pitch concept notes for funding.
- More info can be found in the [full proposal](#).

CoP socio-economic data

CIMMYT are strongly engaged already in the CGIAR Platform for Big Data in Agriculture, leading a Community of Practice on crop modelling (Matthew Reynolds), and another on economic data (Gideon Kruseman).

The **Community of Practice on socio-economic data** aims at bringing together CGIAR centers, academia, not-for-profit research and development organizations and private sector partners interested, able and willing to tackle to major issues related to socio-economic (survey) data and moving forward to making the data interoperable in order to enhance the impact of the research for development of the CGIAR and the use of CGIAR socio-economic data for partners in development.

Some of the issues we could/should/will be tackling in the CoP include, but not be limited to:

- Ontologies for socio-economic data, both in terms of variables as well as in terms of the classifications of valid non-numeric variable values. Related to the ontologies are the protocols for defining the concordances between classifications. This is a key area for making data interoperable. It covers not only the definitions but also the way this structural metadata is stored. This metadata should in principal be machine readable in order to allow the development of both generic and custom ETL procedures.
- Can we define a set of standard questions for surveys that allow the development of very large scale dynamic data sets that will allow enhanced analytics in the area of agricultural development economics
- Best practices in data capture.

In order to **map out our key stakeholders, we would like to have those interested respond to our [short survey](#)**.

Once we've mapped out participation, there will be a number of surveys going to participants to gauge interests and set priorities.

Very much looking forward to working with you and your organization on the Big Data in Agriculture Platform.

Please feel free to pass this on to anybody you think might be interested in participating in the CoP on socio-economic data.

Work plan

The first quarter of the year 2017 is almost over so the work will take place in the three remaining quarters. In the table below

Time slot	Desired outcome	Output/deliverable	activities	participants
Q1	The Community of Practice on socio-economic data aims at bringing together CGIAR centers, academia, not-for-profit research and development organizations and private sector partners interested, able and willing to tackle to major issues related to socio-economic (survey) data	Blog(s) on the CoP on socio-economic data	Spreading the information on the CoP across our networks	Coordinated by Gideon Kruseman, all CoP members participate
Q2	Initial priorities set regarding to socio-economic (survey) data and moving forward to making the data interoperable in order to enhance the impact of the research for development of the CGIAR and the use of CGIAR socio-economic data for partners in development.	Report on initial priorities Blog(s) on the CoP on socio-economic data initial priorities	Series of surveys, remote meetings and possibly in-situ meetings with CoP members to gauge interests and priorities. Conduct a CGwide gap analysis to determine strengths and weaknesses of each CG centre in terms of socio-economic data	Coordinated by Gideon Kruseman, all CoP members participate
Q3	Moving forward to making the data interoperable in order to enhance the impact of the research for development of the CGIAR and the use of CGIAR socio-economic data for partners in development.	Forums established on at least two priority issues (examples mentioned earlier)	- Promotion of active discussions on key priority issues. - Coordination with other groups within the Platform and within and outside the CGIAR to create synergies. - Promote harmonization of research practices in the social and economic science community –i.e. common research methods, core	Coordinated by Gideon Kruseman. Specific roles on key issues TBD

			data sets, ontologies etc.- to facilitate data sharing and broader R4D goals.	
Q4	(see above)	Report and blogs on progress on at least two priority issues. Workplan 2018	Discussions on key priority issues. Work planning for 2018	Coordinated by Gideon Kruseman. Specific roles on key issues TBD

Budget

The total budget we have to work with is 100K USD. 25K is earmarked for the coordination time of the CoP Coordinator (Gideon Kruseman at CIMMYT). The remainder of the budget can be used to achieve the deliverables of the CoP in order to achieve its aims. This includes travel costs, publication costs, ICT and depending on the priorities determined in Q2 specific operational costs.

Given the limited budget, the CoP aims to raise additional funds with donors for specific projects related to the key priority issues.

- 1) Share the concept notes with respective participants to get feedback -and consensus as far as possible- and refine the proposals.
- 2) Present the concept notes to international and regional development agencies.
- 3) Consider in what context some initial funding for these proposals could be made available from *CGIAR Platform for Big Data in Agriculture* and relevant CRPs, to help with leverage of funds from other bodies.