



BIG DATA EUROPE

# Second Community Workshop for Societal Challenge 2 – Food and Agriculture

CLORA headquarters, Brussels, 30 September 2016

## Agenda

09:00 - 09:30	<i>Arrivals &amp; coffee</i>	
09:30 - 10:00	<b>Welcome &amp; Introduction</b>	Kimmo Rossi, Nikos Manouselis
10:00 - 11:00	<b>Session 1. Towards a Data Ecosystem</b>	
10:00 - 10:10	Big Data & IoT actions in DG CONNECT: policy context and EC expectations	Kimmo Rossi
10:10 - 10:20	Report from DG AGRI data events & consultation process for next WP	Iman Boot
10:20 - 10:30	GODAN status & report from 2016 Summit	Martin Parr
10:30 - 10:40	Open Data Charter Agricultural Sector Package	Andre Jellema
10:40 - 10:50	Syngenta/GODAN Data Ecosystem Paper & WG	Graham Mullier
10:50 - 11:00	The Chania Declaration & what's next	Johannes Keizer
11:00 - 11:10	Q & A session	all
11:10 - 11:30	<b>Coffee break for networking &amp; discussion</b>	
11:30 - 12:30	<b>Session 2. Scientific communities &amp; open research data</b>	
11:30 - 11:40	DG RTD, the European Science Cloud (EOSC) & H2020 Open Research Data	Wim Haentjens

11:40 - 11:50	Towards an e-infrastructure roadmap for open science in agriculture (e-ROSA)	Odile Hologne
11:50 - 12:00	CGIAR Big Data in Agriculture Platform	Medha Devare
12:00 - 12:10	The GODAN Action project	Sander Janssen
12:10 - 12:20	Positioning Europe in the global context	Nikos Manouselis
12:20 - 12:30	Q & A session	all
<b>12:30 - 13:30</b>	<b>Light lunch break</b>	
<b>13:30 - 14:30</b>	<b>Session 3. Tools &amp; services</b>	
13:30 - 13:40	The EU Open Data Portal	Agnieszka Zajac
13:40 - 13:50	Publications Office semantic repository (Cellar)	Marc Kuster
13:50 - 14:00	VocBench and its use in Eurovoc	Willem Van Gemert
14:00 - 14:10	The Global Agricultural Concept Scheme (GACS)	Johannes Keizer
14:10 - 14:20	The Big Data Europe initiative	Simon Scerri
14:20 - 14:30	Q & A session	all
<b>14:30 - 15:00</b>	<b>Coffee break for networking &amp; discussion</b>	
<b>15:00 - 16:00</b>	<b>Session 4. Where do we want to go for a data revolution in agriculture &amp; food?</b>	
	Views & position statements from the participants	
<b>16:00 - 17:00</b>	<b>Wrap up &amp; action points</b> What does the EC need from the community? What does the community need from the EC? Who commits to do what next?	

## Expectations and Background

The first BDE workshop for Societal Challenge 2 “food and agriculture” gathered mainly the research community (Bioversity, CYMMIT, CRA Italy, INRA, Wageningen University, Global Plant Council) and some representatives from international organizations (FAO) and the industry (Bayer).

Based on lessons learnt from the first workshop, especially the weak involvement of EC Units, the objective of this second workshop was to get together a selected group of EC representatives that are working on agendas or initiatives of relevance (from DG AGRI, [DG CONNECT](#), [DG RTD](#) as well as the [EU Publications Office](#)) to meet a number of experts from key non-EC organizations and companies working on data for food and agriculture (FAO, CGIAR, GODAN, INRA, Alterra Wageningen, TNO, Syngenta, AgroKnow) to share recent developments.

Therefore, the Big Data Europe Consortium, with its partners for societal challenge 2 [FAO](#) and [AgroKnow](#) and with support from the Global Open Data in Agriculture & Nutrition ([GODAN](#)) initiative and the Institut National de la Recherche Agronomique ([INRA](#)), organized the second workshop on big data for food and agriculture in Brussels, with a strong focus on sharing experiences with EC representatives as well as providing some input for their agendas.

The one-day workshop took place at the CLORA premises in Brussels on 30/09/2016 and followed the [DG AGRI](#) “Digitising the Agri-food Sector” meeting that had taken place the previous two days also in Brussels. Organizing the workshop back-to-back with the DG AGRI meeting gave us the opportunity to have some members of the BDE project and some key partners invited to the DG AGRI meeting and to get a full report of the meeting by Iman Boot during the BDE workshop (more on this in the conclusions).

Some of the key expectations we had for this workshop were:

- Making our EC colleagues aware of key initiatives at the global level that involve several actors from different stakeholder groups (inter-governmental organizations, research institutes and academy, research networks, industry and SMEs) on data infrastructures and data sharing in the field of food and agriculture;
- Learning from our EC colleagues about the most recent developments in terms of their funding agendas and projects relevant to big and open data in agriculture and food.
- Discussing how the community is building a common roadmap for the data infrastructure elements that the agri-food sector needs (covering aspects such as physical structures, software, standards, and governance) and highlighting the commitment of major sector representatives to support coordination and collaboration, in a way that represents international perspectives and includes both the public and private sector.

## Setting the scene

The agenda (see above) was organized around three main sessions:

**1) Towards a Data Ecosystem**, where participants from DG CONNECT, DG AGRI, GODAN, Syngenta and FAO presented their latest endeavors to involve their respective stakeholders in envisioning what a data ecosystem for food and agriculture should look like and what has to be done to get there.

Kimmo Rossi (DG CONNECT) and Iman Boot (DG AGRI) presented the policy context and the funding agendas of their respective Directorates, including a report of the results of the recent DG AGRI workshop. Martin Parr and Andre Jellema (GODAN) presented some key activities of

the GODAN initiative both in advocacy and convening the stakeholders and in designing an “open data charter” for our sector for the [ODC initiative of the G20](#). Graham Mullier (Syngenta UK) presented the results of Syngenta’s collaboration with GODAN partners on what components should make up a data ecosystem for our sector and Johannes Keizer (FAO) presented the “Chania declaration”, the latest commitment to a common agenda signed by several organizations in May 2016.

**2) Scientific communities & open research data**, where participants from the research community presented their current activities under different funding schemes (EC and beyond).

Wim Haentjens (DG RTD) and Odile Hologne (INRA) presented two projects on research data funded by the EC: the already advanced European Open Science Cloud and the now-starting eROSA project (“Towards an e-infrastructure roadmap for open science in agriculture”). Medha Devare ([CGIAR](#)) and Sander Janssen ([Alterra](#), [WUR](#)) presented publicly funded activities for the improvement of global and institutional opening up of data: the efforts and aspirations of the CGIAR in opening up their data (also learning from other domains) and the GODAN Action project, funded by the Department for International Development of the UK government (DfID). Nikos Manouselis (AgroKnow) then gave his perspective on how European initiatives can guide / lead the changes in the food and agriculture data industry, through already existing mechanisms, like the alignment between public and private sectors, the approach of having different infrastructures communicate, the attention given to business models and SMEs. EC projects and GODAN can also help coordination in order to avoid duplication.

**3) Tools & services**, where participants presented relevant global and regional infrastructural components for data management and data sharing.

Agnieszka Zajac and Marc Kuster (EU Publications Office) presented respectively the EU Open Data Portal, illustrating the technologies implemented to make EU datasets open, and the CELLAR semantic repository to make legal documents from the EC semantically accessible. Willem Van Gemert (EU Publications Office) and Johannes Keizer (FAO) presented two platforms for vocabularies: the vocabulary management tool VocBench (used for both AGROVOC and Eurovoc) and the vocabulary framework Global Agricultural Concept Scheme (GACS), a common environment for hosting, linking, referencing and publishing vocabularies for food and agriculture.

## Discussions

Presentations were followed by Q&A sessions and a final discussion took place at the end of the workshop.

Among the most relevant issues raised by presentations and questions asked and discussed at Q&A time were:

- Importance of public-private partnerships
- DG AGRI included “combining and exchanging data” among the five priority areas for the next biennium work program
- Beyond technical difficulties (formats, connectivity, storage), the very important issue of privacy (interesting note: there will be a workshop dedicated to data ownership and privacy early next year)
- Other reason for reluctance to share: lack of incentives; besides the “stick”, study more the carrot options? Need for reward mechanisms, business models; incentive schemes; consider also the value of analytics platforms (example of sequence data): being able to use these platforms incentivates the sharing of data
- Learning from other sectors (e.g. health: example of idea of algorithms moving over the internet instead of data...)
- Challenges: a) open data benefits only the big actors? b) new monopolies and bigger digital divide?
- There doesn't seem to be a mature ecosystem with which to interlink data: so far, work with data is ad hoc, not systematic and routinely repeatable
- Thinking beyond Europe, as agricultural data is a global issue
- The opportunity of the European Open Science Cloud, high priority for the Commissioner and also discussed in the broader community with Barend Mons to see what role they can have in the EOSC (e.g. CGIAR will try and bring their data to the EOSC to see what can be done)
- Issues of integration / interaction between the EOSC and other platforms: OpenAire, CORDIS, aginfra+; possible DG CONNECT project to build an infrastructure for the EOSC?
- The upcoming eROSA CSA project as an opportunity to collectively shape a data infrastructure / ecosystem for food and agriculture
- Standards still an issue, very fragmented landscape, too many standards, need for recommendations; need for a clearinghouse?
- Role of the private sector: data from science is something where government has an easier role; the role of government with private sector is weaker; the industry sees the point in collaborating, but the public sector has to make sure that the industry doesn't govern
- Complexity, need for multi-dimensional analysis
- Comparison of investments between the EC and the big industry; comparison of investments in different domains: Food 2030 (about the challenges in food production and safety): recognized investment gap with respect with other fields (and e.g. regarding food much lower investment than in US or other areas)
- Issue of how to involve member countries in these initiatives
- Can Europe lead the changes in this industry? We have a culture of collaboration, we have some world-class research institutions on board, some elements of the infrastructure exist; some standards (if not data standards) exist. How? Alignment between public and private sectors; to have different infrastructures communicate; to do business; to coordinate to avoid duplication

- EU Open Data Portal for data published by the EC: possible to extend / re-use the platform for broader data sharing? Explore synergies with similar initiatives at the global level

The final discussion collected feedback from all participants and brought forward some more details on the issues already discussed:

- On incentives: open data linked to open evidence
- On public-private: importance of pre-competitive spaces
- On unfair benefits: corrective measures to avoid open data benefitting big actors more
- On coordination / consultation: Importance of having these priorities in the DG AGRI program
- On infrastructure and standards: need for reference frames
- On coordination / consultation: good to see so many initiatives having so much in common (standards, tools...)
- On coordination / consultation: importance for the broader community of having an audience in the EC, and of having something useful to contribute to the EC
- On standards: we are still banging our heads against the same issues: standards, identifying the gaps, interoperability. Why are we not making more progress?
- On coordination / consultation: importance of this learning exercise between EC and other actors; Importance of having the three relevant DGs together
- On incentives: more focus on impact
- On public-private: GODAN can have a role on negotiating pre-competitive spaces
- On coordination / consultation: GODAN needs to do more to engage and have a solid partnership with the DGs; DG RTD ready to hear specific recommendations from this community; importance of widening the scope for those who work within/for Europe; convene a broader group around these issues? GODAN offers its convening power; also consider the upcoming call under Food2030, there is an action around convening.

In the end, grouping the key observations raised in the final discussion, a few broad areas clearly emerged as priority for all participants:

1. Convening mechanisms for making these consultations systematic (and seeing if the EC can also build the case for further systematic investment); also to help align the technology investments (avoiding duplication)
2. New incentives and business models for data sharing (think of automated mechanisms besides “stick and carrot” incentives)
3. Work still needed on interoperability and standards, especially on demonstrating impact and collecting use cases and stories to more clearly understand what is needed / missing; think of evolution of standards into Standards-as-a-Service.
4. Need for shared data services (big data platforms, e.g. big-data empowered learning mechanisms) and agreed reference datasets (as standards and for learning mechanisms)

# Conclusions

A first conclusion is about the important results achieved thanks to the fact that the workshop was organized back-to-back with the DG AGRI meeting:

- we managed to have some members of the BDE project and some key partners invited to the DG AGRI meeting (Johannes Keizer from FAO, Odile Hologne from INRA, Sander Janssen from Alterra) to influence the agenda for the next biennium;
- we could have a full report of the meeting during the BDE workshop, from which it appeared that our main objective had been achieved: open data (more precisely “combining and exchanging data”) is one of the 5 identified priorities for DG AGRI for the next biennium (the others being: farm optimization, food traceability, environmental issues / sustainability, creating new businesses).

The workshop itself met the expectations that we had (as set out in the background chapter), in that:

- We managed to make our EC colleagues aware of key initiatives at the global level on data infrastructures and data sharing in the field of food and agriculture, and comments from the EC representatives confirmed that they appreciated the sharing of information and agreed that there should be some regular convening mechanism for these consultations.
- We learned a lot from our EC colleagues about the most recent developments in terms of their funding agendas and projects relevant to big and open data in agriculture and food.
- Many elements of the “common roadmap” for a data infrastructure / ecosystem that the global community is trying to build (from physical structures to software, standards and governance) came up in the discussion and there was a general agreement on the need for coordination and collaboration, in a way that represents international perspectives and includes both the public and private sector.

In particular, as concerns the impact of this workshop on our work in the Big Data Europe project, the areas of priority n. 3 and especially n. 4 above show how the project has relevance for both the EC DGs and the broader food and agriculture community, and we will make sure that the pilots and the big data platform instance for SC2 are built bearing in mind all the observations of participants in this workshop.

---

All presentations are available on Slideshare: <http://www.slideshare.net/tag/BDESC2W2>.

All photos are available on Flickr:

<https://www.flickr.com/photos/133018547@N06/albums/72157675012276145>.

# List of participants

	Name	Organization
1	Iman Boot	DG AGRI
2	Christopher Brewster	TNO
3	Ana Cuadrado Galvan	DG AGRI
4	Medha Devare	CGIAR
5	Wim Haentjens	DG RTD
6	Odile Hologne	INRA
7	Sander Janssen	Alterra (WUR)
8	André Jellema	Data-Impact
9	Pythagoras Karampiperis	Agroknow
10	Johannes Keizer	FAO / GODAN Secretariat
11	Marc Kuster	Publications Office of the EU
12	Nikos Manouselis	Agroknow
13	Graham Mullier	Syngenta UK
14	Pilar Ocon-Garces	DG CONNECT
15	Martin Parr	CABI / GODAN secretariat
16	Valeria Pesce	FAO / GFAR
17	Jane Richardson	EFSA
18	Kimmo Rossi	DG CONNECT
19	Simon Scerri	Fraunhofer IAIS
20	Derek Scuffell	Syngenta UK
21	Willem Van Gemert	Publications Office of the EU
22	Agnieszka Zajac	Publications Office of the EU
23	Panagiotis Zervas	Agroknow

*Report by Valeria Pesce (FAO/GFAR), Nikos Manouselis (AgroKnow)*