



RESEARCH DATA ALLIANCE

# Agrisemantics WG Breakout session

S.Aubin, C. Caracciolo, B. Whitehead  
Philadelphia, 2d April 2019

[http://bit.ly/AgriSem\\_Phillly\\_19](http://bit.ly/AgriSem_Phillly_19)

# Agenda

**11:30 - 11:45** Introduction to the WG group

**11:45 - 12:30** Group output: presentation and discussion

**12:30 - 13:00** What next?

# What we've been working on

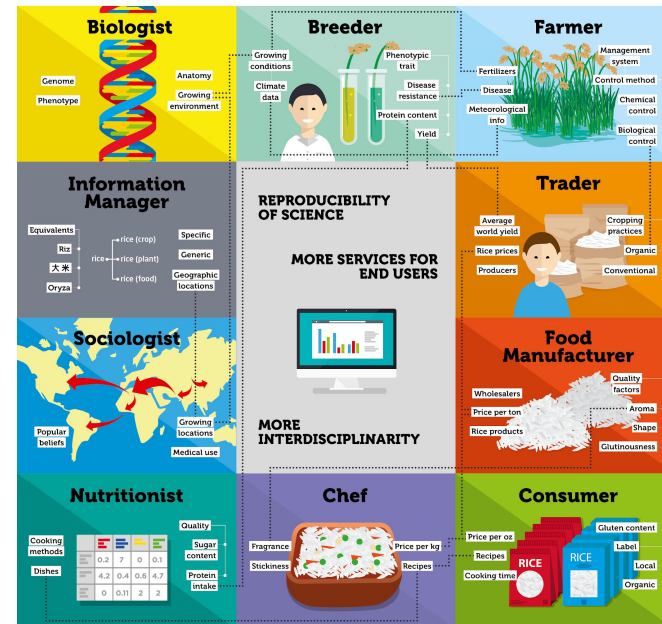
Semantics is everywhere, and applies to any information object, but it may not always be explicitly stated or expressed in machine processable formats.

**Semantic interoperability** applies when one is programmatically able to know if similar items from distinct datasets or information systems actually refer to the same objects of the world

**Semantic resources** : controlled vocabularies, taxonomies, thesauri, ontologies...

## SEMANTICS - THE WAY TO RECONCILE POINTS OF VIEW AND DATA

THE EXAMPLE OF "RICE"




<http://dx.doi.org/10.7490/f1000research.1114998.1>

# Introduction to Agrisemantics

**Objective of the group** : Envision the seamless use and creation of semantic resources supporting agricultural and food data findability and interoperability

## Life of the WG

- Launched : January 2017, now wrapping up 
- 106 registered members (public & private organisations, worldwide)
- Meetings
  - Panel @MTSR 2016-11
  - P9 Barcelona 2017-04: sharing knowledge and vision
  - P10 Montreal 2017-09: gather evidence on requirements from the ag-community
  - P11 Berlin 2018-03: turning requirements into recommendations
- +/- regular online meetings, collaborative writing

# Productions of Agrisemantics

## 3 Deliverables

1. A report on **Semantics Landscape for Agricultural Data**  
(applications, research trends, resources, toolkits)
2. A set of **20 use cases and a list of community requirements**  
(access, reusability, tools and services for creation and management, use in applications, standards and best practices)
3. A document on **Recommendations** to facilitate the uptake of semantics for agricultural data (version submitted to the RDA TAB for endorsement)

## A paper and presentation @MTSR2018

Caracciolo, C., Aubin, S., Whitehead, B., & Zervas, P. (2019). Semantics for Data in Agriculture: A Community-Based Wish List. In Contributions to Management Science (pp. 340–345). Springer International Publishing. [https://doi.org/10.1007/978-3-030-14401-2\\_32](https://doi.org/10.1007/978-3-030-14401-2_32)

# Step 1: landscaping the situation

**+ semantics in applications** for search, information extraction, data organization, integrating third party data, reasoning...

**An active, diverse, global research community** at the crossroads of ag/food and computer sciences

**Initiatives and projects:** the RDA Agrisemantics WG, GODAN, the eROSA project, EOSC...

**Standards & principles** to rely on: W3C standards, FAIR principles, open licenses

**Little reuse** of semantic resources. Only a few are accessible, machine readable; few licenses stated; poor documentation

**Uneven coverage** of subdomains of agriculture and food

**Fit-for-purpose assessment difficulties**

**Tools difficult to choose** by non experts:

- many tools for editing (specific for given semantic resource types);
- few tools for visualization, documentation, quality assessment
- no mainstream mapping tools

<http://bit.ly/AgSemLandscape>

# Step 2: what's needed?

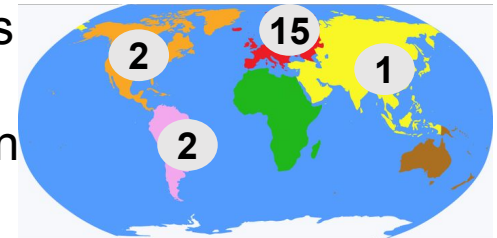
One-to-one interviews and a questionnaire



<http://bit.ly/AgSemReqUC>

## 20 use cases collected from

- 15 from research organizations
- 3 international organizations
- 1 professional
- 1 governmental organization

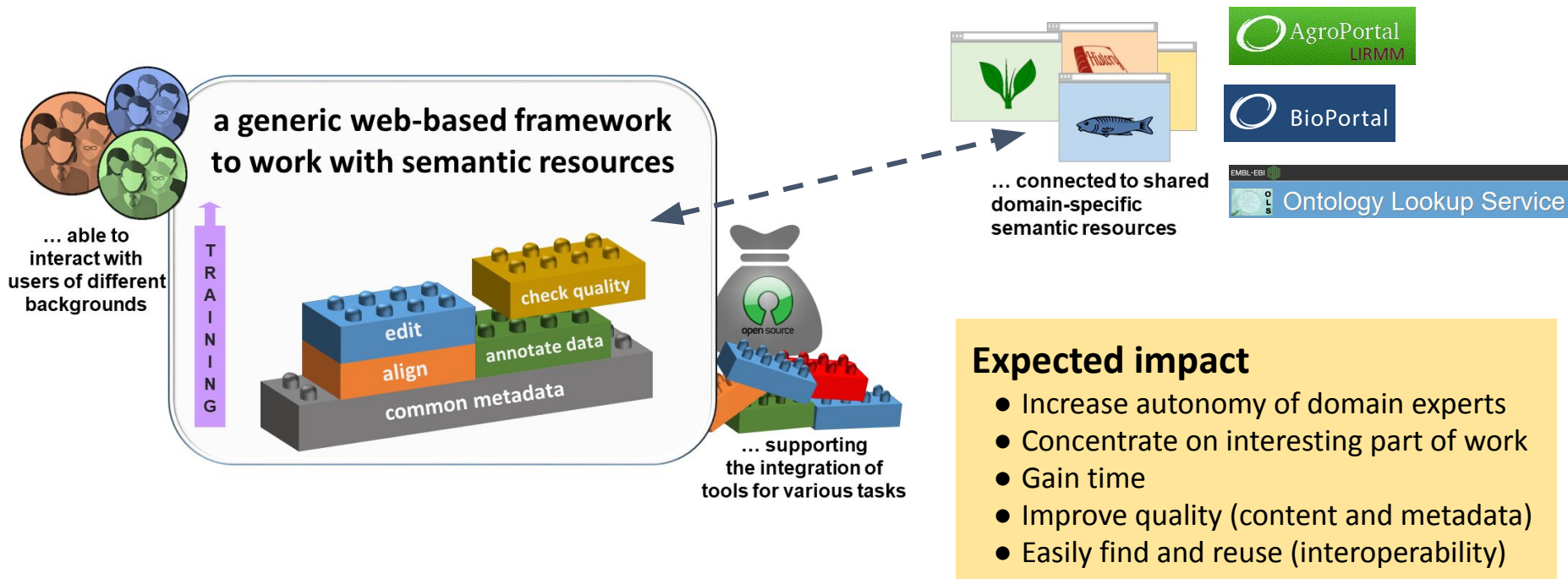


## Findings

- need to publish and share semantic resources according to Semantic Web standards (W3C)
- requirements are not specific to agriculture
- what is specific to our community?
  - terminology adopted
  - access to training (and types of training)
  - expectations about interfaces and functionalities

# Step 3: what can be done or improved?

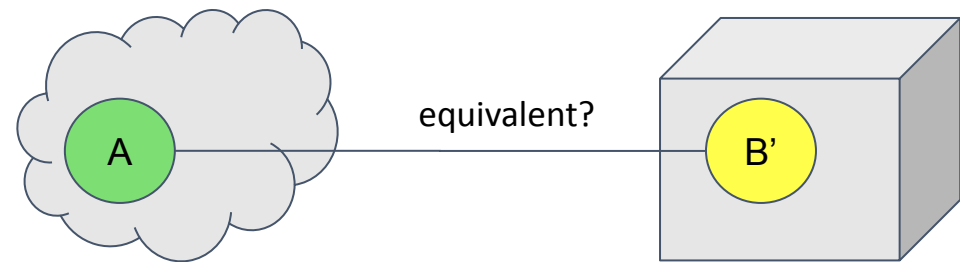
1. **A generic framework to work with Semantic Resources (SR).** This framework should be available online and be easily adaptable





# Step 3: what can be done or improved?

2. **Tools and resources for alignment**, i.e build the knowledge graph of agriculture and nutrition
  - alignments between 2 semantic resources
    - i. create and use modular controlled vocabularies for common objects of agriculture
    - ii. connect ontologies to them and to each others
  - between a semantic resource and a dataset in “usual” formats
  - standards needed to express and share alignements
  - alignment tasks integrated in the user environment



## Expected impact

- More modular SRs, easier to reuse
- Avoid duplicated efforts
- Increase alignment quality and usability
- Semantize more data

# Step 3: what can be done or improved?

3. Adopt **better practices**, build FAIR semantic resources
  - common metadata models for the description of semantic resources and alignments
  - global persistent identifiers (ARK, DOIs, URIs...)
  - adapt/develop usage metrics
  - cite SRs used with data, within information systems



## Expected impact

- Increase SRs and alignments findability
- Ensure usage tracking
- Generalize citation
- Augment trust, thus reuse

# Step 3: what can be done or improved?

4. **Create courses on semantics and integrate them into curricula** in computer sciences, data & information management and analytics, agronomy, bioinformatics, etc.



## Expected impact

- Show where semantics are valuable
- Increase mutual understanding of semantic experts and other actors of data
- Make semantics enter mainstream tools and research e-infrastructures
- Boost performances of semantic technologies and tools

# Who can adopt?

## Policy makers & funders

Promote and support

- Enable a **generic, extensible, web-based framework** to work with semantic resources
- Foster initiatives that **increase the discoverability** of semantic resources and services
- Promote the integration of **semantics into mainstream tools** and services
- Sustain the **creation and long-term maintenance of vocabularies** of strategic importance
- Develop **courses and training** on semantics

## Software developers

Integrate, make it simple

- **Choose standard technologies**, open licences, and shared I/O formats
- Build tools that **support known best practices**, e.g. FAIR principles
- Implement **automatic generation of metadata**
- Create **user-friendly tools** that use understandable terminology
- **Guide users** in choosing modelling approaches and file formats

## Semantics professionals

Share, document, reuse

- Publish SRs in **repositories that handle version control**
- Provide **persistent identifiers** for your resources
- **Reuse existing resources** (e.g., concepts, vocabularies, metadata scheme) when possible
- Promote standards for **resource alignment**
- Develop metrics to **assess resource usage**

## Data producers & managers

Semantize and document

- **Stay abreast of developments** in semantic technologies
- Make explicit (FAIR) **which SR(s) are used within your dataset**
- Develop **semantically enabled data types** for common features, e.g. measure units, parameters for experimental or observational data, soil properties, etc.
- **Provide documentation** when aligning datasets (for the processes and the result)

# Who can adopt?

## Policy makers & funders

Promote and support

- Enable a **generic, extensible, web-based framework** to work with semantic resources
- Foster initiatives that **increase the discoverability** of semantic resources and services
- Promote the integration of **semantics into mainstream tools** and services
- Sustain the **creation and long-term maintenance of vocabularies** of strategic importance
- Develop **courses and training** on semantics

## Software developers

Integrate, make it simple

- **Choose standard technologies**, open licences, and shared I/O formats
- Build tools that **support known best practices**, e.g. FAIR principles
- Implement **automatic generation of metadata**
- Create **user-friendly tools** that use understandable terminology
- **Guide users** in choosing modelling approaches and file formats

## Semantics professionals

Share, document, reuse

- Publish SRs in **repositories that handle version control**
- Provide **persistent identifiers** for your resources
- **Reuse existing resources** (e.g., concepts, vocabularies, metadata scheme) when possible
- Promote standards for **resource alignment**
- Develop metrics to **assess resource usage**

## Data producers & managers

Semantize and document

- **Stay abreast of developments** in semantic technologies
- Make explicit (FAIR) **which SR(s) are used within your dataset**
- Develop **semantically enabled data types** for common features, e.g. measure units, parameters for experimental or observational data, soil properties, etc.
- **Provide documentation** when aligning datasets (for the processes and the result)

# You!

# Questions or remarks on recommendations?

Full report on <http://bit.ly/AqSemRecom>

# What next? (non-exclusive options)

1. Within RDA :
  - a. “Finish and disband”
  - b. Enter maintenance phase (no fixed end)
    - i. identify adopters (members of RDA and others): how should we proceed?
    - ii. support adopters of the original recommendations
  - c. Start a new WG
    - i. to produce a new version or significant update of the recommendations, or take additional work
    - ii. dedicated to the implementation of specific recommendations
  - d. Start a new Interest Group
    - i. to serve as a platform for communication and coordination around WG topic

See <https://rd-alliance.org/groups/creating-and-managing-rda-groups/creating-or-joining-rda-working-group.html>

2. Outside
  - a. Become a GO FAIR implementation network <https://www.go-fair.org/implementation-networks/>

# Breakout discussions

- adoption of Agrisemantics recommendations (by other groups)
- ideas for a new RDA interest/working group
- non-RDA initiatives
- ...



Visit our poster during breaks!

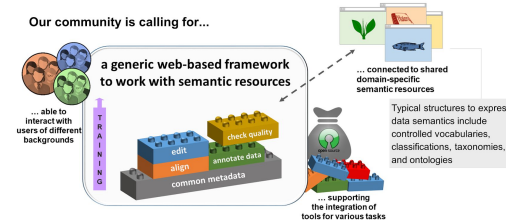
Connect Agrisemantics to other groups

Today at 4:00, attend the **RDA Business & Output Session** (Regency Ballroom AB): group output presentation and participation to panel

# Thank you!

**Agriculture is a widely interdisciplinary field** - from farm to fork and from cell to territory. In a given study or analysis, the data needed may come from different sources and communities, and at different scales of observation. Our stand is that **semantics is key to interoperability**, i.e., programmatically reuse data produced in other applications. Semantic technologies (linked open data, shared vocabularies, ontologies, ...) have been adopted by many actors in the agricultural sector. However, much is still to be done in order to make these technologies widespread. We produced recommendations on how to make semantic technologies more accessible and used in our field.

Our community is calling for...



What can you do if you are...

**Policy makers & funders**

Promote and support

- Enable a **generic, extensible, web-based framework** to work with semantic resources
- Foster initiatives that **increase the discoverability** of semantic resources and services
- Promote the **integration of semantics into mainstream tools** and services
- Sustain the **creation and long-term maintenance of vocabularies** of strategic importance
- Develop **courses and training** on semantics

**Software developers**

Integrate, make it simple

- Choose **standard technologies**, open licenses, and shared I/O formats
- Build tools that **support known best practices**, e.g. FAIR principles
- Implement **automatic generation of metadata**
- Create **user-friendly tools** that use understandable terminology
- Guide users in choosing modelling approaches and file formats

**Semantics professionals**

Share, document, reuse

- Publish SRs in **repositories that handle version control**
- Provide **persistent identifiers** for your resources
- **Reuse existing resources** (e.g., concepts, vocabularies, metadata scheme) when possible
- Promote standards for **resource alignment**
- Develop metrics to **assess resource usage**

**Data producers & managers**

Semantize and document

- **Stay abreast of developments** in semantic technologies
- Make explicit (FAIR) which SR(s) are used **within your dataset**
- Develop **semantically enabled data types** for common features, e.g. measure units, parameters for experimental or observational data, soil properties, etc.
- Provide **documentation** when aligning datasets (for the processes and the result)

**The Agrisemantics WG:** The group gathers together researchers and practitioners at the intersection between semantic technologies and agriculture, sharing the goal of enhancing agricultural data interoperability by means of semantics. The group started its activity by producing a landscape report (1) of how semantic resources are used in the area, then moved on to collect specific use cases (2) around problems and bottlenecks that people dealing with semantics for agricultural data encounter in their work. The final output of the group is a set of recommendations on what should be available in order to make semantic technologies more useful and used in applications.

Read more on <https://www.rd-alliance.org/groups/agrisemantics-wg.html>