

# On how RDA has contributed to many aspects of INRA's Open Science policy



An adoption story written by Sophie Aubin, Michaël Alaux, Esther Dzalé, Odile Hologne and Cyril Pommier

*As a participant to groups and meetings from almost the beginning of RDA, INRA (French National Institute for Agricultural Research) became a natural adopter of recommendations from agricultural related working groups and several others. In fact, RDA has contributed to many aspects of INRA's Open Science policy either directly or indirectly. Being part of the Interest Group on Agricultural Data has been particularly fruitful in terms of strategic and technical advances to increase scientific data discoverability*

*and interoperability at community level on Wheat Data for instance, or regarding semantic-based solutions for data handling and analyzing. Other RDA outputs are of importance for INRA, naming only those by the Libraries for Research Data IG which help defining the role of librarians from INRA in data management and sharing. RDA has also been the place to develop INRA's research networks and federate to address major issues induced by data sharing.*

**Esther Dzalé-Yeumo**, Unit Director INRA

## The challenge

Agri-food research produces and consumes high amounts of data which are heterogeneous in terms of source, scale, format, purpose, sensibility, etc. Yet, they need to be exchanged, understood, merged to address the interdisciplinary nature of researches in agronomy, nutrition and health, climate impact, biodiversity preservation, etc. This cannot be achieved without the adoption of standards for data representation, description and exchange protocols, inside and across disciplinary communities. Beyond local adoption by researchers, teams, and projects, standards and best practices have to be implemented in mainstream tools and shared e-infrastructure dedicated to data preservation, sharing and analysis. The forum offered by RDA helps INRA in both making informed technical and methodological choices and putting forth its needs to produce more and better (open) research data.

## The RDA outputs adopted

1. **Wheat Data Interoperability Guidelines** (a website and a publication in F1000Research 2018) have been adopted and promoted to improve data interoperability inside different communities publishing wheat data. It allowed to improve FAIRness in several working groups of the Wheat Initiative (e.g. WheatIS, Phenotype), projects (e.g. European FP7 Whealbi project) and infrastructures (e.g. Elixir, Emphasis). INRA, like many other stakeholders, adopted these recommendations.
2. **Agrisemantics Working Group recommendations** are being adopted or scheduled in future projects. "Specifically support engineering and research work, or any other initiative, which develop tools and resources for alignment (semantic mapping)." This is needed to address some complex cases of data interoperability. "Promote the adoption of common metadata models to describe semantic resources and alignments, and of global identifiers to ensure usage tracking, and citation". To avoid effort duplication in (re)developing semantic resources for describing largely shared agricultural objects. Resources producers need incentives and tools to share their work.
3. The eROSA Roadmap for a pan-European e-Infrastructure for Open Science in Agricultural and Food Sciences (led by INRA) significantly reflects outputs of several RDA groups, including Data Fabric's "Recommendations for Implementing a Virtual Layer for Management of the Complete Life Cycle of Scientific Data".
4. "23 Things: Libraries for Research Data" by the Libraries for Research Data IG also inspired the INRA STI Department in developing data services for researchers.

Find out more at:

[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit

[rd-alliance.org](http://rd-alliance.org)

or write us at

[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)

5. **The FAIRsharing Registry and Recommendations:** Interlinking Standards, Databases and Data Policies: INRA launched an institutional repository (Data INRA) early in 2018. As recommended by the FAIRsharing working group, we describe this resource in the FAIRsharing registry and linked it with existing resources. We did the same with the GNPIS repository which is domain specific repository. We also contribute to AgroPortal with semantic resources that are further referred as standards in FAIRsharing.

6. **The Data Citation of Evolving data:** INRA recommends the implementation of the Data Citation working group by its databases and repositories. In particular, the Data INRA repository implements many points of the recommendations, namely the R1 – Data Versioning: Apply versioning to ensure earlier states of data sets can be retrieved, R11 – Landing Page: Make the PIDs resolve to a human readable landing page that provides the data (via query re-execution) and metadata, including a link to the superset (PID of the data source) and citation text snippet and the R12 – Machine Actionability: Provide an API /machine actionable landing page to access metadata and data via query re-execution.

RDA has been the place to develop INRA's research networks and federate to address major issues induced by data sharing.

**Esther Dzalé-Yeumo**, Unit Director INRA

## RDA added value for INRA

Open Science brings a lot of new issues for Scientific Organizations like INRA, without any extra competency or financial means. In this context, technical issues of course, but also legal, and governance-related ones cannot be addressed locally. They truly gain to be thought and discussed inside and across disciplines, with other than usual research partners. Initiatives like RDA have a real impact in that sense on the implementation of OS directives and recommendations in an organisation like INRA, which has to deal with tones of data with limited manpower on IT activities. It contributes to making faster and better informed choices, because shared with others; to avoiding error repetition by several actors, and thus

reducing costs; and raising ambitions to reach our goals, for the benefit of a greater number of researchers and of science. INRA has been involved in RDA since the very beginning as we attended the first plenary in March 2013 in Goteborg. Since then, INRA has identified RDA as an opportunity to contribute to the identification and / or construction of data standards. INRA has been consistently actively advocating for RDA and contributing to many working and interest groups, particularly the IGAD interest group, the Wheat Data Interoperability working group and the Agrisemantics working group. Beyond these, groups, INRA has been consistently monitoring and testing the outputs of many other groups of interest.

## The impact of adoption

ThreedatastandardsincludedintheWDIrecommendations have especially proven their importance and originality: MIAPPE, BrAPI and Crop Ontology. Those three standards already allow a consistent formatting of data for MIAPPE, their description with Crop Ontology and their machine enabled exchange with BrAPI. Those success stories are at the center both of the Elixir Emphasis collaboration strategy and the Elixir Plant community. They have proven

their usability not only for wheat but also for other crops as well as forest trees. Two training sessions for the biologists and several presentations were warmly welcomed and demonstrated a good potential of adoption. The standard for genomic data have been useful to share the wheat reference sequence data recently published by IWGSC (Ramírez-González et al., 2018).

## The adoption process

Find out more at:

[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit

[rd-alliance.org](http://rd-alliance.org)

or write us at

[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)

The standards promoted in the guidelines have been highly improved between 2016 and 2018 with a new version of MIAPPE currently under review. Likewise, the integration between MIAPPE, BrAPI and Crop Ontology have been strengthened. In 2018, INRA and partners working on Wheat developed training courses for data managers based on WDI recommendations. Two days of training for Phenotype EWG biologists were organized by the WheatIS (led by INRA) and supported by the

Wheat Initiative. Two major European infrastructures, Elixir (European bioinformatics platforms description on demand) and Emphasis (European Plant Phenotyping Infrastructure) have co-written a common strategy for data standards and information systems (including MIAPPE, BrAPI, Crop ontology). The outputs adopted by INRA contribute to make data more Findable, Accessible, Interoperable and Reusable.

## Lessons learned

For WDI as well as for Agrisemantics, INRA participated with chairs and members, which really facilitates the process of adoption. The recommendations fit the needs of our organization and projects, and participants to the groups are already familiar with them when the adoption process starts. Persuading colleagues and partners to also adopt the WG outputs is also facilitated.

Participating in RDA plenaries and attending sessions on many different issues also fastens and facilitates awareness on new solutions to address our objectives. We strongly recommend the same adoption. Indeed, to achieve a good level of interoperability at a global level, we believe that it is necessary to increase the number of adopters. It might be challenging and even intimidating to adopt some recommendations as a whole. Fortunately, a partial adoption is possible for most of the recommendations and is still better than nothing.

## About INRA

Ranked the number one agricultural institute in Europe and number two in the world, INRA (French National Institute for Agricultural Research) carries out mission-oriented research for high-quality and healthy foods, competitive and sustainable agriculture and a preserved and valorised environment. One of the missions of INRA is to produce and enable access to knowledge to the international community of researchers and practitioners in agriculture but also towards policy makers and society. INRA has also to develop innovations and know-how of service to society.

These practical applications contribute to developing agricultural, industrial or service companies. INRA is engaged towards Open Science in all its aspects including FAIR principles, shared e-infrastructures, and its active participation to several RDA Interest and Working Groups. The stakeholder community of INRA is the researchers in agriculture, nutrition and environment who work at INRA in the first place and their partners. As INRA is involved in many research projects (from local to international), e-infrastructures and collaboration networks, partners are many and varied, researchers and practitioners, public and private. ).

**Contact: Esther Dzalé-Yeumo**, Unit Director INRA



Find out more at:

[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit

[rd-alliance.org](http://rd-alliance.org)

or write us at

[enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)

Find out more at:  
[www.rd-alliance.org/recommendations-outputs](http://www.rd-alliance.org/recommendations-outputs)

Visit [rd-alliance.org](http://rd-alliance.org) or write us at [enquiries@rd-alliance.org](mailto:enquiries@rd-alliance.org)