



INRA institutional data policy : data sharing and data management



Odile Hologne, Head of the Department of Scientific Information
Esther Dzalé Yeumo Kaboré, Project manager

Agenda

- ❖ Introduction : Inra, challenges
- ❖ Towards a data sharing policy
- ❖ Implementation of the data policy
- ❖ International partnership
- ❖ Conclusion

National Institute for agricultural research

FOOD

AGRICULTURE

ENVIRONMENT



4
transdisciplinary
research priorities

French National Institute for Agricultural Research (INRA) produces scientific knowledge and works for economic and social innovation in the areas of food, agriculture and the environment.

Improving the
economic, social &
environmental
performances of
agriculture

Reducing
greenhouse gases
emissions
& adapting to climate
change

Developing healthy
& sustainable food
systems

PERMANENT
STAFF

RESEARCH
UNITS

EXPERIMENTAL
UNITS

SCIENTIFIC
DIVISIONS

RESEARCH
CENTRES

8 500

200

49

13

17

Agriculture, Food, environment : big challenges

- Some Societal challenges :
 - Feed the world
 - Climate change
 - Sustainable agriculture
 - Health and nutrition
- Imply to deal with data :
 - Data driven science
 - Big data
 - Data management, sharing, and re-use
- With different point of views :
 - Politics
 - Technics (IT, libraries)
 - Scientific Disciplines
 - Intellectual property, ethics
 - Economics





Towards a data sharing policy

Key moments

- ❖ 2009 : political awareness of Inra CEO
 - ✓ RCUK data sharing policy
 - ✓ Big data in omics
- ❖ 2012 : Report of Inra scientific council « data management and sharing »
 - ✓ 9 recommendations, 1st : define Inra Policy
- ❖ 2013 : INRA data sharing policy
 - ✓ political decision
 - ✓ working groups



Gaspin, C., Pontier, D., Colinet, L., Dardel, F., Franc, A., Hologne, O., Le Gall, O., Maurin, N., Perrière, G., Pichot, C., Rodolphe, F. (2012). Rapport du groupe de travail sur la gestion et le partage des données <http://prodinra.inra.fr/record/206746>

How to define our policy ?

- ❖ Nothing from the french funding agency
- ❖ Review of most of the data sharing and management policy
- ❖ Inra : A policy about sharing, not only management
 - ✓ a top down approach at the beginning

11 data sharing and management principles

Guiding principles for responding to the challenge of open science¶

- 1.→ The data produced within the framework of research projects financed by public funds (Europe, ANR, regional funds...) must be made public (with or without a period of exclusivity);¶
- 2.→ INRA is setting up a system which is part of an open international ecosystem of research data. Setting up this system and providing support to scientists on how to use it constitute one of the priorities of INRA's Information Systems' master plan.¶

Principles in relation with intellectual property, ethics, copyright¶

- 3.→ In agreement with the ethics charter, the research data originating from INRA research teams or research infrastructures are the property of INRA or the joint property, if they are obtained in the framework of partnership;¶
- 4.→ The transfer of copyright cannot be granted to a third party without guarantee of open access rights and open re-use of data;¶
- 5.→ The dissemination of data, obtained by INRA or conveyed by third parties, must respect the legal and ethical rules which prevail, in particular in experimentation, in the use of personal data or in the use of data acquired from third parties;¶
- 6.→ All the data disseminated must include the names of the authors, teams and research units involved from INRA and from any organization associated with INRA. The status of the data in terms of open access, re-use and citation must be specified explicitly;¶

Principles relative to good practices and standards: data must be published in coherence with the national or international practices of the discipline in which they are produced.

7. → The data published must meet the appropriate international standards concerning format, metadata and persistent identifiers to be citable;
8. → The use of recognized national or international repositories, specific to certain research or disciplinary fields should be favored;

Principle of referencing/monitoring/traceability within INRA

9. → All the data published in internal or external repositories must be referenced within the INRA referencing directory, as well as a possible international referencing directory;

Principles in relation with financing the cost of scientific data management

10. → Any project submitted to a subsidising body (ANR, European Commission, etc.), if it leads to the production of data financed as part of the project, must:
 - a. → integrate the life cycle of the data produced;
 - b. → include a management plan of the data produced and the associated cost;
 - c. → specify the intellectual property rights on the data.
11. → If the period of conservation of the data is not specified in the framework of the financing of the project, the data should be retained for a period of at least 10 years from the date of the publication which is based on these data.

Next steps

- ❖ Approval of the data policy by the Inra top management at the beginning of 2013
- ❖ Launching of the implementation project, different working groups to build data infrastructures (bottom up phase):
 - ✓ by disciplines (3 fields): omics; observation; survey and personal information
 - ✓ technical : standards; citation ...
 - ✓ Intellectual property
 - ✓ Skills



Implementation of the data policy

Assessment of the situation at INRA

- ❖ INRA produces hundreds of terabytes data/year¹
- ❖ Data produced by INRA are diverse, heterogeneous and geographically scattered
- ❖ A lot of raw data are locked in researcher's PCs or scattered accross multiple infrastructures both institutional and external

1. According to an internal survey conducted in 2011

A working group: deliver services that will meet the policy objectives



Researchers

Librarians

Computer specialists

1. Chairpersons: Lydie Soler, Esther Dzalé Yeumo

Members: Alaux Michael, Bossy Robert, Carrere Sébastien, Cocaud Sylvie, Dervaux Stéphane, Fily Marie Françoise, Guinet Nicolas, Michotey Célia, Pommier Cyril, Steinbach Delphine

A one-stop data identification service

- ❖ Use of DataCite DOIs
 - ✓ A unique prefix for INRA's datasets
 - ✓ Generation of the DOIs, metadata quality management
- ❖ Use of certified platforms for the storage and preservation
- ❖ Tailored assistance to researchers
 - ✓ Granularity
 - ✓ Metadata
 - ✓ Ethic, legal and IP issues

Data stewardship

- ❖ Tools and services to help with the description, deposit, and ongoing management of research data outputs.
 - ✓ A data repository for deposit and discovery of data collections (co-)generated by INRA's researchers
 - ✓ A registry of research data assets to record the location and description of INRA's data assets

Data management support

- ❖ Tailored DMP assistance
- ❖ Training



International partnership

Our strategy

- ❖ Technical axis : involvement in RDA « Research Data Alliance »
- ❖ « Agri » axis : involvement in GODAN « Global Open Data for Agriculture and Nutrition »
- ❖ political axis: member of Science Europe



Research Data Alliance

- ❖ Launched in mars 2013 by the european commission, NSF and Australia
- ❖ Creation of 2 groups (interest, working)



The Research Data Alliance aims to accelerate and facilitate research data sharing and exchange

Agricultural Data Interoperability IG



Status: Recognised & Endorsed

The Agricultural Data Interest Group is a domain oriented interest group to work on all issues related to data important for the development of global agriculture. The interest group aims to represent all stakeholders producing, managing, aggregating, sharing and consuming data for agricultural research and innovation.

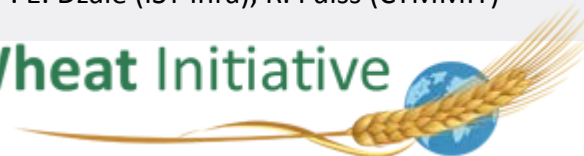
Wheat Data Interoperability WG



The Wheat Data Interoperability Working Group aims to provide a common framework for describing, representing linking and publishing Wheat data with respect to open standards.

Co chair : E. Dzalé (IST Inra), R. Fulss (CYMMIT)

Wheat Initiative



- ❖ GODAN « global open data for agriculture and nutrition », issu du G8+5 « open data for agriculture »

GODAN is an initiative that seeks to support global efforts to make agricultural and nutritional data available, accessible, and usable for unrestricted use worldwide.

- ❖ CIARD et 

- ❖ Meeting à la FAO en avril

- ✓ projet de Data Journal
- ✓ Global Agricultural Concept Scheme (aligne CABT)

What is the RING

Le CIARD RING est un **registre de sources et de services d'informations dans le domaine de l'agriculture..** [plus >>]

Tous les fournisseurs de services d'informations sur la recherche agricole pour le développement sont invités à enregistrer leurs services sur le présent portail [voir [comment](#)]



Working group on Research data

SCIENCE
EUROPE



Science Europe will:

- Promote the importance of data sharing principles – in reflecting the needs of the various disciplines, and in generally furthering research and innovation to gain maximum societal benefit – as well as guiding the definition and implementation of consistent data-sharing policies and practices;
- Contribute to the establishment of an 'ecosystem' of globally and disciplinarily interoperable, trustworthy and sustainable research data infrastructures, and explore appropriate funding structures adapted to national and organisational capabilities;
- Foster the development of relevant training and career paths, acknowledging that data-intensive research requires new and additional types of skill;
- Collaborate in developing appropriate incentive measures for scientists to archive and share their data, by promoting data management plans and support for research data collection;
- Advocate that data and scientific software contributions are treated as valuable research outputs and should play a significant role in the evaluation of research;
- Seek clarity on the legal conditions framing the envisaged re-use of research data and the possible harmonisation and changes necessary to realise this; and
- Identify where protected environments, or 'safe havens', for data are necessary, and promote the creation of policies, technical concepts and, ultimately, safe infrastructure for such cases.

ANR	Martine Garnier
INRA	Odile Hologne
CNRS	Francis André
CEA	Delphine Vidart-Dufort
Ifremer	Jean-François Masset
Inserm	Anita Burgun
IRD	Jean-Pierre Finance



SCIENCE
EUROPE
Shaping the future of research



Knowledge Exchange

Research Data Management: funding and necessary infrastructures

INTRODUCTION

The new demands in the area of research data management (RDM) together with the exponential growth in volume and complexity of research data make it clear that the existing funding in these areas will not keep pace with increasing costs. The questions arise which stakeholders are involved in maintaining a sustainable infrastructure and covering the costs of RDM activities, what are their respective roles and responsibilities and who has to pay for what.

The Science Europe Working Group on Research Data and Knowledge Exchange developed this questionnaire to get a clearer picture of the situation on the funding of research data management and infrastructure from Research Funding Organisations' (RFOs) and Research Performing Organisations' (RPOs) points of view in the first phase. This questionnaire aims at studying the landscape in European countries by exploring:

- What kind of funding strategies and funding models there are currently in place.
- What RFOs' and RPOs' policies with regard to research data management are.
- What the RFOs' and RPOs' views on the roles and responsibilities of different stakeholders are.
- Which funding models could be used as best practices.
- Whether there are gaps in the funding flows/models.
- What the RFOs' and RPOs' views on future developments are.

The outcomes of the survey will then serve as a background for possible further studies and in any case help raising awareness of the current challenges. The overall aim of this activity is to provide knowledge, eventually together with possible further studies, leading to a set of recommendations for policy-makers both on the national and European level.

Conclusion

- ❖ Need to combine different approaches :
 - ✓ top down, bottom up
 - ✓ specificities of scientific disciplines and technical issues
 - ✓ alone / partnership – France / international
- ❖ Key role for libraries :
 - ✓ policy definition, advocacy
 - ✓ development of new services



Thank you for your attention...

- odile.hologne@versailles.inra.fr
- esther.dzale@versailles.inra.fr