# What is the problem?

Wheat Data Interoperability

Chairs: Esther Dzalé Yeumo, Richard Fulss

Editor: Herman Stehouwer

Editors: Herman Stehouwer

The Wheat Data Interoperability group is working within the global context of a large societal challenge. I.e., given the following facts:

Interoperability of all wheat-related data

* Wheat is the most widely grown crop in the world.
* Wheat provides 20% of the world’s daily protein and calories.
* Wheat is the second most important crop in the developing world after rice.
* Wheat production has not satisfied demand in recent years.
* It is expected that by 2050 the demand for wheat will increase by 60%.

To respond to these facts, and to produce an adequate amount of wheat, the yield increase must go from 1% a year to 1.6% a year.

In order to tackle this issue many initiatives are doing experiments, which are producing a large quantity of heterogeneous data. There is a lack of harmonization in terms of data.

There is a strong need to make wheat-related data interoperable.

What are the goals?

The goals of the group are quite straightforward, to:

* Create a one stop shop for relevant information related to wheat data management 🡪 raise awareness, avoid duplicated efforts, foster adoption of common practices
* Facilitate the use of common data exchange formats 🡪 ease data sharing/submission to international repositories
* Foster a standardized description of datasets with consistent use of ontologies and metadata 🡪 increase the identification, the findability and the usability of the datasets



Note that the group did not start from scratch, the community has a large amount of assets which are used as a basis. The requirements for the work are based on the real needs of the wheat community.

What is the solution?

The needs of the wheat community are



addressed in three ways:

1. By building an interactive cookbook with recommendations and guidelines on data use
2. By filling a repository with wheat-related linked vocabularies
3. By building a number of interoperability prototypes and assessing the gain of interoperability through them

What is the impact?

The impact of this work is quite clear, i.e. improvement of management, sharing, discovery, reusability, and interoperability of data and datasets

within the wheat community.

Going forward the standardization and harmonization of wheat data will reduce variability and increase the relevance of wheat data related tools.

The outputs of this group are used as a building block in the Wheat Information System (Wheat IS) of the Wheat Initiative. Wheat IS builds the framework to establish a global wheat information system.

When can we use this?

* The guidelines produced by the group as well as the bioportal of wheat-related linked vocabularies (<http://wheat.agroportal.lirmm.fr/ontologies>) are directly usable now.

Following the guidelines and linking into existing vocabularies will give wheat-related data a larger relevance and impact going forward.

See also [datastandards.wheatis.org](http://datastandards.wheatis.org) for direct links to resources.

The URGI research unit at INRA will implement the recommendations about phenotype data and share some public datasets as RDF in a dedicated SPARQL endpoint. The targeted wheat dataset is the public INRA Breeding Network, available in GnpIS (<https://urgi.versailles.inra.fr/ephesis/ephesis/viewer.do#dataResults/trialSetIds=5,6,7>), a platform involved in the Elixir European project. For more information, please contact Cyril Pommier: cyril.pommier@versailles.inra.fr