Coordinating global research for wheat

Wheat IS EWG

Annual report and action plan

NAME OF EXPERT WORKING GROUP
Wheat Information System

LEADERSHIP & AUTHORSHIP	
Chair	Hadi Quesneville (INRA, France)
Vice-Chair(s)	Mario Caccamo (NIAB, UK), Dave Edwards (University of Western Australia, Australia), Gerard Lazo (USDA-ARS, USA)
Report Authors	Hadi Quesneville (INRA, France)

MEETINGS HELD (please attach minutes of these meetings using the template provided)						
Face-to-Face Meetings	Name	Date	Location	Duration	# EWG members expected	Estimated Cost (€)
	Annual WheatIS EWG meeting	9 Jan 2015	San Diego, USA	4 hours	17	0
	RDA-WI workshop	29-30 June 2015	Paris	2 d	16	9172,85
	EWG meeting	23 Sept 2015	Sydney			320
Other Meetings						
	Online	Sept 2016	Webex	2 hours	6	0
Priorities identified for 2016- 2017						

Short-term actions

- Enforce networking and collaboration among the nodes.
- Promote the dspace/iRODS platform to access and exchange data.
- Increase the data that will be searchable in the network of the WheatIS nodes

Medium-term actions

- *WheatIS provide integrated DBs*. Relevant data for the wheat community are integrated in a single place.

Expected outputs and deliverables for 2016 - 2017

2016

- Distributed search portal including more wheatIS nodes
- WheatIS nodes storage federation

2017

- Centralized access to integrated data in an information system

Timeline of activities for 2016 - 2017

AIMS OF THE EWG

List objectives/aims of the EWG as stated in the proposal

This expert group work together (i) to define data standards and data exchange protocols, and (ii) to develop a framework to support an integrated Wheat Information System. Our aim is to provide the international wheat research community with an easy access to wheat genetic, phenotype, and environmental information, genomic data and bioinformatics tools.

A central node, called WheatIS core, will provide a single entry point for the WheatIS users. The WheatIS core will be built upon the resources provided and shared by the nodes. It will provide access to data and information through a web portal. This portal (i) will give access to a data file repository storing files with their associated metadata, (ii) through a google-like search engine, will allow to find data available in the WheatIS core and its nodes using keywords (iii) several dedicated integrative databases, e.g. for genomic, genetic, and phenotype information, comparative genomics, and functional genomics, will be available. Analysis tools will be available for download from the web portal. Some WheatIS nodes will provide computing resources for data analysis.

2015 ACTIVITY REPORT

Reminder of **EWG action plan for the duration of the EWG, with flow-chart indicating timeline and outputs**

Our action plan.

Short-term actions

- *WheatIS nodes network.* Enforce networking and collaboration among the nodes.
- *Building a web platform to access and exchange data*. We focus on data standards and data file repository.
- *Integrated virtual portal*. A portal to access a network of DBs. Data will be searchable in a network of WheatIS nodes

Medium-term actions

- *WheatIS provide integrated DBs*. Relevant data for the wheat community are integrated in a single place.

Objectives identified for 2015

We identified as main objectives for 2015:

Data standards

Definition of standards for data sharing. Part of the activity is to follow and participate in international initiatives (e.g. transPLANT European project, crop and trait ontologies, Research Data Alliance, etc) and to propose standards when they are missing.

Objective: Provide a "cookbook" giving recommendation on the standards to use for sharing wheat data.

Distributed search

A search engine will be set up on the WheatIS web portal, allowing to dynamically search remote WheatIS nodes databases. Users will be able to connect to the WheatIS portal and type a keyword, or a term, that will be searched remotely in each node. Results will be provided as a brief summary of the matching data (e.g. identifier, name, short description) with links to access the remotely hosted data. This tool will allow the researchers to discover available data in the WheatIS core and its nodes, from simple keywords.

Objective: Search tools put in production.

Data file submission

The WheatIS will offer a space allowing data submission and the exchange of standardized data files with their associated metadata. The submission process will check for required metadata information and the correct use of data standards and ontologies, as defined by the EWG. This space will be also used for data exchange between WheatIS nodes or with wheat scientists. Data correctly formatted and respecting defined standards will be provided here promoting data interoperability.

Objective: Data submission tool put in production

WheatIS nodes identification

Identification of organisations offering wheat data to the community, willing to be considered as a WheatIS node and to contribute to the WheatIS effort.

Objective: Identification of new nodes with the associated wheat and bioinformatics resources to be contacted for inclusion as a node.

WheatIS portal

A web site giving access to WheatIS resources through links to the nodes.

Objective: A website in production.

Progress against aims in for 2015

Standards definition

A cookbook recommending standards to use for wheat data sharing was finished (section «data standards » at wheatis.org).

Distributed search

A first version is in production with a new interface (section «search» at wheatis.org). Not all WheatIS nodes are included for the moment, but still offer valuable search capabilities using faceting technologies.

A package installing locally an element of the data discovery tool (SolR server) is available to help nodes to expose their data through the system.

Data file submission

A data submission tool is available (section «submit data» at wheatis.org). Workflows for submitting phenotypes, genotypes and SNP discovery data are available.

WheatIS nodes identification

Two new nodes have been identified. *Genome Canada* and IGR-PAN, Institute of Plant Genetics (from the Polish Academy of Sciences) appear to be possible candidates.

WheatIS portal

A first version was put in production.

Outputs and deliverables in 2015
New chair and co-chairs election with the same chair and co-chairs elected? A "WheatIS contributor" status to recognize people who work for the WheatIS but not in the EWG, was created with a mailing list. A project proposal, called "*wheatual*", answering a H2020 European Call, was submitted but rejected. A letter of support was written for a TGAC project submitted to a BBSRC Call.
Contribution to Wheat Initiative objectives (http://www.wheatinitiative.org/about/objectives)
The EWG contribute to the "Core theme 5: Shared Methods, Resources and Capabilities". We work on the deployment of a "Wheat Information System" to provide the international wheat community with a single and comprehensive access point for wheat information and informatics tools. This action is considered as a top priority to be completed in the first 5 years.
Links established with other Wheat Initiative activities
Some nodes already established a close collaboration with the IWGSC EWG to manage and integrate the sequence data they provided. In particular the new version of wheat survey sequence have been integrated in different nodes databases and made available to the scientific community. The « Quality » EWG was contacted to organise jointly a two-day workshop focused on the naming of quality genes in the WheatIS databases. All EWGs have been contacted to give their feedbacks on data standards (section « data standards » at wheatis.org)
Additional activities
Contact for sharing experiences with other Plant information system was taken during a meeting at PAG: "Plant information system coordination meeting".

ACTION PLAN FOR NEXT TWO YEARS						
Meeting planned for 2016						
Face-to-Face Meetings	Name	Date	Location	Duration	# EWG members expected	Estimated Cost (€)
	Annual WheatIS EWG meeting	Jan, 2017	San Diego, USA	2 hours	20	1000
	Working group meeting with Quality EWG	2016 TBD	TBD	2 half-days	10	7500
	Working group meeting with another EWG	2016 TBD	TBD	2 half-days	10	7500
Other Meetings						
	Online	May, 2016	Webex	2 hours	10	0
Priorities identified for 2016- 2017						

Short-term actions

- Enforce networking and collaboration among the nodes.
- Promote the dspace/iRODS platform to access and exchange data.
- Increase the data that will be searchable in the network of the WheatIS nodes

Medium-term actions

- *WheatIS provide integrated DBs*. Relevant data for the wheat community are integrated in a single place.

Expected outputs and deliverables for 2016 - 2017

2016

- Distributed search portal including more wheatIS nodes
- WheatIS nodes storage federation

2017

- Centralized access to integrated data in an information system

Timeline of activities for 2016 - 2017

Meetings organised in 2013-2015

	2013	2014		2015	
	Meeting 1	Meeting 1	Meeting 2	Meeting 1	Meeting 2
Location	Versailles, France	San Diego, USA	Workshop on Wheat Data	San Diego, USA	Paris (RDA)
Date/duration	1-2/12/2013	13/01/2014	01-02/10/2014	09/01/2015	29-30/06/2015
# attendees	15	±20		16	16

Meetings planned in 2016

	Meeting 1
Location	San Diego
Date/duration	08 January 2016