HAL open science

AgroPortal: A vocabulary and ontology repository for agronomy

Clement Jonquet, Anne Toulet, Elizabeth Arnaud, Sophie Aubin, Esther Dzale Yeumo, Vincent Emonet, John Graybeal, Marie-Angélique Laporte, Mark A. Musen, Valeria Pesce, et al.

To cite this version:

Clement Jonquet, Anne Toulet, Elizabeth Arnaud, Sophie Aubin, Esther Dzale Yeumo, et al.. AgroPortal: A vocabulary and ontology repository for agronomy. Computers and Electronics in Agriculture, 2018, 144, pp.126-143. 10.1016/j.compag.2017.10.012 . lirmm-01679502

HAL Id: lirmm-01679502

https://hal-lirmm.ccsd.cnrs.fr/lirmm-01679502

Submitted on 17 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Computers and Electronics in Agriculture 144 (2018) 126–143

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

AgroPortal: A vocabulary and ontology repository for agronomy

Clément Jonquet[a,b,f,*], Anne Toulet[a,b], Elizabeth Arnaud[c], Sophie Aubin[d], Esther Dzalé Yeumo[d], Vincent Emonet[a], John Graybeal[f], Marie-Angélique Laporte[c], Mark A. Musen[f], Valeria Pesce[g], Pierre Larmande[b,e]

[a] Laboratory of Informatics, Robotics and Microelectronics of Montpellier (LIRMM), University of Montpellier & CNRS, France
[b] Computational Biology Institute (IBC) of Montpellier, France
[c] Bioversity International, Montpellier, France
[d] INRA Versailles, France
[e] UMR DIADE, IRD Montpellier, France
[f] Center for BioMedical Informatics Research (BMIR), Stanford University, USA
[g] Global Forum on Agricultural Research (GFAR), Food and Agriculture Organization (FAO) of the United Nations, Rome, Italy

ARTICLE INFO

Keywords:
Ontologies
Controlled vocabularies
Knowledge organization systems or artifacts
Ontology repository
Metadata
Mapping
Recommendation
Semantic annotation
Agronomy
Food
Plant sciences
Biodiversity

ABSTRACT

Many vocabularies and ontologies are produced to represent and annotate agronomic data. However, those ontologies are spread out, in different formats, of different size, with different structures and from overlapping domains. Therefore, there is need for a common platform to receive and host them, align them, and enabling their use in agro-informatics applications. By reusing the National Center for Biomedical Ontologies (NCBO) BioPortal technology, we have designed AgroPortal, an ontology repository for the agronomy domain. The AgroPortal project re-uses the biomedical domain's semantic tools and insights to serve agronomy, but also food, plant, and biodiversity sciences. We offer a portal that features ontology hosting, search, versioning, visualization, comment, and recommendation; enables semantic annotation; stores and exploits ontology alignments; and enables interoperation with the semantic web. The AgroPortal specifically satisfies requirements of the agronomy community in terms of ontology formats (e.g., SKOS vocabularies and trait dictionaries) and supported features (offering detailed metadata and advanced annotation capabilities). In this paper, we present our platform's content and features, including the additions to the original technology, as well as preliminary outputs of five driving agronomic use cases that participated in the design and orientation of the project to anchor it in the community. By building on the experience and existing technology acquired from the biomedical domain, we can present in AgroPortal a robust and feature-rich repository of great value for the agronomic domain.

1. Introduction

Agronomy, food, plant sciences, and biodiversity are complementary scientific disciplines that benefit from integrating the data they generate into meaningful information and interoperable knowledge. Undeniably, data integration and semantic interoperability enable new scientific discoveries through merging diverse datasets (Goble and Stevens, 2008). A key aspect in addressing semantic interoperability is the use of ontologies as a common and shared means to describe data, make them interoperable, and annotate them to build structured and formalized knowledge. Biomedicine has always been a leading domain encouraging semantic interoperability (Rubin et al., 2008). The domain has seen success stories such as the Gene Ontology (Ashburner et al., 2000), widely used to annotate genes and their products. And other disciplines have followed, developing among others the Plant Ontology (Cooper et al., 2012), Crop Ontology (Shrestha et al., 2010), Environment Ontology (Buttigieg et al., 2013), and more recently, the Agronomy Ontology (Devare et al., 2016), TOP Thesaurus (Garnier et al., 2017), Food Ontology (Griffiths et al., 2016), the IC-FOODS initiative's ontologies (Musker et al., 2016), and the animal traits ontology (Hughes et al., 2014). Ontologies have opened the space to various types of semantic applications (Meng, 2012; Walls et al., 2014), to data integration (Wang et al., 2015), and to decision support (Lousteau-Cazalet et al., 2016). Semantic interoperability has been identified as a key issue for agronomy, and the use of ontologies declared a way to address it (Lehmanna et al., 2012).

Communities engaged in agronomic research often need to access specific sets of ontologies for data annotation and integration. For instance, plant genomics produces a large quantity of data (annotated genomes), and ontologies are used to build databases to facilitate cross-

* Corresponding author at: 161 Rue Ada, 34090 Montpellier, France.
E-mail address: jonquet@lirmm.fr (C. Jonquet).

https://doi.org/10.1016/j.compag.2017.10.012
Received 27 October 2016; Received in revised form 12 May 2017; Accepted 13 October 2017
0168-1699/ © 2017 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

species comparisons (Jaiswal, 2011). More recently, the focus of many scientific challenges in plant breeding has switched from genetics to phenotyping, and standard traits/phenotypes vocabularies have become necessary to facilitate breeders' data integration and comparison. In parallel with very specific crop dictionaries (Shrestha et al., 2010), important organizations have produced large reference vocabularies such as Agrovoc (Food and Agriculture Organization) (Sachit Rajbhandari, 2012), the NAL Thesaurus (National Agricultural Library), and the CAB Thesaurus (Centre for Agricultural Bioscience International).[1] These thesauri are primarily used to index information resources and databases. As more vocabularies and ontologies[2] are produced in the domain, the greater the need to discover them, evaluate them, and manage their alignments (d'Aquin and Noy, 2012).

However, while great efforts have taken place in the biomedical domain to harmonize content (e.g., the *Unified Medical Language System* (UMLS), mostly for medical terminologies) (Bodenreider, 2004) and ontology design principles (e.g., the OBO Foundry, containing mostly biological and biomedical ontologies) (Smith et al., 2007), ontologies in agriculture are spread out around the web (or even unshared), in many different formats and artifact types, and with different structures. Agronomy (and its related domains such as food, plant sciences, and biodiversity) needs an one-stop shop, allowing users to identify and select ontologies for specific tasks, as well as offering generic services to exploit them in search, annotation or other scientific data management processes. The need is also for a community-oriented platform that will enable ontology developers and users to meet and discuss their respective opinions and wishes. This need was clearly expressed by stakeholders in various roles (developers, database maintainers, and researchers) across many community meetings, such as: 1st International Workshop for Semantics for Biodiversity in 2013 (http://semantic-biodiversity.mpl.ird.fr) (Larmande et al., 2013); the "Improving Semantics in Agriculture" workshop in 2015 (Baker et al., 2015); or several meetings of the Agricultural Data Interest Group (IGAD) of the Research Data Alliance.

These motivations prompted us to build a vocabulary and ontology repository to address these needs. In this paper, we present the AgroPortal project, a community effort started by the Montpellier scientific community to build an ontology repository for the agronomy domain. Our goal is to facilitate the adoption of metadata and semantics to facilitate open science in agronomy. By enabling straightforward use of agronomical ontologies, we let data managers and researchers focus on their tasks, without requiring them to deal with the complex engineering work needed for ontology management. AgroPortal offers a robust and reliable service to the community that provides ontology hosting, search, versioning, visualization, comment, and recommendation; enables semantic annotation; stores and exploits ontology alignments; and enables interoperation with the semantic web. Our vision is to facilitate the integrated use of all vocabularies and ontologies related to agriculture, regardless of their source, format, or content type.

In order to capitalize on what is already available in other communities, we have reused the openly available NCBO BioPortal technology (http://bioportal.bioontology.org) (Noy et al., 2009; Whetzel et al., 2011) to build our ontology repository and services platform. BioPortal was originally dedicated to health, biology and medicine and has some content related to agriculture, but the portal does covers few of the facets of agronomy, food, plant sciences and biodiversity, let alone environment and animal sciences. Therefore, many in the agronomy community do not see themselves as users targeted by BioPortal. For instance, the Crop Ontology is listed on the NCBO BioPortal (along with other top-level plant-related ontologies), but is not currently fully accessible and described through this portal; none of the crop specific ontologies are available. In addition to its core repository of ontology mission, the NCBO technology also offers many applicable tools, including a mapping repository, an annotator, an ontology recommender, community support features, and an index of annotated data. All these services are reused and customized within AgroPortal to benefit its target user community.[3] Furthermore, our vision was to adopt, as the NCBO did, an open and generic approach where users can easily participate to the platform, upload content, and comment on others' content (ontologies, concepts, mappings, and projects). As explained below, we determined that the NCBO technology (Whetzel and Team, 2013) implemented the greatest number of our required features, while recognizing the technical challenges of adopting such a various and complex software.

In the following sections, we offer extensive descriptions of AgroPortal's features. We will focus on how they address community requirements expressed within five agronomic driving use cases involving important research organizations in agriculture such as Bioversity International (CGIAR), French INRA, and United Nations FAO. The rest of the paper is organized as follows: In Section 2, we review related work in ontology repositories in relation to our domain of interest. Section 3 describes the requirements of AgroPortal's initial five driving agronomic use cases. Section 4 presents our platform by extensively describing its content, as well as its features (both inherited from the NCBO BioPortal, and added by us). Section 5 analyzes how our initial five driving use case results benefit from AgroPortal. Finally, Section 6 provides a discussion of the contributions of AgroPortal, and Section 7 presents our conclusions.

2. Background and related work

With the growing number of developed ontologies, ontology libraries and repositories have been of interest in the semantic web community. Ding and Fensel (2001) presented in 2001 a review of ontology libraries that introduced the notion of "library." Then Hartman et al. Baclawski and Schneider (2009) introduced the concept of ontology repository, with advanced features such as search, metadata management, visualization, personalization, and mappings. By the end of the 2000′s, the Open Ontology Repository Initiative (Baclawski and Schneider, 2009) was a collaborative effort to develop a federated infrastructure of ontology repositories.[4] d'Aquin and Noy (2012) provided the latest review of ontology repositories in 2012.

In the biomedical or agronomic domains there are several standards or knowledge organization systems libraries (or registries) such as FAIRSharing (http://fairsharing.org) Sansone et al., 2012, the FAO's VEST Registry (http://aims.fao.org/vest-registry), and the agINFRA linked data vocabularies (vocabularies.aginfra.eu) (Pesce et al., 2013). They usually register ontologies and provide a few metadata attributes about them. However, because they are registries not focused on vocabularies and ontologies, they do not support the level of features that an ontology repository offers. In the biomedical domain, the OBO Foundry (Smith et al., 2007) is a reference community effort to help the

[1] http://aims.fao.org/agrovoc, https://agclass.nal.usda.gov and http://www.cabi.org/cabthesaurus

[2] In this paper, we often use the word "ontologies" or "vocabularies and ontologies" to include ontologies, vocabularies, terminologies, taxonomies and dictionaries. We acknowledge the differences (not discussed here) in all these types of Knowledge Organization Systems (KOS) or knowledge artifacts. The reader may refer to McGuinness's discussion (McGuinness, 2003). While being an "ontology repository", AgroPortal handles all these artifact types, if they are compatibly formatted. While AgroPortal thereby enables horizontal use of these artifact types with common user interface and application programming interface, it does not leverage the full power of ontologies (e.g., reasoning), instead map all the imported artifact types to a "common simplified model."

[3] Except the "NCBO Resource Index" component, a database of 50+ biomedical resources indexed with ontology concepts (Jonquet et al., 2011) that we have not reused in AgroPortal because we work with the AgroLD use case to fulfill the mission of interconnecting ontologies and data.

[4] At that time, the effort already reused the NCBO technology that was open source, but not yet packaged in an appliance as it is today.

biomedical and biological communities build their ontologies with an enforcement of design and reuse principles that have made the effort very successful. The OBO Foundry web application is not an ontology repository per se, but relies on other applications that pull their data from the foundry, such as the NCBO BioPortal (Noy et al., 2009), OntoBee (Xiang et al., 2011), the EBI Ontology Lookup Service (Côté et al., 2006) and more recently AberOWL (Hoehndorf et al., 2015). In addition, there exist other ontology libraries and repository efforts unrelated to biomedicine, such as the Linked Open Vocabularies (Vandenbussche et al., 2014), OntoHub (Till et al., 2014), and the Marine Metadata Initiative's Ontology Registry and Repository (Graybeal et al., 2012).

Some of the known ontology repositories could be candidates for hosting agronomical ontologies. However, all of these portals either are too generic, or too narrowly focused on health, biology or medicine, and despite any existing thematic overlaps, scientific lineage and partnerships, we have identified, as established in Section 1, the crucial need for a community platform where agronomy will actually be the primary focus. To avoid building a new ontology repository from scratch, we have considered which of the previous technologies are reusable. While all of them are open source, only the NCBO BioPortal[5] and OLS[6] are really meant for reuse, both in their construction, and in their provided documentation. At the start of our project in 2014, AberOWL was not yet published and OntoBee (released in 2011) had not changed between 2011 and 2014 (a new release took place thereafter (Ong et al., 2016). Of the two candidate technologies at the time, we will show, that the NCBO technology was the one implementing highest number of requested features.[7]

In the biomedical domain, the NCBO BioPortal is a well-known open repository for biomedical ontologies originally spread out over the web and in different formats. There are 656 public ontologies in this collection as of Nov. 2017, including relevant ones for agronomy. By using the portal's features, users can browse, search, visualize and comment on ontologies both interactively through a user web interface, and programmatically via web services. Within BioPortal, ontologies are used to develop an annotation workflow (Jonquet et al., 2009) that indexes several biomedical text and data resources using the knowledge formalized in ontologies to provide semantic search features that enhance information retrieval experience (Jonquet et al., 2011). The NCBO BioPortal functionalities have been progressively extended in the last 12 years, and the platform has adopted semantic web technologies (e.g., ontologies, mappings, metadata, notes, and projects are stored in an RDF[8] triple store) (Salvadores et al., 2013).

An important aspect is that NCBO technology (Whetzel and Team, 2013) is domain-independent and open source. A BioPortal virtual appliance[9] is available as a server machine embedding the complete code and deployment environment, allowing anyone to set up a local ontology repository and customize it. It is important to note that the NCBO Virtual Appliance has been quite regularly reused by organizations which needed to use services like the NCBO Annotator but, for privacy reason, had to process the data in house. Via the Virtual Appliance, NCBO technology has already been adopted for different ontology repositories in related domains and was also chosen as foundational software of the Open Ontology Repository Initiative (Baclawski and Schneider, 2009). The Marine Metadata Interoperability Ontology Registry and Repository (Rueda et al., 2009) used it as its backend storage system for over 10 years, and the Earth Sciences Information Partnership earth and environmental semantic portal (Pouchard and Huhns, 2012) was deployed several years ago. More recently, the SIFR BioPortal (Jonquet et al., 2016) prototype was created at University of Montpellier to build a French Annotator and experiment multilingual issues in BioPortal (Jonquet et al., 2015). Although we cannot know all the applications of other technologies, the visibly frequent reuse of the NCBO technology definitively confirmed it was our best candidate. There are two other major motivations for AgroPortal to reuse the products of biomedicine: (i) to avoid re-developing tools that have already been designed and extensively used and contribute to long term support of the commonly used technology; and (ii) to offer the same tools, services and formats to both communities, to facilitate the interface and interaction between their domains. This alignment will enhance both technical reuse (for example, enabling queries to either system with the same code), and semantic reuse (knowing the same semantic capabilities and practices apply to both sets of ontologies).

More specifically to the plant domain, the Crop Ontology web application (www.cropontology.org) (Matteis et al., 2013) publishes online sets of ontologies and dictionaries required for describing crop germplasm, traits and evaluation trials. As of Nov. 2017, it contains 28 crop-specific phenotype and trait ontologies, in addition to ontologies related to the crop germplasm domain. Besides its role as a repository, the Crop Ontology web application offers community-oriented features such as an CSV template (TDv5) for trait submission, and addition and filtering of new terms. A web Application Programming Interface (API) provides all necessary services to third party users like the Global Evaluation Trials Database, currently storing 35,000 trial records. Efforts have been made to structure and formalize the crop-specific ontologies following semantic web standards (using the Web Ontology Language (OWL)), as well as offering collaborative ontology enrichment and annotation features. The current Crop Ontology web application facilitates the ontology-engineering life cycle (Noy et al., 2010), starting with collaborative construction, publishing, use and modification. However, it would require important improvements such as: versioning, community features, multilingual aspects, visualization, data annotation, and mapping services. For instance, it is important to support the alignment (or mapping) of terms within and across different ontologies both within the Crop Ontology itself (in different crop branch) and with other top level ontologies commonly used in plant biology, like the Plant Ontology, Plant Trait Ontology, Plant Environment Ontology, Plant Stress Ontology all maintained and extended within the Planteome project (Jaiswal et al., 2016).

The Planteome platform (www.planteome.org) is reusing the Gene Ontology project AmiGO technology (Carbon et al., 2009) to build a database of searchable and browsable annotations for plant traits, phenotypes, diseases, genomes, gene expression data across a wide range of plant species. The project focuses on developing reference ontologies for plant and on integrating annotated data within the platform. Their objective is slightly different than AgroPortal's objective, and the scope is not as large as the one we envision for AgroPortal.

3. Driving agronomic use cases requirements

The AgroPortal project was originally driven by five agronomic use cases that were the principal sources of ontologies and vocabularies. In this section, we present their requirements in terms of ontology repository functionalities – summarized in Table 1. The results for each use case will be presented in Section 5.

[5] The technology has always been open source, and the appliance has been made available since 2011. However, the product became concretely and easily reusable after BioPortal v4.0 end of 2013.

[6] The technology has always been open source but some significant changes (e.g., the parsing of OWL) facilitating the reuse of the technology for other portals were done with OLS 3.0 released in December 2015.

[7] It is beyond the scope of this paper to draw a complete comparison of ontology portals. The reader may refer to d'Aquin and Noy (2012).

[8] The Resource Description Framework (RDF) is the W3C language to described data. It is the backbone of the semantic web. SPARQL is the corresponding query language. By adopting RDF as the underlying format, AgroPortal can easily make its data available as linked open data and queryable through a public SPARQL endpoint. To illustrate this, the reader may consult the Link Open Data cloud diagram (http://lod-cloud.net) that since 2017 includes ontologies imported from the NCBO BioPortal (most of the Life Sciences section).

[9] www.bioontology.org/wiki/index.php/Category:NCBO_Virtual_Appliance

Table 1
Summary of agronomic use case requirements for AgroPortal.

#	Requirement	Use case	Example
1	One-stop-shop to store, browse, search, visualize agronomical ontologies	LovInra	Facilitate the adoption of semantic web standards by INRA' scientist, with a focus on agriculture
		VEST	The registry targets specifically the agriculture community and requires content-based services. The organization of ontologies by group and categories is also necessary
2	Unique ontology access point and application programming interface (API) to ontologies	AgroLD	Automatically retrieve the most recent version of ontologies currently hosted either on OBO Foundry or Cropontology.org. At the beginning of the project, a SPARQL endpoint for ontologies was also needed
		VEST	Access point to automatically obtain metadata about all the ontologies
3	Directly accessible to scientists to upload their ontologies or vocabularies	LovInra, VEST	INRA's researchers and VEST users need to upload their resources to a platform themselves
4	Ontology-based annotation service	AgroLD	Annotate text data from database fields to create RDF triples
		LovInra, Crop Ontology	Identify plant phenotypes in text descriptions
5	Handle different level of semantic description and the corresponding standard formats (SKOS and OWL)	LovInra	INRA's develop different type of knowledge organization systems include: ontologies (AFEO, Biorefinery, OntoBiotope) but also thesauri (AnAEE, GACS)
		VEST	Many resources in agronomy are in SKOS format.
6	Store and retrieve mappings between ontologies	ALL	All use cases have expressed the need to have a place to store, describe and retrieve alignments
7	Store mappings between ontologies and external resources	AgroLD	Publish AgroLD mapping annotations to reference ontologies such as SIO, EDAM, PO
		Others	Reference thesauri like Agrovoc have adopted linked open data practices and offer mappings to multiple semantic web resources (not necessarily ontologies)
8	Automatically generate mappings between ontologies	ALL	All use cases have expressed the need to automatically align ontologies one another
9	Query and search annotated data from ontologies	AgroLD	Identify AgroLD data elements when browsing ontologies in AgroPortal.
10	Offer a unique sub-endpoint specific to a community or group	WDI	Visualize and use only the 22 vocabularies identified by the WDI working group
		LovInra	Clearly identify resources (co-)developed by INRA's researchers
		Crop Ontology	Handle as a collection the Crop Ontology project, which is composed of multiple crop-specific trait ontologies. Possible alternative to cropontology.org
11	Provide rich metadata description for ontologies (using semantic web standards)	WDI	Clearly describe access rights and license information for ontologies
		LovInra	Clearly describe the type of resources (ontology, thesaurus, vocabulary, etc.) and their format and syntax
		VEST	Facilitate an automatic interconnection with VEST, including aligning the metadata fields
12	Get community feedback	WDI	Inform the community about the WDI guidelines and get their feedback on the selected ontologies
		Crop Ontology	Offer breeders a way to suggest new trait and comment existing ones
		VEST	Enable a large community of "standard" developers to provide feedback and comments on the use (or non-use) of ontologies and vocabularies in AgroPortal
13	Multilingual ontology support	VEST, Others	Increasingly vocabularies have labels in different languages (e.g., Agrovoc, GACS, NALt). Distinguish between these labels in lexical-based services (search, annotation)
		Others	IRSTEA develops vocabularies only in French
14	Dereference URIs for ontologies	LovInra, Crop Ontology	When opening in a web browser a URI created by INRA or CO, display the corresponding class or property page
15	Mechanism to identify and select the relevant ontologies for a given task	LovInra, VEST	Facilitate the identification of relevant agronomical ontologies for non-experts
16	Enable private access to ontologies during working and/or development phases	LovInra	Access and test the AnAEE Thesaurus or GCAS before they release; work on certain versions of OntoBiotope not public in OpenMinted project
17	Export ontologies in different formats, including downgrading them to CSV	Crop Ontology	Breeders may need simpler formats, as they may not be able to use advanced semantic web formats
18	Store the project/ontology relationships	VEST, AgBioData	Select and maintain a list of ontologies used by model organism databases

3.1. Agronomic Linked Data (AgroLD)

Agronomic research aims to effectively improve crop production through sustainable methods. To this end, there is an urgent need to integrate data at different scales (e.g., genomics, proteomics and phenomics). However, available agronomical information is highly distributed and diverse. Semantic web technology offers a remedy to the fragmentation of potentially useful information on the web by improving data integration and machine interoperability (Schmachtenberg et al., 2014). This has been often illustrated in data integration and knowledge management in the biomedical domain (Belleau et al., 2008; Jonquet et al., 2011; Jupp et al., 2014; Groth et al., 2014). To further build on this line of research in agronomy, we have developed the Agronomic Linked Data knowledge base (www.agrold.org) (Venkatesan et al., 2015). Launched in May 2015, it serves as a platform to consolidate distributed information and facilitate formulation of research hypotheses. AgroLD offers information on genes, proteins, Gene Ontology Associations, homology predictions, metabolic pathways, plant traits, and germplasm, on the following species: rice, wheat, arabidopsis, sorghum and maize. We provide integrated agronomic data, as well as the infrastructure to aid domain experts answering relevant biological questions (for example, "identify wheat proteins that are involved in root development"). AgroLD relies on RDF and SPARQL technologies for information modelling and retrieval, and uses OpenLink Virtuoso (version 7.1) triple store. Database contents were parsed and converted into RDF using a semi-automated pipeline implemented in Python (https://github.com/SouthGreenPlatform/AgroLD).

The conceptual framework for knowledge in AgroLD is based on well-established ontologies in plant sciences such as Gene Ontology, Sequence Ontology, Plant Ontology, Crop Ontology and Plant Environment Ontology. AgroLD needs a dedicated application programming interface to these ontologies, as well as a means to annotate database fields (header and values) with ontology concepts. In addition, it requires a system to store mappings annotations between key entities in the AgroLD knowledge base and reference ontologies. In the long-term vision for AgroPortal and AgroLD, the former might be an entry point to the knowledge stored in AgroLD, enabling users to easily query and locate data annotated with ontologies.

3.2. RDA Wheat Data Interoperability (WDI) working group

Wheat is a major source of calories and protein, especially for consumers in developing countries, and thus plays an important socio-economical role. The International Wheat Initiative (www.wheatinitiative.org) has identified easy access and interoperability of all wheat related data as a top priority, to make the best possible use of genetic, genomic and phenotypic data in fundamental and applied wheat science. For example, the identification of causative genes for an important agronomic trait is key to effective marker-assisted breeding and reverse genetics. It requires integrating information from many different sources such as gene function annotations, biochemical pathways, gene expression data, as well as comparative information from related organisms, gene knock-out and the scientific literature (Hassani-Pak et al., 2013). However, the disparate nature of the formats and vocabularies used to represent and describe the data has resulted in a lack of interoperability.

The Wheat Data Interoperability working group was created in March 2014 within the frame of the Research Data Alliance (https://rd-alliance.org) and under the umbrella of the International Wheat Initiative, in order to provide a common framework for describing, linking and publishing wheat data with respect to existing open standards. The working group conducted a survey to identify and describe the most relevant vocabularies and ontologies for data description and annotation in the wheat domain (Dzalé-Yeumo et al., 2017). For some data types like DNA sequence variations, genome annotations, and gene expressions, the survey showed good consensus regarding data exchange formats. However, the survey did not show good consensus about data exchange formats and data description practices for phenotypes and germplasm, suggesting the need for harmonization and standardization.

Finally, this group identified 22 relevant vocabularies and ontologies for which, beyond the consensus issue, other problems were identified: (i) format and location heterogeneity: ontology formats included OBO format, OWL, and even SKOS (or SKOS-XL); (ii) heterogeneity: these ontology coverages ranged from describing generic experimental crop study (e.g., Crop Research), to narrow wheat-related topics (Wheat Trait, Wheat Anatomy and Development), to top-level concepts in biomedicine (BioTop). The need to offer a dedicated repository of linked vocabularies and ontologies relevant for wheat having been identified, the NCBO technology was seen as a likely tool to address this needs and desired features.

3.3. INRA Linked Open Vocabularies (LovInra)

What does a specialist in cattle developmental biology really need to easily identify, evaluate and exploit a few potential vocabularies of interest? Whether familiar with semantics technology or not, she needs a place that reflects her scientific environment and community, where those with similar concerns can share comments and content. As an example, INRA develops models to predict feed efficiency and meat quality for beef production, using experimental data collected during decades at INRA and externally. To meet the challenge of data integration, INRA developed the Animal Trait Ontology for Livestock (ATOL). In part thanks to AgroPortal, ATOL developers have identified the Animal Disease Ontology (ADO), developed by another team at INRA, as a possible resource to expand the perimeter of actionable data. This raised the question: How many complementary or competing resources to ATOL exist?

With this vision in mind, LovInra is a service offered by the French National Institute for Agricultural Research (INRA) Scientific and Technical Information department to identify and evaluate knowledge organization sources produced by INRA's scientists, so that the agricultural community and possibly a larger public can benefit from them. Many such resources developed within specific projects remain unknown to the research community despite their value. They are often developed by subject matter experts who are not semantic experts, and who often do not have the resources (knowledge, time, or money) to share their results. Further, they span multiple semantic levels, from simple lexical descriptions, to hierarchies, to complex semantic relations. To achieve this goal, the vocabularies must be published with respect to open standards and linked to other existing resources. INRA adopted the semantic web's practices and standards (RDF, SKOS, OWL, SPARQL) to enable the methodological and technical practices needed by INRA's scientists to standardize, document and publish the vocabularies created in their projects. Examples of INRA's projects developing vocabularies or ontologies includes: (i) the AnAAE Thesaurus for the semantic description of the study of continental ecosystems developed by the AnaEE-France infrastructure;[10] (ii) the OntoBiotope ontology of microorganism habitats used collaboratively in multiple projects such as OpenMinted as well as for the BioNLP shared tasks; (iii) the Agri-Food Experiment Ontology (AFEO) ontology network which cover various viticultural practices, and winemaking products and operations.

Beyond its evaluation and standardization role, LovInra also serves to assign, deference, and provide programmatic access to INRA URIs (for example, http://opendata.inra.fr/ms2o/Observation), using its triple store and web interface (http://lovinra.inra.fr). Although the current service, which includes description of resource metadata and direct access to source files, is necessary for internal use, it does not meet external dissemination objectives. In addition, the LovInra registry does not support any content-based features, such as searching, browsing, visualizing, mappings and annotation. We see AgroPortal as a possible solution to the entire range of INRA's unmet semantic needs above, complementing the services already provided by LovInra.

3.4. The Crop Ontology project

Communities engaged in germplasm evaluation trials need to access specific sets of ontologies for plant data annotation and integration. The Crop Ontology project (www.cropontology.org) (Shrestha et al., 2010) of the Integrated Breeding Platform (IBP) is AgroPortal's fourth use case. The main goals of this project are: (i) to publish online fully documented lists of breeding traits and standard variables used for producing standard field books and (ii) to support data analysis and integration of genetic and phenotypic data through harmonized breeders' data annotation (Shrestha et al., 2012). Crop breeders, data managers, modelers, and computer scientists created a community of practice to discuss their variables, methods and scales of measurement, and field books. They seek to develop the most complete crop-specific trait ontologies according to the Crop Ontology template and guidelines.

The Crop Ontology website, released in 2010, provides 28 crop-specific trait ontologies, in addition to ontologies describing germplasm material and evaluation trials. The website publishes each crop-specific trait ontology online, making it available for download from the user interface or through an API in various formats: CSV, OBO, RDF/SKOS. Partners like the Oat Global, the US Department of Agriculture (USDA), INRA and the Polish Genomic Network have uploaded ontologies.[11] The project requires a specific dedicated infrastructure that deals with the adopted multi-trait ontologies approach, and supports search and versioning of ontologies. Plus, the Crop Ontology breeders need an interface to suggest new crop traits (i.e., new terms in the trait ontologies) and simple formats (such as CSV) to export the "trait dictionary" locally.

[10] Analysis and Experimentation on Ecosystems is European research infrastructure dedicated to the experimental manipulation of managed and unmanaged terrestrial and aquatic ecosystems (www.anaee.com).

[11] In addition, the Crop Ontology is used by several third-party projects like the Next Generation Breeding (Nextgen) databases, the Integrated Breeding Platform's breeding management system, and the global repository of the Agricultural Trials or EU-SOL.

3.5. GODAN Map of Agri-Food Data Standards

Recently, a new project under the umbrella of the GODAN[12] initiative called *GODAN Action* identified as one of its outputs a global map of standards used for exchanging data in the field of food and agriculture. To avoid duplicating effort, and to reuse previous community work, the project reviewed possible sources of standards that could be integrated. Two existing suitable platforms were identified: the FAO Agricultural Information Management Standards VEST Registry (http://aims.fao.org/vest-registry – now merged inside the new Map of Standards presented Section 5.5) and the then-new AgroPortal project.

The VEST Registry, created by FAO in 2011, was a metadata catalog of around 200 knowledge organization sources and tools. It had a broader coverage than the AgroPortal in two facets, knowledge types and domains. (i) Types of vocabularies or standards covered: the VEST Registry covered all types of knowledge artifacts, not just vocabularies or ontologies formally defined in RDFS, OWL, SKOS, or OBO. For instance, the VEST registry would cover data exchange format specification defined in XML or text description. (ii) Domain coverage: Besides standards used specifically for food and agriculture data, the directory included resources used in neighboring disciplines (like climate and environment, sciences). The VEST Registry was conceived as a metadata catalog, providing descriptions and categorization of standards and linking to the original website or download of the standard, but it did not exploit the content of the vocabularies or ontologies, only their metadata descriptions. It did not support any alignment between the sources either. To interconnect the VEST and AgroPortal, rich and unambiguous metadata would be crucial, as well as good classification of resources per categories and types.

3.6. Other requirements identified

In addition to these five first driving use cases, other projects or organizations have identified AgroPortal as a relevant application to host, share and serve their ontologies:

> IRSTEA's projects, such as the French Crop Usage thesaurus about crops cultivated in France, and the French Agroecology Knowledge Management ontology for design innovative crop systems. These two projects produce ontologies only in French and needed a host for their work.
> The Agrovoc thesaurus (Sachit Rajbhandari, 2012), which is the most worldwide used multilingual vocabulary developed by FAO. Agrovoc contains more than 32 K concepts covering topics related to food, nutrition, agriculture, fisheries, forestry, environment and other related domains. Agrovoc Linked Open Data version contains multiple mappings to other vocabularies or resources that a resource hosting Agrovoc must incorporate.
> The Consortium of Agricultural Biological Databases (www.agbiodata.org), a group of database developers and curators maintaining model organism databases. The group wants to identify which databases use which ontologies, and recommend a list of ontologies based on that information.

4. A portal for agronomic related ontologies

In 2014, the *Computational Biology Institute of Montpellier* project identified the need for an ontology-based annotation service for the AgroLD and Crop Ontology use cases above. This large bioinformatics project in France had a specific plant/agronomy data work package. In parallel, we started reusing NCBO technology (Whetzel and Team, 2013) in the context of the SIFR[13] project, in which we develop a French version of the Annotator (Jonquet et al., 2016). We then implemented a connector to BioPortal within WebSmatch (an open environment for matching complex schemas from many heterogeneous data sources (Coletta et al., 2012) enabling calls either to the NCBO Annotator web service, or any other NCBO-based Annotator (Castanier et al., 2014). Once we had a portal prototype hosting a few specific ontologies, interest in it grew when we presented it to several interlocutors (for examples, Bioversity International, INRA, IRD, CIRAD, FAO, RDA, Planteome). Driven additionally by the other use cases presented in Section 3, we extended our reuse of the NCBO technology to the full stack, and publish it under the brand AgroPortal.

We now have an advanced prototype platform (illustrated in figures on following pages) whose latest version v1.4 was released in July 2017 at http://agroportal.lirmm.fr.[14] The platform currently hosts 77 ontologies (Table 2), with more than 2/3 of them not present in any similar ontology repository (like NCBO BioPortal), and 11 private ontologies. We have identified 93 other candidate ontologies (Table 3) and we work daily to import new ones while involving/informing the original ontology developers. The platform already has more than 90 registered users. For an overview of AgroPortal ontology analytics, see Fig. 5 (Annex).

4.1. Ontology organization and sources

Developers generally upload their ontologies when they think the ontologies have reached a sufficient maturity and relevance to make them publicly available. Sometime, like in the AnaEE thesaurus, or OntoBiotope, developers use/used the portal as a staging location before the ontology goes public. If the initiative comes from our side, we usually always interact with the developers before importing any new resources: the original ontology developers always stay the only authority for the ontologies in the portal. Because of the features offered by AgroPortal (Sections 4.2 and 4.3), we think it is reasonable to incorporate ontologies that are already listed on other platforms (OBO Foundry, FAIRSharing, VEST registry, or LovInra). However, in those cases we follow these practices:

> Developers can configure the entry in AgroPortal to automatically pull new version of ontologies. We synchronize the ontology in AgroPortal with the one at the original location via a nightly update[15] so the latest version is always available. For instance, all the ontologies in the OBO-FOUNDRY group are systematically updated using their PURL (e.g., for the Plant Ontology: http://purl.obolibrary.org/obo/po.owl).
> We always inform the ontology developers of their ontology publication on AgroPortal if they did not submit their ontology directly, and offer them to claim administration role on the ontology if desired. While we often edit ontology descriptions, we ask the ontology developers to validate our edits and complete them.
> We try to avoid duplicating ontologies already hosted in the NCBO BioPortal, unless required by a specific use case. Of course, overlap exists between our domain of interest and biomedicine. Our general approach is to let ontology developers decide if their ontology should be incorporated in the AgroPortal while it is already in the NCBO BioPortal. The long-term vision for AgroPortal and BioPortal is an interconnected network of "bioportals" that will enable easy access to ontologies for anyone independently from where they are hosted and that could extend to ontology repository types beyond the NCBO technology.

[12] Global Open Data for Agriculture and Nutrition: http://www.godan.info.

[13] Semantic Indexing of French Biomedical Data Resources (SIFR) project - http://www.lirmm.fr/sifr.

[14] https://github.com/agroportal/documentation/wiki/Release-notes

[15] Except for three ontologies (GO, BIOREFINERY & TRANSMAT) that are updated only weekly for scalability reasons.

Table 2
Examples of ontologies uploaded in AgroPortal. Acronyms in parenthesis are the identifier on AgroPortal e.g., http://agroportal.lirmm.fr/ontologies/AEO has the acronym AEO (Size = approximate number of classes or concepts).

Title	Format	Source	Group	Size
IBP rice trait ontology (CO_320)	OWL	cropontology.org	CROP, AGBIODATA, AGROLD	~2K
IBP wheat trait ontology (CO_321)	OWL	cropontology.org	CROP, AGBIODATA, AGROLD, WHEAT	~1K
IBP wheat anatomy & development ontology (CO_121)	OBO	cropontology.org	CROP, WHEAT	~80
IBP crop research (CO_715)	OBO	cropontology.org	CROP, AGBIODATA, WHEAT	~250
Multi-crop passport ontology (CO_020)	OBO	cropontology.org	CROP	~90
Biorefinery (BIOREFINERY)	OWL	Inra	LOVINRA, WHEAT, AGBIODATA	~300
Matter transfer(TRANSMAT)	OWL	Inra	LOVINRA, WHEAT, AGBIODATA	~1.1 K
Plant ontology (PO)	OWL	OBO Foundry	OBOF, AGROLD, WHEAT, AGBIODATA	~2K
Plant trait ontology (TO)	OWL	OBO Foundry	OBOF, AGROLD, WHEAT, AGBIODATA	~4.4 K
Durum wheat (DURUM_WHEAT)	OWL	Inra	LOVINRA	~130
Agricultural experiments (AEO)	OWL	Inra	LOVINRA	~60
Environment ontology (ENVO)	OWL	OBO Foundry	WHEAT, OBOF	~6.3 K
NCBI organismal classification (NCBITAXON)	RRF	UMLS	WHEAT, AGROLD	~900 K
AnaEE thesaurus (ANAEETHES)	SKOS	Inra	LOVINRA	~3.3 K
French crop usage (CROPUSAGE)	SKOS	Irstea	None	~300
Agrovoc (AGROVOC)	SKOS	FAO (UN)	WHEAT, AGBIODATA	~32 K
Food ontology (FOODON)	OWL	OBO Foundry	OBOF	~10 K
National agricultural library thesaurus (NALT)	SKOS	NAL (USDA)	WHEAT, AGBIODATA	~67 K
Global agricultural concept scheme (GACS)	SKOS	FAO-NAL-CABI	None	~580 K
Agronomy ontology	OWL	CGIAR	OBOF	~430
Biological collections ontology	OWL	OBO Foundry	OBOF	~160
Flora phenotype ontology	OWL	AberOWL	None	~28 K

Table 3
Selection of candidate ontologies of interest for the agronomic community, not present in the NCBO BioPortal.

Title	Organization or source
CAB thesaurus	CABI
Chinese agricultural thesaurus	CAAS
Wine ontology	INRA
Oat, Barley, Brachiaria, Potato (etc.) trait ontologies	Crop Ontology
Plant disease ontology	INRA
Agriculture activity ontology	CAVOC
Agriculture and forestry ontology	Univ. of Helsinki
IC-FOODS ontologies (~10)	UC Davis
agINFRA soil vocabulary	FAO, GFAR
Plant-pathogen interactions ontology	CBGP
Plant phenology ontology	OBO Foundry
Thesaurus of plant characteristics	CEFE
Livestock product trait ontology	Iowa State Univ.
Livestock breed ontology	Iowa State Univ.

Within AgroPortal, each time an ontology is uploaded into the portal, it is assigned a group and/or category. Groups associate ontologies from the same project or organization, for better identification of the provenance. We have created a group for each use case, except the fifth one that is not a source of ontologies, and another one for the OBO Foundry. For each group we have deployed a specific slice (a restriction of the user interface to a specific group of ontologies) as explained later. Categories indicate the topic(s) of the ontology, providing another way to classify ontologies in the portal independently from their groups or provenance. As of now we have defined 20 general categories such as Farms and Farming Systems, Plant Phenotypes and Traits, Plant Anatomy and Development, Agricultural Research, and Technology and Engineering. These categories were established in cooperation with FAO Agricultural Information Management Standards (AIMS), which has maintained the VEST Registry since 2011.

Groups and categories, along with other metadata, can be used on the "Browse" page of AgroPortal to filter out the list of ontologies (cf. Fig. 3). Of course, groups and categories are customizable, and will be adapted in the future to reflect the evolution of the portal's content and community feedback. The portal's architecture provides URIs for any portal objects, including groups and categories. For example, the URI http://data.agroportal.lirmm.fr/categories/FARMING identifies the group "Farms and Farming Systems." External applications can use those URIs to organize ontologies or tag them.

4.2. Features from AgroPortal inherited from the NCBO BioPortal

The main features offered by the NCBO BioPortal are described in Noy et al. (2009), Whetzel et al. (2011). They include:[16]

Ontology library. The core mission of the AgroPortal is to serve as a one-stop shop for ontology descriptions and files. The portal also allows users to specify the list of ontologies that shall be displayed in their user interface when logged-in. While not replacing source code repository such as for instance GitHub, highly used by the community, the portal stores all ontology versions as they are submitted or automatically pulled, and can display their metadata and differences from one version to the next, although only the latest ontologies are referenced for queries. Ontologies can either be harvested from specified locations, or directly uploaded by users. Ontologies are semantically described (cf. metadata), and a browsing user interface allows to quickly identify, with faceted search, the ontologies of interest based on their descriptions and metadata.

Search across all the ontologies. AgroPortal search service indexes the ontology content (classes, properties and values) with Lucene, and offers an endpoint to search across the ontologies by keyword or identifier. For example, a keyword search on "abiotic factor"[17] will identify the occurrence of this term (or similar terms if none match exactly) in all the ontologies of the portal, and sort the results by relevance to the query and ontology popularity in the portal (number of views) (Noy et al., 2013). For the above search, the first three results are Abiotic factor (CO_715_0000078), Abiotic stress (CO_320:Abiotic_stress), and abiotic stress trait (TO_0000168).

Ontology browsing and content visualization. The ontology 'classes' and 'properties' tab lets users visualize a class or property within is hierarchy, as well as see the related content (labels, definition, mappings, any other relations). An important point is that each

[16] The features of the portal inherited from the NCBO BioPortal are more extensively described in other publications that are referenced here. We provide here only a small summary as well as relevant agronomy related examples. In addition, the documentation of the portal is also available: https://github.com/agroportal/documentation.

[17] http://agroportal.lirmm.fr/search?q = Abiotic%20factor

AgroPortal content page can be accessed by a direct URL, that can be potentially used to dereference an ontology URI. Dereferencing (or resolving) means to obtain a concrete representation of the identified resource (e.g., a web page), for instance, http://agroportal.lirmm.fr/ontologies/EOL/?p = classes&conceptid = http://opendata.inra.fr/EOL/EOL_0000014 directly points to the class 'water salinity' in Environment Ontology for Livestock. For each ontology, a JavaScript widget allowing autocomplete with class names is also automatically generated and can be used by external web applications to facilitate the edition of data fields restricted to ontology concepts.

Ontology versioning. AgroPortal handles versioning through the concept of "submission." Once an "Ontology" (an empty skeleton with minimal metadata) has been added once to the portal, "submission" objects can be attached. A new submission is created every time that ontology is re-submitted by a user, or pulled from its original location URL. Many ontologies are not necessarily maintained in a versioning system which offers a pull URL. It is up to the developer to decide when to manually uploading the new file, thereby creating a new submission (version) in AgroPortal. However, when the ontology is configured with a pull URL, the new ontology will be pulled in automatically (and versioned as a new submission) any night that it has changed. For example, the Matter Transfer Ontology for instance is developed by INRA using the @ Web application (http://pfl.grignon.inra.fr/atWeb).[18] Although only the latest version is indexed and therefore available for searching, browsing and annotation, all the previous versions are downloadable, and a difference comparison can be viewed for each submission.

Ontology mappings. Another key role of AgroPortal is to store mappings (or alignments) between ontologies (Ghazvinian et al., 2009). Indeed, because ontologies' contents overlap, it is crucial to maintain their interconnections—mappings—alongside the ontologies themselves. AgroPortal implements a mapping repository where each class-to-class mapping added to the portal is a first-class citizen and can be: stored, described, retrieved and deleted. The portal automatically creates some mappings when two classes share the same URI or CUI properties,[19] or when they share a common normalized preferred label or synonym. Although basic lexical mapping approaches can be inaccurate and should be used with caution (Faria et al., 2014; Pathak and Chute, 2009), they usually work quite well with the LOOM mapping algorithm used in AgroPortal (Ghazvinian et al., 2009). Other mappings can be explicitly uploaded from external sources, and in that case a mapping is reified as a resource described with provenance information (e.g., automatic or manual, who added it) and one or several tags to classify the mapping (e.g., owl:sameAs, skos:exactMatch, skos:broaderMatch, gold:translation). Such information helps users decide if they want to use these mappings.

Community feedback. While not being a state-of-the-art Web 2.0 social platform for ontologies, the AgroPortal features a few community features (Noy et al., 2009) such as: (i) *Ontology reviews*: for each ontology, a review can be written by a logged-in user from the ontology "Summary" page. It helps keep track of the quality. (ii) *Manual mapping creation*: On each ontology class, a logged-in user can create a mapping to another class (whether the class is inside the AgroPortal, or in the NCBO BioPortal or another resource (cf. next Section)) (Noy et al., 2008). While this is illustrative, and may stimulate propositions, the real strength of the portal comes from using the API to automatically import mappings. (iii) *Notes* can be attached in a forum-like mode to a specific ontology or class, in order to discuss the ontology (its design, use, or evolution) or allow users to propose changes to a certain class (for instance, see http://agroportal.lirmm.fr/ontologies/CO_321/?p = notes). Ontology developers (or any registered users) can subscribe to email notifications to be informed each time user feedback is added to their ontologies of interest.

Ontology-based annotation. AgroPortal features a text annotation service that will identify ontology classes inside any text (Jonquet et al., 2009) and can filter the results per ontologies and UMLS Semantic Types (McCray, 2003).[20] The text annotation service provides a mechanism to employ ontology-based annotation in curation, data integration, and indexing workflow; it has been used to semantically index several data resources such as in the NCBO Resource Index (Jonquet et al., 2011).[21] The workflow is based on a highly efficient syntactic concept recognition tool (using concept names and synonyms) (Dai et al., 2008), and on a set of semantic expansion algorithms that leverage the semantics in ontologies (e.g., is_a relations and mappings). The Annotator is illustrated Fig. 1. It is also used to recommend ontologies for given text input, as described hereafter.

Ontology recommendation. The NCBO (in collaboration with LIRMM & University of Coruña) has recently released a new version of the Recommender system in BioPortal (Martinez-Romero et al., 2017), which has also been installed in AgroPortal. This service suggests relevant ontologies from the parent repository for annotating text data. The new recommendation approach evaluates the relevance of an ontology to biomedical text data according to four different criteria: (1) the extent to which the ontology covers the input data; (2) the acceptance of the ontology in the community; (3) the level of detail of the ontology classes that cover the input data; and (4) the specialization of the ontology to the domain of the input data. This new version of a service originally released in 2010 (Jonquet et al., 2010) combines the strengths of its predecessor with a range of adjustments and new features that improve its reliability and usefulness. To our knowledge, the AgroPortal Recommender is the first ontology recommendation service made for the agronomy community to identify which ontologies are relevant for (i) a given corpus of text or (ii) a list of keywords. For instance, if used with the 'Plant height' text example, from Fig. 1. the service will help users to identify Trait Ontology and multiple sources from the Crop Ontology as relevant for this text.

Register ontology related projects. The AgroPortal provides a project list edited by its users that materialize the ontology-project relation. For instance, the relation between the Planteome project and the six ontologies it uses is described at http://agroportal.lirmm.fr/projects/Planteome, in a format that can be used by AgroPortal to illustrate the ontologies that are most used. This information can then be employed for instance to sort ontologies by number of projects that use them.

In addition, all the previous features are available through two endpoints allowing automatic querying of the content of the portal: (i) a REST web service API (http://data.agroportal.lirmm.fr/

[18] There are 328 submissions as of March 2017: http://data.agroportal.lirmm.fr/ontologies/TRANSMAT/submissions. The latest one is always available under http://data.agroportal.lirmm.fr/ontologies/TRANSMAT/latest_submission

[19] Uniform Resource Identifiers (URIs) are the standard way to identify resources (classes, properties, instances) on the semantic web when using RDF-based languages such as OWL or SKOS. Concept Unique Identifiers (CUIs) are identifiers used in the UMLS Metathesaurus. They are heavily used in the biomedical domain, but not very relevant within AgroPortal, where only two sources (the Semantic Network and the NCBI Taxonomy) are extracted from the UMLS.

[20] This feature originally developed for the NCBO Annotator (Jonquet et al., 2009) allows to filter the annotation results using the upper level 127 UMLS semantic type (http://agroportal.lirmm.fr/ontologies/STY) with which each concept in the UMLS are tagged. Because this was very useful on the NCBO BioPortal, we are considering an equivalent network and mechanism in the AgroPortal.

[21] The 'Resource Index' feature is not used in AgroPortal. Our vision is to accomplish this with the AgroLD partner project.

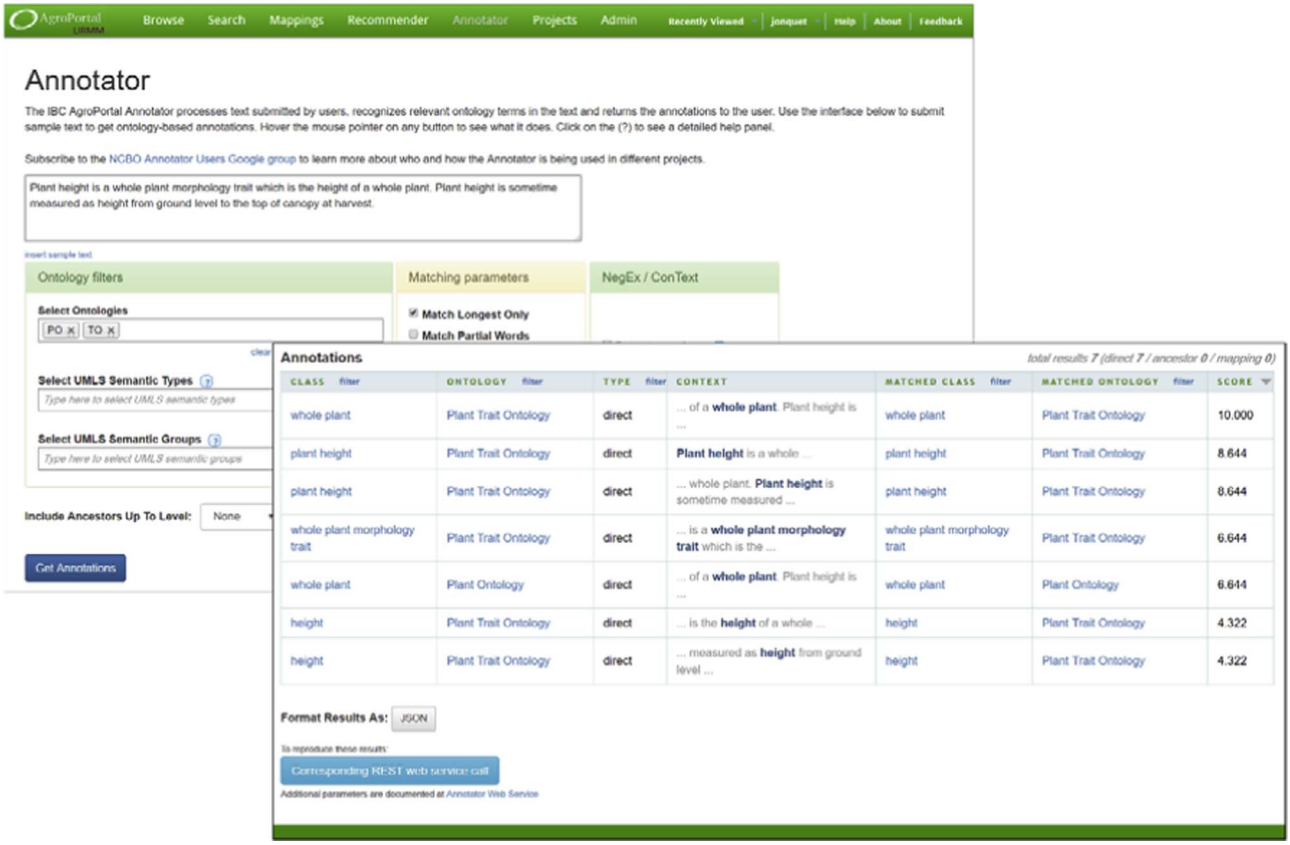

Fig. 1. AgroPortal Annotator with scored results. (web service call: http://services.agroportal.lirmm.fr/annotator?text = Plant height is a whole plant morphology trait which is the height of a whole plant. Plant height is sometime measured as height from ground level to the top of canopy at harvest.&ontologies = PO,TO&longest_only = true &whole_word_only = true&score = cvalue).

documentation) that returns XML or JSON-LD, making it easy to use AgroPortal within any web based application (Whetzel et al., 2011); and (ii) a SPARQL endpoint (http://sparql.agroportal.lirmm.fr/test), which is the standard mechanism to query RDF data (Salvadores et al., 2012).

We also like to point out that by adopting the NCBO technology, including its web service APIs (Whetzel and Team, 2013), an important number of external applications developed by the biomedical semantics community become available at very low cost for the agronomy community because of backward compatibility. This includes spreadsheet annotation tools such as OntoMaton (Maguire et al., 2013) Weboulous (Jupp et al., 2015), RightField (Wolstencroft et al., 2010) and WebSmatch (Coletta et al., 2012; Castanier et al., 2014); Zooma, a tool similar to the Annotator developed by the European Bioinformatics Institute (www.ebi.ac.uk/spot/zooma); the UIMA wrapper to use the Annotator web service in other NLP applications (Roeder et al., 2010); the ontology wrapper OntoCAT (Adamusiak et al., 2010); the Galaxy platform tools (Miñarro-Giménez et al., 2012); the visualization tool FlexViz (Falconer et al., 2009); and finally all the different API clients (Java, Ruby, Perl, etc.) developed by the NCBO (https://github.com/ncbo) or other organizations (e.g. REDCap or Protégé plugins). To some extent, other ontology platforms such as the AberOWL, which features reasoning capabilities that AgroPortal does not yet offer (Slater et al., 2016), can automatically pull content from the AgroPortal.

4.3. New AgroPortal features developped since the beginning of the project

While assuring community support, day-to-day maintenance and monitoring of the portal and keeping it up-to-date with the NCBO technology, we have worked on customizations and specific services. These services target the agronomic community, but that could in some cases be used for any domains. With the vision of collaborative development of BioPortal and AgroPortal, when relevant and possible, we push new features back to the main NCBO code branch where BioPortal users or the appliance itself can benefit. The AgroPortal open source code and documentation are accessible on GitHub: https://github.com/agroportal.

Multilingualism in AgroPortal. In the context of the SIFR project and in consultation with the NCBO, we are working on making BioPortal multilingual (Jonquet et al., 2015). This is still work in progress, although we have already added relevant metadata properties to: (i) identify the natural language in which labels are available; and (ii) link monolingual ontologies to their translations. We have also changed the representation of multilingual translation mappings. For the moment, we have chosen to consider English as the main language of AgroPortal (i.e., the one use to display content as well as indexed for Search, Annotator and Recommender services). Multilingual ontologies (i.e., with labels in multiple languages) are parsed, but only the English content is explicitly used. Non-English monolingual ontologies are attached as "views" of a main ontology that is solely described with metadata (no content). For instance, the French Agroecology Knowledge Management ontology, used in a French collaborative network (http://agroportal.lirmm.fr/ontologies/GECO) is only described with metadata but has attached a specific view (http://agroportal.lirmm.fr/ontologies/GECO-FR) with the real content in French.

Mapping related features. In order to interconnect AgroPortal with the NCBO BioPortal or any other repositories, we have changed the model of AgroPortal mappings to store mappings to ontologies (i) in another instance of the BioPortal technology ('inter-portal'), (ii) in any 'external' resources. Hence, any AgroPortal class can be linked to any class in other knowledge resource (e.g., DBPedia, WordNet, AgroLD) or the NCBO BioPortal itself). Mappings are described with

provenance data and typed with a property from a standard semantic web vocabulary (e.g., OWL, SKOS, GOLD). For instance:

- o The class 'plant organ' in the Plant Ontology has been manually mapped to the 'Plant organ' entity in the DBPedia knowledge base. The mapping tag used is skos:exactMatch which means that the classes represent the same entity, while not supporting a logical substitution (as with owl:sameAs).
- o The class 'biomass' in the Biorefinery ontology has been manually mapped to the class 'Biomass' in MeSH on the NCBO BioPortal, and automatically mapped to the class 'biomass' in the AnaEE Thesaurus.
- o The class 'zooplankton' in the AnaEE Thesaurus has been mapped to 'zooplankton' in the Ontology for MIRNA Target (http://purl.obolibrary.org/obo/OMIT_0015869), which is not available in AgroPortal.

Semantic annotation with scoring. Within the SIFR project we develop new features and natural language based enhancement that target all the Annotator deployments (the NCBO, AgroPortal or SIFR one). For instance, to facilitate the use of annotation for semantic indexing, we have implemented three scoring methods for the Annotator. They are based on term frequency and especially useful with multi-word terms. We demonstrate the results of these new scoring measures in Melzi and Jonquet (2014). For instance, when considering annotating the text:[22] "*Plant height is a whole plant morphology trait which is the height of a whole plant. Plant height is sometime measured as height from ground level to the top of canopy at harvest.*" with the AgroPortal Annotator, the scoring method gives more importance to the concept 'plant height' (score = 8.64) than to the concept 'height' (score = 4.32), whose lexical form is actually more frequent in the text. The user interface of the Annotator is illustrated in Fig. 1.

Ontology formats. We have worked on the full support of different formats such as (i) SKOS (SKOS-XL is not handled yet), which is highly used in agronomy (AnaEE Thesaurus, Agrovoc, CAB Thesaurus and NAL Thesaurus all use SKOS); and (ii) the Crop Ontology Trait Dictionary template v5, adopted for instance by the Breeding API and Crop Ontology (import/export in this format is currently done outside of AgroPortal).

Ontology metadata. To facilitate the ontology identification and selection process, which has been assessed as crucial to enable ontology reuse (Park et al., 2011), we implemented a new metadata model to better support descriptions of ontologies and their relations, respecting recent metadata specifications, vocabularies, and practices used in the semantic web community (Xiang et al., 2011). We reviewed the most common and relevant vocabularies (23 in total) to describe metadata for ontologies, including Dublin Core, VoID, Ontology Metadata Vocabulary, and the Data Catalog Vocabulary. We then grouped those properties into a unified and simplified model of 127 properties (distilled from an initial list of 346 properties that will be parsed by the portal)[23] that includes the 45 properties originally offered by the NCBO BioPortal, and describe all the new properties with standard vocabularies.[24] This gives us, for example, a model to describe the type of the semantic resource uploaded to the portal (for example, thesaurus, ontology, taxonomy, or terminology). Our work provided three important new features for AgroPortal (Toulet et al., 2016):

- o Once an ontology is uploaded, AgroPortal automatically extracts most of the ontology metadata if they are included in the original file, and automatically populates some of them if possible (e.g., metrics, endpoints, links, examples). Ontology developers can manually update those extracted or calculated values if desired. In addition, we have entirely redesigned AgroPortal's ontology submission page to facilitate editing the metadata. Whenever possible, the user interface facilitates the selection of the metadata values, while in the backend those values are stored with standard URIs. For instance, the user interface will offer a pop-up menu to select the relevant license (CC, BSD, etc.) while the corresponding URI will be taken from the RDFLicense dataset (http://rdflicense.appspot.com). Knowledge organization systems types are taken from the KOS Types Vocabulary from the Dublin Core initiative.[25] An example using the OntoBiotope ontology metadata page in AgroPortal is shown in Fig. 2.
- o AgroPortal ontology browse page (Fig. 3) offers three additional ways to filter ontologies in the list (content, natural language, formality level) as well as three new options to sort this list. We believe these new features facilitate the process of selecting relevant ontologies.
  - o We have begun facilitating the comprehension of the agronomical ontology landscape by displaying diagrams and charts about all the ontologies on the portal (average metrics, most used tools, leading contributors & organization, and more). We have created a new AgroPortal 'landscape' page that displays metadata "by property" –as opposed as "by ontology" as in Fig. 2 (http://agroportal.lirmm.fr/landscape).

For each ontology available and uploaded in the portal, we collaborate with the ontology developers to extensively describe their metadata. Information is generally found either in other registries (e.g., LovInra, VEST Registry, the OBO Foundry) or identified in the publication, web site, documentation, etc. found about the ontologies. With these curated metadata, all users can confidently select and review ontologies; any submission of the ontology can include more authoritative and more complete metadata, available to any user including the original provider, and for other linked open data users and applications; and AgoPortal's users can better understand the landscape of ontologies in the agronomy and related domains.

5. Driving agronomic use case results

Now that AgroPortal has been extensively presented, we focus on the results of each use case, and illustrate the value added by this portal and its semantic content.

5.1. Agronomic Linked Data (AgroLD)

The OWL versions of the ontologies available in AgroPortal were retrieved from that single repository. Although AgroPortal is not the main original location for these ontologies (they are accessible on the OBO Foundry and Cropontoloy.org) it was convenient to find them all in one place, and to use a unique and consistent API. Plus, we also used the AgroPortal Annotator web service to annotate more than 50 datasets and produced 22% additional triples, which were validated manually (Fig. 4). Building such an annotation service for all these ontologies was one of the driving needs for AgroPortal. Encoding the original data in RDF allowed us to establish an annotation for every appropriate case, using owl:sameAs relations, between the data element (e.g., Protein in the SouthGreen database) defined with a new URI (http://www.southgreen.fr/agrold/resource/Protein) and an ontology term (e.g., the term 'polypeptide' in the Sequence Ontology (http://purl.obolibrary.org/obo/SO_0000104). Note that we have decided to use owl:sameAs in this case as the resources are logically equivalent and this is a common practice in linked open data to

[22] Two appended definitions from the Trait Ontology and from the Crop Ontology.

[23] https://github.com/agroportal/documentation/tree/master/metadata

[24] For instance, the call http://data.agroportal.lirmm.fr/ontologies/PR/latest_submission?display = all will display the JSON-LD format of all the metadata properties (populated or not) for the Protein Ontology.

[25] http://wiki.dublincore.org/index.php/NKOS_Vocabularies (ANSI/NISO Z39.19-2005).

Fig. 2. AgroPortal's Ontolgy metadata page for ONTOBIOTOPE (http://agroportal.lirmm.fr/ontologies/ONTOBIOTOPE). The red box corresponds to the new metadata fields added in AgroPortal ontology model extracted by the portal, or provided by the adminstrators or by the ontology developers. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

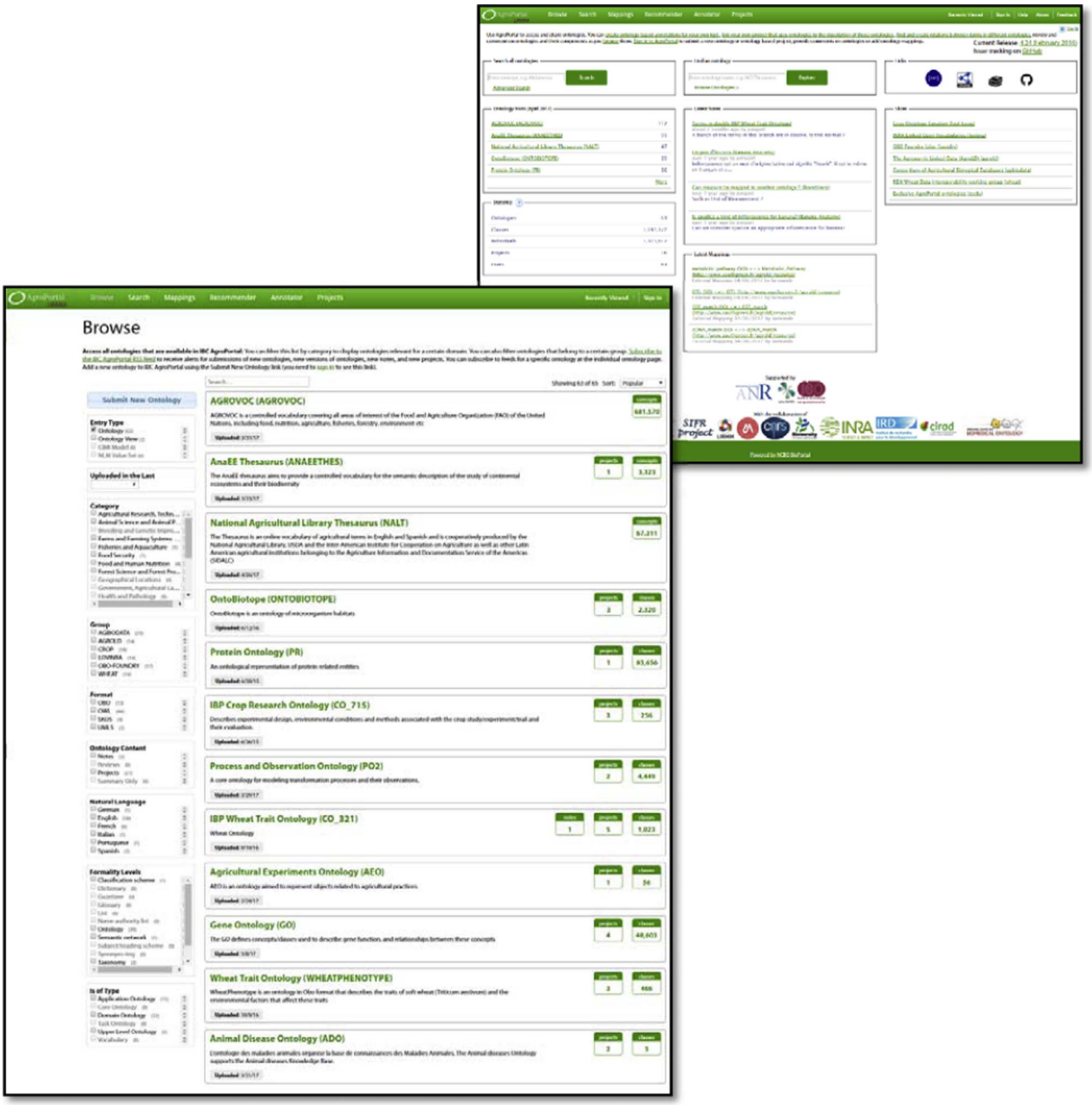

Fig. 3. Screeshots from the AgroPortal user interface (http://agroportal.lirmm.fr). The welcome page (back) provides a rapid overview of the content of the portal and enables a user to quickly search for and in ontologies. The browse ontology page (front) provides the list of ontologies and offer multiple sorting or faceted filtering of this list to facilitate the identification of the ontologies of interest.

interlink datasets; similar annotations have been made for properties using owl:equivalentProperty or rdfs:subPropertyOf (when an equivalent property did not exist). Now that AgroPortal handles 'external mapping' as described in Section 4.3, we have been able to upload all our annotations (to 23 classes and 21 properties) to fully connect the concepts from the different ontologies, and create annotations, directly within AgroPortal.[26]

As a result, AgroLD has incorporated the data from various databases (Table 4), and produced 37 million RDF triples (Venkatesan et al., 2015). The data source selection followed the needs and priorities of the IBC project's work-package 5. It included important data sources such as GOA, Gramene, Oryza Tag Line, and GreenPhylDB. AgroLD can now gather genomic and phenotypic information to answer biological questions such as: "find proteins involved in plant disease resistance and high grain yield traits." Such queries would be hard or impossible to resolve without the appropriate ontologies integrated to support the conclusion. The reader may refer to http://agrold.org/sparqleditor.jsp for more examples of queries in AgroLD.

5.2. RDA Wheat Data Interoperability (WDI) working group

We created and maintain explicit sub-parts within AgroPortal called

[26] The previous example ('polypeptide' in SO) is available here in the mapping tab: http://agroportal.lirmm.fr/ontologies/SO?p = classes&conceptid = http%3A%2F%2Fpurl.obolibrary.org%2Fobo%2FSO_0000104

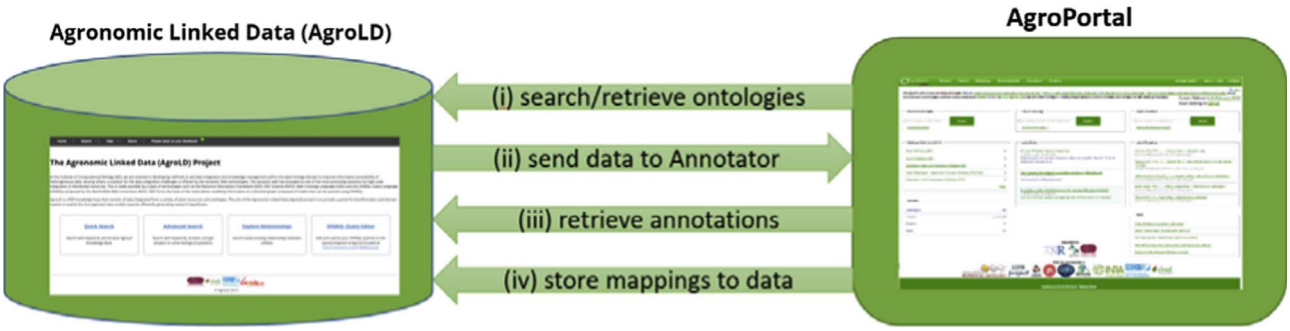

Fig. 4. Interaction between AgroPortal and AgroLD. (i) AgroPortal provides a unique endpoint to retrieve heterogenous ontologies; (ii) AgroLD's annotation pipleline send data to the AgroPortal Annotator and (iii) retrieves annotations with ontology terms used to build AgroLD; finally (iv) AgroPortal offers a link from the ontologies to data stored in AgroLD with the 'inter portal' mapping mechanism.

Table 4
Plant species and data sources in AgroLD. The number of tuples gives an idea of the number of elements we have annoated from the data sources and the number of RDF triples produced. The crops and ontologies are refered as: R = rice, W = wheat, A = Arabidopsis, S = sorghum, M = maize GO = Gene Ontology, PO = Plant Ontology, TO = Plant Trait Ontology, EO = Environment Ontology, SO = Sequence Ontology, CO = Crop Ontoloy (specific trait ontologies).

Data sources	URL s	# tuples	Crops	Ontologies used	# triples produced
GO associations	geneontology.org	1160 K	R, W, A, M, S	GO, PO, TO, EO	2700 K
Gramene	gramene.org	1718 K	R, W, M, A, S	GO, PO, TO, EO	5172 K
UniprotKB	uniprot.org	1400 K	R, W, A, M, S	GO, PO	10000 K
OryGenesDB	orygenesdb.cirad.fr	1100 K	R, S, A,	GO, SO	2300 K
Oryza Tag Line	oryzatagline.cirad.fr	22 K	R	PO, TO, CO	300 K
TropGeneDB	tropgenedb.cirad.fr	2 k	R	PO, TO, CO	20 K
GreenPhylDB	greenphyl.org	100 K	R, A	GO, PO	700 K
SniPlay	sniplay.southgreen.fr	16 K	R	GO	16000 K
TOTAL					37000 K

slices.[27] The wheat slice in AgroPortal (http://wheat.agroportal.lirmm.fr) allows the community to share common definitions for the words they utilize to describe and annotate data, which in turn makes the data more machine-readable and interoperable. Furthermore, each slice enables ontology developers to make their ontologies more visible to targeted agronomic research communities; as of today, AgroPortal's Wheat group contains 20 of the 23 ontologies identified by the WDI.[28] Each ontology has been carefully described (with licenses, authority, availability, and so on), and a new metadata property (omv:endorsedBy) is used to show the ontology's endorsement by the WDI working group.

This work has been reported in the WDI's set of guidelines for wheat data description (http://ist.blogs.inra.fr/wdi) (Dzalé-Yeumo et al., 2017), and used since then as a reference to identify and select ontologies related to wheat. Among AgroPortal's registered users, a dozen are members of the RDA WDI working group. In the future, the slice will be maintained/managed by the WheatIS consortium to organize new wheat-related ontologies and store the alignments between them. AgroPortal's adoption by the WDI working group leveraged several advanced features of the platform as customized by the AgroPortal team. The result directly enhanced the community's processes and capabilities, provided customized access to information of particular interest to this community, and achieved wide uptake in the working group.

5.3. INRA Linked Open Vocabularies (LovInra)

To augment the visibility of INRA's semantic resources, and achieve their mapping to resources within and external to INRA, the institute has chosen AgroPortal to publish and host INRA's resources and encourage adoption of semantic web standards. If a semantic resource is declared on the LovInra service, it is immediately uploaded and fully described on AgroPortal. Resources that are not on the LovInra service can be directly uploaded by their developers to the portal, an important consideration for such a big organization. AgroPortal assigns the new resources to the correct group and slice, and properly tags them (SKOS vocabularies, OWL/SKOS termino-ontological resources, or OBO/OWL ontologies).

The LovInra group/slice contains 16 ontologies relating to process modeling, biotopes, animal breeding, and plant phenotypes. AgroPortal has become a major element of the LovInra service and is heavily encouraged and supported by INRA. It has started to play a key resource role allowing the group's users to: (i) have a comprehensive view of the portal's ontologies (topics, types, community, etc.); (ii) quickly find a resource, and understand its content and structure by browsing it and annotating documents; (iii) discover additional vocabularies that could be used; and (iv) have access to projects linked to vocabularies, and understand how they were created or used by the projects, possibly exchanging shared experience or insights.

5.4. The Crop Ontology project

Currently, the AgroPortal hosts 19 crop-specific trait ontologies developed within the Crop Ontology project: Wheat, Rice, Cassava, Groundnut, Chickpea, Banana, Sweet potato, Cowpea, Soybean, Lentil, Pigeon pea, Sorghum, Pear millet, Maize, Groundnut, Castor bean, Mungbean, and Cassava. Additional ontologies will be integrated in the future with the help of the crop ontology curators. Similarly to the

[27] Slices are a mechanism supported by the platform to allow users to interact (both via API or UI) only with a subset of ontologies in AgroPortal. If browsing the slice, all the portal features will be restricted to the chosen subset, enabling users to focus on their specific use cases. On AgroPortal, slices and groups are synchronized, so every group (described Section 4.1) has a corresponding slice displaying only the ontologies from that group.

[28] Among the missing ones are, CAB Thesaurus, that we are currently working on integrating; CheBI that we have decided not to upload yet; and Wheat Inra Phenotype Ontology (that is currently being merged with CO_321).

LovInra or WDI use cases, these ontologies are grouped within the portal and can be browsed in a dedicated slice (http://crop.agroportal.lirmm.fr). Parsers for specific trait template have been developed, and in the future any of this community's formats (OBO, OWL, and CSV) shall be used to import and export trait ontologies directly within AgroPortal.[29]

Moreover, in the context of the Planteome project (www.planteome.org), the alignment (or mapping) of terms within and across different plant related ontologies have been created: both within the crop ontologies themselves (in different crop branch) or with other reference ontologies commonly used in plant biology (e.g., PO, TO, EO). In the future, AgroPortal will formally store the alignments between all these ontologies.[30]

Finally, hosting ontologies on AgroPortal offers new functionalities to the crop ontology community such as versioning, an open SPARQL endpoint, community notes, and the annotation service, while still supporting the uses of the current web site.[31] For instance, new traits or mappings between them can be suggested directly by breeders using AgroPortal's community features, while not directly impacting the original ontology. Each time a suggestion is made to an ontology, the breeders interested in the corresponding crop can be notified of the suggestions and comments of their peers.

5.5. GODAN Map of Agri-Food Data Standards

The GODAN Action project wanted to build a broadly scoped global map of standards while leveraging detailed information and content about them that could be maintained in an ontology or vocabulary. To achieve this, the new map of standards was built on top of the existing VEST Registry, but added bidirectional mechanisms linking the VEST Registry with AgroPortal. The combined system automatically imports resource descriptions from the AgroPortal into the VEST, and links records from the VEST back to the AgroPortal entries, in order to provide access to the AgroPortal content and related services. The new registry, called *Map of Agri-Food Data Standards* (http://vest.agrisemantics.org), was released in 2016 under two umbrellas: the GODAN Action project, and the new RDA AgriSemantics working group,[32] which launched at the end of 2016. The Map of Standards leverages the AgroPortal's new metadata model and application programming interface to populate the entries in the Map using a single web service call. In addition to searching by metadata, the AgroPortal's Recommender will help the agronomy community identify ontologies or vocabularies of interest.

The synchronization and interlinking of the two platforms is for the moment semi-automatic, with the content of AgroPortal being regularly imported into the global map. Users can register or edit the description of a vocabulary in the Map, and if the vocabulary is in a compatible format, they are offered, the option to add the vocabulary directly into AgroPortal. In the future, this process will be fully automatized.

6. Discussion

6.1. General reflection on research scenarios supported by AgroPortal

AgroPortal (like the NCBO BioPortal before it) adopted a vision where multiple knowledge artifacts are made available in a common place (though not combined), and cast to a common model. While doing so, the portal arguably limits the full power of ontologies, constraining their use to features supported by the common model. We see two general scenarios of use for our portal:

> The portal provides basic ontology library services for users with a "vertical need" —those who want to do very precise things (e.g., reasoning, using specific relations) using only suitable ontologies (developed by the same communities and in the same format). Such users may just use the portal to find and download ontologies, and work in their own environment.
> The portal provides many semantic services (for examples, lexical analysis, search, text annotation, and use of hierarchical knowledge) to users with "horizontal needs" —those who wants to work with a wide range of ontologies and vocabularies useful in their domain but developed by different communities, overlapping and in different formats. Such users greatly appreciate the unique endpoints (web application and programmatic for REST and SPARQL queries) offered by the portal under a simplified common model.

We believe there are existing resource to address the first need in agronomy (e.g., OBO Foundry, FAIRSharing, VEST registry), although without containing all the relevant ontologies and vocabularies. However, we argue the second need is unmet by any of the available platforms. If we want semantic resources like ontologies and vocabularies to achieve widespread adoption, we must facilitate their use for non-ontological experts who still want to use multiple heterogeneous semantic resources.

6.2. Implementation of the requirements

As presented and illustrated on examples, most of the requirements listed in Section 3 have been addressed at least partially thanks to the original BioPortal features (e.g., requirements #1-#6, #8, #10, #15, #16, #18), our new implementations (#5, #7, #11, #15), and our applying the platform to the community needs (#1, #10, #11, #17, #18). Some requirements are not yet completely achieved and/or evaluated, for instance:

> (#4) The AgroPortal Annotator has been used by the AgroLD use case, but not by other ones. We have not yet evaluated the capability of the service to automatically identify entities such as plant phenotypes in text.
> (#8) Automatically generating mappings is an important issue for a portal on ontologies. Although it is convenient to have some simple lexical mappings automatically generated by AgroPortal with the LOOM algorithm (Ghazvinian et al., 2009), we find that this is not enough to correctly interlink the multiple vocabularies and ontologies developed by the community. We are integrating other state-of-the-art ontology matchers such as YAM++ (Ngo and Bellahsene, 2012) as well as designing specific mapping curation interfaces. At the same time, identifying and harvesting into AgroPortal the mappings already produced by the community is a huge task, not yet begun.
> (#9) We have not automatically linked databases of annotated agronomical data using ontology concepts (from within AgroPortal). While the original BioPortal has the NCBO Resource Index (Jonquet et al., 2011), we plan to rely on external annotated resources such as AgroLD (Venkatesan et al., 2015) to interlink with data. To store this information, we will build on our rich mapping model in AgroPortal as presented Section 4.3. As another example, being part of the map of standards will allow ontologies in AgroPortal to link directly to

[29] Most of these conversions are still achieved outside of AgroPortal. The automatically generated CSV output format is not yet compliant with the Crop Ontology trait template (v5).

[30] For instance, something to capture that plant height for wheat (CO_321:0000024) is somehow linked to the general plant height trait (TO_0000207) that is itself a morphology trait (TO:0000398). This work is ongoing, and the data is not yet publicly released.

[31] In the future, to offer to breeders a simple and customized interface while avoiding duplication effort, we will consider serving the Crop Ontology website use cases by directly accessing AgroPortal's backend through the REST API.

[32] https://www.rd-alliance.org/groups/agrisemantics-wg.html

datasets that use them such as the CIARD RING directory (http://ring.ciard.net) (Pesce et al., 2011), as that was previously indexed with some of the VEST content. The CIARD RING can be queried via SPARQL or REST API and the links between vocabularies and datasets can therefore be retrieved by any system. Such a feature, has been requested and will be among the next features of AgroPortal. In the long-term vision, AgroPortal will directly query the CIARD RING, AgroLD, or any relevant data sources like Bio2RDF or Planteome, so that a user browsing ontologies can get direct access to the data to which these ontologies link.

(#12) Although community feedback is an important aspect for working group and communities, we have not successfully engaged yet our user groups to add reviews, notes, or comments about the ontologies. A complete rethinking of this issue is a future challenge for AgroPortal.

(#13) The roadmap to make the technology fully multilingual has been identified, but not yet fully implemented.

(#15) AgroPortal can be used as a destination for dereferenced URIs. In the future, we shall discuss these strategic questions with our collaborators.

6.3. Future and perspectives

Considering the need for a repository of ontologies for agronomy, food, plant sciences, and biodiversity, we expect broad community adoption of the AgroPortal. The endorsement of associated partners (IRD, CIRAD, INRA, IRSTEA) illustrates the impact and interest not just in France, but also internationally (e.g., FAO, Bioversity International, IC-FOODS consortium, NCBO, Planteome, RDA working groups). More recently, two other RDA working groups (Rice Data Interoperability[33] and AgriSemantics[34]) have expressed interest in using AgroPortal as a backbone for data integration and standardization.

In the future, we will identify more potential users for the portal and support new research scenarios. For instance, within the RDA AgriSemantics WG, we are interested in using AgroPortal to host the future Global Agricultural Concept Scheme (GACS) (Baker et al., 2016), which will result from the integration and alignment of Agrovoc, NAL Thesaurus and CAB Thesaurus. The portal is considered by the GACS working group as a candidate to host the three source vocabularies (it already includes two of them), as well as the GACS itself. GACS beta version 3.1 is currently available in AgroPortal, but no specific customization has been performed. In addition, we will be offering our services to these projects:

the new IC-FOODS project (International Center for Food Ontology, Operability, Data & Semantics - www.ic-foods.org) that will be developing ontologies related to food, nutrition, eating behaviors (Musker et al., 2016);

ecologists developing the Thesaurus of Plant characteristics (Garnier et al., 2017);

the French IRESTA organization, to facilitate the use of ontologies in the design of the future government-led open data repository for agriculture project (AgGate).[35]

To foster interest in agronomy and the semantic web and identify potential AgroPortal applications, we launched in 2016 a series of AgroHackathons (www.agrohackathon.org) that focused among other things on AgroPortal and AgroLD. Finally, in the next future, we plan to achieve a community survey evaluation to capture the feedback of our community, review the requirements, and drive the future directions of the project.

7. Conclusion

In this paper we have presented AgroPortal, an open vocabulary and ontology repository for agronomy. We have discussed five use cases already using the portal to support their work on data interoperability, and demonstrated that beyond these use cases the portal offers services of value to the broader community. The thematic boundaries of the portal are evolving (agriculture also includes animals, and is strongly related to environmental science), and over time the community will communicate what they expect to find in such a repository.

The community outreach challenge of such a project is huge. It involves identifying already existing resources, whether already shared or not, encouraging their developers to make them available, and finally harvesting them into the single ontology repository, capable of providing many services across the heterogeneous content. We recognize that this challenge was highly facilitated by previous important efforts such as the NCBO BioPortal, OBO Foundry, Planteome, and Crop Ontology projects. In addition, we are conscious that by adopting an open library approach, knowledge "conflicts" or redundancies as well as convergences and consolidations will appear. We believe the AgroPortal will help the scientific community to fully understand these issues, and address them as appropriate.

The technological challenges of such a project are also huge; therefore, we have built upon technology previously developed in the biomedical domain. We see here an opportunity to capitalize technology and scientific outcomes of the last twelve years in a closely related domain. We illustrated in the context of five important driving agronomic use cases how AgroPortal can enable new science for the community developing and using agronomical ontologies and vocabularies worldwide. In addition, the AgroPortal platform offers a terrain for pursuing important informatics and semantic web issues, such as semantic annotation, multilingual ontologies, metadata description, ontology engineering and alignment, and ontology recommendation, and will.

Ultimately, we believe AgroPortal provides powerful services, standards, and information that will greatly facilitate the adoption of open data in agriculture and benefit the extended agronomic community, the semantic web and data science communities, and the biomedical community that in many ways laid the groundwork that AgroPortal now leverages.

Acknowledgment

This work is partly achieved within the Semantic Indexing of French biomedical Resources (SIFR – www.lirmm.fr/sifr) project that received funding from the French National Research Agency (grant ANR-12-JS02-01001), the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 701771, the NUMEV Labex (grant ANR-10-LABX-20), the Computational Biology Institute of Montpellier (grant ANR-11-BINF-0002), as well as by the University of Montpellier and the CNRS. We also thank the National Center for Biomedical Ontologies for their help and time spent with us in deploying the AgroPortal.

Author contributions

CJ conceived of the project, provided the scientific direction and led the writing of this manuscript. VE & AT respectively implemented/maintained the portal and managed the content with help of the community. JG and MAM helped and gave directions in realizing the project in collaboration with NCBO, and JG provided extensive final review and editing. Then, EA, SA, MAL, EDY, VP & PL respectively presented each of the use cases. All authors declare no conflict of interest and approved the final manuscript.

[33] https://rd-alliance.org/groups/rice-data-interoperability-wg.html
[34] https://rd-alliance.org/groups/agrisemantics-wg.html
[35] https://www.economie.gouv.fr/files/files/PDF/rapport-portail-de-donnees-agricoles.pdf

Appendix

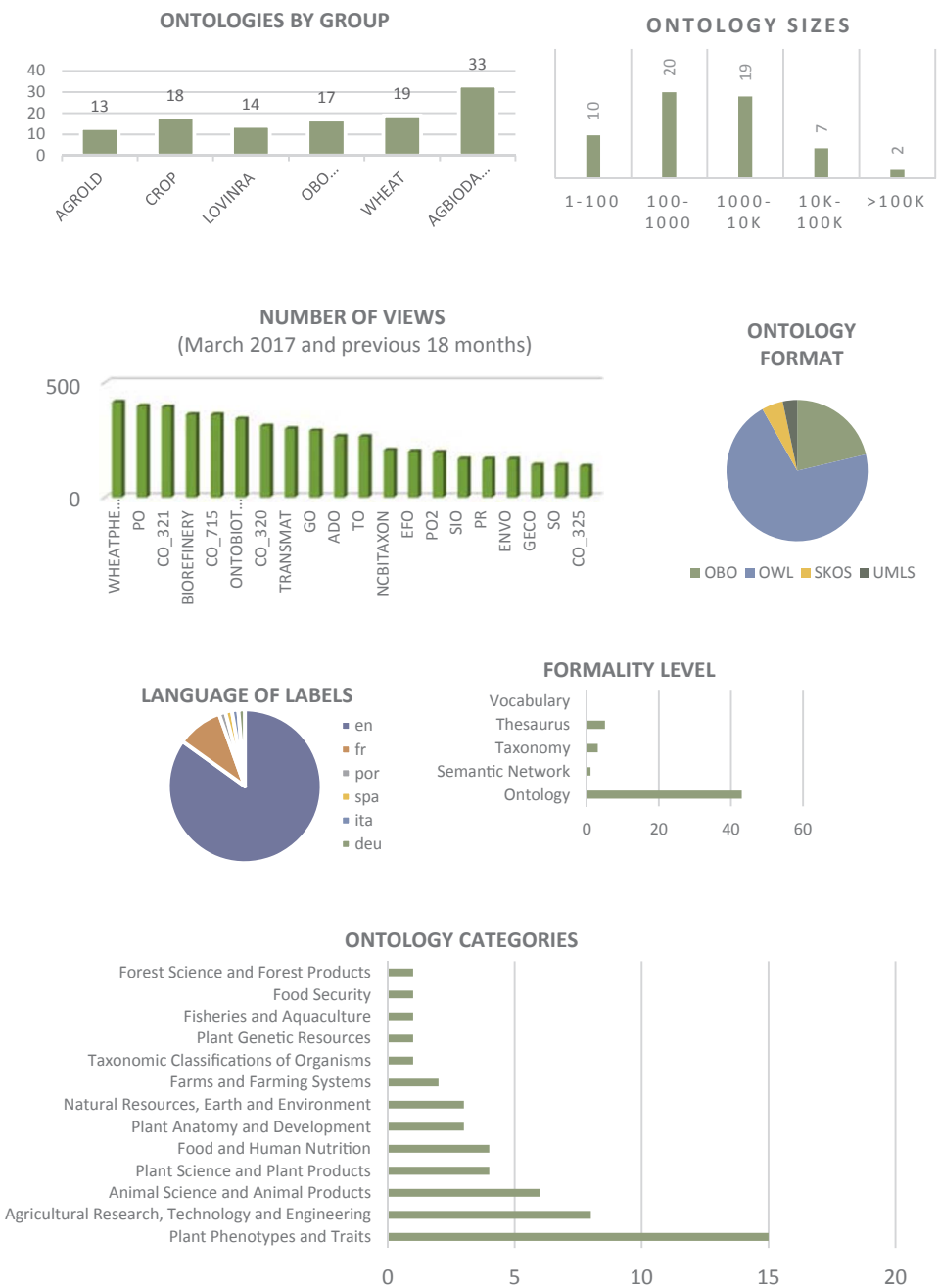

Fig. 5. AgroPortal public ontology analytics (May 2017). Updated versions of these statistics are automatically generated from AgroPortal's new metadata model, and made available on its Landscape page (http://agroportal.lirmm.fr/landscape).

References

Goble, C., Stevens, R., 2008. State of the nation in data integration for bioinformatics. Biomed. Inf. 41, 687–693.

Rubin, D.L., Shah, N.H., Noy, N.F., 2008. Biomedical ontologies: a functional perspective. Brief. Bioinform. 9 (1), 75–90.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al., 2000. Gene ontology: tool for the unification of biology. Nat. Genet. 25 (1), 25–29.

Cooper, L., Walls, R.L., Elser, J., Gandolfo, M.A., Stevenson, D.W., Smith, B., Preece, J., Athreya, B., Mungall, C.J., Rensing, S., Hiss, M., Lang, D., Reski, R., Berardini, T.Z., Li, D., Huala, E., Schaeffer, M., Menda, N., Arnaud, E., Shrestha, R., Yamazaki, Y., Jaiswal, P., 2012. The plant ontology as a tool for comparative plant anatomy and genomic analyses. Plant Cell Physiol. 54, e1.

Shrestha, R., Arnaud, E., Mauleon, R., Senger, M., Davenport, G.F., Hancock, D., Morrison, N., Bruskiewich, R., McLaren, G., 2010. Multifunctional crop trait ontology for breeders' data: field book, annotation, data discovery and semantic enrichment of the literature. AoB Plants, vol. 2010, May.

Buttigieg, P.L., Morrison, N., Smith, B., Mungall, C.J., Lewis, S.E., 2013. The environment ontology: contextualising biological and biomedical entities. Biomed. Semantics 4, 43.

Devare, M., Aubert, C., Laporte, M.-A., Valette, L., Arnaud, E., Buttigieg, P.L., 2016. Data-driven agricultural research for development - a need for data harmonization via semantics. In: Jaiswal, P., Hoehndorf, R. (Eds.), 7th International Conference on Biomedical Ontologies, ICBO'16, vol. 1747 of CEUR Workshop Proceedings, Corvallis, Oregon, USA, pp. 2, August.

Garnier, E., Stahl, U., Laporte, M.-A., Kattge, J., Mougenot, I., Kühn, I., Laporte, B., Amiaud, B., Ahrestani, F.S., Bönisch, G., Bunker, D.E., Cornelissen, J.H.C., Díaz, S., Enquist, B.J., Gachet, S., Jaureguiberry, P., Kleyer, M., Lavorel, S., Maicher, L., Pérez-Harguindeguy, N., Poorter, H., Schildhauer, M., Shipley, B., Violle, C., Weiher, E., Wirth, C., Wright, I.J., Klotz, S., 2017. Towards a thesaurus of plant characteristics: an ecological contribution. Ecology 105, 298–309.

Griffiths, E., Brinkman, F., Buttigieg, P.L., Dooley, D., Hsiao, W., Hoehndorf, R., 2016. FoodON: a global farm-to-fork food ontology - the development of a universal food vocabulary. In: Jaiswal, P., Hoehndorf, R., (Eds.), 7th International Conference on

Biomedical Ontologies, ICBO'16, vol. 1747 of CEUR Workshop Proceedings, Corvallis, Oregon, USA, pp. 2, August.
Musker, R., Lange, M., Hollander, A., Huber, P., Springer, N., Riggle, C., Quinn, J.F., Tomich, T.P., 2016. Towards designing an ontology encompassing the environment-agriculture-food-diet-health knowledge spectrum for food system sustainability and resilience. In: Jaiswal, P., Hoehndorf, R. (Eds.), 7th International Conference on Biomedical Ontologies, ICBO'16, vol. 1747 of CEUR Workshop Proceedings, Corvallis, Oregon, USA, pp. 5, August.
Hughes, L.M., Bao, J., Hu, Z.-L., Honavar, V., Reecy, J.M., 2014. Animal trait ontology: The importance and usefulness of a unified trait vocabulary for animal species. Anim. Sci. 86, 1485–1491.
Meng, X.-X., 2012. Special issue – agriculture ontology. Integrative Agriculture, vol. 11, pp. 1, May.
Walls, R.L., Deck, J., Guralnick, R., Baskauf, S., Beaman, R., Blum, S., Bowers, S., Buttigieg, P.L., Davies, N., Endresen, D., Gandolfo, M.A., Hanner, R., Janning, A., Krishtalka, L., Matsunaga, A., Midford, P., Morrison, N., Tuama, Éamonn Ó., Schildhauer, M., Smith, B., Stucky, B.J., Thomer, A., Wieczorek, J., Whitacre, J., Wooley, J., 2014. Semantics in support of biodiversity knowledge discovery: an introduction to the biological collections ontology and related ontologies. PLoS One 9, 13.
Wang, Y., Wang, Y., Wang, J., Yuan, Y., Zhang, Z., 2015. An ontology-based approach to integration of hilly citrus production knowledge. Comput. Electron. Agric. 113, 24–43.
Lousteau-Cazalet, C., Barakat, A., Belaud, J.-P., Buche, P., Busset, G., Charnomordic, B., Dervaux, S., Destercke, S., Dibie, J., Sablayrolles, C., Vialle, C., 2016. A decision support system for eco-efficient biorefinery process comparison using a semantic approach. Comput. Electron. Agric. 127, 351–367.
Lehmanna, R.J., Reichea, R., Schiefera, G., 2012. Future internet and the agri-food sector: State-of-the-art in literature and research. Comput. Electron. Agric. 89, 158–174.
Jaiswal, P., 2011. Plant Reverse Genetics: Methods and Protocols, ch. Gramene Database: A Hub for Comparative Plant Genomics. Humana Press, pp. 247–275.
Sachit Rajbhandari, J.K., 2012. The AGROVOC concept scheme – a walkthrough. Integrative, Agriculture 11, 694–699.
d'Aquin, M., Noy, N.F., 2012. Where to publish and find ontologies? a survey of ontology libraries. Web Semantics 11, 96–111.
Bodenreider, O., 2004. The unified medical language system (UMLS): integrating biomedical terminology. Nucleic Acids Res. 32, 267–270.
Smith, B., Ashburner, M., Rosse, C., Bard, J., Bug, W., Ceusters, W., Goldberg, L.J., Eilbeck, K., Ireland, A., Mungall, C.J., Consortium, T.O., Leontis, N., Rocca-Serra, P., Ruttenberg, A., Sansone, S.-A., Scheuermann, R.H., Shah, N.H., Whetzel, P.L., Lewis, S., 2007. The OBO foundry: coordinated evolution of ontologies to support biomedical data integrationNat. Biotechnol. 25, 1251–1255.
Larmande, P., Arnaud, E., Mougenot, I., Jonquet, C., Libourel, T., Ruiz, M., (Eds.), 2013. Proceedings of the 1st International Workshop on Semantics for Biodiversity, Montpellier, France, May.
Baker, T., Caracciolo, C., Jaques, Y., (Eds.), 2015. Report on the Workshop "Improving Semantics in Agriculture, (Rome, Italy), Food and Agriculture Organization of the UN, July.
Noy, N.F., Shah, N.H., Whetzel, P.L., Dai, B., Dorf, M., Griffith, N.B., Jonquet, C., Rubin, D.L., Storey, M.-A., Chute, C.G., Musen, M.A., 2009. BioPortal: ontologies and integrated data resources at the click of a mouse. Nucleic Acids Res. 37, 170–173.
Whetzel, P.L., Noy, N.F., Shah, N.H., Alexander, P.R., Nyulas, C., Tudorache, T., Musen, M.A., 2011. BioPortal: enhanced functionality via new web services from the national center for biomedical ontology to access and use ontologies in software applications. Nucleic Acids Res. 39, 541–545.
Whetzel, P.L., Team, N., 2013. NCBO technology: powering semantically aware applications. Biomed. Semantics 49 4S1.
Ding, Y., Fensel, D., 2001. Ontology library systems: the key to successful ontology re-use. In: 1st Semantic Web Working Symposium, SWWS'01, Stanford, CA, USA, pp. 93–112, CEUR-WS.org, August.
Baclawski, K., Schneider, T., 2009. The open ontology repository initiative: Requirements and research challenges. In: Tudorache, T., Correndo, G., Noy, N., Alani, H., Greaves, M., (Eds.), Workshop on Collaborative Construction, Management and Linking of Structured Knowledge, CK'09, vol. 514 of CEUR Workshop Proceedings, Washington, DC., USA, pp. 10, CEUR-WS.org, October.
Sansone, S.-A., Rocca-Serra, P., Field, D., Maguire, E., Taylor, C., Hofmann, O., Fang, H., Neumann, S., Tong, W., Amaral-Zettler, L., Begley, K., Booth, T., Bougueleret, L., Burns, G., Chapman, B., Clark, T., Coleman, L.-A., Copeland, J., Das, S., de Daruvar, A., de Matos, P., Dix, I., Edmunds, S., Evelo, C.T., Forster, M.J., Gaudet, P., Gilbert, J., Goble, C., Griffin, J.L., Jacob, D., Kleinjans, J., Harland, L., Haug, K., Hermjakob, H., Sui, S.J.H., Laederach, A., Liang, S., Marshall, S., McGrath, A., Merrill, E., Reilly, D., Roux, M., Shamu, C.E., Shang, C.A., Steinbeck, C., Trefethen, A., Williams-Jones, B., Wolstencroft, K., Xenarios, I., Hide, W., 2012. Toward interoperable bioscience data. Nat. Genet. 44, 121–126.
Pesce, V., Geser, G., Protonotarios, V., Caracciolo, C., Keizer, J., 2013. Towards linked agricultural metadata: directions of the agINFRA project. In: 7th Metadata and Semantics Research Conference, AgroSem track, Thessaloniki, Greece, pp. 12, November.
Xiang, Z., Mungall, C., Ruttenberg, A., He, Y., 2011. Ontobee: a linked data server and browser for ontology terms. In: Bodenreider, O., Martone, M.E., Ruttenberg, A., (Eds.), 2nd International Conference on Biomedical Ontology, ICBO'11, vol. 833 of CEUR Workshop Proceedings, Buffalo, NY, USA, p. 3, July.
Côté, R.G., Jones, P., Apweiler, R., Hermjakob, H., 2006. The ontology lookup service, a lightweight cross-platform tool for controlled vocabulary queries. BMC Bioinf. 7, 7.
Hoehndorf, R., Slater, L., Schofield, P.N., Gkoutos, G.V., 2015. Aber-OWL: a framework for ontology-based data access in biology. BMC Bioinf. 16, 1–9.
Vandenbussche, P.-Y., Atemezing, G.A., Poveda-Villalón, M., Vatant, B., 2014. Linked open vocabularies (LOV): a gateway to reusable semantic vocabularies on the web. Semantic Web.
Till, M., Kutz, O., Codescu, M., 2014. Ontohub: A semantic repository for heterogeneous ontologies, In: Theory Day in Computer Science, DACS'14, (Bucharest, Romania), p. 2, September.
Graybeal, J., Isenor, A.W., Rueda, C., 2012. Semantic mediation of vocabularies for ocean observing systems. Comput. Geosci. 40, 120–131.
Ong, E., Xiang, Z., Zhao, B., Liu, Y., Lin, Y., Zheng, J., Mungall, C., Courtot, M., Ruttenberg, A., He, Y., 2016. Ontobee: a linked ontology data server to support ontology term dereferencing, linkage, query and integration. Nucleic Acids Res. 45, D347–D352.
Jonquet, C., Shah, N.H., Musen, M.A., 2009. The open biomedical annotator. In: American Medical Informatics Association Symposium on Translational BioInformatics, AMIA-TBI'09, San Francisco, CA, USA, pp. 56–60, March.
Jonquet, C., LePendu, P., Falconer, S., Coulet, A., Noy, N.F., Musen, M.A., Shah, N.H., 2011. NCBO resource index: ontology-based search and mining of biomedical resources, web semantics. In: 1st Prize of Semantic Web Challenge at the 9th International Semantic Web Conference, ISWC'10, Shanghai, China, vol. 9, pp. 316–324, September.
Salvadores, M., Alexander, P.R., Musen, M.A., Noy, N.F., 2013. BioPortal as a dataset of linked biomedical ontologies and terminologies in RDF. Semantic Web 4 (3), 277–284.
Rueda, C., Bermudez, L., Fredericks, J., 2009. The MMI ontology registry and repository: a portal for marine metadata interoperability. In: MTS/IEEE Biloxi - Marine Technology for Our Future: Global and Local Challenges, OCEANS'09, Biloxi, MS, USA, pp. 6, October.
D.A., Pouchard, L., Huhns, M., 2012. Lessons learned in deploying a cloud-based knowledge platform for the ESIP Federation. In: American Geo-physical Union Fall Meeting, poster session, San Francisco, USA, December.
Jonquet, C., Annane, A., Bouarech, K., Emonet, V., Melzi, S., 2016. SIFR BioPortal: Un portail ouvert et générique d'ontologies et de terminologies biomédicales françaises au service de l'annotation sémantique. In: 16th Journées Francophones d'Informatique Médicale, JFIM'16, Genève, Suisse, pp. 16, July.
Jonquet, C., Emonet, V., Musen, M.A., 2015. Roadmap for a multilingual BioPortal. In: In: Gracia, J., McCrae, J., Vulcu, G. (Eds.), 4th Workshop on the Multilingual Semantic Web, MSW4'15, vol. 1532. CEUR Workshop Proceedings, Portoroz, Slovenia, pp. 15–26.
Matteis, L., Chibon, P., Espinosa, H., Skofic, M., Finkers, R., Bruskiewich, R., Arnaud, E., 2015. Crop ontology: vocabulary for crop-related concepts. In: In: Larmande, P., Arnaud, E., Mougenot, I., Jonquet, C., Libourel, T., Ruiz, M. (Eds.), 1st International Workshop on Semantics for Biodiversity, vol. 1. CEUR Workshop Proceedings, Montpellier, France, pp. 37–46.
Noy, N.F., Tudorache, T., Nyulas, C., Musen, M.A., 2010. The ontology life cycle: Integrated tools for editing, publishing, peer review, and evolution of ontologies. In: AMIA Annual Symposium, Washington DC, USA, pp. 552–556, November.
Jaiswal, P., Cooper, L., Elser, J.L., Meier, A., Laporte, M.-A., Mungall, C., Smith, B., Johnson, E.K., Seymour, M., Preece, J., Xu, X., Kitchen, R.S., Qu, B., Zhang, E., Arnaud, E., Carbon, S., Todorovic, S., Stevenson, D.W., 2016. Planteome: A resource for Common Reference Ontologies and Applications for Plant Biology. In: 24th Plant and Animal Genome Conference, PAG'16, San Diego, USA, January.
Carbon, S., Ireland, A., Mungall, C.J., Shu, S., Marshall, B., Lewis, S., 2009. AmiGO: online access to ontology and annotation data. Bioinformatics 25 (2), 288–289.
Schmachtenberg, M., Bizer, C., Paulheim, H., 2014. Adoption of the linked data best practices in different topical domains. In: Mika, P., Tudorache, T., Bernstein, A., Welty, C., Knoblock, C., Vrandecic, D., Groth, P., Noy, N., Janowicz, K., Goble, C., (Eds.), 13th International Semantic Web Conference, ISWC'14, vol. 8796 of Lecture Notes in Computer Science, Riva del Garda, Italy, Springer, pp. 245–260, October.
Belleau, F., Nolin, M.-A., Tourigny, N., Rigault, P., Morissette, J., 2008. Bio2RDF: towards a mashup to build bioinformatics knowledge systems. Biomed. Inf. 41, 706–716.
Jupp, S., Malone, J., Bolleman, J., Brandizi, M., Davies, M., Garcia, L., Gaulton, A., Gehant, S., Laibe, C., Redaschi, N., Wimalaratne, S.M., Martin, M., Le Novère, N., Parkinson, H., Birney, E., Jenkinson, A.M., 2014. The EBI RDF platform: linked open data for the life sciences. Bioinformatics 30, 1338–1339.
Groth, P., Loizou, A., Gray, A.J., Goble, C., Harland, L., Pettifer, S., 2014. API-centric linked data integration: the open PHACTS discovery platform case study. Web Semantics 29, 12–18.
Venkatesan, A., Hassouni, N.E., Philippe, F., Pommier, C., Quesneville, H., Ruiz, M., Larmande, P., 2015. Exposing French agronomic resources as linked open data. In: 8th Semantic Web Applications and Tools for Life Sciences International Conference, SWAT4LS'15, vol. 546 of CEUR Workshop Proceedings, Cambridge, UK, pp. 205–207, December.
Hassani-Pak, K., Zorc, M., Taubert, J., Rawlings, C., 2013. QTLNetMiner - candidate gene discovery in plant and animal knowledge networks. In: 21st Plant & Animal Genome Conference, poster session, San Diego, USA, pp. P0980, January.
Dzalé-Yeumo, E., et al., 2017. Developing data interoperability using standards: A wheat community use case, F1000 Research, 6–1843, October 2017. (In preparation).
Shrestha, R., Matteis, L., Skofic, M., Portugal, A., McLaren, G., Hyman, G., Arnaud2, E., 2012. Bridging the phenotypic and genetic data useful for integrated breeding through a data annotation using the Crop Ontology developed by the crop communities of practice. Frontiers Physiol. 3.
Coletta, R., Castanier, E., Valduriez, P., Frisch, C., Ngo, D., Bellahsene, Z., 2012. Public data integration with Websmatch. In: Raschia, G., Theobald, M., (Eds.), 1st International Workshop on Open Data, WOD'12, Nantes, France, pp. 5–12, ACM, May.
Castanier, E., Jonquet, C., Melzi, S., Larmande, P., Ruiz, M., Valduriez, P., 2014. Semantic

annotation workflow using bio-ontologies. In: Workshop on Crop Ontology and Phenotyping Data Interoperability, Montpellier, France, CGIAR, April.
Noy, N.F., Alexander, P.R., Harpaz, R., Whetzel, P.L., Fergerson, R.W., Musen, M.A., 2013. Getting lucky in ontology search: a data-driven evaluation framework for ontology ranking. In: 12th International Semantic Web Conference, ISWC'13, vol. 8218 of Lecture Notes in Computer Science, Sydney, Australia, pp. 444–459, Springer, October 2013.
Ghazvinian, A., Noy, N.F., Jonquet, C., Shah, N.H., Musen, M.A., 2009. What four million mappings can tell you about two hundred ontologies. In: Bernstein, A., Karger, D.R., Heath, T., Feigenbaum, L., Maynard, D., Motta, E., Thirunarayan, K., (Eds.), 8th International Semantic Web Conference, ISWC'09, vol. 5823 of Lecture Notes in Computer Science, Washington DC, USA, pp. 229–242, Springer, November 2009.
Faria, D., Jiménez-Ruiz, E., Pesquita, C., Santos, E., Couto, F.M., 2014. Towards annotating potential incoherences in bioportal mappings. In: 13th International Semantic Web Conference, ISWC'13, vol. 8797 of Lecture Notes in Computer Science, Riva del Garda, Italy. Springer, pp. 17–32.
Pathak, J., Chute, C.G., 2009. Debugging mappings between biomedical ontologies: preliminary results from the NCBO BioPortal mapping repository. In: Smith, B., (Ed.), International Conference on Biomedical Ontology, Buffalo, NY, USA, pp. 95–98, July.
Ghazvinian, A., Noy, N.F., Musen, M.A., 2009. Creating mappings for ontologies in biomedicine: simple methods work. In: American Medical Informatics Association Annual Symposium, AMIA'09, Washington DC, USA, pp. 198–202, November.
Noy, N.F., Dorf, M., Griffith, N.B., Nyulas, C., Musen, M.A., 2009. Harnessing the power of the community in a library of biomedical ontologies. In: Clark, T., Luciano, J.S., Marshall, M.S., Prud'hommeaux, E., Stephens, S., (Eds.), Workshop on Semantic Web Applications in Scientific Discourse, SWASD'09, vol. 523 of CEUR Workshop Proceedings, Washington DC, USA, pp. 11, CEUR-WS.org, November.
Noy, N.F., Griffith, N.B., Musen, M.A., 2008. Collecting community-based mappings in an ontology repository. In: Sheth, A.P., Staab, S., Dean, M., Paolucci, M., Maynard, D., Finin, T.W., Thirunarayan, K. (Eds.), 7th International Semantic Web Conference, ISWC'08, vol. 5318 of Lecture Notes in Computer Science, Karlsruhe, Germany, October. Springer, pp. 368–371.
McCray, A.T., 2003. An upper-level ontology for the biomedical domain. Comp. Funct. Genomics 4, 80–84.
Dai, M., Shah, N.H., Xuan, W., Musen, M.A., Watson, S.J., Athey, B.D., Meng, F., 2008. An efficient solution for mapping free text to ontology terms. In: AMIA Symposium on Translational BioInformatics, AMIA-TBI'08, San Francisco, CA, USA, March.
Martinez-Romero, M., Jonquet, C., O'Connor, M.J., Graybeal, J., Pazos, A., Musen, M.A., 2017. NCBO ontology recommender 2.0: an enhanced approach for biomedical ontology recommendation. Biomed. Semantics 8 (21).
Jonquet, C., Musen, M.A., Shah, N.H., 2010. Building a biomedical ontology recommender web service, biomedical semantics, vol. 1. Selected in Pr. R. Altman's 2011 Year in Review at AMIA TBI.
Salvadores, M., Horridge, M., Alexander, P.R., Fergerson, R.W., Musen, M.A., Noy, N.F., 2012. Using SPARQL to query bioportal ontologies and metadata. In: Cudré-Mauroux, P., Heflin, J., Sirin, E., Tudorache, T., Euzenat, J., Hauswirth, M., Parreira, J., Hendler, J., Schreiber, G., Bernstein, A. (Eds.), 11th International Semantic Web Conference, ISWC'12, vol. 7650 of Lecture Notes in Computer Science, Boston, MA, USA, November. Springer, pp. 180–195.
Maguire, E., González-Beltrán, A., Whetzel, P.L., Sansone, S.-A., Rocca-Serra, P., 2013. OntoMaton: a Bioportal powered ontology widget for Google spreadsheets. Bioinformatics 29, 525–527.
Jupp, S., Welter, D., Burdett, T., Parkinson, H., Malone, J., 2015. Collaborative ontology development using the webulous architecture and google app. In: Malone, J., Stevens, R., Forsberg, K., Splendiani, A., (Eds.), 8th International Conference on Semantic Web Applications and Tools for Life Sciences, SWAT4LS'15, vol. 1546, Cambridge, UK, pp. 120–121, December.
Wolstencroft, K., Horridge, M., Owen, S., Mueller, W., Bacall, F., Snoep, J., Krebs, O., Goble, C., 2010. Rightfield: embedding ontology term selection into spreadsheets for the annotation of biological data. In: Polleres, A., Chen, H., (Eds.), 9th International Semantic Web Conference, Posters & Demonstrations, ISWC'10, vol. 658 of CEUR Workshop Proceedings, Shanghai, China, pp. 141–144, November.
Roeder, C., Jonquet, C., Shah, N.H., Jr, W.A.B., Hunter, L., 2010. A UIMA wrapper for the NCBO annotator. Bioinformatics 26, 1800–1801.
Adamusiak, T., Burdett, T., van der Velde, K.J., Abeygunawardena, N., Antonakaki, D., Parkinson, H., Swertz, M., 2010. OntoCAT – a simpler way to access ontology resources. In: ISMB Conference, Poster session, ISMB'10, Nature Precedings, Boston, MA, USA, July.
Miñarro-Giménez, J.A., Mikel, J.T.F.-B., Aranguren, E., Antezana, E., 2012. NCBO-galaxy: bridging the BioPortal web services and the Galaxy platform. In: Cornet, R., Stevens, R., (Eds.), 3rd International Conference on Biomedical Ontologies, ICBO'12, vol. 897 of CEUR Workshop Proceedings, Graz, Austria, pp. 2, July.
Falconer, S.M., Callendar, C., Storey, M.-A., 2009. FLEXVIZ: visualizing biomedical ontologies on the web. In: Smith, B., (Ed.), International Conference on Biomedical Ontology, ICBO'09, Buffalo, NY, USA, pp. 2, July.
Slater, L., Gkoutos, G.V., Schofield, P.N., Hoehndorf, R., 2016. Using AberOWL for fast and scalable reasoning over BioPortal ontologies. Biomed. Semantics 7, 49.
Melzi, S., Jonquet, C., 2014. Scoring semantic annotations returned by the NCBO Annotator. In: Paschke, A., Burger, A., Romano, P., Marshall, M., Splendiani, A., (Eds.), 7th International Semantic Web Applications and Tools for Life Sciences, SWAT4LS'14, vol. 1320 of CEUR Workshop Proceedings. Berlin, Germany, pp. 15, CEUR-WS.org, December.
Park, J., Oh, S., Ahn, J., 2011. Ontology selection ranking model for knowledge reuse. Expert Syst. Appl. 38 (5), 5133–5144.
Toulet, A., Emonet, V., Jonquet, C., 2016. Modèle de métadonnées dans un portail d'ontologies. In: Diallo, G., Kazar, O., (Eds.), 6èmes Journées Francophones sur les Ontologies, JFO'16, Bordeaux, France, October. Best paper award.
Ngo, D., Bellahsene, Z., 2012. YAM++: a multi-strategy based approach for ontology matching task. In: ten Teije, A., Völker, J., Handschuh, S., Stuckenschmidt, H., d'Acquin, M., Nikolov, A., Aussenac-Gilles, N., Hernandez, N. (Eds.), 18th International Conference on Knowledge Engineering and Knowledge Management,EKAW'12, vol. 7603 of Lecture Notes in Computer Science, Galway City, Irland, October. Springer, pp. 421–425.
Pesce, V., Maru, A., Keizer, J., 2011. The CIARD RING, an infrastructure for interoperability of agricultural research information services. Agric. Inf. Worldwide 4 (1), 48–53.
Baker, T., Caracciolo, C., Arnaud, E., 2016. Global agricultural concept scheme (GACS)a hub for agricultural vocabularies. In: Jaiswal, P., Hoehndorf, R., (Eds.), 7th International Conference on Biomedical Ontologies, ICBO'16, Poster Session, vol. 1747 of CEUR Workshop Proceedings, Corvallis, Oregon, USA, pp. 2, August.
McGuinness, D.L., 2003. Spinning the Semantic Web: Bringing the World Wide Web to its Full Potential, ch. Ontologies Come of Age. MIT Press, pp. 171–194.