



Background

The **advancement** of empirical and information **technologies** in the recent years have drastically **increased** the amount of **data** in fields of Life Sciences, and **Agronomic Sciences**. To **understand** the **complexity** of a given system it is **important** to **link** (integrate) diverse **datasets**. A promising **solution** towards data **integration** challenges is offered by the **Semantic Web** technologies¹. The Semantic Web was proposed, to **remedy** the **fragmentation** of all potentially useful **information** on the **web**. Currently, the **bio-medical** domain has **accepted** the **Semantic Web** technologies as a means to **manage** (integrate) **knowledge**. Although we are witnessing an increased usage of ontologies within the **Agronomic Sciences**, the **data** in this domain is **highly distributed** in nature. **Utilizing** these data **resources** more effectively and taking **advantage** of associated **cross-disciplinary** research **opportunities** poses a major **challenge** to both domain **experts** and **information technologists**.

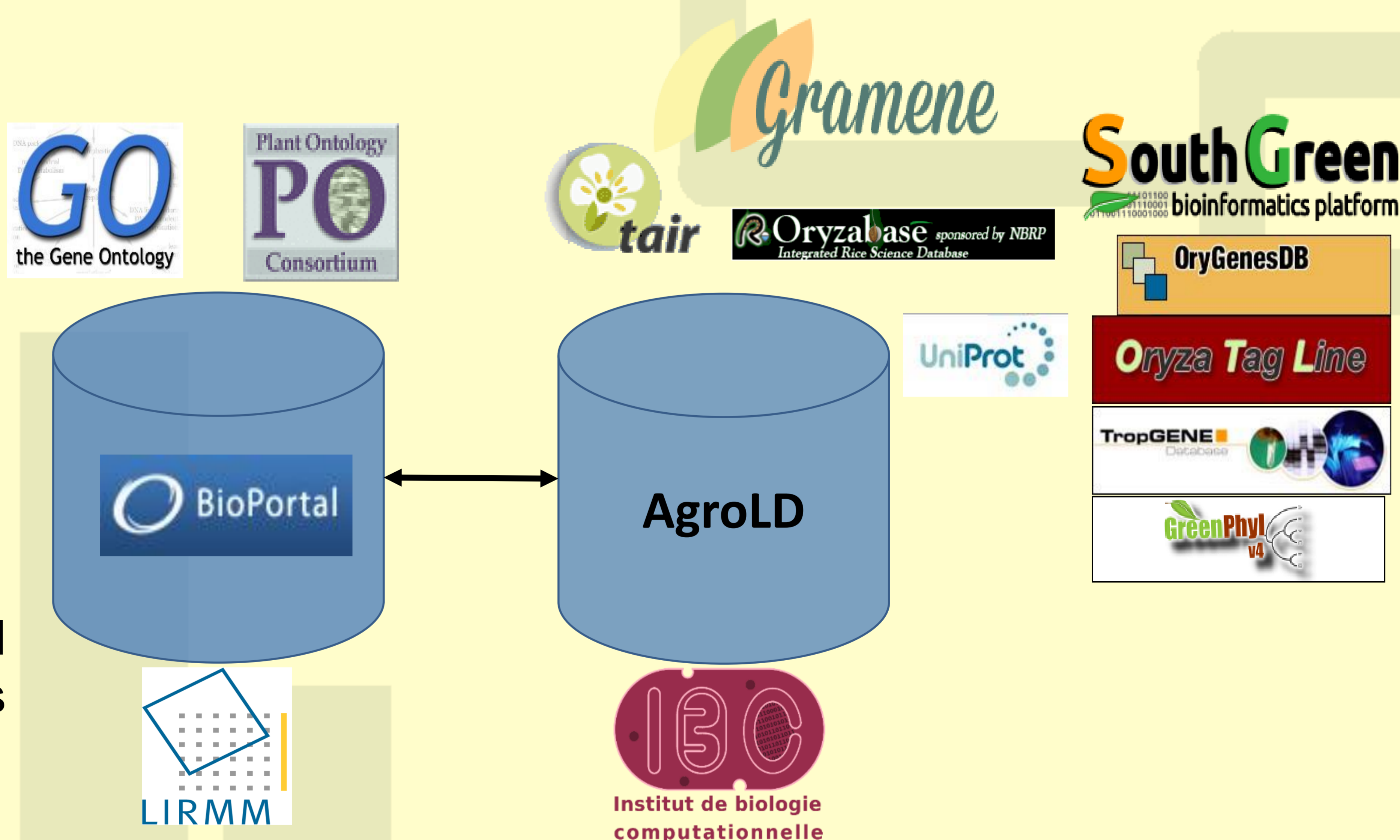
Agronomic Linked Data (AgroLD)

Aim

- Build an **RDF** knowledge base to house **data sources** pertain to **plant data**.
- Enable **answering** of **complex** domain relevant **questions** that were **unapproachable** using **traditional** methods.

Current status

- A **Plant specific instance** of **BioPortal** for managing and browsing ontologies has been **set-up**.
- **Data integration** pipeline for modeling the **various** domain specific **resources** into **RDF** (work in progress).
- **AgroLD** is being developed in stages:
 - **Stage 1**: **Data** (molecular level) pertaining to ***A.thaliana*** and ***Oryza* sps.** exposed as **RDF**.
 - **Stage 2**: The **knowledge base** will be **expanded** to incorporate other important species such as ***Wheat***.



Outlook to the future

- Integration of more data sources (e.g.: PPI, germplasm, and gene markers).
- Collaboration with **biologists** and **bioinformaticians** to provide proof of concept.
 - **Pluggable** with **workflow** systems e.g: **Galaxy**² and **VirtualPlants**³.
 - Work with **plant biologists** to construct **complex queries** for **hypothesis generation** and **validation** (wet lab).

We are open to discussions and collaborations. Feel free to get in touch:

Dr. Aravind Venkatesan: aravind.venkatesan@lirmm.fr

Dr. Pierre Larmande: peirre.larmande@ird.fr

¹ Berners-Lee, T. and Hendler, J. (2001). 'Publishing on the semantic web'. *Nature*, 410, 1023-4.

² Giardine B, et al., Galaxy: a platform for interactive large-scale genome analysis. *Genome Res* 2005, 15:1451-1455.

³ Katari MS, et al. VirtualPlant: A software platform to support Systems Biology research. *Plant Physiol.* 2010;152(2):500-515.