We aim to develop an accurate and complex knowledge model of Arabidopsis regulatory networks to describe regulation networks in other plants and/or organisms.

**Arabidopsis knowledge model**

- Collaboration between information extraction specialists and biology experts
- Fined-grained and complex model: 16 biological entities and 10 relations (Figure 2)

**Motivation**

- *Arabidopsis thaliana* is a plant model in biology
- Understanding the molecular network underlying seed development regulations is important for fundamental research, agriculture, and industry
- Regulatory networks for reserve accumulation and seed maturation are complex (Figure 1)
- Modeling the overall seed development process requires the reconstruction of regulatory networks at different scales (e.g., genetic level, environmental factors and phenotypes)
- A large part of the information needed is spread throughout thousands of articles

We propose a knowledge model of *Arabidopsis* regulatory networks to extract information from text with Natural Language Processing and Machine Learning.

**Annotated sentence**

*FUS3* and *LEC2* are implicated in repression of GA biosynthesis during seed development.

**Conclusion & Perspectives**

- An accurate and complex knowledge model was developed to describe seed development of *Arabidopsis thaliana*
- We are currently training Machine-Learning methods using the *Arabidopsis thaliana* corpus
- We aim at proposing for the first time an annotation task on *Arabidopsis thaliana* at BioNLP Shared Task 2016
- We plan to use the knowledge model to describe regulation networks in other plants and/or organisms

**References**


**Figure 1**: A model of genetic (framed) and molecular (in blue cyan) interactions involved in the control of seed development and maturation in Arabidopsis thaliana (from Santos-Mendoza et al., 2008) (Numbers are biological processes).

**Figure 2**: Schematic representation of the knowledge model.

**Figure 3**: Annotation with AlvisAE.

**Figure 4**: Example of an annotated sentence.