Expert knowlege integration for smart tools modeling: application to wheat use

MUNCH M. (1), KANSOU K. (2), BAUDRIT C. (1)
1 Universitde Bordeaux - I2M, Bordeaux, France
2 INRAE BIA, Nantes, France

The quality of wheat grain is generally determined with regard to its potential value for breadmaking through selected technological tests. With growing constraints affecting the production of wheat worldwide, it is critical to optimize the use value of wheat, in particular wheat produced locally. A proper evaluation of the wheat quality should associate a multifactorial characterization of the wheat with a specific use in the industry, e.g. for French bread, rye bread, biscuits and so on.

This work is part of the ANR Evagrain project, whose ambition is to design a decision support system (DSS) capable of computing an evaluation of the wheat use value from wheat grain characteristics. The evaluation takes the form of a prediction of the end-product properties. With this in mind, a database describing the characteristics of more than 100 wheat grains produced in France has been compiled by public and industrial laboratories of the cereal science field.
To annotate and structure the data with a standard vocabulary, we developed a wheat quality ontology through experts' interviews. Our ontology is based on @Web, a core ontology for the annotation of heterogeneous data sources.

More specifically, our ontology is structured so that to annotate each measurement of a specific wheat with the method (which device, protocol, etc.) and the sample information (origin, type, operator, variety etc.). This ontology will allow to feed automatically the DSS with upcoming data all along the project and beyond. It is also an essential medium to query and compare heterogeneous datasets by supplying all the necessary metadata.

Using this newly structured dataset, it becomes possible to learn complex models able to show the links between the different variables of interest. Moreover, by integrating expert knowledge, we aim to provide explanations for these links. To do so, we translate expert inputs as constraints used for guiding the learning of a Bayesian network (BN). As probabilistic graphical models, BNs are well suited to deal both with uncertainty (which is common in technological tests where measures can be imprecise and external factors have to be considered), and allow complex reasoning such as predictions and reverse-engineering.