

Création d'un système d'information pour visualiser et administrer les ontologies ATOL, EOL et AHOL

Matthieu Reichstadt, Alexandre Journaux, Catherine Hurtaud, Marie-Christine Meunier Salaün, Gaëlle Maxin, Alice Fatet, Pierre-Yves Le Bail, Jérôme Bugeon

Livestock Ontologies est un projet regroupant la construction et la diffusion de 3 ontologies, ATOL (Animal Trait Ontology for Livestock), EOL (Environment Ontology for Livestock) et AHOL (Animal Health Ontology for Livestock). Le projet ATOL a été initié en 2009 et mis pour la première fois en ligne en 2012 (<https://www.atol-ontology.com/>). Il regroupe un ensemble de 7 chercheurs INRAE, responsables chacun d'une des branches de l'ontologie des caractères phénotypiques des animaux (bien-être, croissance et qualité des produits carnés, foie gras, glande mammaire et production de lait, œuf, nutrition, reproduction et fertilité).

Le projet EOL traite de toute la partie environnement et systèmes d'élevage des animaux de rente. Basé sur le même concept qu'ATOL, il comporte 4 branches principales (environnement d'élevage, structure d'élevage, système d'élevage, alimentation du bétail).

Enfin le projet AHOL traite de toute la partie santé des animaux de rente. Son organisation est différente des deux autres puisque AHOL regroupe 3 ontologies (maladies, agents pathogènes, symptômes), qui sont reliées entre elles autour de la maladie (une maladie fait entrer en jeu des agents pathogènes et génère des symptômes).

Ces ontologies étaient initialement basées sur l'utilisation de fichiers Excel, puis du logiciel Protégé collaboratif (<https://protege.stanford.edu/>). L'équipe du projet a réalisé que l'évolution de ces ontologies passait par une refonte totale du système et a exprimé le besoin d'un outil permettant un travail collaboratif et une mise en ligne rapide des mises à jour. Ainsi à partir de 2017, nous avons mis en place un nouveau Système d'Informations (SI) afin de traiter les 3 ontologies en parallèle.

Ce SI se découpe en réalité en 2 parties :

- Une partie visualisation (accessible depuis le site web du projet)
- Une partie administration, elle-même découpée en trois parties (une par ontologie)

Ceci simplifiait à la fois l'utilisation et l'implémentation des ontologies par les scientifiques associés au projet, mais permettait également d'ajouter une couche de sécurité et un aspect collaboratif à ces applications. De plus, nous avons estimé que l'ouverture de la visibilité des ontologies aux autres systèmes d'information développés dans notre CATI (SICPA) apporterait une grande plus-value aux travaux réalisés. Ainsi il a été mis en place une liaison entre AHOL et l'application SICPA Sanitaire.

Le principe de notre SI est qu'il est entièrement développé en PHP/Symfony, associé à une base de données mysql. Il est ainsi entièrement orienté web.

La partie « visualisation » permet à toute personne de visualiser l'intégralité des différentes ontologies, et donne accès à l'arborescence de chacun d'entre elles. En sélectionnant un des traits, l'utilisateur a accès à toutes les informations présentes en base de données (définition, espèces concernées, méthodes de mesure, synonymes, parentalité). Il peut également télécharger l'ontologie ou une sous-branche de celle-ci en différents formats (SQL, OWL) afin de pouvoir travailler en toute indépendance dessus.

La partie « administration » de chaque ontologie permet aux curateurs (les scientifiques responsables des différentes branches des ontologies) de gérer leur branche en toute autonomie, d'importer des éléments depuis des fichiers annexes, ou d'exporter leurs données sous différents formats. Elle est sécurisée et installée sur le serveur de notre CATI. Nous avons mis également en place un système de versioning des ontologies afin de garder une trace de toute modification effectuée.