

Stage de Master 2

Navigation dans les règles d'implication extraites de connaissances agroécologiques en santé animale et végétale pour l'aide à la décision

Encadrants

Marianne Huchard (marianne.huchard@lirmm.fr)

Pierre Martin (pierre.martin@cirad.fr)

Arnaud Sallaberry (Arnaud.Sallaberry@lirmm.fr)

Pierre Silvie (pierre.silvie@cirad.fr)

Pascal Poncelet (Pascal.Poncelet@lirmm.fr)

*Navigation in implication rules extracted from agroecological knowledge
in animal and plant health for decision making*

Durée: **6 mois (+ 1 mois CDD possible suite au stage)**

Date début de stage : **février 2022**

Laboratoire d'accueil : **LIRMM, CNRS, Univ. Montpellier**

Financé dans le cadre de l'appel #Digitag

<https://www.hdigitag.fr/fr/aap-masters2-vos-sujets-avant-le-18-octobre-2021/>

Résumé du sujet en français

Pour un producteur agricole, implémenter les pratiques agroécologiques au sein de son exploitation lui impose de disposer d'un système d'aide à la décision (SAD) pour les identifier, ce SAD comportant une base de connaissances suffisamment étoffée et un système de navigation adapté à ses besoins. La base de connaissances Knomana, par exemple, rassemble plus de 46000 descriptions d'usage de plantes à effet pesticide et antibiotique en santé végétale, animale et humaine (Silvie et al. 2021). La plateforme de visualisation RCaviz (<https://info-demo.lirmm.fr/rcaviz/>) permet de naviguer dans ce type de base de connaissances. Basé sur l'Analyse de Concepts Relationnels (RCA), une méthode de classification de données relationnelles, RCaviz permet de naviguer dans des structures conceptuelles de type graphe et d'identifier facilement, par exemple, une plante locale susceptible de protéger une culture contre un bioagresseur invasif, ou des plantes en partie équivalentes pour un problème sanitaire donné.

Outre les structures conceptuelles, RCA propose de représenter les connaissances sous forme de règles d'implication, un formalisme qui, proche du langage naturel, est bien adapté aux utilisateurs du monde rural (e.g. « F_Meliaceae => no-food » : les plantes de la famille Meliaceae ne sont pas consommées). Les évaluations conduites pour l'approche One Health (i.e. Mahrach et al. 2021 ; Saoud et al. 2021) ont démontré la viabilité et l'utilité de cette solution. Pour autant, identifier la règle la mieux appropriée dans un grand ensemble est difficile.

L'objectif du stage est de développer un prototype logiciel de visualisation de connaissances, exprimées sous forme de règles d'implications produites par la librairie FCA4J (<http://www.lirmm.fr/fca4j>). Cette application présentera les règles en regard de mesures

d'intérêt ou d'après une formulation symbolique donnée par l'utilisateur, e.g. les règles relatives à un certain ensemble de conditions, pour facilement les exploiter.

English version

For a farmer, implementing agroecological practices on his farm requires having a decision support system (DSS) allowing him to identify them. The development of such a DSS requires a sufficiently extensive knowledge base and a knowledge navigation system (SN) adapted to her/his needs, allowing to consider innovative practices, where appropriate.

The Knomana knowledge base, for instance, brings together more than 46,000 descriptions of usage of pesticides and antibiotics plants in plant, animal and human health (Silvie et al. 2021). The visualization platform RCAviz (<https://info-demo.lirmm.fr/rcaviz/>) has been developed to navigate in this type of knowledge base. Based on Relational Concept Analysis (RCA), a classification method dedicated to relational data, RCAviz makes it possible to navigate graph-type conceptual structures and easily identify, for example, a local plant capable of protecting a crop against an invasive pest, or partly equivalent plants to solve a given sanitary problem. In addition to conceptual structures, RCA enables to represent knowledge as implications rules, a formalism which, by approaching natural language, is better suited to users in the rural world (e.g. « F_Meliaceae => no-food » : Meliaceae family plants are not consumed as food). The evaluations conducted within the One Health approach (i.e. Mahrach et al. 2021; Saoud et al. 2021) have demonstrated the viability and usefulness of this solution.

The objective of the internship is to develop a software prototype for visualizing knowledge, expressed in the form of rules of implications produced by the FCA4J library (<http://www.lirmm.fr/fca4j>). This application will allow the rules to be presented in relation to measures of interest or according to a symbolic formulation given by the user, for example the rules relating to a certain set of conditions. This will definitively allow the user to easily exploit them.

Mots clés en français

Agroécologie, usage des plantes, Analyse de Concepts Formels, Règle d'implication, Visualisation

Agroecology, plant use, Formal Concept Analysis, Implication rule, Visualization

Profil attendu

Etudiant de Master 2 (informatique ou bioinformatique) ayant des compétences solides en programmation d'application web (notamment javascript) et en analyse de données, ainsi qu'un intérêt pour l'ingénierie des connaissances, l'analyse visuelle (visual analytics), et à trouver des solutions alternatives aux pesticides et antibiotiques de synthèse pour l'agriculture biologique.

-

Student in Master studies (computer science or bioinformatics) with strong skills in programming web applications (including javascript) and data analysis, with an interest for knowledge engineering, visual analytics, and to find alternatives to chemical pesticides and antibiotics in organic agriculture.