

# Sujet de stage de master recherche 2ème année: Définition d'un langage pour représenter les évolutions d'une ontologie

Contacts : [dibie@agroparistech.fr](mailto:dibie@agroparistech.fr), [liliana.ibanescu@agroparistech.fr](mailto:liliana.ibanescu@agroparistech.fr), [fatiha.sais@lri.fr](mailto:fatiha.sais@lri.fr)

Le contexte méthodologique de ce stage concerne l'enrichissement d'entrepôts de données thématiques ouverts sur le Web de données (Linked open data). Un moyen pour enrichir ces entrepôts est d'exploiter des documents à l'aide de technologies du Web Sémantique. Nous proposons pour cela d'utiliser une ontologie avec laquelle les données seront annotées. Dans son cycle de vie, une ontologie peut être amenée à évoluer pour diverses raisons : prendre en considération de nouvelles données qui deviennent pertinentes dans un domaine d'application, tirer partie de nouvelles ontologies mises à disposition sur le Web concernant le domaine d'application, prendre en compte de nouveaux besoins des utilisateurs ou encore améliorer la qualité des annotations qui dépend de la qualité de l'ontologie.

Nous nous intéressons dans ce stage à l'évolution de cette ontologie et plus précisément à la définition d'un langage de représentation simple et facilement compréhensible qui permet : (1) de caractériser les différentes évolutions possibles d'une ontologie qui passe alors d'un état à un autre et (2) d'identifier les répercussions de ces évolutions sur les instances de l'ontologie, comme par exemple la base d'annotations. Le stage portera plus précisément sur l'alignement d'une ontologie dédiée à la représentation de relations n-aires entre des données quantitatives expérimentales. Cette ontologie se caractérise par ses concepts fortement interdépendants car reliés entre eux par des relations n-aires. Les différentes approches présentées dans la littérature qui s'intéressent à l'évolution d'ontologie s'appuient sur la notion de changements qu'ils présentent sous la forme d'opérations élémentaires ou composés à appliquer sur les différentes entités de l'ontologie. Aucun langage n'a, à notre connaissance, été proposé pour décrire simplement les évolutions possibles des entités d'une ontologie et de ses instances. C'est l'objectif de ce stage.

Ce stage sera réalisé dans le cadre d'une collaboration entre deux équipes en informatique, l'équipe LaHDAK du LRI de l'Université Paris Sud (Orsay) qui travaille en particulier sur la réconciliation de référence, et, l'équipe « Ingénierie des connaissances » de l'unité INRA Mét@risk qui travaille en particulier sur l'évolution d'ontologie et l'annotation de données.

Les étapes du travail seront les suivantes:

1. Etude bibliographique sur l'évolution d'ontologies
2. Proposition d'un langage simple et semi formel pour décrire les évolutions possibles des entités d'une ontologie
3. Extension du langage pour décrire les répercussions des évolutions des entités d'une ontologie sur ses instances
4. Evaluation du langage en l'utilisant pour décrire les évolutions possibles d'une ontologie sur l'évaluation des risques microbiologiques dans les aliments emballés et sur sa base de données annotés

## Bibliographie

Kondylakis, H., G. Flouris, et D. Plexousakis (2009). Ontology and Schema Evolution in Data Integration: Review and Assessment. In OTM Conferences (2), pp. 932–947.  
Buche P., Dibie-Barthélemy J., Ibanescu L., Soler L. Fuzzy Web Data Tables Integration Guided by a Terminology-Ontological Resource. IEEE TKDE 25(4): 805-819 (2013).

Lieu du stage: LRI – Université Paris Sud (Orsay), durée de 6 mois, stage rémunéré.