



■ mOeX : Évolution de la connaissance

LIG/mOeX
INRIA & Univ. Grenoble Alpes
<https://moex.inria.fr>

Jérôme EUZENAT
Jerome.Euzenat@inria.fr

Membres

- Manuel ATENCIA, MCF Univ. Grenoble Alpes
- Jérôme DAVID, MCF Univ. Grenoble Alpes
- Jérôme EUZENAT, DR INRIA

Thématique générale de l'équipe

mOeX travaille dans le domaine de la représentation des connaissances et plus spécifiquement de la connaissance vivante, c'est-à-dire hétérogène et évolutive.

Nos travaux ont longtemps été motivés par le développement du web sémantique. Nous avons en particulier fortement contribué à la mise en correspondance d'ontologies [4]. Aujourd'hui, mOeX s'intéresse spécifiquement à deux sujets :

- le liage des données et en particulier l'étude des clés de liage,
- l'évolution culturelle de la connaissance.

Liage de données et clés de liage

Le besoin d'accès aux données par la société a conduit à la publication, par des acteurs divers (gouvernement, universités, acteurs culturels), de vastes corpus de données exprimées dans les formalismes du web sémantique (principalement RDF). Une part importante de la valeur ajoutée des données liées réside dans les liens identifiant la même entité dans différents jeux de données. Par exemple, cela permet d'identifier les mêmes ouvrages dans différentes sources de données bibliographiques. Les liens permettent d'exploiter conjointement les données de ces sources.

Nous appelons la tâche consistant à extraire ces liens, liage de données (data inter-linking). Elle consiste, à partir de deux sources de données en RDF, à engendrer un ensemble de liens entre identifiants des deux sources faisant référence à la même entité. Elle est pilotée par une spécification de liage. Différents types de spécifications sont disponibles. La plus répandue consiste à calculer une distance entre les descriptions des ressources et à estimer que plus ils sont proches, plus ils ont de chance d'identifier la même ressource.

Un autre type de spécification est ce que nous appelons 'clé de liage' (*link key*). Les clés de liage généralisent les clés de bases de données dans deux directions : elles fonctionnent avec des données représentés en RDF, et elles s'appliquent à deux jeux de données indépendants.

Un exemple de clé de liage est :

$\{\langle \text{auteur}, \text{creator} \rangle\} \{\langle \text{titre}, \text{title} \rangle\} \times \langle \text{Livre}, \text{Book} \rangle$

qui signifie que si deux instances des classes *Livre* et *Book* respectivement, ont les mêmes valeurs pour les propriétés *auteur* et *creator* et au moins une valeur commune pour les propriétés *titre* et *title*, alors elles dénotent la même ressource.

On s'intéresse principalement à deux opérations sur les clés de liage : leur extraction à partir de deux sources de données, et le raisonnement prenant en compte les clés de liages comme axiomes.

Extraction de clés de liage Une première méthode a été conçue pour extraire les clés de liage entre deux classes [1]. Elle commence par



extraire des clés candidates puis les évalue à l'aide de mesures adaptées.

Nous avons reformulé l'étape d'extraction de clés candidates en un problème d'analyse formelle de concepts (FCA) pour le cas simple des bases de données. Nous avons étendu ce travail pour prendre en compte les attributs non fonctionnels et les dépendances entre clés de liage (lorsque les conditions d'une clé nécessitent de déterminer l'égalité de deux instances d'autres classes, ce qui utilise une autre clé. Pour prendre en compte les références circulaires, il est devenu nécessaire d'adapter les techniques d'analyse relationnelle de concepts (RCA) au cas des clés de liage [2].

Nous travaillons aussi à l'élaboration d'algorithmes de recherche pour des clés de liage composées ou incluant de expressions de classes et de propriétés non atomiques. Cela augmente la complexité du problème à résoudre. En ce qui concerne l'évaluation des clés obtenues, nous développons de nouvelles mesures de qualité.

Raisonnement avec des clés de liage Les clés de liages peuvent être considérées comme des axiomes d'une logique de description. Il est alors possible de les utiliser pour déduire de nouveaux liens, de nouvelles assertions voire d'autres clés de liage. Pour ce faire, nous avons étendu la méthode des tableaux appliquée à ces logiques, en particulier *ALC*.

Une perspective intéressante de ces travaux est d'être capable d'utiliser les procédures de raisonnement lors de l'extraction.

Collaborations Ces travaux sont développés dans le cadre du projet ANR [Elker](#) en collaboration avec les équipes LORIA/Orpailleur (Amedeo Napoli, Miguel Couceiro), LIASD (Myriam Lamolle, Chan Le Duc) et LIG/Slide (Marie-Christine Rousset). Nous collaborons aussi dans le cadre du PEPS RegleX-LD avec

IRIT/Melodi (Cássia Trojahn, Élodie Thiéblin), LRI/LaHDAK (Nathalie Pernelle, Fatiha Saïs) et INRA/LInK (Liliana Ibanescu).

Évolution culturelle de la connaissance

mOeX aborde l'évolution des représentations de connaissance chez les individus et les populations. Notre ambition est de répondre en particulier aux questions suivantes :

- Comment des populations d'agents adaptent-elles leurs connaissances à leur environnement et aux autres populations ?
- Comment cette connaissance doit-elle évoluer lorsque l'environnement change et que de nouvelles populations se rencontrent ?
- Comment les agents peuvent-ils préserver la diversité des connaissances et cette diversité est-elle bénéfique ?

Nous les étudions principalement dans un contexte informatique contrôlé.

À cette fin, nous combinons des méthodes de représentation des connaissances et d'évolution culturelle expérimentale. Les premières fournissent des modèles formels de la connaissance ; les secondes un cadre bien défini pour étudier l'évolution située.

Nous considérons la connaissance comme une culture et étudions les propriétés globales des opérateurs d'adaptation locale appliquées par des populations d'agents en :

- testant expérimentalement les propriétés des opérateurs d'adaptation dans diverses situations en utilisant l'évolution culturelle expérimentale, et
- déterminant théoriquement ces propriétés en modélisant la manière dont les opérateurs façonnent la représentation de la connaissance.

Simulation de l'évolution de la connaissance

Nous simulons l'évolution d'alignements entre



ontologies utilisées par des agents pour communiquer. Les agents ne connaissent que leurs ontologies et leurs alignements avec les autres et agissent de manière totalement décentralisée. Ils jouent à un jeu dans lequel ils utilisent les alignements pour reclassifier des objets d'une ontologie à une autre. Lorsque les agents détectent une erreur de reclassification ils réagissent en appliquant un opérateur d'adaptation à l'alignement utilisé.

De telles expériences de réparation ont révélé que des agents dotés d'opérateurs très simples peuvent effectivement réparer des réseaux aléatoires d'ontologies [5, 3]. En faisant varier les modalités d'application des opérateurs (expansion, relation, génération, renforcement), il est possible s'approcher des alignements idéaux entre les ontologies et de partir d'alignements vides.

Nous développons ce type d'expériences avec des populations d'agents différentes (caractérisées, par exemple, par les ontologies qu'elles utilisent). Cela nous permet d'étudier des phénomènes comme la synchronisation entre agents de la même population ou la transmission verticale de connaissance entre générations d'agents.

Modèles théoriques Nous développons une logique épistémique dynamique afin de saisir la dynamique du jeu de réparation d'alignement. Les ontologies sont modélisées en tant que connaissances et les alignements en tant que croyances. La dynamique du jeu est réalisée par des annonces (publiques) et les opérateurs d'adaptation par des mises à jour conservatives, c'est-à-dire des modalités qui transforment les modèles. Cela nous a permis d'établir formellement certaines limites logiques des opérateurs proposés, en particulier leur redondance pour un raisonneur logique complet. Ces résultats sont valables pour un agent du jeu mais pas nécessairement pour l'autre qui peut

ne pas connaître les classes par lesquelles l'alignement est réparé, ni les relations entre-elles. Le premier problème peut être traité en déclarant que les agents connaissent la signature des deux ontologies (hypothèse de signature publique), mais cela ne permet pas aux ontologies d'évoluer. Nous étudions actuellement la sémantique partielle comme une solution plus dynamique à ce problème.

Méthodologie expérimentale Nous décrivons nos expériences afin que d'autres et nous-mêmes puissions les répéter. Elles sont exécutées dans notre logiciel *Lazy lavender* qui utilise des scripts pour les spécifier, exécuter et analyser. Bien que ces descriptions soient encore ad hoc, c'est un premier pas vers l'exécution et l'analyse automatique des expériences. En particulier, *Lazy lavender* engendre une spécification de conteneur Docker pouvant rejouer les expériences dans les mêmes conditions ou après une mise à jour de la plateforme. La publication des expériences sur notre [wiki](#) est également facilitée par ce processus.

Collaborations Ce travail est jusqu'ici développé indépendamment, mais la chaire sur l'évolution de la connaissance de l'institut 3IA de Grenoble ([MIAI](#)) devrait permettre de renforcer les liens avec d'autres équipes locales (Tyrex : Pierre Genevès et Nabil Layaida ; Yves Demazeau) et étrangères.

Références

- [1] Manuel Atencia, Jérôme David, and Jérôme Euzenat. Data interlinking through robust linkkey extraction. In *Proc. 21st european conference on artificial intelligence (ECAI), Praha (CZ)*, pages 15–20, 2014.
- [2] Manuel Atencia, Jérôme David, Jérôme Euzenat, Amedeo Napoli, and Jérémy Vizini. Link key candidate extraction with re-



AfIA

Association française
pour l'Intelligence Artificielle

- lational concept analysis. *Discrete applied mathematics*, 2019. to appear.
- [3] Jérôme Euzenat. Interaction-based ontology alignment repair with expansion and relaxation. In *Proc. 26th International Joint Conference on Artificial Intelligence (IJCAI), Melbourne (VIC AU)*, pages 185–191, 2017.
- [4] Jérôme Euzenat and Pavel Shvaiko. *Ontology matching*. Springer, Heidelberg (DE), 2nd edition, 2013.
- [5] Jérôme Euzenat. First experiments in cultural alignment repair (extended version). In *Proc. ESWC 2014 satellite events revised selected papers*, number 8798 in Lecture notes in computer science, pages 115–130, 2014.