Évaluation automatique d'alignements complexes : une approche basée sur des instances

Elodie Thieblin, Ollivier Haemmerlé, Cassia Trojahn dos Santos

Évaluation automatique d'alignements complexes : une approche basée sur des instances

Elodie Thiéblin, Ollivier Haemmerlé, Cassia Trojahn
Institut de Recherche en Informatique de Toulouse

elodie@thieblin.fr, ollivier.haemmerle@irit.fr, cassia.trojahn@irit.fr

Résumé

Un alignement d'ontologies est un ensemble de correspondances entre les entités de différentes ontologies. La plupart des travaux sur l'évaluation d'alignements se sont concentrés sur l'évaluation d'alignements simples (i.e., dont les correspondances sont entre une seule entité de l'ontologie source et une seule entité de l'ontologie cible). L'émergence d'outils d'alignement capables de générer des alignements complexes (i.e., dont des correspondances incluent des constructeurs logiques ou des fonctions de transformation) a fait apparaître le besoin d'outils d'évaluation automatique de ces alignements. Ce papier propose i) un système d'évaluation automatique d'alignements complexes fondé sur des requêtes et des instances, et ii) un jeu de données sur l'organisation de conférences. Ce jeu de données est composé d'ontologies peuplées et d'un ensemble de questions de compétence pour alignement sous la forme de requêtes SPARQL. Des alignements de l'état de l'art sont évalués sur le jeu de données et les difficultés sur l'évaluation d'alignements sont discutées.

Mots-clés

alignement d'ontologies, alignement complexe, évaluation, jeu de données d'évaluation

Abstract

Ontology matching is the task of generating a set of correspondences (i.e., an alignment) between the entities of different ontologies. While most efforts on alignment evaluation have been dedicated to the evaluation of simple alignments (i.e., those linking one single entity of a source ontology to one single entity of a target ontology), the emergence of matchers providing complex alignments (i.e., those composed of correspondences involving logical constructors or transformation functions) requires new strategies for addressing the problem of automatically evaluating complex alignments. This paper proposes i) a benchmark for complex alignment evaluation composed of an Automatique evaluation system that relies on queries and instances, and ii) a dataset about conference organisation. This dataset is composed of populated ontologies and a set of competency questions for alignment as SPARQL queries. State-of-the-art alignments are evaluated and a discussion on the difficulties of the evaluation task is provided.

Keywords

ontology matching, complex alignment, evaluation, benchmark

1 Introduction

Cet article est une version française de [1]. Un alignement d'ontologies est un ensemble de correspondances entre les entités de différentes ontologies. Cela sert de base pour de nombreuses tâches comme l'intégration de données, l'évolution d'ontologies ou la réécriture de requêtes. Bien que ce champ de recherche se soit bien développé, la plupart des travaux se concentrent sur la génération de correspondances simples (i.e., lier une entité d'une ontologie source à exactement une entité d'une ontologie cible). Toutefois, les correspondances simples ne permettent pas de couvrir toute l'hétérogénéité des ontologies à aligner. Les correspondances complexes expriment des relations plus expressives entre les ontologies. Par exemple, l'information selon laquelle un article a été accepté dans une conférence peut s'exprimer par une classe *ekaw:Accepted_Paper* ou par une restriction sur la propriété *cmt:hasDecision* sur la classe *cmt:Acceptance*. La correspondance $\langle$*ekaw:Accepted_Paper,* $\exists$*cmt:hasDecision.cmt:Acceptance,* $\equiv$*, 1*$\rangle$ exprime l'équivalence entre les deux représentation d'un "article accepté" avec une confiance de 1.

Le besoin pour des alignements complexes a été décrit assez tôt [2, 3], et de nombreux outils de génération d'alignements complexes ont suivi, comme le présente l'état de l'art sur le sujet [4]. En revanche, peu d'initiatives se sont concentrées sur l'évaluation de ces alignements. La plupart des outils d'alignement complexe ont été évalués manuellement [5], en général seulement en termes de précision, ou sur des jeux de données spécifiques à leur outil [6], sur lesquels un rappel est calculé.

Bien que de nombreux efforts soient déployés sur l'évaluation automatique d'alignements, notamment dans les campagnes de l'Ontology Alignment Evaluation Initiative (OAEI) [1], la plupart se concentrent sur les alignements simples. Récemment, la première tâche d'alignement complexe a été proposée à l'OAEI [7], ouvrant de nouvelles perspectives à l'évaluation automatique d'alignements complexes.

1. http ://oaei.ontologymatching.org/

Dans cet articles, un jeu de données constitué d'ontologies aux instances contrôlées et partagées, ainsi qu'un ensemble de questions de compétences sous forme de requêtes SPARQL sont proposés, ainsi qu'un système d'évaluation automatique des alignements. Tandis que les outils et données d'évaluation d'alignements [8,9] se fondent sur des alignements de référence et mesurent la similarité entre l'alignement évalué et le référent, nous proposons un ensemble de questions de compétences pour alignements (CQA) comme référence. Une CQA exprime, via une requête SPARQL, la connaissance qu'un alignement devrait couvrir entre les ontologies source et cible [10]. Nous proposons deux métriques d'évaluation, la *couverture de CQA*, fondée sur des paires de requêtes SPARQL équivalentes mesure à quel point l'alignement évalué couvre les besoins des requêtes ; la *précision intrinsèque* compare les instances des membres de la correspondance. La précision intrinsèque équilibre la couverture de CQA comme la précision équilibre le rappel.

Cet article

- discute des défis de l'évaluation automatique d'alignements complexes par rapport aux approches d'évaluation existantes ;
- propose une approche automatique pour évaluer l'alignement complexe ;
- propose un jeu de données d'ontologies peuplées de manière contrôlée et des questions de compétence pour alignement associées ;
- évalue des alignements de l'état de l'art sur le jeu de données et discute des résultats.

Le système et le jeu de données sont publiés sous licence LGPL [2].

2 Contexte

2.1 Alignement complexe d'ontologies

Un alignement A lie deux ontologies : une ontologie source o et une cible o' [11]. A est directionnel, noté $A_{o\to o'}$. $A_{o\to o'}$ est un ensemble de correspondances $\langle e, e', r, n\rangle$. Chaque correspondance exprime une relation r (e.g., equivalence ($\equiv$), subsomption ($\sqsupseteq$, $\sqsubseteq$)) entre ses deux membre e et e', et n exprime le degré de confiance $[0..1]$ dans cette correspondance. Un membre peut être une entité simple de l'ontologie (classe, propriété sur les données, sur les objets ou individu) de respectivement o et o' ou une construction plus complexe composée d'entités et de constructeurs ou de fonctions de transformation. Nous considérons deux types de correspondances fondées sur leurs membres [12] :

- une correspondance est **simple** si e et e' sont des entités simples (representés avec une IRI) : $\langle$*ekaw:Paper, cmt:Paper,* $\equiv$*, 1*$\rangle$
- une correspondance est **complexe** si e et/ou e' implique un constructeur ou une fonction de transformation $\langle$*ekaw:Accepted_Paper,* $\exists$ *cmt:hasDecision.cmt:Acceptance,* $\equiv$*, 1*$\rangle$ et $\langle$*concatenation(edas:hasFirstName, " ", edas:hasLastName), cmt:name,* $\to$*, 1*$\rangle$

Une correspondance simple est notée (s:s). Une correspondance complexe peut-être (s:c), si son membre source est une entité simple, (c:s) si son membre cible est une entité simple, ou (c:c) si ses deux membres sont complexes.

2.2 Questions de compétence pour alignement (CQA)

En conception d'ontologies, les questions de compétences (CQ) ont été introduites comme *les besoins en connaissance sous la forme de questions auxquels l'ontologie doit pouvoir répondre* [13]. Comme définie dans [10], une question de compétence pour alignement (CQA) est une question de compétence qui devrait être ouverte par deux ontologies ou plus, i.e., elle exprime la connaissance qu'un alignement devrait couvrir (si les deux ontologies permettent elles-mêmes d'y répondre).

La première différence entre une CQA et une CQ est que le champ d'une CQA est limité par l'intersection des champs des ontologies source et cible. La seconde est que ce champ maximal et idéal d'un alignement n'est pas connu a priori puisqu'il est le but même de l'alignement.

Comme pour les CQ [14], une CQA peut être exprimée en langage naturel ou comme une requête SPARQL SELECT. Inspirée de la notion d'arité de prédicat [14], l'*arité d'une CQA* représente l'arité des réponses attendues à une CQA [10] :

- Une question *unaire* attend un ensemble d'instances ou valeurs, e.g., "Quels sont les articles acceptés ?" *(paper1), (paper2).*
- Une question *binaire* attend un ensemble de paires d'instances ou valeurs, e.g., "Quelle est la décision des articles ?" *(paper1, accept), (paper2, reject).*
- Une question *n-aire* attend un tuple de taille n, e.g., "Quelle est la décision associée à la relecture des articles ?" *(paper1, review1, weak accept), (paper1, review2, reject).*

3 État de l'art

L'évaluation de système d'alignement se fait sur un jeu de données d'évaluation, généralement composé d'un ensemble d'ontologies, d'un alignement de référence et potentiellement d'autres entrées variées (requêtes, instances, alignement partiel, etc.). L'alignement généré est évalué par un **outil d'évaluation** qui lui donne un score. Différentes dimensions d'évaluation peuvent être considérées :

Ressources Consommation de l'outil (mémoire, temps, CPU)

Entrées contrôlées Évaluation de l'outil en faisant varier ses paramètres d'entrée, comme dans les tâches GeoLink et Hydrography [7]

Sortie Évaluation de l'alignement lui-même, soit sur ses propriétés intrinsèques [15, 16], soit sur sa conformité à un alignement de référence.

Orienté tâche Évaluation de l'alignement sur son application à une tâche donnée [17, 18]

2. https://framagit.org/IRIT_UT2J/conference-dataset-population

3.1 Métriques d'évaluation

Les travaux ayant proposé des outils d'alignement complexe ont été évalués manuellement en termes de précision [5, 6, 19, 20], ou sur des jeux de données spécifiques (parfois manuellement créés) à leur outil pour calculer un rappel [6, 20].

Les métriques *accuracy* et *top−x accuracy* [21] ont été appliquées dans des évaluations où le nombre de correspondances recherchées est prédéfini, e.g., une seule correspondance est attendue pour chaque entité de l'ontologie cible. L'accuracy est le pourcentage de "questions" (ou sous-tâches) prédéfinies ayant une réponse correcte. Une "question" dans ce contexte pourrait être une entité de l'ontologie cible à aligner et les "réponses" les correspondances ayant cette entité comme membre cible. Certains outils génèrent des réponses différentes pour chaque question, e.g., une liste ordonnée de correspondances pour chaque entité cible. Dans ce cas, la top$-x$ accuracy est le pourcentage de questions pour lesquelles la réponse correcte se trouve dans les x premières réponses à la question.

L'approche [22], pour évaluer les correspondances complexes entre ontologies agronomiques, se fonde sur une comparaison manuelle de requêtes de référence et de requêtes automatiquement réécrites grâce à l'alignement.

3.2 Approches et jeux de données

La première tâche d'alignement complexe de l'OAEI [7] consiste en quatre jeux de données sur des domaines variés avec des stratégies d'évaluation variées :

Complex conference Un alignement consensuel contenant des correspondances (s :c) a été créé, fondé sur la méthodologie de la réécriture de requête [12]. Chaque correspondance est manuellement classifiée comme vrai positif ou faux positif par rapport à l'alignement de référence.

Hydrography and GeoLink Ce jeu porte sur des ontologies sur l'hydrographie et sur la GeoScience [23]. Les alignements sont évalués sur les tâches suivantes : i) trouver toutes les entités qui apparaissent dans une correspondance, ii) trouver la construction correcte reliant les entités et iii) trouver les correspondances complexes à partir de rien. En 2018, uniquement la première tâche a été implémentée [24]. En 2019, une métrique proche de la précision et rappel relaxés [25] a été appliquée aux tâches i et ii.

Taxon un ensemble de CQA sur des bases de connaissances agronomiques sont réécrites en utilisant l'alignement évalué. Chaque requête réécrite est classifiée manuellement comme sémantiquement équivalente ou non. Chaque correspondance est également manuellement classifiée comme vrai positif ou faux positif sans référence.

3.3 Évaluation orientée tâche

Certaines applications des alignements comme l'évolution d'ontologie ou la réponse aux questions (query answering) ont des contraintes différentes en termes de couverture et de temps d'exécution [11]. La tâche *OA4QA* [26] s'est spécialisée sur la réponse aux questions. Cette tâche a utilisé une version artificiellement peuplée du jeu de données *Conférence* et un ensemble de requêtes manuellement créées sur ces instances. Une requête exprimée avec l'ontologie *cmt* est exécutée sur l'ontologie fusionnée *cmt* $\cup$ *ekaw* $\cup$ *A*, où *A* est l'alignement entre *cmt* et *ekaw*. Les résultats de la requête étaient comparés à ceux sur l'ontologie *cmt* $\cup$ *ekaw* $\cup$ *ra1*, *ra1* étant l'alignement de référence.

[27] propose une évaluation "end-to-end" de requêtes réécrites à partir d'un alignement évalué. Les résultats des requêtes sont manuellement classifiés sur une échelle à 6 points par rapport à leur pertinence pour un utilisateur. Cette évaluation a été menée avec deux systèmes de réécriture de requêtes. Si un membre source e n'apparaît pas dans les correspondances, le système "vers le haut" cherchera une classe parente de e dans les membres sources et le système "vers le bas" cherchera une classe enfant de e.

La réécriture de requêtes est une des applications importantes pour l'alignement complexe. Évaluer de tels alignements sur cette tâche est par conséquent pertinent. La réécriture de requêtes fondée sur des alignements simples peut se contenter d'une approche naïve consistant à remplacer l'IRI d'un membre source par celle du membre cible dans la requête [28]. La tâche n'est pas aussi aisée pour les alignements complexes où la sémantique de l'alignement elle-même doit être prise en considération. [29] présente une approche de réécriture pour des requêtes SPARQL CONSTRUCT spécifiques mais la plupart des systèmes de réécriture de requêtes se fondent sur des correspondances simples ou (s :c) et échouent à gérer les correspondances (c :c) très expressives.

3.4 Positionnement

Pour l'évaluation des alignements complexes, des travaux se concentrent sur une évaluation manuelle en termes de précision [5, 19], calculer le rappel sur des patrons récurrents entre ontologies [6, 20], ou se fondent sur un échantillon de correspondances de référence [30]. Tandis que ces approches se concentrent sur la comparaison de correspondances, nous transférons le problème sur de la comparaison d'instances. Nous proposons une évaluation qui considère des requêtes comme référence et dont les métriques sont la couverture de requêtes (comme rappel) et la précision intrinsèque (comme la précision, mais sans alignement de référence). Par conséquent, notre approche nécessite des jeux de données peuplées de manière contrôlée.

Comme [26], la référence pour l'évaluation est un ensemble de requêtes et non un alignement. Proche de notre approche, [26] se fonde sur le jeu de données *Conférence* peuplé artificiellement. Toutefois, leurs requêtes sont exécutées sur une ontologie fusionnée et leurs alignements sont limités aux correspondances simples. Dans notre cas, les requêtes sont exécutées sur les différentes ontologies peuplées. Similairement à [27], les requêtes sont réécrites automatiquement avec l'alignement évalué. Toutefois, notre approche est entièrement automatique et ne se fonde pas sur une classifica-

tion manuelle des requêtes réécrites.
La Table 1 résume les approches d'évaluation d'alignements complexes proches de notre proposition (**CQA** benchmark, en gras dans la Table 1).

4 Évaluation automatique d'alignements complexes

La plupart des métriques d'évaluation pour les alignements simples ou complexes se fondent sur une comparaison syntaxique ou sémantique, mais très peu sur la comparaison au niveau des instances.
La comparaison *syntaxique* des alignements mesure l'effort à fournir pour transformer une correspondance évaluée en celle de référence. Toutefois, cela ne prend pas en compte le cas où une correspondance est sémantiquement équivalente à celle de référence mais utilise des constructeurs ou des niveaux de factorisation différents. Une comparaison syntaxique dépend également du langage et de la manière dont les correspondances sont exprimées. Par exemple, $\langle$ *o:Author* , $\exists o'$*:authorOf.*$\top$, $\equiv$ $\rangle$ est sémantiquement équivalent à $\langle$ *o:Author* , $\exists o'$*:writtenBy*$^-.\top$, $\equiv$ $\rangle$ mais ces correspondances utilisent des IRI et des constructeurs différents et sont donc syntaxiquement différentes. Un problème de factorisation reviendrait à comparer $\langle$ *o:paperWrittenBy , dom(*o'*:Paper)* $\sqcap$ o'*:writes*$^-$, $\equiv$ $\rangle$ et $\langle$ *o:paperWrittenBy , (*o'*:writes* $\sqcap$ *range(*o'*:Paper))*$^-$, $\equiv$ $\rangle$ qui sont deux correspondances équivalentes. Le constructeur *inverse* est factorisé dans la seconde correspondance. La comparaison syntaxique de requêtes fait face au même problème : des requêtes syntaxiquement différentes peuvent être sémantiquement équivalentes.
La comparaison *sémantique* compare le sens des formules. La précision et le rappel sémantiques comparent l'ensemble des axiomes inférés de la fusion des ontologies avec l'alignement évalué à la fusion des ontologies avec l'alignement de référence. Pour les alignements complexes, cela a l'avantage que tout élément traduisible en OWL puisse être évalué de cette manière. Toutefois, l'expressivité de l'alignement évalué fusionné avec les ontologies est limité à $\mathcal{SROIQ}$ (le fragment décidable de OWL [31]). Les correspondances avec des fonctions de transformation ne peuvent pas être comparées de cette manière non plus. La comparaison sémantique de requêtes proposée par [32] se fonde sur l'imbrication de requêtes, qui peut nécessiter des inférences. Cette comparaison ne peut également pas s'appliquer aux fonctions de transformation.
La comparaison *fondée sur les instances* (de correspondances ou de résultats de requêtes) est une alternative aux deux méthodes sus-mentionnées. Elle a l'inconvénient de nécessiter que les ontologies soient peuplées et de manière contrôlée.
Nous proposons deux métriques d'évaluation. La *couverture de CQA* (ou CQA Coverage) mesure à quel point un alignement permet de traduire un ensemble de CQA. La *précision intrinsèque* (ou intrinsic precision) compare les instances des membres des correspondances. La précision intrinsèque équilibre la couverture de CQA comme la précision équilibre le rappel en recherche d'information.

4.1 Déroulé d'évaluation

Figure 1 présente le déroulé d'évaluation adopté par notre approche. Les étapes du déroulé sont :

① **Sélection Ancre** Cette étape consiste à renvoyer une paire d'objets comparables $\langle x_i, x_{rj} \rangle$. x_i est un objet lié à A_{eval} and x_{rj} lié à $reference$. Dans le cas où la référence est une requête ou une paire de requêtes équivalentes, x_i peut être une requête dérivée de A_{eval} et x_{rj} une requête de référence.

② **Comparaison** L'étape de comparaison a pour but de renvoyer une relation $rel(x_i, x_{rj})$ pour chaque paire obtenue précédemment $\langle x_i, x_{rj} \rangle$. La relation peut être une équivalence (*i.e.*, $x_i \equiv x_{rj}$), une subsomption, une intersection, une disjonction, *etc.* Une valeur de similarité peut être associée à la relation. Dans notre approche, la comparaison est fondée sur les instances.

③ **Score** Cette étape associe un score à chaque relation trouvée précédemment $rel(x_i, x_{rj})$.

④ **Agrégation** Les scores sont localement et globalement agrégés pour obtenir le $score\ final$. Les agrégations peuvent être faites par : meilleur candidat, moyenne, moyenne pondérée, *etc.* L'agrégation locale agrège les scores pour un objet donné. Il peut y avoir différentes agrégations locales. Par exemple, il peut y avoir une agrégation par rapport à l'objet évalué et par rapport à l'objet de référence. L'agrégation globale se fait sur les scores localement agrégés pour obtenir le score final.

4.2 Couverture de CQA

La référence est un ensemble de CQA équivalentes sous la forme de requêtes SPARQL. Un alignement évalué A_{eval} sera utilisé pour réécrire chaque CQA source. La requête réécrite sera comparée avec la requête CQA cible. La comparaison des requêtes est fondée sur les instances et une valeur est associée à la relation entre les deux requêtes fondée sur l'intersection des instances renvoyées la requête évaluée et par la CQA cible. Différentes fonctions de scores sont appliquées en fonction de la relation entre requêtes. Une agrégation de meilleur candidat est faite localement sur les requêtes de référence. Une moyenne des scores localement agrégés est faite pour obtenir le score final.

4.2.1 Ancrer la CQA source

La référence dans cette évaluation est un ensemble de CQA sous la forme de requêtes SPARQL sur les ontologies cible et source. Chaque CQA source cqa_s a un équivalent cible cqa_t. Dans l'étape de sélection d'ancre, chaque cqa_s est réécrit en utilisant l'alignement évalué A_{eval}. L'étape de réécriture renvoie toutes les requêtes cibles possibles générées par le système de réécriture à partir de A_{eval} dans l'ensemble Q_t. Pour chaque requête q_t dans Q_t, une paire (q_t, cqa_t) est formée.
Deux systèmes de réécritures ont été considérés. Aucun de ces système ne considère la relation d'une correspondance

TABLE 1 – Comparaison des approches d'évaluation d'alignements complexes. Le *Type de corresp.* la forme la plus expressive que permet de gérer l'approche – (c :c) est plus complexe que (s :c), qui est plus complexe que (s :s).

Benchmark	*Type d'évaluation*	*Type de référence*	*Type de corresp.*
OA4QA [26]	Automatique (precision/recall)	Requête	(s :s)
Query rewrite [27]	Manuelle	Requête	(s :s)
Patterns evaluation [6]	Manuelle	Alignement	(s :c)
Patterns evaluation [20]	Manuelle	Alignement	(s :c)
Thieblin 2018 [12]	Manuelle	Alignement	(s :c)
GeoLink 2018 [23]	Automatique (precision/recall) /Manuelle	Alignement	(c :c)
Hydrography 2018 [23]	Automatique (precision/recall)/Manuelle	Alignement	(c :c)
GeoLink 2019	Automatique (relaxed precision/recall)	Alignement	(c :c)
Hydrography 2019	Automatique (relaxed precision/recall)	Alignement	(c :c)
Taxon [22]	Manuelle	Requête	(c :c)
CQA benchmark	**Automatique basée sur instances (couverture de CQA/ précision intrinsèque)**	**Requête**	**(c :c)**

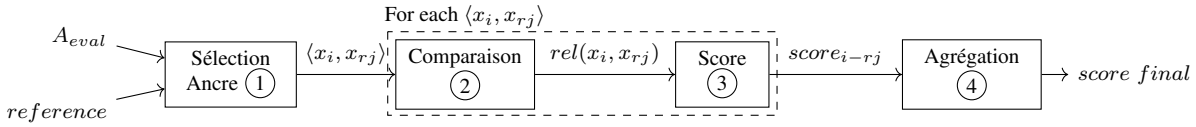

FIGURE 1 – Déroulé de l'évaluation de l'alignement A_{eval} avec une $reference$ générique.

ni la valeur de confiance lui étant attribuée. Le premier système [33] réécrit chaque triplet du patron de graphe de cqa_s en utilisant A_{eval}. Quand le prédicat ou l'objet du triplet apparaît comme membre source d'une correspondance dans A_{eval}, le membre cible de la correspondance est transformé en patron de graphe SPARQL et remplace le triplet dans la requête. Ce système ne gère que les correspondances (s :c). Si un triplet peut être réécrit avec plusieurs correspondances, toutes les combinaisons possibles sont ajoutées à Q_t. Par exemple,

```
Q1 = SELECT ?s WHERE{?s a
ekaw:Accepted_Paper.}
```

contient *ekaw:Accepted_Paper*, membre source de $c_1 = \langle$*ekaw:Accepted_Paper,* $\exists$*cmt:hasDecision.cmt:Acceptance,* $\equiv$*, 1*$\rangle$.

La requête réécrite en utilisant la correspondance c_1 :

```
Q2 = SELECT ?s WHERE{?s cmt:hasDecision
?o. ?o a cmt:Acceptance.}
```

Ce système de réécriture ne peut malheureusement par fonctionner dans l'autre sens : `Q2` ne pourra pas être réécrite avec c_1.

Le second système est fondé sur les instances et a été développé dans le contexte de cet article. Les instances I_s^{cqa} de cqa_s sont récupérées de l'ontologie source. Pour chaque correspondance c de A_{eval}, le membre source est transformé en requête pour récupérer les instances I_s de l'ontologie source. Si $I_s \equiv I_s^{cqa}$, alors le membre cible de c est transformé en requête et ajouté à Q_t.

Par exemple, `Q1` récupère un ensemble d'instances de papier accepté de l'ontologie *ekaw*. Cet ensemble d'instances est comparé aux membres sources de chaque correspondance de A_{eval}. Dans ce cas, *ekaw:Accepted_Paper* décrit les mêmes instances que le membre source de c_1. Son membre cible peut être traduit en requête SPARQL et devient `Q2`.

Ce système permet de réécrire des requêtes dans les deux sens, donc de gérer les correspondances (c :c). `Q2` pourra être réécrite avec l'inverse de c_1 (l'inverse d'une correspondance est son équivalent où le membre source devient le membre cible et vice-versa). Toutefois, ce système ne permet pas de combiner les correspondances entre elles.

Les deux systèmes de réécriture de requêtes sont utilisés.

4.2.2 Comparaison

Les instances I_t^{cqa} de cqa_t et I_t de q_t sont récupérées de l'ontologie cible. I_t et I_t^{cqa} sont comparés. Nous calculons deux scores par cette comparaison : précision de requête (QP) et de rappel de requête (QR). Ces valeurs sont associées à la relation $rel(q_t, cqa_t)$.

$$QP = \frac{|I_t \cap I_t^{cqa}|}{|I_t|} \qquad QR = \frac{|I_t \cap I_t^{cqa}|}{|I_t^{cqa}|}$$

$$rel(q_t, cqa_t) = \begin{cases} \equiv \text{ if } QR = 1 \text{ and } QP = 1 \\ \sqsubseteq \text{ if } QR \leq 1 \text{ and } QP = 1 \\ \sqsupseteq \text{ if } QR = 1 \text{ and } QP \leq 1 \\ overlap \text{ if } 0 < QR \leq 1 \text{ and} \\ 0 < QP \leq 1 \\ \bot \text{ if } QR = 0 \text{ and } QP = 0 \end{cases}$$

4.2.3 Score

Les valeurs de précision de requête et rappel de requête entre cqa_t et q_t sont transformées en un score de F-mesure de requête par une moyenne harmonique.

$$Fmesure = 2 \times \frac{QR \times QP}{QR + QP}$$

La F-mesure de requête équilibrant précision et rappel de requête a été choisie par rapport aux autres métriques (classique, relaxée).

4.2.4 Agrégation

Comme l'étape de réécriture renvoie toutes les requêtes possibles sans prendre en compte la relation de la correspondance, du bruit est introduit. De plus, la même requête peut être générée par les deux systèmes de réécriture. Par conséquent, pour chaque cqa_t, la requête q_t ayant le meilleur score est gardée. Cette agrégation par meilleur candidat permet de limiter l'impact du bruit introduit par l'étape de réécriture. Si cqa_s n'a pas pu être réécrit par l'alignement, son score est 0. L'agrégation globale est faite par moyenne sur les scores de chaque CQA.

4.3 Précision intrinsèque

La couverture de CQA agrège les résultats par CQA et non par requête réécrite à cause du bruit introduit par les systèmes de réécriture de requêtes. Cela induit qu'un alignement ayant toutes les correspondances possibles (correctes et erronées) entre une ontologie source et cible obtiendrait un bon score de couverture de CQA. Pour contrebalancer cela, nous proposons de mesurer la précision intrinsèque fondée sur les instances d'un alignement. Pour chaque correspondance c_i de A_{eval}, les instances I_s représentées par le membre source de la correspondance sont comparées aux instances I_t représentées par le membre cible. Chaque correspondance est classifiée comme *équivalence*, *sumbsomption*, *intersection*, *disjonction* en fonction de la relation entre I_s et I_t, ou *vide* if $I_s = I_t = \emptyset$. Une correspondance peut être *vide* si ses deux membres sont des entités insatisfiables ou des entités non peuplées.

Des scores de précision sont donnés pour chaque classe de correspondance : *équivalence* mesure le pourcentage de correspondances dont les membres sont exactement peuplés avec les mêmes instances, *subsomption* celles dont un membre subsume l'autre, *intersection* celles ayant une intersection non nulle et *non-disjointes* considère toutes les correspondances sauf les *disjonctions*.

5 Jeu de données fondé sur des CQA

5.1 Méthodologie de création du jeu de données

Le but est de créer un jeu de données sur lequel les outils d'alignement d'ontologies peuvent fonctionner et sur lequel l'évaluation décrite dans la section précédente peut être exécutée. Ce jeu de données doit donc contenir des ontologies peuplées ainsi qu'un ensemble de CQA exprimées sous forme de requêtes SPARQL sur ces ontologies.

Nous proposons la méthodologie suivante :

1. Créer un ensemble de CQA fondé sur un scénario applicatif. Seules les CQA unaires et binaires sont considérées dans ce travail.
2. Créer un format pivot (qui servira de pierre de rosette entre les différentes ontologies) couvrant toutes les CQA précédemment définies.
3. Pour chaque ontologie du jeu de données, créer des requêtes SPARQL INSERT correspondant au format pivot.
4. Instancier le format pivot avec des données réelles ou artificielles.
5. Peupler les ontologies avec le format pivot instancié en utilisant les requêtes SPARQL INSERT.
6. Faire tourner un raisonneur pour vérifier la consistance des ontologies peuplées. Si une erreur est levée, changer son interprétation de l'ontologie et itérer sur les point 3 à 5.
7. En se fondant sur les requêtes SPARQL INSERT, traduire les CQA couvertes par deux ontologies ou plus comme requêtes SPARQL SELECT.

Dans cette méthodologie, l'interprétation des ontologies est identique pour le peuplement et pour la création des CQA. La création des CQA peut être faite avec des utilisateurs et des experts du domaine, comme recommandé dans la méthodologie NeOn [34] pour les questions de compétence. Les CQA peuvent aussi dériver des questions de compétence utilisées pour concevoir les ontologies du jeu de données. Dans notre implémentation, un expert a créé les CQA, et elles ont été validées par un second expert qui a jugé qu'elles étaient assez exhaustives pour couvrir un scénario d'organisation de conférence.

5.2 Jeu de données conférence

Le jeu de données conférence [35] a été largement utilisé [9], en particulier lors des campagnes de l'OAEI. Il se compose de 16 ontologies sur le domaine de l'organisation de conférences et d'un alignement de référence simple entre 7 d'entre elles. Ces ontologies ont été développées individuellement. Des alignements complexes de référence ont été proposés pour 5 ontologies de ce jeu de données [12]. Dans la première tâche complexe de l'OAEI, une évaluation a été proposée sur un alignement complexe consensuel entre 3 ontologies (*cmt*, *conference*, *ekaw*) [7].

Nous avons peuplé les 5 ontologies des alignements de référence de [12] : *cmt*, *conference* (Sofsem), *confOf* (confTool), *edas* et *ekaw* (Table 2).

TABLE 2 – Number of entities by type of each ontology.

	cmt	conference	confOf	edas	ekaw
Classes	30	60	39	104	74
Obj. prop.	49	46	13	30	33
Data prop.	10	18	23	20	0

Bien que ce jeu de données ait été beaucoup utilisé, il a seulement été partiellement peuplé. Dans la tâche OA4QA, les classes sur lesquelles portaient les 18 requêtes de l'évaluation ont été peuplées mais la création de ces instances n'a pas été documentée.

5.3 Peuplement des ontologies de conférence

Un scénario d'organisation de conférence a été envisagé en examinant un cas réel : l'édition 2018 de la conférence ESWC. La liste de CQA en résultant a été étendue en explorant le champ couvert par les ontologies de conférence. Le format pivot a été instancié une première fois avec les données du site Web d'ESWC. L'étape du raisonneur a levé plusieurs erreurs comme la relation *cmt:hasAuthor* qui est fonctionnelle et ne représente donc que le premier auteur d'un article. Ayant identifié ces erreurs dans notre interprétation, nous avons repris la méthodologie étape par étape.

Deux erreurs qui n'ont pu être résolues par un changement d'interprétation ont nécessité une légère modification des ontologies.

— *cmt* : la relation *cmt:acceptPaper* entre un *Administrator* et un *Paper* était définie fonctionnelle et inverse fonctionnelle. Cela levait une erreur quand un administrateur acceptait plus d'un article. Elle a été modifiée pour n'être plus que inverse fonctionnelle.
— *conference* : *conference:Contribution_1st_author* était disjointe de *conference:Contribution_co-author*, ce qui levait une incohérence lorsqu'une personne était à la fois autrice d'un article et co-autrice d'un autre. La disjonction a été retirée.

Une ontologie n'est couverte que par CQA. Cela peut résulter en des populations hétérogènes pour certains concepts pourtant équivalents. Par exemple, *ekaw* et *cmt* ont toutes deux une classe *Document* mais la question *"Quels sont les documents ?"* n'est pas une CQA tandis que leurs sous-classes *Paper, Review, WebSite, Proceedings* sont le focus de CQA. Tandis que *ekaw* comporte les 4 sous-classes, *cmt* n'a que *Paper* et *Review*. La classe *cmt:Document* ne contiendra donc pas de site Web ou de proceedings tandis que *ekaw:Document* si.

Pour être proches de la réalité dans notre jeu de données, nous avons analysé les propriétés d'ISWC 2018 et ESWC 2017 à partir de Scholarly Data en complément d'ESWC 2018 et extrait des statistiques.

La première instanciation du format pivot à partir du site Web manquait d'informations pour représenter l'organisation d'une conférence. La liste des CQA a été étendue pour couvrir ce scénaria. En ont résulté une extension du format pivot et des requêtes SPARQL INSERT. Des données générées artificiellement à partir des statistiques des 3 conférences ont instancié le format pivot pour des conférences plus ou moins grandes (plus une conférence est grande, plus elle a de papiers soumis, de personnes, de membres du program committee, de workshop, etc.).

Pour permettre aux outils d'alignement fondés sur des statistiques d'être testés sur ce jeu de données, nous avons peuplé les mêmes ontologies avec des instances se recoupant plus ou moins. Chaque ontologie est peuplée avec plusieurs conférences, chaque conférence n'ayant aucune instance commune avec les autres. Cela permet de quantifier la partie d'instances communes à deux ontologies peuplées. 6 jeux de données en résultent, peuplés avec 25 conférences artificielles :

— 0 % : 5 conférences différentes par ontologie
— 20 % : 1 conférence commune pour toutes les ontologies et 4 différentes par ontologie
— ...
— 100 % : 5 conférences communes pour toutes les ontologies

Le pourcentage donné dans le nom des jeux de données est le pourcentage de conférences communes par ontologie. Comme la taille de chaque conférence est différente, le pourcentage des autres instances (articles, personnes, etc.) n'y sera pas identique.

5.4 CQA pour évaluation

Pour l'évaluation d'alignements, seules les CQA pouvant être couvertes par deux ontologies ou plus sont intéressantes. Nous avons élagué la liste des CQA pour enlever :

— les CQA couvertes par une seule ontologie
— les CQA se traduisant de la même manière pour toutes les ontologies (e.g., label d'une instance avec *rdfs :label*).

Table 3 présente le nombre de CQA initiales couvertes par chaque ontologie et ayant été gardées pour l'évaluation. 278 requêtes SPARQL SELECT en résultent.

TABLE 3 – Nombre de CQA initiales et d'évaluation couvertes par chaque ontologie

	cmt	conference	confOf	edas	ekaw	total
init.	46	90	67	60	84	152
eval.	34	73	54	52	65	100

6 Évaluation

Nous avons évalué 5 alignements avec notre approche. Nous indiquons le nombre de paires d'ontologies sur les 10 de notre jeu de données qui sont couvertes par ces alignements. Ces alignements ne contiennent pas de correspondances (c :c).

Query_rewriting [12] créé manuellement, contient 431 correspondances dont 191 complexes sur les 10 paires.

Ontology_merging [12] créé manuellement, contient 313 correspondances dont 54 complexes sur les 10 paires.

ra1 l'alignement simple de référence du jeu de données *conférence* [9] sur les 10 paires.

Ritze_2010 généré automatiquement [19], sur 4 paires. Il ne contient qu'une seule correspondance par paire.

Faria_2018 généré automatiquement [36] entre 3 paires.

L'alignement ra1 a été utilisé comme entrée pour Ritze_2010 et Faria_2018. Ra1 donc été ajouté à ces alignements pour le calcul de couverture de CQA. La couverture de CQA a été calculée sur tous les jeux de données pour calculer la déviation standard des scores de précision

de requête, rappel de requête et F-mesure de requête. La déviation standard est inférieure à 2×10^{-3} sur les 6 jeux de données et les 5 alignements.

Table 4 présente les résultats de l'évaluation sur le jeu de données 100%. Ritze_2010 et Faria_2018 ont une meilleure couverture que ra1 qu'ils incluent. Les correspondances qu'ils contiennent sont un complément aux correspondances simples pour couvrir les CQA. ra1 a une meilleure précision d'équivalence (0.56) que les autres alignements créés manuellement car il contient uniquement des correspondances avec des équivalences. Ce score est bas pour un alignement de référence pour la raison évoquée §5.3. Au vu des problèmes de peuplement, les scores d'intersection et non-disjointes donnent un bon aperçu de la précision d'un alignement. Nous avons choisi d'agréger la couverture de CQA et les scores de précisions d'intersection et non-disjointes respectivement en une moyenne harmonique (MH).

TABLE 4 – Couverture de CQA(C), précision intrinsèque(P), moyenne harmonique (MH)

	Query rew.	Onto. merg.	ra1	Faria 2018	Ritze 2010
C. CQA	**0.69**	0.63	0.42	0.41	0.48
P. équi.	0.42	0.43	0.56	0.65	0.75
P. subs.	0.80	0.80	0.83	0.71	0.75
P. inter.	0.90	0.86	**0.92**	0.71	0.75
P. non-dis.	0.94	0.91	**0.96**	0.71	0.75
MH inter.	**0.78**	0.73	0.58	0.52	0.59
MH non-dis.	**0.80**	0.74	0.58	0.52	0.59

Les alignements Query_rewriting et Ontology_merging obtiennent les meilleurs résultats, ce qui conforte leur rôle d'alignement complexe de référence sur ce jeu de données. Bien que ra1 obtienne la meilleure précision, sa couverture de CQA faible (0.42) montre que nombre de CQA nécessitent des alignements pour être traduites dans ce jeu de données. Faria_2018 et Ritze_2010 ne s'appliquent pas au même nombre de paires d'ontologies et ne peuvent donc pas être comparés aux autres.

Dans les résultats de l'OAEI 2018 [24], la précision mesurée pour Faria_2018 était de 0.54 (cf. Table 5). La précision intrinsèque donne les mêmes résultats pour la paire *cmt-ekaw*. Pour les autres paires, la différence est significative.

Pour la paire *conference-ekaw*, $\langle \exists$*conference:was_a_track-workshop_chair_of.* *conference:Tutorial, ekaw :Tutorial_Chair,* $\equiv$*, 0.369*$\rangle$ était considéré correct dans l'évaluation OAEI 2018. Pourtant, un axiome de l'ontologie *conference* restreint le domaine de *conference:was_a_track-workshop_chair_of* à *conference:Track* $\sqcup$ *conference:Workshop*. Cela a été pris en compte dans le peuplement de l'ontologie et la correspondance a été évaluée comme étant disjointe par le système d'évaluation.

TABLE 5 – Comparaison des scores de précisoin de Faria_2018 dans l'évaluation OAEI 2018 et basée sur nos métriques.

pair	OAEI 2018	P. equiv.	P. non-disj.
cmt-conference	0.4	1.00	1.00
cmt-ekaw	0.86	0.86	0.86
conference-ekaw	0.36	0.09	0.27
Average	0.54	0.65	0.71

7 Conclusions et perspectives

Nous avons présenté une approche d'évaluation automatique d'alignements complexes et son jeu de données associé. Par rapport à l'évaluation d'alignements simple, il est difficile de comparer les membres d'une correspondance évaluée à celle d'un alignement de référence. Tandis que les métriques d'évaluation fondées sur de la comparaison syntaxique échoueraient à couvrir l'ensemble des combinaisons possibles de constructeurs et IRI, les approches sémantiques restreignent l'expressivité des correspondances et des alignements à celle supportée par les raisonneurs, délaissant notamment les fonctions de transformation. La comparaison fondée sur les instances permet un compromis. Notre proposition déplace le problème sur la comparaison d'instances dans une tâche de réécriture de requête ciblant des besoins utilisateur. Nous avons proposé deux métriques d'évaluation. La couverture de CQA mesure à quel point un alignement permet de traduire un ensemble de requêtes SPARQL et la précision intrinsèque compare les instances des membres d'une correspondance. La précision intrinsèque équilibre la couverture de CQA de la manière dont la précision équilibre le rappel en recherche d'information.

Un jeu de données artificiel sur les ontologies de conférence a été proposé. Son peuplement a été contrôlé et guidé par des CQA.

Les systèmes de réécriture de requêtes sont limités aux correspondances (s :c), gérer les correspondances (c :c) reste une question ouverte. Nous avons proposé un système permettant de les prendre en compte seulement individuellement et non de les combiner. La réécriture de requêtes fondée sur les instances pourrait toutefois être une nouvelle piste à explorer dans cette voie. Nous n'avons pas évalué l'impact des systèmes de réécriture de requêtes sur l'évaluation, mais plutôt tenté de limiter leur impact avec une sélection du meilleur candidat.

L'approche proposée a été appliquée à des alignements existants et dans la campagne OAEI 2019. Notre approche d'évaluation attribue une bonne précision aux alignements de référence et une meilleure couverture de CQA pour les alignements complexes que simples. Toutefois, l'évaluation manuelle dans l'OAEI et celle proposée ici présentent des différences significatives.

L'évaluation d'alignements complexes et un défi trop large pour être résolu avec une seule approche et d'autres pistes, plus sémantiques, sont notamment à explorer.

Références

[1] É. Thiéblin, O. Haemmerlé, and C. Trojahn, "Automatic evaluation of complex alignments : An instance-based approach," *Semantic Web*, vol. 12, no. 5, pp. 767–787, 2021.

[2] P. R. S. Visser, D. M. Jones, B. T. J. M. Capon, and M. J. R. Shave, "An analysis of ontological mismatches : Heterogeneity versus interoperability," in *AAAI 1997 Spring Symposium on Ontological Engineering*, Stanford, USA, 1997, pp. 164–72.

[3] A. Maedche, B. Motik, N. Silva, and R. Volz, "Mafra — a mapping framework for distributed ontologies," in *Knowledge Engineering and Knowledge Management : Ontologies and the Semantic Web*, A. Gómez-Pérez and V. R. Benjamins, Eds. Berlin, Heidelberg : Springer Berlin Heidelberg, 2002, pp. 235–250.

[4] É. Thiéblin, O. Haemmerlé, N. Hernandez, and C. Trojahn, "Survey on complex ontology matching," *Semantic Web*, vol. 11, no. 4, pp. 689–727, 2020.

[5] D. Ritze, C. Meilicke, O. Sváb-Zamazal, and H. Stuckenschmidt, "A pattern-based ontology matching approach for detecting complex correspondences," in *4th International Workshop on Ontology Matching (OM-2009) collocated with the 8th International Semantic Web Conference (ISWC-2009)*, ser. CEUR Workshop, P. Shvaiko, J. Euzenat, F. Giunchiglia, H. Stuckenschmidt, N. F. Noy, and A. Rosenthal, Eds., vol. 551, 2009.

[6] B. Walshe, R. Brennan, and D. O'Sullivan, "Bayesrecce : A bayesian model for detecting restriction class correspondences in linked open data knowledge bases," vol. 12, no. 2, p. 25–52, Apr. 2016.

[7] É. Thiéblin, M. Cheatham, C. T. dos Santos, O. Zamazal, and L. Zhou, "The first version of the OAEI complex alignment benchmark," in *ISWC 2018 Posters & Demonstrations, Industry and Blue Sky Ideas Tracks co-located with 17th International Semantic Web Conference ISWC 2018*, ser. CEUR Workshop, M. van Erp, M. Atre, V. López, K. Srinivas, and C. Fortuna, Eds., vol. 2180, 2018.

[8] J. Euzenat, M. Rosoiu, and C. Trojahn, "Ontology matching benchmarks : Generation, stability, and discriminability," *Journal of Web Semantics*, vol. 21, pp. 30–48, 2013.

[9] O. Zamazal and V. Svátek, "The Ten-Year OntoFarm and its Fertilization within the Onto-Sphere," *Journal of Web Semantics*, vol. 43, pp. 46–53, 2017.

[10] É. Thiéblin, "Do competency questions for alignment help fostering complex correspondences ?" in *EKAW Doctoral Consortium 2018*, ser. CEUR Workshop, L. Hollink and F. Osborne, Eds., vol. 2306, 2018.

[11] J. Euzenat and P. Shvaiko, *Ontology Matching*, 2nd ed. Springer Berlin Heidelberg, 2013.

[12] É. Thiéblin, O. Haemmerlé, N. Hernandez, and C. Trojahn, "Task-oriented complex ontology alignment : Two alignment evaluation sets," in *The Semantic Web - 15th International Conference, ESWC 2018*, ser. Lecture Notes in Computer Science, A. Gangemi, R. Navigli, M. Vidal, P. Hitzler, R. Troncy, L. Hollink, A. Tordai, and M. Alam, Eds., vol. 10843. Springer, 2018, pp. 655–670.

[13] M. Grüninger and M. S. Fox, "Methodology for the design and evaluation of ontologies," in *Workshop on Basic Ontological Issues in Knowledge Sharing*, vol. 15, 1995.

[14] Y. Ren, A. Parvizi, C. Mellish, J. Z. Pan, K. van Deemter, and R. Stevens, "Towards competency question-driven ontology authoring," in *The Semantic Web : Trends and Challenges - 11th International Conference, ESWC 2014*, ser. Lecture Notes in Computer Science, V. Presutti, C. d'Amato, F. Gandon, M. d'Aquin, S. Staab, and A. Tordai, Eds., vol. 8465. Springer, 2014, pp. 752–767.

[15] C. Meilicke and H. Stuckenschmidt, "Incoherence as a basis for measuring the quality of ontology mappings," in *3rd International Workshop on Ontology Matching*, ser. CEUR Workshop, P. Shvaiko, J. Euzenat, F. Giunchiglia, and H. Stuckenschmidt, Eds., vol. 431, 2008.

[16] A. Solimando, E. Jiménez-Ruiz, and G. Guerrini, "Detecting and correcting conservativity principle violations in ontology-to-ontology mappings," in *The Semantic Web - ISWC 2014 - 13th International Semantic Web Conference*, ser. Lecture Notes in Computer Science, P. Mika, T. Tudorache, A. Bernstein, C. Welty, C. A. Knoblock, D. Vrandecic, P. Groth, N. F. Noy, K. Janowicz, and C. A. Goble, Eds., vol. 8797. Springer, 2014, pp. 1–16.

[17] A. Isaac, H. Matthezing, L. van der Meij, S. Schlobach, S. Wang, and C. Zinn, "Putting ontology alignment in context : Usage scenarios, deployment and evaluation in a library case," in *The Semantic Web : Research and Applications, 5th European Semantic Web Conference, ESWC 2008*, ser. Lecture Notes in Computer Science, S. Bechhofer, M. Hauswirth, J. Hoffmann, and M. Koubarakis, Eds., vol. 5021. Springer, 2008, pp. 402–417.

[18] A. Isaac, D. Kramer, L. van der Meij, S. Wang, S. Schlobach, and J. Stapel, "Vocabulary matching for book indexing suggestion in linked libraries - A prototype implementation and evaluation," in *The Semantic Web - ISWC 2009, 8th International Semantic Web Conference, ISWC 2009*, ser. Lecture Notes in Computer Science, vol. 5823. Springer, 2009, pp. 843–859.

[19] D. Ritze, J. Völker, C. Meilicke, and O. Sváb-Zamazal, "Linguistic analysis for complex ontology matching," in *5th International Workshop on Ontology Matching (OM-2010)*, ser. CEUR Workshop,

P. Shvaiko, J. Euzenat, F. Giunchiglia, H. Stuckenschmidt, M. Mao, and I. F. Cruz, Eds., vol. 689, 2010.

[20] R. Parundekar, C. A. Knoblock, and J. L. Ambite, "Discovering concept coverings in ontologies of linked data sources," in *The Semantic Web - ISWC 2012 - 11th International Semantic Web Conference*, ser. Lecture Notes in Computer Science, vol. 7649. Springer, 2012, pp. 427–443.

[21] Y. An, X. Hu, and I. Song, "Learning to discover complex mappings from web forms to ontologies," in *21st ACM International Conference on Information and Knowledge Management, CIKM'12*, X. Chen, G. Lebanon, H. Wang, and M. J. Zaki, Eds. ACM, 2012, pp. 1253–1262.

[22] É. Thiéblin, F. Amarger, N. Hernandez, C. Roussey, and C. T. dos Santos, "Cross-querying LOD datasets using complex alignments : An application to agronomic taxa," in *Metadata and Semantic Research - 11th International Conference, MTSR*, ser. Communications in Computer and Information Science, E. Garoufallou, S. Virkus, R. Siatri, and D. Koutsomiha, Eds., vol. 755. Springer, 2017, pp. 25–37.

[23] L. Zhou, M. Cheatham, A. Krisnadhi, and P. Hitzler, "A complex alignment benchmark : Geolink dataset," in *The Semantic Web - ISWC 2018 - 17th International Semantic Web Conference,*, ser. Lecture Notes in Computer Science, D. Vrandecic, K. Bontcheva, M. C. Suárez-Figueroa, V. Presutti, I. Celino, M. Sabou, L. Kaffee, and E. Simperl, Eds., vol. 11137. Springer, 2018, pp. 273–288.

[24] A. Algergawy, M. Cheatham, D. Faria, A. Ferrara, I. Fundulaki, I. Harrow, S. Hertling, E. Jiménez-Ruiz, N. Karam, A. Khiat, P. Lambrix, H. Li, S. Montanelli, H. Paulheim, C. Pesquita, T. Saveta, D. Schmidt, P. Shvaiko, A. Splendiani, É. Thiéblin, C. Trojahn, J. Vatascinová, O. Zamazal, and L. Zhou, "Results of the ontology alignment evaluation initiative 2018," in *13th International Workshop on Ontology Matching*, ser. CEUR Workshop, P. Shvaiko, J. Euzenat, E. Jiménez-Ruiz, M. Cheatham, and O. Hassanzadeh, Eds., vol. 2288, 2018, pp. 76–116.

[25] M. Ehrig and J. Euzenat, "Relaxed precision and recall for ontology matching," in *Integrating Ontologies '05, K-CAP 2005 Workshop on Integrating Ontologies*, ser. CEUR Workshop, B. Ashpole, M. Ehrig, J. Euzenat, and H. Stuckenschmidt, Eds., vol. 156, 2005.

[26] A. Solimando, E. Jiménez-Ruiz, and C. Pinkel, "Evaluating ontology alignment systems in query answering tasks," in *ISWC 2014 Posters & Demonstrations Track a track within the 13th International Semantic Web Conference, ISWC*, ser. CEUR Workshop, M. Horridge, M. Rospocher, and J. van Ossenbruggen, Eds., vol. 1272, 2014, pp. 301–304.

[27] L. Hollink, M. van Assem, S. Wang, A. Isaac, and G. Schreiber, "Two variations on ontology alignment evaluation : Methodological issues," in *The Semantic Web : Research and Applications, 5th European Semantic Web Conference, ESWC 2008*, ser. Lecture Notes in Computer Science, S. Bechhofer, M. Hauswirth, J. Hoffmann, and M. Koubarakis, Eds., vol. 5021. Springer, 2008, pp. 388–401.

[28] J. David, J. Euzenat, F. Scharffe, and C. Trojahn, "The alignment API 4.0," *Semantic Web*, vol. 2, no. 1, pp. 3–10, 2011.

[29] J. Euzenat, A. Polleres, and F. Scharffe, "Processing ontology alignments with SPARQL," in *Second International Conference on Complex, Intelligent and Software Intensive Systems (CISIS-2008)*, F. Xhafa and L. Barolli, Eds. IEEE Computer Society, 2008, pp. 913–917.

[30] H. Qin, D. Dou, and P. LePendu, "Discovering executable semantic mappings between ontologies," in *On the Move to Meaningful Internet Systems 2007 : CoopIS, DOA, ODBASE, GADA, and IS, OTM*, ser. Lecture Notes in Computer Science, vol. 4803. Springer, 2007, pp. 832–849.

[31] I. Horrocks, O. Kutz, and U. Sattler, "The even more irresistible SROIQ," in *Tenth International Conference on Principles of Knowledge Representation and Reasoning*, P. Doherty, J. Mylopoulos, and C. A. Welty, Eds. AAAI Press, 2006, pp. 57–67.

[32] J. David, J. Euzenat, P. Genevès, and N. Layaïda, "Evaluation of query transformations without data : Short paper," in *Companion of The Web Conference 2018 WWW*, P. Champin, F. Gandon, M. Lalmas, and P. G. Ipeirotis, Eds. ACM, 2018, pp. 1599–1602.

[33] É. Thiéblin, F. Amarger, O. Haemmerlé, N. Hernandez, and C. T. dos Santos, "Rewriting SELECT SPARQL queries from 1 :n complex correspondences," in *11th International Workshop on Ontology Matching*, ser. CEUR Workshop, vol. 1766, 2016, pp. 49–60.

[34] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The neon methodology for ontology engineering," in *Ontology Engineering in a Networked World*, M. C. Suárez-Figueroa, A. Gómez-Pérez, E. Motta, and A. Gangemi, Eds. Springer, 2012, pp. 9–34.

[35] O. Šváb, V. Svátek, P. Berka, D. Rak, and P. Tomášek, "Ontofarm : Towards an experimental collection of parallel ontologies," in *4th International Semantic Web Conference (ISWC). Poster*, 2005.

[36] D. Faria, C. Pesquita, B. S. Balasubramani, T. Tervo, D. Carriço, R. Garrilha, F. M. Couto, and I. F. Cruz, "Results of AML participation in OAEI 2018," in *13th International Workshop on Ontology Matching*, ser. CEUR Workshop, P. Shvaiko, J. Euzenat, E. Jiménez-Ruiz, M. Cheatham, and O. Hassanzadeh, Eds., vol. 2288, 2018, pp. 125–131.