

# État des lieux de l'utilisation de OWL2 : Analyse et proposition pour capturer les utilisations de la sémantique OWL2 dans les graphes de connaissances RDF

Pierre-Henri Paris\*, Fayçal Hamdi\*  
Samira Si-said Cherfi\*

\*Conservatoire National des Arts et Métiers  
Laboratoire CEDRIC, équipe ISID  
Paris, France

<https://cedric.cnam.fr/index.php/labo/Isid>  
pierre-henri.paris@upmc.fr, faycal.hamdi@cnam.fr, samira.cherfi@cnam.fr

**Résumé.** Le nombre et la taille des graphes de connaissances RDF sont en constante augmentation. Par conséquent, le traitement des données pour des agents (automatisés ou humains) devient de plus en plus difficile. Si plusieurs outils peuvent être utilisés pour une tâche donnée, mais qu'ils dépendent chacun à des degrés divers de la sémantique disponible dans le graphe de connaissances, alors il est important d'avoir un aperçu en amont du graphe pour sélectionner le meilleur outil pour cette tâche. Nous avons mené, à grande échelle, une étude approfondie pour vérifier la présence de sémantique dans les graphes de connaissances publiés actuellement dans le Web de données (*Linked Data*). Bien que certains graphes de connaissances utilisent la sémantique OWL 2, beaucoup ne le font pas ou partiellement. Nous proposons donc une approche qui, en se basant sur des statistiques, instancie une ontologie facilitant la sélection de l'outil le mieux adapté à une tâche donnée en fonction de l'utilisation de la sémantique OWL 2.

## 1 Introduction

Avec l'augmentation du nombre et de la taille des graphes de connaissances RDF, la difficulté pour requêter ou utiliser ces données s'accroît. Pour une tâche donnée, plusieurs types d'approches peuvent être envisagés. Certaines approches reposant principalement sur la sémantique disponible dans les graphes, d'autres au contraire ne l'utilisant que peu ou pas du tout. Bien entendu, entre ces deux extrêmes, des approches peuvent tirer parti de la sémantique, sans toutefois se reposer totalement sur cette dernière. Par exemple, si la tâche est d'interconnecter plusieurs graphes de connaissances entre eux, les approches peuvent utiliser une combinaison de techniques telles que les statistiques, d'autres graphes de connaissances externes, la sémantique ou encore des algorithmes de partitionnement de données. Par ailleurs, des approches se reposant principalement sur la sémantique peuvent surpasser d'autres types d'approches si la sémantique est très présente dans le graphe de connaissances. Mais si la