

THÈSE DE DOCTORAT DE

NANTES UNIVERSITÉ

ÉCOLE DOCTORALE N° 641

*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*

Spécialité : *Informatique*

Par

Ginwa FAKIH

Relaxation of SPARQL queries in the presence of RDF reification

Thèse présentée et soutenue à Nantes , le 12 September 2025

Unité de recherche : Laboratoire des Sciences du Numérique de Nantes

Rapporteurs avant soutenance :

Sébastien FERRÉ Professeur, Université de Rennes
Khalid BELHAJJAME Professeur, CNAM

Composition du Jury :

Président :	Colin DE LA HIGUERA	Professeur, Nantes Université
Examineurs :	Khalid BELHAJJAME	Professeur, CNAM
	Zoltan MIKLOS	Professeur, Université de Rennes 1
	Mounira HARZALLAH	Maître de conférences HDR, Nantes Université
Dir. de thèse :	Patricia SERRANO ALVARADO	Maître de conférences HDR, Nantes Université

Invité(s) :

Benjamin MOREAU Docteur, Opendatasoft

Titre : Relâchement de requêtes SPARQL en présence de réification RDF

Mot clés : Web sémantique, données liées, RDF, SPARQL, relâchement des requêtes, ontologie, réification, métadonnées

Résumé : La réponse aux requêtes sur des graphes de connaissances RDF retourne souvent peu ou pas de résultats. Le relâchement de requête permet d'atténuer les contraintes pour obtenir des résultats similaires. Deux méthodes principales existent : utiliser la hiérarchie des classes et propriétés, ou les similarités entre entités et littéraux calculées via des fonctions de mappage. Ces approches ciblent respectivement des requêtes spécifiques au domaine et centrées sur les entités, mais les requêtes réelles combinent souvent ces éléments. De plus, les ensembles de données RDF peuvent contenir des métadonnées représentées à l'aide de la réification, ce qui complexifie davantage la relaxation de re-

quêtes. Cela motive le besoin d'un modèle hybride intégrant les deux approches de relaxation tout en prenant en compte les métadonnées. Les contributions de cette thèse sont les suivantes : (1) une analyse de l'état de l'art sous l'angle de la réification RDF ; (2) un modèle hybride combinant les relaxations basées sur l'ontologie et sur les entités à travers une mesure de similarité unifiée permettant un classement équitable des résultats issus de relaxations hétérogènes ; et (3) un benchmark d'évaluation prolongeant LUBM4OBDA en introduisant sept nouvelles requêtes, accompagnées de leurs versions relâchées, classées selon des jugements humains pour une évaluation équitable.

Title: Relaxation of SPARQL queries in the presence of RDF reification

Keywords: Semantic Web, Linked Data, RDF, SPARQL, query relaxation, ontology, reification, metadata

Abstract: Query answering over RDF knowledge graphs can often return too few or even no results. Query relaxation addresses this issue by modifying the constraints of a query to retrieve k-similar results. Two main approaches exist: using the hierarchy of classes and properties, or using the similarities between entities and literals, computed through pre-established mapping functions. These approaches typically target different query types, domain-specific queries versus entity-centric ones, but real-world queries often combine these different terms. Moreover, RDF datasets may contain metadata represented using reification, which further complicates query re-

laxation. This motivates the need for a hybrid model that integrates both relaxation approaches while taking metadata into account. The contributions of this thesis are: (1) a comprehensive analysis of the state of the art in query relaxation through the lens of RDF reification; (2) a hybrid model combining ontology-based and entity-based relaxations through a unified similarity measure that fairly ranks results obtained from heterogeneous relaxations; and (3) an evaluation benchmark that extends LUBM4OBDA by introducing seven new queries, along with their corresponding relaxed versions, ranked based on human judgment for fair evaluation.

