



**HAL**  
open science

## Interfacer un framework d'interrogation du LOD avec un framework de validation des résultats

Béatrice Markhoff, Jacques Chabin, Vivien Leonard

### ► To cite this version:

Béatrice Markhoff, Jacques Chabin, Vivien Leonard. Interfacer un framework d'interrogation du LOD avec un framework de validation des résultats. JIRC 2020 (Journées Informatique en Région Centre-Val de Loire), Feb 2020, Bourges, France, Feb 2020, Bourges, France. hal-03553186

**HAL Id: hal-03553186**

**<https://hal.science/hal-03553186v1>**

Submitted on 2 Feb 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Interfacer un framework d'interrogation du LOD avec un framework de validation des résultats

Béatrice MARKHOFF<sup>†</sup>, Vivien LEONARD, Jacques CHABIN<sup>‡</sup>

beatrice.markhoff@univ-tours.fr, vivien.leonard@etu.univ-Tours.fr, jacques.chabin@univ-orleans.fr

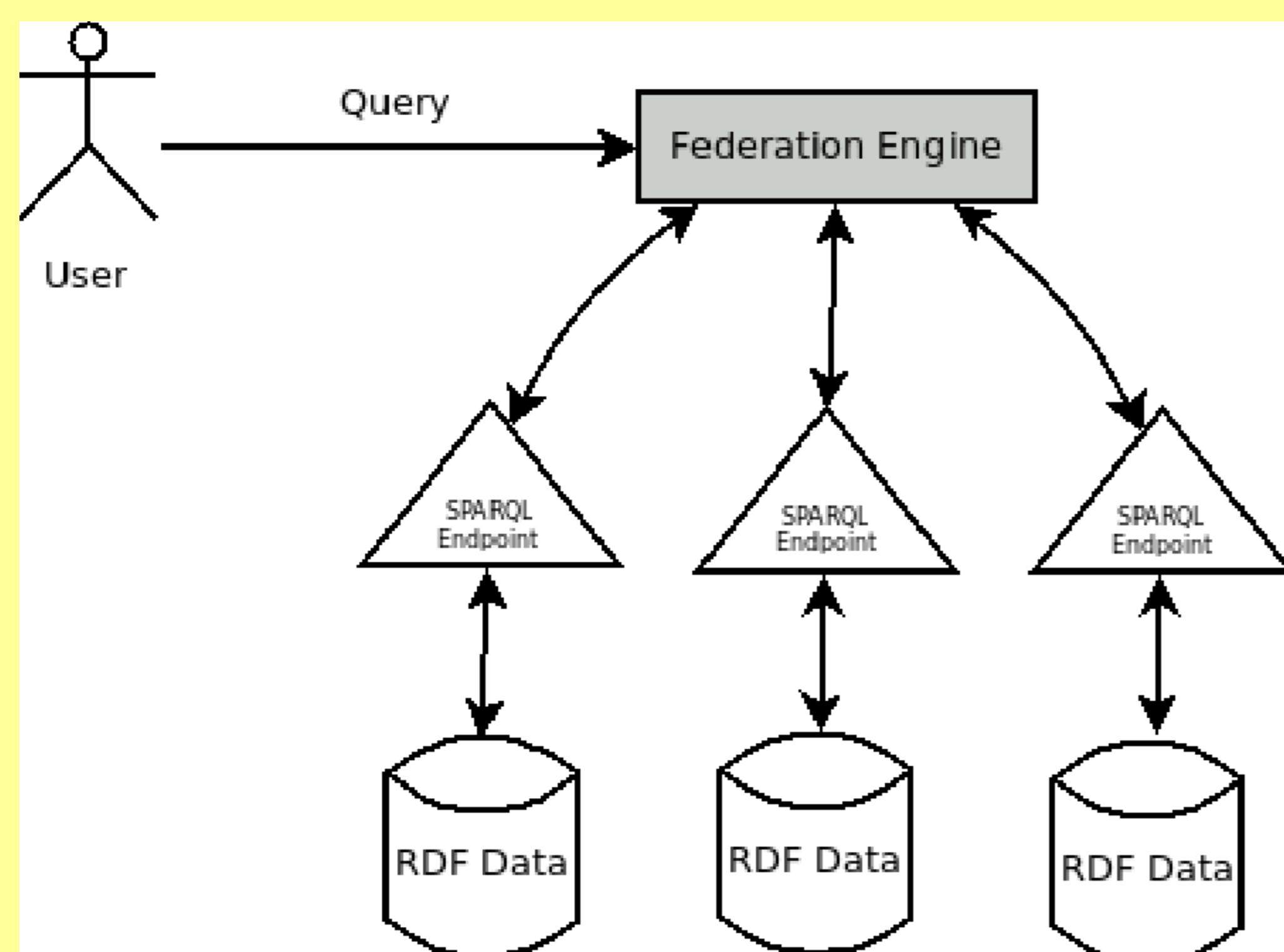
<sup>†</sup> Laboratoire d'Informatique Fondamentale et Appliquée de Tours. <sup>‡</sup> Laboratoire d'Informatique Fondamentale d'Orléans.

## Context scientifique

Le Linked Open Data (LOD), encore appelé web de données, est en croissance constante depuis une dizaine d'années et représente une mine d'informations exploitables par des logiciels. Les données y sont structurées en RDF et beaucoup de fournisseurs de données offrent des points d'accès SPARQL pour interroger leurs jeux de données. La qualité des données fournies dans le LOD est souvent discutable, variable, et surtout le consommateur de ces données ne peut pas agir directement sur les sources des données pour en augmenter la qualité.

### FedX

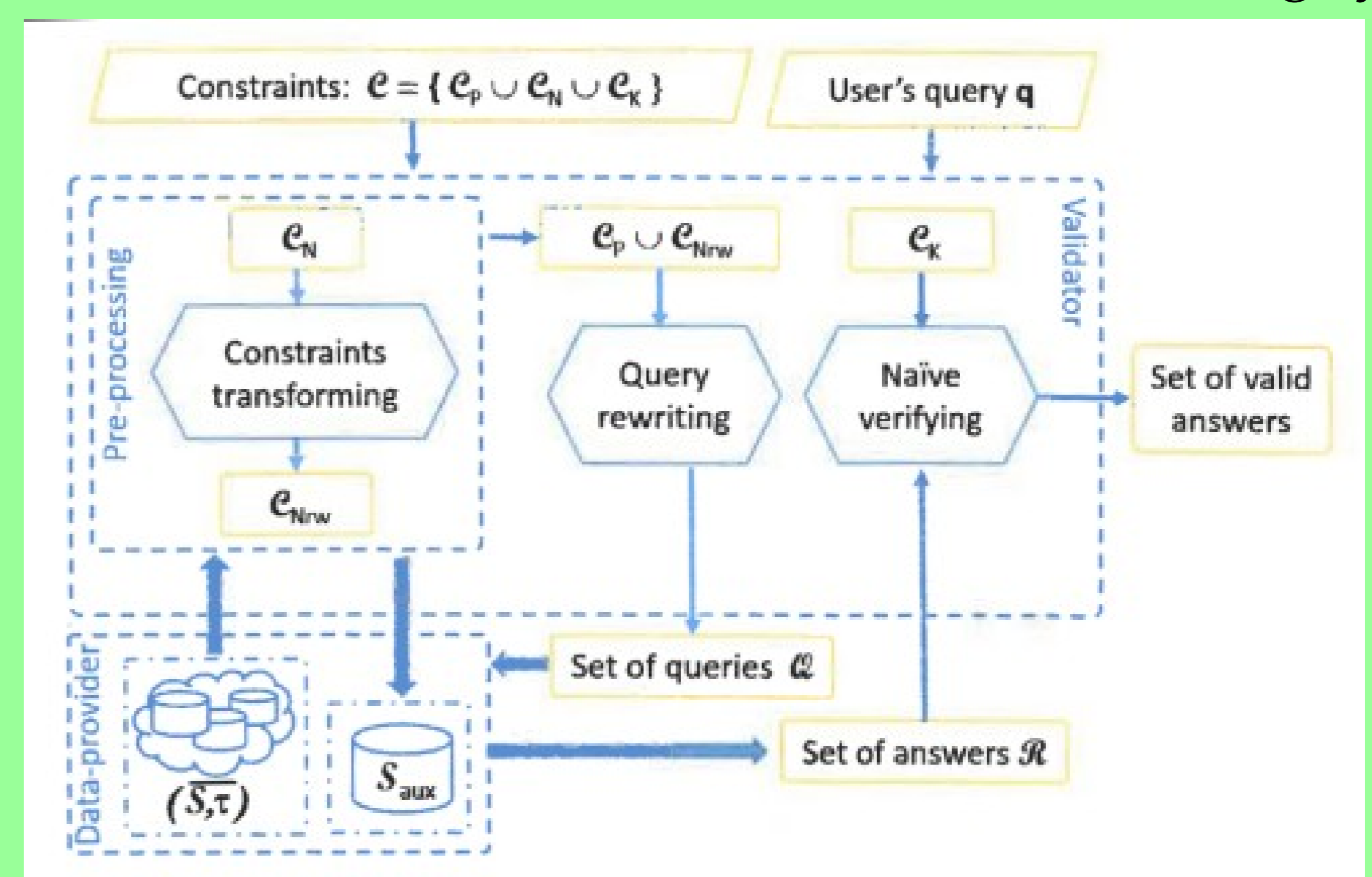
Pour prendre en charge la nature distribuée d'un système rassemblant plusieurs points d'accès SPARQL, nous utilisons un système de requête fédéré.



[Fedx] : Andreas Schwarte, Peter Haase, Katja Hose, Ralf Schenkel, and Michael Schmidt. Fedx: Optimization techniques for federated query processing on linked data. In *International Semantic Web Conference*, 2011.

### QCR

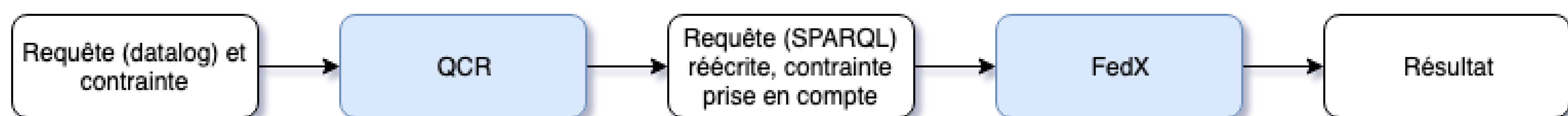
Un système permettant de spécifier des contraintes de qualité côté consommateur de données et de valider les résultats des requêtes en fonction de ces contraintes a été conçu dans le cadre de la thèse de Thanh Binh Nguyen.



[Sofsem] : Jacques Chabin, Mirian Halfeld Ferrari, Béatrice Markhoff and Thanh Binh Nguyen. Validating data from semantic web providers. In *SOFSEM 2018: Theory and Practice of Computer Science - 44th International Conference on Current Trends in Theory and Practice of Computer Science, Krems, Austria, January 29 - February 2, 2018, Proceedings*, pages 682–695, 2018.

[Gaal] : Jean-François Baget, Michel Leclère, Marie-Laure Mugnier, Swan Rocher, and Clément Sipietier. Gaal: A toolkit for query answering with existential rules. In *Rule Technologies: Foundations, Tools, and Applications*, pages 328–344, 2015.

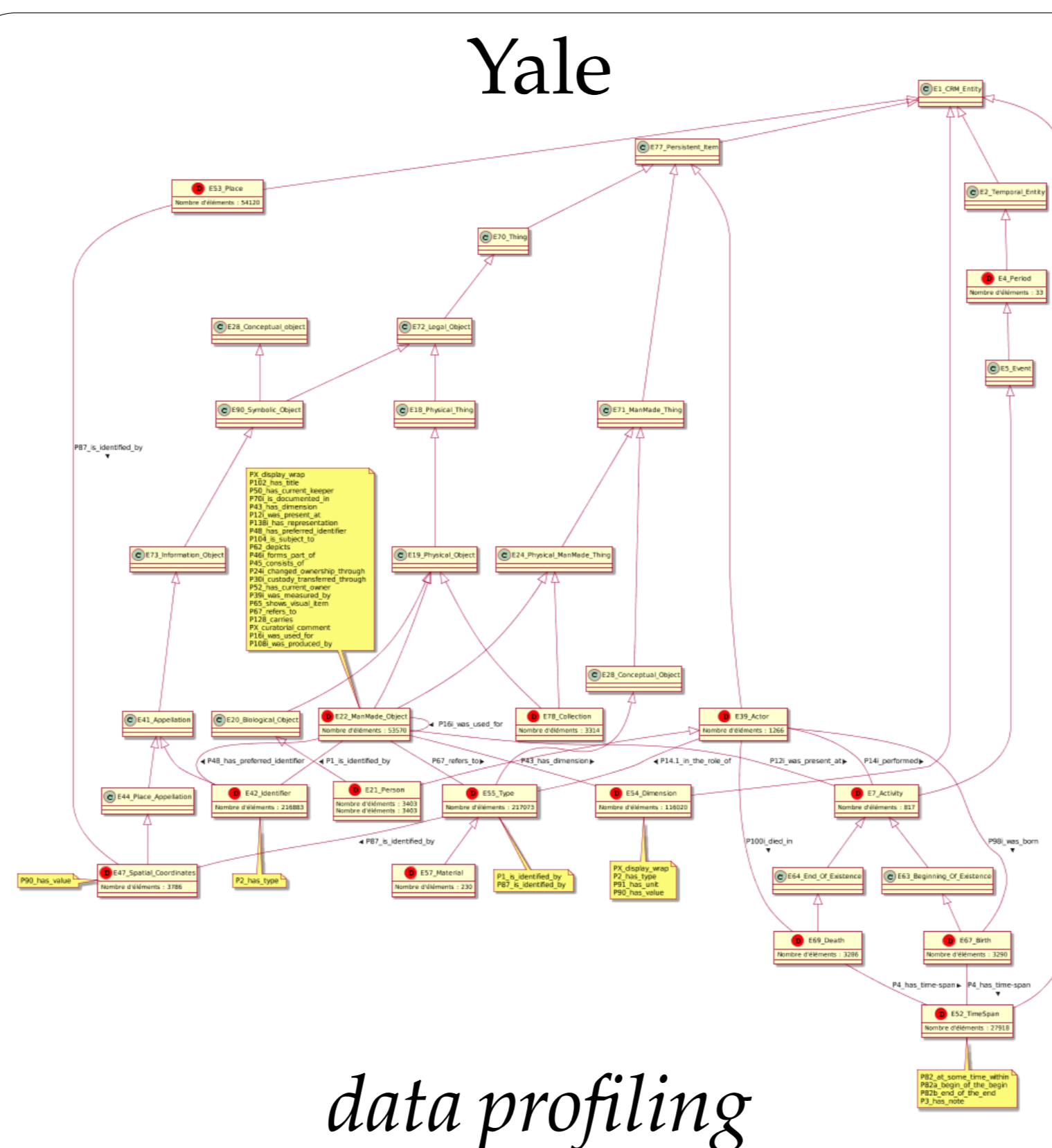
## Interfaçage



### Création d'un scénario d'utilisation

Ontologie : CIDOC-CRM (héritage culturel)

- Yale (Musées britanniques)
- Doremus (Musique classique)
- OpenArcheo (Archéologie)
- Archaeology Data Service (Archéologie)



#### • Contraintes :

E21 Person(X) -> P14 carried out by(Y,X) .

#### • Requête :

?(X) <- E21 Person(X)

#### • Requête réécrite :

?(X) <- E21 Person(X), P14 carried out by(Y,X) .

#### • Requête SPARQL :

```
SELECT DISTINCT ?X
WHERE {
  ?X a <http://erlangen-crm.org/current/E21.Person>.
  ?Y <http://erlangen-crm.org/current/P14.carried.out.by> ?X.
}
```