
LUBM4OBDA: Benchmarking OBDA Systems with Inference and Meta Knowledge

Julián Arenas-Guerrero*, María S. Pérez and Oscar Corcho

Universidad Politécnica de Madrid, Spain

E-mail: julian.arenas.guerrero@upm.es

**Corresponding Author*

Received 02 November 2022; Accepted 30 January 2024;
Publication 22 February 2024

Abstract

Ontology-based data access focuses on enabling query evaluation over heterogeneous relational databases according to the model represented by an ontology. The relationships between the ontology and the data sources are commonly defined with declarative mappings, which are used by systems to perform SPARQL-to-SQL query translation or to generate RDF dumps from the relational databases. Besides the potential homogenization of data because of using an ontology, some additional advantages of this paradigm are that it may allow applying reasoning thanks to the ontology, as well as querying for meta knowledge, which describes statements with information such as provenance or certainty. In this paper, (i) we adapt a widely used RDF graph store benchmark, namely LUBM, for ontology-based data access, (ii) extend the benchmark for the evaluation of queries that exploit meta knowledge, and (iii) apply it for performance evaluation of state-of-the-art declarative mapping systems. Our proposal, the LUBM4OBDA Benchmark, considers inference capabilities that are not covered by previous

Journal of Web Engineering, Vol. 22_8, 1163–1186.

doi: 10.13052/jwe1540-9589.2284

© 2024 River Publishers

ontology-based data access benchmarks, and it is the first one for the evaluation of meta knowledge and the RDF-star data model. The experimental evaluation shows that current virtualization systems cannot handle some advanced inference tasks, and that optimizations are needed to scale RDF-star materialization.

Keywords: OBDA, semantic web, ontology, data integration.

1 Introduction

Relational databases (RDBs) are widely used by organizations to manage their data. Some application scenarios can benefit from exploiting this data as knowledge graphs [22], to potentially homogenize data using ontologies, applying reasoning over ontologies, get new insights, or for semantic data integration use cases [39, 41]. *Ontology-based data access* [40] (OBDA) is a paradigm for making data sources such as RDBs (or local schemas) available through a standardized and common view (or global schema). This view is given in the form of an ontology, which also enriches the database with context by providing background knowledge, reasoning capabilities, and interlinking with other knowledge bases. The relationships between the global and local schemas are usually defined through *mappings* [27,39] which populate the ontology with instances generated from the RDB. As a result, it is possible to perform semantic queries independently of the structure of the original RDB.

In the Semantic Web community, a set of standards has been proposed: OWL [21] and RDFS [8] are the languages used to encode ontologies, RDF [13] allows representing data, and queries are expressed using SPARQL [18]. The mappings are specified using R2RML [14], the W3C recommendation to map RDBs to RDF that follows the *global as view* [27] approach, in which the mappings define, for each element in the global schema, a query over the local schemas. In addition, RML [24] is a well-known superset of R2RML to map not only RDBs, but also other types of data sources such as CSV, JSON or XML.

OBDA may be performed in two different manners. The first is query translation or *virtualization* [41], in which the mappings are used to translate (or unfold) SPARQL queries to SQL [32] *on-the-fly*, which produces the result set. The other is data translation, or *materialization* [4], which uses the mappings to transform the data into RDF. The generated RDF can then be loaded into a triplestore (i.e., an RDF graph store) to answer SPARQL

queries. Hence, the OBDA materialization approach consists of two phases, first an offline phase in which the knowledge graph is created by a materialization system and loaded into a triplestore, and then an online phase in which it is queried.

Recently, *meta knowledge* [34] has gained attention in the Semantic Web community. Meta knowledge allows describing (or annotating) RDF triples with information such as provenance or certainty. Different approaches have been proposed [16, 20, 28, 29, 33, 34], even extensions of RDF and SPARQL (*RDF-star* and *SPARQL-star* [19]) to facilitate the annotation of RDF triples. This growing interest in statement-level metadata is observed in the recent creation of an RDF-star Working Group¹ and the number of systems that have adopted RDF-star and SPARQL-star², including several triplestores and two OBDA systems, Ontop [38] for virtualization, and Morph-KGC [3] for materialization. While a benchmark to evaluate meta knowledge in triplestores has been proposed [30], a benchmark for OBDA systems is still missing.

This work extends a widely used RDF store benchmark, namely the Lehigh University Benchmark [17] (LUBM), to evaluate OBDA systems. Our proposal, the LUBM4OBDA Benchmark, allows evaluating inference capabilities that are not covered by previous OBDA benchmarks. LUBM4OBDA additionally considers meta knowledge, and it is the first benchmark for the evaluation of meta knowledge in OBDA systems. We have conducted a performance study of state-of-the-art OBDA systems using LUBM4OBDA considering virtualization and materialization.

This work makes the following contributions in the OBDA field:

- It extends the LUBM Benchmark to OBDA. We provide mappings to a relational representation of the original RDF data. This enables not only the comparison of OBDA systems but also with triplestores.
- It further extends the LUBM Benchmark to consider meta knowledge. We provide mappings and additional queries for well-known meta knowledge approaches, namely, standard reification, singleton properties and RDF-star.
- It evaluates the proposed benchmark with state-of-the-art open-source systems. We consider both, the virtualization approach with two relational database management systems (RDBMSs) and one virtual knowledge graph system, and the materialization approach with two

¹<https://www.w3.org/groups/wg/rdf-star>

²<https://w3c.github.io/rdf-star/implementations.html>

knowledge graph generation systems. The performance evaluation reveals the inference limitations of current systems, and the need for additional developments and optimizations for RDF-star in OBDA.

The rest of the paper is structured as follows: Section 2 introduces meta knowledge approaches and previous OBDA benchmarks; Section 3 presents the LUBM4OBDA Benchmark and its elements (ontology, dataset, mappings, and queries); Section 4 reports our evaluation with the LUBM4OBDA Benchmark using virtualization and materialization OBDA systems in the state of the art; finally, Section 5 discusses the prominence of our proposal and the results obtained, and Section 6 wraps up with some conclusions.

2 Preliminaries

In this section, we first briefly describe some popular meta knowledge approaches in RDF that we include in our benchmark. Then we present previous works on OBDA benchmarking.

2.1 Meta Knowledge

Meta knowledge [34], also known as reification or statement-level metadata, allows describing RDF triples, for instance by providing provenance or certainty information for them. A large amount of meta knowledge alternatives has emerged since the inception of RDF. In our work we consider those selected by the REF Benchmark [30]: *standard reification*,³ *singleton property* and *RDF-star*.

The W3C RDF Primer [28] introduced standard reification as an approach enabling meta knowledge in RDF using a built-in vocabulary. In standard reification, an RDF triple is assigned a unique identifier (typically a blank node), which can then be described with the built-in vocabulary. Specifically, the identifier is typed with `rdf:Statement`, and the subject, predicate, and object of the triples are described using the properties `rdf:subject`, `rdf:predicate` and `rdf:object` respectively. Meta knowledge is possible by further describing the identifier with additional properties such as provenance.

³Similar to previous works [20], here we use the term *reification* to refer to meta knowledge, and *standard reification* to refer to the meta knowledge approach presented in the W3C RDF Primer [28].

Listing 1 Example of standard reification.

```
_:50482 rdf:type rdf:Statement .
_:50482 rdf:subject person:Alice .
_:50482 rdf:predicate ex:knows .
_:50482 rdf:object person:Bob .
_:50482 person:statedBy person:Charlie .
```

Singleton properties [29] use unique predicates (conventionally with IRI fragments) in the RDF triples to be described. The property `rdf:singletonPropertyOf` is then used to associate the unique predicates with the original ones. It is then possible to use the unique predicates in the RDF triples to describe the statement using the unique predicates as the subject of additional triples.

Listing 2 Example of singleton property.

```
person:Alice ex:knows#1 person:Bob .
ex:knows#1 rdf:singletonPropertyOf ex:knows .
ex:knows#1 person:statedBy person:Charlie .
```

RDF-star [19] is an extension of RDF to describe triples in a more compact way. RDF-star enables meta knowledge through the notion of *quoted triple*. A quoted triple is a triple appearing in the subject or object of another triple. In this manner, the quoted triple represents the statement being described, and the triple containing it provides the meta knowledge. SPARQL-star [19] is the extension of the SPARQL language to query RDF-star graphs.

Listing 3 Example of RDF-star.

```
person:Alice ex:knows person:Bob .
<< person:Alice ex:knows person:Bob >> person:statedBy person:Charlie .
```

2.2 Related Work

OBDA systems have been evaluated even before the inception of R2RML. Table 1 shows a comparison of existing OBDA benchmarks. The Berlin SPARQL Benchmark [7] (BSBM) is a benchmark in the e-commerce domain that aims at comparing the performance of native RDF stores and virtualization OBDA systems. The evaluation of BSBM included the D2R Server [6], which uses a custom mapping language, but after the standardization of R2RML in 2012, most of the current systems have adopted this W3C Recommendation. Recently, the work presented in [11] extends BSBM to R2RML and uses this benchmark to compare state-of-the-art virtualization systems.

Table 1 Comparison of OBDA benchmarks

Benchmark	Domain	Inference	Meta knowledge	Mappings
BSBM	E-commerce	No	No	SQL base tables
NPD	Energy	Yes	No	R2RML views
GTFS-Madrid	Transport	No	No	SQL base tables
LUBM4OBDA	University	Yes	Yes	R2RML views and SQL base tables

Similar to the LUBM4OBDA Benchmark, the dataset in BSBM can be scaled in size; however, the reasoning capabilities covered by BSBM are limited, the mappings are simple (they only consider SQL base tables) and it does not allow for meta knowledge evaluation.

The NPD Benchmark [25], in the energy domain, is based on the Norwegian Petroleum Directorate FactPages. The benchmark consists of an ontology, a dataset that can be scaled in size using VIG [26], the mappings and a set of queries. The benchmark focuses on reasoning and query rewriting over the ontology, which uses the OWL 2 QL profile.⁴ Mappings use R2RML views [5], i.e., SQL queries over the RDB. The LUBM4OBDA Benchmark considers inference capabilities not included in NPD, has more complex mappings (e.g., in the number of joins in the R2RML views), and additionally provides meta knowledge benchmarking. ForBackBench [1] is a recent reasoning OBDA benchmark, but it focuses on the evaluation of specific OBDA algorithms rather than on the end-to-end evaluation of systems.

The GTFS-Madrid-Bench [12] is a benchmark in the transport domain, which uses the linked GTFS vocabulary⁵ (derived from the GTFS format) to describe public transport routes, stations, etc. Similar to the NPD Benchmark, the GTFS-Madrid-Bench also uses VIG to scale the relational datasets (but it also considers additional data formats such as CSV or XML). The mappings are simpler than the ones of the LUBM4OBDA Benchmark, since they only take into account SQL base tables. The query set in GTFS-Madrid-Bench does not consider inference, and the benchmark does not cover meta knowledge either.

Unlike the former benchmarks, which were built from scratch, we extend an existing benchmark for RDF graph stores, namely LUBM [17]. The reason for this is that LUBM requires more complex mappings (e.g., in the number of joins, or the use of both SQL base tables and R2RML views) and considers

⁴<https://www.w3.org/TR/owl2-profiles/>

⁵<https://github.com/OpenTransport/linked-gtfs>

reasoning tasks that were not considered in the aforementioned benchmarks. Furthermore, our proposal considers meta knowledge, following the design decisions (regarding meta knowledge approaches and query expressivity) of a recent benchmark [30] for the evaluation of reification approaches in RDF graph stores.

The requirements for OBDA benchmarks have been defined in [25]. These requirements are grouped in three layers: conceptual (the ontology), data (sources) and mapping. Since LUBM was originally designed to evaluate graph stores, it does not meet the requirements for the data and mapping layers. In our work, we reuse a recent data generator [37] of a relational representation of LUBM which addresses the data requirements, and we solve the mapping requirements by creating them. Although the LUBM ontology is small [25], we show that it considers reasoning tasks beyond those in the aforementioned benchmarks; hence LUBM4OBDA can be used as a complementary reasoning benchmark to those.

3 The LUBM4OBDA Benchmark

The LUBM4OBDA Benchmark consists of an ontology, a generator for relational data, a set of mappings, and two sets of queries, one involving inference and the other involving meta knowledge. All resources are available online at GitHub⁶ and archived at Zenodo.⁷ In the following we introduce these elements.

3.1 Ontology

LUBM4OBDA considers the original Univ-Bench ontology [17] from LUBM. The ontology is expressed in OWL Lite and it describes the university domain, with departments, research groups, professors, students, etc. The ontology is composed of 43 classes, 25 object properties, and 7 data properties.

3.2 Dataset

LUBM4OBDA uses the synthetic data generator of the Linköping GraphQL Benchmark [37] (LinGBM), a benchmark for GraphQL servers based on

⁶<https://github.com/oeg-upm/lubm4obda>

⁷<https://doi.org/10.5281/zenodo.7110221>

LUBM. LinGBM extends the Univ-Bench Artificial data generator [17] to generate SQL data dumps. The number of universities is used as the scaling parameter to create SQL data dumps (MySQL or PostgreSQL). To evaluate meta knowledge, we have additionally created a new column associated to the property `ub:undergraduateDegreeYear` in the table associated to the entity `ub:graduateStudent`, thus annotating the undergraduate degrees with the year in which they were awarded. The database schema consists of 14 tables that have between 1 and 8 columns. We provide a docker container⁸ to facilitate the use of the data generator. We also made available SQL data dumps for scaling factors 1, 10, 100 and 1000 in Zenodo.

3.3 Mappings

We have created the mappings from the relational representation of the data to the Univ-Bench ontology. In addition to the mappings for the original RDF representation of the data, we have created three additional mappings that consider the meta knowledge approaches described in Section 2.1 (i.e., standard reification, singleton property and RDF-star). The mappings use the R2RML and RML languages and their respective extensions for RDF-star, R2RML-star [38] and RML-star [15]. The mappings contain a high number of joins with respect to previous OBDA benchmarks (up to four joins are required in R2RML views as shown in Listing 4). Differences in case sensitivity among relational database management systems lead to the need for specific mappings for different RDBMSs. That is the case of the NPD Benchmark, which has different mapping versions for MySQL and PostgreSQL (GTFS-Madrid-Bench and BSBM have a unique mapping for all RDBMSs, but this is because they only consider SQL base tables). To keep the benchmark simple, we have used aliases within the R2RML views so that different RDBMSs use the same mappings. The original mappings (i.e., without meta knowledge) have a total of 22 triples maps, 36 object maps and 13 `rr:class` properties, and they consider both SQL base tables and R2RML views.

Listing 4 Triples map that links authors to publications.

```
<#CoAuthorOfPublication>
  rr:logicalTable [
    rr:sqlQuery """
      SELECT publication.name AS publicationname,
```

⁸<https://hub.docker.com/r/oegdataintegration/lubm4obda>


```

graduateStudent.name AS graduatestudentname,
faculty.name AS facultyname,
department.nr AS departmentnr,
department.subOrganizationOf AS universitynr
FROM coAuthorOfPublication
INNER JOIN publication ON coAuthorOfPublication.publicationID=
    publication.nr
INNER JOIN faculty ON publication.mainAuthor=faculty.nr
INNER JOIN department ON faculty.worksFor=department.nr
INNER JOIN graduateStudent ON coAuthorOfPublication.
    graduateStudentID=graduateStudent.nr
""
];
rr:subjectMap [
  rr:template "http://www.department{departmentnr}.university{
    universitynr}.edu/{facultyname}/{publicationname}"
];
rr:predicateObjectMap [
  rr:predicate ub:publicationAuthor;
  rr:objectMap [
    rr:template "http://www.department{departmentnr}.university{
    universitynr}.edu/{graduatestudentname}"
  ];
].

```

3.4 Queries

LUBM4OBDA has two query sets: the original set from LUBM, and the new set exploiting meta knowledge. The former is composed of 14 SPARQL queries that consider different input sizes, selectivity, complexity (in terms of the number of classes and properties), class hierarchy information, and logical inference (e.g., subsumption or realization). A complete description of these queries can be found in the LUBM paper [17]. The meta knowledge query set consists of 4 SPARQL queries, with different versions for standard reification, singleton property and RDF-star (SPARQL-star is used to query RDF-star). We have selected the same meta knowledge approaches as the REF Benchmark [30], i.e., those that add the meta knowledge as triples. Listings 5–9 depict the SPARQL-star version of the meta knowledge queries. Queries 15 and 16 are similar to queries 1 and 2 from the original set respectively, but they also retrieve the meta knowledge about the year in which an undergraduate degree was awarded. Query 17 has high selectivity and asks for the annotation values. Query 18 asks for all triples that are

annotated with `ub:yearOfAward` in 2017. This can be done in SPARQL-star in two ways: binding the subject, object, and predicate of the quoted triple to different variables (query 18a) or binding the quoted triples to a single variable in the query (query 18b).

Listing 5 Query 15.

```
SELECT ?x ?y ?z
WHERE {
  ?x a ub:GraduateStudent .
  ?x ub:takesCourse <http://www.department0.university0.edu/
graduateCourse0> .
  << ?x ub:undergraduateDegreeFrom ?y >> ub:yearOfAward ?z .
}
```

Listing 6 Query 16.

```
SELECT ?x ?y ?z ?w
WHERE {
  ?x a ub:GraduateStudent .
  << ?x ub:undergraduateDegreeFrom ?y >> ub:yearOfAward ?w .
  ?x ub:memberOf ?z .
  ?y a ub:University .
  ?z a ub:Department .
  ?z ub:subOrganizationOf ?y .
}
```

Listing 7 Query 17.

```
SELECT ?x
WHERE {
  << <http://www.department14.university0.edu/graduateStudent106> ub:
undergraduateDegreeFrom <http://www.university1.edu> >> ub:
yearOfAward ?x .
}
```

Listing 8 Query 18a.

```
SELECT ?x ?y ?z
WHERE {
  << ?x ?y ?z >> ub:yearOfAward "2017"^^xsd:integer .
}
```

Listing 9 Query 18b.

```
SELECT ?x
WHERE {
  ?x ub:yearOfAward "2017"^^xsd:integer .
}
```

Table 2 Number of triples of the LUBM4OBDA datasets used in the evaluation for each meta knowledge alternative (and the basic one without meta knowledge). The considered data scaling factors are 1, 10, 100 and 1000.

Data scaling factor	Without meta knowledge	Standard reification	Singleton property	RDF-star
LUBM4OBDA 1	111.994	119.490	115.742	113.868
LUBM4OBDA 10	1.418.673	1.514.749	1.466.711	1.442.692
LUBM4OBDA 100	14.949.936	15.960.184	15.455.060	15.202.498
LUBM4OBDA 1000	148.971.695	159.062.159	154.016.927	151.494.311

4 Experimental Evaluation

In this section, we present a performance evaluation using the LUBM4OBDA Benchmark. The evaluation considers OBDA systems of the two possible approaches, virtualization and materialization.

The RDBMSs used are MySQL v8.0.28 and PostgreSQL v14.2. We consider the scaling factors 1, 10, 100 and 1000 of the LUBM4OBDA dataset. Table 2 contains the number of triples of the datasets that have been used. The evaluation was performed on an Intel® Core™ i7-1165G7 (2.80GHz) and a memory of 40 GB RAM DDR4 (3200 MHz). All the times reported are the average time of 10 executions.

4.1 Virtualization

We have used Ontop v4.2.1 [9, 42] together with the aforementioned RDBMSs to evaluate virtualization with the LUBM4OBDA Benchmark. We selected Ontop based on the overview of systems presented in [41], in which Ontop stands out as the only open-source system with ontology inference capabilities.

Figure 1 shows the execution times obtained for the queries. Ontop was able to provide results for most of them, except for queries 10, 11 and 12. However, the system does not return complete results for queries 6–9, which involve the explicit relationship between UndergraduateStudent and Student, and the implicit (subsumption) relationship between GraduateStudent and Student. Ontop is only able to provide results for the explicit relationship. This behavior is also observed in query 10, which only involves the mentioned implicit relationship and does not return results. In addition, query 11 involves a transitive property (`owl:TransitiveProperty`) which the system is not able to handle. Query 12 requires realization, i.e., “*inference of the most specific concepts that an individual is an instance of*” [17]. As these queries do not produce results, we have omitted them in Figure 1.

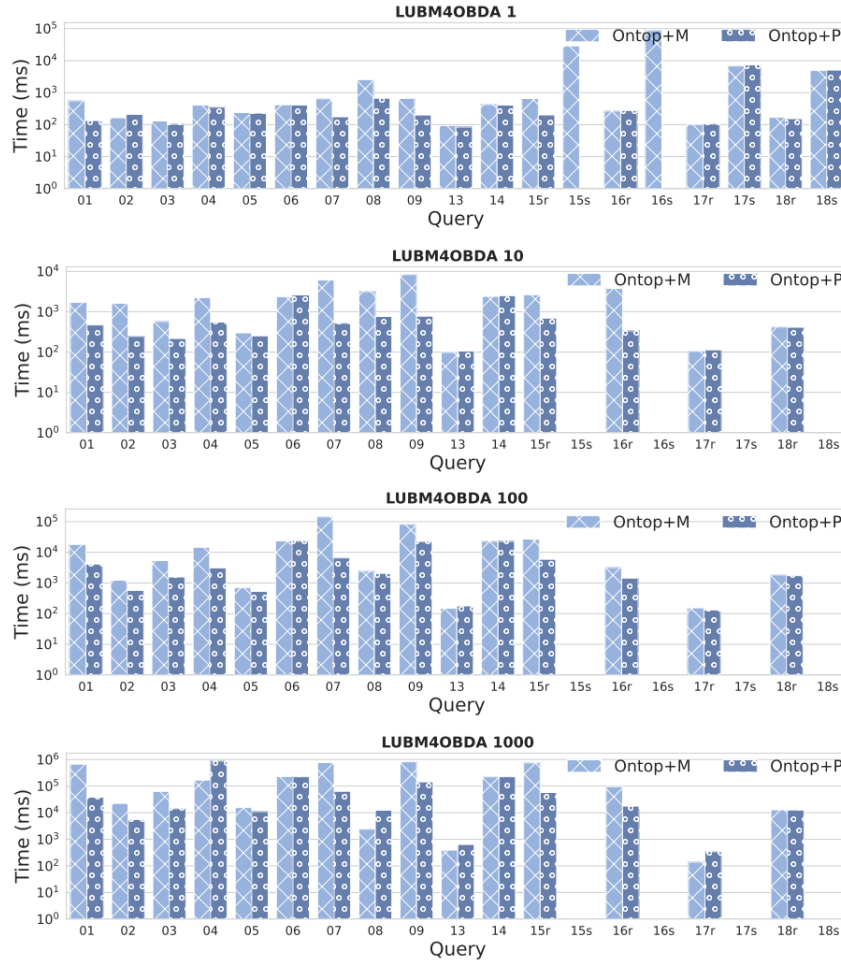


Figure 1 Query execution times of Ontop with MySQL and PostgreSQL for LUBM4OBDA with data scaling factors 1, 10, 100 and 1000. Time is reported in milliseconds using a logarithmic scale. The absence of a bar indicates that an error has been obtained or that the mapping is not supported. The legend specifies the OBDA system along with the underlying RDBMS (*M* denotes MySQL and *P* denotes PostgreSQL). An *r* is used along with the query number to refer to standard reification, and an *s* is used for singleton property.

Regarding queries with meta knowledge, it is observed that a singleton property results in an error in most of the cases. Indeed, Ontop is only successful for scaling factor 1, and it even fails in that case for queries 15 and 16 over PostgreSQL. Inspecting the unfolded SQL query generated by

Ontop for the query 18, it is observed that the system generates a subquery for each row in the `graduateStudent` table. The subqueries are combined using UNION ALL (i.e., the size of the unfolded query is dependent on the data), which results in a large SQL query that the system is not able to process. By contrast, standard reification is successful in all meta knowledge queries, with reasonable execution times. In this case, the unfolded query for query 18 retrieves data from the `graduateStudent` table without the need of subqueries. Ontop does not support SPARQL-star; however, a recent work has proposed R2RML-star [38] as an extension of R2RML to query RDBs using SPARQL-star and it extended Ontop to process it. Unfortunately, the source code is not publicly available; hence we do not consider the RDF-star meta knowledge approach here. It must be noted that “*the query is then supported provided that the queries and mappings map in an exact way and no quoted triples are bound to variables*” [38], i.e., the version of query 18 that uses a variable to retrieve quoted triples (i.e., query 18b) is not yet supported by the system.

In general, it is observed that PostgreSQL is faster in most cases. For instance, for queries 1 and 7 the difference between both RDBMSs for scaling factor 1000 exceeds one order of magnitude. However, there are some cases where MySQL slightly outperforms PostgreSQL (e.g., query 8 for scaling factor 1000). It is also observed that Ontop scales well for increasing dataset sizes, executing all queries without producing timeouts. As the data size increases, it seems that MySQL reduces the difference with respect to PostgreSQL, even outperforming it in some cases (e.g., queries 4, 8 or 13 for scaling factor 1000).

4.2 Materialization

We consider Ontop v4.2.1 [9, 42], with its materialization mode, and two state-of-the-art materialization systems, Morph-KGC v2.1.1 [2] and SDM-RDFizer v4.5.2 [23]. Figure 2 shows the total materialization time taken by each system.

For the original mappings (i.e., without meta knowledge) it is observed that Morph-KGC is the system that materializes the RDF knowledge graph faster. This is due to the implementation of the *mapping partitioning* strategy [2], which creates groups of mappings that are processed in parallel. Ontop is the second best performing system, followed by SDM-RDFizer. SDM-RDFizer implements techniques for efficient local join execution which are needed when processing data files (e.g., CSV or XML). However, the

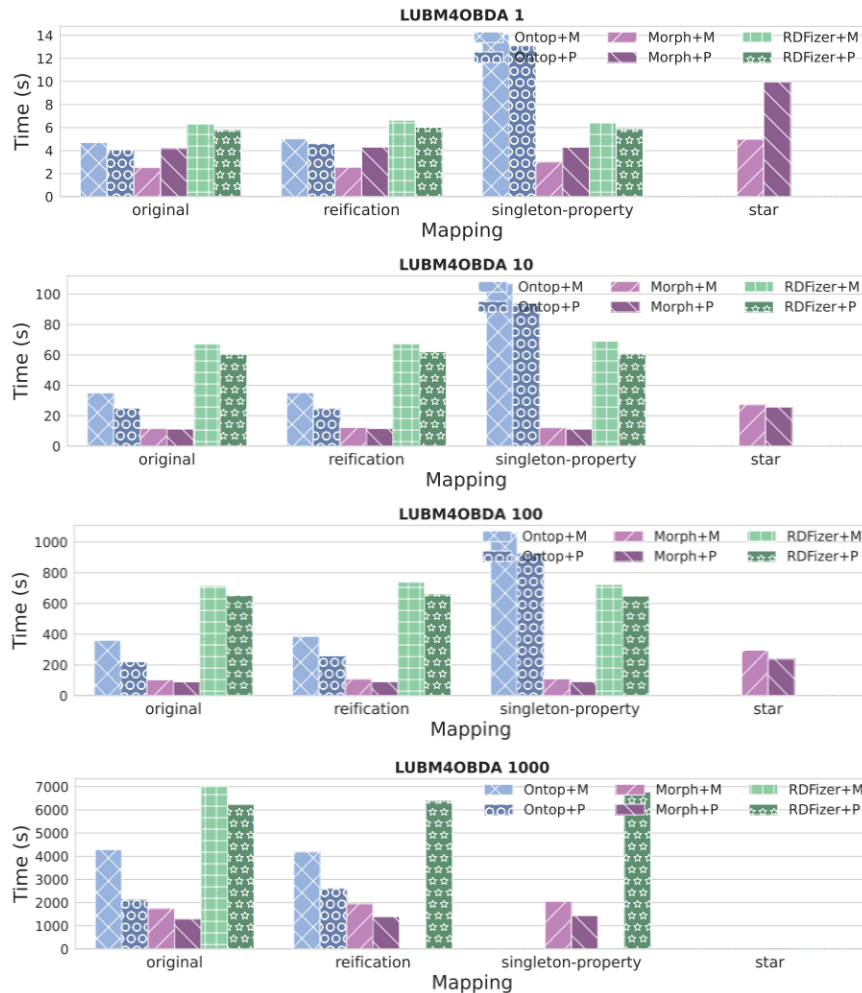


Figure 2 Materialization times for Ontop, Morph-KGC and SDM-RDFizer with MySQL and PostgreSQL for LUBM4OBDA with data scaling factors 1, 10, 100 and 1000. Time is reported in seconds. The absence of a bar indicates that an error has been obtained. The legend specifies the OBDA system along with the underlying RDBMS (*M* denotes MySQL and *P* denotes PostgreSQL).

LUBM4OBDA Benchmark pushes down joins in R2RML views, which is the common case when working with RDBs.

Regarding standard reification, we see that the results are similar to those of the mappings without meta knowledge. The most notable change is

that SDM-RDFizer produces an out-of-memory error for the largest scaling factor when using MySQL as the underlying RDBMS. The same happens for the singleton property, but in this case, Ontop was not able to materialize the knowledge graph for PostgreSQL and MySQL either. This is because Ontop exploits the fact that predicates are usually constants. For the singleton property, this does not hold (predicates are dynamically generated), and Ontop generates a SPARQL query for every different predicate, which are then translated to SQL and evaluated over the RDBMS.

Morph-KGC is the only open-source system currently supporting RDF-star. As can be observed in Figure 2, the materialization times of Morph-KGC for RDF-star are much higher than in the rest of the cases. This is because mapping partitioning has not yet been extended for RML-star, and therefore the system does not use parallelization in this case. Indeed, for the largest scaling factor, the system results in an out-of-memory error, since the entire knowledge graph has to be stored in memory.

Similarly to virtualization, the underlying RDBMSs have an impact on execution times. Although for small datasets MySQL sometimes outperforms PostgreSQL, we see that for large data sources PostgreSQL performs significantly better.

5 Discussion

In this section we discuss the results obtained in the evaluation of the LUBM4OBDA Benchmark, and show how our proposal can contribute to future research and implementations in OBDA.

Since the inception of R2RML, many open-source (e.g., Ontop or Morph-RDB [31]) and commercial (e.g., Ultrawrap [36] or Mastro [10]) systems have implemented virtualization techniques. However, the inference capabilities of these systems are diverse; Ontop and Mastro provide reasoning capabilities over OWL 2 ontologies, Ultrawrap is limited to RDFS, and Morph-RDB does not provide reasoning at all. In the evaluation of systems with the LUBM4OBDA Benchmark, we have shown that a cutting-edge and mature system like Ontop is still missing some advanced inference capabilities. Given that Ontop already covers the reasoning capabilities included in the NPD Benchmark [25], the LUBM4OBDA Benchmark can be used to further enhance the system by providing a testbed to evaluate new inference tasks. For instance, transitivity could be implemented in Ontop using SQL recursion [35] and evaluated with LUBM4OBDA. As for meta knowledge in virtualization, our evaluation shows that singleton properties are not currently

efficient in OBDA; this meta knowledge approach does not scale to large volumes of data and even fails to process queries over small-sized knowledge graphs. As the adoption of RDF-star increases, more OBDA systems will provide support for R2RML-star and RML-star. Although LUBM4OBDA does not consider all possible modelling patterns of meta knowledge (e.g., it does not consider quoted triples in the objects) our evaluation shows that the benchmark represents a challenge for current systems. The LUBM4OBDA Benchmark can contribute to new research and improvements of these systems, such as enabling mappings from a variable binding to a quoted triple, as it is the only testbed of the state of the art that provides support for RDF-star.

Regarding materialization, our evaluations show that systems specifically designed for that task (i.e., Morph-KGC and SDM-RDFizer) have a similar performance for the standard reification and singleton properties approaches. The generation of RDF-star is still in its infancy, and only one open-source system, namely Morph-KGC, provides support for it. As shown in Section 4.2, the impact of mapping partitioning in the materialization process is notable. The efficient generation of RDF-star with declarative mappings will require the extension of this optimization to R2RML-star and RML-star. The LUBM4OBDA Benchmark can be used by researchers and developers to test mapping partitioning and other potential optimizations addressing the generation of RDF-star graphs.

6 Conclusions

In this paper, we have presented the LUBM4OBDA Benchmark, a new benchmark for ontology-based data access over relational databases. Unlike previous OBDA benchmarks that were built from scratch, we have reused a widely used benchmark for RDF graph stores, namely LUBM [17]. We have shown that our proposal evaluates inference capabilities (present in LUBM) that were not considered in previous benchmarks, such as realization, subsumption relationship, or transitivity. We have further extended the benchmark with meta knowledge considering standard reification, singleton property, and RDF-star. Finally, we have conducted a performance study of state-of-the-art OBDA systems (with both virtualization and materialization approaches), showing that Ontop does not currently support all the ontology inference capabilities included in the benchmark, and that more efficient strategies are needed for the materialization of RDF-star graphs.

Acknowledgements

This work was funded partially by the project Knowledge Spaces: Técnicas y herramientas para la gestión de grafos de conocimientos para dar soporte a espacios de datos (Grant PID2020-118274RB-I00, funded by MCIN/AEI/10.13039/501100011033) and also received partial financial support in the frame of the Euratom Research and Training Programme 2019–2020 under grant agreement No 900018 (ENTENTE project).

References

- [1] Alhazmi, A., Blount, T., Konstantinidis, G., 2022. ForBackBench: A Benchmark for Chasing vs. Query-Rewriting. *Proceedings of the VLDB Endowment*, 15(8), pp. 1519–1532. doi: 10.14778/3529337.3529338
- [2] Arenas-Guerrero, J., Chaves-Fraga, D., Toledo, J., Pérez, M.S., Corcho, O., 2024. Morph-KGC: Scalable knowledge graph materialization with mapping partitions. *Semantic Web*. doi: 10.3233/SW-223135.
- [3] Arenas-Guerrero, J., Iglesias-Molina, A., Chaves-Fraga, D., Garijo, D., Corcho, O., Dimou, A., 2024. Declarative generation of RDF-star graphs from heterogeneous data. Submitted to *Semantic Web*. URL: <https://www.semantic-web-journal.net/system/files/swj3602.pdf>.
- [4] Arenas-Guerrero, J., Scrocca, M., Iglesias-Molina, A., Toledo, J., Pozo-Gilo, L., Doña, D., Corcho, O., Chaves-Fraga, D., 2021. Knowledge Graph Construction with R2RML and RML: An ETL System-based Overview, in: *Proceedings of the 2nd International Workshop on Knowledge Graph Construction*, CEUR Workshop Proceedings. URL: <http://ceur-ws.org/Vol-2873/paper11.pdf>.
- [5] Arenas-Guerrero, J., Alobaid, A., Navas-Loro, M., Pérez, M.S., Corcho, O., 2023. Boosting Knowledge Graph Generation from Tabular Data with RML Views, in: *Proceedings of the 20th Extended Semantic Web Conference*, Springer Nature Switzerland. pp. 484-501. doi: 10.1007/978-3-031-33455-9_29.
- [6] Bizer, C., Cyganiak, R., 2006. D2R Server – Publishing Relational Databases on the Semantic Web, in: *International Semantic Web Conference*.
- [7] Bizer, C., Schultz, A., 2009. The Berlin SPARQL Benchmark. *International Journal on Semantic Web and Information Systems* 5(2), pp. 1–24. doi: 10.4018/jswis.2009040101.

- [8] Brickley, D., Guha, R., 2014. RDF Schema 1.1. W3C Recommendation. URL: <https://www.w3.org/TR/rdf-schema/>.
- [9] Calvanese, D., Cogrel, B., Komla-Ebri, S., Kontchakov, R., Lanti, D., Rezk, M., Rodriguez-Muro, M., Xiao, G., 2017. Ontop: Answering SPARQL queries over relational databases. *Semantic Web* 8, pp. 471–487. doi: 10.3233/SW-160217.
- [10] Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Poggi, A., Rodriguez-Muro, M., Rosati, R., Ruzzi, M., Savo, D.F., 2011. The MASTRO system for ontology-based data access. *Semantic Web* 2(1), pp. 43–53. doi: 10.3233/SW-2011-0029.
- [11] Chaloupka, M., Necasky, M., 2020. Using Berlin SPARQL Benchmark to Evaluate Relational Database Virtual SPARQL Endpoints. Submitted to *Semantic Web*. URL: <https://www.semantic-web-journal.net/system/files/swj2473.pdf>.
- [12] Chaves-Fraga, D., Priyatna, F., Cimmino, A., Toledo, J., Ruckhaus, E., Corcho, O., 2020. GTFS-Madrid-Bench: A Benchmark for Virtual Knowledge Graph Access in the Transport Domain. *Journal of Web Semantics* 65, 100596. doi: 10.1016/j.websem.2020.100596.
- [13] Cyganiak, R., Wood, D., Lanthaler, M., 2014. RDF 1.1 Concepts and Abstract Syntax. W3C Recommendation. URL: <https://www.w3.org/TR/rdf11-concepts/>.
- [14] Das, S., Sundara, S., Cyganiak, R., 2012. R2RML: RDB to RDF Mapping Language. W3C Recommendation. URL: <http://www.w3.org/TR/r2rml/>.
- [15] Delva, T., Arenas-Guerrero, J., Iglesias-Molina, A., Corcho, O., Chaves-Fraga, D., Dimou, A., 2021. RML-star: A Declarative Mapping Language for RDF-star Generation, in: *International Semantic Web Conference, P&D*, pp. 1–5.
- [16] Dividino, R., Sizov, S., Staab, S., Schueler, B., 2009. Querying for provenance, trust, uncertainty and other meta knowledge in RDF. *Journal of Web Semantics* 7, 204–219. doi: 10.1016/j.websem.2009.07.004.
- [17] Guo, Y., Pan, Z., Heflin, J., 2005. LUBM: A benchmark for OWL knowledge base systems. *Journal of Web Semantics* 3, 158–182. doi: 10.1016/j.websem.2005.06.005.
- [18] Harris, S., Seaborne, A., 2013. SPARQL 1.1 Query Language. W3C Recommendation. URL: <https://www.w3.org/TR/sparql11-query/>.
- [19] Hartig, O., 2017. Foundations of RDF* and SPARQL* (An Alternative Approach to Statement-Level Metadata in RDF), in: *Proceedings of*

- the 11th Alberto Mendelzon International Workshop on Foundations of Data Management and the Web, CEUR Workshop Proceedings. URL: <http://ceur-ws.org/Vol-1912/paper12.pdf>.
- [20] Hernández, D., Hogan, A., Krötzsch, M., 2015. Reifying RDF: What Works Well With Wikidata?, in: Proceedings of the 11th International Workshop on Scalable Semantic Web Knowledge Base Systems, CEUR Workshop Proceedings. pp. 32–47. URL: http://ceur-ws.org/Vol-1457/SSWS2015_paper3.pdf.
- [21] Hitzler, P., Krötzsch, M., Parsia, B., Patel-Schneider, P.F., Sebastian, R., 2012. OWL 2 Web Ontology Language. W3C Recommendation. URL: <https://www.w3.org/TR/owl2-primer/>.
- [22] Hogan, A., Blomqvist, E., Cochez, M., d’Amato, C., Melo, G. D., Gutierrez, C., ... & Zimmermann, A., 2021. Knowledge graphs. *ACM Computing Surveys*, 54(4), pp. 1–37. doi: 10.1145/3447772.
- [23] Iglesias, E., Jozashoori, S., Chaves-Fraga, D., Collarana, D., Vidal, M.E., 2020. SDM-RDFizer: An RML Interpreter for the Efficient Creation of RDF Knowledge Graphs, in: Proceedings of the 29th ACM International Conference on Information and Knowledge Management, Association for Computing Machinery. pp. 3039–3046. doi: 10.1145/340531.3412881.
- [24] Iglesias-Molina, A., Van Assche, D., Arenas-Guerrero, J., De Meester, B., Debruyne, C., Jozashoori, S., Maria, P., Michel, F., Chaves-Fraga, D., Dimou, A., 2023. The RML Ontology: A Community-Driven Modular Redesign After a Decade of Experience in Mapping Heterogeneous Data to RDF., in: Proceedings of the 22nd International Semantic Web Conference, Springer Nature Switzerland. pp. 152–175. doi: 10.1007/978-3-031-47243-5_9.
- [25] Lanti, D., Rezk, M., Xiao, G., Calvanese, D., 2015. The NPD Benchmark: Reality Check for OBDA Systems, in: Proceedings of the 18th International Conference on Extending Database Technology, OpenProceedings.org. pp. 617–628. URL: <https://openproceedings.org/2015/conf/edbt/paper-350.pdf>.
- [26] Lanti, D., Xiao, G., Calvanese, D., 2019. VIG: Data scaling for OBDA benchmarks. *Semantic Web* 10, 413–433. doi: 10.3233/SW-180336.
- [27] Lenzerini, M., 2002. Data Integration: A Theoretical Perspective, in: Proceedings of the 21st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems, Association for Computing Machinery. pp. 233–246. doi: 10.1145/543613.543644.

- [28] Manola, F., Miller, E., McBride, B., et al., 2004. RDF primer. W3C Recommendation. URL: <https://www.w3.org/TR/rdf-primer/>.
- [29] Nguyen, V., Bodenreider, O., Sheth, A., 2014. Don't like RDF Reification? Making Statements about Statements Using Singleton Property, in: Proceedings of the 23rd International Conference on World Wide Web, Association for Computing Machinery. pp. 759–770. doi: 10.1145/2566486.2567973.
- [30] Orlandi, F., Graux, D., O'Sullivan, D., 2021. Benchmarking rdf metadata representations: Reification, singleton property and rdf, in: 2021 IEEE 15th International Conference on Semantic Computing, pp. 233–240. doi: 10.1109/ICSC50631.2021.00049.
- [31] Priyatna, F., Corcho, O., Sequeda, J., 2014. Formalisation and experiences of R2RML-based SPARQL to SQL query translation using morph, in: Proceedings of the 23rd International Conference on World Wide Web, Association for Computing Machinery. pp. 479–490. doi: 10.1145/2566486.2567981.
- [32] Rodríguez-Muro, M., Rezk, M., 2015. Efficient SPARQL-to-SQL with R2RML Mappings. *Journal of Web Semantics* 33, 141–169. doi: 10.1016/j.websem.2015.03.001.
- [33] Sahoo, S.S., Bodenreider, O., Hitzler, P., Sheth, A., Thirunarayan, K., 2010. Provenance Context Entity (PaCE): Scalable Provenance Tracking for Scientific RDF Data, in: Proceedings of the 22nd Scientific and Statistical Database Management, Springer Berlin Heidelberg. pp. 461–470. doi: 10.1007/978-3-642-13818-8_32.
- [34] Sen, S., Katoriya, D., Dutta, A., Dutta, B., 2021. RDFM: An alternative approach for representing, storing, and maintaining meta-knowledge in web of data. *Expert Systems with Applications* 179, 115043. doi: 10.1016/j.eswa.2021.115043.
- [35] Sequeda, J.F., Arenas, M., Miranker, D.P., 2014. OBDA: Query Rewriting or Materialization? In Practice, Both!, in: Proceedings of the 13th International Semantic Web Conference, Springer International Publishing. pp. 535–551. doi: 10.1007/978-3-319-11964-9_34.
- [36] Sequeda, J.F., Miranker, D.P., 2013. Ultrawrap: SPARQL execution on relational data. *Journal of Web Semantics* 22, pp. 19–39. doi: 10.1016/j.websem.2013.08.002.
- [37] Sijin, C., Hartig, O., 2022. LinGBM: A Performance Benchmark for Approaches to Build GraphQL Servers, in: Proceedings of the 23rd International Conference on Web Information Systems Engineering. doi: 10.1007/978-3-031-20891-1_16

- [38] Sundqvist, L., 2022. Extending VKG Systems with RDF-star Support. URL: <https://ontop-vkg.org/publications/2022-sundqvist-rdf-star-ontop-msc-thesis.pdf>.
- [39] Van Assche, D., Delva, T., Haesendonck, G., Heyvaert, P., De Meester, B., Dimou, A., 2022. Declarative RDF graph generation from heterogeneous (semi-)structured data. *Journal of Web Semantics* 75, 100753. doi: 10.1016/j.websem.2022.100753.
- [40] Xiao, G., Calvanese, D., Kontchakov, R., Lembo, D., Poggi, A., Rosati, R., Zakharyashev, M., 2018. Ontology-Based Data Access: A Survey, in: *Proceedings of the 27th International Joint Conference on Intelligence, IJCAI, International Joint Conferences on Artificial Intelligence Organization*. pp. 5511–5519. doi: 10.24963/ijcai.2018/777.
- [41] Xiao, G., Ding, L., Cogrel, B., Calvanese, D., 2019. Virtual Knowledge Graphs: An Overview of Systems and Use Cases. *Data Intelligence* 1, pp. 201–223. doi: 10.1162/dint_a_00011.
- [42] Xiao, G., Lanti, D., Kontchakov, R., Komla-Ebri, S., Güzel-Kalaycı, E., Ding, L., Corman, J., Cogrel, B., Calvanese, D., Botoeva, E., 2020. The Virtual Knowledge Graph System Ontop, in: *Proceedings of the 19th International Semantic Web Conference, Springer International Publishing*. pp. 259–277. doi: 10.1007/978-3-030-62466-8_17.

Biographies



Julián Arenas-Guerrero received his B.Sc. degree in computer science engineering from the Universidad Complutense de Madrid, and his M.Sc. degree in artificial intelligence from the Universidad Politécnica de Madrid. He is currently pursuing a Ph.D. degree in artificial intelligence with the Universidad Politécnica de Madrid. His research interests include data integration, semantic web and knowledge graphs.



María S. Pérez is currently a Full Professor with the Universidad Politécnica de Madrid. She is also a part of the Board of Directors of BDVA and also a member of the Research and Innovation Advisory Group, a EuroHPC Joint Undertaking. She has coauthored four books and seven book chapters. She has published more than 100 articles in international journals and conferences. She has been involved in the organization of several workshops and conferences and has edited several proceedings books and special issues. She has participated in a number of EU projects (ENTENTE, BigStorage, Wf4Ever, PlanetData, SEMDATA, SCALUS, SemsorGrid4Env, SEALS, OntoGrid, and RAIL) and Spanish R&D projects (CABAHLA, España Virtual, myBigData, GeoBuddies, 4V, and Datos 4.0). Her research interests include data science, big data, machine learning, storage, high performance, and large-scale computing. She has served as a program committee member for many relevant conferences.



Oscar Corcho is currently a Full Professor with the Universidad Politécnica de Madrid where he is also co-director of the Ontology Engineering Group. He previously worked as a Marie Curie researcher at the University

of Manchester and as a research manager at the company iSOCO. He has published several books, among which “Ontological Engineering” stands out; it is used as an academic text in several Spanish and foreign universities, as well as more than 100 articles in journals, conferences and workshops. He is a member of the editorial committee of several journals, and regularly participates in the program committees of the most relevant conferences in the semantic web field, having directed the program committee of some of them.

