

Article

A GeoSPARQL Compliance Benchmark

Milos Jovanovik ^{1,2,*} , Timo Homburg ³  and Mirko Spasić ^{2,4} 

¹ Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje, 1000 Skopje, North Macedonia

² OpenLink Software Ltd., Croydon, Surrey CR0 0XZ, UK; mirko@matf.bg.ac.rs

³ i3mainz—Institute for Spatial Information & Surveying Technology, Mainz University of Applied Sciences, 55128 Mainz, Germany; timo.homburg@hs-mainz.de

⁴ Faculty of Mathematics, University of Belgrade, 11000 Belgrade, Serbia

* Correspondence: milos.jovanovik@finki.ukim.mk

Abstract: GeoSPARQL is an important standard for the geospatial linked data community, given that it defines a vocabulary for representing geospatial data in RDF, defines an extension to SPARQL for processing geospatial data, and provides support for both qualitative and quantitative spatial reasoning. However, what the community is missing is a comprehensive and objective way to measure the extent of GeoSPARQL support in GeoSPARQL-enabled RDF triplestores. To fill this gap, we developed the GeoSPARQL compliance benchmark. We propose a series of tests that check for the compliance of RDF triplestores with the GeoSPARQL standard, in order to test how many of the requirements outlined in the standard a tested system supports. This topic is of concern because the support of GeoSPARQL varies greatly between different triplestore implementations, and the extent of support is of great importance for different users. In order to showcase the benchmark and its applicability, we present a comparison of the benchmark results of several triplestores, providing an insight into their current GeoSPARQL support and the overall GeoSPARQL support in the geospatial linked data domain.

Keywords: GeoSPARQL; geospatial data; benchmark; RDF; SPARQL



Citation: Jovanovik, M.; Homburg, T.; Spasić, M. A GeoSPARQL Compliance Benchmark. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 487. <https://doi.org/10.3390/ijgi10070487>

Academic Editors: Rob Brennan, Brian Davis, Armin Haller, Beyza Yaman and Wolfgang Kainz

Received: 22 May 2021
Accepted: 10 July 2021
Published: 16 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The geospatial Semantic Web [1] as part of the Semantic Web [2] represents an ever-growing semantically interpreted wealth of geospatial information. The initial research [3] and the subsequent introduction of the OGC GeoSPARQL standard [4] formalized geospatial vector data representations (WKT [5] and GML [6]) in ontologies, and extended the SPARQL query language [7] with support for spatial relation operators.

Several RDF storage solutions have since adopted GeoSPARQL to various extents as features of their triplestore implementations [8,9]. These varying levels of implementation may lead to some false assumptions of users when choosing an appropriate triplestore implementation for their project. For example, some implementations allow for defining a coordinate reference system (CRS) [10] in a given WKT geometry literal as stated in the GeoSPARQL standard (e.g., GraphDB). Other implementations do not allow a CRS definition and instead only support the world geodetic system WGS84 (e.g., RDF4J) [11]. Such implementations, even though incomplete according to the GeoSPARQL standard, still cover many geospatial use-cases and can be useful in many scenarios. However, they are not useful, for example, for a geospatial authority that needs to work with many different coordinate system definitions.

The requirements of GeoSPARQL compliant triplestores have been clearly spelled out in the GeoSPARQL standard [4]. However, the Semantic Web and GIS community lack a compliance test suite for GeoSPARQL, which we contribute in this publication. We hope that our contribution may be added to the list of OGC conformance tests (OGC Test Suites:

<https://cite.openegeospatial.org/teamengine/> (accessed on 22 May 2021)), as they lack a suitable test suite for GeoSPARQL.

Our paper is organized as follows. In Section 2, we discuss existing approaches that worked towards evaluating geospatial triplestores, and Section 3 introduces the test framework of the benchmark and describes how the compliance tests were implemented. Section 4 describes the application of the defined test framework against different triplestore implementations, and we discuss the results in Section 5. In Section 6, we lay out the limitations of our approach, before concluding the work in Section 7.

2. Related Work

Most standards define requirements which need to be fulfilled to satisfy the standard definition. However, not all standards expose explicit descriptions on how to test compliance with their requirements or a test suite that tests the overall compliance to the standard.

GeoSPARQL [4], as an extension of the SPARQL [7] query language, defines an ontology model to represent vector geometries, their relations and serializations in WKT and GML, a set of geometry filter functions, an RDFS entailment extension, a query rewrite extension to simplify geospatial queries and further geometry manipulation functions.

First, it is important that we distinguish between performance benchmarks and compliance benchmarks. Performance benchmarks try to evaluate the performance of system, usually by employing a set of queries. Performance benchmarks may also consider semantically equivalent implementations that are not following the syntax specified by a given standard. On the other hand, compliance benchmarks are not concerned with the efficiency or overall performance of a system, but rather with its ability to fulfill certain requirements.

Several benchmark implementations targeting geospatial triplestores, such as the Geographica Series [12,13] or [9], try to evaluate the performance of geospatial function implementations. Both approaches originate from the Linked Data community. Additionally, Ref. [14] shows that the geospatial community is interested in benchmarking geospatial triplestores as well. Their benchmark includes a newly created dataset and tests GeoSPARQL filter functions. While the aforementioned benchmarks might reveal if functions are implemented, they do not necessarily reveal an incorrect implementation of a given function.

The Tests for Triplestores (TFT) benchmark [15] includes a GeoSPARQL subtest. However, the subtest used here is based on the six example SPARQL queries and the example dataset defined in Annex B of the GeoSPARQL standard [4]. Although these examples are a good starting point, they are of informative nature and are intended as guidelines. Therefore, any benchmark based solely on them does not even begin to cover all possible requirements or the multiple ways in which they have to be tested, in order for a system to be deemed as compliant with the standard.

Recently, the EuroSDR group reused the benchmark implementation of [14] to implement a small GeoSPARQL compliance benchmark (EuroSDR GeoSPARQL Test: <https://data.pldn.nl/eurohdr/geosparql-test> (accessed on 22 May 2021)). This compliance benchmark consists of 27 queries testing a selection of GeoSPARQL functions on a test dataset. In contrast to our benchmark, this implementation does not explicitly test all requirements defined in the GeoSPARQL standard. In particular, GML support, RDFS entailment support and the query rewrite extension, among others, have not been tested in this benchmark.

3. GeoSPARQL Compliance Benchmark

The GeoSPARQL compliance benchmark is based on the requirements defined in the GeoSPARQL standard [4]. The 30 requirements defined in the standard are grouped into six categories and refer to the core GeoSPARQL ontology model and a set of extensions which systems need to implement, and which need to be tested in our benchmark:

1. Core component (CORE): Defines the top-level spatial vocabulary components (Requirements 1–3);

2. Topology vocabulary extension (TOP): Defines the topological relation vocabular (Requirements 4–6);
3. Geometry extension (GEOEXT): Defines the geometry vocabulary and non-topological query functions (Requirements 7–20);
4. Geometry topology extension (GTOP): Defines topological query functions for geometry objects (Requirements 21–24);
5. RDFS entailment extension (RDFSE): Defines a mechanism for matching implicit (inferred) RDF triples that are derived based on RDF and RDFS semantics, i.e., derived from RDFS reasoning (Requirements 25–27);
6. Query rewrite extension (QRW): Defines query transformation rules for computing spatial relations between spatial objects based on their associated geometries (Requirements 28–30).

Each of the specified requirements may be tested using a set of guidelines which are loosely defined in the abstract test suite in Annex A of the GeoSPARQL standard [4]. While the abstract test suite defines the test purpose, method and type to verify if a specific requirement has been fulfilled, it does not define a concrete set of SPARQL queries and a test dataset which may be used for reference. We contribute the test dataset and the set of SPARQL queries to verify each requirement in this publication.

In the GeoSPARQL compliance benchmark, each requirement is tested by one or more SPARQL queries, where there is a single expected answer or a set of expected answers. The number of queries used to test a requirement, as well as the number of expected answers per query, depends on the nature of the requirement. For some of them, it is sufficient to have a single query and a single expected answer to test whether the system under testing complies with it. In contrast, other requirements have sub-requirements—for example, requirements which refer to multiple properties or functions, requirements referring to functions which can be used with geometries with different serializations, or requirements which need a broader coverage of cases, to make sure they are fully met. In these cases, multiple queries are used. Multiple logically equivalent expected answers are used when the answer of a SPARQL query can be technically expressed in different formats or literal serializations.

This approach of using queries and expected answers as tests allows us to measure the compliance of any RDF storage system by using the HOBBIT benchmarking platform [16,17].

The output of the benchmark is a percentage which measures the overall compliance of the tested system with the GeoSPARQL standard. It measures the number of supported requirements of the system, out of the 30 specified requirements, as a percentage.

3.1. Benchmark Dataset

The GeoSPARQL standard defines an example dataset for testing in its Annex B [4], which can be used with the set of six example test queries defined in the same section. This example dataset contains six geometries. We wanted to use this dataset, but given that we aimed to test all requirements of the standard, we had to substantially extend the dataset both with new geometries and additional properties of the existing geometries. Figure 1 shows the geometries included in our extended dataset, while Listing 1 contains an RDF excerpt of the dataset, in Turtle syntax that represents the geometry A (Point).

The extended benchmark dataset contains 13 geometries of Polygon, Point and LineString types, all expressed as both WKT and GML literals. The total size of the RDF dataset is over 300 triples. The dataset is available as part of the benchmark code [18,19], in RDF/XML, GeoJSON [20] and GML representations.

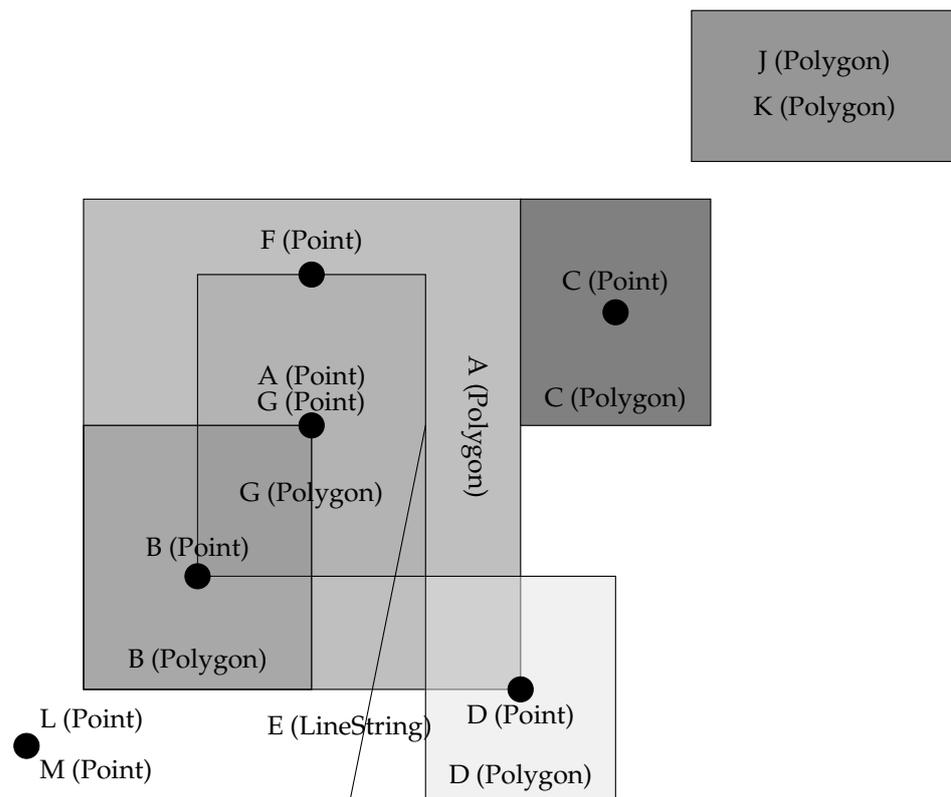


Figure 1. Abstract view of the geometries which are part of the benchmark dataset. Geometries A, B, C, D, G, J and K represent Polygon geometries and (aside from J and K) all have a center Point geometry as well. Geometry E represents a LineString geometry, while geometries F, L and M represent Point geometries. Geometries H and I are empty geometries and not visible in this figure. All geometries are represented in the CRS84 geodetic system, except for geometry M which is represented in EPSG:4326. Each geometry is represented both using WKT and GML literals.

3.2. Benchmark Queries

We provide here an overview of the approach we had in writing the queries used by the benchmark to test the requirements of the GeoSPARQL standard. The requirements are presented in order of the GeoSPARQL extension definitions presented in Section 3. The benchmark queries are available as part of the benchmark code [18], along with a summary table that maps the requirements to the relevant sets of queries, i.e., tests and sub-tests. The details about how each test and sub-test is scored are presented in Section 3.3.

Req. 1 Implementations shall support the SPARQL Query Language for RDF [7], the SPARQL Protocol for RDF [21] and the SPARQL Query Results XML Format [22].

We test requirement 1 with a single, basic SPARQL query which selects the first triple where geometry A is the subject. To get consistent results across different systems, we have to use a specific subject and have to order the results.

Req. 2 Implementations shall allow the RDFS [23] class `geo:SpatialObject` to be used in SPARQL graph patterns.

Req. 3 Implementations shall allow the RDFS class `geo:Feature` to be used in SPARQL graph patterns.

Requirements 2 and 3 are tested with single SPARQL queries, which select the first entity of type `geo:SpatialObject` and `geo:Feature`, respectively. In order to get consistent results for both queries across different systems, we order the results.

Req. 4 Implementations shall allow the properties `geo:sfEquals`, `geo:sfDisjoint`, `geo:sfIntersects`, `geo:sfTouches`, `geo:sfCrosses`, `geo:sfWithin`, `geo:sfContains`, `geo:sfOverlaps` to be used in SPARQL graph patterns.

Listing 1: An RDF excerpt of the benchmark dataset, in Turtle syntax, which represents a 2D point geometry.

```

1 @prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
2 @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
3 @prefix ex: <http://example.org/ApplicationSchema#> .
4 @prefix geo: <http://www.opengis.net/ont/geosparql#> .
5 @prefix sf: <http://www.opengis.net/ont/sf#> .
6 @prefix gml: <http://www.opengis.net/ont/gml#> .
7
8 ex:APointGeom
9   a                geo:Geometry,
10                  sf:Point,
11                  gml:Point ;
12   geo:isEmpty      false ;
13   geo:isSimple      true ;
14   geo:dimension     2 ;
15   geo:spatialDimension 2 ;
16   geo:coordinateDimension 2 ;
17   geo:asWKT         """
18                   <http://www.opengis.net/def/crs/OGC/1.3/CRS84> Point(-83.4 34.3)
19                   """^^geo:wktLiteral ;
20   geo:asGML         """
21                   <gml:Point xmlns:gml="http://www.opengis.net/ont/gml"
22                   srsName="http://www.opengis.net/def/crs/OGC/1.3/CRS84">
23                   <gml:pos>-83.4 34.3</gml:pos>
24                   </gml:Point>
25                   """^^geo:gmlLiteral .

```

Req. 5 Implementations shall allow the properties [geo:ehEquals](#), [geo:ehDisjoint](#), [geo:ehMeet](#), [geo:ehOverlap](#), [geo:ehCovers](#), [geo:ehCoveredBy](#), [geo:ehInside](#), [geo:ehContains](#) to be used in SPARQL graph patterns.

Req. 6 Implementations shall allow the properties [geo:rcc8eq](#), [geo:rcc8dc](#), [geo:rcc8ec](#), [geo:rcc8po](#), [geo:rcc8tppi](#), [geo:rcc8tpp](#), [geo:rcc8ntpp](#), [geo:rcc8ntppi](#) to be used in SPARQL graph patterns.

We test each of the requirements 4, 5 and 6 with eight different queries (one per property), to test the sub-requirements for each property specified. Since the queries for requirements 28, 29 and 30 require the use of these same properties to test the system's compliance to the GeoSPARQL RIF [24] rules, we use an approach where the explicit RDF triples needed to test requirements 4, 5 and 6 involve geometries which are the top result when using the ordering of the query results. For this purpose, the queries for requirements 4, 5 and 6 order the results, and select the top result only, to ensure they test the existence of the explicit and materialized RDF triple in the dataset.

Req. 7 Implementations shall allow the RDFS class [geo:Geometry](#) to be used in SPARQL graph patterns.

Req. 8 Implementations shall allow the properties [geo:hasGeometry](#) and [geo:hasDefaultGeometry](#) to be used in SPARQL graph patterns.

Req. 9 Implementations shall allow the properties [geo:dimension](#), [geo:coordinateDimension](#), [geo:spatialDimension](#), [geo:isEmpty](#), [geo:isSimple](#), [geo:hasSerialization](#) to be used in SPARQL graph patterns.

The tests for requirements 7, 8 and 9 are done by selecting all entities of type [geo:Geometry](#) (Req. 7), or by selecting the object/value of geometry A denoted by the property in question (Req. 8 and 9). Since requirement 8 specifies two distinct properties,

and requirement 9 specifies six such properties, the tests for these requirements consist of two and six queries, respectively.

Req. 10 All RDFS Literals of type `geo:wktLiteral` shall consist of an optional URI identifying the coordinate reference system followed by Simple Features Well Known Text (WKT) describing a geometric value. Valid `geo:wktLiteral` instances are formed by concatenating a valid, absolute URI as defined in [25], one or more spaces (Unicode U+0020 character) as a separator, and a WKT string as defined in Simple Features [5].

We test requirement 10 by selecting and checking the datatype of a correctly defined WKT literal from the dataset, to make sure the system under testing supports the specified format of WKT literals and their datatype.

Req. 11 URI `<http://www.opengis.net/def/crs/OGC/1.3/CRS84>` shall be assumed as the spatial reference system for `geo:wktLiterals` that do not specify an explicit spatial reference system URI.

We test requirement 11 by first defining two geometries in the dataset: J and K, which represent the same polygon, but geometry K has a WKT literal with an explicitly specified reference system, while geometry J does not contain the URI and only contains the polygon points in the literal value:

```
J: Polygon((-77.089005 38.913574,-77.029953 38.913574,
-77.029953 38.886321,-77.089005 38.886321,-77.089005 38.913574))
K: <http://www.opengis.net/def/crs/OGC/1.3/CRS84>
Polygon((-77.089005 38.913574,-77.029953 38.913574,
-77.029953 38.886321,-77.089005 38.886321,-77.089005 38.913574))
```

Then, we test whether these two geometries, i.e., their corresponding WKT literals, are geometrically equal. This ensures that a correct answer to this test means that the underlying system assumes `CRS84` as the default spatial reference system for WKT literals which do not specify one explicitly.

Req. 12 Coordinate tuples within `geo:wktLiterals` shall be interpreted using the axis order defined in the spatial reference system used.

In order to test requirement 12, we define two new geometries in the dataset: L and M, which represent the same point. Geometry L has a WKT literal which specifies the point using the `CRS84` coordinate system, while geometry M uses the `EPSG:4326` coordinate system [26]. Compared to one another, these coordinate systems use an inverted axis order:

```
L: <http://www.opengis.net/def/crs/OGC/1.3/CRS84> Point(-88.38 31.95)
M: <http://www.opengis.net/def/crs/OGC/1.3/CRS84> Point(31.95 -88.38)
```

In order to test whether the system interprets the axis order correctly, i.e., according to the spatial reference system, we test if the two geometries are equal based on the system under testing.

Req. 13 An empty RDFS Literal of type `geo:wktLiteral` shall be interpreted as an empty geometry.

We define two new geometries, H and I, for the purpose of testing requirement 13. Geometry H represents a `LineString` geometry which has a WKT literal, which is an empty string. Geometry I represents an explicitly defined empty `LineString` geometry:

```
H:
I: LineString EMPTY
```

Additionally, as most of the other geometries, these two geometries have a `Point` representation as well. In the case of geometry H, it is again represented by an empty value of the WKT literal, while geometry I has an explicitly defined empty `Point` geometry in its WKT literal:

H:
I: Point EMPTY

The test then consists of two parts, where both check if the WKT literals of H and I are equal. The two parts refer to the separate testing of the equality of the LineString geometries and the Point geometries. Both parts should be correct in order for requirement 13 to be fulfilled and thus fully scored by the benchmark.

Req. 14 Implementations shall allow the RDF property `geo:asWKT` to be used in SPARQL graph patterns.

We test requirement 14 by simply selecting the `geo:asWKT` value of geometry A and checking it against the expected literal value.

Req. 15 All `geo:gmlLiterals` shall consist of a valid element from the GML schema that implements a subtype of `GM_Object` as defined in [27].

For the purpose of testing requirement 15, we select all the values of the `geo:asGML` property, regardless of the RDF subject, and check whether all of them contain a valid `GM_Object` subtype in the value and whether its datatype is `geo:gmlLiteral`. The ordered list of results is then checked against the expected answers, which include all valid GML literals from the dataset.

Req. 16 An empty `geo:gmlLiteral` shall be interpreted as an empty geometry.

Similarly to requirement 13, we test compliance to requirement 16 by providing an empty string as a GML literal value in one geometry—geometry H, and an explicitly defined empty LineString in a GML literal—geometry I:

H:
I: <LineString><posList></posList></LineString>

Just like with requirement 13, here we use a Point representations as well. In the case of geometry H, it is again represented by an empty value of the GML literal, while geometry I has an explicitly defined empty Point geometry in its GML literal:

H:
I: <Point><pos></pos></Point>

The test for requirement 16 consists of two parts, as well, where both check if the GML literals of H and I are equal. The two parts refer to the separate testing of the equality of the LineString geometries and the Point geometries. Both parts should be correct in order for requirement 16 to be fulfilled.

Req. 17 Implementations shall document supported GML profiles.

Requirement 17 is the only non-technical requirement of the GeoSPARQL standard, and therefore cannot be automatically checked and tested. This is the only requirement omitted by the benchmark tests. To keep it simple, we assume that all GeoSPARQL implementations fulfill this requirement and provide proper documentation for supported GML profiles, which we believe to be a reasonable assumption.

Req. 18 Implementations shall allow the RDF property `geo:asGML` to be used in SPARQL graph patterns.

Similarly to requirement 14, we test requirement 18 by simply selecting the `geo:asGML` value of geometry A and checking it against the expected literal value.

Req. 19 Implementations shall support `geof:distance`, `geof:buffer`, `geof:convexHull`, `geof:intersection`, `geof:union`, `geof:difference`, `geof:symDifference`, `geof:envelope` and `geof:boundary` as SPARQL extension functions, consistent with the definitions of the corresponding functions (distance, buffer, convexHull, intersection, difference, symDifference, envelope and boundary respectively) in Simple Features [5].

In order to test requirement 19, we use separate tests for the nine functions in question, i.e., we check each function separately. To test the full compliance of each function, we run three sub-tests for them: (a) we test the function with geometry parameters which are expressed as WKT literals, (b) we test it with geometry parameters expressed as GML literals, and (c) we test it with a combination of WKT and GML literals. If the function uses a single parameter, we only use the (a) and (b) sub-tests. If it uses two parameters, we use the (a), (b) and (c) sub-tests, where (c) consists of two queries in which WKT is the first and GML is the second parameter of the function (denoted as WKT-GML), and vice versa (denoted as GML-WKT). With this, the test for each function consists of either two sub-tests (WKT and GML), or of four sub-tests (WKT-WKT, GML-GML, WKT-GML and GML-WKT). This ensures that the compliance score for each function is thoroughly checked. The scoring details for these tests are presented in Section 3.3.

With this, the entire test for requirement 19 consists of tests for the nine functions, each with two or four sub-tests, for a total of 28 SPARQL queries.

Req. 20 Implementations shall support `geof:getSRID` as a SPARQL extension function.

We test requirement 20 by using the `geof:getSRID` function in two queries: one with the WKT literal of geometry A, and the other with the GML literal of geometry A. In both cases, we check if the system correctly returns <http://www.opengis.net/def/crs/OGC/1.3/CRS84> as an answer.

Req. 21 Implementations shall support `geof:relate` as a SPARQL extension function, consistent with the relate operator defined in Simple Features [5].

For testing requirement 21, we use a relate operator which denotes the contains relation (expressed as T****FF* in DE-9IM [28]), and test it on geometries A and B, where A contains B in the dataset. Given that the `geof:relate` function uses two parameters, there are four queries for this test: WKT-WKT, GML-GML, WKT-GML and GML-WKT.

Req. 22 Implementations shall support `geof:sfEquals`, `geof:sfDisjoint`, `geof:sfIntersects`, `geof:sfTouches`, `geof:sfCrosses`, `geof:sfWithin`, `geof:sfContains`, `geof:sfOverlaps` as SPARQL extension functions, consistent with their corresponding DE-9IM intersection patterns [28], as defined by Simple Features [5].

Req. 23 Implementations shall support `geof:ehEquals`, `geof:ehDisjoint`, `geof:ehMeet`, `geof:ehOverlap`, `geof:ehCovers`, `geof:ehCoveredBy`, `geof:ehInside`, `geof:ehContains` as SPARQL extension functions, consistent with their corresponding DE-9IM intersection patterns, as defined by Simple Features [5].

Req. 24 Implementations shall support `geof:rcc8eq`, `geof:rcc8dc`, `geof:rcc8ec`, `geof:rcc8po`, `geof:rcc8tppi`, `geof:rcc8tpp`, `geof:rcc8ntpp`, `geof:rcc8ntppi` as SPARQL extension functions, consistent with their corresponding DE-9IM intersection patterns [28], as defined by Simple Features [5].

We test requirements 22, 23 and 24 by applying a separate set of tests for each of the twenty-four functions specified. Each function is tested by employing four queries: one with two WKT literals (WKT-WKT), one with two GML literals (GML-GML), and two with a combination of WKT and GML literals (WKT-GML and GML-WKT). Each of the queries tests if the relation implemented by the tested function is correct for the used geometries from the dataset, and each of them returns a `xsd:boolean` answer. The geometries used for the tests of each function are carefully selected in order to provide an unambiguous assessment of whether the function is supported and correctly implemented in the system under testing.

Req. 25 Basic graph pattern matching shall use the semantics defined by the RDFS Entailment Regime [29].

For the purpose of testing requirements 25, 26 and 27, we use queries which require the system to select both materialized RDF triples, as well as inferred RDF triples, based on the specifics of each requirement.

Therefore, we test requirement 25 using three separate queries: the first one selects all instances of the `geo:Feature` class, where we expect the system to select instances of the subclasses of the class, as well, e.g., `my:PlaceOfInterest`; the second and the third one select all instances with the `geo:hasGeometry` and `geo:hasDefaultGeometry` properties, but expect the results to contain entities which use subproperties of these properties, as well, e.g., `my:hasExactGeometry`.

Req. 26 Implementations shall support graph patterns involving terms from an RDF-S/OWL [30] class hierarchy of geometry types consistent with the one in the specified version of Simple Features [5].

For requirement 26, we use two separate queries: they select all instances of `sf:Surface` and `sf:Curve`, respectively, but expect the results to contain all instances of their subclasses as well, such as `sf:LineString` and `sf:Polygon`.

Req. 27 Implementations shall support graph patterns involving terms from an RDF-S/OWL class hierarchy of geometry types consistent with the GML schema that implements `GM_Object` using the specified version of GML [27].

To test requirement 27, we use a single query which selects all instances of `gml:Surface`, but the expected results include all instances of its subclass, `gml:LineString`.

Req. 28 Basic graph pattern matching shall use the semantics defined by the RIF Core Entailment Regime [W3C SPARQL Entailment] for the RIF rules [31] `geor:sfEquals`, `geor:sfDisjoint`, `geor:sfIntersects`, `geor:sfTouches`, `geor:sfCrosses`, `geor:sfWithin`, `geor:sfContains`, `geor:sfOverlaps`.

Req. 29 Basic graph pattern matching shall use the semantics defined by the RIF Core Entailment Regime [W3C SPARQL Entailment] for the RIF rules [31] `geor:ehEquals`, `geor:ehDisjoint`, `geor:ehMeet`, `geor:ehOverlap`, `geor:ehCovers`, `geor:ehCoveredBy`, `geor:ehInside`, `geor:ehContains`.

Req. 30 Basic graph pattern matching shall use the semantics defined by the RIF Core Entailment Regime [W3C SPARQL Entailment] for the RIF rules [31] `geor:rcc8eq`, `geor:rcc8dc`, `geor:rcc8ec`, `geor:rcc8po`, `geor:rcc8tppi`, `geor:rcc8tpp`, `geor:rcc8ntpp`, `geor:rcc8ntppi`.

We test the requirements 28, 29 and 30 with eight different queries each, in order to test the sub-requirements for each individual rule specified. The queries used here are similar to the queries for requirements 4, 5 and 6, with the difference that the tests for requirements 28, 29 and 30 require both materialized RDF triples and inferred RDF triples to be selected for the query response. To ensure that the system selects all such entities and therefore supports the semantics defined in the RIF core entailment regime for the RIF rules, the tests require an ordered list of entities fulfilling the query request.

3.3. Benchmark Results

The benchmark can test whether the benchmarked system provides a correct or an incorrect answer on each of the 206 benchmark queries. In order to transform these individual results into an overall result, we calculate two benchmark results from a given experiment:

- **Correct answers:** The number of correct answers out of all GeoSPARQL queries, i.e., tests.
- **GeoSPARQL compliance percentage:** The percentage of compliance with the requirements of the GeoSPARQL standard.

The former is straightforward—it is the number of correct answers the system provided, out of the 206 test queries. The latter is calculated from the perspective of the 30 requirements and measures the overall compliance of the benchmarked system with the GeoSPARQL standard. It measures the amount of supported requirements of the system, out of the 30 specified requirements, where the weight of each requirement is uniformly distributed, i.e., each requirement contributes 3.33% to the total result.

If a requirement contains multiple sub-test queries, its 3.33% are uniformly distributed among them. Therefore, for instance, each of the eight sub-requirements of requirement 4 contributes with 12.5% to the parent test score, i.e., with 0.4167% ($3.33\% \times 12.5\%$) to the total benchmark compliance percentage score. This means that a single requirement from the GeoSPARQL standard can be fully supported, partially supported or not supported at all.

The only exceptions to this rule of uniform distribution of the weights between tests on the same level are the sub-test queries which test GeoSPARQL functions with different serializations of literals as parameters, i.e., requirements 19–24. When we test a function for compliance to the standard while using (a) WKT-only literals, (b) GML-only literals and (c) a combination of WKT and GML literals, the score is uniformly distributed between these three logical groups, each contributing with 33.33% to the parent test score. However, (c) is practically tested using two queries: one where WKT is the first and GML is the second parameter of the function (denoted as WKT-GML), and vice versa (denoted as GML-WKT). These two queries technically contribute with 16.67% to the parent test score each, so that the total contribution from the logical group (c) remains 33.33%. With this, the technical weight of the queries themselves is 33.33% for the WKT-only query, 33.33% for the GML-only query, 16.67% for the WKT-GML query and 16.67% for the GML-WKT query. Technically, on a query level, this is an exception of the uniform distribution rule we practice, but, logically, on a group level, it still holds.

Given that requirement 17 is non-technical, and therefore not tested as part of the benchmark, each system gets its 3.33% score points automatically, when it provides at least one correct answer to the benchmark tests.

3.4. Benchmark Considerations

When creating the benchmark, we needed to take certain considerations and interpretations which were implicitly given in the GeoSPARQL standard. We elaborate on these in this subsection.

3.4.1. Geometry Literals

Many results of query functions defined in the GeoSPARQL standard return a `ogc:geomLiteral` as a result, following the GeoSPARQL standard definition. This means that, according to the standard, a function such as:

```
geof:boundary(ogc:geomLiteral):ogc:geomLiteral
```

may take either a WKT, a GML 2.0, or a GML 3.2 literal as an argument, and may return either a WKT, a GML 2.0, or a GML 3.2 literal as a result. The dataset we use for our benchmark includes WKT and GML 3.2 formatted literals. However, we provide query answers in WKT, GML 2.0 and GML 3.2 to support all possible outcomes from a system tested by the benchmark.

The decision to include only GML 3.2 and not GML 2.0 literals in our dataset was taken because GML 2.0 has been de-facto superseded by GML 3.2. GML 2.0 is not even supported as an export option in current GIS software, such as QGIS, for instance. In addition, in all systems, we benchmarked that the only GML variant that was supported was GML 3.2.

3.4.2. Variations between Literal Serializations

Within the same literal type, different semantically equivalent representations of geometries are possible. WKT serializations may include a CRS URI, but they may also omit it (if it is missing, WGS84 CRS is assumed), and they may differ in the amount and positioning of whitespaces. GML literals may differ in the order of attributes and definition of namespaces. To be flexible about these variations, we apply a normalization process before comparing the results from the tested system with the expected answer. WKT literals are trimmed and their whitespaces are removed, and GML literals are converted to canonicalized XML with normalized namespace definitions.

3.4.3. Alternative Answers

The GeoSPARQL standard defines the results of GeoSPARQL functions as `ogc:geomLiteral` values but does not define which geometry types these literals should serialize. Therefore, functions may not only return results in different literal types, but also in different geometry representations even within the same literal serialization. One example is the `geof:boundary` function which could return a `sf:LinearRing` or a `sf:Polygon` geometry as a result. Even supposedly simple return values such as an `xsd:boolean` may be represented as either the `xsd:boolean` literals with value *true* and *false* or *1* and *0*.

In order to deal with these scenarios, we define alternative query answers for each of the aforementioned possibilities. This means that each test consists of a single query which is issued to the system under testing, and a set of several alternative correct answers, which are logically equivalent, but may be technically represented in different serializations.

3.5. Implementation

We have implemented the benchmark as a benchmark for the HOBBIT platform (Public instance of the HOBBIT Platform: <http://master.project-hobbit.eu> (accessed on 22 May 2021)), intended for holistic benchmarking of big Linked Data [17]. The HOBBIT platform allows for users to define and execute benchmarks, on one hand, and provide and add triplestore systems, on the other. A user can run an experiment on the platform by selecting the desired benchmark and the target triplestore system to be tested. The platform then loads the benchmark as a set of Docker containers (benchmark controller, data generator, task generator and evaluation module), loads the system as a Docker container (benchmarked system), and then runs the benchmark according to its logic, programmed in the controller (Figure 2). The results of each experiment are stored in the platform and are made publicly available on the Web.

In our case, the GeoSPARQL compliance benchmark first loads the dataset into the benchmarked system, then reads all the test queries and sends them to the benchmarked system for execution. The evaluation module reads the single expected answer or the set of expected alternative answers for each query, and compares whether the benchmarked system returns a correct or an incorrect answer, saving the result into the evaluation store. After all tests are done, the evaluation module calculates two summarized results: (1) the number of correct answers, out of all possible tests, and (2) the percentage of compliance to the requirements of the GeoSPARQL standard, as described in Section 3.3.

We decided to use the HOBBIT platform for our benchmark due to its plug-in nature, in which additional systems can be added by interested users, which will then be able to run an experiment with the benchmark over their own system. A user can also run our GeoSPARQL compliance benchmark over any triplestore system which is already available on the platform. Additionally, the public nature of the platform allows for greater transparency and reproducibility of the results of each benchmark, including our GeoSPARQL compliance benchmark.

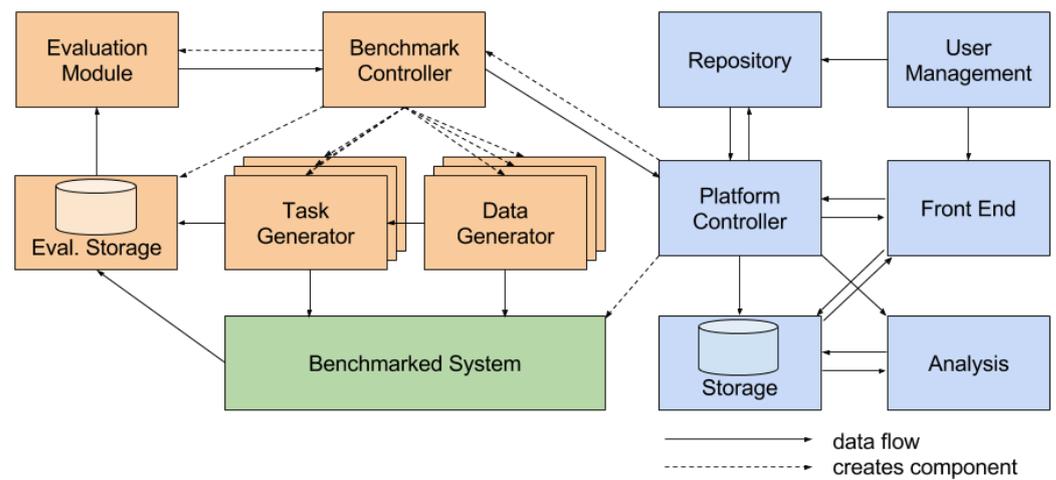


Figure 2. The HOBBIT benchmarking platform.

4. Experimental Setup

In order to showcase the usability and usefulness of the GeoSPARQL compliance benchmark, we set out to run a number of experiments over some of the most commonly used triplestores. The set of chosen triplestores is shown in Table 1.

Table 1. Triplestores which have been tested using the GeoSPARQL compliance benchmark.

Triplestore	Version	Reference
Apache Marmotta	3.4.0	[32]
Blazegraph	3.1.5	[33]
Eclipse RDF4J	3.4.0	[34]
GeoSPARQL Fuseki	3.17.0	[9,35]
Jena Fuseki	3.14.0	[36]
Ontotext GraphDB	9.3.3	[37]
OpenLink Virtuoso	7.2	[38,39]
Stardog	7.4.0	[40]
TriplyDB	3.5	[41]

For each experiment, a system adapter has been created and published on a public HOBBIT platform instance, as well as in the HOBBIT GitLab repository (HOBBIT Platform Triplestores: <https://git.project-hobbit.eu/triplestores> (accessed on 22 May 2021)). This allows for the reproduction of the experiments and the results. Each triplestore version from Table 1 was the most recent available stable version of the implementation at the time of testing. If a triplestore requires a license file (e.g., Stardog, TriplyDB), its tests are reproducible on the HOBBIT platform only until the embedded license of the integrated system is valid. When the license expires, any interested party needs to submit their own instance of the system to the platform in order to test it. For each of the triplestores which have been tested, a system adapter implementation has been created which handles the initial configuration of the triplestore, e.g., setting up a repository which contains the data to be tested, enabling geospatial query support, etc. If possible, this adapter implementation was added to the triplestore implementation in a joint Docker image or two Docker images—the adapter implementation and the triplestore implementation—were created for testing. It needs to be stated that not all of the aforementioned triplestores claim to support GeoSPARQL. In fact, Blazegraph and Jena Fuseki do not support GeoSPARQL. We included them in our experiments in order to show which GeoSPARQL requirements are already supported by a non-GeoSPARQL implementation of an RDF triplestore which at least supports the SPARQL query language.

5. Results and Discussion

5.1. Overall Results

The results of the experiments with our benchmark and the systems listed in Table 1 are shown in Table 2 and in Figure 3, and are available online on the HOBBIT platform (Results on the HOBBIT platform: <https://master.project-hobbit.eu/experiments/1612476122572,1612477003063,1612476116049,1625421291667,1612477500164,1612661614510,1612637531673,1612828110551,1612477849872> (accessed on 8 July 2021)). They show that none of these widely used RDF storage solutions fully comply with the GeoSPARQL standard. Aside from that, we can point out that one of them stands out with a significantly better GeoSPARQL compliance score than the others, and, more generally, the top four stand out from the rest. The triplestores in positions 5–8 share an almost identical result.

Table 2. Results from the GeoSPARQL compliance benchmark.

Triplestore	Correct Answers (out of 206)	GeoSPARQL Compliance
GeoSPARQL Fuseki 3.17	177	82.75%
Ontotext GraphDB 9.3.3	80	69.75%
OpenLink Virtuoso 7.2	73	63.46%
TriplyDB 3.5	73	63.46%
Eclipse RDF4J 3.4.0	47	58.33%
Stardog 7.4.0	46	56.67%
Blazegraph 2.1.5	46	56.67%
Jena Fuseki 3.14	46	56.67%
Apache Marmotta 3.4.0	40	46.67%

Experiment									
Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark	GeoSPARQL Compliance Benchmark
Challenge Task									
System	GeoSPARQL Jena Fuseki 3.17.0	GraphDB 9.3.3	Virtuoso v7.2 Open-Source Edition	TriplyDB 3.5	RDF4J 3.4.0	Stardog 7.4.0	Blazegraph 2.1.5	Jena Fuseki 3.14.0	Apache Marmotta 3.4.0
URI	http://w3id.org								
KPIs									
Correct Answers	177	80	73	73	47	46	46	46	40
GeoSPARQL Compliance Percentage	82.74691009	69.75308990	63.45679092	63.45679092	58.33333206	56.66666793	56.66666793	56.66666793	46.66666793

Figure 3. Results from the GeoSPARQL compliance benchmark, from the public instance of the HOBBIT platform.

In order to see the reasons for these variations more closely, we made a breakdown of the compliance results into the six extensions defined in the GeoSPARQL standard. These results are shown in Table 3. As we can see from this table, the triplestores in positions 5–8 share the same result due to demonstrating full compliance with the CORE, TOP and RDFSE extensions of the GeoSPARQL benchmark, but not with the other extensions. The reason why almost all benchmarked triplestores comply with CORE, TOP and RDFSE is simple: these requirements are designed in such a way that they are satisfied “out-of-the-box” by most RDF- and SPARQL-compliant storage solutions. They refer to the use of specific classes (CORE) and properties (TOP) in SPARQL query patterns, as well as RDFS reasoning (RDFSE), which are features supported in most triplestores nowadays. Since RDFS reasoning was not activated in the Marmotta version we benchmarked, it has no

compliance for RDFSE so its score comes only from its compliance with CORE and TOP, thus is lower than the scores of the other systems.

The bottom three systems are explicitly not GeoSPARQL-compliant, but we included them in our experiments as baseline tests. As we can see, they all demonstrated compatibility with either two or with three extensions of the GeoSPARQL standard (Table 3), and scored 56.67% or 46.67% of the GeoSPARQL compliance score (Table 2). This, however, does not mean that the benchmark score should start at 56.67% or 46.67%, since a benchmarked RDF storage system may fail these tests too.

Table 3. Support of the different GeoSPARQL extension by the tested triplestores. *Full* indicates full support, comprised of correct query answers only, *Full/E* indicates that support is implemented but erroneous, *Partial [GML/WKT]* indicates that support is partially implemented, *None* indicates that support for this GeoSPARQL extension is not present.

Triplestore	CORE	TOP	GEOEXT	GTOP	RDFSE	QRW
GeoSPARQL Fuseki	Full	Full	Full/E	Full	Full	Full/E
Ontotext GraphDB	Full	Full	Partial [WKT]	Partial [WKT]	Full	None
OpenLink Virtuoso	Full	Full	Partial [WKT]	Partial [WKT]	Full	None
TriplyDB	Full	Full	Partial [WKT]	Partial [WKT]	Full	None
Eclipse RDF4J	Full	Full	Partial [WKT CRS84]	Partial [WKT CRS84]	Full	None
Stardog	Full	Full	None	None	Full	None
Blazegraph	Full	Full	None	None	Full	None
Jena Fuseki	Full	Full	None	None	Full	None
Apache Marmotta	Full	Full	None	None	None	None

5.2. Discussion on the Results for Each Triplestore

First, we tested RDF triplestores which claim GeoSPARQL support. We wanted to check how extensive their compliance with the GeoSPARQL standard is, and this list included: GeoSPARQL Fuseki, GraphDB, Virtuoso, TriplyDB, RDF4J and Stardog.

GeoSPARQL Fuseki is the triplestore with the highest GeoSPARQL compliance score in our experiments. It is the only system with full GML and WKT support and the only system with a full implementation of all GeoSPARQL extensions (Table 3). However, GeoSPARQL Fuseki produced incorrect results in many functions covered by the query rewrite extension and in a few functions covered by the geometry extension. In addition, just like all other triplestores we tested, GeoSPARQL Fuseki fails to handle empty WKT and empty GML literals.

GraphDB provides a full implementation of all but the query rewrite extension. However, GraphDB can only handle WKT literals but not GML literals. This leads to a substantially lower score in our benchmark, as many queries require either a GML literal as input, or a combination of a GML and a WKT literal in order to be executed. Most functions with WKT-only literals in the GEOEXT and GTOP extension tests produced correct results.

Virtuoso provides support for WKT literals, but not GML literals. Similarly to GraphDB, it provides full implementation for all GeoSPARQL extensions, except for the query rewrite extension. However, it has an additional issue: even though it returns logically correct results for the tests for the functions in requirement 19 (part of the GEOEXT extension), the literals are transformed from WKT literals to an internal literal type which is Virtuoso-specific. This renders a mismatch between the provided and expected answer, and lowers the benchmark score for Virtuoso.

TriplyDB is a Linked Data solution which uses Virtuoso Open-Source and Jena Fuseki for the storage of RDF data on the back-end. We used a Virtuoso-based version in our experiments. Given that TriplyDB preprocesses the data during ingestion, preprocesses the SPARQL queries before execution, and postprocesses the SPARQL results after execution, it might provide different results and demonstrate different behavior compared to using Virtuoso and Fuseki on their own. However, in the case of our benchmark, TriplyDB 3.5 scored exactly the same as Virtuoso Open-Source 7.2, meaning that the geospatial support of TriplyDB is identical to Virtuoso.

The RDF4J triplestore implements all the GeoSPARQL functions of the GEOEXT extension and the GTOP extension for WKT literals. However, RDF4J fails almost all of the GeoSPARQL tests from these extensions because it does not support CRS URIs in WKT literals. While the GeoSPARQL standard acknowledges that the integration of CRS URIs in WKT Literals is optional, they are used in various use-cases, especially at geospatial authorities, and we expect them to be supported in every triplestore which claims GeoSPARQL support. Thus, WKT literals with explicit CRS URIs are included in most of the tests of the benchmark. In addition, RDF4J lacks support for GML literals and the query rewrite extension.

The Stardog triplestore provides an implementation covering WKT literals and implements five geospatial functions which are similar to the GeoSPARQL functions, but not fully compatible. More specifically, out of their five geospatial functions (`geof:within`, `geof:area`, `geof:nearby`, `geof:distance` and `geof:relate`), only the `geof:distance` function follows the signature of the GeoSPARQL function with the same URI. However, our tests for this function include WKT literals with explicit CRS URIs, which Stardog does not support, so the test for this function fails. The tests for the other functions fail either because functions with those URIs do not exist in the GeoSPARQL standard, or because of a function signature mismatch. Thus, Stardog only scores in tests which cover the CORE, TOP and RDFSE extensions.

Next, we tested triplestores which do not claim to support GeoSPARQL, but claim support for other geospatial extensions. We expected that they will provide full support for the GeoSPARQL CORE, TOP and RDFSE extensions which do not rely on the implementation of additional geospatial operators. They thereby constitute as baseline tests for our approach, and this list included: Blazegraph, Jena Fuseki, Apache Marmotta and Parliament.

Blazegraph supports some non-GeoSPARQL spatial functions in its GeoSpatial Search Extension (<https://github.com/blazegraph/database/wiki/GeoSpatial> (accessed on 22 May 2021)). This extension allows the definition of `Points` via WKT literals but is otherwise not GeoSPARQL-compliant. Blazegraph therefore fails the GEOEXT, GTOP and QRW tests, as expected.

Jena Fuseki includes a customized spatial extension *Jena Spatial* (<https://jena.apache.org/documentation/query/spatial-query.html> (accessed on 22 May 2021)) which is planned to be replaced by the GeoSPARQL Fuseki implementation we tested. Jena Fuseki can cope with WKT literals and defines a custom set of functions, none of which match the function signatures defined in the GeoSPARQL standard. Hence, Jena Fuseki only gets awarded a full score in the CORE, GTOP and RDFSE extensions.

Apache Marmotta has a GeoSPARQL implementation which was created in a Google Summer of Code project (<http://marmotta.apache.org/kiwi/geosparql.html> (accessed on 22 May 2021)). At the time of testing, the extension was not included in the last stable version of this triplestore; therefore, the version we tested was not GeoSPARQL-compliant. Even though Marmotta supports RDFS reasoning, we were unsuccessful in our attempts to activate it for the instance we worked with, so even though we expected it to achieve the same score as the other triplestores which do not support GeoSPARQL, it only scored as compliant with CORE and TOP.

Finally, we want to acknowledge that we also tested the Parliament 2.7.10 triplestore. Parliament validates WKT and GML literals before they are added to the graph, and fails to load a dataset if a validation error occurs. In our test, Parliament failed to parse GML 3.2 literals and the empty WKT literals. As a result, the benchmark dataset could not be loaded, and we could not conduct the experiment with the Parliament triplestore.

6. Limitations of the Benchmark

The GeoSPARQL compliance benchmark does not test every GeoSPARQL function with every available geometry type and their combinations. We do that with WKT and GML serializations but not different geometry types. The reason for this is that the amount

of possible combinations of geometries would be inconceivably too large and the benefit of testing them far too low. WKT defines 27 geometry types, GML defines at least as many which would need to be considered both in their GML 2.0 and in their GML 3.2 variants, to be complete. Instead, our dataset consists of *Points*, *LineStrings* and *Polygons*, which are the most widely used geometry types. With this, we believe we strike a good balance between the benchmark being too extensive and being sufficiently precise in measuring a system's compliance with the GeoSPARQL standard.

Regarding the GeoSPARQL compliance percentage score: as we already stated, this score measures the number of supported requirements of the system, out of the 30 specified requirements, where the weight of each requirement is uniformly distributed, i.e., each requirement contributes 3.33% to the total result. The reason we decided to use uniform distribution instead of assigning requirement-specific weights is because adding weights to different requirements would be somewhat arbitrary. Given that the authors of the GeoSPARQL standard have not discussed or put any variable significance between the different requirements, gives us a signal that, at least for the time being, we should treat them as equally important. While that practically is not the case, and different stakeholders may have different significance implicitly assigned to them, we do not think there is a better universal way to address this.

7. Conclusions

This paper introduces a GeoSPARQL compliance benchmark which aims to measure the extent to which an RDF triplestore complies with the requirements specified in the GeoSPARQL standard. By doing a series of tests for each requirement, the benchmark is able to assess whether the benchmarked system fully or partially supports a given requirement, or not at all. The results from the 206 individual tests are transformed into a GeoSPARQL compliance percentage which aims to provide a metric of the amount of requirements covered by the benchmarked system.

In order to showcase the usefulness and usability of the benchmark, as part of the HOBBIT platform, we ran a series of experiments with eight of the most commonly used RDF triplestores. The overall results show that GeoSPARQL support varies greatly between the tested triplestores. While the CORE, TOP and RDFSE extensions are supported in almost every triplestore—as they only depend on SPARQL and RDFS functionalities and are not GeoSPARQL-specific—the GEOEXT and GTOP extensions show varying levels of implementation. Some triplestores, such as GraphDB or Virtuoso, chose to only implement support for WKT literals, RDF4J supports only WKT literals without CRS URIs and only GeoSPARQL-Jena provides a full GeoSPARQL-compliant implementations of all functions with both GML and WKT compatibility. GeoSPARQL-Jena is also the only implementation tested in our benchmark which implements the QRW extension of GeoSPARQL.

In conclusion, we can see that the GeoSPARQL standard, almost nine years after its initial release, is often only partially supported by major triplestore vendors. We hope that the contribution of our GeoSPARQL benchmark can help to motivate implementers to improve their RDF storage solutions, give customers a guideline as to which implementation is most suitable for their given use-case, and provide a starting point for a further standard-conform expansion of the geospatial Semantic Web.

Future Work

Recently, the OGC GeoSPARQL Working Group has been reactivated [42,43] to define GeoSPARQL 2.0, a successor to the GeoSPARQL standard. It is a good practice of emerging OGC standards to first be defined, then reviewed, and at the same time also implemented as a proof-of-concept. During the course of this implementation, compliance testing becomes increasingly common as can be seen by the establishment of the OGC Team Engine (<https://cite.opengeospatial.org/teamengine/> (accessed on 22 May 2021)), a compliance test suite which enterprises may use to get official OGC compliance certifications for their software implementations. Given that currently no OGC-endorsed OGC GeoSPARQL

compliance test exists, we would welcome a collaboration with the OGC and would like to extend our test suite to cover the changes which will be defined in GeoSPARQL 2.0.

Given that our benchmark is a compliance benchmark, we plan to develop a complementary performance benchmark, which would test the performance of the tested RDF triplestore for each GeoSPARQL functions it supports. This would enable a more holistic approach in the evaluation of geospatial RDF storage solutions. Despite the emergence of many performance benchmarks for geospatial RDF triplestores (outlined in Section 2), none of the existing benchmarks tests for every function defined in GeoSPARQL on a given test dataset. Our performance benchmark would be able to utilize the results from the compliance benchmark, and target the supported GeoSPARQL functions. The HOBBIT benchmarking platform provides an excellent environment for performance benchmarks, given that they all share the same infrastructure for the experiments and all results are reproducible.

Author Contributions: Conceptualization, Milos Jovanovik; methodology, Milos Jovanovik, Timo Homburg and Mirko Spasić; software, Milos Jovanovik, Timo Homburg and Mirko Spasić; validation, Milos Jovanovik, Timo Homburg and Mirko Spasić; formal analysis, Milos Jovanovik, Timo Homburg and Mirko Spasić; investigation, Milos Jovanovik, Timo Homburg and Mirko Spasić; resources, Milos Jovanovik, Timo Homburg and Mirko Spasić; data curation, Milos Jovanovik, Timo Homburg and Mirko Spasić; writing—original draft preparation, Milos Jovanovik and Timo Homburg; writing—review and editing, Milos Jovanovik, Timo Homburg and Mirko Spasić; visualization, Milos Jovanovik and Timo Homburg; supervision, Milos Jovanovik; project administration, Milos Jovanovik. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially supported by Eurostars Project SAGE (GA no. E!10882).

Data Availability Statement: The code of the benchmark is publicly available on GitHub, at <https://github.com/OpenLinkSoftware/GeoSPARQLBenchmark> (accessed on 8 July 2021). The results from the executed experiments are available on the public instance of the HOBBIT platform, at <https://master.project-hobbit.eu> (accessed on 8 July 2021). More specifically, the results from Figure 3 are available at <https://master.project-hobbit.eu/experiments/1612476122572,1612477003063,1612476116049,1625421291667,1612477500164,1612661614510,1612637531673,1612828110551,1612477849872> (accessed on 8 July 2021), where each experiment is linked and can be viewed separately. The detailed logs from each experiment are also publicly available for download from the same web location. The HOBBIT platform provides reproducibility of our results, by allowing users to run their own experiments with the GeoSPARQL compliance benchmark and any system(s) they are interested in benchmarking.

Conflicts of Interest: Milos Jovanovik and Mirko Spasić work for OpenLink Software, which is the vendor of Virtuoso, one of the benchmarked triplestores. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

CORE	Core Component
CRS	Coordinate Reference System
GEOEXT	Geometry Extension
GML	Geography Markup Language
GTOP	Geometry Topology Extension
OGC	Open Geospatial Consortium
QRW	Query Rewrite Extension
RDF	Resource Description Framework
RDFS	RDF Schema
RDFSE	RDFS Entailment Extension
SPARQL	SPARQL Protocol and RDF Query Language
TOP	Topology Vocabulary Extension
WKT	Well-Known Text

References

1. Fonseca, F. The Geospatial Semantic Web. In *The Handbook of Geographic Information Science*; Blackwell Publishing: Malden, MA, USA, 2008; pp. 367–376.
2. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web. *Sci. Am.* **2001**, *284*, 34–43. [CrossRef]
3. Battle, R.; Kolas, D. GeoSPARQL: Enabling a GeoSpatial Semantic Web. *Semant. Web J.* **2011**, *3*, 355–370. [CrossRef]
4. Perry, M.; Herring, J. OGC GeoSPARQL—A Geographic Query Language for RDF Data. OGC Standard, Open Geospatial Consortium, Wayland, MA, USA, 2012. Available online: <https://www.ogc.org/standards/geosparql> (accessed on 22 May 2021).
5. Herring, J. OpenGIS® Implementation Standard for Geographic Information—Simple Feature Access—Part 1: Common Architecture. OpenGIS Implementation Standard, Open Geospatial Consortium, Wayland, MA, USA, 2011. Available online: <https://www.ogc.org/standards/sfa> (accessed on 22 May 2021).
6. Portele, C. OGC Geography Markup Language (GML)—Extended Schemas and Encoding Rules. OpenGIS Implementation Standard, Open Geospatial Consortium, Wayland, MA, USA, 2012. Available online: <https://www.ogc.org/standards/gml> (accessed on 22 May 2021).
7. Prud'hommeaux, E.; Seaborne, A. SPARQL Query Language for RDF. W3C Recommendation, W3C, 2008. Available online: <https://www.w3.org/TR/2008/REC-rdf-sparql-query-20080115/> (accessed on 22 May 2021).
8. Battle, R.; Kolas, D. Enabling the Geospatial Semantic Web with Parliament and GeoSPARQL. *Semant. Web* **2012**, *3*, 355–370. [CrossRef]
9. Albiston, G.L.; Osman, T.; Chen, H. GeoSPARQL-Jena: Implementation and Benchmarking of a GeoSPARQL Graphstore. *Semant. Web J.* **2019**, under review.
10. Janssen, V. Understanding Coordinate Reference Systems, Datums and Transformations. *Int. J. Geoinformatics* **2009**, *5*, 41–53.
11. Decker, B.L. *World Geodetic System 1984*; Technical Report; Defense Mapping Agency Aerospace Center: St Louis, MO, USA, 1986.
12. Garbis, G.; Kyzirakos, K.; Koubarakis, M. Geographica: A Benchmark for GeoSpatial RDF Stores (long version). In Proceedings of the International Semantic Web Conference, Sydney, NSW, Australia, 21–25 October 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 343–359.
13. Ioannidis, T.; Garbis, G.; Kyzirakos, K.; Bereta, K.; Koubarakis, M. Evaluating Geospatial RDF stores Using the Benchmark Geographica 2. *arXiv* **2019**, arXiv:1906.01933.
14. Huang, W.; Raza, S.A.; Mirzov, O.; Harrie, L. Assessment and Benchmarking of Spatially Enabled RDF Stores for the Next Generation of Spatial Data Infrastructure. *ISPRS Int. J. Geo-Inf.* **2019**, *8*, 310. [CrossRef]
15. Rafes, K.; Nauroy, J.; Germain, C. TFT, Tests For Triplestores. In Proceedings of the Semantic Web Challenge, Part of the International Semantic Web Conference, Riva del Garda, Italy, 19–23 October 2014.
16. Ngomo, A.C.N.; Garcia Rojas, A.; Fundulaki, I. HOBBIT: Holistic Benchmarking for Big Linked Data. *ERCIM News* **2016**.
17. Röder, M.; Kuchelev, D.; Ngonga Ngomo, A.C. HOBBIT: A Platform for Benchmarking Big Linked Data. *Data Sci.* **2020**, *3*, 15–35. [CrossRef]
18. Jovanovik, M.; Homburg, T.; Spasić, M. GeoSPARQL Compliance Benchmark. Available online: <https://github.com/OpenLinkSoftware/GeoSPARQLBenchmark> (accessed on 8 July 2021).
19. Jovanovik, M.; Homburg, T.; Spasić, M. Software for the GeoSPARQL Compliance Benchmark. *Softw. Impacts* **2021**, *8*, 100071. [CrossRef]
20. Butler, H.; Daly, M.; Doyle, A.; Gillies, S.; Hagen, S.; Schaub, T. *The GeoJSON Format*; Technical Report 7946; IETF: Fremont, CA, USA, 2016.
21. Clark, K.; Feigenbaum, L.; Torres, E. SPARQL Protocol for RDF. W3C Recommendation, W3C, 2008. Available online: <https://www.w3.org/TR/2008/REC-rdf-sparql-protocol-20080115/> (accessed on 22 May 2021).
22. Beckett, D.; Broekstra, J. SPARQL Query Results XML Format. W3C Recommendation, W3C, 2008. Available online: <https://www.w3.org/TR/2008/REC-rdf-sparql-XMLres-20080115/> (accessed on 22 May 2021).
23. Brickley, D.; Guha, R. RDF Schema 1.1. W3C Recommendation, W3C, 2014. Available online: <https://www.w3.org/TR/2014/REC-rdf-schema-20140225/> (accessed on 22 May 2021).
24. Kifer, M.; Boley, H. RIF Overview (Second Edition). W3C Note, W3C, 2013. Available online: <https://www.w3.org/TR/2013/NOTE-rif-overview-20130205/> (accessed on 22 May 2021).
25. Berners-Lee, T.; Masinter, L.M.; Fielding, R.T. *Uniform Resource Identifiers (URI): Generic Syntax*; Technical Report 2396; IETF: Fremont, CA, USA, 1998.
26. Nicolai, R.; Simensen, G. The New EPSG Geodetic Parameter Registry. In Proceedings of the 70th EAGE Conference and Exhibition Incorporating SPE EUROPEC 2008, Rome, Italy, 9–12 January 2008; European Association of Geoscientists & Engineers: Houten, The Netherlands, 2008. [CrossRef]
27. Portele, C. OpenGIS® Geography Markup Language (GML) Encoding Standard. OpenGIS Standard, Open Geospatial Consortium, Wayland, MA, USA, 2007. Available online: <https://www.ogc.org/standards/gml> (accessed on 22 May 2021).
28. Strobl, C. Dimensionally Extended Nine-Intersection Model (DE-9IM). In *Encyclopedia of GIS*; Shekhar, S., Xiong, H., Zhou, X., Eds.; Springer: Cham, Switzerland, 2017; pp. 470–476. [CrossRef]
29. Glimm, B.; Ogbuji, C. SPARQL 1.1 Entailment Regimes. W3C Recommendation, 2013. Available online: <https://www.w3.org/TR/2013/REC-sparql11-entailment-20130321/> (accessed on 22 May 2021).

30. McGuinness, D.; van Harmelen, F. OWL Web Ontology Language Overview. W3C Recommendation, W3C. 2004. Available online: <https://www.w3.org/TR/2004/REC-owl-features-20040210/> (accessed on 22 May 2021).
31. Boley, H.; Hallmark, G.; Kifer, M.; Paschke, A.; Polleres, A.; Reynolds, D. RIF Core Dialect. W3C Recommendation, W3C. 2010. Available online: <https://www.w3.org/TR/2010/REC-rif-core-20100622/> (accessed on 22 May 2021).
32. Apache Marmotta. Available online: <http://marmotta.apache.org> (accessed on 22 May 2021).
33. Blazegraph. Available online: <https://blazegraph.com> (accessed on 22 May 2021).
34. Eclipse RDF4J. Available online: <https://rdf4j.org> (accessed on 22 May 2021).
35. GeoSPARQL Fuseki. Available online: <https://jena.apache.org/documentation/geosparql/geosparql-fuseki> (accessed on 22 May 2021).
36. Jena Fuseki. Available online: <https://jena.apache.org/documentation/fuseki2/> (accessed on 22 May 2021).
37. GraphDB. Available online: <https://graphdb.ontotext.com> (accessed on 22 May 2021).
38. Erling, O. Virtuoso, a Hybrid RDBMS/Graph Column Store. *IEEE Data Eng. Bull.* **2012**, *35*, 3–8.
39. Virtuoso. Available online: <https://virtuoso.openlinksw.com> (accessed on 22 May 2021).
40. Stardog. Available online: <https://www.stardog.com> (accessed on 22 May 2021).
41. TriplyDB. Available online: <https://triplydb.com> (accessed on 8 July 2021).
42. Abhayaratna, J.; van den Brink, L.; Car, N.; Atkinson, R.; Homburg, T.; Knibbe, F.; McGlenn, K.; Wagner, A.; Bonduel, M.; Holten Rasmussen, M.; et al. OGC Benefits of Representing Spatial Data Using Semantic and Graph Technologies. OGC White Paper, Open Geospatial Consortium, Wayland, MA, USA. 2020. Available online: <http://docs.ogc.org/wp/19-078r1/19-078r1.html> (accessed on 22 May 2021).
43. Abhayaratna, J.; van den Brink, L.; Car, N.; Homburg, T.; Knibbe, F. OGC GeoSPARQL 2.0 SWG Charter. Available online: https://github.com/opengeospatial/geosemantics-dwg/tree/master/geosparql_2.0_swg_charter (accessed on 22 May 2021).