

## Parámetros del *benchmarking*.

**Objetivo:** Plantear las métricas a usar y la forma de cómo se llevará a cabo el *benchmarking*.

**Resultados esperados:** Documento con una lista de parámetros que serán usados en el *benchmarking*.

### Definiciones

*Benchmarking*: Proceso metódico y continuo para evaluar servicios, productos y/o procesos de trabajo de organizaciones o empresas las cuales son bien reconocidas por llevar a cabo las mejores prácticas con el objetivo que mejorar o llevar a cabo estudios de referencia [1].

Plan de ejecución: Secuencia de operaciones en sistemas de administración de bases de datos para llevar a cabo las sentencias que se requieran [2].

### Desarrollo

En el presente documento se explican las bases del *benchmarking* que se realizará una vez completado el módulo de consultas federadas para el software Apache Marmotta.

En [3], Michael Schmidt y sus colegas proponen un *benchmarking* que mida las capacidades de los gestores de consultas federadas. Fue elaborado basándose en otros *benchmarking* populares y citados, los cuales son [4] y [5], cuyo objetivo era evaluar consultas centralizadas. Este documento será la base para construir el *benchmarking* del actual proyecto.

Tal como lo establece Michael Schmidt y su equipo, no existe una evaluación que cubra todos los aspectos de un sistema de consultas. Sin embargo, ellos abordan la heterogeneidad a nivel de datos el cual, es uno de los retos que los sistemas de consultas tienen que abordar.

La heterogeneidad a nivel de datos se identifica en los siguientes aspectos

- Distribución física: Cómo se comportan los sistemas de consultas frente a *datasets* cuya ubicación física son distintas.
- Interfaz de acceso a datos: El acceso a los datos pueden ser repositorios nativos, *SPARQL endpoints* y datos en la nube del *Linked Data*.
- Existencia de fuentes de datos: De manera local, si se tiene el conocimiento sobre los campos que hay en el repositorio por lo que evaluar este campo es posible. Sin embargo, este aspecto no es posible llevarse a cabo en la nube del *Linked Data* ya que a priori no se puede averiguar si la nube de datos contiene o no las fuentes de datos solicitadas en la consulta.
- Estadísticas de datos: De manera local, se puede conocer información de los repositorios mediante sus histogramas sobre propiedades, cantidad de datos y su

distribución de datos. Sin embargo, cuando los repositorios no son locales, puede haber poca o nula información acerca de ellos.

Aunado a lo anterior, también existen retos a nivel de consultas. Michael y sus colegas identificaron los siguientes aspectos a cubrir en la evaluación de consultas.

- Lenguaje de consulta: La expresividad entre lenguajes de consulta puede variar. Unos son simples, mientras que otros basan su sistemas en lenguajes de consultas muy específicos.
- Completitud de datos: Existen sistemas que pueden lidiar con una gran cantidad de información, pero también hay otros que no. Cabe decir que en la nube del *Linked Data* no es posible determinar la completitud de datos ya que un *triple store* puede desembocar en otros *triple store*, dando como resultado una interminable búsqueda de datos o quizá, un *triple store* no está referenciado correctamente a otros *triple store* resultando en resultados truncados.
- Lista de clasificación: Los sistemas pueden tener establecido una lista de mejores consultas basadas en sus propias métricas.

*FedBench* está basado en 3 componentes los cuales, pueden ser modificados, extendidos y personalizado para cubrir un escenario en específico. Los componentes son:

- Múltiples *datasets*: Elaboración de *datasets* y estadísticas. Para este proyecto, los *datasets* serán sobre datos geoespaciales.
- Múltiples conjuntos de consultas: Elaboración de las consultas.
- *Framework* de evaluación.

El artículo continua con una profundización de los 3 componentes mencionados, pero con sus propios datos y consultas. En el artículo se propone que el *benchmarking* propuesto es posible cambiar en función de las necesidades de la comparación, por lo que para este proyecto terminal se adaptará los componentes de la evaluación para llevarlo a cabo con el software de Apache Marmotta una vez implementado el módulo.

Después de realizar un análisis en [3], se determinó que la información del artículo anterior no era suficiente para delimitar el *benchmarking* del proyecto actual por lo que se buscaron otras alternativas.

Gabriela Montoya y su equipo en España desarrollaron un artículo que cubriera las limitaciones del *benchmarking* que Michael había desarrollado. En [6] se especifica más sobre *FedBench* y sobre los componentes que impactan el rendimiento de sistemas de consultas federadas.

Con base en [6], existen 2 variables que influyen en los sistemas de consultas federadas y son

- Variables independientes: Son aquellas características que debe de ser especificadas con el fin de asegurar que sean replicables en los escenarios de evaluación. Las variables independientes que se proponen usar son:
  - Consulta: Variable que definen las consultas en términos de su estructura, evaluación y expresividad del lenguaje. Se basa en 3 aspectos
    - Forma del plan de ejecución: La forma en que los planes de ejecución se generan en los sistemas pueden variar, por lo que tiene un importante impacto en las 3 variables dependientes.
    - Número de patrones de tripletas básicas en la consulta: Este aspecto afecta directamente al plan de ejecución debido a que depende del número de tripletas que se recuperan en función de las tripletas especificadas en la consulta.
    - Instancias y posición en las tripletas: Este aspecto se refiere a la relación que existe entre el sujeto, el objeto y el predicado en las tripletas. Puede tener un gran impacto en la evaluación ya que, por ejemplo, el predicado de una tripla puede no estar instanciado dando lugar a que diversos *endpoints* respondan cuando una consulta requiera información de dicho predicado.
  - Datos: Esta variable está compuesta por 4 elementos
    - Tamaño del RDF *dataset*.
    - Características estructurales del *dataset* las cuales se refieren al número de predicados, objetos y sujetos. Estas características influyen en el número de tripletas que son transferidas y por consecuencia, el tiempo del plan de ejecución total.
    - Tipo de partición: Se refiere a cómo está fragmentado el *dataset*. La partición puede ser horizontal y vertical. Dependiendo del tipo de partición puede influir en las 3 variables dependientes ya que la cantidad de información puede ser menor, al tener menor cantidad de filas, o parcial al solo considerar ciertas propiedades.
    - Distribución de datos: Se refiere a cómo las particiones están distribuidas en los *endpoints*, también afecta las variables dependientes. Pueden estar completamente distribuido, centralizado o mixto.
  - Plataforma: Esta variable agrupa a otras relacionadas a la infraestructura de cómputo durante la evaluación. Los elementos que componen esta variable son.
    - Memoria RAM.
    - Número de procesadores.

- Administración de memoria caché.
- *Endpoint*: Esta variable se refiere a los elementos asociados al número y capacidades de *SPARQL endpoints* usados en la evaluación. Los elementos son
  - Número de *SPARQL endpoints* a visitar por la consulta cuando esta sea llevada a cabo.
  - Tipo de *SPARQL endpoints* usados en la evaluación.
  - Distribución de las transferencias de datos la cual es la distribución estadística de tiempos de transmisión hechos por los *SPARQL endpoints*.
  - Latencia de red: retraso en el envío de paquetes a través de la red.
  - Retraso inicial del *SPARQL endpoint*.
- Variables dependientes: Son las características que están influenciadas por las variables independientes. Las variables dependientes son las que serán medidas en la evaluación. Las variables dependientes que se proponen son
  - Tiempo de selección del *endpoint*: Tiempo transcurrido entre la realización de la consulta y la generación de la consulta federada, basado en el protocolo *SPARQL* 1.1 y los *endpoint* donde se llevarán a cabo las subconsultas.
  - Tiempo de ejecución: Esta variable comprende 3 elementos
    - Tiempo entre la realización de la consulta y la primera respuesta.
    - Distribución de tiempo en la recepción de respuestas de las consultas.
    - Tiempo total de ejecución.
  - Completitud de respuestas: Cantidad de respuestas recibidas comparadas con los datos disponibles en los *endpoints* seleccionados.

El *benchmarking* permitirá describir el módulo de consultas federadas y hacer una comparación respecto a otros sistemas similares. Opcionalmente, ya que no es el objetivo de la actividad, se podría optimizar el módulo de consultas federadas en el *software* Apache Marmotta utilizando los resultados obtenidos del *benchmarking*

La tabla 1 muestra la relación existente entre las variables independientes y las dependientes.

Tabla 1 Relación existente entre variables dependientes e independientes

Variables independientes		Variables dependientes		
		Tiempo de selección del <i>SPARQL endpoint</i>	Tiempo de ejecución	Complejidad de la respuesta
Consulta	Forma del plan de ejecución	Si	Si	Si
	Número de patrones de tripletas básicas en la consulta	Si	Si	Si
	Instancias y posición en las tripletas	Si	Si	No
Datos	Tamaño del RDF <i>dataset</i>	No	Si	No
	Características estructurales del <i>dataset</i>	No	Si	No
	Tipo de partición	Si	Si	Si
	Distribución de datos	Si	Si	Si
Plataforma	Memoria RAM	Si	Si	No
	Número de procesadores	Si	Si	No
	Administración de memoria caché	Si	Si	No
<i>SPARQL endpoint</i>	Número de <i>SPARQL endpoints</i>	Si	Si	Si
	Tipo de <i>SPARQL endpoints</i>	Si	Si	No
	Distribución de las transferencias	Si	Si	Si
	Latencia de red	Si	Si	Si
	Retraso inicial del <i>SPARQL endpoint</i>	Si	Si	No

En resumen, los 3 parámetros a evaluar en el *benchmarking* son

- Tiempo de selección del *SPARQL endpoint*
- Tiempo de ejecución.
- Completitud de respuesta.

Cabe decir que el *benchmarking* a realizar en el proyecto terminal 2, estará contextualizado en datos geoespaciales por lo que las consultas [7] serán modificadas y adaptadas, tal y como *FedBench* lo sugiere.

Una lista de *triple store* a consultar pueden ser los mostrados en la tabla 2, aunque no son definitivos.

Tabla 2 Triple stores y su cantidad de tripletas disponibles

Sitio Web	Total tripletas
<a href="http://www.geonames.org/ontology/">http://www.geonames.org/ontology/</a>	93,896,732
<a href="http://lod.taxonconcept.org/">http://lod.taxonconcept.org/</a>	34,394,994
<a href="http://eunis.eea.europa.eu">http://eunis.eea.europa.eu</a>	20,229,105
<a href="http://data.fishesoftexas.org">http://data.fishesoftexas.org</a>	528,897
<a href="http://lod.geospecies.org/">http://lod.geospecies.org/</a>	2,163,922
<a href="http://ja.dbpedia.org/">http://ja.dbpedia.org/</a>	70,790,371
<a href="https://www.geospatialworld.net/">https://www.geospatialworld.net/</a>	27,017
<a href="http://linkedgedata.org/">http://linkedgedata.org/</a>	3,000,000,000
<a href="http://www.best-project.nl">http://www.best-project.nl</a>	1,931,860
<a href="http://webenemasuno.linkeddata.es/">http://webenemasuno.linkeddata.es/</a>	9,262,184
<a href="http://geo.linkeddata.es/">http://geo.linkeddata.es/</a>	21,564,199
<a href="http://www.rdfabout.com/demo/census/">http://www.rdfabout.com/demo/census/</a>	1,002,848,918
<a href="http://www4.wiwiw.fu-berlin.de/eurostat/">http://www4.wiwiw.fu-berlin.de/eurostat/</a>	8,850
<a href="http://www4.wiwiw.fu-berlin.de/factbook/">http://www4.wiwiw.fu-berlin.de/factbook/</a>	38,640
<a href="http://data.ordnancesurvey.co.uk/">http://data.ordnancesurvey.co.uk/</a>	2,788,872
<a href="http://statistics.data.gov.uk/">http://statistics.data.gov.uk/</a>	343,733
<a href="https://www.pdok.nl/introductie?articleid=1948911">https://www.pdok.nl/introductie?articleid=1948911</a>	1,000,000,000
<a href="http://mapasinteractivos.didactalia.net/comunidad/mapasflashinteractivos">http://mapasinteractivos.didactalia.net/comunidad/mapasflashinteractivos</a>	90,226
<a href="http://thesaurus.iiia.cnr.it/index.php/vocabularies/earth">http://thesaurus.iiia.cnr.it/index.php/vocabularies/earth</a>	133,315
<a href="http://purl.org/NET/ThISTWebPage">http://purl.org/NET/ThISTWebPage</a>	396,885
<a href="http://www.morelab.deusto.es">http://www.morelab.deusto.es</a>	3,561
<a href="http://transparency.270a.info/">http://transparency.270a.info/</a>	41,592
<a href="http://worldbank.270a.info/">http://worldbank.270a.info/</a>	175,000,000
<a href="http://id.ecs.soton.ac.uk/docs/">http://id.ecs.soton.ac.uk/docs/</a>	50,000
<a href="http://eprints.ecs.soton.ac.uk/">http://eprints.ecs.soton.ac.uk/</a>	1,000,000
<a href="http://telegraphis.net/data/">http://telegraphis.net/data/</a>	15,257
<a href="http://nuts.psi.enakting.org/">http://nuts.psi.enakting.org/</a>	40,000
<a href="http://nuts.geovocab.org/">http://nuts.geovocab.org/</a>	300,000

<a href="http://www.openmobilenetwork.org">http://www.openmobilenetwork.org</a>	19,048,586
<a href="https://lod-cloud.net/dataset/linkedcrowdsourceddata">https://lod-cloud.net/dataset/linkedcrowdsourceddata</a>	17,043
<a href="http://lobid.org/resource">http://lobid.org/resource</a>	702,686,189
<a href="http://lobid.org/organisation">http://lobid.org/organisation</a>	776,735
<a href="http://purl.org/NET/marccodes/">http://purl.org/NET/marccodes/</a>	8,816
<a href="http://purl.org/smartlink">http://purl.org/smartlink</a>	2,829
<a href="http://education.data.gov.uk/">http://education.data.gov.uk/</a>	6,619,847
<a href="http://ec.europa.eu/eurostat/ramon/rdfdata/">http://ec.europa.eu/eurostat/ramon/rdfdata/</a>	68,000
<a href="http://wiki.knoesis.org/index.php/SSW_Datasets">http://wiki.knoesis.org/index.php/SSW_Datasets</a>	1,730,284,735
<a href="http://www.opencalais.com/">http://www.opencalais.com/</a>	4,500,000
<a href="http://openlylocal.com/">http://openlylocal.com/</a>	10,000,000
<a href="http://theviewfrom.org/">http://theviewfrom.org/</a>	368
<a href="http://transport.data.gov.uk/">http://transport.data.gov.uk/</a>	329,527,661
<a href="http://www.uk-postcodes.com/">http://www.uk-postcodes.com/</a>	27,137,408
<a href="https://lod-cloud.net/dataset/geoecuador">https://lod-cloud.net/dataset/geoecuador</a>	1,278,288
<a href="http://www.kanzaki.com/works/2011/stat/ra/">http://www.kanzaki.com/works/2011/stat/ra/</a>	1,648,890
<a href="http://greek-lod.math.auth.gr/fire-brigade/">http://greek-lod.math.auth.gr/fire-brigade/</a>	4,134,837
<a href="http://greek-lod.math.auth.gr/police/">http://greek-lod.math.auth.gr/police/</a>	145,368
<a href="http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/">http://www.linkedopenservices.org/services/geo/SpatialResources/point/ICAO/</a>	75,184
<a href="http://eurostat.linked-statistics.org/">http://eurostat.linked-statistics.org/</a>	8,000,000,000
<a href="http://geowordnet.semanticmatching.org/">http://geowordnet.semanticmatching.org/</a>	53,390,969
<a href="http://kent.zpr.fer.hr:8080/educationalProgram/">http://kent.zpr.fer.hr:8080/educationalProgram/</a>	2,373,863
<a href="http://datos.fundacionctic.org/en/">http://datos.fundacionctic.org/en/</a>	535,230
<a href="http://seobook.blog.com">http://seobook.blog.com</a>	40,798
<a href="http://data-gov.ie/">http://data-gov.ie/</a>	12,390,973
<a href="http://www.fao.org/countryprofiles/geoinfo.asp?lang=en">http://www.fao.org/countryprofiles/geoinfo.asp?lang=en</a>	22,495
<a href="http://ontologycentral.com/2009/01/eurostat/">http://ontologycentral.com/2009/01/eurostat/</a>	40,000,000
<a href="http://data.lenka.no/">http://data.lenka.no/</a>	8,399
<a href="http://airports.dataincubator.org/">http://airports.dataincubator.org/</a>	748,363
<a href="http://www.josemalvarez.es/web/2011/11/01/nomenclator-asturias-2010/">http://www.josemalvarez.es/web/2011/11/01/nomenclator-asturias-2010/</a>	4,508,050
<a href="http://www.josemalvarez.es/web/2011/11/16/product-scheme-classifications/">http://www.josemalvarez.es/web/2011/11/16/product-scheme-classifications/</a>	1,842,053
<a href="http://kasabi.com/dataset/yahoo-geoplanet">http://kasabi.com/dataset/yahoo-geoplanet</a>	49,734,022
<a href="https://lod-cloud.net/dataset/ww1lod">https://lod-cloud.net/dataset/ww1lod</a>	40,160
<a href="http://greek-lod.math.auth.gr/kalikratis/">http://greek-lod.math.auth.gr/kalikratis/</a>	31,068
<a href="http://data.oceandrilling.org/linkeddata.html">http://data.oceandrilling.org/linkeddata.html</a>	24,489,366
<a href="http://data.oceandrilling.org/codices/">http://data.oceandrilling.org/codices/</a>	24,489,366
<a href="http://ecowlim.tfri.gov.tw/lode">http://ecowlim.tfri.gov.tw/lode</a>	23,015,257
<a href="http://www.ldf.fi/dataset/warsa">http://www.ldf.fi/dataset/warsa</a>	12,677,886
<a href="http://www.linkedopendata.it/datasets/grp">http://www.linkedopendata.it/datasets/grp</a>	153,935
<a href="http://smcjournals.dataincubator.org">http://smcjournals.dataincubator.org</a>	15,801
<a href="http://climb.dataincubator.org/">http://climb.dataincubator.org/</a>	174,120
<a href="http://pleiades.stoa.org/">http://pleiades.stoa.org/</a>	2,600,000
<a href="http://chronos.org/janusAmp/">http://chronos.org/janusAmp/</a>	24,489,366

<a href="http://aemet.linkeddata.es/">http://aemet.linkeddata.es/</a>	62,932,032
<a href="http://environment.data.gov.uk/lab">http://environment.data.gov.uk/lab</a>	8,735,962
<a href="http://metoffice.dataincubator.org">http://metoffice.dataincubator.org</a>	26,000,000
<a href="http://ieevis.tw.rpi.edu">http://ieevis.tw.rpi.edu</a>	19,935,340
<a href="http://kasabi.com/dataset/renewable-energy-generators">http://kasabi.com/dataset/renewable-energy-generators</a>	99,917
<a href="http://onki.fi/en/browser/overview/kunnat">http://onki.fi/en/browser/overview/kunnat</a>	11,380
<a href="http://eurobarometer.publicdata.eu">http://eurobarometer.publicdata.eu</a>	1,193,494
<a href="http://my-family-lineage.com/wiki/Main_Page">http://my-family-lineage.com/wiki/Main_Page</a>	3,000,000
<a href="http://opendatacommunities.org/datasets/imd-score-2010">http://opendatacommunities.org/datasets/imd-score-2010</a>	194,892
<a href="http://riese.joanneum.at/data/">http://riese.joanneum.at/data/</a>	5,000,000

Los datos de la tabla 2 se extrajeron del portal web de la nube del *Linked Data*<sup>1</sup> usando un *script* en Python para extraer los sitios web que tengan datos geoespaciales. El script es el que se muestran en la figura 1 y 2. El código puede visitarlo en el GitHub<sup>2</sup> del estudiante.

```
# Importando librerías necesarias para el Web Scrapper
import requests
import urllib.request
import time
from bs4 import BeautifulSoup
import re # Librería para expresiones regulares
from selenium import webdriver # Librería para renderizar el HTML
from requests_html import HTMLSession
import numpy
import csv

# https://lod-cloud.net/clouds/lod-cloud.svg

def get_geourls(url_set):
    session = HTMLSession();
    resp = session.get(url_set);
    resp.html.render();
    soup = BeautifulSoup(resp.html.html, "html.parser");
    total_etiquetas = len(soup.findAll(class_="badge badge-primary"));
    # Bandera que indica que las etiquetas del triple store contiene geo-datos
    flag_geo = False;
    # Se analiza cada etiqueta
    labels = [];
    for m in range(total_etiquetas):
        # Se indica que etiquetas HTML se deben de extraer
        one_a_tag = soup.findAll(class_="badge badge-primary")[m];
        # Se guardan dichas etiquetas
        labels.append(one_a_tag.contents[0]);
        # Se extraen las etiquetas de cada burbuja
        # y se chequea que tenga asociado la preposición 'geo'
        if (re.findall('geo', one_a_tag.contents[0])):
            flag_geo = True;
    # En dado caso que el triple store contenga geodatos, se continua con la extracción de datos
    if (flag_geo):
        print("Datos Geoespaciales");
        # Con base al DOM de la página, se lleva a cabo la siguiente búsqueda y extracción de datos
        # Sitio Web
        try:
            website = soup.findAll(class_="website")[0].findAll('a')[0]['href'];
        except:
            print("Triple store sin website");
            website = NA;
        print(website);
        # Total de tripletas
        total_triples = soup.findAll(class_="table table-striped")[0].findAll('td')[1].contents[0];
        resp.close();
        session.close();
        return [website, total_triples, labels];
    # En caso negativo, se regresa de la función para continuar con el siguiente enlace
    else:
        print("No geodatos");
        resp.close();
        session.close();
        return;
```

Fig. 1 Web scrapper para la nube de Linked Data

<sup>1</sup> <https://lod-cloud.net>

<sup>2</sup> <https://github.com/Osw1997/PT1-scripts.git>



```

def get_urls(cantidad):
    # URL donde extraer la info
    url = 'https://lod-cloud.net/clouds/lod-cloud.svg';
    #url = 'http://web.mta.info/developers/turnstile.html';
    # Se establece la conexión a la URL
    response = requests.get(url);
    # Se utiliza la herramienta BeautifulSoup para traducir el HTML en una
    # estructura 'BeautifulSoup'
    soup = BeautifulSoup(response.text, 'html.parser');
    # Se cuentan cuantas burbujas hay que analizar
    total_burbujas = len(soup.findAll('a'));
    print(total_burbujas)
    # Array de todas las URLs a analizar
    url_burbujas = [];
    # Se analiza cada burbuja
    for m in range(total_burbujas):
        # Se indica que etiquetas HTML se deben de extraer
        one_a_tag = soup.findAll('a')[m];
        # Se extraen las URLs de cada burbuja
        link = one_a_tag['href'];
        url_burbujas.append(link);
    with open("LinkedDataURL.csv", "w") as f:
        for url in url_burbujas:
            f.write("%s\n" % url);

def get_geoinfo():
    url_burbujas = [];
    with open('LinkedDataURL.csv', newline='') as csvfile:
        spamreader = csv.reader(csvfile, delimiter=' ', quotechar='|')
        url_burbujas = list(spamreader)
    print(url_burbujas);
    # Se extraen las páginas que contienen geodatos
    info = [['URL', 'Website', 'Total tripletas', 'Etiquetas'],];
    separador = ',';
    for indx, url in enumerate(url_burbujas):
        print(indx);
        info_per_url = [0,0,0,0]
        info_per_url[0] = url[0]
        data = get_geourls(url[0]);
        if (data is not None):
            info_per_url[1] = data[0]
            info_per_url[2] = data[1]
            info_per_url[3] = separador.join(data[2])
            info.append(info_per_url);
    print(info);
    print(type(info));

    with open("LinkedData_completo.csv", "w") as f:
        writer = csv.writer(f)
        writer.writerows(info)

# Inicio web scrapping
get_geoinfo()

```

Fig. 2 Web scrapper para la nube de Linked Data - continuación

## Referencias

- [1] UNAM, «Benchmarking,» UNAM, 23 Noviembre 2018. [En línea]. Available: [http://depa.fquim.unam.mx/amyd/archivero/TRABAJO7\\_2848.pdf](http://depa.fquim.unam.mx/amyd/archivero/TRABAJO7_2848.pdf). [Último acceso: 19 Octubre 2019].
- [2] D. Tow, *SQL Tuning: Generating Optimal Execution Plans*, California: O'Really Media, In, 2003.
- [3] M. Shmidt, O. Görlitz, P. Haase, G. Ladwig, A. Shwarte y T. Tran, «FedBench: A Benchmark Suite for Federated Semantic Data Query Processing,» de *International Semantic Web Conference*, Berlín, 2011.
- [4] M. Schmidt, T. Hornung, G. Lausen y C. Pinkel, «SP<sup>2</sup> Bench: a SPARQL performance benchmark,» de *2009 IEEE 25th International Conference on Data Engineering*, Freiburg, Alemania, 2009.
- [5] C. Bizer y A. Schultz, «The berlin sparql benchmark,» *International Journal on Semantic Web and Information Systems*, vol. 5, n° 2, pp. 1-24, 2009.
- [6] G. Montoya, M. E. Vidal, O. Corcho, E. Ruckhaus y C. Buil-Aranda, «Benchmarking Federated SPARQL Query Engines: Are Existing Testbeds Enough?,» de *International Semantic Web Conferernce*, Berlín, 2012.
- [7] C. Buil-Aranda y O. Corcho García, *Federated Query Processing for the*, Madrid: IOS Press, 2012.