

# SEMTEST: Un demostrador pedagógico de enriquecimiento semántico multicapa basado en Linked Open Data para la docencia en Información y Documentación

## SEMTEST: A pedagogical multi-layer semantic enrichment demonstrator based on Linked Open Data for teaching in Information and Documentation Science

Manuel Blázquez-Ochando; María-Antonia Ovalle-Perandones

Citación recomendada:

**Blázquez-Ochando, Manuel; Ovalle-Perandones, María-Antonia** (2026). "SEMTEST: Un demostrador pedagógico de enriquecimiento semántico multicapa basado en Linked Open Data para la docencia en Información y Documentación [SEMTEST: A pedagogical multi-layer semantic enrichment demonstrator based on Linked Open Data for teaching in Information and Documentation Science]". *Infonomy*, 4(4) e26022.  
<https://doi.org/10.3145/infonomy.26.022>

Artículo recibido: 06-05-2026  
Artículo aprobado: 07-06-2026



**Manuel Blázquez-Ochando**  
<https://orcid.org/0000-0002-4108-7531>  
<https://directorioexit.info/ficha1215>  
Universidad Complutense de Madrid  
Facultad de Ciencias de la Documentación  
Departamento de Biblioteconomía y Documentación  
Santísima Trinidad, 37  
28010 Madrid, España  
[manublaz@ucm.es](mailto:manublaz@ucm.es)





**María-Antonia Ovalle-Perandones**

<https://orcid.org/0000-0002-6149-4724>

<https://directorioexit.info/ficha509>

Universidad Complutense de Madrid

Facultad de Ciencias de la Documentación

Departamento de Biblioteconomía y Documentación

Santísima Trinidad, 37

28010 Madrid, España

[maovalle@ucm.es](mailto:maovalle@ucm.es)

## Resumen

Se presenta la segunda versión de SEMTEST que es una herramienta web interactiva diseñada para la enseñanza del enriquecimiento semántico de consultas en el ámbito de la Biblioteconomía y la Documentación. A partir de un término libre introducido por el usuario, la aplicación ejecuta un pipeline de nueve fases secuenciales que ilustra, de forma completamente transparente y trazable, la transición desde la Web de documentos (*Wikipedia*, HTML) hacia la Web de Datos (*DBpedia*, *Wikidata*), integrando asimismo recursos bibliográficos abiertos (*Open Library*). La principal novedad de esta versión reside en la profundización en la representación del conocimiento mediante la incorporación explícita de todas las capas de la arquitectura de la Web Semántica –desde los identificadores URI e IRI y la sintaxis XML hasta las taxonomías RDFS, los esquemas SKOS, las ontologías OWL y la interrogación SPARQL– tanto en el proceso de extracción y clasificación de propiedades RDF como en la visualización del grafo resultante y en los formatos de exportación estándar. El artículo describe con detalle la arquitectura técnica del demostrador, sus mecanismos de acceso a APIs REST y endpoints SPARQL, la lógica de clasificación pedagógica de propiedades, y la generación de grafos de conocimiento multinivel que reflejan relaciones jerárquicas, asociativas y de instancia entre entidades. Se analizan asimismo sus posibilidades pedagógicas en el contexto de la formación en organización del conocimiento y Web Semántica. Se concluye que SEMTEST constituye un recurso didáctico eficaz para desmitificar las tecnologías semánticas y evidenciar su continuidad conceptual con los procesos documentales tradicionales.

## Palabras clave

Web Semántica; Linked Open Data; Enriquecimiento semántico; RDF; SPARQL; *DBpedia*; *Wikidata*; SKOS; OWL; Grafos de conocimiento; Docencia en Documentación; Demostrador pedagógico; API; Recuperación de información; BIBFRAME; FOAF; BIBO; URI.

## Abstract

SEMTEST is presented as an interactive web tool designed for teaching semantic query enrichment in the field of Library and Information Science. Starting from a free term entered by the user, the application executes a nine-phase sequential pipeline that transparently and traceably illustrates the transition from the Document Web (*Wikipedia*, HTML) to the Web of Data (*DBpedia*, *Wikidata*), also integrating open

bibliographic resources (*Open Library*). The main novelty of this version lies in the deepening of knowledge representation through the explicit incorporation of all layers of the Semantic Web architecture –from URI/IRI identifiers and XML syntax to RDFS taxonomies, SKOS schemas, OWL ontologies, and SPARQL querying– both in the RDF property extraction and classification process and in the visualization of the resulting graph and standard export formats. The article describes in detail the technical architecture of the demonstrator, its mechanisms for accessing REST APIs and SPARQL endpoints, the logic of pedagogical property classification, and the generation of multi-level knowledge graphs reflecting hierarchical, associative, and instance relationships between entities. Its pedagogical possibilities in the context of knowledge organisation and Semantic Web training are also analysed. It is concluded that SEMTEST constitutes an effective didactic resource for demystifying semantic technologies and demonstrating their conceptual continuity with traditional documentary processes.

### Keywords

Semantic Web; Linked Open Data; Semantic Enrichment; RDF; SPARQL; *DBpedia*; *Wikidata*; SKOS; OWL; Knowledge Graphs; Documentation Teaching; Pedagogical Demonstrator; API; Information Retrieval; BIBFRAME; FOAF; BIBO; URI.

### 1. Introducción

La propuesta de la Web Semántica como extensión formal de la Web hipertextual – formulada inicialmente por **Berners-Lee** (1998) y desarrollada con Miller (**Berners-Lee; Miller**, 2002) como un espacio donde la información disponible adquiere un significado bien definido que permite a los ordenadores y a las personas trabajar en cooperación– continúa siendo un horizonte de referencia en la investigación sobre organización del conocimiento y recuperación de información. Desde ese momento fundacional, el Consorcio W3C ha desplegado una arquitectura de capas –denominada habitualmente *Semantic Web Layer Cake*– que articula un conjunto de estándares interdependientes: la identificación unívoca de recursos mediante URI e IRI como capa base, la sintaxis XML en la siguiente capa, el modelo de datos RDF para el intercambio de información, los vocabularios RDFs para las taxonomías, los sistemas de organización del conocimiento representados con SKOS, las ontologías en OWL y el lenguaje de consulta SPARQL como mecanismo de interrogación (**Berners-Lee**, 1998; W3C, 2013a; **Balloni et al.**, 2012; **Codina; Pedraza**, 2017).

Esta arquitectura ha dado lugar a un ecosistema operativo de datos abiertos enlazados (*Linked Open Data* o LOD) de escala global. El crecimiento de la nube o *cloud* LOD es sostenido. Así, **Bizer et al.** (2009) establecieron los principios fundacionales de la publicación de datos enlazados, y **Hogan et al.** (2021) documentan cómo los grafos de conocimiento –desde *DBpedia* hasta *Wikidata*, pasando por repositorios especializados– han llegado a constituir infraestructuras de conocimiento globales con miles de millones de tripletas RDF. El concepto de Web 2.0 como plataforma participativa que describió **O'Reilly** (2007) fue, en cierta medida, el prelude social de este ecosistema: la Web semántica extiende esa participación al nivel de los datos estructurados y las relaciones formales entre entidades.

En el ámbito de las bibliotecas y los servicios de información, esta transformación se materializa en iniciativas como BIBFRAME (*Library of Congress*, 2012), que propone una transición desde los registros MARC hacia un modelo RDF capaz de expresar relaciones entre obras, instancias, personas y conceptos. Casi paralelamente, en el modelo de referencia IFLA LRM (**Riva et al.**, 2017), se reformulan las entidades bibliográficas en términos compatibles con los grafos de conocimiento. **Gaitanou et al.** (2024), en su revisión sistemática de literatura sobre datos enlazados en bibliotecas publicada en el *Journal of Information Science*, confirman que la interoperabilidad semántica y la capacidad de conectar colecciones bibliotecarias al espacio LOD son objetivos consolidados en la literatura, aunque la distancia entre los principios teóricos y la implementación práctica sigue siendo considerable.

Una dimensión fundamental de esta brecha es la formativa. La enseñanza de las tecnologías semánticas en los estudios de Biblioteconomía y Documentación enfrenta una dificultad estructural: la distancia entre la comprensión teórica de los estándares y su aplicación práctica en contextos reales. **Méndez y Greenberg** (2012) identificaron con precisión cuatro comunidades de usuarios de las tecnologías LOD —desarrolladores web, comunidad de la Web Semántica del W3C, comunidades de metadatos y usuarios de *knowledge organization systems* (KOS) o sistemas de organización del conocimiento— y señalaron que los profesionales de la Documentación, con sus competencias en organización del conocimiento, constituyen un puente natural entre los vocabularios controlados tradicionales y los grafos formales, aunque ese puente no siempre se visibiliza pedagógicamente. **Lehmann et al.** (2015), en la descripción técnica de *DBpedia* publicada en la revista *Semantic Web*, muestran cómo la ontología *DBpedia* —con sus más de 320 clases y 1.650 propiedades extraídas de las infoboxes de *Wikipedia*— es precisamente la materialización en RDF de estructuras conceptuales que los documentalistas reconocen en sus propias prácticas de catalogación y análisis de materias.

La interoperabilidad semántica y la capacidad de conectar colecciones bibliotecarias al espacio LOD son objetivos consolidados en la literatura, aunque la distancia entre los principios teóricos y la implementación práctica sigue siendo considerable

El problema específico que SEMTEST aborda es la opacidad de los sistemas de enriquecimiento semántico disponibles. Los grafos de conocimiento como *DBpedia* (**Lehmann et al.**, 2015) y *Wikidata* (**Vrandečić; Krötzsch**, 2014) exponen sus datos mediante endpoints SPARQL públicos y APIs REST documentadas; sin embargo, para el estudiante de Documentación sin formación técnica previa, estas interfaces resultan opacas: se introduce un término y se obtienen resultados sin comprender los mecanismos subyacentes, las cabeceras HTTP enviadas, las estructuras JSON de respuesta, la diferencia entre un URI y una URL o el significado de una propiedad como `rdfs:subClassOf`. Esta opacidad impide que el estudiante establezca la conexión entre los conceptos que ya conoce —términos genéricos y específicos en un tesoro, control de autoridades, análisis de materias— y sus equivalentes formales en el entorno LOD.

Para abordar este problema se desarrolló SEMTEST (Semantic Enrichment Test), un demostrador interactivo en PHP y JavaScript que convierte el proceso de enriquecimiento semántico de una consulta, en un recorrido didáctico visible y trazable. La versión 2, la que aquí se presenta, supone una renovación integral de la herramienta original: reemplaza el scraping HTML frágil por APIs JSON oficiales, amplía los *timeouts* y cabeceras cURL para garantizar la fiabilidad de las peticiones, alinea explícitamente cada fase del pipeline con una capa de la arquitectura de la Web Semántica, introduce la clasificación pedagógica de propiedades RDF por tipo ontológico, y transforma el grafo de visualización de una red plana en un grafo multicapa con exportación estándar en *Turtle*.

SEMTEST (Semantic Enrichment Test) es un demostrador interactivo en PHP y JavaScript que convierte el proceso de enriquecimiento semántico de una consulta en un recorrido didáctico visible y trazable

El artículo tiene como objetivos de estudio describir el diseño, la arquitectura técnica y las funciones pedagógicas de SEMTEST, analizando en qué grado permite a los estudiantes de Documentación visualizar, experimentar y comprender los mecanismos de la Web Semántica a partir de sus propias consultas, y en qué medida hace visible la continuidad conceptual entre los procesos documentales tradicionales de organización del conocimiento y las tecnologías de la Web de Datos. La herramienta está disponible en acceso abierto en:

<https://mblazquez.es/lab/semTest2>

## 2. Arquitectura y pipeline semántico

### 2.1. Principios técnicos de diseño

SEMTEST está desarrollado en PHP 8.x (*backend*) y JavaScript/D3.js v7 (*frontend*). La arquitectura *backend* se basa en un patrón de peticiones encadenadas: cada fase del pipeline realiza una o varias llamadas HTTP a APIs externas mediante la biblioteca cURL de PHP, procesa la respuesta JSON o HTML y la envía al *frontend*, donde se renderiza de forma progresiva mediante AJAX. El resultado es que el usuario ve cómo se construye el enriquecimiento semántico en tiempo real, fase a fase, con una barra de progreso que refleja el estado de cada operación.

El principio rector de diseño es la **transparencia radical**: cada fase no solo muestra los datos recuperados, sino también la URL exacta llamada, las cabeceras HTTP enviadas —incluyendo el *User-Agent*, el *Accept* y el *Content-Type*—, la respuesta JSON en bruto y el fragmento de código PHP responsable de la petición. Esta transparencia convierte a SEMTEST en una herramienta de aprendizaje activo: el estudiante no lee sobre APIs ni sobre tripletas RDF; lo ve funcionar sobre sus propias consultas.

El cambio técnico más profundo respecto a la versión original reside en la **sustitución completa del scraping HTML por APIs JSON oficiales**. La versión inicial construía URLs de búsqueda de *Wikipedia* como páginas web con parámetros obsoletos (*profile=advanceddefault*), obtenía los backlinks mediante el scraping de *Especial:LoQueEnlazaAquí* —cuya estructura HTML cambió con las actualizaciones de *MediaWiki*— y construía las URLs de *DBpedia* de forma manual sin resolver

ambigüedades de nombre. La versión 2 reemplaza todas estas llamadas frágiles por sus equivalentes en las APIs oficiales: `action=query&list=search` para la búsqueda en *Wikipedia*, `prop=linkshere` con `pageID` estable para los backlinks, y el *DBpedia* Lookup Service para la resolución canónica de URIs. La tabla 1 sintetiza las principales diferencias de la versión 2 respecto de la original.

Tabla 1. Comparativa de versiones, mejoras y evolución

Función	Versión original	SEMTEST v2
Búsqueda <i>Wikipedia</i>	Scraping HTML con <code>profile=advanceddefault</code>	<code>action=query&amp;list=search&amp;srwhat=text</code>
Contenido artículo	Scraping XPath del HTML	<code>prop=extracts links + action=parse&amp;prop=sections</code>
Backlinks	Scraping de <code>Especial:LoQueEnlazaAquí</code>	<code>prop=linkshere&amp;lhnamespace=0</code> con <code>pageID</code>
Imágenes Commons	Scraping HTML con URLs relativas	<code>action=query&amp;list=search + prop=imageinfo&amp;iiprop=url</code>
<i>DBpedia</i> URI	URL manual <code>/page/Término</code>	Lookup API → URI canónico → datos RDF
<i>Wikidata</i>	Scraping HTML de la página de entidad	<code>wbsearchentities + wbgetentities&amp;props=claims labels</code>
SPARQL	No existía	Llamada directa con <code>Accept: application/sparql-results+json</code>
<i>Open Library</i>	No existía	API <code>search.json</code>
Grafo	Red plana D3.js de un solo nivel	Grafo multicapa jerárquico con 5 capas semánticas
Exportación	JSON interno	JSON-LD, Turtle (.ttl), D3-JSON, SVG
Timeouts cURL	1 s conexión / 4 s total	10 s conexión / 30 s total

A nivel de gestión de conexiones, se introducen tres funciones cURL diferenciadas según el tipo de recurso:

- `getcurl_api()`: para llamadas a APIs JSON (*Wikipedia*, *Wikidata*, *Open Library*). Timeouts de 10 s de conexión y 30 s de respuesta total; cabecera `User-Agent: SEMTEST/2.1 (https://mblazquez.es; manublaz@ucm.es)`; cabecera `Accept: application/json`.
- `getcurl_sparql()`: para el endpoint SPARQL de *DBpedia*. Idéntica configuración a la anterior, pero con la cabecera `Accept: application/sparql-results+json`, diferente del tipo MIME genérico y requerido por el protocolo SPARQL 1.1 (W3C, 2013a).

- `getcurl1()`: para el acceso HTML a páginas de recursos en *DBpedia* cuando se necesita la representación de navegador del recurso.

Esta diferenciación técnica tiene valor pedagógico directo: el estudiante comprende por qué el endpoint SPARQL requiere una cabecera diferente, y por qué *Wikipedia* bloquea peticiones sin *User-Agent*. La infraestructura viva de la Web Semántica no es solo un objeto de estudio, sino el entorno operativo real de la herramienta.

## 2.2. Las nueve fases del pipeline

### Fase 1: Búsqueda y desambiguación (*capa URI/Unicode de la Semantic Web Layer*)

Cuando el usuario introduce un término en el formulario —puede ser desde un concepto abstracto como "Inteligencia artificial" hasta un nombre propio como "Pablo Picasso" o un término técnico como "Ontología"— SEMTEST lanza inmediatamente una consulta estructurada contra la API JSON de *Wikipedia* en español utilizando `action=query&list=search&srwhat=text&srsearch=TÉRMINO&srlimit=10`. El sistema recupera hasta diez artículos relacionados, ordenados por relevancia, mostrando para cada uno su título, identificador de página (`pageID`), tamaño en bytes, número de palabras y un fragmento de texto (snippet) que permite evaluar la pertinencia del resultado.

Esta fase tiene un objetivo pedagógico preciso: ilustrar la diferencia entre la URL de navegación ([https://es.wikipedia.org/wiki/Inteligencia\\_artificial](https://es.wikipedia.org/wiki/Inteligencia_artificial)) y el `pageID` numérico (un entero como `3249853`), que es el identificador interno estable de *Wikipedia*, independiente de cualquier cambio en el título del artículo. El `pageID` es el equivalente funcional de un URI en el ecosistema *MediaWiki*: identifica de forma unívoca un recurso independientemente de su localización. Esta distinción entre identificación y localización —entre URI y URL en el vocabulario de la Web Semántica (Berners-Lee, 2006)— es uno de los conceptos más difíciles de transmitir de forma abstracta, pero resulta inmediatamente comprensible cuando el estudiante ve que el mismo artículo puede ser accedido por título ("Inteligencia artificial", que puede cambiar) o por su `pageID` (que permanece fijo). El primer resultado de la lista se convierte en el artículo de referencia (top-1) para todas las fases subsiguientes del pipeline.

### Fase 2: Extracción de contenidos (*capa XML/HTML de la Semantic Web Layer*)

Con el `pageID` del artículo de referencia, SEMTEST realiza dos llamadas adicionales a la API *MediaWiki*. La primera, con `prop=extracts|links`, recupera el texto completo del artículo en formato plano (sin marcas HTML) y los términos enlazados internamente, que pueden alcanzar hasta 500 entidades del espacio de nombres principal (namespace 0). La segunda, con `action=parse&prop=sections`, extrae la tabla de secciones del artículo (sus encabezados), que actúan como descriptores temáticos implícitos del concepto.

Lo que el estudiante experimenta en esta fase es la materialización del paso descrito en la arquitectura de la Web Semántica: los datos que existían encerrados en el formato HTML de la Web de documentos son liberados y estructurados mediante la API. Los términos enlazados internamente de *Wikipedia* constituyen una primera red de conceptos relacionados —una embrionaria lista de términos relacionados comparable a los TR de un tesoro—, y los encabezados de sección funcionan como un análisis

temático emergente del concepto. La herramienta muestra al estudiante tanto el JSON de respuesta como el código PHP de la llamada, permitiéndole ver que la misma información que la Web 1.0 almacenaba como HTML no estructurado puede recuperarse como datos estructurados mediante la API oficial.

### **Fase 3: Backlinks** (*relaciones documentales pre-semánticas*)

Mediante `prop=linkshere&lhpageid=PAGEID&lhnamespace=0`, SEMTEST recupera los artículos de *Wikipedia* que enlazan hacia el artículo de referencia. En términos documentales, estos *backlinks* son el equivalente de las relaciones de remisión inversa ("Véase también" o "Use en") en un tesoro: delimitan el campo de aplicación y los dominios de uso del concepto identificando qué otros artículos lo mencionan como parte relevante de su contenido. La comparación pedagógica entre los enlaces de *Wikipedia* (relaciones no tipadas, editoriales y contextuales) y las relaciones que se verán en las fases 5 y 6 (relaciones explícitamente tipadas con propiedades RDF como `dbo:birthPlace` o `skos:broader`) es uno de los contrastes conceptuales más fructíferos de la herramienta. El estudiante comprende intuitivamente que los *backlinks* de *Wikipedia* son el correlato no estructurado de lo que en un grafo de conocimiento se expresa mediante propiedades con dominio y rango definidos.

### **Fase 4: Multimedia con Wikimedia Commons** (*extensión contextual*)

SEMTEST consulta la API de Wikimedia Commons con:

`action=query&list=search&srnamespace=6&srsearch=TÉRMINO` para recuperar archivos de imagen relacionados con el concepto. Una segunda llamada con `prop=imageinfo&iiprop=url&iiurlwidth=180` obtiene las URLs de las miniaturas. Los resultados se presentan como una galería visual interactiva donde cada imagen enlaza a su página en Commons, que incluye metadatos de autoría, licencia y categorías.

Esta fase introduce el concepto de recurso no textual identificado mediante URI: en el ecosistema LOD, una imagen no es un simple archivo binario sino una entidad con propiedades descritas en RDF, con su propia URI canónica y con relaciones de categorización que la conectan al grafo de conocimiento de *Wikipedia*. El tipo de licencia (*Creative Commons*) también es relevante, ya que cada licencia tiene su propia URI en el vocabulario de Creative Commons (<https://creativecommons.org/ns>), un ejemplo concreto de cómo los metadatos de derechos se expresan mediante vocabularios formales.

### **Fase 5: DBpedia y clasificación RDFS** (*capa de datos RDF de la Semantic Web Layer e información/esquema*)

Esta es la fase más rediseñada y pedagógicamente más rica de SEMTEST. Opera en dos pasos. El primero utiliza el *DBpedia Lookup Service*:

(<https://lookup.dbpedia.org/api/search>) para resolver el URI canónico del recurso en el grafo *DBpedia*, manejando así ambigüedades, redirecciones y diferencias ortográficas entre el término de la consulta y la denominación exacta usada en *DBpedia*. El segundo lanza una consulta SPARQL al endpoint para recuperar las tripletas del recurso.

*DBpedia* es, en términos técnicos, la serialización en RDF de la información enciclopédica de *Wikipedia*: extrae las propiedades de las cajas de información (*infoboxes*) y las categorías de los artículos y las publica como tripletas RDF (Lehmann et al., 2015). La *DBpedia Ontology*, con más de 320 clases y 1.650 propiedades, proporciona el esquema mediante el que estas tripletas se organizan. Esto hace de *DBpedia* un grafo de conocimiento especialmente adecuado para la enseñanza: el estudiante reconoce la información porque la ha visto en *Wikipedia*, pero ahora la ve estructurada formalmente.

La novedad central de SEMTEST en esta fase es la **clasificación automática de las propiedades RDF recuperadas en tres grupos visuales diferenciados**:

- *Propiedades de datatype (owl:DatatypeProperty)*: literales tipados como `dbo:birthDate` ("1881-10-25"^^xsd:date), `dbo:activeYears` (85), o `rdfs:label` ("Pablo Picasso"@es). El sistema muestra al estudiante la diferencia entre un literal simple (una cadena de texto) y un literal tipado (un valor con su tipo de dato XML Schema), concretando el concepto de tipado semántico que permite a los agentes software razonar sobre los datos.
- *Propiedades de objeto (owl:ObjectProperty)*: propiedades como `dbo:birthPlace` que apuntan al URI `<http://dbpedia.org/resource/Málaga>`, o `dct:subject` que apunta a: `<http://dbpedia.org/resource/Category:Spanish_painters>`. Estas propiedades crean la red de datos propiamente dicha: cada URI objeto es a su vez un nodo del grafo con sus propias propiedades y relaciones.
- *Relaciones jerárquicas y de tipo (rdf:type, rdfs:subClassOf)*: la propiedad `rdf:type` indica que el recurso es una instancia de una clase ontológica (por ejemplo, `dbo:Artist`, `dbo:Person`). SEMTEST intenta reconstruir la jerarquía de superclases, mostrando que `dbo:Artist rdfs:subClassOf dbo:Person` y `rdfs:subClassOf owl:Thing`, lo que permite ilustrar el mecanismo de inferencia: si una entidad es de tipo `dbo:Artist`, se puede inferir automáticamente que también es de tipo `dbo:Person`.

La interfaz muestra simultáneamente el URI canónico del recurso (`http://dbpedia.org/resource/Pablo_Picasso`) y la URL de página (`https://dbpedia.org/page/Pablo_Picasso`), concretando la distinción fundamental entre identificación y localización que Berners-Lee (2006) establece en el segundo principio de los datos enlazados. Esta distinción, detallada en el estándar RFC 3986 de la IETF, es abstracta en su formulación teórica, pero resulta inmediatamente comprensible cuando el estudiante observa las dos URLs en paralelo y entiende que una identifica al concepto Pablo Picasso y la otra localiza un documento HTML sobre él.

#### **Fase 6: Wikidata, SKOS y OWL (capa de conocimiento de la Semantic Web Layer)**

*Wikidata* es la base de conocimiento estructurado de los proyectos Wikimedia, mantenida activamente por una comunidad global y no derivada automáticamente de *Wikipedia* (Vrandečić; Krötzsch, 2014). A diferencia de *DBpedia*, que es un espejo estructurado de *Wikipedia*, *Wikidata* tiene su propio modelo de datos con una semántica más

rica: sus declaraciones (*statements*) incluye referencias bibliográficas, rangos temporales de validez y valores alternativos, representando así la pluralidad de perspectivas sobre un mismo hecho.

SEMTEST utiliza [wbsearchentities](#) para localizar el identificador Q del concepto (por ejemplo, Q5593 para Pablo Picasso) y [wbgetentities&props=claims|labels](#) para recuperar todas sus declaraciones. Una resolución en lote de etiquetas en español e inglés convierte los identificadores Q de las entidades relacionadas en nombres legibles para el usuario.

La novedad pedagógica de esta fase es la **agrupación de las declaraciones por tipo de relación semántica**:

- Propiedades de clasificación ontológica (P31 *instance of*, P279 *subclass of*): ilustran la jerarquía de clases en OWL.
- Propiedades mereológicas (P361 *part of*, P527 *has part*): muestran las relaciones todo-parte que el tesoro expresa como relaciones jerárquicas partitivas.
- Propiedades de identidad externa ([owl:sameAs](#) implícito): *Wikidata* incluye identificadores URNs externos como VIAF (P214), ISNI (P213) u ORCID (P496), mostrando cómo el principio de reconciliación de identidades se implementa en el ecosistema LOD mediante identificadores persistentes (**Herb et al.**, 2023; **McCloud**, 2015).
- Propiedades biográficas y relacionales: personas, instituciones, obras y conceptos relacionados.

La comparación entre P279 (*subclass of*) de *Wikidata* y [rdfs:subClassOf](#) de RDFS hace visible que lo que en un tesoro son relaciones TG/TE (término genérico/término específico), normalizadas en la norma UNE-ISO 25964, se convierte en una ontología formal en una declaración procesable por un razonador OWL (*W3C*, 2012). Esta equivalencia conceptual –entre el oficio manual del tesauro y la declaración formal del ontólogo– es uno de los efectos pedagógicos más relevantes de SEMTEST, y conecta directamente con la argumentación de **Méndez y Greenberg** (2012) sobre el potencial de los profesionales de la Documentación como mediadores entre Sistemas de Organización del Conocimiento y grafos de conocimiento.

La fase 6 incluye además un visor del *iframe* del *Wikidata* Query Service, que ejecuta una consulta SPARQL visual sobre las relaciones de la entidad, proporcionando al estudiante una primera experiencia directa del lenguaje de interrogación antes de que la fase 7 lo aborde en profundidad.

**Fase 7: Consulta SPARQL directa** (*capa de interrogación de la Semantic Web Layer*)

Esta fase ejecuta una consulta SPARQL [SELECT](#) parametrizada directamente contra <https://dbpedia.org/sparql>. La consulta recupera hasta 60 tripletas del recurso consultado, filtradas por idioma (español, inglés y literales sin etiqueta) y por tipo de valor (literales y URIs). Los resultados se presentan en una tabla estructurada de tres columnas: propiedad (nombre abreviado con tooltip al URI completo), valor (enlazado cuando es URI) y tipo de dato. Puede comprobarse una muestra de ello en la tabla 2.

Tabla 2. Estructura de la consulta SPARQL generada por SEMTEST

```
PREFIX dbr: <http://dbpedia.org/resource/>
PREFIX dbo: <http://dbpedia.org/ontology/>
SELECT ?predicate ?object
WHERE {
  <URI_RECURSO> ?predicate ?object .
  FILTER (isLiteral(?object) && langMatches(lang(?object), "es"))
  UNION
  { <URI_RECURSO> ?predicate ?object . FILTER (isURI(?object)) }
}
LIMIT 60
```

La herramienta muestra la query completa al estudiante, con sus cláusulas PREFIX, SELECT, WHERE, FILTER y LIMIT comentadas, permitiéndole asociar cada elemento con la sintaxis SPARQL 1.1 documentada por el W3C (2008; 2013a). El aspecto técnico pedagógicamente más valioso es la cabecera HTTP `Accept: application/sparql-results+json`, que es diferente del tipo MIME genérico `application/json`: SEMTEST muestra explícitamente esta cabecera en la interfaz, ilustrando que el protocolo SPARQL tiene su propio tipo de respuesta estandarizado y que una llamada sin esa cabecera devuelve un error o una respuesta inesperada.

El contraste entre la Fase 5 (acceso HTML al recurso *DBpedia*) y la Fase 7 (acceso RDF/SPARQL al mismo recurso) materializa en pantalla la diferencia conceptual entre la Web de documentos y la Web de datos: el mismo recurso Pablo Picasso existe como página HTML legible por humanos y como conjunto de tripletas RDF consultables por máquinas mediante SPARQL. Los cuatro principios de Linked Data de **Berners-Lee** (2006) —uso de URIs, uso de HTTP URIs resolubles, provisión de información útil mediante RDF y SPARQL, e inclusión de enlaces a otras URIs— quedan ilustrados en las fases 1 a 7 del pipeline.

### **Fase 8: *Open Library* y enriquecimiento bibliográfico (vocabularios bibliográficos)**

La consulta a la API de *Open Library* se articula con las variables “q” y “limit”, tal como sigue (<https://openlibrary.org/search.json?q=TÉRMINO&limit=12>) recuperándose hasta doce obras relacionadas con el concepto consultado. Para cada obra se recupera título, autores, año de primera publicación y portada (desde <https://covers.openlibrary.org>). Los resultados se presentan en una galería de tarjetas bibliográficas visuales, cada una enlazada a su página completa en *Open Library*.

La novedad de SEMTEST en esta fase es la **capa de correspondencia semántica**: la interfaz muestra al estudiante cómo cada campo del resultado JSON podría serializarse en RDF utilizando vocabularios bibliotecarios estándar. El título de una obra corresponde a `dc:title` o `bibo:title`; el autor, a `dc:creator` con su URI VIAF; el año de publicación, a `dc:date`; el tipo de documento (libro, artículo, etc.), a una subclase de `bibo:Document`.

La *Bibliographic Ontology* (*BIBO*, disponible en <http://purl.org/ontology/bibo>) permite describir cualquier tipo de aspecto bibliográfico en RDF, y los datos planos de *Open*

*Library* son exactamente el tipo de información que *BIBO* fue diseñado para estructurar formalmente, de acuerdo a las especificaciones oficiales del *DCMI* (2015). Esta fase establece el puente que conecta el enriquecimiento semántico automatizado de las fases anteriores con los instrumentos clásicos de la gestión bibliográfica. El catálogo bibliográfico —que el documentalista conoce como el espacio de los registros MARC— tiene su equivalente en el ecosistema LOD en forma de repositorios de datos RDF que implementan vocabularios como *BIBO*, *Dublin Core* o el modelo IFLA LRM (Riva et al., 2017). *Open Library*, con más de 20 millones de registros accesibles mediante API abierta, es un punto de acceso privilegiado a esa dimensión bibliográfica del grafo de conocimiento.

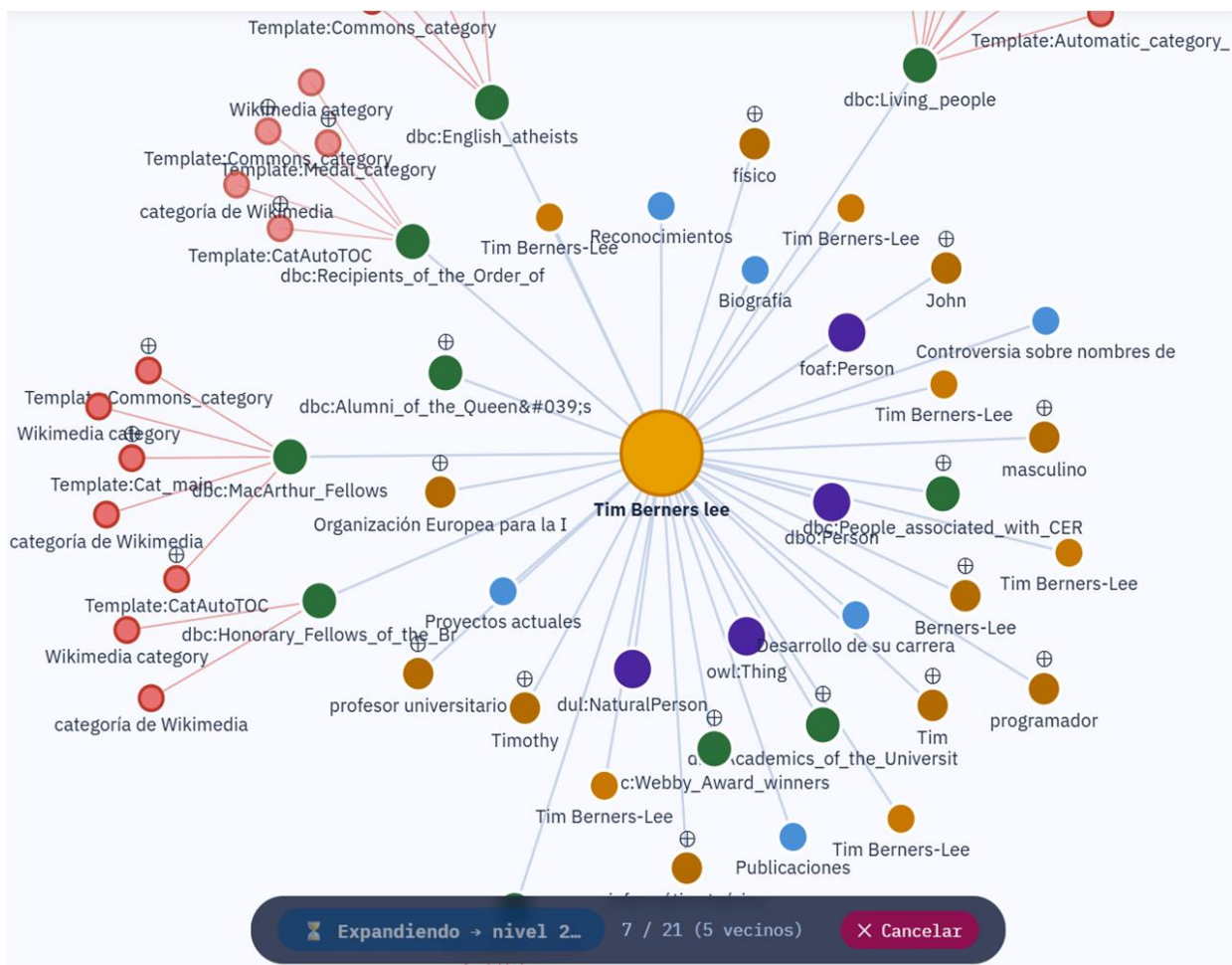


Figura 1. Grafo semántico multicapa generado por SEMTEST, mostrando la diferenciación de nodos por capa (RDFS, SKOS, Instancias y capa Bibliográfica)

### Fase 9: Grafo semántico multicapa y exportación (*síntesis y representación formal*)

La fase final sintetiza todos los resultados anteriores en dos productos integrados: un grafo interactivo y una exportación en cuatro formatos estándar. El grafo de SEMTEST está construido con D3.js v7 utilizando un layout de fuerza dirigida (*force-directed layout*) modificado para implementar un comportamiento jerárquico por capas. A diferencia de la versión original, que generaba una red plana de nodos sin diferenciación semántica, SEMTEST organiza los nodos en **cinco capas semánticas diferenciadas** por color, tamaño y posición (véase la Figura 1):

1. **Capa central** (nodo amarillo grande): la entidad consultada con su URI canónica. Es el sujeto de todas las tripletas recuperadas en las fases anteriores.
2. **Capa RDFS** (nodos azul oscuro): nodos de clase (`rdf:type`) y superclases inferidas (`rdfs:subClassOf`). Representan la posición de la entidad en la jerarquía taxonómica de la *DBpedia* Ontology.
3. **Capa SKOS** (nodos verdes): categorías y materias (`dct:subject`) conectadas mediante relaciones `skos:broader` cuando se detectan en el esquema de categorías de *Wikipedia*. Representan la posición de la entidad en el vocabulario controlado categorial.
4. **Capa de instancias** (nodos azul claro): entidades relacionadas vía `dbo:wikiPageWikiLink`, propiedades de objeto de *DBpedia* o propiedades significativas de *Wikidata*.
5. **Capa bibliográfica** (nodos naranjas): obras de *Open Library* conectadas mediante `dc:subject` al concepto consultado.

Las relaciones inferidas a partir de `rdfs:subClassOf` se representan con líneas discontinuas de color distinto a las relaciones explícitas, ilustrando visualmente la diferencia entre datos afirmados y datos inferidos —la base del razonamiento ontológico en OWL. El grafo es completamente interactivo: admite *zoom*, desplazamiento (*pan*) y arrastre de nodos individuales. Al hacer clic sobre cualquier nodo se abre la URL del recurso correspondiente en el ecosistema LOD.

Cuando se llega al punto que permite realizar la exportación, ésta incluye cuatro formatos:

- **JSON-LD**: formato de datos enlazados basado en JSON con contexto explícito que referencia vocabularios Schema.org, *DBpedia* y *Wikidata*. Permite incrustar los metadatos del grafo en páginas web como datos estructurados legibles por buscadores.
- **Turtle (.ttl)**: serialización RDF con prefijos estándar (`@prefix rdf;`, `@prefix rdfs;`, `@prefix skos;`, `@prefix owl;`, `@prefix dct;`, `@prefix bibo;`, `@prefix foaf;`). El archivo generado puede cargarse directamente en *Protégé*, *Apache Jena Fuseki* o cualquier triplestore para visualización y consulta posterior.
- **D3-JSON**: la estructura de nodos y aristas del grafo en formato JSON, reutilizable para otros proyectos de visualización.
- **SVG**: la imagen vectorial del grafo tal como aparece en pantalla, descargable para publicación académica o presentación.

La posibilidad de exportar en *Turtle* y cargarlo en *Protégé* cierra el ciclo formativo de SEMTEST: el estudiante puede partir de una consulta libre en lenguaje natural, seguir paso a paso cómo esa consulta se enriquece semánticamente a través de nueve fases, ver el grafo multinivel resultante y finalmente importar ese grafo en una herramienta de gestión ontológica para explorar sus propiedades e inferencias. El proceso completo ilustra la arquitectura de la Web Semántica de extremo a extremo.

### 3. Resultados y aplicación didáctica

#### 3.1. Qué puede ver el estudiantado

Desde el punto de vista pedagógico, SEMTEST no solo describe las tecnologías semánticas: las hace **observables**. Describimos a continuación los principales fenómenos que el estudiante puede observar directamente mediante la herramienta.

**La diferencia entre la Web de documentos y la Web de datos.** Al ejecutar la misma búsqueda en la Fase 1 (*Wikipedia*, respuesta HTML estructurada como JSON) y en la Fase 7 (*DBpedia* SPARQL, respuesta de tripletas RDF), el estudiante puede comparar dos modos radicalmente distintos de representar la misma información sobre una entidad: la Web de documentos proporciona texto legible por humanos, mientras que la Web de datos proporciona declaraciones formales procesables por máquinas. Esta distinción, que en la literatura teórica requiere definiciones abstractas sobre la diferencia entre datos, información y conocimiento (**Bizer et al.**, 2009; **Hogan et al.**, 2021), se convierte en una experiencia directa cuando el estudiante ve la misma entidad representada de ambas formas en la misma sesión de consulta.

**El continuo KOS en funcionamiento real.** Al pasar de las categorías *DBpedia* (Fase 5) a las propiedades P279 de *Wikidata* (Fase 6) y finalmente al grafo multicapa con nodos SKOS (Fase 9), el estudiante puede seguir en tiempo real la progresión desde una taxonomía (categorías de *Wikipedia* como estructura jerárquica simple) hacia un tesoro conceptual (relaciones SKOS con relaciones jerárquicas, asociativas y de equivalencia) y hacia una ontología formal OWL (clases con inferencia automática mediante `rdfs:subClassOf`). Este continuo, que **Méndez y Greenberg** (2012) describen como el eje de articulación entre los KOS bibliotecarios y los vocabularios LOD, se materializa visualmente en el grafo de cinco capas.

**La distinción URI/URL en contexto operativo.** En las Fases 5 y 7, SEMTEST muestra en paralelo el URI canónico [http://dbpedia.org/resource/Pablo\\_Picasso](http://dbpedia.org/resource/Pablo_Picasso) y la URL de página [https://dbpedia.org/page/Pablo\\_Picasso](https://dbpedia.org/page/Pablo_Picasso). El estudiante puede comprobar que al hacer clic sobre el URI en un navegador, el servidor *DBpedia* devuelve un redireccionamiento 303 hacia la URL de página: este comportamiento, conocido como *content negotiation* o negociación de contenido, es precisamente el mecanismo de desreferenciación descrito por **Berners-Lee** (2006). Ver el proceso en acción —no solo leerlo— elimina la ambigüedad conceptual entre identificar y localizar, entre nombrar una cosa y acceder a información sobre ella.

**La sintaxis SPARQL sobre datos reales.** La Fase 7 muestra al estudiante una consulta SPARQL real, ejecutada contra un endpoint público, con resultados visibles. Puede modificar la consulta cambiando el URI del sujeto por otro recurso *DBpedia* y ver cómo los resultados cambian. Esta posibilidad de experimentación directa —sin necesidad de instalar ningún software, directamente desde el navegador— reduce drásticamente la barrera de entrada al aprendizaje de SPARQL, que de otro modo requeriría un entorno configurado y un conjunto de datos cargado en local.

**Los identificadores de autoridad como datos enlazados.** En la Fase 6, *Wikidata* devuelve las propiedades P214 (VIAF ID), P213 (ISNI) y P496 (ORCID) de personas, si las tienen asignadas. SEMTEST convierte estos identificadores en enlaces activos hacia

los registros de autoridad correspondientes en VIAF (<http://viaf.org/viaf/ID>) e ISNI (<https://isni.org/isni/ID>). El estudiante ve cómo el control de autoridades bibliotecario —que conoce como práctica documental— se implementa en el ecosistema LOD mediante identificadores persistentes que cumplen la doble cualidad de ser únicos y resolverse como HTTP URIs (Herb et al., 2023; McCloud, 2015). Esta conexión entre la práctica bibliotecaria y el principio 1 y 2 de Berners-Lee (2006) es una de las más valoradas pedagógicamente.

### 3.2. Qué puede experimentar el estudiantado

Además de la posibilidad de observar los resultados, SEMTEST ofrece al estudiante tres modos de experimentación activa. Se detallan en los próximos párrafos:

**Variación del término de consulta.** El estudiante puede consultar términos de muy distinta naturaleza: personas (Cervantes, Marie Curie), lugares (Madrid, Atenas), disciplinas (Biblioteconomía, Arquitectura), obras (Don Quijote, Guernica) o conceptos abstractos (Epistemología, Resiliencia). Al hacerlo, observa cómo cambia la riqueza del grafo resultante dependiendo de la cobertura que *DBpedia* y *Wikidata* tienen del concepto: los recursos muy documentados generan grafos densamente conectados con decenas de propiedades, mientras que los recursos menos documentados producen grafos dispersos. Esta variación ilustra una limitación real del ecosistema LOD: la calidad y completitud de los datos enlazados depende de la comunidad que los mantiene.

Los recursos muy documentados generan grafos densamente conectados con decenas de propiedades, mientras que los recursos menos documentados producen grafos dispersos. Esta variación ilustra una limitación real del ecosistema LOD: la calidad y completitud de los datos enlazados depende de la comunidad que los mantiene

**Inspección del JSON en bruto y el código fuente.** Para cada fase, el estudiante puede desplegar dos paneles adicionales: el JSON de respuesta de la API (sin procesar, exactamente como lo devuelve el servidor) y el código PHP de la función correspondiente. Esta transparencia permite al estudiante verificar por sí mismo qué datos se han recuperado, compararlos con lo que muestra la interfaz y comprender qué transformaciones se han aplicado entre la respuesta cruda y la presentación visual.

**Exportación y carga en *triplestore*.** El estudiante puede descargar el archivo *Turtle* generado por SEMTEST y cargarlo en *Apache Jena Fuseki* o *Protégé*, aplicaciones que se utilizan habitualmente durante la formación en Vocabularios y Esquemas Semánticos. Al cargar el archivo, puede ejecutar sus propias consultas SPARQL sobre el grafo local construido a partir de su consulta, comprobando empíricamente las reglas de inferencia RDFS y la estructura del grafo antes de visualizarlo en el árbol ontológico de *Protégé*.

### 3.3. Qué problemas resuelve

Por otra parte, SEMTEST aborda tres problemas pedagógicos específicos que dificultan la enseñanza de las tecnologías semánticas en los estudios de Documentación y en los que profundizamos.

**El problema de la abstracción.** Los conceptos de URI, tripleta RDF, propiedad de objeto, razonamiento ontológico o consulta SPARQL son difíciles de comprender en abstracto. SEMTEST los hace concretos: el URI de Pablo Picasso en *DBpedia*, la tripleta `<db:Pablo_Picasso> <dbo:birthPlace> <db:Málaga>`, la distinción entre una `owl:DatatypeProperty` y una `owl:ObjectProperty` visibles en la clasificación de la Fase 5. La abstracción se vuelve comprensible cuando el estudiante la ve funcionando sobre entidades que conoce.

**El problema de la caja negra.** La mayoría de las implementaciones de enriquecimiento semántico disponibles comercialmente o como servicio web operan como cajas negras: el usuario no sabe qué APIs se consultan, con qué cabeceras, con qué estructura de consulta ni qué se hace con la respuesta. SEMTEST elimina esta opacidad mostrando cada llamada HTTP con sus cabeceras exactas, cada respuesta JSON en bruto y cada fragmento de código responsable. Esta transparencia es especialmente valiosa en un contexto donde los grandes modelos de lenguaje tienden a invisibilizar aún más las infraestructuras subyacentes del conocimiento estructurado.

#### El problema de la desconexión disciplinar.

Los estudiantes de Documentación conocen los tesauros, el control de autoridades, la catalogación analítica y el análisis de materias, pero frecuentemente no establecen la conexión entre esas prácticas y las tecnologías LOD. SEMTEST construye explícitamente esa conexión: los TG/TE del tesauro son `rdfs:subClassOf` en el grafo; el identificador de autoridad es un URI persistente en *Wikidata*; el encabezamiento de materia es un nodo en la capa SKOS del grafo; el registro bibliográfico es una instancia de `bibo:Book`

con propiedades *Dublin Core*. Esta correspondencia conceptual, que Méndez y Greenberg (2012) consideran el potencial más valioso de los profesionales de la Documentación en el ecosistema LOD, se hace visible en la pantalla del estudiante durante una consulta ordinaria.

SEMTEST representa un tipo de recurso pedagógico que la literatura sobre enseñanza de la Web Semántica identifica como necesario pero escaso: no un tutorial de una tecnología concreta, sino un entorno integrado donde múltiples capas de la arquitectura semántica interactúan sobre datos reales

### 4. Discusión

SEMTEST representa un tipo de recurso pedagógico que la literatura sobre enseñanza de la Web Semántica identifica como necesario pero escaso: no un tutorial de una tecnología concreta, sino un entorno integrado donde múltiples capas de la arquitectura semántica interactúan sobre datos reales. Hogan *et al.* (2021), en su revisión comprensiva de los grafos de conocimiento publicada en *ACM Computing Surveys*, señalan que la comprensión de los grafos de conocimiento requiere familiarizarse si-

multáneamente con modelos de datos, lenguajes de consulta, esquemas de validación y mecanismos de inferencia, una combinación que ninguna herramienta aislada cubre completamente. SEMTEST no cubre todo el espectro descrito por **Hogan et al.** (2021), pero sí ofrece un punto de entrada integrado que relaciona todos esos elementos en el contexto de una consulta real.

La revisión sistemática de **Gaitanou et al.** (2024) sobre datos enlazados en bibliotecas confirma que la mayoría de las publicaciones sobre implementación de LOD en el entorno bibliotecario se centran en aspectos técnicos de publicación de datos —migración de MARC a BIBFRAME, generación de RDF, configuración de *triplestores*— con menor atención a los aspectos formativos. SEMTEST complementa esa literatura aportando un demostrador orientado explícitamente a la formación, donde el objetivo no es publicar datos sino comprenderlos.

Desde el punto de vista de las limitaciones, la herramienta depende de la disponibilidad y estabilidad de APIs externas sobre infraestructuras vivas. El endpoint SPARQL de *DBpedia* tiene una disponibilidad variable y puede devolver errores de *timeout* en períodos de alta carga. Las APIs de *Wikidata* imponen límites de tasa de petición que SEMTEST gestiona mediante reintentos y *caching* de resultados. *Open Library* actualiza periódicamente su API y puede cambiar la estructura de sus respuestas JSON. Estas dependencias son, simultáneamente, una limitación técnica y un recurso pedagógico: el estudiante aprende que la Web Semántica es un ecosistema dinámico donde la robustez de las aplicaciones depende de la calidad de implementación de las peticiones HTTP —de ahí la importancia de los *timeouts* ampliados, las cabeceras correctas y los gestores de error que SEMTEST implementa de forma visible.

Como líneas de trabajo futuro, se plantean dos extensiones principales. La primera es la incorporación de un módulo de razonamiento básico en el lado del cliente mediante la librería *rdflib.js*, que permita al estudiante ejecutar consultas SPARQL de forma interactiva sobre el grafo local generado y comprobar empíricamente las reglas de inferencia RDFS (*rdfs:subClassOf*, *rdfs:domain*, *rdfs:range*). La segunda es la integración de un módulo de alineación con el tesoro de la UNESCO (disponible como LOD en <http://vocabularies.unesco.org/thesaurus>) que permita contrastar las categorías de *DBpedia* y los descriptores de *Wikidata* con los términos del tesoro institucional, materializando en pantalla la equivalencia entre las relaciones *skos:broader/skos:narrower* del tesoro y las jerarquías ontológicas del grafo de conocimiento.

La mayoría de las publicaciones sobre implementación de LOD en el entorno bibliotecario se centran en aspectos técnicos de publicación de datos —migración de MARC a BIBFRAME, generación de RDF, configuración de *triplestores*— con menor atención a los aspectos formativos. SEMTEST complementa esa literatura aportando un demostrador orientado explícitamente a la formación, donde el objetivo no es publicar datos sino comprenderlos

## 5. Conclusiones

SEMTEST demuestra que los procesos de análisis documental y enriquecimiento conceptual que los profesionales de la información realizan manualmente mediante tesauros, listas de encabezamientos de materia y registros de autoridad tienen su correlato automatizado y formalizado en la Web Semántica: las categorías *DBpedia* son taxonomías expresadas en RDFS; las propiedades *Wikidata* son relaciones de tesoro expresadas en SKOS y OWL; los registros de *Open Library* son instancias bibliográficas describibles con *BIBO* y *Dublin Core*; los identificadores VIAF e ISNI son URIs persistentes que implementan el control de autoridades en el ecosistema LOD. La diferencia no es conceptual sino de escala, velocidad y formalización.

SEMTEST demuestra que los procesos de análisis documental y enriquecimiento conceptual que los profesionales de la información realizan manualmente mediante tesauros, listas de encabezamientos de materia y registros de autoridad tienen su correlato automatizado y formalizado en la Web Semántica

El carácter transparente de SEMTEST —que muestra cabeceras HTTP, URLs exactas, JSON en bruto y código fuente en cada paso— elimina el efecto caja negra y sitúa al estudiante en condiciones de evaluar críticamente los servicios LOD disponibles: qué devuelven, cómo se accede a ellos, cuáles son sus limitaciones técnicas y cómo se articulan en la arquitectura de la Web Semántica. Esta capacidad de evaluación crítica es una competencia de especial relevancia para el futuro profesional de la información, particularmente en el contexto actual donde la proliferación de sistemas de inteligencia artificial tiende a invisibilizar las infraestructuras subyacentes del conocimiento estructurado.

La capacidad de exportar el grafo generado en formatos estándar como *Turtle* para su ingestión en herramientas de gestión ontológica (ej. *Protégé*), culmina el ciclo pedagógico. Esta funcionalidad permite a los estudiantes pasar de la visualización a la experimentación activa con reglas de inferencia RDFS y SKOS sobre un *dataset* localmente construido, cerrando la brecha entre la arquitectura teórica de la Web Semántica y las herramientas prácticas de la gestión ontológica.

A nivel técnico, la dependencia de la herramienta respecto a la estabilidad y los *timeouts* de las APIs externas de LOD (*DBpedia*, *Wikidata*) se transforma en un valioso recurso didáctico. El estudiante no solo aprende sobre el modelo de datos RDF, sino también sobre la dinámica operativa real del ecosistema de datos abiertos enlazados, incluyendo sus limitaciones de robustez e interoperabilidad en un entorno vivo.

Como se avanzaba en el primer epígrafe, de nuevo, SEMTEST está disponible en acceso abierto en <https://mblazquez.es/lab/semTest2>

## 6. Referencias

**Balloni, Antonio-José; De-Souza-Bermejo, Paulo-Henrique; Holm, Jeanne; Tonelli, Adriano O.** (2012). Governance, sociotechnical systems and knowledge society: challenges and reflections. In: *Organizational Integration of Enterprise Systems and Resources: Advancements and Applications*, pp. 22-41. IGI Global.

**Berners-Lee, Tim** (1998). *Semantic Web Road map*. W3C.  
<https://www.w3.org/DesignIssues/Semantic.html>

**Berners-Lee, Tim** (2006). *Linked Data: Design Issues*. W3C.  
<https://www.w3.org/DesignIssues/LinkedData.html>

**Berners-Lee, Tim; Miller, Eric** (2002). The semantic Web lifts off. *ERCIM News*, n. 51, pp. 9-11.  
[https://www.ercim.eu/publication/Ercim\\_News/enw51/berners-lee.html](https://www.ercim.eu/publication/Ercim_News/enw51/berners-lee.html)

**Bizer, Christian; Heath, Tom; Berners-Lee, Tim** (2009). Linked Data: The Story So Far. *International Journal on Semantic Web and Information Systems*, v. 5, n. 3.  
<https://doi.org/10.4018/jswis.2009081901>

**Codina, Lluís; Pedraza-Jiménez, Rafael** (2017). *Taxonomías y ontologías: Qué son y cómo se aplican en medios de comunicación*.  
<https://www.lluiscodina.com/taxonomias-ontologias>

**DCMI; D'Arcus, Bruce; Giasson, Frédérick** (2015). *Bibliographic Ontology Specification (BIBO)*.  
<https://www.dublincore.org/specifications/bibo/bibo>

**Gaitanou, Panorea; Andreou, Ioanna; Sicilia, Miguel-Angel; Garoufallou, Emmanouel** (2024). Linked data for libraries: Creating a global knowledge space, a systematic literature review. *Journal of Information Science*, v. 50, n. 1, pp. 204-244.  
<https://doi.org/10.1177/01655515221084645>

**Herb, Ulrich; De-Castro, Pablo; Rothfritz, Laura; Schöpfel, Joachim** (2023). Building the plane as we fly it: the promise of Persistent Identifiers. Zenodo.  
<https://doi.org/10.5281/zenodo.7258286>

**Hogan, Aidan; Blomqvist, Eva; Cochez, Michael; D'Amato, Claudia; De-Melo, Gerard; Gutierrez, Claudio; Kirrane, Sabrina; Labra-Gayo, José-Emilio; Navigli, Roberto; Neumaier, Sebastian; Ngonga-Ngomo, Axel-Cyrille; Polleres, Axel; Rashid, Sabbir M.; Rula, Anisa; Schmelzeisen, Lukas; Sequeda, Juan; Staab, Steffen; Zimmermann, Antoine** (2021). Knowledge graphs. *ACM Computing Surveys*, v. 54, n. 4, art. 71.  
<https://doi.org/10.1145/3447772>

**Lehmann, Jens; Isele, Robert; Jakob, Max; Jentzsch, Anja; Kontokostas, Dimitris; Mendes, Pablo N.; Hellmann, Sebastian; Morsey, Mohamed; Van-Kleef, Patrick; Auer, Sören; Bizer, Christian** (2015). DBpedia: A large-scale, multilingual knowledge base extracted from Wikipedia. *Semantic Web*, v. 6, n. 2, pp. 167-195.

<https://doi.org/10.3233/SW-140134>

Library of Congress (2012). *Bibliographic Framework Initiative (BIBFRAME)*.

<https://www.loc.gov/bibframe>

**McCloud, Joanna M.** (2015). *Persistent Identifiers: Weighing the Benefits of In-house Systems versus External Registries*. Johns Hopkins University.

<http://jhir.library.jhu.edu/handle/1774.2/59082>

**Méndez, Eva; Greenberg, Jane** (2012). Linked data for open vocabularies and HIVE's global framework / Datos enlazados para vocabularios abiertos: marco global de HIVE. *El Profesional de la Información*, v. 21, n. 3, pp. 236-244.

<https://doi.org/10.3145/epi.2012.may.03>

**O'Reilly, Tim** (2007). *What is Web 2.0: Design patterns and business models for the next generation of software*. MPRA Paper 4578, University Library of Munich.

**Riva, Pat; Le-Bœuf, Patrick; Žumer, Maja** (2017). *Modelo de referencia bibliotecaria de la IFLA (IFLA LRM)*. International Federation of Library Associations and Institutions.

[https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017\\_rev201712-es.pdf](https://www.ifla.org/wp-content/uploads/2019/05/assets/cataloguing/frbr-lrm/ifla-lrm-august-2017_rev201712-es.pdf)

**Vrandečić, Denny; Krötzsch, Markus** (2014). Wikidata: A free collaborative knowledge base. *Communications of the ACM*, v. 57, n. 10, pp. 78-85.

<https://doi.org/10.1145/2629489>

W3C (2008). *SPARQL Query Language for RDF*. W3C Recommendation.

<https://www.w3.org/TR/rdf-sparql-query>

W3C (2012). *OWL 2 Web Ontology Language: Document Overview (Second Edition)*. W3C Recommendation.

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

W3C (2013a). *SPARQL 1.1 Query Language*. W3C Recommendation.

<https://www.w3.org/TR/sparql11-query>

W3C (2013b). *SPARQL 1.1 Overview*. W3C Recommendation.

<https://www.w3.org/TR/sparql11-overview>