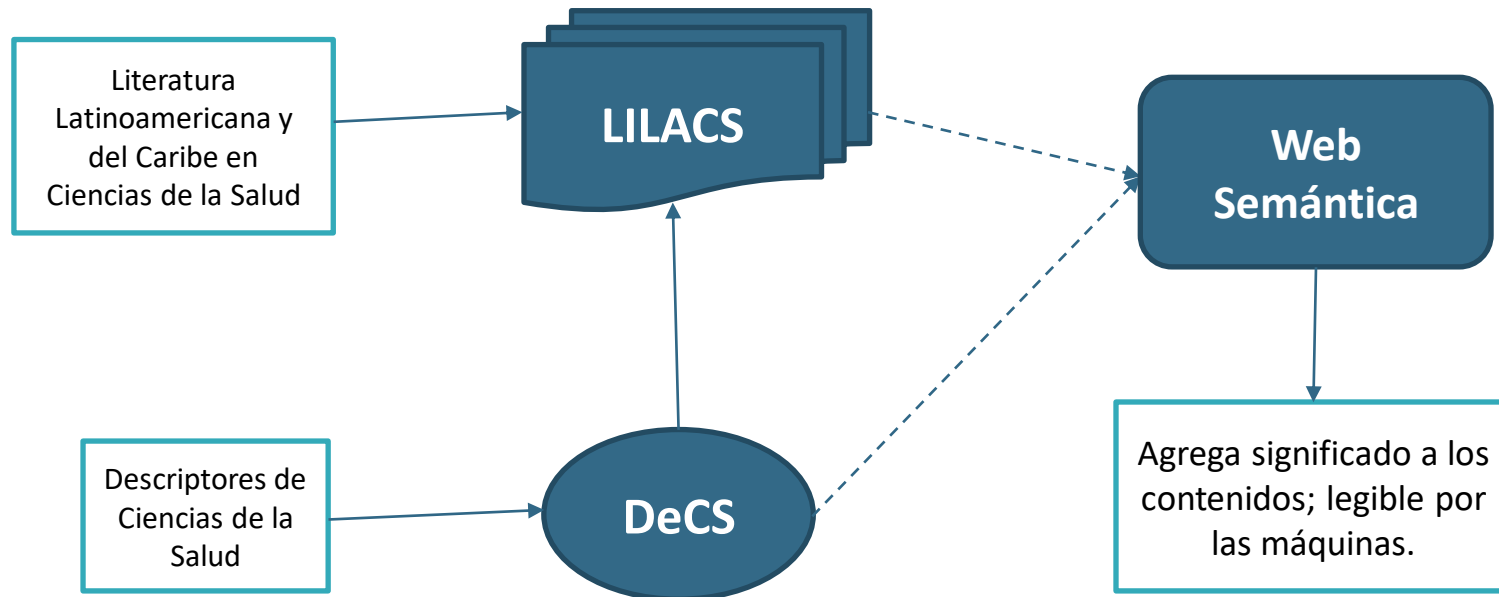




**CENTRO NACIONAL DE
INFORMACIÓN DE CIENCIAS
MÉDICAS**

LILACS-DeCS y Web Semántica

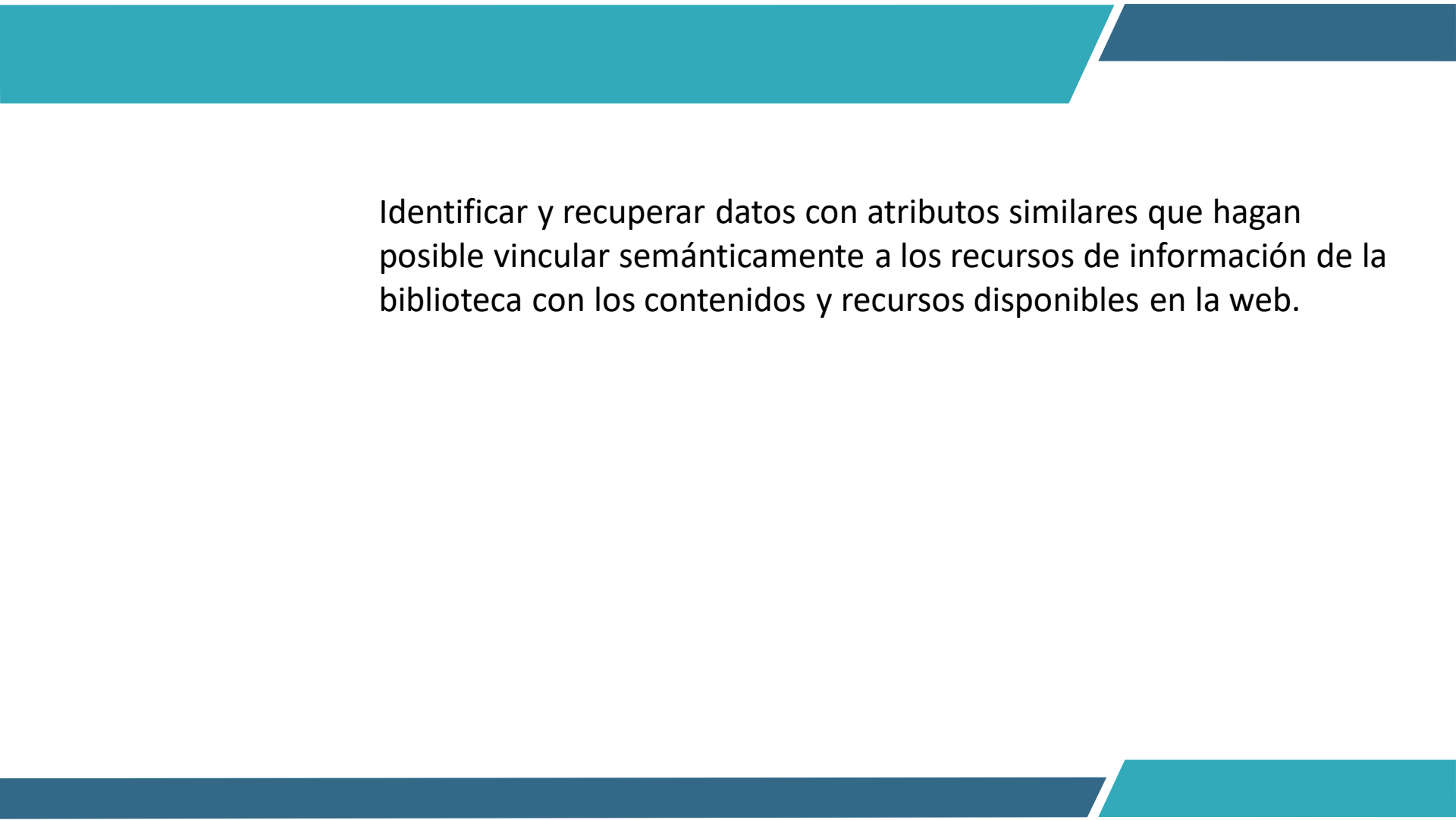
Introducción del Proyecto



Proyecto - Objetivo



- LILACS y DeCS están contruidos con la metodología LILACS como estándar, pero esta representación del conocimiento no puede ser reconocida por las máquinas y estos recursos de información quedan relativamente aislados.
- El proyecto propone:
 - Representar LILACS y DeCS con estándares de la Web Semántica
 - Publicar sus datos como datos enlazados.
- Con el objetivo de
 - Que puedan ser recolectados y procesados por aplicaciones,
 - Fomentar su conexión con otros conjuntos de datos,
 - Aumentar su presencia y relevancia en Internet,
 - Vincular semánticamente los recursos de LILACS con los contenidos y recursos disponibles en la web



Identificar y recuperar datos con atributos similares que hagan posible vincular semánticamente a los recursos de información de la biblioteca con los contenidos y recursos disponibles en la web.

Contexto



- Web semántica y datos enlazados
- Datos Bibliográficos
- Nube de Datos Abiertos Enlazados
- Estándares de la Web Semántica
- Ventajas y Desventajas

Web Semántica, datos enlazados

- Web semántica y datos enlazados son estándares del consorcio WWW
- La Web semántica agrega una **capa de datos con significado y de relaciones** entre los contenidos de la Web
- Los datos pueden ser consumidos no solo por los humanos sino también por las máquinas
- Se pueden procesar automáticamente, por ejemplo con inteligencia artificial.
- Datos enlazados es la forma de representar la información en la red semántica, con estándares, de manera descentralizada

Web tradicional




La máquina entiende la capa de presentación: HTML, CSS, Java Script



El usuario entiende el contenido

Web Semántica

Los datos estructurados, legibles para las máquinas, permiten a **buscadores basados en texto** como Google, Bing o Yahoo evaluar los datos semánticos, hacer **búsquedas mas precisas** y darles **diferentes formas de presentación**.

carbonara 

[Todo](#) [Imágenes](#) [Vídeos](#) [Maps](#) [Noticias](#) [Más ▾](#) [Herramientas de búsqueda](#)


Aproximadamente 64.200.000 resultados (0,64 segundos)

Salsa Carbonara: Consejos, preparación y recetas variadas
[www.salsacarbonara.net/ ▾](http://www.salsacarbonara.net/)
22 sept. 2015 - Salsa Carbonara: Te enseñamos a preparar una estupenda Salsa Carbonara con la amplia variedad de recetas que te mostramos en nuestra ...
Pasta a la Carbonara - Salsa Carbonara Tradicional - Salsa Carbonara con Huevo

Pasta a la Carbonara - Salsa Carbonara
 [!\[\]\(c029974817362751c5ce461e263b8135_img.jpg\)
★★★★★ Valoración: 3,3 - 212 reseñas
Karlos Arguiñano prepara la famosa y típica receta de espaguetis con salsa carbonara, un plato sencillo de reproducir y delicioso.
Espaguetis a la carbonara - Pasta casera a la carbonara](http://www.salsacarbonara.net/SalsaCarbonara/PastaALaCarbonara/PastaALaCa... ▾
Pasta a la Carbonara: Te describimos paso a paso la receta para cocinar ...</p><p>Receta de Espaguetis con salsa carbonara - Karlos Arguiñano
<a href=)

Carbonara - Wikipedia, la enciclopedia libre
<https://es.wikipedia.org/wiki/Carbonara ▾>
Carbonara es un plato de pasta italiano originario del Lacio, y más concretamente de Roma. La receta original se basa en huevos, queso (Pecorino Romano o ...
Historia - Características - Variantes - Véase también

Salsa Carbonara
[salsacarbonara.org/ ▾](http://salsacarbonara.org/)
Blog de recetas con salsa carbonara o variantes, con imágenes y explicaciones paso a paso.



Más imágenes

Carbonara






Plato

Carbonara es un plato de pasta italiano originario del Lacio, y más concretamente de Roma. La receta original se basa en huevos, queso, aceite de oliva virgen extra, pancetta o guanciale y pimienta negra.
Wikipedia

Origen: Italia

Otras personas también buscan

Ver otros 15

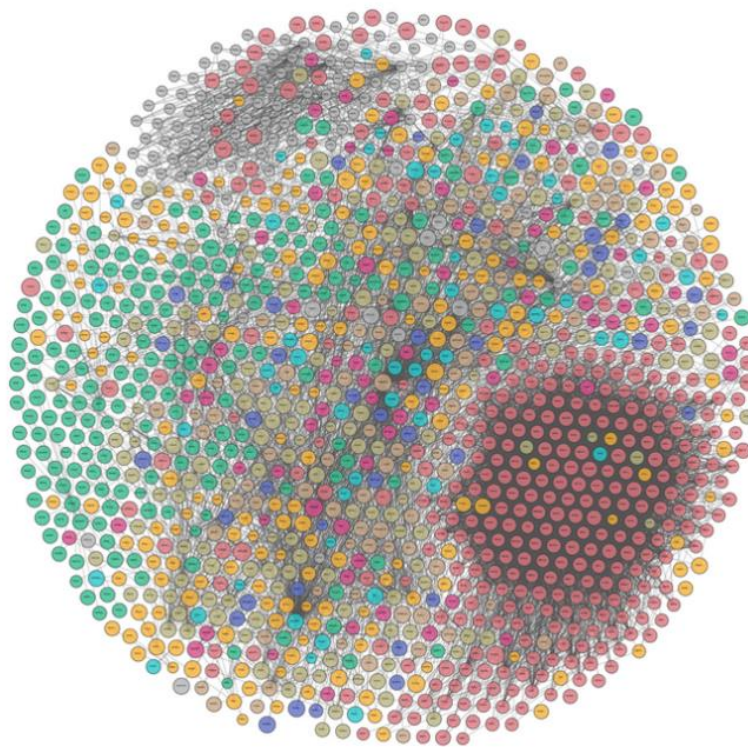
				
Espagueti	Boloñesa	Pancetta	Amatriciana	Queso parmesano

Denos tu opinión

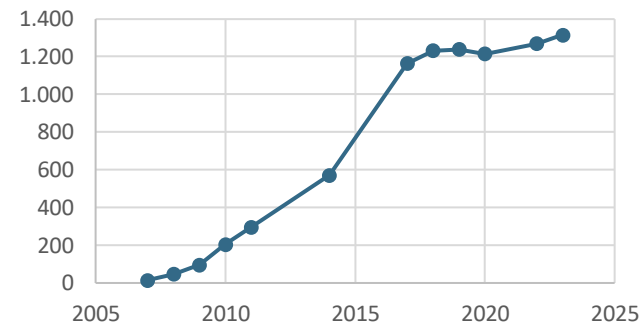
Datos Bibliográficos

- El aporte de los datos bibliográficos a la Web semántica se considera como un componente esencial, por la calidad y el tratamiento profesional de los datos
- Las bibliotecas se han distinguido por la producción sistemática y organizada de contenidos
- Han desarrollado estándares y metodologías para describir y representar sus colecciones, pero estos recursos de información quedan relativamente aislados.
- **El objetivo de representar los datos bibliográficos en forma de datos enlazados es Incrementar su impacto en el espacio global de la Web.**

Nube de Datos Abiertos Enlazados



1,314 Datasets



Datos Bibliográficos:datos enlazados

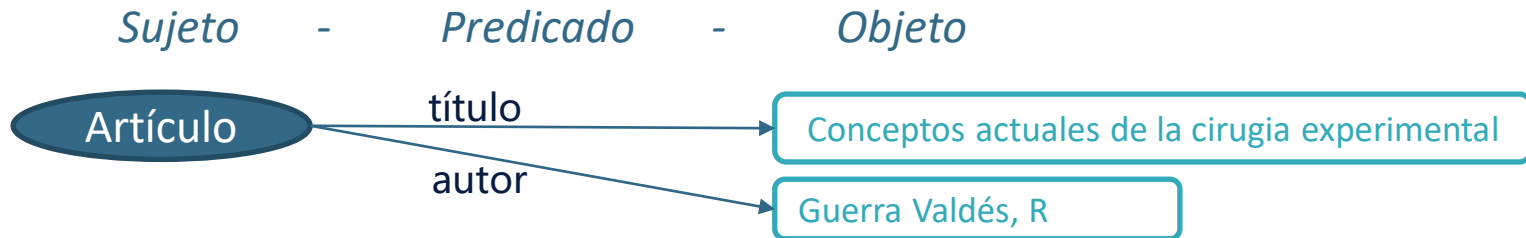
- 1.WorldCat:** Catálogo global de colecciones de bibliotecas. Publicado por OCLC.
- 2.VIAF:** Registro Internacional de Autoridad, contiene registros bibliográficos de autores. Publicado por OCLC.
- 3.FAST:** Tesauro y esquema de clasificación multilingüe. OCLC ha publicado sus datos como datos enlazados.
- 4.DBPedia:** Proyecto de Web Semántica a partir de la extracción de datos estructurados de Wikipedia
- 5.MeSH RDF:** Tesauro, publicado por la Biblioteca Nacional de Medicina de los Estados Unidos.

Ventajas y Desventajas

Ventajas	<ul style="list-style-type: none">• Modelo estándar a nivel de Internet• Contenido capaz de ser entendido por las máquinas• Resultados de búsqueda más precisos, por significado; enriquecidos con atributos y relaciones a otros conjuntos• Capacidad para Multilenguaje• Construcción de registro de autoridades: Autores, Instituciones• Capacidad de establecer relaciones en Internet con otros conjuntos de datos de Bibliotecas y Autoridades
Desventajas	<ul style="list-style-type: none">• Es costoso y laborioso adaptar los documentos de Internet, para poder ser procesados de forma semántica• Afortunadamente los datos de LILACS y DeCS son datos estructurados y se puede realizar la conversión de forma automática

Estándares

RDF: Método para descomponer el conocimiento en piezas pequeñas (triples de conocimiento):



URI: Identificador Único de Recurso en toda la Web

<https://id.nlm.nih.gov/mesh/D000001.html>

Ontología: define, a través de clases, los *nombres de las entidades, propiedades y relaciones entre ellas*; para describir y representar un dominio específico.

Book, Person, Author

JSON-LD, RDF/XML, N3, Turtle: formas de serialización de RDF

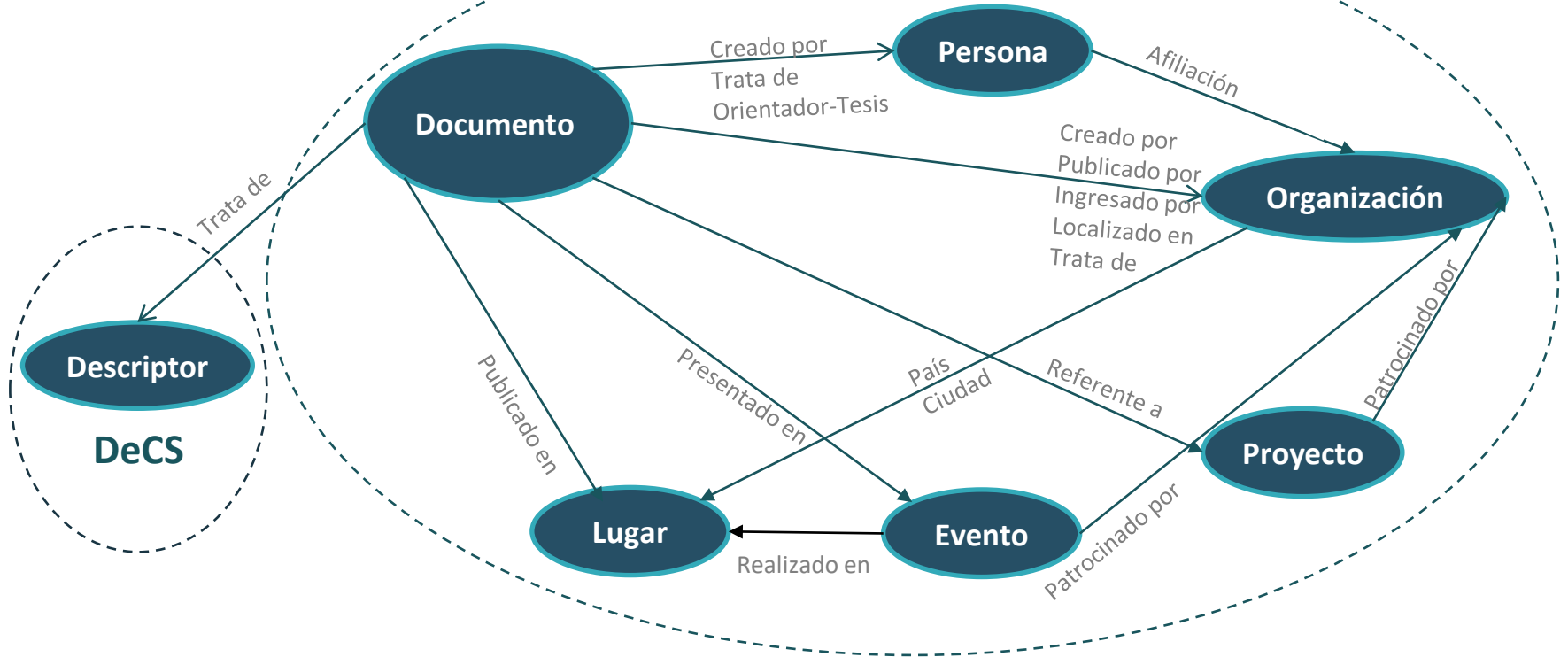
Solución



1. Desarrollar el modelo de entidades y relaciones de LILACS.
2. Producir los datos basados en estas entidades.
3. Asociar las entidades con un identificador persistente (URI).
4. Validación de Entidades
5. Publicar como datos enlazados

LILACS – Entidades

LILACS



LILACS – Ontologías

1. Schema.org

- Desarrollada por Google, Yahoo, Bing y Yandex
- cubre los aspectos generales de vida diaria,
- con los elementos básicos de muchos dominios específicos
- incluye casi todas la entidades de LILACS

2. Lilacs

- Creada para entidades y propiedades específicas de LILACS, que no pueden ser representadas con schema.org

LILACS

Documento

Persona

Institución

Evento

Lugar

SCHEMA.ORG

CreativeWork

Person

Organization

Event

Place

Propiedades

Tipo de Literatura

Nivel de Tratamiento

Schema.org Clases

Clases		Propiedades	
CreativeWork	Person	author	name
Book	Organization	contributor	description
Chapter	Event	editor	keywords
Article	PublicationEvent	affiliation	inLanguage
Thesis	Place	about	url
AudioObject	PostalAddress	publisher	datePublished
VideoObject	Language	publication	issn
Map		provider	isbn
Collection		releasedEvent	bookEdition
BookSeries		isPartOf	bookFormat
Periodical		hasPart	numberOfPages
PublicationVolume			pageStart
PublicationIssue			pageEnd
			volumeNumber
			issueNumber
			inSupportOf
			fileFormat

“Hemos argumentado que el desarrollo de modelos para la descripción de recursos bibliotecarios puede basarse en Schema.org—no sólo porque es el vocabulario más ampliamente adoptado para exponer datos estructurados en la Web y es reconocido por los principales motores de búsqueda del mundo—sino por la Una razón mucho más simple: es un estándar sofisticado, publicado por primera vez durante una época de inestabilidad histórica, en la evolución de los estándares para la descripción de recursos bibliotecarios..”

OCLC Research

Producción de datos de Entidades

Migración del modelo LILACS a Modelo Datos enlazados (Schema.org)

- Las propiedades literales se traducen a la *propiedad* en el nuevo modelo
- Las propiedades que son entidades:
 - Se le asigna el tipo (*@type*) correspondiente
 - Se guarda como entidad

Producción de datos de Entidades

LILACS	Document	Organization	Person
	@type: Article	@type: Organization	@type: Person
	id:1	id: 3	id:2
V1: CU1	provider: @type: Organization branchCode: CU1	branchCode: CU1	
V10: Guerra Valdés, R	author: @type:Person name: Guerra Valdés, R		name: Guerra Valdés, R
V12: Conceptos actuales de cirugía experimental	name: [{@value: Conceptos actuales de cirugía experimental, @language: es}, {@value: Actual concepts of experimental surgery, @language: en}]		
V13: Actual concepts of experimental surgery			

Asignar URIs a las Entidades

Estructura de la URI:

- protocolo
- dominio
- nombre de entidad en plural
- identificador de entidad

<https://lilacs-ld.org/documents/1>

Asignar URIs a las Entidades

Document	Organization	Person
@type: Article	@type: Organization	@type: Person
id:1	id: 3	id:2
@id: https://lilacs-ld.org/documents/1	@id: https://lilacs-ld.org/organizations/3	@id: https://lilacs-ld.org/persons/2
provider: @type: Organization branchCode: CU1 @id: https://lilacs-ld.org/organizations/3	branchCode: CU1	
author: @type:Person name: Guerra Valdés, R @id: https://lilacs-ld.org/persons/2		name: Guerra Valdés, R

Validación de Entidades

id	Data	isValid	sameAs	inDocument
25	{"name": "Hospital Universitario Gineco- Obstétrico Mariana Grajales", "@type": "Organization"}	1	24	[51]
24	{"name": "Hospital Universitario Gineco-Obstétrico Mariana Grajales", "@type": "Organization"}	1		[52]
7	{"name": "ICBP Victoria de Giron", "@type": "Organization"}	1	1	[12]
1	{"name": "ICBP Victoria de Girón", "@type": "Organization"}	1		[1,2,7,19,21]
13	{"name": "ICBP Victroria de Girón", "@type": "Organization"}	1	1	[32]
6	{"name": "Wx", "@type": "Organization"}	0		[36]

Publicar como datos enlazados

Api RESTFul para LILACS-Id

- Permite la consulta y administración de los datos:
 - Recuperar, filtrar, buscar, paginar, ordenar, obtener atributos específicos
 - Insertar
 - Modificar
 - Eliminar
- Entidades: Document, Person, Organization, Event, Project
- Datos a Insertar o Modificar a traves de parámetro *data* en formato json
- Respuesta en formato json

Peticiones:

<http://lilacs-ld.org/documents?name=cirugia&q=Perros&type=Article>

<http://lilacs-ld.org/persons/123?fields=name,email>

POST <http://lilacs-ld.org/documents/?data=>

PUT <http://lilacs-ld.org/persons/?id=5&data=>

DELETE <http://lilacs-ld.org/persons/123>

Requerimientos

- **MySQL 5.7.8+**, con tipo de dato JSON activado
 - Este tipo de datos garantiza la validación y el acceso eficiente a los elementos del JSON.
 - Cuenta con funciones para:
 - Crear valores JSON (JSON_ARRAY, JSON_OBJECT)
 - Realizar búsquedas (JSON_SEARCH, JSON_CONTAINS, JSON_KEYS, ...)
 - Modificar (JSON_REPLACE, JSON_APPEND, JSON_INSERT, JSON_REMOVE, ...)
 - Obtener atributos del JSON (JSON_TYPE, JSON_DEPTH, JSON_VALID, ...)
 - Crear un índice
 - Este tipo de datos permite almacenar datos mas complejos en una columna, lista de valores, cantidad variable de atributos. De esta forma permite la representación NoSQL de los datos
- **Php 5.6.19+** (soporta campo JSON de MySQL)

Proyección



La continuación del Proyecto incluye:

- La creación de la capa semántica a partir de los datos de fi-admin
- La inclusión de esta capa en los documentos de la BVS para aumentar su presencia en Internet
- La posibilidad de enriquecer los resultados de la Búsqueda en la BVS con relaciones a otros conjuntos de datos
- Implementar SPARQL para realizar consultas tipo SQL a los contenidos RDF

Conclusiones



Exponer LILACS y DeCS como un conjunto de datos semánticos, con datos enlazados favorece que los buscadores encuentren estos contenidos de forma mas rápida y eficiente.

Se pueden mejorar los sistemas de búsqueda de la BVS y relacionar los datos actualmente aislados en otras aplicaciones. Los datos enlazados adquieren mayor significado para los usuarios a través de su vinculación con datos similares.

Agregar datos adicionales a los resultados de la búsqueda a partir de las bases de datos nacionales de los Ministerios de Salud, directorios de usuarios, bases de datos de categorías científicas, así como directorios internacionales de autores, artículos, libros, tesis y producción científica es una proyección de este proyecto.

Referencias

- **Web semántica, datos enlazados y bibliotecas**

Disponible en: <https://acimed.sld.cu/index.php/acimed/article/view/1058>

- **Nube de Datos Abiertos Enlazados (The Linked Open Data Cloud)**

Disponible en: <https://lod-cloud.net/>

- **Library Linked Data in the Cloud. (OCLC Research)**

Disponible en: <https://www.oclc.org/research/publications/books/library-linked-data-in-the-cloud/chapter5.html>

- **Los datos estructurados en la Web Semántica**

Disponible en: <https://www.ionos.es/digitalguide/paginas-web/creacion-de-paginas-web/los-datos-estructurados-en-la-web-semantica/>

- **Four ways WorldCat.org improves library visibility**

Disponible en: <https://blog.oclc.org/next/four-ways-worldcat-org-improves-library-visibility/>