

Integración semántica de los recursos de información en una memoria corporativa

Erik Alarcón Zamora

Enero 2014. México, D.F.

Asesores:

Dra. Reyna Carolina Medina Ramírez

Dr. Héctor Pérez Urbina

Contenido I

1 Marco Introdutorio

- Contexto y Motivación
- Descripción del Problema

2 Integración Semántica de una Memoria Corporativa

- Marco de Referencia
- Arquitectura de la Integración Semántica
- Casos de Uso
- Representación el Conocimiento
- Enriquecer el conocimiento en el modelo semántico
- Buscar y recuperar la información en el modelo semántico

3 Referencias

Memoria Corporativa I

Definición

La representación explícita, tácita, consistente y persistente del conocimiento de una organización. [Gandon, 2002]

Finalidad

Una memoria corporativa conserva y mantiene el conocimiento de una organización [Dieng et al., 1998], para *facilitar el acceso, intercambio y difusión de éste*.

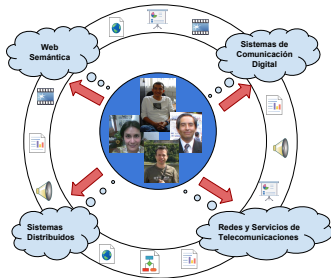
Caso de Estudio

El grupo de investigación del área de Redes y Telecomunicaciones (RyT) de la Universidad Autónoma Metropolitana Unidad Iztapalapa (UAM-I).

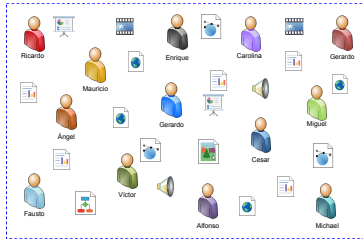
Memoria Corporativa II

Recurso de Información

Un elemento que representa y encapsula una parte del conocimiento de una organización (investigaciones, colaboraciones, proyectos, cursos, temas de interés, objetos e ideas).



(a) Conocimiento del área de RyT



(b) Memoria Corporativa del área de RyT

Memoria Corporativa III

Diversidad en formato



pdf, doc, odp, html, txt, xsl, wav, png, mp3, mp4, mpeg, mov, ppt, mov

Diversidad en contenido



p2p, middleware, estado global, replicación, concurrencia, sincronización

Diversidad en estructura



estructurados

semi-estructurados

sin estructura

Homonimia

radio ---> Química, Comunicación, Anatomía o Geometría

Sinonimia

herramienta = aparato = instrumento = mecanismo = artificio

Tecnologías Semánticas I

Definición

Un conjunto de metodologías, lenguajes, aplicaciones, herramientas y estándares para suministrar u obtener el significado de las palabras, información y las relaciones entre éstos. [Alfred et al., 2010]



Tecnologías Semánticas II

Resource Description Framework (RDF)

Marco genérico para describir el conocimiento e información explícita de los recursos mediante sus características y relaciones. [Bouzid et al., 2012]

Recurso

Puede ser cualquier cosa (persona, lugar, documento, entidades del mundo real o conceptos abstractos) que tiene un identificador único de recursos (URI).

Propiedad

Un aspecto significativo, característica, o relación que se describe de un recurso (relación binarias).

Tecnologías Semánticas III

Clase

Una colección de objetos que comparten características comunes.

Literal

Un valor de datos como cadenas o enteros particulares.

Declaración

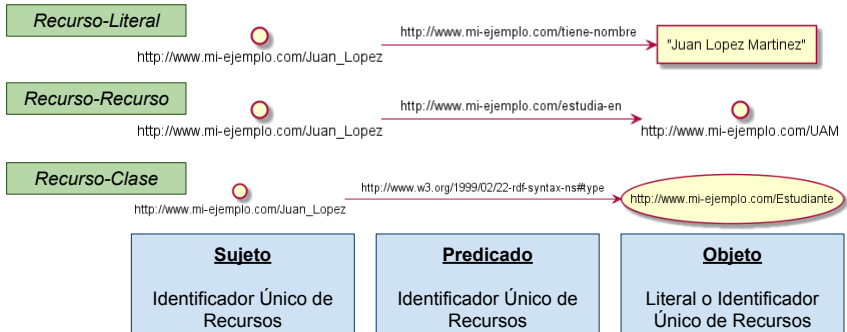
Una afirmación de un hecho explícito sobre un recurso, en términos de una propiedad y el valor asignado a ésta.

- Juan se llama “Juan López Martínez”.
- Juan estudia en la UAM.
- Juan es un estudiante.

Tecnologías Semánticas IV

Tripleta RDF

La forma básica para representa una declaración en un modelo semántico.

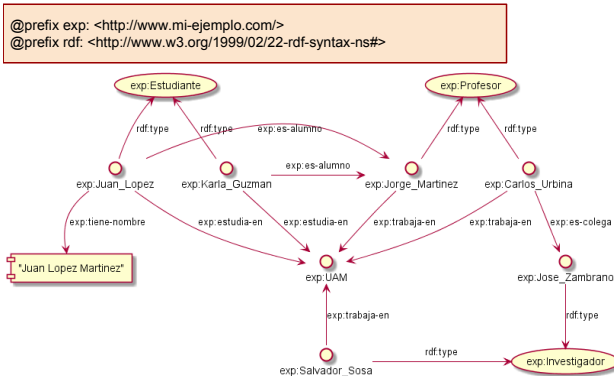


Tecnologías Semánticas V

Grafo RDF

Un grafo estructurado y dirigido compuesto por nodos, aristas y etiquetas para representar las tripletas.

Tecnologías Semánticas VI



Tecnologías Semánticas VII

SPARQL

Lenguaje de consulta y protocolo de acceso a RDF, para la búsqueda y recuperación de la información en un grafo RDF.

Motor de Búsqueda SPARQL

Programa que interpreta una consulta SPARQL, la compara con el grafo RDF y recupera los valores de la misma.

###Lista de prefijos

PREFIX exp: <http://www.mi-ejemplo.com/>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

Variables a recuperar

SELECT ?x

WHERE {

Patrones tripletas

?x exp:estudia-en exp:UAM.

?x rdf:type exp:Estudiante. }

Tecnologías Semánticas VIII

Ontología

Una definición formal, explícita y compartida de los conceptos, así como las relaciones de un determinado dominio. [Gruber, 1993]

Componente Asertivo (ABox)

Este componente está constituido por las declaraciones (descripciones o hechos verdaderos) que afirman que los individuos son instancias de una clase o propiedad.

Componente Terminológico (TBox)

Este componente describe las clases y propiedades relevantes, así como las reglas de inferencia que permiten aprovechar la manera en que las instancias se relacionan entre sí.

Tecnologías Semánticas IX

Reglas de inferencia o Axiomas

Los *axiomas* o *reglas de inferencia* [Gruber, 1993] son expresiones para enriquecer el conocimiento explícito en un grafo RDF.

Funcionalidad Axiomas

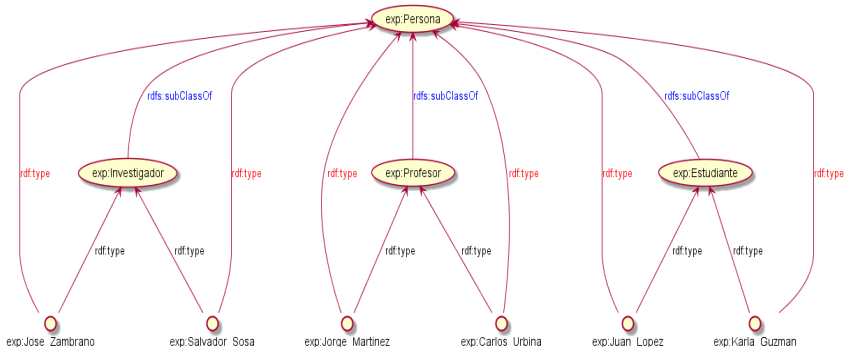
Describir relaciones entre clases, definir propiedades en términos de otras, definir relaciones entre conceptos, definir restricciones de cómo las propiedades se relacionan, por mencionar algunos.

Razonador

Un programa que deduce declaraciones a partir de los axiomas y declaraciones explícitas en la ontología.

Tecnologías Semánticas X

```
@prefix exp: <http://www.mi-ejemplo.com/>  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```



Integración Semántica

Definición

La búsqueda y recuperación significativa de información existente en los recursos de información para responder una consulta dada por un usuario.

Etapas

Representar el *conocimiento* de los *recursos de información* en un *modelo semántico*.

Buscar y recuperar información existente en la memoria corporativa mediante la interrogación del modelo semántico.

Pregunta Investigación

¿Las tecnologías semánticas son viables para solucionar la integración semántica de los recursos de información de una memoria corporativa?



Objetivos

Objetivo Principal

Contribuir a la integración semántica de los recursos de información en una memoria corporativa, mediante el uso de las tecnologías semánticas.

Objetivos Particulares

- 1 Desarrollar una **marco de referencia** para la *integración semántica* de los *recursos de información* existentes en una *memoria corporativa*.
- 2 Implementar un **modelo semántico** que representa el *conocimiento explícito e implícito* de los *recursos de información*.
- 3 Implementar un **prototipo de interfaz gráfica de usuario** que permita a los usuarios una interacción amigable para la integración semántica de los recursos de información.
- 4 Evaluar los **resultados devueltos** y **tiempos de procesamiento** en la *integración semántica* para el dominio de redes y telecomunicaciones.

Metodología

Marco de Referencia

- 1 Identificar los *casos de uso* para encontrar los principales *recursos de información* existentes en la memoria, así como los criterios de búsqueda asociados a éstos.
- 2 Construir el diagrama de casos de uso.
- 3 Evaluar herramientas semánticas para: edición de descripciones semánticas, edición de reglas de inferencia, gestión de modelos semánticos.
- 4 Recopilar los recursos de información de acuerdo a los casos de uso.
- 5 Adquirir el conocimiento o información de los recursos de información con base en las características y relaciones de los mismos.
- 6 Construir el diagrama de clases.

Metodología

Marco de Referencia

Modelo Semántico

- 7 Describir el conocimiento explícito de los *recursos de información* recopilados en un modelo semántico.
- 8 Identificar las reglas de inferencia a introducir en el modelo, con base en el diagrama de clases.
- 9 Escribir las reglas de inferencia para enriquecer el modelo semántico con conocimiento implícito, mediante el uso del editor de reglas de inferencia.
- 10 Identificar las preguntas en lenguaje natural a partir de los casos de uso.
- 11 Diseñar las consultas en el *lenguaje estándar de búsqueda* que correspondan a las preguntas en lenguaje natural.

Metodología

Marco de Referencia

- 12 Emplear un proceso que permita hacer explícito el conocimiento implícito.
- 13 Buscar y recuperar información en la memoria corporativa, interrogando el modelo semántico.

Prototipo de interfaz gráfica de usuario

- 14 Diseñar un prototipo para interacción (búsqueda y navegación) amigable y transparente de los usuarios de la memoria con el modelo semántico.
- 15 Proponer funcionalidades básicas del prototipo.
- 16 Indicar cuáles son las interfaces para los usuarios (pantallas).
- 17 Describir las especificaciones de estas interfaces.
- 18 Implementar el prototipo y realizar pruebas del mismo.

Metodología

Evaluar los resultados devueltos

- 19 Evaluar la calidad de los resultados (recursos relevantes recuperados) con y sin inferencia, mediante el uso de métricas que se emplean en la recuperación de la información: exhaustividad y precisión.
- 20 Identificar aquellos recursos (total de recursos relevantes) que responden las preguntas del paso 10 de este listado.
- 21 Consultar al modelo semántico y comparar los recursos relevantes recuperados con los recursos relevantes que se identificaron en el paso 20 de este listado.
- 22 Calcular la exhaustividad y precisión.

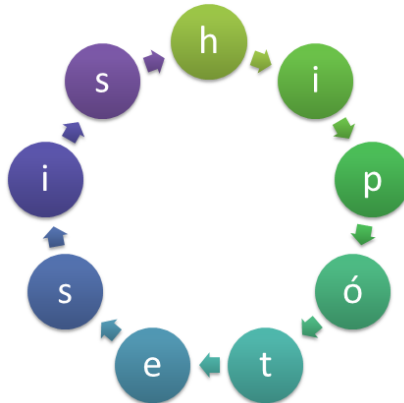
Metodología

Evaluar los tiempos de procesamiento

- 23 Evaluar los *tiempos promedios* que toma la herramienta electa de gestión de los modelos semánticos, para consultar los modelos con/sin inferencia.
- 24 Elaborar un script que calcule '*n*' veces el *tiempo de procesamiento* al consultar un modelo semántico (con o sin inferencia). Las consultas se hacen a las preguntas identificadas del paso 10 de este listado.

Hipótesis

El uso de las *tecnologías semánticas* es adecuado para lograr la *integración semántica de recursos de información* en una *memoria corporativa*.



Aportaciones

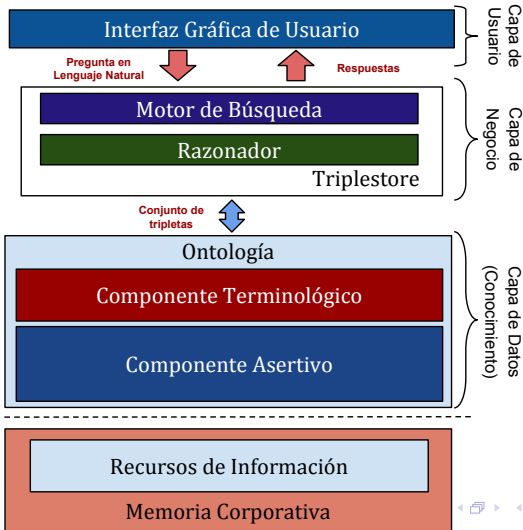
- 1 Un *marco de referencia* para lograr la integración semántica de recursos de información.
- 2 Un modelo semántico que representa el conocimiento de una memoria corporativa, el cual tiene tres ramas principales (Personas, Recursos Digitales y Conceptos del Redes y Telecomunicaciones).
- 3 Un prototipo (interfaz gráfica de usuario) para la interacción amigable (búsqueda y consulta de información) de los usuarios al modelo semántico.
- 4 Los resultados de nuestra evaluación experimental.
- 5 Un par de scripts para la generación automática y controlada de descripciones (conocimiento explícito) de los *recursos de información*, con el fin de poblar la base de conocimiento.

Marco de Referencia

Etapas

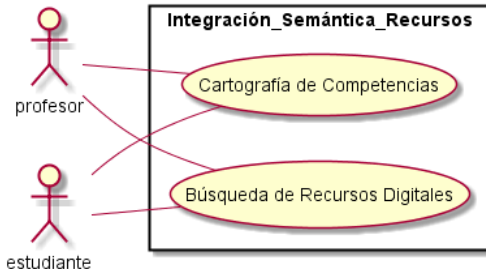
- 1 ***Representación del conocimiento explícito de los recursos*** consiste en identificar los recursos de información de la memoria corporativa, así como representar las características y relaciones (conocimiento explícito) de estos recursos en un modelo semántico.
- 2 ***Enriquecimiento del conocimiento en el modelo*** consiste en introducir reglas de inferencia (axiomas) para completar y enriquecer el modelo semántico con conocimiento implícito del dominio de la memoria corporativa.
- 3 ***Búsqueda y recuperación de la información en el modelo*** consisten en identificar las principales consultas de los usuarios, así como interrogar el modelo semántico para recuperar la información que responda a estas consultas.

Arquitectura de la Integración Semántica



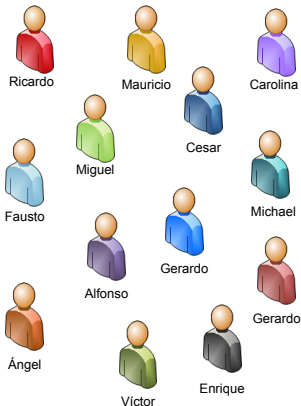
Casos de Uso

- ***Cartografía de Competencias*** consiste en la búsqueda y recuperación de información significativa de las personas a partir de las características personales y profesionales de las mismas.
- ***Búsqueda de Recursos Digitales*** consiste en la búsqueda y recuperación de información significativa de los documentos y archivos multimedia a partir del contenido de los mismos.

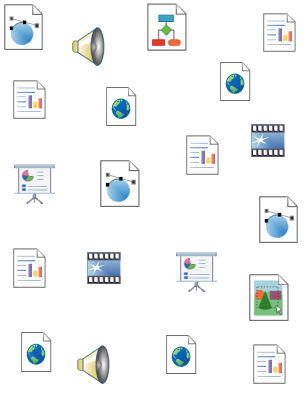


Identificar los principales recursos de información

Cartografía de Competencias



Búsqueda Recursos Digitales



Memoria Corporativa del área de RyT

Diagrama de clases que muestra la estructura de un sistema de gestión de personas y organizaciones. Las clases y sus atributos son:

- Persona**: Nombre: String, Edad: Int, Número telefónico: String, Email: URL, Sitio Web: URL.
- Lenguaje**: (Sin atributos visibles).
- Género**: (Sin atributos visibles).
- Estudiante**: (Sin atributos visibles).
- Investigador**: Línea de investigación: String.
- Profesionalista**: (Sin atributos visibles).
- Empleado**: (Sin atributos visibles).
- Organización**: (Sin atributos visibles).
- CompetenciaProfesional**: (Sin atributos visibles).
- TemaODaRYT**: (Sin atributos visibles).
- Profesor**: (Sin atributos visibles).
- Universidad**: (Sin atributos visibles).

Relaciones y asociaciones:

- Persona** tiene género **Género**.
- Persona** es un **Estudiante**.
- Persona** es un **Investigador**.
- Persona** es un **Profesionalista**.
- Persona** es un **Profesor**.
- Persona** tiene asesor **Profesionalista**.
- Persona** trabaja en **Organización**.
- Persona** estudia en **Universidad**.
- Persona** es competente en **CompetenciaProfesional**.
- Persona** es habilidoso en **TemaODaRYT**.
- Profesionalista** es un **Empleado**.
- Profesionalista** tiene colega **Profesionalista**.
- Profesionalista** tiene colega **Profesor**.
- Empleado** trabaja en **Organización**.
- Profesor** trabaja en **Universidad**.

Representar el conocimiento e información mediante el estándar RDF I

Actividades en la representación del conocimiento

- 1 Asignar un identificador único de recursos para cada recurso de información en la memoria corporativa.
- 2 Asignar los identificadores únicos de recursos a las propiedades.
- 3 Reconocer los valores de las propiedades: otro recurso o literal.
- 4 Generar las tripletas RDF asociadas a las descripciones de los recursos de información.

```
@prefix sirp: <http://arte.izt.uam.mx/ontologies/personRyT.owl#>  
@prefix sird: <http://arte.izt.uam.mx/ontologies/digiResourceRyT.owl#>  
@prefix redes: <http://mcyti.izt.uam.mx/arios/odaryt.owl#>
```

Representar el conocimiento e información mediante el estándar RDF II

Recursos de la Cartografía de Competencias

Nombre	Identificador
Alfonso Prieto	<i>sirp:AlfonsoPrieto</i>
Michael Pascoe	<i>sirp:MichaelPascoe</i>
Reyna Carolina Medina	<i>sirp:CarolinaMedinaRamirez</i>
Ricardo Marcelin	<i>sirp:RicardoMarcelinJimenez</i>
Miguel López	<i>sirp:MiguelLopez</i>
Víctor Manuel Ramos	<i>sirp:VictorRamosVictorRamos</i>
Fausto Marcos Casco	<i>sirp:FaustoCasco</i>
Cesar Jalpa	<i>sirp:CesarJalpa</i>
Enrique Rodríguez	<i>sirp:EnriqueRodriguez</i>

Representar el conocimiento e información mediante el estándar RDF III

Recursos de la Búsqueda de Recursos Digitales

Nombre	Identificador
Ontology engineering	<i>sird:RR-4396-2002-pdf</i>
A Description Logic Primer	<i>sird:A-DescriptionLogicPrimer-2012-pdf</i>
Introduction to Ontologies and OWL	<i>sird:Introduction2Ontol-2005-pdf</i>
The Semantic Web - An Overview	<i>sird:TheSemanticWeb-AnOverview-2011-flv</i>
What is Linked Data?	<i>sird:What-isLinkedData-2012-flv</i>

Representar el conocimiento e información mediante el estándar RDF IV

Propiedades de la Cartografía de Competencias

Característica o relación	Identificador
Nombre	<i>sirp:has-name</i>
Sitio Web	<i>sirp:has-webSite</i>
Lugar de trabajo	<i>sirp:worksIn</i>
Línea de investigación	<i>sirp:researchesOn</i>
Colega	<i>sirp:has-colleague</i>
Competencias	<i>sirp:competentIn</i>
Habilidades en temas de Redes y Telecom.	<i>sirp:expertiseIn</i>

Representar el conocimiento e información mediante el estándar RDF V

Propiedades de la Búsqueda de Recursos Digitales

Característica o relación	Identificador
Título	<i>sird:has-title</i>
Autor	<i>sird:has-author</i>
Ruta Archivo	<i>sird:has-filePath</i>
Año de creación	<i>sird:has-yearOfCreation</i>
Lenguaje Fuente	<i>sird:has-languageSource</i>
Temas de Redes y Telecom.	<i>sird:has-topic</i>

Representar el conocimiento e información mediante el estándar RDF VI

@prefix **sirp**: <http://arte.izt.uam.mx/ontologies/personRyT.owl#> .

@prefix **xsd**: <http://www.w3.org/2001/XMLSchema#> .

@prefix **redes**: <http://mcyti.izt.uam.mx/arios/odaryt.owl#> .

sirp:RicardoMarcelinJimenez

a **sirp:Teacher** ;

sirp:has-name "Ricardo Marcelin Jiménez"^^**xsd:string**;

sirp:has-email "calu@xanum.uam.mx"^^**xsd:anyURI**;

sirp:has-webSite "http://cbi.izt.uam.mx/electrica/profs/ricardo_marcelin.html"^^**xsd:anyURI**;

sirp:has-gender **sirp:Male**;

sirp:worksIn **sirp:UAM**;

sirp:researchesOn "El almacenamiento distribuido, las redes inalámbricas de sensores y la simulación de eventos discretos."^^**xsd:string**;

sirp:expertiseIn **redes:Distributed_Systems**, **redes:Distributed_Storage**, **redes:MDS_Codes**,
redes:Performance_evaluation, **redes:Semantic_Annotations**, **redes:Image_compression**,
redes:Routing_Protocols, **redes:Distributed_Algorithms**, **redes:Wireless_Sensor_Networks**,
redes:N_and_ST;

sirp:competentIn **sirp:Article_Reviewing_Skills**, **sirp:Thesis_Supervision_Skills**,
sirp:Oral_And_Written_Communication_Skills, **sirp:Area_Expert**, **sirp:Analysis_Skills**,
sirp:Decision_Making_Skills, **sirp:Research_Skills**, **sirp:Problem_Solving_Skills**,
sirp:Synthesis_Skills, **sirp:Abstraction_Skills**, **sirp:Counseling_Skills_for_Social_Service**,
sirp:IT_And_Communication_Skills;

sirp:has-colleague **sirp:MiguelLopez**, **sirp:CarolinaMedinaRamirez**;

sirp:reads **sirp:Spanish**, **sirp:English**;

sirp:writes **sirp:Spanish**, **sirp:English**;

sirp:speaks **sirp:Spanish**, **sirp:English**.



Representar el conocimiento e información mediante el estándar RDF VII



```
@prefix sird: <http://arte.izt.uam.mx/ontologies/digiResourceRyT.owl#> .  
@prefix sirp: <http://arte.izt.uam.mx/ontologies/personRyT.owl#> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
@prefix redes: <http://mcyti.izt.uam.mx/arios/odaryt.owl#> .
```

sird:What-isLinkedData-2012-flv

a sird:Video ;

sird:has-title "What is Linked Data?"^^xsd:string ;

sird:has-author "Manu Sporny"^^xsd:string ;

sird:has-filePath "http://www.youtube.com/watch?v=4x_xzT5eF5Q"^^xsd:anyURI ;

sird:has-fileExtension sird:flv ;

sird:has-languageSource sirp:English ;

**sird:has-topic redes:Linked_data, redes:Data, redes:Web, redes:Semantic,
redes:Knowledge_graph ;**

**sird:has-synopsis "A short non-technical introduction to Linked Data, Google's
Knowledge Graph, and Facebook's Open Graph Protocol"
^^xsd:string ;**

sird:has-yearOfCreation "2012"^^xsd:int .

Enriquecer el conocimiento en el modelo semántico

Axiomatización

Para cada *caso de uso* debe encontrarse el respectivo conjunto de axiomas (TBox).

Lo que es obvio para un humano, no lo es para una maquina.

Lenguajes

Especificaciones para describir clases, propiedades e individuos.

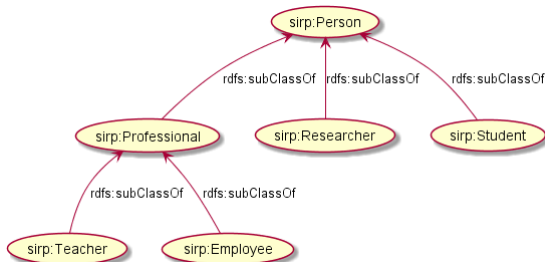
- *RDF Schema* **RDF(S)**
- *Web Ontology Language* **OWL**

```
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
@prefix owl: <http://www.w3.org/2002/07/owl#>
```

Herencia de Clases I

Subclase (rdfs:subClassOf)

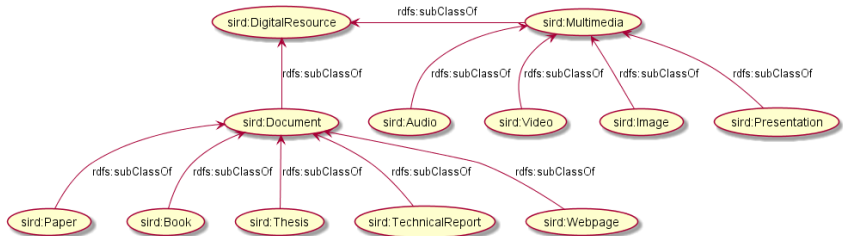
Afirma que una *clase A* se subsume por una *clase B*, es decir, la clase A es un caso particular de la *clase B*. En este caso, las instancias de la clase A son instancias de la clase B.



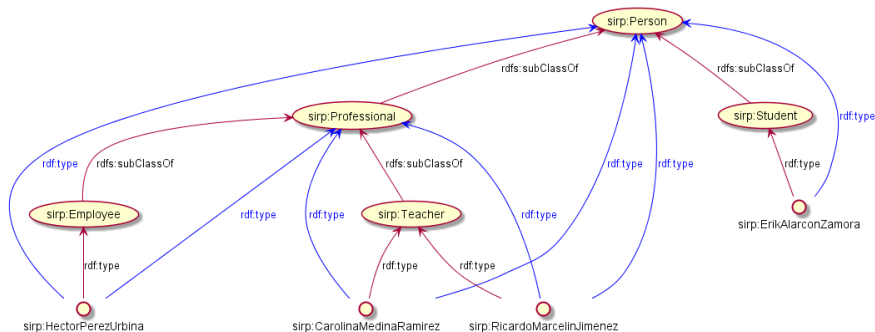
Herencia de Clases II

Subclase (rdfs:subClassOf)

Afirma que una *clase A* se subsume por una *clase B*, es decir, la clase A es un caso particular de la *clase B*. En este caso, las instancias de la clase A son instancias de la clase B.



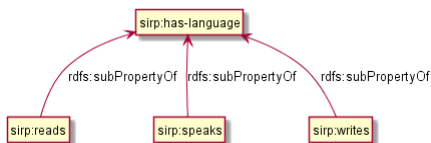
Herencia de Clases III



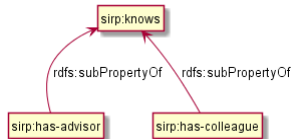
Herencia de Propiedades I

Subpropiedad (rdfs:subPropertyOf)

Afirma que todos los recursos que se relacionan por la *propiedad* X, también se relacionan por la *propiedad* Y.

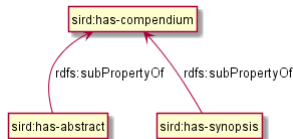


(c) Jerarquía de propiedades para las habilidades lingüísticas

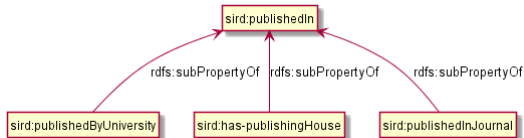


(d) Jerarquía de propiedades para las relaciones profesionales entre personas

Herencia de Propiedades II

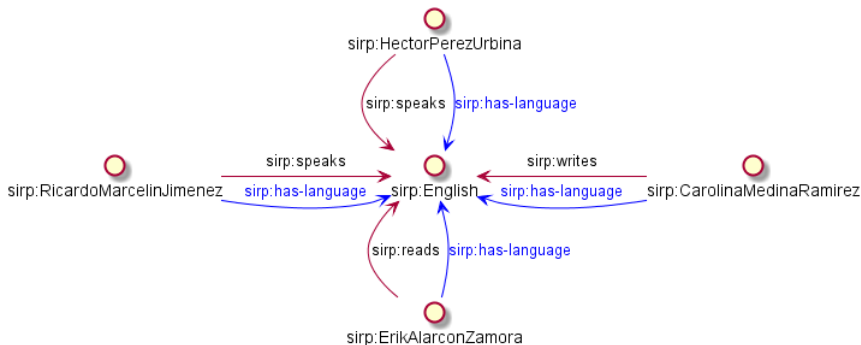


(e) Jerarquía de propiedades para describir el contenido de un recurso digital



(f) Jerarquía de propiedades para vincular a una organización con un recurso digital

Herencia de Propiedades III



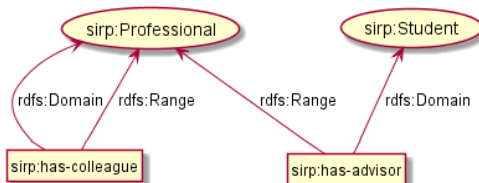
Dominio y Rango en las Propiedades I

Dominio (rdfs:domain)

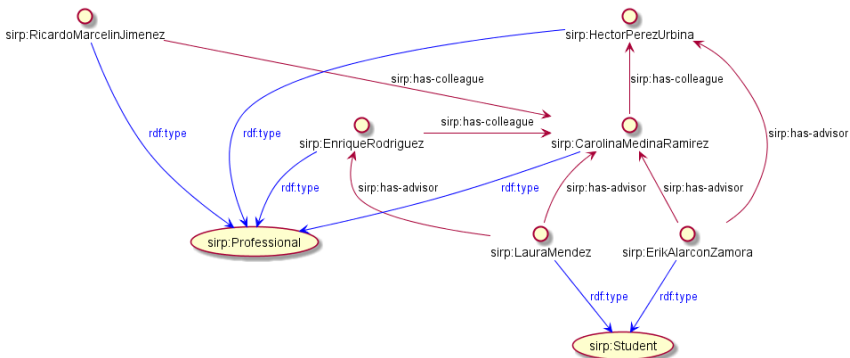
Especifica qué clase se aplica a una propiedad.

Rango (rdfs:range)

Especifica los valores (clase o tipo de literal) que puede asumir una propiedad.



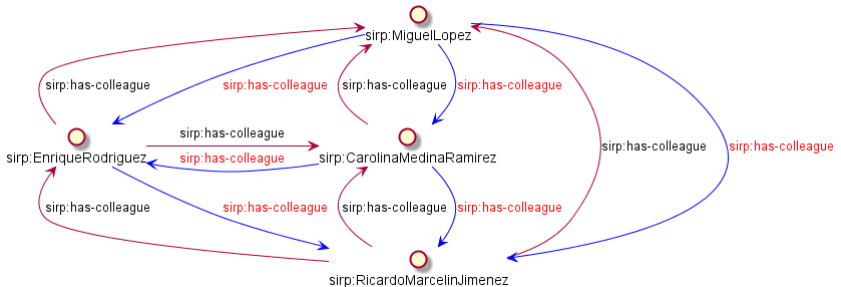
Dominio y Rango en las Propiedades II



Características en las propiedades

Propiedad simétrica (owl:SymmetricProperty)

Afirma que la *propiedad X* es su propia propiedad inversa, es decir, si la *propiedad X* relaciona al *individuo A* con el *individuo B*, entonces, esta propiedad debe relacionar al *individuo B* con el *individuo A*.



Buscar y recuperar la información en el modelo semántico

Objetivo

La búsqueda y recuperación de la información para responder las preguntas o necesidades informativas de los usuarios del área de Redes y Telecomunicaciones (RyT).

Actividades

- 1 Identificar las preguntas en lenguaje natural.
- 2 Transformar las preguntas a una consultas SPARQL.
- 3 Ejecutar las consultas mediante un motor de búsqueda SPARQL.

Identificación de las preguntas en lenguaje natural I

Cartografía de Competencias

- Q1.1.— **¿Cuáles son los nombres, correos, sitios web, géneros y edades de las personas del área de RyT?**
- Q1.2.— ¿Cuáles son los nombres, sitios web y los lugares donde laboran las personas del RyT?
- Q1.3.— ¿Quiénes son mayores de 20 años y menores de 45 años?
- Q1.4.— ¿Cuáles son los nombres y sitios web de los profesionistas del área de RyT?
- Q1.5.— ¿Quiénes trabajan en la Clark & Parsia y son del sexo Masculino?
- Q1.6.— ¿Quiénes son estudiantes y leen en inglés?
- Q1.7.— ¿Quiénes hablan, leen y escriben en inglés?

Identificación de las preguntas en lenguaje natural II

- Q1.8.— ¿Qué estudiantes saben algo de inglés?
- Q1.9.— ¿Qué profesores tienen la capacidad de síntesis?
- Q1.10.— ¿Qué profesionistas tienen conocimiento en los temas de Web Semántica?
- Q1.11.— ¿Qué profesores tienen conocimientos en Sistemas Distribuidos?
- Q1.12.— ¿Quiénes tienen conocimiento en Java, OWL, RDF, Threads, C, OpenMP?
- Q1.13.— ¿Qué estudiantes tienen algún conocimiento en los subtemas de Sistemas Operativos?
- Q1.14.— ¿Quiénes trabajan en una Universidad?
- Q1.15.— ¿Quiénes laboran en la UAM y tienen algún conocimiento en Web Semántica?
- Q1.16.— ¿Qué personas tienen como asesor a Carolina Medina?

Identificación de las preguntas en lenguaje natural III

- Q1.17.— ¿Quiénes son los colegas de Ricardo Marcelin?
- Q1.18.— **¿Cuáles son los nombres y correos de las personas que conocen a Carolina Medina Ramírez?**
- Q1.19.— ¿Qué personas son profesores-investigadores?

Búsqueda de Recursos Digitales

- Q2.1.— ¿Cuáles son los títulos, rutas, extensión, idioma de todos los recursos digitales de RyT?
- Q2.2.— ¿Cuáles libros tratan sobre algunos temas de Sistemas Distribuidos?
- Q2.3.— ¿Qué recursos fueron publicados por la UAM?
- Q2.4.— ¿Qué documentos sirven para dar un curso de Sistemas P2P?

Identificación de las preguntas en lenguaje natural IV

- Q2.5.— ¿Qué recursos multimedia son mayores al año 2009?
- Q2.6.— **¿Cuáles documentos tratan sobre Ontologías?**
- Q2.7.— ¿Qué recursos fueron publicados en una Revista científica?
- Q2.8.— ¿Qué recursos tienen en su descripción las palabras "linked data"?
- Q2.9.— ¿Cuáles documentos en Inglés y mayores al año 2000 son de autoría de Erik Alarcon Zamora?
- Q2.10.— ¿Cuál son las tesis de Samuel Hernandez Maza?

Transformar las preguntas a una consultas SPARQL I

¿Cuáles son los nombres, correos, sitios web, géneros y edades de las personas del área de RyT?



```
PREFIX sirp: <http://arte.izt.uam.mx/ontologies/personRyT.owl#>

SELECT ?name ?mail ?ws ?gender ?age
WHERE
{
    ?x sirp:has-name ?name;
    sirp:has-email ?mail;
    sirp:has-webSite ?ws;
    sirp:has-gender ?gender;
    sirp:has-age ?age.
}
```

Transformar las preguntas a una consultas SPARQL II

¿Cuáles son los **nombres** y **sitios Web** de las personas que conocen a **Carolina Medina Ramírez**?



```
PREFIX sirp: <http://arte.izt.uam.mx/ontologies/personRyT.owl#>
```

```
SELECT ?name ?ws  
WHERE  
{
```

```
{?x sirp:has-colleague sirp:CarolinaMedinaRamirez.} UNION  
{sirp:CarolinaMedinaRamirez sirp:has-colleague ?x.} UNION  
{?x sirp:knows sirp:CarolinaMedinaRamirez.} UNION  
{sirp:CarolinaMedinaRamirez sirp:knows ?x.} UNION  
{?x sirp:has-advisor sirp:CarolinaMedinaRamirez.}  
?x sirp:has-name ?name;  
sirp:has-webSite ?ws.
```

*

Transformar las preguntas a una consultas SPARQL III

¿Cuáles son los nombres y sitios Web de las personas que conocen a Carolina Medina Ramírez?



```
PREFIX sirp: <http://arte.izt.uam.mx/ontologies/personRyT.owl#>

SELECT ?name ?ws
WHERE
{
  ?x sirp:knows sirp:CarolinaMedinaRamirez;
    sirp:has-name ?name;
    sirp:has-webSite ?ws.
}
```

Transformar las preguntas a una consultas SPARQL IV

¿Cuáles **documentos** tratan sobre **Ontologías**?



```
PREFIX sird: <http://arte.izt.uam.mx/ontologies/digiResourceRyT.owl#>  
PREFIX redes: <http://mcyti.izt.uam.mx/arios/odaryt.owl#>  
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
SELECT ?x
```

```
WHERE
```

```
{
```

```
{?x rdf:type sird:Paper.} UNION  
{?x rdf:type sird:Book.} UNION  
{?x rdf:type sird:TechnicalReport.} UNION  
{?x rdf:type sird:Thesis.} UNION  
{?x rdf:type sird:Webpage.} UNION  
{?x rdf:type sird:Document.}  
?x sird:has-topic redes:Ontology.
```



```
}
```


Transformar las preguntas a una consultas SPARQL V

¿Cuáles **documentos** tratan sobre **Ontologías**?



```
PREFIX sird: <http://arte.izt.uam.mx/ontologies/digiResourceRyT.owl#>  
PREFIX redes: <http://mcyti.izt.uam.mx/arios/odaryt.owl#>  
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
SELECT ?x  
WHERE  
{  
  ?x rdf:type sird:Document;  
    sird:has-topic redes:Ontology.  
}
```

Funcionamiento de un motor de búsqueda I

Cartografía de Competencias

- 1 Interpreta esta consulta SPARQL dada por un usuario.
- 2 Compara los patrones de la cláusula WHERE con todos los triples en el modelo.
- 3 Recupera la información de las variables resultado.
- 4 Regresa la información de las variables al usuario.

Funcionamiento de un motor de búsqueda II

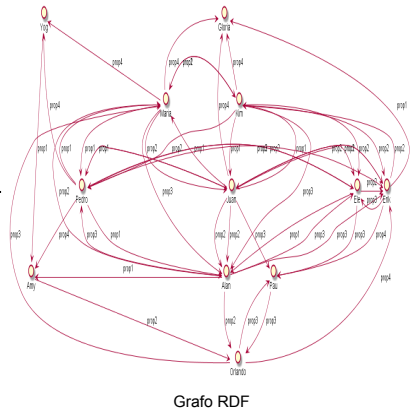
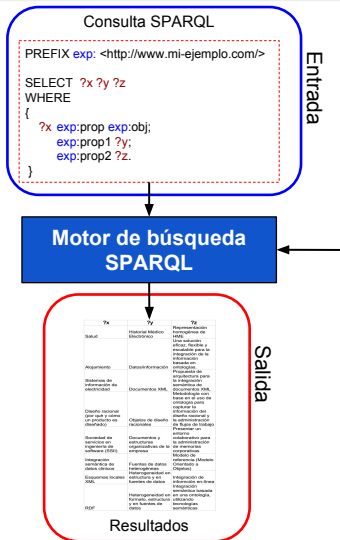


Figura : Grafo RDF sin inferencia



Ejemplo de consulta en un modelo II

```
PREFIX sird: <http://arte.izt.uam.mx/ontologies/digiResourceRyT.owl#>  
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>  
  
SELECT ?x  
WHERE  
{  
  {?x rdf:type sird:Paper.} UNION  
  {?x rdf:type sird:Book.} UNION  
  {?x rdf:type sird:TechnicalReport.} UNION  
  {?x rdf:type sird:Thesis.} UNION  
  {?x rdf:type sird:Webpage.} UNION  
  {?x rdf:type sird:Document.}  
}
```

(a) Consulta sin inferencia

?x
<i>sird:tesis01-pdf</i>
<i>sird:tesis04-pdf</i>
<i>sird:book02-doc</i>
<i>sird:book04-docx</i>
<i>sird:paper01-pdf</i>
<i>sird:paper02-odp</i>

(b) Resultados de la consulta

Ejemplo de consulta en un modelo III

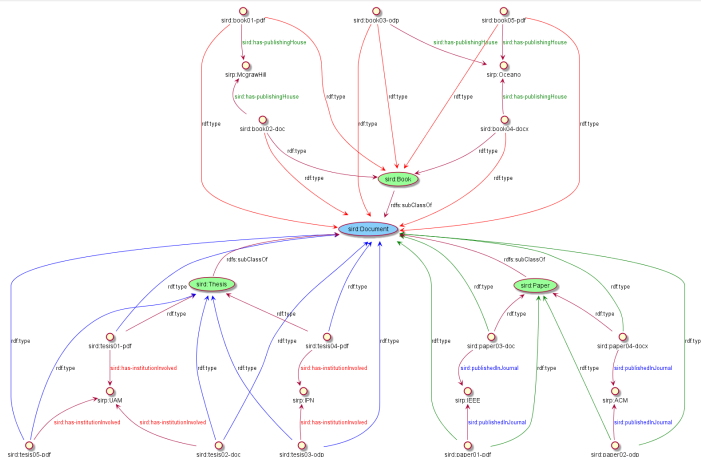


Figura : Grafo RDF con inferencia

Ejemplo de consulta en un modelo IV

```
PREFIX sird: <http://arte.izt.uam.mx/ontologies/digiResourceRyT.owl#>  
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
```

```
SELECT ?x  
WHERE  
{ ?x rdf:type sird:Document. }
```

(a) Consulta con inferencia

?x
<i>sird:tesis01-pdf</i>
<i>sird:tesis02-doc</i>
<i>sird:tesis03-odp</i>
<i>sird:tesis04-pdf</i>
<i>sird:tesis05-pdf</i>
<i>sird:book01-pdf</i>
<i>sird:book02-doc</i>
<i>sird:book03-odp</i>
<i>sird:book04-docx</i>
<i>sird:book05-pdf</i>
<i>sird:paper01-pdf</i>
<i>sird:paper02-odp</i>
<i>sird:paper03-doc</i>
<i>sird:paper04-docx</i>

(b) Resultados de la consulta

Prototipo (Aplicación) I

Dificultad

La búsqueda y recuperación en este modelo no son sencillas, porque se requiere que un usuario tenga un grado especializado de conocimientos en el uso de las tecnologías semánticas y los vocabularios en las ontologías.

Referencias I

- [Alfred et al., 2010] Alfred, S., Arpah, A., Lim, L. H. S., and Sarinder, K. K. S. (2010). Semantic technology: An efficient approach to monogenean information retrieval. In *Computer and Network Technology (ICCNT), 2010 Second International Conference on*, pages 591–594.
- [Bouzid et al., 2012] Bouzid, S., Cauvet, C., and Pinaton, J. (2012). A survey of semantic web standards to representing knowledge in problem solving situations. In *Information Retrieval Knowledge Management (CAMP), 2012 International Conference on*, pages 121–125.
- [Dieng et al., 1998] Dieng, R., Corby, O., Giboin, A., and Ribière, M. (1998). Methods and Tools for Corporate Knowledge Management. Technical Report RR-3485, INRIA.
- [Fujino and Fukuta, 2012] Fujino, T. and Fukuta, N. (2012). A sparql query rewriting approach on heterogeneous ontologies with mapping reliability. In *Advanced Applied Informatics (IIAIAI), 2012 IIAI International Conference on*, pages 230–235.
- [Gandon, 2002] Gandon, Fabien, L. (2002). Ontology Engineering: a Survey and a Return on Experience. Technical Report RR-4396, INRIA.

Referencias II

- [Gruber, 1993] Gruber, T. R. (1993).
A translation approach to portable ontology specifications.
Knowl. Acquis., 5(2):199–220.