



## **Representación semántica de conocimiento operativo y de dominio para repositorios institucionales**

*Dra. María Auxilio Medina Nieto, Dr. J. Alfredo Sánchez*

*Dra. Ofelia D. Cervantes Villagómez,*

*Dra. Reyna Carolina Medina Ramírez*

25 de Abril del 2017

# Representación semántica de conocimiento operativo y de dominio para repositorios institucionales

## Índice

	Página
1. Introducción	3
2. Marco legal	4
3. Audiencia	5
4. Objetivo	6
5. Beneficios del modelo semántico	6
6. Atributos de un repositorio institucional	7
7. Descripción de la ontología onto4AIR	9
7.1 Determinación del dominio y alcance	9
7.2 Reutilización de ontologías existentes	10
7.3 Enumeración de términos importantes	11
7.4 Definición de clases y su jerarquía	13
7.5 Definición de propiedades de las clases	15
7.6 Definición de las facetas de las propiedades	17
7.7 Creación de instancias e individuos	21
8. Métricas de onto4AIR	23
9. Conclusiones y trabajo a futuro	24
10. Referencias	25

# Representación semántica de conocimiento operativo y de dominio para repositorios institucionales

## 1. Introducción

La producción científica y académica en las Instituciones de Educación Superior (IES) se registran en el Curriculum Vitae Único del Consejo Nacional de Ciencia y Tecnología (CONACYT) o en el sistema de información del Programa de Desarrollo Profesional Docente (PRODEP). Algunas IES almacenan parte de esta producción en un repositorio institucional (RI), definida como una “plataforma digital centralizada que sigue estándares internacionales de interoperabilidad, almacena, mantiene, preserva y disemina la información científica, tecnológica y de innovación, la cual se deriva de las investigaciones, productos educativos, académicos y desarrollos tecnológicos [CONACYT 2016]”.

Para las IES, contar con un RI ofrece beneficios como preservación de la producción académica y científica, integración de requisitos para estandarizar, diseminar e interoperar con otros RIs, incremento en la visibilidad de la institución y de los autores, establecimiento de políticas de distribución de acceso abierto para los documentos y los metadatos, acceso a texto completo de documentos en formato digital del propio RI u otros. De acuerdo con [Suber 2010], la literatura en acceso abierto (AA), que es la que se almacena regularmente en un RI, se caracteriza por ser digital, en línea, libre de cargo y de las restricciones que imponen las licencias y los derechos de autor.

Este documento propone una representación semántica del conocimiento operativo y de dominio para RIs conforme a los lineamientos generales [SIICYT 2017a] y técnicos [SIICYT 2017b] del Repositorio Nacional y los Repositorios Institucionales. Los lineamientos técnicos tienen como objetivo “establecer los mecanismos y las acciones a seguir para la preservación del patrimonio académico, científico, tecnológico y de innovación, que cumplan con el proceso de evaluación por pares; así como de sus datos primarios relevantes [SIICYT 2017b]”.

“Los lineamientos técnicos se aplican al Sistema Nacional de Ciencia, Tecnología e Innovación (SNCTI) y a todo aquel que lleve a cabo producción académica, científica,

tecnológica y de innovación, financiado total o parcialmente con recursos públicos o haya utilizado infraestructura pública [SIICYT 2017b]”; éstos se resumen como sigue:

1. Adecuación de la iniciativa de Acceso Abierto de Budapest, 10 años después
2. Estándares internacionales que dan soporte al esquema de metadatos Dublin Core, las directrices OpenAIRE<sup>1</sup> 3.0 para repositorios de literatura y para archivos de datos
3. Exposición de recursos con el protocolo OAI-PMH<sup>2</sup>
4. Requerimientos del certificado DINI<sup>3</sup>
5. El documento “síntesis de reglas de metadatos LA-Referencia”

La representación semántica en este documento es una ontología denominada onto4AIR, la cual integra las definiciones de los términos de [SIICYT 2017b] y otros de la plataforma de código abierto Dspace; a la fecha de elaboración de este documento, representa la más utilizada por los RIs registrados en la Red Mexicana de Repositorios Institucionales (REMEDI) [REMEDI 2016].

Una ontología es “una especificación explícita y formal de una conceptualización compartida [Gruber 1993]”. Previamente, [Neches et al. 1991] utilizaron el concepto para hacer referencia a los términos básicos y las relaciones de un área, tema o dominio de interés que incluye las reglas para combinar los términos y las relaciones entre sí. [Studer et al., 1998] consideran que los elementos de una ontología son: 1) clases o conceptos, 2) relaciones taxonómicas, 3) relaciones entre individuos y 4) reglas (axiomas y restricciones entre los individuos).

Onto4AIR forma parte de la estrategia denominada *Linked Open Data for all institutional repositories (LOD4AIR)*, descrita en [Medina et al., 2017]; constituye un vocabulario controlado, una representación de conocimiento formal, no ambigua de los términos y sus relaciones que puede compartirse y reutilizarse entre usuarios y computadoras. Los conceptos tratan temáticas como tipos de usuario, organización de contenidos, políticas

---

<sup>1</sup> El significado de las siglas OpenAIRE corresponde a la expresión *Open Access Infrastructure for Research in Europe*

<sup>2</sup> Siglas de *Open Archives Initiative Protocol for Metadata Harvesting*

<sup>3</sup> En lugar del certificado alemán DINI, en México se emplea OpenAIRE

de distribución de acceso abierto, reglas de operación, procedimientos de interoperabilidad y mantenimiento, así como términos técnicos definidos en [SIICYT 2017b].

El documento está organizado como sigue. La Sección 2 presenta el marco legal de onto4AIR; la audiencia y el objetivo del documento se incluyen en la sección 3 y 4, respectivamente. Los beneficios de onto4AIR en términos de expresividad de un modelo se listan en la sección 5. La representación de un RI que emplea el modelo a partir de un conjunto de atributos se incluye en la sección 6. La sección 7 describe los componentes de onto4AIR. Finalmente, la sección 8 presenta las conclusiones y el trabajo a futuro.

## 2. Marco legal

El marco legal para la “representación semántica de conocimiento operativo y de dominio para RIs”, adopta los documentos citados en la “Convocatoria 2016 para desarrollar RIs de acceso abierto a la información científica, tecnológica y de innovación”; éste se resume como sigue:

- *Plan Nacional de Desarrollo 2013 – 2018.*- Establece una estrategia que contribuye a fortalecer la infraestructura científica y tecnológica del país con líneas de acción para apoyar su incremento en el sistema de centros públicos de investigación y en las instituciones públicas de investigación científica y tecnológica, a nivel estatal y regional.
- *Ley de Ciencia y Tecnología. Capítulo X.*- Instruye al CONACYT a expedir las disposiciones para asegurar la implementación de la política de acceso abierto a la información científica, tecnológica y de innovación.
- *Programa Especial de Ciencia, Tecnología e Innovación. Estrategia 5.3. Línea de acción 5.3.1.*- Se enfoca a la creación de repositorios de acceso abierto estandarizados para las IES y los Centros Públicos de Investigación (CI).

## 3. Audiencia

Este documento está dirigido a usuarios con conocimientos técnicos en ontologías y tecnologías de la web semántica, relacionados con la producción de documentos científicos y académicos.

## 4. Objetivo

El objetivo de este documento es apoyar a los usuarios en la apropiación del conocimiento operativo y de dominio en IES sobre RIs que satisfacen los lineamientos generales y técnicos del Repositorio Nacional y Repositorios Institucionales, mediante la representación de conceptos en un modelo semántico u ontología.

## 5. Beneficios del modelo semántico

Onto4AIR, como vocabulario controlado, es una representación de conocimiento formal, no ambigua de los términos y sus relaciones que puede compartirse y reutilizarse entre usuarios y computadoras. En comparación con un modelo de datos relacional, en donde se definen entidades, relaciones y objetos, un modelo semántico tiene mayor expresividad, por ejemplo:

- La representación de propiedades incluye herencia, composición o asociación
- Junto con las restricciones de integridad, se modelan restricciones de cardinalidad a las que se les asocia un valor máximo o mínimo o la propiedad exactamente igual
- Las relaciones entre instancias o tipos de datos, es decir, propiedades de datos y de objeto, pueden tener asociado un dominio y rango, así como valores fijos
- Se pueden definir reglas que utilicen cuantificadores universales o existenciales, éstas se utilizan en la validación de consistencia
- Es factible tener instancias que pertenezcan a una, ninguna o más clases si las últimas no son disjuntas
- Definición única para el uso de los metadatos Dublin Core no calificado en la versión 2.0 del protocolo OAI-PMH

A corto plazo, Onto4AIR pretende apoyar la apropiación del conocimiento operativo y de dominio, así como la implantación de los RIs en las IES; a mediano plazo, su adopción amplía la posibilidad de revisar de manera semi-automática la integridad y consistencia de los datos, lo cual representaría un aumento en su valor y veracidad.

## 6. Atributos de un repositorio institucional

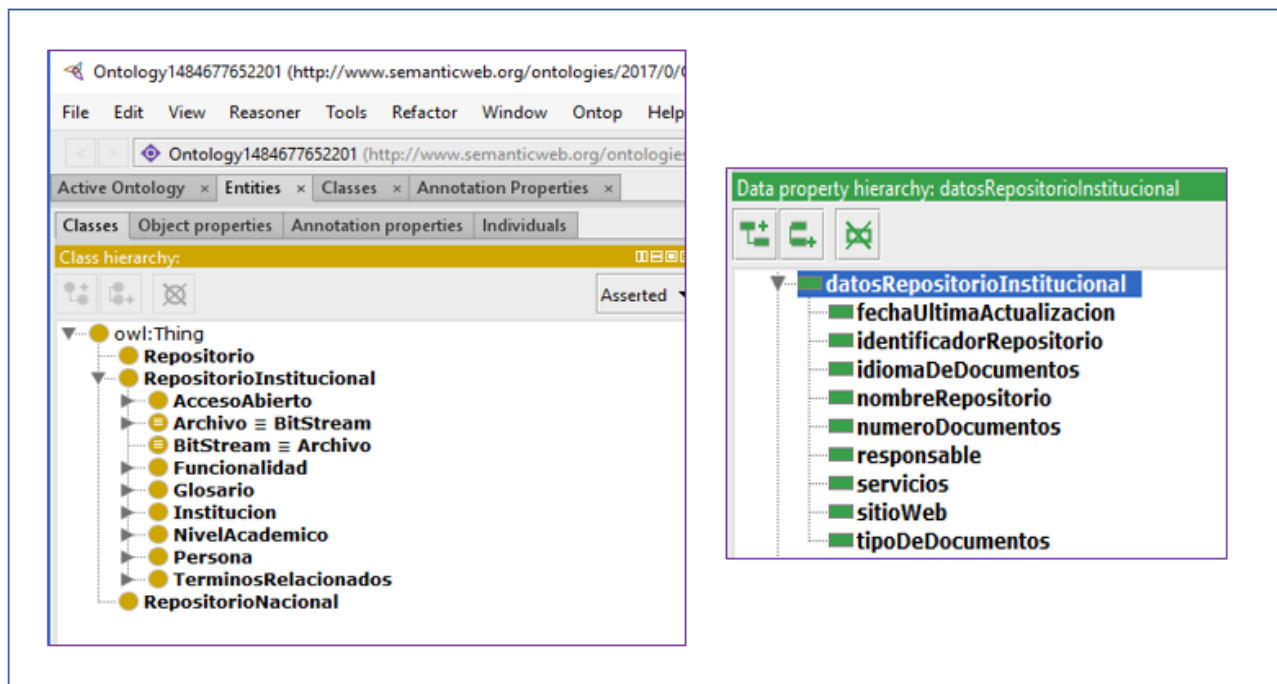
La Tabla 1 propone una lista de atributos (o propiedades) que describen a un RI en Onto4AIR. De acuerdo con [SIICYT 2017a] y [SIICYT 2017b], un atributo es *mandatorio* si es un requisito incluirlo, *recomendado* si aporta información deseable u *opcional* cuando puede o no aparecer.

**Tabla 1.** Descripción de un RI.

Atributo	Descripción	Tipo de atributo
fechaUltimaActualizacion	Fecha de la última actualización de los datos	Mandatorio
identificadorRepositorio	Cadena que distingue a un RI de otro en la misma institución	Mandatorio
idiomaDeDocumentos	Abreviación que indica los idiomas en los que están escritos los documentos, por ejemplo, español: es, inglés: en.	Recomendado
nombreRepositorio	Cadena que contiene el nombre del repositorio, éste aparecerá en página de inicio.	Mandatorio
numeroDocumentos	Número de documentos almacenados.	Recomendado
responsable	Persona encargada del RI.	Mandatorio
servicios	Lista de servicios, por ejemplo: <ul style="list-style-type: none"> <li>• Búsqueda por autor, fecha, tema y palabras clave de los documentos</li> <li>• Interfaz para subir documentos</li> <li>• Envío de cartas de donación a los autores</li> </ul>	Opcional
sitioWeb	URL que contiene la dirección web de la página de inicio	Mandatorio
tipoDeDocumentos	Enumeración (lista) que indica el tipo de documentos que almacena el RI, por ejemplo, reportes técnicos, artículos, carteles o tesis.	Recomendado

Los elementos principales de un RI se modelan en onto4AIR mediante clases, (ver la Figura 1 de la izquierda). La clase `Repositorio` se incluye como parte de los términos de referencia, es análoga a la clase `RepositorioInstitucional`, pero no se definen como equivalentes para que se puedan incluir otro tipo de repositorios si el usuario así lo desea. Observe también que la clase `RepositorioNacional`, es también un `RepositorioInstitucional`, sin embargo, se modeló como una clase independiente para hacer énfasis que de acuerdo con los lineamientos técnicos, cualquier RI requerirá interoperar con el `RepositorioNacional`.

Del lado derecho de la Figura 1, se muestran los atributos de la Tabla 1, éstos se modelan como sub-propiedades de datos de la propiedad `datosRepositorioInstitucional`.



**Figura 1.** Sub-clases y atributos de `RepositorioInstitucional`.

En onto4AIR se asume lo siguiente:

1. La calidad de los documentos está respaldada por la *revisión por pares* (método de validación de documentos científicos escritos). Generalmente, los evaluadores son profesores o investigadores de las IES.
2. La distribución de los documentos se realiza mediante la “ruta dorada”, esto implica que las políticas de acceso abierto se asignan inmediatamente después de que los documentos se publican en el RI.

## 7. Descripción de Onto4AIR

La construcción de la ontología, se llevó a cabo siguiendo la metodología 101, propuesta en [Noy y McGuinness 2001], la cual está compuesta por los pasos siguientes:

1. Determinar el dominio y alcance
2. Reutilizar ontologías existentes
3. Enumerar términos importantes
4. Definir clases y su jerarquía
5. Definir las propiedades de las clases, es decir, describir la estructura interna de los conceptos
6. Definir las características (o facetas) de las propiedades
7. Crear instancias o individuos

La ontología de este documento está editada en la versión 5.1.0 de Protégé, en una computadora con Windows, por tanto, la implementación del paso 5 implica la definición de propiedades de datos y de objeto, en tanto, el paso 6, se refiere a seleccionar las facetas (características) de las propiedades, las cuales son: a) funcional, b) simétrica, c) inversa funcional, d) asimétrica, e) reflexiva e f) irreflexiva. Las secciones siguientes describen cómo se han implementado estos pasos en onto4AIR.

### 7.1 Determinación del dominio y alcance

Onto4AIR modela un RI conforme a la definición propuesta en [SIICYT 2017a] y [CONACYT 2016], y su operación en una IES, incluye las definiciones de los lineamientos técnicos [SIICYT 2017b] y términos de la plataforma *Dspace*. Su alcance comprende la identificación de los tipos de usuario, la organización de los contenidos, la descripción de políticas y la operatividad esperada. La Tabla 2 muestra las preguntas de competencia a las que responde la ontología.

**Tabla 2.** Preguntas de competencia para onto4AIR.

Elemento	Preguntas de competencia
Organización de contenidos	¿Qué tipo de documentos almacena el RI? ¿Cómo se organizan los documentos? ¿En qué formato están? ¿Cuáles son los datos que describen a cada tipo de documento? (metadatos) ¿Los metadatos pasan por un proceso de curación?
Políticas	¿Qué es el acceso abierto? ¿Qué tipo de licencia de acceso abierto tienen los documentos? ¿Qué tipo de licencia de acceso abierto tienen los metadatos?
Tipos de usuarios	¿Quiénes utilizan el RI? ¿Quién o quiénes son los propietarios de los documentos? ¿Los documentos pasan por un proceso de revisión? ¿Quién es el responsable del RI?
Operatividad	¿Qué tareas pueden hacer los usuarios en el RI? ¿Quiénes intervienen en la elaboración de un documento? ¿Con qué frecuencia se respaldará el RI?

## 7.2 Reutilización de ontologías existentes

La Tabla 3 muestra los vocabularios requeridos por la versión 2.0 del protocolo OAI-PMH, el cual representa el mecanismo de interoperabilidad en [SIICYT 2017b] y de la red REMERI. La diseminación y cosecha de metadatos en los RIs se basa en este protocolo.

El uso de vocabularios controlados u ontologías como los de las Tabla 4 y 5, podría enriquecer el valor de los datos en los RIs, estos vocabularios describen a los usuarios y sus relaciones, un modelo de datos o valores específicos para metadatos en OAI-PMH.

**Tabla 3.** Vocabularios requeridos para implementar la versión 2.0 del protocolo OAI-PMH

Prefijo	Nombre de la ontología
dc	Dublin Core
dcterms	Dublin Core terms

**Tabla 4.** Ontologías utilizadas frecuentemente en aplicaciones de la web semántica

Prefijo	Nombre de la ontología
foaf	Friend of a friend (FOAF)
rdfs	RDF schema

**Tabla 5.** Ejemplo de ontologías para elementos Dublin Core [DUBLIN CORE 2008].

<b>Término DC</b>	<b>Nombre de la ontología</b>
dc:subject	Sistema de clasificación de la ACM (ACM CCS 2012).
dc:language	Lexvo (pertenece a <i>LOD Cloud</i> , <a href="http://lod-cloud.net/">http://lod-cloud.net/</a> )
dc:creator	FOAF, ORCID <sup>4</sup>

### 7.3 Enumeración de términos importantes

Las Tablas 6 y 7 muestran las clases principales para un RI. Las definiciones de las clases se modelan con la propiedad `rdfs:isDefinedBy`, el valor es de tipo `rdf:PlainLiteral` y el idioma es español (`es`). La definición de algunas de las clases proviene de [SIICYT 2017b]; se utilizó `rdfs:comment` para separar la definiciones de los comentarios. La tercer columna muestra los términos relacionados con las clases principales o sub - clases.

**Tabla 7.** Clases principales de Onto4AIR. Parte 1.

<b>Clase principal</b>	<b>Definición</b>	<b>Conjunto de términos</b>
AccesoAbierto	En este tipo de acceso, cualquier persona con conexión a internet, puede tener acceso sin restricciones económicas, técnicas o legales a información científica, académica, de desarrollo tecnológico, innovación o cultural producida con recursos públicos o que en su producción se usó infraestructura pública.	Acceso a recursos de información académica, científica, tecnológica y de innovación, estándares internacionales, licencia, <i>Creative Commons</i> , mandao institucional, openAIRE, política, política de acceso, política de comunicación pública, de distribución y mantenimiento.
Archivo	Recurso de información que se almacena en Dspace.	<i>Bitstream</i> , colección, metadato, recursos de información académica, científica, tecnológica y de innovación, producto del desarrollo y la innovación, publicación científica, comunidad, datos de las investigaciones.

<sup>4</sup> Para mayor información sobre el identificador único para investigadores, consultar: <https://orcid.org>

**Tabla 7.** Clases principales de Onto4AIR. Parte 2.

<b>Clase principal</b>	<b>Definición</b>	<b>Conjunto de términos</b>
Funcionalidad	Agrupar las tareas que realizan los usuarios del RI.	Almacenamiento, catalogación, catálogo CONACYT, diseminación, interoperabilidad, preservación, recuperación.
Glosario	Conjunto de definiciones de lineamientos técnicos que no se asocian con las clases principales del RI.	API, comité, SIICYT, SNCTI, subcomité.
Institución	Organismo público o privado fundado para desempeñar una labor científica o académica.	Centro de investigación, CONACYT, institución de educación superior, cuerpo académico, programa académico, LA-Referencia.
NivelAcademico	Niveles o grados en instituciones de educación superior.	Diplomado, doctorado, especialidad, licenciatura, maestría.
Persona	Es un individuo de la especie humana.	Autor, usuario, académico, administrador, director, rector, administrativo, alumno, discente, estudiante, depositario,
TerminosRelacionados	Conjunto de términos y siglas que no se relacionan directamente con la operatividad del RI, pero que se encuentran en los lineamientos técnicos.	Almetría, embargo, enmiendas de autor, evaluación por pares, ResearchGate, SRU, SWORD, XML.

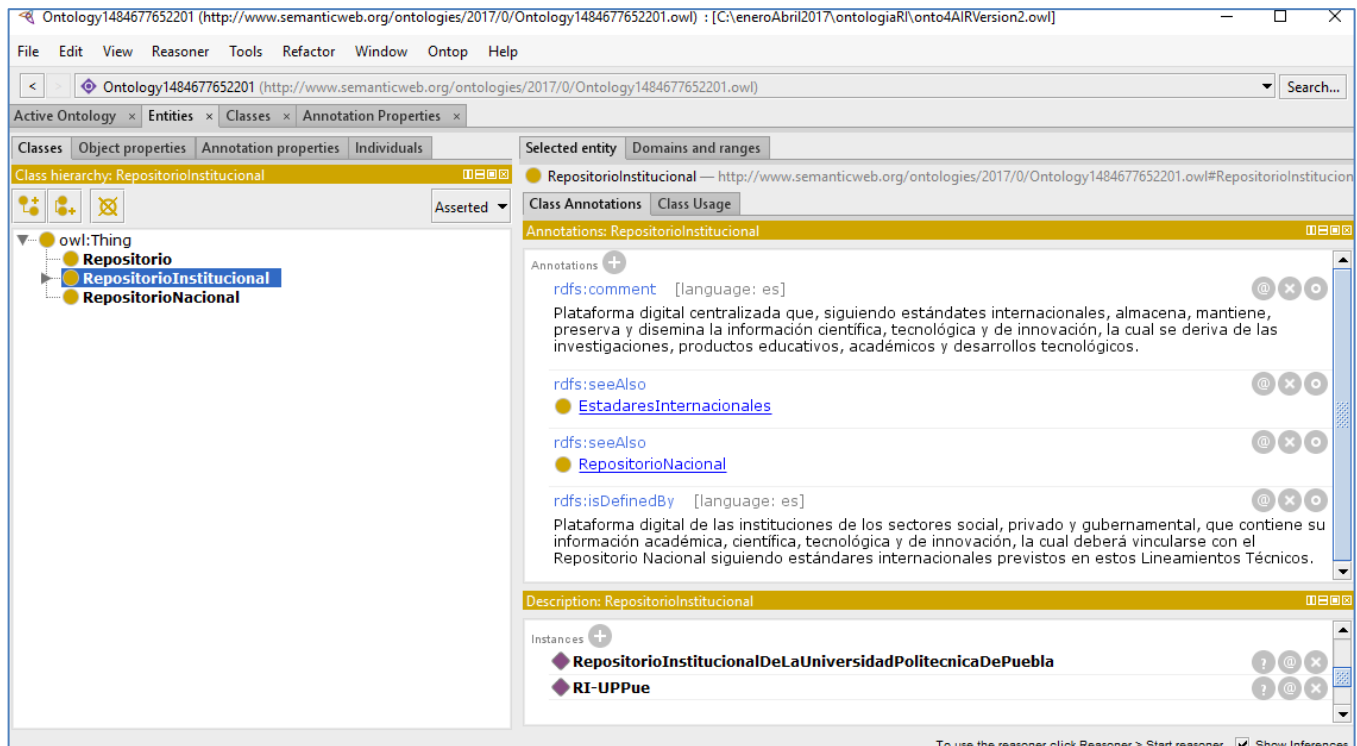
Las clases de las tablas 6 y 7 son disjuntas, lo cual implica que las instancias sólo podrán pertenecer a una de estas clases.

Por convención, la notación propuesta para los elementos de onto4AIR es:

- *Clases.* Términos en singular, inician con mayúscula, generalmente son sustantivos. Si están formados por varias palabras, se utiliza la primer letra en mayúscula como separador.
- *Propiedades.* Términos que inician con minúscula, usualmente contienen un verbo. Se emplean mayúsculas como separadores.

## 7.4 Definición de clases y su jerarquía

La Figura 2 muestra la clase principal de la ontología, `RepositorioInstitucional`; en el marco de la derecha, se muestra la definición, un comentario y dos objetos relacionados.



**Figura 2.** Descripción de la clase principal: `RepositorioInstitucional`.

Las Figura 3 y 4 muestran las subclases de `RepositorioInstitucional` que corresponden a las clases de las Tablas 6 y 7.

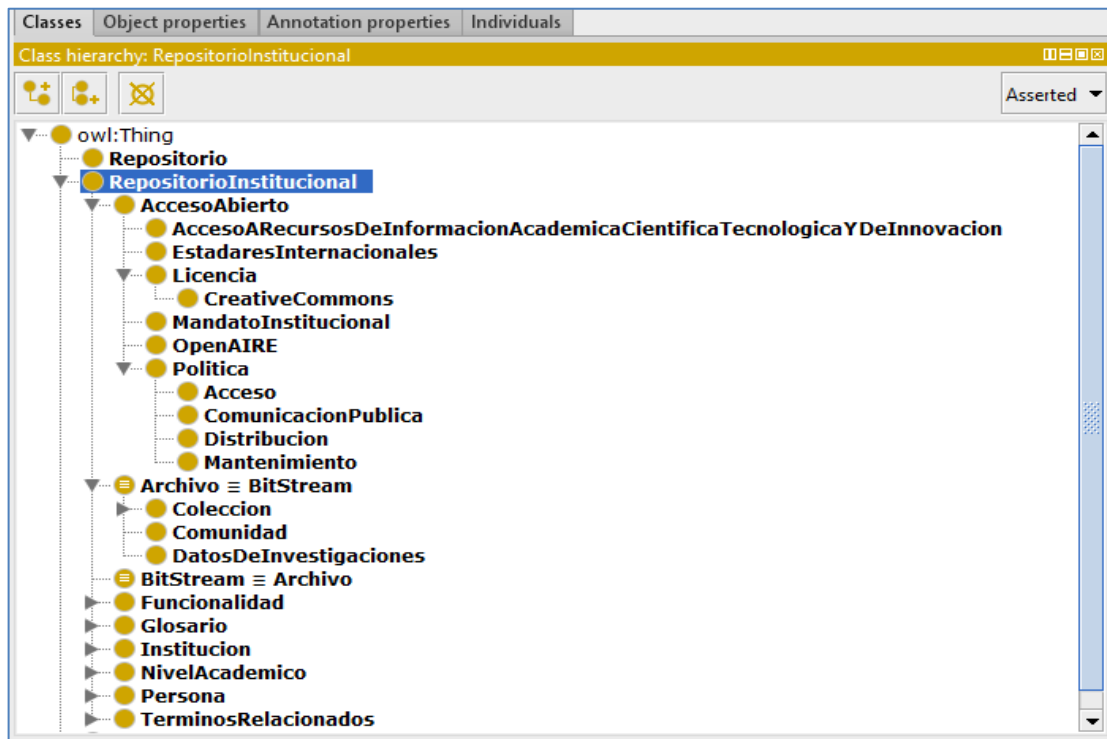


Figura 3. Jerarquía de clases de onto4AIR. Parte 1.

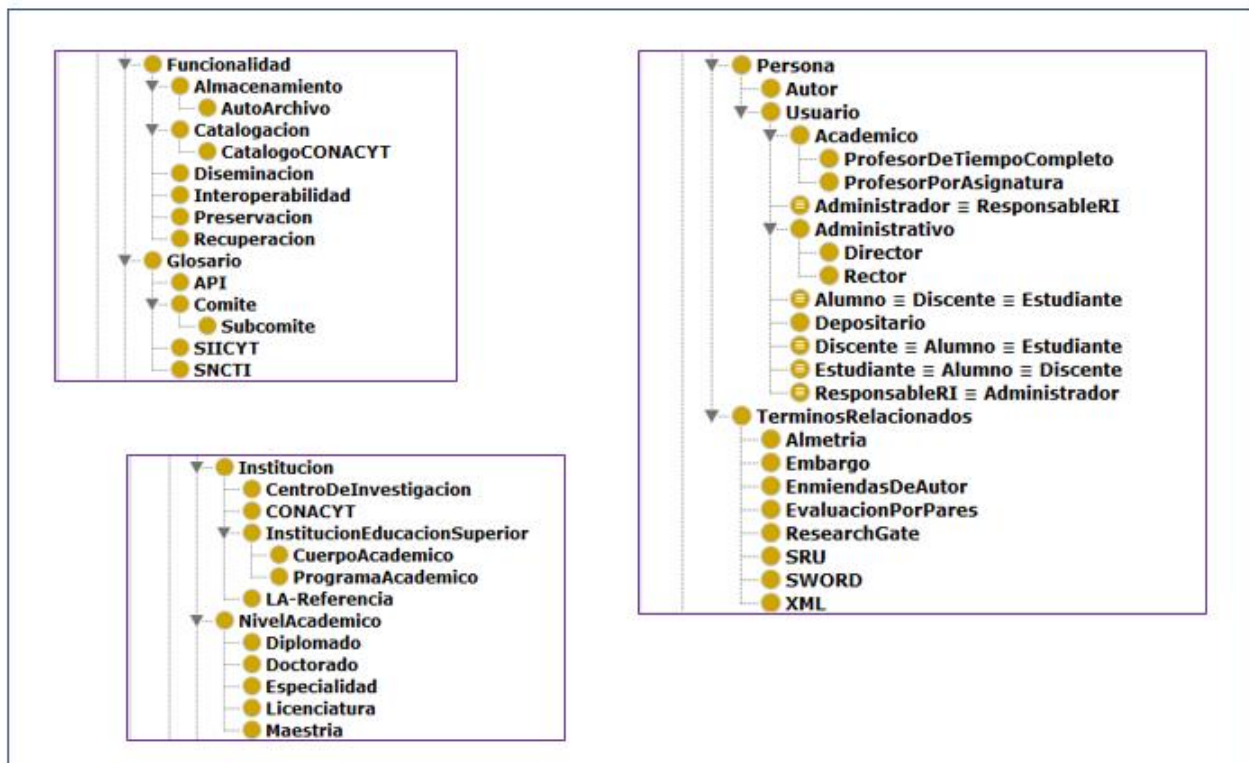


Figura 4. Jerarquía de clases de onto4AIR. Parte 2.

En la ontología, el uso de sinónimos se modela mediante la relación de equivalencia entre clases. Por ejemplo, una instancia de la clase `Estudiante`, es también instancia de la clase `Alumno` y `Discente`. Esta relación se utiliza también para incluir términos en otro idioma como `BitStream`, equivalente a `Archivo`, se refiere a cualquier tipo de documento que se pueda procesar en `Dspace`. Otras características de la jerarquía son:

- Un conjunto de `Archivos` o `BitStream` se almacenan en una `Colección`, ésta contiene metadatos y recursos de información académica, científica, tecnológica y de innovación de dos tipos principales: 1) producto del desarrollo tecnológico y la innovación y 2) publicación científica.
- Se consideran publicaciones científicas a artículos, capítulos de libro, disertaciones, libros, memorias de congreso, reportes de investigación, revista de acceso abierto y tesis de posgrado.
- `ResponsableRI` es el usuario con todos los privilegios para mantener a un RI, es equivalente a un `Administrador`; coincide con el súper-usuario de la base de datos de `Dspace`.

*Observación:* la políticas de acceso pueden estar limitadas por la definición de usuarios y permisos de la plataforma tecnológica que soporte al RI.

## 7.5 Definición de propiedades de las clases

En el nivel 2 de `onto4AIR`, a diferencia de las clases `AccesoAbierto`, `Funcionalidad`, `Glosario`, `NivelAcademico` y `TerminosRelacionados`, `Archivo`, `Institución` y `Persona` tienen un conjunto de atributos que se modelan como *propiedades de datos*. Estas propiedades se caracterizan porque el rango que corresponde a un tipo de datos definido en `XML Schema`; son funcionales si su valor es único. La Figura 5 muestra las propiedades de datos para `Archivo`, `Institución` y `Persona`, éstas utilizan el prefijo `datos` seguido del nombre de la clase. Observe que junto con estas clases, `onto4AIR` agrupa propiedades de datos de sus sub-clases.

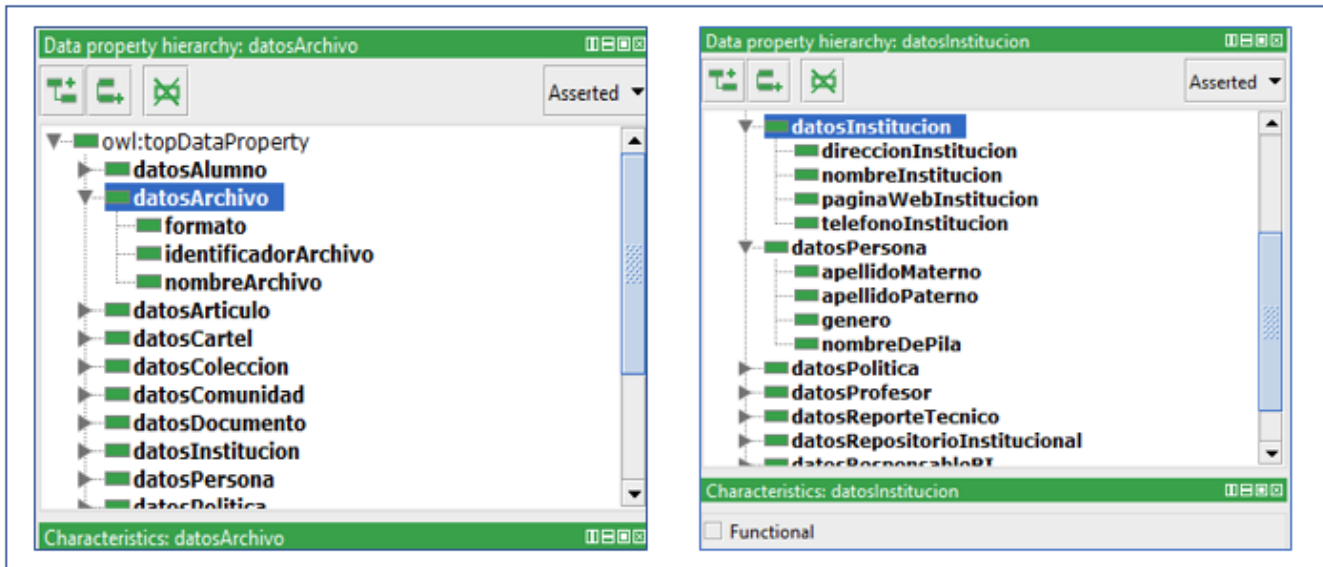


Figura 5. Propiedades de datos de clases en el nivel 2 de onto4AIR.

Las propiedades de datos se asocian con dominio y rango, se indica también si admiten un solo valor o más señalando la casilla de Funcionalidad como se muestra en la Figura 5. Notar que las filas en donde el valor de rango no tiene un tipo de datos, corresponde a las propiedades de datos que agrupan a propiedades específicas para las clases.

Propiedad	Funcional	Clase	Rango
datosRepositorioInstitucional	<input type="checkbox"/>	RepositorioInstitucional	
fechaUltimaActualizacion	<input checked="" type="checkbox"/>	RepositorioInstitucional	xsd:dateTime
nombreRepositorio	<input checked="" type="checkbox"/>	RepositorioInstitucional	xsd:string
servicios	<input type="checkbox"/>		xsd:string
idiomaDeDocumentos	<input type="checkbox"/>	RepositorioInstitucional	xsd:string
responsable	<input type="checkbox"/>		
identificadorRepositorio	<input checked="" type="checkbox"/>	RepositorioInstitucional	xsd:string
sitioWeb	<input checked="" type="checkbox"/>		xsd:anyURI
numeroDocumentos	<input checked="" type="checkbox"/>	RepositorioInstitucional	xsd:integer
tipoDeDocumentos	<input checked="" type="checkbox"/>		xsd:string
datosArchivo	<input type="checkbox"/>	Archivo	
nombreArchivo	<input checked="" type="checkbox"/>	Archivo	xsd:string
identificadorArchivo	<input checked="" type="checkbox"/>	Archivo	xsd:anyURI
formato	<input checked="" type="checkbox"/>	Archivo	xsd:string

Figura 7. Ejemplo de dominio, rango y característica de funcionalidad para algunas propiedades de datos.

## 7.6 Definición de las facetas de las propiedades

La Figura 8 muestra las propiedades de objeto (o predicados), definidas como la relación entre dos instancias (conocidas también como individuos u objetos). Los nombres de las propiedades están del lado izquierdo, en tanto las facetas se muestran en las columnas que contienen cajas de selección. Si la caja está marcada, se asume que la propiedad cumple con la faceta correspondiente.

Dado que las facetas representan relaciones binarias, en términos matemáticos, las facetas pueden satisfacer o no algunas de las propiedades siguientes: funcional (Func), simétrica (Sym), inversa funcional (Inv Func), transitiva (Trans), asimétrica (ASym), reflexiva (Refl) e irreflexiva (Irrefl).

Object Property	Func	Sym	Inv Func	Trans	ASym	Refl	Irrefl
owl:topObjectProperty	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
perteneceA	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
adscritoA	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
esServicioDe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
soportaServicioDe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
esSinodalDe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
tieneAlumno	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
esEstudianteDe	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
tieneRolDe	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
tieneIntegrante	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
administra	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
parteDe	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
esAutorDe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
esAsesorDe	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>
incluyeColeccion	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
contieneA	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
esAdministradoPor	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
tieneSinodal	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
esIntegranteDe	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>

Figura 8. Facetas de las propiedades de objeto

La propiedad `esAutorDe` se divide en dos sub-propiedades: `esPrimerAutorDe` y `esCoautorDe`, con el propósito de identificar al primer autor de cada documento. Estas subpropiedades se declaran disjuntas.

Las Figuras 9 y 10 muestran las clases que corresponden al dominio y rango de cada una de las propiedades de datos, mismas que están representadas como flechas. Notar que en las propiedades `tieneSinodal` y `perteneceA`, el dominio es la unión de dos clases.

Observación: Maestrías e Ingenierías representan en la Figura 9 los nombres de dos comunidades.

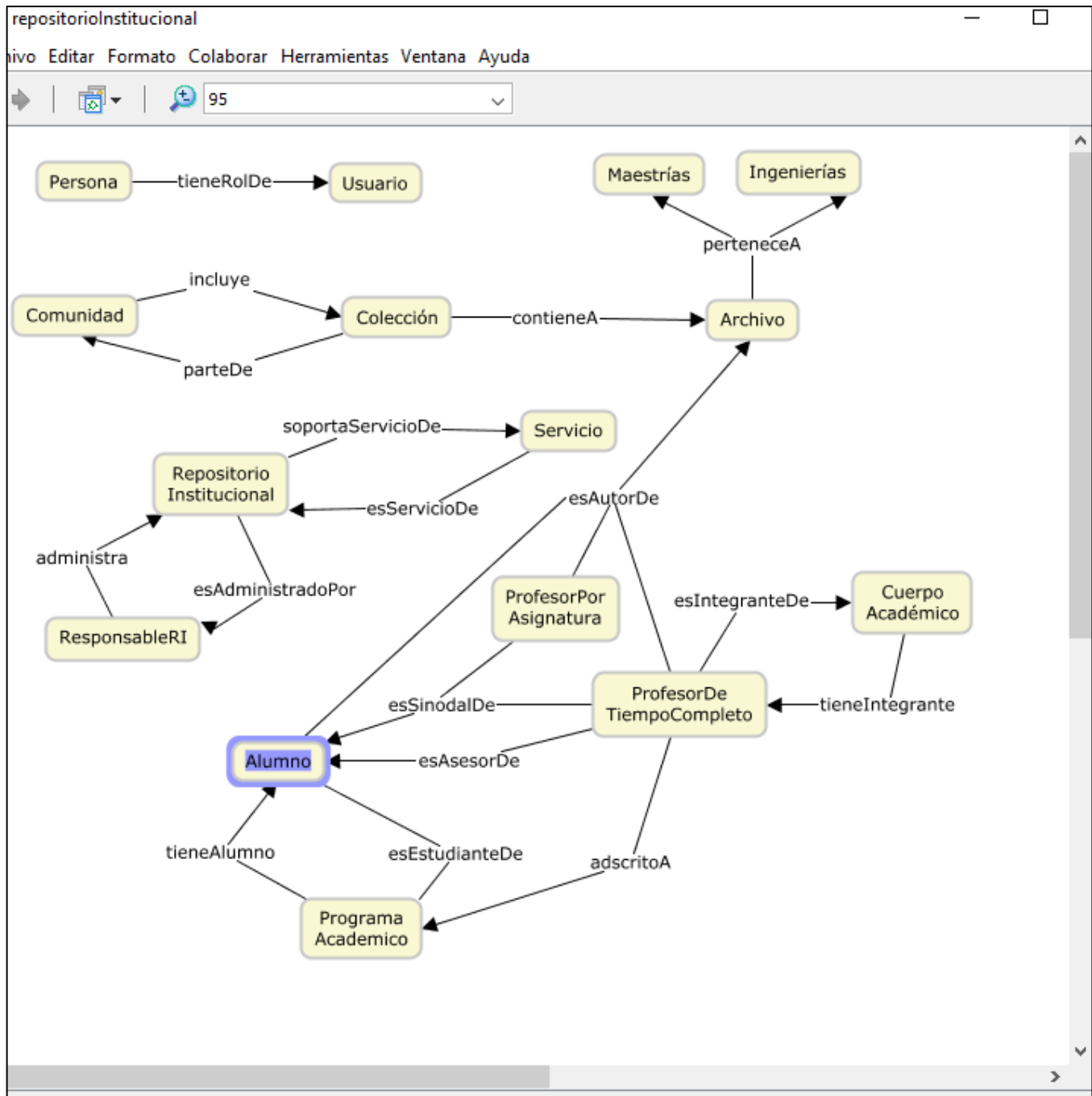


Figura 9. Dominio y rango para propiedades de objeto.

Object Property	Domain	Range	Inverse
<input checked="" type="checkbox"/> perteneceA	<input checked="" type="checkbox"/> Archivo	Ingenierías or Maestrías	
<input checked="" type="checkbox"/> adscritoA	<input checked="" type="checkbox"/> ProfesorDeTiempoCompleto	ProgramaAcademico	
<input checked="" type="checkbox"/> esServicioDe	<input checked="" type="checkbox"/> Servicio	RepositorioInstitucional	soportaServicioDe
<input checked="" type="checkbox"/> soportaServicioDe	<input type="checkbox"/> RepositorioInstitucional	Servicio	esServicioDe
<input checked="" type="checkbox"/> tieneAlumno	<input type="checkbox"/> ProgramaAcademico	Alumno	esEstudianteDe
<input checked="" type="checkbox"/> esSinodalDe	<input checked="" type="checkbox"/> ProfesorDeTiempoCompleto or ProfesorPorAsignatura	Alumno	
<input checked="" type="checkbox"/> esEstudianteDe	<input checked="" type="checkbox"/> Alumno	ProgramaAcademico	tieneAlumno
<input checked="" type="checkbox"/> tieneIntegrante	<input type="checkbox"/> CuerpoAcademico	ProfesorDeTiempoCompleto	esIntegranteDe
<input checked="" type="checkbox"/> tieneRolDe	<input checked="" type="checkbox"/> Persona	Usuario	
<input checked="" type="checkbox"/> administra	<input checked="" type="checkbox"/> ResponsableRI	RepositorioInstitucional	esAdministradoPor
<input checked="" type="checkbox"/> esAutorDe	<input checked="" type="checkbox"/> Estudiante or ProfesorDeTiempoCompleto or ProfesorPorAsignatura	Archivo	
<input checked="" type="checkbox"/> esPrimerAutorDe			
<input checked="" type="checkbox"/> esCoautorDe			
<input checked="" type="checkbox"/> parteDe	<input checked="" type="checkbox"/> Coleccion	Comunidad	incluyeColeccion
<input checked="" type="checkbox"/> esAsesorDe	<input checked="" type="checkbox"/> ProfesorDeTiempoCompleto	Alumno	
<input checked="" type="checkbox"/> incluyeColeccion	<input type="checkbox"/> Comunidad	Coleccion	parteDe
<input checked="" type="checkbox"/> contieneA	<input type="checkbox"/> Coleccion	Archivo	
<input checked="" type="checkbox"/> esAdministradoPor	<input type="checkbox"/> RepositorioInstitucional	ResponsableRI	administra
<input checked="" type="checkbox"/> tieneSinodal	<input type="checkbox"/> Alumno	ProfesorDeTiempoCompleto or ProfesorPorAsignatura	
<input checked="" type="checkbox"/> esIntegranteDe	<input checked="" type="checkbox"/> ProfesorDeTiempoCompleto	CuerpoAcademico	tieneIntegrante

Figura 9. Matriz de dominio, rango y propiedades inversas para las propiedades de objeto.

La Figura 10 muestra el grafo generado en la herramienta *OntoGraph*, dependiendo de la selección del usuario, es posible mostrar sólo algunas de las propiedades de objeto.

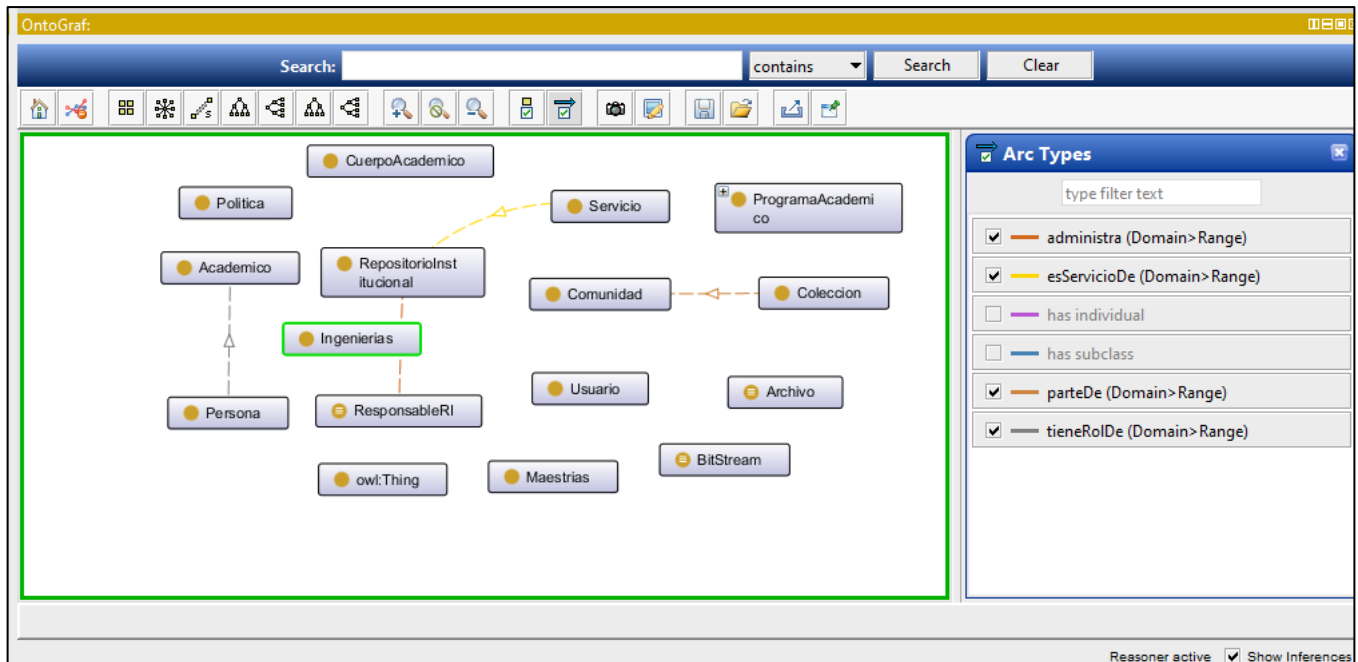
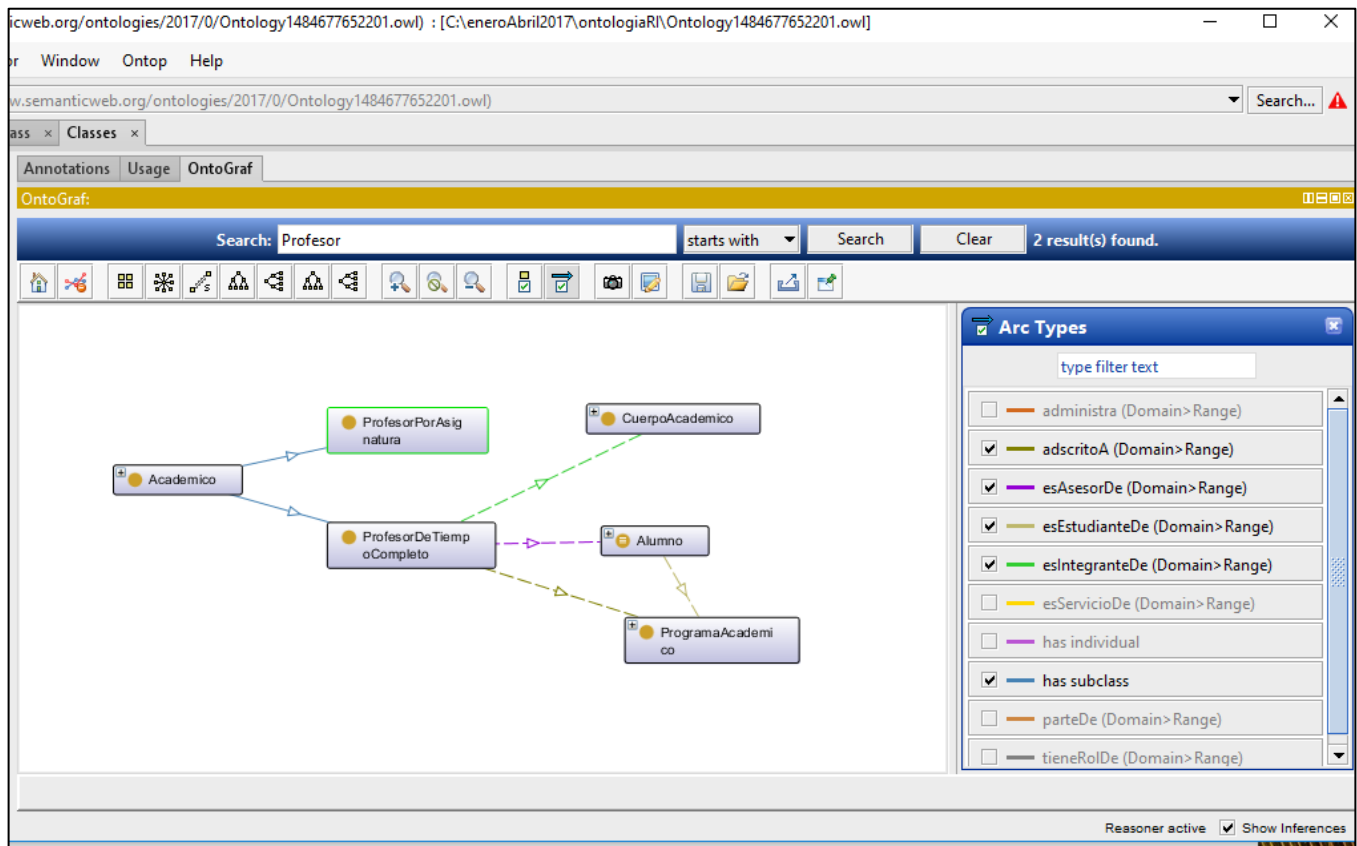


Figura 10. Grafo con propiedades de objeto en *Ontograph*.

*OntoGraph* se emplea también para realizar búsquedas sencillas, por ejemplo, suponga que se desea buscar el término *Profesor*, el resultado se muestra en la Figura 11.



**Figura 11.** Resultado de la búsqueda del término `Profesor` utilizando Ontograh.

La búsqueda de términos se extiende a cualquier elemento de la ontología. Al dar un click sobre una clase o arista del grafo, se despliega un texto que contiene información relacionada con la búsqueda. Por ejemplo, la Figura 12 muestra información del término `Profesor`.

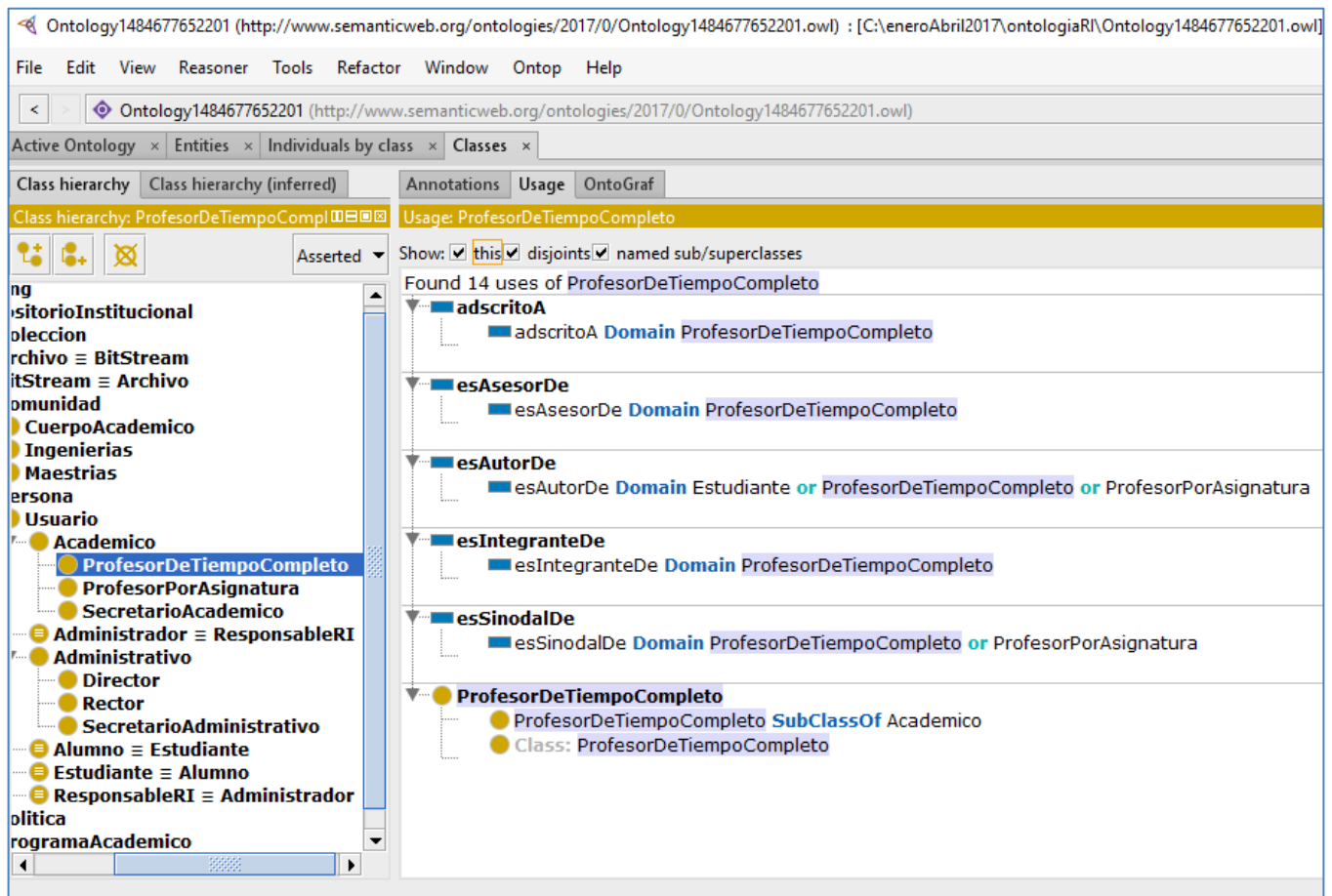


Figura 12. Información relacionada con el término `Profesor`.

## 7.7 Creación de instancias e individuos

La creación de instancias o individuos, tarea conocida también como *población de la ontología*, permitirá a los usuarios finales realizar consultas a onto4AIR. La Figura 13 muestra algunos ejemplos de instancias. La Figura 14, muestra una vista general conceptos y sus relaciones.

The screenshot shows a Semantic Web browser interface. On the left, a list of individuals is displayed, with 'MariaAuxilioMedinaNieto' selected. The main area is divided into two panes: 'Description: MariaAuxilioMedinaNieto' and 'Property assertions: MariaAuxilioMedinaNieto'. The description pane shows the type 'ProfesorDeTiempoCompleto'. The property assertions pane lists several object and data properties, such as 'adscritoA' (pointing to 'MaestriaEnGestionEInnovacionTecnologica') and 'esAsesorDe' (pointing to 'FidelaAlejandraSolisGutierrez' and 'DeserethLopezViveros').

Figura 13. Ejemplos de instancias para la ontología.

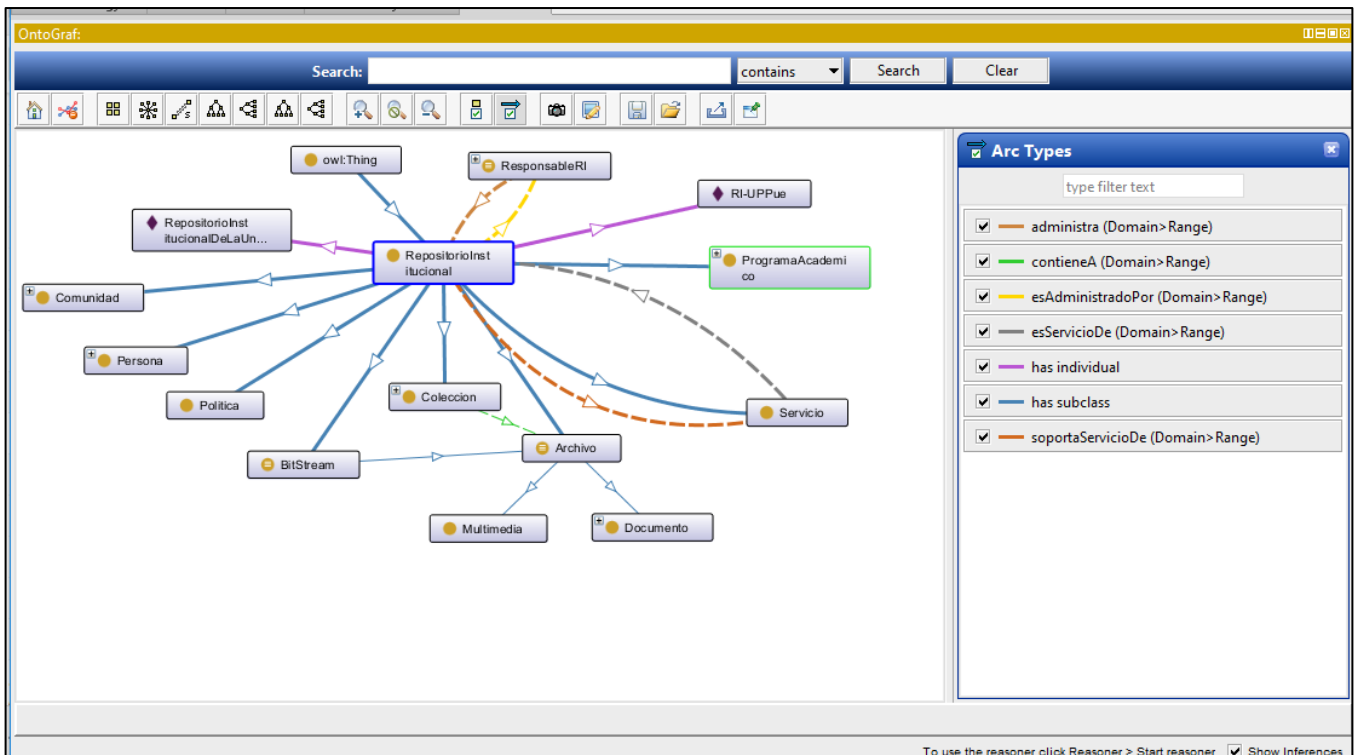
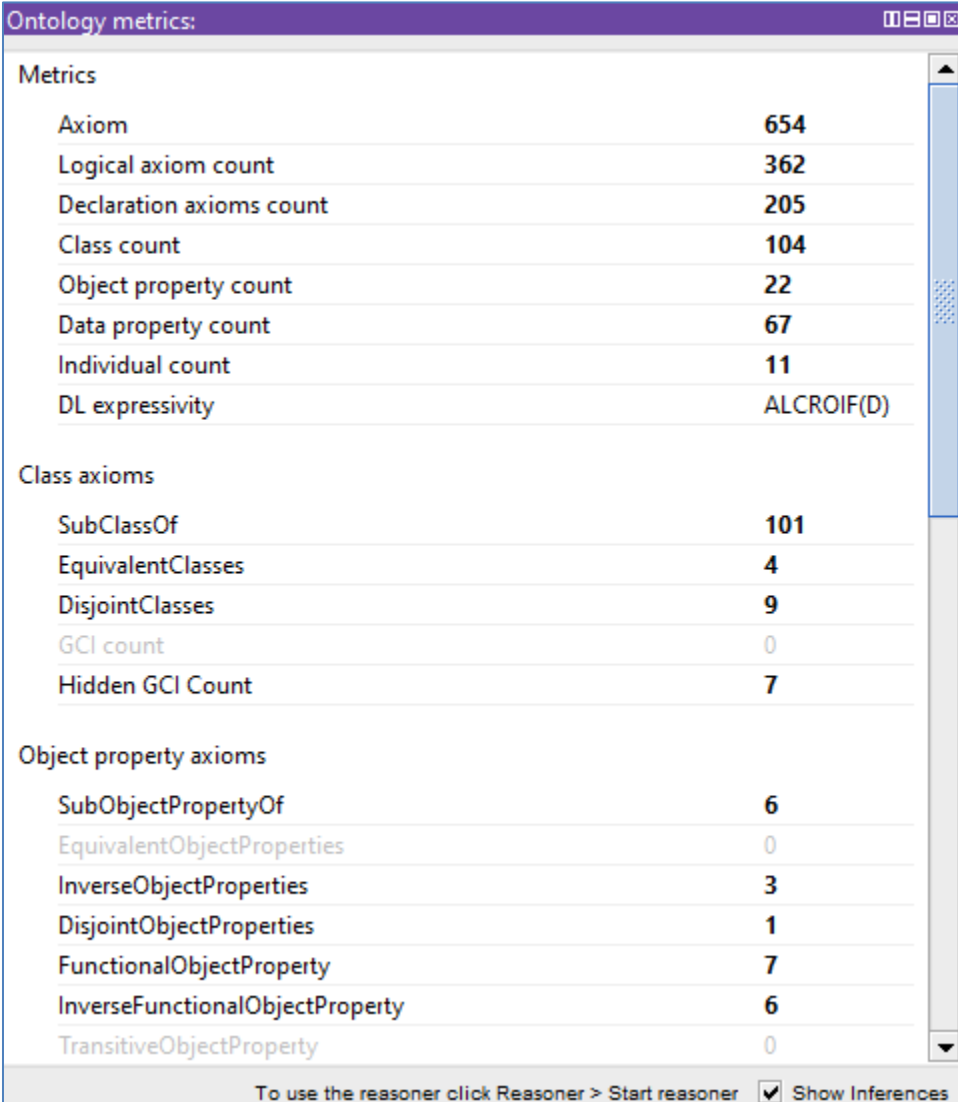


Figura 14. Vista general del repositorio.

## 8. Métricas de onto4AIR

Onto4AIR se forma de la principal, `RepositorioInstitucional` que incluye 9 subclases en el segundo nivel de la jerarquía de clases, mismas que agrupan un total de 104 clases, enlazadas entre sí mediante 22 propiedades de objeto. Cada propiedad de objeto tiene definido rango y dominio, así como un conjunto de facetas. Las Figuras 15 y 16 contienen éstas y otras métricas, que resumen de forma cuantitativa las características de este modelo semántico.



Metrics	
Axiom	654
Logical axiom count	362
Declaration axioms count	205
Class count	104
Object property count	22
Data property count	67
Individual count	11
DL expressivity	ALCROIF(D)

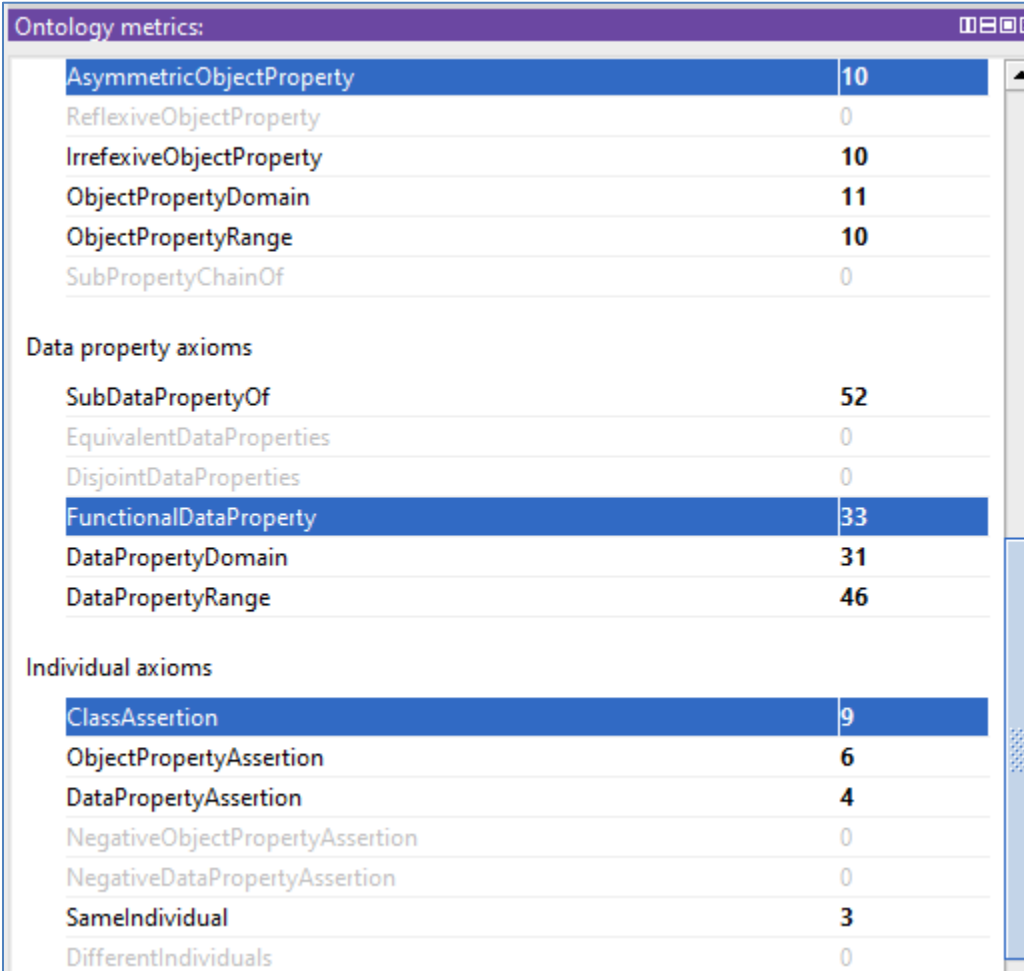
Class axioms	
SubClassOf	101
EquivalentClasses	4
DisjointClasses	9
GCI count	0
Hidden GCI Count	7

Object property axioms	
SubObjectPropertyOf	6
EquivalentObjectProperties	0
InverseObjectProperties	3
DisjointObjectProperties	1
FunctionalObjectProperty	7
InverseFunctionalObjectProperty	6
TransitiveObjectProperty	0

To use the reasoner click Reasoner > Start reasoner  Show Inferences

Figura 15. Métricas de onto4AIR. Parte 1.



Ontology metrics:	
AsymmetricObjectProperty	10
ReflexiveObjectProperty	0
IrreflexiveObjectProperty	10
ObjectPropertyDomain	11
ObjectPropertyRange	10
SubPropertyChainOf	0
Data property axioms	
SubDataPropertyOf	52
EquivalentDataProperties	0
DisjointDataProperties	0
FunctionalDataProperty	33
DataPropertyDomain	31
DataPropertyRange	46
Individual axioms	
ClassAssertion	9
ObjectPropertyAssertion	6
DataPropertyAssertion	4
NegativeObjectPropertyAssertion	0
NegativeDataPropertyAssertion	0
SameIndividual	3
DifferentIndividuals	0

Figura 15. Métricas de onto4AIR. Parte 2.

## 9. Conclusiones y trabajo a futuro

Este documento presentó la ontología onto4AIR, un modelo semántico de RIs para IES. Onto4AIR integra conceptos del dominio relacionado con el acceso abierto, los repositorios institucionales, sus requerimientos operativos y técnicos en un vocabulario común, no ambiguo, formal y consistente que puede ser compartido entre usuarios y computadoras. Los términos representados provienen de un estudio tipo exploratorio de los documentos presentados en el marco legal y del análisis de los lineamientos generales [SIICYT 2017a] y técnicos [SIICYT 2017b] del Repositorio Nacional y los Repositorios Institucionales, éstos se agruparon en una jerarquía de clases y se relacionaron entre sí mediante propiedades de objeto y datos.

Los usuarios y las aplicaciones, durante las consultas, pueden recuperar además de las definiciones de los términos, las relaciones de éstos con otros elementos, comentarios o buscar únicamente sobre términos conectados a través de una relación específica.

Como trabajo a futuro, se propone la validación de onto4AIR considerando dos enfoques: 1) la percepción del modelo mediante la aplicación de pruebas de concepto a un conjunto de usuarios potenciales y 2) la población de la ontología a partir de datos provenientes de uno o más RIs de REMERI. A partir de las instancias introducidas, se podrá estimar la capacidad del modelo para responder consultas.

## 10. Referencias

- [Bauer y Kaltenböck 2012] Linked Open Data: The Essentials – A Quick Start Guide for Decision Makers (2012) by Florian Bauer, Martin Kaltenböck.
- [CONACYT 2016] Estrategia de acceso abierto a la información científica, tecnológica y de innovación repositorios institucionales. Términos de referencia para desarrollar los repositorios institucionales de acceso abierto a la información científica, tecnológica y de innovación. Fecha de consulta: 30 de Abril del 2016. Disponible en:  
<http://www.conacyt.mx/index.php/el-conacyt/convocatorias-y-resultados-conacyt/convocatorias-direccion-adjunta-de-planeacion-y-evaluacion/convocatoria-2015-para-desarrollar-los-repositorios-institucionales-de-acceso-abierto-a-la-informacion-cientifica-tecnologica-y-de-innovacion/10724-terminos-de-referencia-repositorios-institucionales-2015/file>.
- [DUBLIN CORE 2008] Dublin Core. (2008). Dublin Core. Obtenido de Metadata innovation: [en línea] disponible en: <http://dublincore.org/documents/2008/01/14/dcmi-terms/>
- [Gruber 1993] Gruber T R. A Translation Approach to Portable Ontology Specifications[S]. Knowledge Acquisition, 5/1993: 199-200.
- [Medina et. Al., 2017] Medina-Nieto M.A., Sánchez J., Cervantes O., Benitez A. De la Calleja J. 2017. LOD4AIR: a strategy to produce and consume linked open data from OAI-PMH repositories. En “Proceedings of the 27th International Conference on Electronics, Communications and Computers (CONIELECOMP 2017), San Andrés Cholula, Puebla, México. © IEEE. ISBN: 978-1-5090-3621-9/17. Pp. 161-168.
- [Musen 2015] Musen, M. A. The Protégé project: A look back and a look forward. AI Matters. Association of Computing Machinery Specific Interest Group in Artificial Intelligence, 1(4), Junio 2015. DOI: 10.1145/2557001.25757003.

- [Neches et al. 1991] Robert Neches, Richard Fikes, Tim Finin, Tom Gruber, Ramesh Patil, Ted Senator, and William R. Swartout. 1991. Enabling technology for knowledge sharing. *AI Mag.* 12, 3 (September 1991), 36-56.
- [Noy y McGuinness 2001] Noy N. F., McGuinness D. L. 2001. *Ontology development 101: A guide to create your first ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05, [en línea]. [Fecha de consulta: 16 de Enero del 2017). Disponible en:  
[http://protege.stanford.edu/publications/ontology\\_development/ontology101.pdf](http://protege.stanford.edu/publications/ontology_development/ontology101.pdf)
- [REMERI 2016] Red Mexicana de Repositorios Institucionales (REMERI). Fecha de consulta: 28 de Octubre del 2016. Disponible en:  
<http://www.remeri.org.mx/portal/index.html>
- [SIICYT 2017a] Lineamientos generales para el repositorio nacional y los repositorios institucionales. Documento disponible desde el sitio web del Sistema Integrado de Información sobre Investigación Científica, Desarrollo Tecnológico e Innovación (SIICYT). Fecha de consulta: 15 de Marzo del 2017. Dirección web:  
[https://www.repositorionacionalcti.mx/docs/Lineamientos\\_generales.pdf](https://www.repositorionacionalcti.mx/docs/Lineamientos_generales.pdf).
- [SIICYT 2017b] Lineamientos técnicos para el repositorio nacional y los repositorios institucionales. Documento disponible desde el sitio web del Sistema Integrado de Información sobre Investigación Científica, Desarrollo Tecnológico e Innovación (SIICYT). Fecha de consulta: 15 de Marzo del 2017. Dirección web:  
<http://www.siicyt.gob.mx/index.php/normatividad/2-conacyt/4-conacyt/1499-lineamientos-tecnicos-para-el-repositorio-nacional-y-los-repositorios-institucionales/file>.
- [Studer et al. 1998] Studer R, Benjamins V.R., Fensel D. 1998. Knowledge engineering: principles and methods. *Data Knowl. Eng.* 25, 1-2 (March 1998), 161-197. DOI=[http://dx.doi.org/10.1016/S0169-023X\(97\)00056-6](http://dx.doi.org/10.1016/S0169-023X(97)00056-6)
- [Suber 2010] Suber, P. 2010. Open Access Overview. Research on Institutional Repositories: Articles and Presentations. Paper 45.  
<http://digitalcommons.bepress.com/repository-research/45>