



**Consultora de Ciencias de la Información
Buenos Aires
Argentina**

Serie

DOCUMENTOS DE TRABAJO

Área: Investigación

**Una aproximación a un análisis comparativo entre tesauros y ontologías en
el área de las Ciencias Astronómicas**

Marcelo de la Puente

Diciembre 2009

N°011

ISSN 1852-6411

Copyright Consultora de Ciencias de la Información

Editor: Patricia Allendez Sullivan. Asistente Editorial: Mariana Sabugueiro

Puente, Marcelo de la

Una aproximación a un análisis comparativo entre tesauros y ontologías en el área de las Ciencias Astronómicas. Buenos Aires: Consultora de Ciencias de la Información, 2009.

ISSN 1852 - 6411

1. Tesauros. 2. Ontologías. 3. Astronomía. 4. Relaciones Semánticas. I. Título

Resumen:

El presente trabajo realiza un breve análisis comparativo entre tesauros y ontologías, sus semejanzas y diferencias, así como la posibilidad de convertir tesauros en ontologías. Se examinan distintos tesauros y una ontología del área de las ciencias astronómicas con el propósito de evaluar parámetros tales como la granularidad, grado de reutilización de los términos por parte de sistemas informáticos, etc.

Introducción

En el presente trabajo, se abordará el tema de un estudio comparativo entre las antiguas formas de organización del conocimiento, nacidas del ámbito de la Bibliotecología y Documentación, como los vocabularios controlados, los tesauros y los nuevos esquemas que provienen de otros campos temáticos, como la Ingeniería Informática, que se aplican a la clasificación e indización de recursos digitales que conocemos con el nombre de ontologías.

Las ontologías guardan, a su vez, relación con el proyecto de la web semántica, es decir la búsqueda del significado a través de las anotaciones de los recursos web, por parte de los buscadores, para mejorar de este modo la relevancia de los resultados de las búsquedas en los mismos.

En este trabajo nos dedicaremos a investigar y estudiar las semejanzas y diferencias entre los tesauros y las ontologías, así como su eficacia en la recuperación de información en el ámbito digital

Esperamos que este tema sea de interés ya que, humildemente consideramos que el trabajo que se realizará en las Unidades de Información del futuro inmediato, se relaciona con esta práctica.

Ontologías

El término “Ontología”, se deriva originalmente de la Filosofía, es la rama de la Filosofía que estudia el Ser en si la esencia de la realidad, pero fue tomado por la Ingeniería Informática en la década de los 90 del siglo pasado con un significado diferente, y se aplicó en el ámbito de la representación del conocimiento. Una ontología se define como una descripción formal de conceptos y de sus relaciones, en un dominio temático determinado, es decir, se especifica un vocabulario relativo a cierto dominio, se definen entidades, clases, propiedades, funciones y las relaciones entre estos componentes. La definición tradicional, según Gruber (1993), es una *explicitación formal de una conceptualización compartida*, esta definición se refiere a lo que puede ser representado computacionalmente. *Conceptualización* compartida se refiere a un modelo abstracto de algún fenómeno en el mundo proveniente de haber identificado los conceptos relevantes de dicho fenómeno. *Explícita* alude al tipo de concepto usado y a que las restricciones para su uso son claramente definidas. *Formal* se refiere a que la ontología debe ser legible por computadora y *Compartida* alude a una base de conocimiento consensual compartida por un grupo determinado. Tienen un rol protagónico en la cuestión de la interoperabilidad semántica entre distintos sistemas de información.

En términos prácticos el desarrollo de una ontología incluye:

- Definición de *clases* en la ontología
- Ordenación de las clases en una *jerarquía taxonómica* (Clases generales-clases particulares)
- Definición de *slots* o atributos y descripción de los valores permitidos para los mismos
- Rellenado de los valores de los atributos con *instancias* o ejemplos

Para la definición de clases dentro de una estructura taxonómica se puede comenzar con los conceptos más generales y terminar con los más específicos. En segundo lugar, se puede comenzar con los más específicos y terminar con los más genéricos y en una tercera aproximación , se puede comenzar con los conceptos de un nivel intermedio de la jerarquía y luego generalizar y especializar.

Este sería un ejemplo de la estructura taxonómica de conceptos en la Astronomy Ontology, la clase *B Giant* (Gigantes B), o estrellas gigantes de tipo B en la clasificación espectral de las estrellas, lo que corresponde a supergigantes azules y a la Superclase de *O-B Stars* (Estrellas azules), que a su vez está contenida en la Superclase *Young Stars* (Estrellas Jóvenes). También se observa el fenómeno de la herencia múltiple en el que una clase puede ser subclase de diferentes clases mayores, es decir, pertenecer a distintas escalas jerárquicas, la clase *B Giant* (Gigantes B) es tanto subclase de *Ob Star* (Estrellas O-B) que tiene como clase mayor a *Young star* (Estrellas jóvenes), como de *Giant Star* (Estrellas gigantes), que es subclase de la Clase mayor *Stars* (Estrellas)

Class: BGiant

<http://archive.astro.umd.edu/ont/astronomy.owl#BGiant>

- [youngStar](#)
 - [ObStar](#)
 - [BStar](#)
 - BGiant
 - [BStar](#)
 - BGiant
- [stellarObject](#)
 - [star](#)
 - [youngStar](#)

- [ObStar](#)
 - [BStar](#)
 - BGiant
 - [BStar](#)
 - BGiant
 - [giantStar](#)
 - BGiant
- [astroMatterObject](#)
 - [star](#)
 - [youngStar](#)
 - [ObStar](#)
 - [BStar](#)
 - BGiant
 - [BStar](#)
 - BGiant
 - [giantStar](#)
 - BGiant
- [ultravioletSource](#)
 - [earlyTypeStar](#)
 - [ObStar](#)
 - [BStar](#)
 - BGiant
 - [BStar](#)
 - BGiant

Los atributos o roles de las clases pueden tener distintos tipos de valores permitidos las llamadas *facet* o restricciones de roles. Se denomina cardinalidad al número de valores permitidos que un atributo puede tener, puede ser simple (un solo valor), múltiple (dos o más valores), En cuando al tipo de valores permitidos pueden ser; string o cadena de caracteres, numérico (solo números), booleano (respuesta binaria sí o no), enumerativo (lista de valores permitidos para un atributo dado) y tipo instancia o ejemplo que permite definir relaciones entre individuos pertenecientes a una clase y define una lista de clases permitidas de la cual, la instancia puede provenir.

Las clases permitidas para los atributos del tipo instancia se denominan el *rango* de un atributo, por ejemplo, en la Astronomy Ontology, ontología de Astronomía, el atributo *has star measurerent* (tiene medición estelar) solo se aplica a la clase Estrellas

Object Property: hasStarMeasurement

<http://archive.astro.umd.edu/ont/astronomy.owl#hasStarMeasurement>

Domains

- [star](#) or [starSimulation](#)

Ranges

- [starMeasurement](#)

El conjunto de clases e instancias terminan constituyendo una base de conocimiento para ese campo en particular. La ontología es un modelo de un campo disciplinario dado y sus conceptos deben reflejar esa realidad.

Al igual que las taxonomías documentales, las ontologías no provienen del campo de la Bibliotecología sino más bien del área Informática, más específicamente del campo interdisciplinario que implicaría la representación de contenidos digitales en la web que abarca disciplinas como la Inteligencia Artificial y la Lingüística, entre otras. Para la inteligencia artificial, las ontologías son recursos construidos que permiten representar el conocimiento en común y compartido sobre un tema particular. La ontología no solo incluye el vocabulario sobre un dominio temático particular sino también las reglas que permiten combinar los términos y relaciones para definir extensiones al vocabulario.

En Lingüística las ontologías se utilizan frecuentemente en traducción automática basada en el conocimiento y en Terminología. Su función en la representación de una base de conocimiento es la de apoyo semántico para las palabras, o sea las palabras son descritas como objetos lingüísticos en una base de datos léxica y son relacionadas con una jerarquía conceptual apoyada en la ontología.

¿Cuál sería el aporte de las ontologías al desarrollo de la web semántica? La web semántica busca la catalogación de la información contenida en los recursos web a través de los metadatos y de las ontologías, es decir, utilizando el significado de las palabras no meramente mediante el uso de palabras claves como en los buscadores más utilizados en Internet. Las ontologías permitirán la organización de la información en forma tal que los agentes de software puedan interpretar el significado de los términos utilizados en una búsqueda y consecuentemente, buscar con una mayor precisión que actualmente. Las aplicaciones podrán extraer automáticamente datos de páginas web, procesarlos y extraer conclusiones de ellos. Un agente inteligente elegirá ontologías especializadas en distintos dominios temáticos en función de los términos de búsqueda usados para satisfacer la consulta de un usuario. Existen ontologías de diferentes dominios temáticos, por ejemplo, una ontología médica SNOMED

<http://www.snomed.org>, ontologías lingüísticas como WorldNet <http://wordnet.princeton.edu/perl/webwn>, ontologías de carácter general como una desarrollada por la ONU, <http://www.unspsc.org>, ontologías del campo empresarial como Enterprise Ontology, etc. Las ontologías pueden ser útiles también en el desarrollo de mecanismos que faciliten la comunicación entre personas y máquinas a través del lenguaje natural, es decir como herramienta de procesamiento del lenguaje natural, como en el caso de la ontología WorldNet.

Dependiendo del grado de formalidad las ontologías se clasifican en informales, semi-informales, semi-formales y formales. Las primeras se expresan en cualquier lenguaje natural, las segundas en una forma estructurada y restringida de un lenguaje natural, las terceras usan lenguajes estructurados como el RDF y las últimas definen los términos mediante lenguajes lógico-matemáticos cuyos símbolos se definen de forma unívoca, o sea, sin ambigüedades, consecuentemente permiten emplear teoremas y demostraciones. Los primeros tipos de ontologías pueden tener términos inconsistentes o ambiguos pero las ontologías formales no los permiten. Las ontologías favorecen la comunicación entre personas y máquinas al favorecer la comprensión común de un dominio se evitan las confusiones conceptuales y terminológicas, permiten a las aplicaciones ponerse de acuerdo en los términos que utilizan

Otra clasificación dada por otros autores es la de ontologías de *nivel más alto*, *dominios*, *tareas* y *aplicaciones*. Las primeras describen objetos más generales y abstractos, como el espacio, el tiempo, la materia etc. Las segundas describen el vocabulario relacionado con un dominio temático, son éstas a las que se restringe nuestro análisis, las de tareas describen actividades útiles para las organizaciones mientras que las de aplicación describen los conceptos de acuerdo a un campo determinado o a unas tareas concretas.

Permiten la interoperabilidad entre distintos sistemas, facilitando la traducción de términos entre distintas aplicaciones con distintos lenguajes de programación y facilitan el razonamiento automático de los agentes de software a través de reglas de inferencia para extraer conclusiones de los datos analizados.

Diferencias y semejanzas con los tesauros

Al efectuar un cuadro comparativo entre las ontologías y los tesauros, podemos observar en primer lugar, que los tesauros provienen del campo de la Bibliotecología y la Documentación, mientras que las ontologías provienen del campo de la Ingeniería Informática. El tesauro es una lista estructurada de conceptos destinados a representar de manera unívoca el contenido de los documentos y de las consultas dentro de un sistema documental determinado y ayuda a la indización de los documentos y de las consultas. El *tesauro* representa la estructuración conceptual de un determinado campo del conocimiento. La organización semántica se presenta a través de la explicitación de las relaciones entre los distintos conceptos y a través de la restricción de los significados de los términos, básicamente, por el control de la sinonimia (sinónimos) y de la polisemia (palabras con escritura idéntica pero con diferente significado), problemas que comprometen la precisión en la recuperación de la información. Las relaciones entre términos son de tres tipos: de equivalencia (sinónimos), jerárquicas (términos generales y específicos) y asociativas (términos relacionados).

En el caso de las *ontologías* léxicas, las relaciones conceptuales entre términos son mucho más variadas y dependen del campo de conocimiento a estructurar, poseen un grado más alto de especificación y de flexibilidad, es decir, se adaptan más fácilmente a su distribución a la reutilización pueden acomodar términos descriptivos variables. Otra diferencia es que las ontologías están diseñadas para compartir y reutilizar la información almacenada, la cual puede ser interpretada tanto por personas como por programas informáticos, mientras

que los tesauros fueron diseñados originalmente para funcionar en el ámbito impreso, para facilitar la comunicación entre humanos. En las ontologías las relaciones entre conceptos se definen formalmente, por lo que si una persona desea reutilizar una ontología desarrollada por otros puede conseguir la información de todos los conceptos que soporta, su jerarquía taxonómica y los axiomas que permiten a los programas realizar inferencias lógicas a partir de las declaraciones formales de las relaciones entre los distintos conceptos, etc.

Las ontologías al permitir una especificación de las relaciones que implica una mayor formalización semántica, es decir un desarrollo semántico en mayor profundidad, que permite una estructura conceptual mucho más específica y consistente al momento de la búsqueda, desde un punto de vista relacional e informático. Según Qin y Palling (2001), existe en las ontologías una mayor profundidad semántica en las relaciones, un desarrollo de las relaciones de tipo clase/subclase y de las relaciones cruzadas mucho más exhaustivo y ello evita muchas inconsistencias conceptuales que aparecen frecuentemente en los tesauros, por ejemplo, en las relaciones jerárquicas abarcan tanto las relaciones de tipo partitivo (Todo-parte) como las de tipo género-especie sin diferenciar entre ambas. En las ontologías los atributos y las relaciones conceptuales pueden ser asignadas específicamente al concepto en cuestión o pueden ser heredadas de un concepto jerárquicamente superior y puede darse el caso de la herencia múltiple, como se ha visto, es decir, cada concepto puede recibir propiedades y relaciones conceptuales de más de un concepto jerárquicamente superior

Granularidad: La granularidad se refiere a la propiedad de las ontologías de representar los distintos niveles jerárquicos con la mayor especificidad posible, es decir se refiere a la profundidad del detalle o especificidad de los conceptos representados,. Este nivel de detalle en la descripción semántica es más profundo en las ontologías en los tesauros. Otra propiedad de los lenguajes documentales es la modularidad. La modularidad muestra la estructura del

conjunto de términos basándose en las proporciones de términos de más de una palabra y en el grado de reutilización de subtérminos. En un estudio realizado por Good y Tennis (2009), se incluyen parámetros como el grado de precoordinación lingüística o combinaciones de uno o más términos (unitérminos, dupletes, tripletes, etc.) y grado de composición de los términos (composición de términos de un conjunto con términos de otro subconjunto para formar un término compuesto) . Se examinan lenguajes de indización de lenguaje natural como las folksonomías y se las compara con distintos tesauros y ontologías del área, biomédica, como el Mesh (Medical Subject Headings), la Gene Ontology (GO) ,Protein Ontology (PRO) entre otros lenguajes. Se destaca que, como grupo los tesauros y las ontologías, tienen un grado de modularidad más elevado que lenguajes como las folksonomías. La presencia o ausencia de una estructura modular indicaría de modo indirecto, la fortaleza de la estructura conceptual de un conjunto de términos, aunque en un lenguaje no se declare explícitamente las relaciones semánticas entre términos. Brinda una base para comparar conjuntos de términos de diferentes lenguajes documentales con métodos estadísticos

Una de las diferencias fundamentales entre los tesauros y las ontologías es que, con el incremento de los recursos digitales y la mayor accesibilidad a la web, se impone la necesidad de estructuras conceptuales mucho más flexibles para navegar y buscar información, sobre todo en el caso de los usuarios que no son expertos en un campo temático dado, ya que usualmente, estos usuarios no tienen un conocimiento de la estructura temática de dichos recursos y no obtienen resultados satisfactorios en el proceso de búsqueda. Las ontologías cumplirían con este rol al permitir servirle al usuario novato, al funcionar tanto como una herramienta de búsqueda como un instrumento de navegación. Este abordaje sería mucho más “naturalista” desde el punto de vista del usuario, ya que le permitiría seguir su propio punto de vista o estructura conceptual del mundo, durante el proceso de búsqueda, en lugar de seguir una estructura más rígidamente jerarquizada, impuesta por los expertos del campo disciplinario.

Para lograr la interoperabilidad semántica entre diferentes sistemas informáticos, se requieren sistemas de información que trabajen automáticamente con el significado de los términos y sus diferentes combinaciones, esto va más allá de los lenguajes documentales tradicionales como los tesauros que solo normalizan la entrada y salida de los datos. Pero los tesauros pueden ser convertidos en Ontologías, como lo señalan Qin y Palling (2001). Uno de los problemas claves para lograr esto, sería el incremento de la precisión en las relaciones semánticas lo que encuentra un primer obstáculo en los problemas de definición existentes en las relaciones jerárquicas en los tesauros, anteriormente mencionado, además de la infrutilización de los tesauros de las relaciones asociativas, que las ontologías explotan más a profundidad, al tratar de incluir roles, propiedades, etc. Las ontologías diferencian claramente entre objetos y propiedades

Uno de los métodos propuestos sería la conversión de los enlaces de los tesauros en relaciones expresadas en lenguajes de etiquetado de ontologías como el DAML/OIL (DARPA Agent Mark Up Language y Web Ontology Language). También, puede convertirse el contenido del tesoro en relaciones del RDF Schema, un lenguaje primitivo de ontologías que proporciona los elementos básicos para la descripción de vocabularios. En este último caso, la ontología se construye a partir del tesoro con las siguientes modificaciones: a) Representación de cada concepto de la base de conocimiento por un identificador unívoco, b) conversión de la jerarquía existente en el tesoro en jerarquía de conceptos, c) aumento del número de conceptos con slots adicionales, d) suma de conocimiento entre posibles valores de los campos y nodos en la base del conocimiento.

Para la conversión al RDF Schema se van definiendo clases con las relaciones características del tesoro, (Objetos del tesoro, Conceptos, Conceptos más generales, Términos y Notas de alcance), así como propiedades específicas (Código de clasificación, Tiene nota de alcance, Es indicado por, Término

preferido, Usado por, Relaciones conceptuales (General, Específico, Más genérico, Relaciones de conceptos relacionados), sin olvidar las propiedades de términos (En lenguaje de) y las de notas de alcance (En lenguaje de, tipo de nota de alcance)

Análisis comparativo de tesauros y ontologías del área de las Ciencias Astronómicas

Se han efectuado diferentes análisis comparativos entre tesauros y ontologías de diversos dominios temáticos en la literatura. En uno de ellos Romá- Ferri y Palomar (2008), comparan diferentes vocabularios controlados y ontologías del área biomédica. Con el propósito de realizar un análisis comparativo entre tesauros y ontologías de un área temática particular, se tomó como campo general la Astronomía y se efectuó un análisis comparativo entre cuatro vocabularios controlados: El *Astronomy Thesaurus*, de la UAI (Unión Astronómica Internacional), *PACS* (Physics and Astronomy Class), una clasificación del área de las ciencias físicas y astronómicas, elaborada por el American Institute of Physics, el *Tesaurus SPINES*, que es un tesoro general sobre ciencia y tecnología y la *Astronomy Ontology*, una ontología del área Astronómica, que cubre clases astronómicas y sus propiedades, elaborada a partir de IVOA Thesaurus, un tesoro de Astronomía y de la ontología astronómica UMD de un grupo dedicado al uso de las nuevas tecnologías al campo astronómico.

Se tomaron en comparación tres parámetros básicos:

- 1- Cobertura conceptual: finalidad, utilidad y amplitud de los conceptos representados
- 2- Estructura jerárquica: organización conceptual de las relaciones semánticas entre los términos, las que representan el conocimiento del

campo temático, granularidad, profundidad o especificidad de los conceptos representados dentro de la estructura

- 3- Relaciones conceptuales: El tipo de relaciones semánticas o enlaces existentes, bien sea dentro de la estructura jerárquica (relación género/especie, partitiva) o por otro tipo de relaciones semánticas (asociativas, etc.)

Realizando un primer análisis comparativo vemos que existen algunas semejanzas y diferencias en el campo temático abarcado, en el tipo de relaciones conceptuales, en el grado de granularidad alcanzado en la estructura jerárquica de cada tipo de vocabulario (Tabla 1). En la clasificación PACS se observan principalmente, relaciones jerárquicas de varios niveles jerárquicos, en el Astronomy Thesaurus y en el tesoro SPINES, las tres tipos de relaciones semánticas características de los tesauros (de equivalencia, jerárquicas y asociativas) y en la Astronomy Ontology es en donde se observa un mayor nivel de profundidad en el tipo de relaciones semánticas definidas, en la que no solo existe un nivel de granularidad detallado en la estructura jerárquica, sino que se observan en los atributos o roles, diferentes tipos de relaciones definidas para dominios particulares y con rangos de valores específicos, por ejemplo la propiedad *has spectra*, tiene como dominio los objetos astronómicos y como rango la clase espectros, la propiedad *has member* (relación es un tipo de), se diferencia claramente de la relación *has part* o *is part of* (relaciones partitivas) , lo que no ocurre en los tesauros o al menos, no explícitamente en la clasificación astronómica. Se define el rango de de una faceta como los valores permitidos para un atributo particular. Las clases cuyas propiedades describen el atributo se llama el *dominio* del rango. En el caso, por ejemplo del atributo *has spectra* (tiene espectro). El dominio de la misma son los objetos astronómicos y el rango es la clase espectro.

En el grado de formalización, lo que se refiere a la capacidad de ser utilizable por computadoras del lenguaje usado en la representación del conocimiento, vemos que es en la ontología es en donde se alcanza un mayor grado en la definición rigurosa y exhaustiva de las propiedades de las clases y subclases, lo que facilita su utilización por sistemas informáticos.

Se tomó como ejemplo de análisis un concepto *Estrellas binarias visuales*, para observar la localización del mismo dentro de la estructura jerárquica de cada sistema, lo que a su vez depende de la granularidad de la estructura de clases. En el tesoro de la UAI, el término depende de la clase mayor *Estrellas binarias*, y es un concepto hermano (al mismo nivel jerárquico) de otros tipos de estrellas binarias como por ejemplo, *Estrellas binarias cercanas*, *Estrellas binarias de contacto*, *Estrellas binarias eclipsantes*, etc. , se encuentra en una jerarquía intermedia en la escala. En el caso de la clasificación PACS encontramos el concepto en un nivel más específico, dentro de la subclase 97 Estrellas, tenemos la subclase 97.80 *Estrellas binarias y múltiples* y dentro de ella, encontramos el concepto buscado 97.80Di *Estrellas binarias visuales*, en una clase más específica. En el Tesoro SPINES solo tenemos el término *Estrellas binarias*, no existen subdivisiones posteriores. En la ontología observamos una definición mucho más detallada dentro de la clase principal *Estrellas binarias*, con 14 subclases menores, una de las cuales es *Estrellas binarias amplias* y dentro de esta encontramos *Estrellas binarias visuales* (Ver Tabla 2), es en este último caso en dónde observamos el mayor nivel de granularidad en la representación del concepto analizado. Además observamos en el caso de la ontología la herencia múltiple de conceptos, en dónde el concepto buscado pertenece a diferentes escalas jerárquicas, la clase mayor *Estrellas binarias* pertenece a su vez a la clase *Sistemas de Estrellas*, y esta a su vez, depende de la clase mayor *Objetos de materia astronómicos*, que a su vez pertenece a la clase más amplia *Objetos materiales*. A su vez el concepto *Estrellas binarias* es subclase de otra clases genérica, como *Sistema binario*, a su vez dentro de la clase mayor, *Cosas*, de la Clase mayor *Sistemas*, (Ver Tabla 2)

Como resultado del análisis efectuado, podemos concluir que existen diferentes niveles de granularidad, de profundidad semántica en la representación de las relaciones entre conceptos y diferentes grados de usabilidad en el sentido de que los diferentes lenguajes analizados puedan ser utilizados por diferentes sistemas informáticos, para lo que se requiere una definición explícita de las relaciones conceptuales, de lo contrario esto acarrearía ambigüedad en la interpretación semántica lo que dificulta la interoperabilidad entre diferentes sistemas informáticos. Solo se observa ese nivel de detalle en las definiciones conceptuales y en las relaciones entre conceptos en la ontología, ya que en los tesauros y en la clasificación existen algunas restricciones en el tipo de relaciones representadas.

En la clasificación, solo se encuentra explicitada la relación de género/ especie o es *un es tipo de*, en los tesauros se registran tres tipos de relaciones, las de sinonimia, jerárquicas y asociativas, pero esta estructura organizativa no es suficiente para especificar las propiedades de cada concepto. Tanto en la estructura jerárquica de la clasificación como en los tesauros, no se efectúa una distinción entre las relaciones de tipo género/especie como de las relaciones partitivas (todo/parte).

En la Ontología de Astronomía, por el contrario, no solo se expresan en las propiedades a las que se le aplican a cada tipo de categoría una gran variedad de relaciones semánticas, se diferencia no solo entre las relaciones de tipo género/ especie, *miembro de* (subclase-clase mayor) y *tiene miembro* (clase mayor-subclase), de las relaciones partitivas, *parte de* (*parte-todo*) y *tiene parte* (*todo-parte*), sino que se definen relaciones mucho más específicas definidas para clases específicas, como por ejemplo, *medidas estelares*, (dominio: estrellas), *medidas galácticas* (dominio: galaxias). En la ontología también están diferenciadas claramente las instancias o ejemplos de las clases principales:

Por ejemplo, Marte como ejemplo de la clase Planeta del Sistema Solar.

Conclusiones

El grado de formalidad de los lenguajes estudiados es variable, pero es la Ontología la que alcanza el grado más elevado. El UAI Thesaurus, PACS y el tesoro SPINES tienen un grado de formalidad reducido y un nivel mucho más bajo de expresividad semántica de las distintas relaciones conceptuales en el campo de la Astronomía, pero podrían tomarse como base para el desarrollo de Ontologías semiformales o con mayor grado de formalidad si se incluyen definiciones en RDF Schema o si se utilizaran lenguajes de etiquetado de ontologías.

El desarrollo de ontologías en el área facilitaría el descubrimiento de nuevas relaciones conceptuales en investigaciones en el campo y puede ser de utilidad para usuarios nuevos en el campo, ya que los tesauros generalmente son utilizados por los profesionales de la disciplina y por los bibliotecarios. Una ontología al lograr una máxima expresividad en las definiciones conceptuales y en las relaciones semánticas, puede guiar mucho más fácilmente a un usuario no familiarizado con la disciplina, así como incrementar el acierto y la precisión en la recuperación de información relevante en búsquedas. Además, debido a la mayor flexibilidad de las ontologías frente a los tesauros se podrían realizar nuevas adaptaciones frente a los cambios y avances en la disciplina, mucho más fácilmente, que en el caso de los tesauros

Bibliografía

Arano, S. (2005) "Los tesauros y las ontologías en la Biblioteconomía y la Documentación". En: Hipertexto.net, 3". Consultado el 19 de diciembre de 2009. Disponible en: <http://www.hipertext.net/web/pag260.htm>

Astronomy Ontology. Consultado el 19 de diciembre de 2009. Disponible en: <http://atlas.astro.umd.edu:8080/astro-onto/ontologies/astronomy.html>.

Astronomy Thesaurus. Consultado el 19 de diciembre de 2009. Disponible en: <http://msowww.anu.edu.au/library/thesaurus/spanish/>

García Giménez, A. (2004) "Instrumentos de representación del conocimiento: Tesoros versus Ontologías". En: Anales de Documentación, 7, 79-95
Consultado el 19 de diciembre de 2009. Disponible en: <http://revistas.um.es/analesdoc/article/viewFile/1691/1741>

Good, B. M.; J.T. Tennis (2009). "Term based comparison metrics for controlled and uncontrolled indexing languages". En: Information Research, 14, 1.
Consultado el 20 de diciembre de 2009. Disponible en: <http://informationr.net/ir/14-1/paper395.html>

Gruber, T. R. (1993a). "A translation approach to portable ontology specifications". Consultado el 19 de marzo de 2009. Disponible en: <http://gicl.mcs.drexel.edu/people/regli/Classes/KBA/Readings/KSL-92-71.pdf>

Gruber, T. R. (2001). What is an Ontology? . Stanford: Knowledge Systems Laboratory. Stanford University. Consultado el 19 de marzo de 2009. Disponible en: <http://www-ksl.stanford.edu/kst/what-is-an-ontology.html>

Noy, N. McGuinness, D. Ontology 101. Consultado el 19 de marzo de 2009. Disponible en: <http://www-ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf>

PACS (Physics and Astronomy Classification Scheme). Consultado el 19 de diciembre de 2009. Disponible en: <http://www.aip.org/pacs/>

Padirac, B. de (1988). SPINES thesaurus: a controlled and structured vocabulary for information processing in the field of science and technology for development. París: UNESCO. Consultado el 19 de diciembre de 2009. Disponible en: http://unesdoc.unesco.org/Ulis/cgi-bin/ulis.pl?catno=79377&set=4AFA0CF8_2_2&gp=0&lin=1&ll=1

Romá- Ferri, M. T.; Palomar, M. (sec-oct.2008). “Análisis de terminologías de salud para su utilización como ontologías computacionales en los sistemas de información clínicos”. En: Gaceta Sanitaria, 22, 5. Consultado el 19 de diciembre de 2009. Disponible en: http://scielo.isciii.es/scielo.php?pid=S0213-91112008000500006&script=sci_arttext

Qin, J.; Paling, S. (2001). “Converting a controlled vocabulary into an ontology: the case of GEM”. En: Information Research 6, 2. Consultado el 19 de diciembre de 2009. Disponible en: <http://informationr.net/ir/6-2/paper94.html>

Tabla I: Análisis de los componentes de los tesauros y ontologías

	1- Cobertura conceptual
Astronomy Thesaurus UAI	Uso de la terminología en la investigación actual en el campo de la Astronomía. La lista de términos primarios fue recopilada de listas de autoridades existentes y fuentes de referencia del campo. Se incluyeron revisiones y enmiendas de varios campos de la astronomía durante su revisión en 1992. No se incluyeron aspectos sobre disciplinas relacionadas, como historia, instrumentación, electrónica, computación, física, cohetes e investigación espacial. Su uso está pensado para astrónomos y bibliotecarios de bibliotecas astronómicas
PACS	Clasificación jerárquica diseñada para categorizar, clasificar y recuperar la literatura de los campos físicos y astronómicos. Incluye Física, Astronomía, Geofísica y Astrofísica
Tesauro SPINES	Vocabulario controlado para el procesamiento de la información en el área de ciencia y técnica. Posee un vocabulario controlado de 10.000 términos, estructurados en más de 77.000 relaciones semánticas que hacen referencia al desarrollo, tanto en sus vertientes científica y tecnológica,

como en el ámbito económico y social a escala nacional e internacional

Astronomy Ontology Incluye conceptos que permiten recuperar información sobre temas astronómicos. Cubre clases astronómicas y propiedades e incluye términos del Tesoro IVOA y de la Ontología Astronómica UMD, elaborada por un grupo de estudio que busca adaptar las nuevas tecnologías de la información al campo astronómico.

2- Estructura jerárquica

Astronomy

Thesaurus UAI

Más de 300 clases principales. Jerarquía compuesta por lista finita de categorías y subcategorías siguiendo un orden lógico determinado por un criterio que aproxima los conceptos más específicos y los unifica dentro de los más generales

PACS

Tiene una estructura jerárquica de 10 categorías genéricas subdivididas en categorías más específicas. La jerarquía incluye principalmente, cinco niveles de profundidad, siendo el quinto nivel el del término más específico. Utiliza una codificación alfanumérica dividida en tres partes

Tesouro SPINES Posee un vocabulario controlado de 10.000 términos, estructurados en más de 77.000 relaciones semánticas

Astronomy Ontology Tiene una estructura jerárquica de 1677 clases, las que se subdividen en subclases más específicas, 34 propiedades de objetos, 9 propiedades de datos y 88 individuos o instancias de las clases. Estructura jerárquica multiaxial (campo temático estructurado en varias jerarquías): herencia múltiple de conceptos más específicos

3- Relaciones conceptuales

Astronomy

Thesaurus UAI Presenta tres tipos de relaciones: de equivalencia (sinonimia), UP (Usado Por) y USE, jerárquicas : término general (TG), término específico (TE), y asociativas: término relacionados (TR)

PACS

Las relaciones se derivan de la propia dependencia jerárquica en la que está estructurada la clasificación. Cada nivel inferior responde a la relación de tipo, (género/especie) de la clase superior

Tesouro SPINES Relaciones de equivalencia, jerárquicas y asociativas

Astronomy Ontology Presenta una amplia estructura jerárquica de 1677 clases principales que se subdividen en varios niveles más específicos, principalmente, por medio de la relación “es un tipo de” , los conceptos jerárquicamente más específicos heredan las propiedades de los conceptos genéricos de los que forman parte. Los atributos o roles denominados *propiedades de objetos* se aplican a determinadas clases que constituyen dominio particular y no a otras. Permiten una estructura semántica mucho más profunda de relaciones que va más allá de las relaciones jerárquicas, como *hasDetection* (detección), *hasRadius* (tiene radio) , *memberOf* (tipo de), *partOf* (parte de), etc. Se diferencian las relaciones jerárquicas de género-especie o es un tipo de, de las relaciones partitivas. Tienen un rango de valores permitidos . Lo mismo ocurre con las *propiedades de datos*. Los *individuos*, son ejemplos o instancias de las clases

Tabla II: Astronomy Thesaurus UAI

ESTRELLAS BINARIAS VISUALES

See also *OPTICAL DOUBLE STARS*.

UF

[estrellas dobles visuales \[2\]](#)

[estrellas binarias muy separadas](#)

BT

[ESTRELLAS BINARIAS](#)

RT

[DISCO OCULTANTE](#)

[MOVIMIENTOS ORBITALES](#)

[SEPARACION](#)

[OBSERVACIONES VISUALES](#)

ESTRELLAS BINARIAS

Only the broad classes are shown below. Specific types will be found under these, or from the Hierarchical List.

UF

[binarias](#)

[sistemas binarios](#)

[estrellas dobles](#)

[estrellas binarias muy separadas](#)

NT

[ESTRELLAS VARIABLES ALGOL](#)

[ESTRELLAS BINARIAS ASTROMETRICAS](#)

[ESTRELLAS VARIABLES CATACLISMICAS](#)

[ESTRELLAS BINARIAS CERCANAS](#)

[ESTRELLAS BINARIAS DE CONTACTO](#)

[ESTRELLAS BINARIAS SEPARADAS](#)

[ESTRELLAS BINARIAS ECLIPSANTES](#)

[ESTRELLAS VARIABLES ELIPSOIDALES](#)
[ESTRELLAS BINARIAS ERUPTIVAS](#)
[ESTRELLAS BINARIAS INTERACTUANTES](#)
[ESTRELLAS RS CANUM VENATICORUM](#)
[ESTRELLAS BINARIAS SEMIDESPENDIDAS](#)
[ESTRELLAS BINARIAS ESPECTROSCOPICAS](#)
[ESTRELLAS BINARIAS VISUALES](#)

RT

PACS Thesaurus

90- Geophysics, Astronomy and Astrophysics

[91. Solid Earth physics](#)

[92. Hydrospheric and atmospheric geophysics](#)

[93. Geophysical observations, instrumentation, and techniques](#)

[94. Physics of the ionosphere and magnetosphere](#)

[95. Fundamental astronomy and astrophysics; instrumentation, techniques, and astronomical observations](#)

[96. Solar system; planetology](#)

[97. Stars](#)

[98. Stellar systems; interstellar medium; galactic and extragalactic objects and systems; the Universe](#)

97.80.-d Binary and multiple stars 97.80.Af Astrometric and interferometric binaries 97.80.Di **Visual binaries** 97.80.Fk Spectroscopic binaries; close binaries 97.80.Gm Cataclysmic binaries (novae, dwarf novae, recurrent novae, and nova-like objects); symbiotic stars (*see also* 97.30.Qt Novae) 97.80.Hn

Eclipsing binaries 97.80.Jp X-ray binaries (*see also* 98.70.Qy X-ray sources and 97.60.Gb Pulsars) 97.80.Kq Multiple stars

Tesaurus SPINES

Tesaurus SPINES

ESTRELLAS BINARIAS

TG [ESTRELLAS](#)

TR [SATÉLITES NATURALES](#)

EN BINARY STARS

FR ETOILES BINAIRES

Astronomy Ontology

Class: binaryStar

<http://archive.astro.umd.edu/ont/astronomy.owl#binaryStar>

- [set](#)
 - [system](#)
 - [systemOfStars](#)
 - binaryStar

- [AlgolVariableStar](#)
- [astrometricBinaryStar](#)
- [closeBinaryStar](#) +
- [contactBinaryStar](#) +
- [detachedBinaryStar](#) +
- [eclipsingBinaryStar](#) +
- [ellipsoidalVariableStar](#) +
- [eruptiveBinaryStar](#) +
- [interactingBinaryStar](#) +
- [RsCanumVenaticorumStar](#)
- [semiDetachedBinaryStar](#)
- [spectroscopicBinaryStar](#) +
- [visualBinaryStar](#)
- [wideBinaryStar](#) +

Class: visualBinaryStar

<http://archive.astro.umd.edu/ont/astronomy.owl#visualBinaryStar>

- [astroMatterObject](#)
 - [systemOfStars](#)
 - [binaryStar](#)
 - visualBinaryStar
 - [wideBinaryStar](#)
 - visualBinaryStar
- [systemOfStars](#)
 - [binaryStar](#)
 - visualBinaryStar
 - [wideBinaryStar](#)
 - visualBinaryStar

- [Thing](#)
 - [binarySystem](#)
 - [binaryStar](#)
 - visualBinaryStar
 - [wideBinaryStar](#)
 - visualBinaryStar
- [binarySystem](#)
 - [binaryStar](#)
 - visualBinaryStar
 - [wideBinaryStar](#)
 - visualBinaryStar
- [system](#)
 - [systemOfStars](#)
 - [binaryStar](#)
 - visualBinaryStar
 - [wideBinaryStar](#)
 - visualBinaryStar

Annotations

- label: "visual binary star" (string)

Superclasses

- [binaryStar](#)
- [wideBinaryStar](#)

created with [OWLDoc](#)

