

Universidad de las Ciencias Informáticas

Facultad 3



Título: Propuesta metodológica para el desarrollo de una ontología para la Clasificación de Mercancías.

Trabajo de Diploma para optar por el título de
Ingeniero en Ciencias Informáticas

Autor: Yusney Martínez Estrada

Tutores: MSc. Nemury Silega Martínez
Ing. Eilys Pacheco Rodríguez

La Habana
18 de junio de 2012

Nunca consideres el estudio como una obligación, sino como una oportunidad para penetrar en el bello y maravilloso mundo del saber.

Albert Einstein

DECLARACIÓN DE AUTORÍA.

Declaro ser autora de la presente tesis y reconozco a la Universidad de las Ciencias Informáticas los derechos patrimoniales de la misma, con carácter exclusivo.

Para que así conste firmo la presente a los ____ días del mes de _____ del año _____.

Yusney Martínez Estrada

Firma del Autor

MSc. Nemury Silega Martínez

Firma del Tutor

Ing. Eilys Pacheco Rodríguez

Firma del Tutor

DATOS DE CONTACTO.

Nemury Silega Martínez

Correo: nsilega@uci.cu

Eilys Pacheco Rodríguez

Correo: epacheco@uci.cu

AGRADECIMIENTOS.

A mis padres por confiar siempre en mí y apoyarme en todas mis decisiones, por su amor y cariño incondicional. Gracias a mi familia por estar siempre al tanto de mí. A mi eterno tutor José Antonio Cobo Rodríguez.

A mi tutora Eilys por correr conmigo siempre que fue necesario. A mi tutor Nemury por brindarme su ayuda y apoyo incluso sin conocerme. Sin él nunca hubiera sido posible este trabajo. Muchas gracias por confiar en mí.

A mis amigos que me han acompañado durante estos años, gracias por ayudarme, soportarme, escucharme y estar allí para mí. Mi negrita Dayana, mis paños de lágrimas e incansables amigas Yurixay, Danarys y Elizabeth, a mi amigo incondicional Yoan. Al amigo y hoy profe Juan Javier. A los que me acompañaron por 11 meses y fueron mis hermanos.

Gracias a ese equipo que me ha acompañado casi desde que empecé en la universidad y que hoy son parte de mi familia, Yanet Parra, Dismey Saavedra, Asnay Guirola, Yuniesky Armentero.

Gracias y mil gracias a Magdiel Santana Hernández, esa personita que ha llegado por último a mi vida pero no por esto deja de ser importante, gracias mi amor por tu apoyo, tu ayuda, por soportarme cada día y responderme siempre con una sonrisa en la cara.

Gracias a todos y a la Revolución por permitirme estudiar en una universidad como esta.

Mil gracias para todos.

DEDICATORIA.

*Este trabajo de diploma va dedicado a todas esas personas que son importantes
en mi vida.*

A mis padres que son mi fuente de inspiración y la razón de todos mis logros.

*Mis abuelos Toni, Mayí y Tata que siempre me han escuchado con amor y me
han dado su apoyo incondicional.*

A mi primo Angel Arturo, al que se que sirvo de ejemplo y guía.

*A toda mi familia en sentido general que cada día ha estado al pendiente de mis
resultados.*

*A todos esos amigos que siempre me han tendido su mano cuando los he
necesitado.*

Todos mis triunfos también son de ustedes.

RESUMEN.

Aprovechando el progreso actual de las tecnologías y las necesidades que tienen las entidades para organizar su información surgen sistemas para la Gestión del Conocimiento. La Aduana General de la República de Cuba (AGR) es una de las que no ha quedado ajena a este proceso, utilizando dicha gestión en el Departamento de Despacho Comercial para realizar la Clasificación de las Mercancías. En el presente trabajo se hace una propuesta de metodología para el desarrollo de ontologías para apoyar la gestión del conocimiento en dicho departamento. Primeramente se analizan algunas de las herramientas y metodologías existentes para construir las, se definen una serie de pasos y se especifican los elementos necesarios como roles, artefactos y fases para su creación. Dicha propuesta se pone en práctica con la implementación de una ontología para la Clasificación de Mercancías, donde se define como dominio las máquinas, aparatos y materiales eléctricos. Se emplean como herramientas: el lenguaje OWL, Protégé como editor y Pellet como razonador.

PALABRAS CLAVE

Gestión del Conocimiento, Ontología, Metodología, Aduana, Clasificación de Mercancía.

TABLA DE CONTENIDOS.

AGRADECIMIENTOS.....	I
DEDICATORIA.....	II
RESUMEN.....	III
CAPÍTULO 1 : FUNDAMENTACIÓN TEÓRICA.....	5
1.1 INTRODUCCIÓN AL CAPÍTULO.....	5
1.2 GESTIÓN DEL CONOCIMIENTO.....	5
1.2.1 <i>Etapas en el proceso de gestión del conocimiento.</i>	6
1.2.2 <i>Procesos Fundamentales de la Gestión del Conocimiento.</i>	7
1.3 FORMAS DE REPRESENTAR EL CONOCIMIENTO.....	8
1.3.1 <i>Ontologías.</i>	10
1.3.1.1 <i>Conceptos de ontología.</i>	10
1.3.1.2 <i>Objetivo de las ontologías.</i>	12
1.3.1.3 <i>Ventajas de utilizar una ontología.</i>	13
1.3.1.4 <i>Tipos de ontología.</i>	14
1.4 METODOLOGÍAS PARA REALIZAR ONTOLOGÍAS.	15
1.4.1 <i>The Cyc</i>	16
1.4.2 <i>Uschold and King's.</i>	16
1.4.3 <i>Grüninger and Fox's.</i>	17
1.4.4 <i>Kactus.</i>	18
1.4.5 <i>Sensus.</i>	19
1.4.6 <i>On-To-Knowledge.</i>	19
1.4.7 <i>Methontology.</i>	20
1.5 LENGUAJES DE PRESENTACIÓN.	23
1.5.1 <i>OIL (Ontology Inference Layer).</i>	24
1.5.2 <i>DAML + OIL (DARPA's Agent Markup Language + OIL).</i>	24
1.5.3 <i>OWL (Web Ontology Language).</i>	25

Tabla de Contenidos.

1.6 HERRAMIENTAS PARA LA CONSTRUCCIÓN DE ONTOLOGÍAS.....	26
1.6.1 KAON.....	26
1.6.2 <i>OntoEdit</i>	27
1.6.3 <i>Protégé</i>	27
1.7 RAZONADORES SEMÁNTICOS.....	28
1.7.1 <i>Cerebra</i>	29
1.7.2 <i>RACER</i>	29
1.7.3 <i>FaCT ++</i>	30
1.7.4 <i>Pellet</i>	30
1.8 CONCLUSIONES PARCIALES.....	31
CAPÍTULO 2 : DESCRIPCIÓN DE METODOLOGÍA PROPUESTA.....	32
2.1 INTRODUCCIÓN AL CAPÍTULO.....	32
2.2 METODOLOGÍA SELECCIONADA.....	32
2.3 PROCESO DE CLASIFICACIÓN DE MERCANCIAS.....	33
2.4 COMPONENTES DE UNA ONTOLOGÍA.....	34
2.5 ROLES EN EL DESARROLLO DE LA ONTOLOGÍA.....	36
2.6 FASES PARA EL DESARROLLO DE LA ONTOLOGÍA.....	36
2.6.1 <i>Análisis de la información del negocio</i>	37
2.6.2 <i>Análisis de los datos del dominio</i>	40
2.6.3 <i>Implementación del Modelo Conceptual</i>	43
2.6.4 <i>Validación de la ontología</i>	44
2.7 CONCLUSIONES PARCIALES.....	46
CAPÍTULO 3 : DESARROLLO DE LA ONTOLOGÍA.....	47
3.1 INTRODUCCIÓN AL CAPÍTULO.....	47
3.2 ANÁLISIS DE LA INFORMACIÓN DEL NEGOCIO.....	47
3.2.1 <i>Entendimiento del negocio de la Clasificación de Mercancía</i>	47
3.2.2 <i>Planificar las tareas a desarrollar y responsables</i>	48
3.2.3 <i>Establecer Dominio y Alcance</i>	48
3.2.4 <i>Preguntas de Competencia</i>	49

Tabla de Contenidos.

3.2.5 Selección del Lenguaje y Editor Gráfico.....	50
3.3 ANÁLISIS DE LOS DATOS DEL DOMINIO.....	50
3.3.1 Definir términos del capítulo seleccionado.....	50
3.3.2 Definir productos, partidas y subpartidas que pertenecen al capítulo.....	51
3.3.3 Definir jerarquía entre los productos, partidas y subpartidas.....	52
3.3.4 Asociar propiedades a los productos.....	53
3.3.5 Definir relación entre productos, partidas y subpartidas.....	55
3.4 IMPLEMENTACIÓN DEL MODELO CONCEPTUAL.....	56
3.4.1 Definir Axiomas Formales.....	56
3.4.2 Crear Instancias.....	57
3.5 VALIDACIÓN DE LA ONTOLOGÍA.....	58
3.5.1 Pruebas de Consistencia.....	59
3.5.2 Pruebas de Razonamiento.....	59
3.5.3 Análisis de Resultados.....	63
3.6 CONCLUSIONES PARCIALES.....	63
CONCLUSIONES.....	64
RECOMENDACIONES.....	65
REFERENCIAS BIBLIOGRÁFICAS.....	66
GLOSARIO DE TÉRMINOS.....	69

INTRODUCCIÓN.

La Aduana General de la República de Cuba (AGR) se encarga de garantizar la seguridad y protección de la sociedad socialista y de la economía nacional, así como de la recaudación fiscal y las estadísticas del comercio exterior.

Pese a la importancia que tienen los procesos que allí se desarrollan todavía prevalecen insuficiencias en algunos de los departamentos que la componen, ejemplo de esto es Depósito de Aduana, Enfrentamiento, Despacho Comercial donde ocurre el proceso de Clasificación de Mercancías, entre otros.

Como solución a algunos de estos problemas surge el Departamento de Soluciones para la Aduana llevando a cabo el desarrollo del producto GINA, el cual tiene como misión fundamental ofrecer servicios integrales de gestión a la AGR obteniendo resultados satisfactorios. No obstante, existen procesos que aún no se realizan de la manera deseada. Uno de los que más carencia presenta es el proceso de Clasificación de Mercancías.

Este proceso comienza en las entidades que importan productos al país, las que son encargadas de la confección de la Declaración de Mercancía (DM). Al llegar la misma al Despacho Comercial se compara la clasificación y la descripción del producto que esta contiene, con la descripción que provee el Sistema Armonizado¹ (SA). Al comprobar que el producto finalmente es el que debe ser y se encuentra en la subpartida definida para él se le da paso a las otras actividades para la aceptación de la DM.

En un diagnóstico realizado al Despacho Comercial donde ocurre el proceso, se pudo comprobar que la clasificación es un proceso difícil de realizar en muchas ocasiones. Las clasificaciones se encuentran en grandes volúmenes de información, los documentos que rigen el proceso están escritos de manera ambigua, la descripción de los productos contenida en la DM muchas veces no tiene la suficiente información, lo que en muchas ocasiones hace que el trabajo se vuelva un poco engorroso.

¹ SA: documento que rige el proceso de clasificación de mercancías.

Para apoyar el proceso de Clasificación de Mercancías es recomendable la utilización de la gestión del conocimiento, disciplina que toma cada vez más fuerza en el desarrollo de software. La misma se encarga de buscar, filtrar y organizar la información con el objetivo de mejorar el entendimiento, comprensión y gestión de los procesos llevados a cabo en las organizaciones. (Salazar, 2000).

Para llevar a cabo dicha gestión es necesario almacenar el conocimiento de una forma entendible para las personas y los sistemas computacionales debido a que de allí será de donde se obtendrá la información para ser empleada en la obtención de nuevos conocimientos. Este puede ser representado de diferentes formas: como imágenes mentales en los pensamientos, como palabras habladas o escritas en algún lenguaje, como imágenes gráficas, como colección de señales electrónicas o magnéticas dentro de un computador. (Salazar, 2000)

Para almacenar el conocimiento es necesario determinar su representación, permitiendo estructurar el conocimiento sobre un dominio de aplicaciones de modo que sea más fácil resolver ciertos problemas. Una de estas formas de representar el conocimiento que mayor impacto ha tenido son las ontologías. Estas permiten mejorar la comunicación entre personas, organizaciones y aplicaciones, lograr la interoperabilidad entre sistemas informáticos y razonamiento automáticamente sin la intervención humana.

Sin embargo el desarrollo de ontologías se reconoce como actividad compleja. La construcción de ontologías no responde a una única aproximación lógica sino que depende, en gran medida, del contexto en el que se construyen, de forma que ha sido norma común que las diferentes organizaciones que desarrollaban ontologías siguiesen sus propios principios, criterios, reglas o métodos, dependiendo del tipo de ontología a desarrollar o de una situación particular. (Peñalvo, 2006)

Ante esta diversidad, se han creado una serie de métodos de diseño y construcción, que tienen por finalidad proporcionar un procedimiento comúnmente aceptado, ampliamente validado y verificado, que garantice la adquisición de un producto exitoso.

En la AGR es necesario recopilar el conocimiento contenido en el Sistema Armonizado (SA) de alguna forma que permita posteriormente realizar consultas sobre el mismo de una manera más sencilla y rápida, con el objetivo de que el proceso de Clasificación de Mercancías se realice con la calidad requerida.

Por lo anteriormente expuesto se identifica el siguiente *Problema de Investigación*: La complejidad del desarrollo de ontologías limita su aplicación en el proceso de Clasificación de Mercancías.

Por lo que la *Idea a Defender* Si se cuenta con una metodología que disminuya la dificultad del desarrollo de ontologías para el proceso de Clasificación de Mercancías, sería más fácil su aplicación.

En correspondencia con el problema, la investigación tiene como *Objeto de Estudio* Las ontologías como forma de representación del conocimiento. Centrando su *Campo de Acción* en las Metodologías para el desarrollo de ontologías.

Para dar solución al problema planteado, se define enmarcar *Objetivo General* en Definir una metodología para el desarrollo de ontologías para la Clasificación de Mercancías.

Para dar cumplimiento al objetivo general se proponen los siguientes *Objetivos Específicos*:

1. Determinar el marco teórico de la investigación que permita identificar los elementos relevantes de la gestión del conocimiento, formas de representar el conocimiento; metodologías empleadas en la elaboración de ontologías.
2. Desarrollar una metodología para el diseño de ontología que permita la Clasificación de Mercancías.
3. Aplicar la metodología desarrollada en un caso de estudio que permita validar la calidad de la misma.

Las *Tareas de investigación* que se proponen son:

1. Análisis correspondiente a la gestión del conocimiento y formas de representar el mismo.

2. Establecer un diagnóstico de las tendencias actuales y metodologías empleadas en la elaboración de ontologías.
3. Realizar un estudio detallado sobre la Clasificación de Mercancías.
4. Definir una metodología que permitirá establecer un ciclo de desarrollo de la ontología de acuerdo al dominio establecido.
5. Desarrollar una ontología relacionada con la Clasificación de Mercancías en la AGR.
6. Evaluar del correcto funcionamiento de la metodología propuesta a través de la validación de la ontología desarrollada.

Al concluir la investigación se debe tener como *Resultado*:

Una metodología que guíe el desarrollo de ontologías relacionadas con la Clasificación de Mercancías.

El presente trabajo consta de introducción, tres capítulos, conclusiones, recomendaciones, bibliografía, referencias bibliográficas, glosario de términos y anexos.

Capítulo 1: Fundamentación Teórica. Se abordan los principales conceptos para el entendimiento del tema, además se hace un estudio sobre las metodologías, y herramientas necesarias en el desarrollo de las ontologías.

Capítulo 2: Descripción de la metodología propuesta. Se establecen las etapas o fases que componen la metodología propuesta, así como los principales artefactos que se deben generar en cada una de las actividades a desarrollar.

Capítulo 3: Aplicación de la metodología. Se pone en práctica la propuesta de metodología para llevar a cabo la implementación de la ontología, con un pequeño caso de estudio de la ontología relacionada con la Clasificación de las Mercancías en la Aduana.

CAPÍTULO 1 : FUNDAMENTACIÓN TEÓRICA.

1.1 Introducción al Capítulo.

En el capítulo se explican algunos elementos y términos significativos en el desarrollo de ontologías. Se refleja la definición y explicación de la Gestión del Conocimiento (GC), las formas de representar el conocimiento que se pueden emplear, dentro de ellas las ontologías, sus diferentes tipos, ventajas y lenguajes usados en su representación. Se hace además una evaluación de varias metodologías y herramientas que existen para la construcción y validación de ontologías.

1.2 Gestión del Conocimiento.

El conocimiento, tal como se entiende hoy día, es un recurso que no tan sólo permite interpretar el entorno, sino que da la posibilidad de actuar. Es un recurso que se halla en las personas y en los objetos –físicos o no– que estas personas utilizan, pero también en las organizaciones a las que pertenecen, en los procesos y en los contextos de dichas organizaciones. (Canals, 2003).

La gestión del conocimiento no es más que: "[...] *el proceso sistemático de buscar, organizar, filtrar y presentar la información con el objetivo de mejorar la comprensión de las personas en una específica área de interés*". (Figuerola, y otros, 2006).

Existen dos tipos de conocimiento: *el explícito y el tácito*. El primero se ha definido como la noción objetiva y racional que puede ser expresada con palabras, números, fórmulas, etc. El segundo es aquel que una persona, comunidad, organización o país, tiene incorporado o almacenado en su mente, en su cultura y es difícil de explicar. (Figuerola, y otros, 2006).

En el proceso de Clasificación de Mercancías es necesario organizar el conocimiento de forma tal que cuando se vaya a realizar la búsqueda de un producto en el SA, la misma se realice de una forma rápida. Uno de los objetivos del presente trabajo es buscar la forma de recolectar el conocimiento contenido de forma explícita en el SA, junto con el tácito que puede estar en los miembros del departamento donde

Capítulo 1: Fundamentación Teórica.

ocurre el proceso de Clasificación de Mercancías, permitiendo así acumular esta información para hacer del proceso, una actividad un poco más sencillo.

1.2.1 Etapas en el proceso de gestión del conocimiento.

En las organizaciones actuales surge la necesidad de llevar a cabo el proceso de gestión del conocimiento permitiendo filtrar contenido importante y aprovechar al máximo el valor intangible que se encuentra en la misma. Para llevar adelante el proceso se han planteado etapas necesarias.

Crear conocimiento: implica exploración, combinación y el descubrimiento de conocimiento mediante el hacer. Los individuos internamente de una organización crean nuevos conocimientos mediante conexiones intuitivas de las ideas existentes o a través de la interacción con otros individuos de la organización. (Rojas Mesa, 2006).

Compartir conocimiento: se produce cuando los individuos internamente de una organización transfieren y comparten el conocimiento. Al compartir conocimiento, éste se incrementa y llega a ser más valioso, se producen asociaciones que hacen que el total del conocimiento alcanzado sea mayor cuantitativa y cualitativamente que la suma de los conocimientos individuales. (Rojas Mesa, 2006).

Aplicar conocimiento: implica convertir el conocimiento en un resultado valioso para la organización. La aplicación de conocimiento puede conducir a la generación de productos y servicios, pero también a la generación de ideas que llevan a una mejor toma de decisiones estratégicas. (Rojas Mesa, 2006).

Sin duda, este es un desafío que no puede estar ausente para las instituciones de educación superior. Las universidades tienen como tarea fundamental crear y difundir conocimiento, más aun en el caso particular de la Universidad de las Ciencias Informáticas (UCI), donde se acumula un número grande de información de empresas y organizaciones, debido a que los productos que en ella se desarrollan tienen por finalidad prestar servicios a las mismas.

La presente investigación propone una serie de pasos para hacer más fácil el proceso de aplicación del conocimiento, aplicando determinadas formas y procesos que se estudian más adelante en la misma. El

Capítulo 1: Fundamentación Teórica.

conocimiento contenido en grandes volúmenes de información, llevarlo a una aplicación informática que de forma sencilla se pueda consultar y así favorecer la toma de decisiones en cuanto a una clasificación de un producto.

1.2.2 Procesos Fundamentales de la Gestión del Conocimiento.

En la Gestión del Conocimiento (GC) existen dos procesos fundamentales, la creación y la transmisión de conocimiento (Canals, 2003):

- **La creación** es un *proceso de desarrollo de competencias y habilidades de los individuos*, donde se propicia el establecimiento de un ambiente que favorezca el surgimiento de nuevas ideas para fomentar la innovación y de esta forma, generar soluciones que contribuyan al progreso de la sociedad en general.
- **La transmisión** se puede ver cuando se intenta poner de forma explícita el conocimiento en una base de datos, en este sentido el conocimiento se coloca allí para que al cabo del tiempo alguien pueda consultarla, de esta forma se está transmitiendo en el tiempo. Cuando se utilizan herramientas de comunicación se intenta transmitir el conocimiento en el espacio. Las intranets, los portales, las bases de datos relacionales y las documentales son ejemplos de estos instrumentos.

Estos dos procesos están totalmente interrelacionados, la creación de conocimiento no es algo que se hace partiendo de la nada, sino que para crearlo se utiliza el conocimiento que proviene de otras personas y de otros lugares, por lo tanto, ha habido un proceso de transmisión previo. La creación y la transmisión son procesos que están muy interrelacionados y que juntos hacen que el conocimiento contenido dentro de las organizaciones mejore. (Canals, 2003).

La GC permite aprovechar el conocimiento adquirido por las organizaciones para utilizarlos adecuadamente y también protegerlos como parte de su propiedad intelectual, individual y colectiva, además de incrementarlos. Cumpliendo estos principios o normas, la GC puede transformar la

Capítulo 1: Fundamentación Teórica.

organización, aportándole nuevos valores a la misma y por ende a la sociedad a la que brinda sus servicios en función de la razón de ser de ella (Canals, 2003).

1.3 Formas de Representar el Conocimiento.

El conocimiento es importante y primordial en todas las aplicaciones, su representación constituye una de las máximas prioridades de la investigación en Inteligencia Artificial, la cual es *“la ciencia e ingeniería de hacer máquinas inteligentes, especialmente programas de cómputo inteligentes”*. (Osorio Galindo, y otros, 2009).

El conocimiento puede ser representado como imágenes mentales en los pensamientos, como palabras habladas o escritas en algún lenguaje, como forma gráfica o en imágenes, como cadenas de caracteres o colecciones de señales eléctricas o magnéticas dentro de un computador. (Salazar, 2000).

Para realizar el almacenamiento del conocimiento se determina su representación, lo cual no es más que el proceso de estructurar el conocimiento sobre un dominio de aplicación de modo que los problemas sean más fáciles de resolver. Esta representación debe expresar con claridad tanto los hechos del dominio, como las relaciones entre los mismos. (Salazar, 2000).

No existe actualmente una forma de representación del conocimiento (FRC) general capaz de ser usada en todo tipo de aplicación informáticas con éxito; las formas disponibles están limitadas a ciertos dominios específicos.

Entre las primeras y más usadas formas de representar el conocimiento se encuentran las redes semánticas definidas como *“una forma de representación del conocimiento lingüístico en que a las interrelaciones entre diversos conceptos o elementos semánticos se les da la forma de un grafo cuyos nodos pueden ser entidades, objetos, eventos, atributos o estados y los arcos son relaciones entre ellos”* (Ortiz, 2008)

Capítulo 1: Fundamentación Teórica.

Las redes semánticas cuentan con diferentes categorías como son: redes *es un*, muchas veces se menciona este tipo como sinónimo de "red semántica". Redes en las que los puntos de unión de los enlaces son parte de la etiqueta del nodo y los grafos conceptuales, en los que existen dos tipos de nodos: de conceptos y de relaciones.

Una red "*es un*" es una jerarquía taxonómica cuya espina dorsal está constituida por un sistema de enlaces de herencia entre los objetos o conceptos de representación, conocidos como nodos. Estos enlaces o arcos pueden estar etiquetados como "*es un*", también "*superclase*", "*subconjunto*", entre otros. Los restantes tipos de redes semánticas son en realidad especializaciones de redes "*es un*".

Las redes *es un* son el resultado de la observación de gran parte del conocimiento humano, se basa en la adscripción de un subconjunto de elementos como parte de otro conocimiento más general.

Son muy flexibles, pero también tienen algunos problemas y desventajas como pueden ser:

- La elección de los nodos y arcos es crucial en la fase de análisis, una vez se ha decidido una estructura determinada.
- Dificultad para expresar cuantificación.

Esto llevó a idear otros esquemas de representación con una estructura más compleja que simples nodos y arcos, que fuesen capaces de dar cabida a éstas y otras situaciones, surgiendo por este motivo los marcos.

Los marcos representan una evolución de las Redes Semánticas, debido a que el nodo fue sustituido por una estructura de datos que representa una situación estereotipada a partir de sus elementos más significativos, llamado marco. Las redes semánticas tienen menos expresividad que los marcos, dado que no pueden representar, por ejemplo, valores por defecto, y restricciones de cardinalidad sobre los atributos. Los sistemas basados en marcos, y las redes semánticas son considerados como una sola familia de formalismos. (Socorro, 2009).

Capítulo 1: Fundamentación Teórica.

El objetivo final de las redes semánticas y los sistemas de marcos es que se pueda inferir información con los datos existentes, siendo así una evolución para las formas de representar el conocimiento. Para lograr este objetivo surgen las ontologías, donde a través de diversas combinaciones de información, logrando la inferencia son capaces de facilitar el análisis de datos.

Las ontologías son una solución recomendada para representar el contenido del Sistema Armonizado, debido a que es necesario después de tener el conocimiento del mismo, lograr crear nuevos conocimientos para obtener mejores resultados en el proceso de Clasificación de Mercancías.

1.3.1 Ontologías.

1.3.1.1 Conceptos de ontología.

La palabra ontología fue tomada de la filosofía, “*es la parte de la filosofía que investiga en qué consiste el ser y cuáles son los ámbitos o regiones del ser fundamentales.*”. (Olleta, 2002). En las últimas décadas, esta palabra se ha convertido en relevante para la comunidad de la Ingeniería del Conocimiento. En el año 1995 los autores Guarino y Giaretta proponen usar “Ontología” con mayúscula para referirse a la Filosofía y “ontología” con minúscula para la Ingeniería del Conocimiento. En el presente trabajo se toma como referencia la definición propuesta por dichos autores.

Las ontologías han sido tradicionalmente usadas como modelo de representación del conocimiento en Inteligencia Artificial. Existen muchas definiciones de lo que es ontología y éstas han evolucionado a lo largo del tiempo. Una de las primeras definiciones de ontología fue propuesta por Neches en 1991 en la cual plantea “*Una ontología define los términos básicos y las relaciones que comprende el vocabulario de un área temática, así como las reglas para combinar términos y relaciones para definir extensiones al vocabulario*” (Corcho, y otros, 2004). Unos años más tarde Gruber plantea que una ontología es “*una especificación explícita y formal de una conceptualización compartida*” (Studer, y otros, 1998). Esta definición se convirtió en la más citada en la literatura y por la comunidad de desarrolladores de ontologías. Sobre la base de esta definición se fueron creando otras haciéndole algunas modificaciones a la original.

Capítulo 1: Fundamentación Teórica.

En 1998 Guarino y Giaretta proponen una nueva definición basándose en lo expuesto por Studer y refinando la propuesta en 1995.

Dado que las ontologías son ampliamente utilizadas para diferentes propósitos, surgen otro grupo de definiciones basadas en el proceso seguido para construir ontologías. Ejemplo de esto es la definición que Uschold y Jasper proporcionaron en 1999 *“Una ontología puede tomar una variedad de formas, pero necesariamente incluirá un vocabulario de términos y una especificación de su significado. Esto incluye definiciones y una indicación de cómo los conceptos están relacionado entre sí, que de forma conjunta imponen una estructura en el dominio y permite restringir las posibles interpretaciones de los términos”*. (Corcho, y otros, 2004).

A través de los años los conceptos de ontología han ido evolucionando, haciéndose cada día más entendibles y aceptables, por lo que su utilización ha ido en aumento en la última década para proporcionar entendimiento común de los diferentes términos y estructuras de un dominio específico destacándose como principal concepto el expuesto por Guarino y Giaretta en 1998 *“un conjunto de axiomas lógicos para dar cuenta del significado del vocabulario”*. (Corcho, y otros, 2004). Pero siendo utilizado para la presente investigación el propuesto por Gruber.

Donde, Conceptualización se refiere a una forma de entender y describir un dominio. Según (Tramullas, 2003), corresponde a una parte del mundo o universo que es objeto de tratamiento, mientras que para desde una perspectiva tecnológica, es un conjunto de reglas informales que restringen la estructura de una parte de la realidad, las que son usadas por los agentes para separar y agrupar objetos y relaciones relevantes. Compartida porque debe ser consensuada y aceptada por un grupo o comunidad científica, explícita porque se define el conocimiento explícito que existe sobre determinada noción y esa definición es a través del lenguaje natural. Formal, porque debe ser legible por computadora, es decir, desarrollada y puesta en marcha a través de lenguajes computacionales. (Guarino, 1998).

La propuesta de la investigación tiene como objetivo crear una metodología que cumpla con estas condiciones. Debe permitir describir el dominio relacionado a la Clasificación de Mercancías, de forma tal que el conocimiento contenido en grandes volúmenes de información y en las experiencias alcanzadas

Capítulo 1: Fundamentación Teórica.

por los trabajadores del departamento donde ocurre este proceso, se represente de forma explícita y sea estructurado para que sea legible por un sistema computacional.

1.3.1.2 Objetivo de las ontologías.

Las ontologías buscan capturar su conocimiento consensuado de una manera genérica, y que pueden ser reutilizados y compartidos a través de aplicaciones de software y por grupos de personas. Por lo general son construidas en cooperación entre diferentes grupos de personas en diferentes lugares.

Partiendo de Noy y McGuinness, se puede decir que las ontologías tienen como principales objetivos:

- 1) Compartir la comprensión común de la estructura de información entre personas o agentes de software, lo que debe revertir de forma positiva y casi necesaria la extracción y recuperación de información, como por ejemplo en páginas web, de contenidos conectados temáticamente.
- 2) Permitir la reutilización del conocimiento perteneciente a un dominio; por ejemplo, a la hora de iniciar la elaboración de una ontología.
- 3) Permite hacer explícitos los supuestos de un dominio. Esta afirmación puede conducir a conclusiones muy interesantes para la representación del conocimiento más allá de consideraciones técnicas, operativas e informáticas.
- 4) Separa el conocimiento de un dominio del que se puede denominar operacional. En ocasiones, el conocimiento que se está representando se puede implicar en diferentes áreas de un proceso.
- 5) Hace posible analizar el conocimiento de un campo, por ejemplo en lo que se refiere al estudio de los términos y relaciones que lo configuran ya sea formalmente o no.

Se apuntan varias utilidades de las ontologías, entre las que destaca en primer lugar, la mejora en la comunicación, enfocadas a reducir la confusión terminológica y conceptual en un único marco de trabajo. En segundo lugar, la interoperabilidad. Las ontologías, en principio, han de potenciar el intercambio de datos en contextos informáticos y digitales gracias a los fundamentos semánticos que se encuentran en ellas (Noy, y otros, 2005).

Capítulo 1: Fundamentación Teórica.

1.3.1.3 Ventajas de utilizar una ontología.

Las ontologías se usan para favorecer la comunicación entre personas, organizaciones y aplicaciones, lograr la interoperabilidad entre sistemas informáticos y razonar automáticamente.

En la ciencia y en la tecnología los problemas derivados de la falta de comprensión común entre personas, organizaciones y aplicaciones adquieren una gran importancia. Las ontologías proporcionan una comprensión común de un determinado dominio y eliminan estas confusiones de conceptos y términos, además permiten que los sistemas interoperen. Dos sistemas son interoperables si pueden trabajar conjuntamente de una forma automática, sin esfuerzo por parte del usuario. (López, 2002).

Las ontologías resultan muy útiles para facilitar el razonamiento automático, es decir, sin intervención humana, usando para ello reglas de inferencia. Un motor de razonamiento puede usar los datos de las ontologías para inferir conclusiones de ellos. Una de las aplicaciones más importantes del razonamiento automático es la validación de datos.

También se usa para establecer relaciones entre ontologías y así poder descubrir relaciones ocultas o inesperadas entre los datos e integrar esquemas de bases de datos.

En la Ingeniería del Software, las ontologías ayudan a la especificación de los sistemas de software. Como la falta de un entendimiento común conduce a dificultades en identificar los requisitos y especificaciones del sistema que se desea desarrollar, las ontologías facilitan el acuerdo entre desarrolladores y usuarios. (López, 2002).

En la solución del problema planteado sobre la Clasificación de Mercancías las ontologías tienen un papel importante porque:

- 1) Proporcionan una forma de representar y compartir el conocimiento del Sistema Armonizado utilizando un vocabulario común.
- 2) Permiten usar un formato estándar de intercambio de conocimiento.

Capítulo 1: Fundamentación Teórica.

- 3) Posibilitan el entendimiento común por las personas encargadas de trabajar con el Sistema Armonizado.

1.3.1.4 Tipos de ontología.

Las ontologías se pueden clasificar de acuerdo a la cantidad y tipo de estructura de la conceptualización en (Castells, 2004):

- **Ontologías terminológicas:** Especifican los términos que son usados para representar conocimiento en el universo del discurso. Suelen ser usadas para unificar vocabulario en un campo determinado.
- **Ontologías de información:** Especifican la estructura de almacenamiento de bases de datos. Ofrecen un marco para almacenamiento estandarizado de información.

Según su dependencia y relación con una tarea específica, las ontologías tienen la siguiente clasificación (Corcho, y otros, 2004):

- **Ontologías de Nivel Superior:** Describen conceptos muy generales y dan nociones generales en las que todos los términos raíz en ontologías existentes deben ser enlazados.
- **Ontologías de Dominio:** Son reutilizables en un determinado dominio específico (medicina, farmacia, ingeniería, derecho, empresa, automóvil, etc.) Estas ontologías proporcionan un vocabulario de conceptos dentro de un dominio y sus relaciones, sobre las actividades que tienen lugar en él, y sobre las teorías y los principios elementales que lo rigen. Los conceptos en ontologías de dominio son por lo general las especializaciones de los conceptos ya definidos en las ontologías de nivel superior, y lo mismo podría ocurrir con las relaciones.
- **Ontologías de Tareas:** Proporcionan un vocabulario sistemático de los términos empleados para resolver los problemas asociados a las tareas que pueden o no pertenecer al mismo dominio.

Capítulo 1: Fundamentación Teórica.

- **Ontologías de Aplicación:** Dependen de la aplicación. Contienen todas las definiciones necesarias para modelar el conocimiento necesario para una aplicación particular. A menudo se extienden y se especializan en el vocabulario del dominio y de las ontologías de tareas para una aplicación dada.

La ontología a implementar en el trabajo de diploma se clasifica según la cantidad y tipo de estructura de la conceptualización en ontología terminológica y según su dependencia y relación con una tarea específica en ontología de dominio. La misma debe ser capaz de representar el dominio relacionado con en proceso de Clasificación de Mercancías, proporcionando así un vocabulario y entendimiento común a través de la representación de los términos del dominio.

1.4 Metodologías para realizar ontologías.

Desde mediados de la década '80 los profesionales de la investigación en la rama de la Inteligencia Artificial comenzaron a mostrar un marcado interés por el desarrollo de las ontologías. Para su elaboración no existían aún directrices definidas, por lo que cada equipo de desarrollo definía sus pasos a seguir para la creación de ontologías desde cero, para la reutilización y otros para el uso de métodos semiautomáticos que reducen el cuello de botella en la adquisición de conocimientos del proceso de desarrollo de ontologías, haciendo de esto un proceso artesanal más que una actividad de ingeniería.

Algunos de los métodos definidos fueron The Cyc en la primera mitad de la década de 1980 por Corporación de Microelectrónica y Tecnología de Computadores (MCC) por sus siglas en inglés. Algunos años más tarde, en 1995, sobre la base de la experiencia adquirida en el desarrollo de ontología de empresa en el dominio de modelado empresarial, fueron propuestas las primeras directrices en el método Uschold y King. En 1997 un nuevo método fue propuesto para la construcción de ontologías basado en la ontología Sensus (Corcho, y otros, 2004).

Estos métodos dieron paso a la creación de metodologías propias para el desarrollo de ontologías. El presente trabajo de tesis analiza algunas de estas metodologías para tomar alguna como base para la solución propuesta.

Capítulo 1: Fundamentación Teórica.

1.4.1 The Cyc

En la mitad de la década de 1980, Corporación de Microelectrónica y Tecnología de Computadores (MCC) comenzó a crear Cyc, una base de conocimiento amplio con conocimiento de sentido común.

Para llevar a cabo Cyc, se utilizó el lenguaje CycL. Este es un lenguaje híbrido que combina marcos con cálculo de predicados. El motor de inferencia CycL permite: la herencia múltiple, clasificación automática, el mantenimiento de los vínculos inversos, la comprobación de restricciones, la búsqueda primero el mejor, etc. Además tiene un sistema de mantenimiento, y un módulo de resolución.

La razón por la que Cyc puede considerarse como una ontología es porque puede ser usado como base para la construcción de diferentes sistemas inteligentes que puedan comunicarse e interactuar.

Consta de tres procesos fundamentales en su desarrollo:

Proceso 1: Proceso manual de codificación de artículos y piezas de conocimiento.

Proceso 2: Conocimientos de programación con la ayuda de herramientas utilizando el conocimiento que ya está almacenado en la base de conocimiento Cyc.

Proceso 3: La codificación del conocimiento llevada a cabo principalmente por las herramientas de utilización de los conocimientos ya almacenados en la base de conocimiento Cyc (Corcho, y otros, 2004).

1.4.2 Uschold and King's.

Uschold and King propuso el primer método para la construcción de ontologías en 1995. Definió algunas directrices en base a su experiencia en el desarrollo de ontología de empresa. Esta ontología se desarrolló como parte del proyecto del Instituto de Aplicaciones de Inteligencia Artificial de la Universidad de Edimburgo con sus socios de IBM, Lloyd's Register, Logica UK Limited, y Unilever (Corcho, y otros, 2004).

Para construir una ontología según el enfoque Uschold and King, se deben llevar a cabo los siguientes procesos (Corcho, y otros, 2004):

Capítulo 1: Fundamentación Teórica.

Proceso 1: Identificar el propósito de la ontología.

Proceso 2: Construir la ontología.

Proceso 3: Evaluar la ontología.

Proceso 4: Documentar la ontología.

Durante el proceso de construcción, se propone capturar el conocimiento, la codificación y la integración de ontologías con otros dentro de la ontología actual.

1.4.3 Grüninger and Fox's.

Basándose en la experiencia del proyecto Tove² en el dominio de empresa, que se desarrolló en la Universidad de Toronto, Grüninger and Fox en 1995 publicaron un enfoque formal para construir y evaluar ontologías. Esta metodología se ha utilizado para construir las ontologías Tove, un entorno de gráfico que permite al usuario explorar una variedad de diseños de empresa.

La metodología Grüninger and Fox se inspira en el desarrollo de sistemas basados en el conocimiento mediante la lógica de primer orden. Estos autores proponen identificar intuitivamente los escenarios principales, las aplicaciones posibles en los que se utiliza la ontología. A continuación, una serie de preguntas en lenguaje natural, llamadas preguntas de competencia, se utilizan para determinar el alcance de la ontología. Estas preguntas y sus respuestas son utilizadas para extraer los conceptos principales y sus propiedades, relaciones y axiomas formales de la ontología. Por otro lado, el conocimiento se expresa formalmente en la lógica de primer orden. Esta es una metodología muy formal que se aprovecha de la solidez de la lógica clásica y puede ser utilizado como una guía para transformar los escenarios informales en los modelos computables.

² Tove: Ontología Toronto Virtual Enterprise

Capítulo 1: Fundamentación Teórica.

La metodología Grüninger and Fox, se puede decir que está bien estructurada para la construcción y la evaluación de ontologías, pero tiene como desventaja que no cuenta con algunas actividades de gestión y apoyo. (Corcho, y otros, 2004).

1.4.4 Kactus.

Este enfoque fue propuesto por Amaya Bernaras y sus colegas en 1996, dentro del proyecto Esprit KACTUS. Uno de los objetivos de este proyecto fue investigar la factibilidad de la reutilización del conocimiento en sistemas técnicos complejos y el papel de las ontologías para apoyarlo.

Este enfoque de desarrollo de ontologías está condicionado por el desarrollo de aplicaciones. Por lo tanto, cada vez que una aplicación se construye, la ontología que representa los conocimientos necesarios para la aplicación es refinada. La ontología puede ser desarrollada por otros y reutilizada, se puede integrar en ontologías de aplicaciones más tarde. Por lo tanto, cada vez que una aplicación se desarrolla, ocurren los siguientes procesos (Corcho, y otros, 2002):

Proceso 1: Especificación de la aplicación.

Proceso 2: Diseño preliminar basado en las categorías ontológicas de nivel superior.

Proceso 3: Refinamiento y estructuración de la ontología.

La ontología de aplicación se puede considerar como la unión entre la ontología de dominio para el diagnóstico y la planificación de la recuperación del servicio.

Dicha metodología surge para la creación de ontologías de aplicación, se crean a partir de un dominio específico y su aplicación. El objetivo de la metodología es refinar dicha ontología para las posibles utilidades que podría tener, al no contar con una ontología base del dominio que se trata en el presente trabajo, ni una aplicación, dichos procesos no solucionan el problema planteado.

Capítulo 1: Fundamentación Teórica.

1.4.5 Sensus.

El método SENSUS se presenta para construir el esqueleto de la ontología de dominio a partir de una ontología enorme, la ontología Sensus. El método propone relacionar términos de dominio específico de la ontología enorme y podar, en la ontología enorme, esos términos que son irrelevantes para la nueva ontología que se quiere construir.

El resultado de este proceso es el esqueleto de la nueva ontología, que se genera de forma automática mediante un procedimiento y la herramienta OntoSaurus. Según este método, para construir una ontología en un dominio específico se deben tener los procesos (Corcho, y otros, 2002):

Proceso 1: Identificar los términos semillas.

Proceso 2: Vincular los términos de semillas de SENSUS.

Proceso 3: Añadir rutas de acceso a la raíz.

Proceso 4: Añadir nuevos términos de dominio.

Proceso 5: Agregar subárboles completos.

El enfoque adoptado en SENSUS compartibilidad promueve el conocimiento como la ontología misma base se utiliza para desarrollar ontologías en dominios específicos. Una gran ventaja de este enfoque es que si dos ontologías se han desarrollado de forma independiente, la amplia cobertura de la ontología (Sensus) actúa como una "bisagra" que une la terminología y organización de ambas.

SENSUS también se ha aplicado para la construcción de una ontología de planificación militar de campaña aérea. Contiene un resumen de los elementos básicos que caracterizan la campaña aérea de planes como la campaña, el escenario, los participantes, comandantes, etc. También incluye las ontologías de las armas, los sistemas de combustible en general, etc. (Corcho, y otros, 2004).

1.4.6 On-To-Knowledge.

El objetivo de On-To-Knowledge es la aplicación de ontologías a la información electrónica disponible para mejorar la calidad de la gestión del conocimiento en las organizaciones grandes y distribuidas. En este

Capítulo 1: Fundamentación Teórica.

proyecto, desarrollaron una metodología y herramientas para el acceso inteligente a grandes volúmenes de fuentes de información semiestructurada y basado en Internet. Incluye una metodología para la construcción de ontologías para ser utilizada por la aplicación de gestión del conocimiento. Por lo tanto, la metodología On-To-Knowledge propone la construcción de la ontología, teniendo en cuenta la forma en que la ontología se utiliza en otras aplicaciones. En consecuencia, las ontologías desarrolladas con esta metodología son muy dependientes de la solicitud.

La metodología también incluye la identificación de metas a ser alcanzadas por las herramientas de gestión del conocimiento, y se basa en un análisis de escenarios de uso.

Se considera On-To-Knowledge como una metodología, debido a que cuenta con un conjunto de técnicas, métodos, principios para cada uno de sus procesos, y debido a que indica la relación entre estos. Los procesos propuestos por esta metodología son (Corcho, y otros, 2004):

Proceso 1: Estudio de viabilidad.

Proceso 2: Alcance, dominio y objetivos.

Proceso 3: Perfeccionamiento.

Proceso 4: Evaluación.

Proceso 5: Mantenimiento.

On-To-Knowledge propone una serie de pasos donde tiene como base fundamental los prototipos y un ciclo de vida incremental y cíclico. Presenta una dependencia de los recursos disponibles para el proyecto por lo que las ontologías desarrolladas con esta metodología son muy dependientes de la solicitud lo que la hace menos utilizada que otras como Methontology.

1.4.7 Methontology.

Esta metodología fue desarrollada dentro del grupo de Ontología en la Universidad Politécnica de Madrid en 1999. Tiene sus raíces en las principales actividades identificadas por el proceso de desarrollo de software y en las metodologías de ingeniería del conocimiento. Esta metodología incluye: la identificación

Capítulo 1: Fundamentación Teórica.

del proceso de desarrollo de ontologías, un ciclo de vida basado en la evolución de prototipos, y técnicas para llevar a cabo cada actividad en la gestión, orientada al desarrollo, y actividades de apoyo. Considera que las actividades realizadas durante el desarrollo de una ontología puede implicar la realización de otras actividades en otras ontologías ya construidas o en construcción (Corcho, y otros, 2004).

Sus pasos principales son (Corcho, y otros, 2004):

Especificación: Consiste en delimitar los objetivos de su creación (compartir información entre personas o por agentes de software, permitir la reutilización del conocimiento de un dominio, hacer explícitas las suposiciones que se efectúan en un dominio, separar el conocimiento del dominio del conocimiento sobre su fabricación o forma de operación o simplemente analizar el conocimiento del dominio). Consiste además en decidir el dominio de actuación de la ontología (para no representar objetos poco relevantes en perjuicio de otros más importantes), quién la usará y para qué, las preguntas a las que deberá responder (ayuda al establecimiento de las dos anteriores) y quién se encargará de su mantenimiento (decidiendo si se limitará a introducir nuevas instancias, se permitirá la modificación de conceptos o atributos, entre otros).

Conceptualización: Consiste en crear un glosario de términos que pertenecen al dominio, definirlos y crear una taxonomía (estableciendo una clasificación o jerarquía entre los conceptos, sus niveles, las relaciones entre ellos, sus instancias, sus propiedades o atributos, e igualmente los axiomas o reglas).

Formalización: Proceso que consistente en convertir el modelo anterior en un modelo formal o semicomputable.

Implementación: Convierte el modelo formalizado en un modelo computable mediante un lenguaje para construcción de ontologías.

Mantenimiento: Labor que puede ir desde el borrado de instancias ya inútiles o la incorporación de nuevas instancias que se han ido produciendo con el tiempo, hasta las tareas de introducción de cambios en el contenido de la información, ya sea redefiniendo atributos, relaciones o incluso conceptos.

Capítulo 1: Fundamentación Teórica.

Methontology concibe el desarrollo de la ontología de modo incremental y un ciclo de vida basado en la evolución de prototipos igual que On-To-Knowledge, pero además posibilita la identificación de procesos, técnicas para llevar a cabo cada actividad en la gestión, orientada al desarrollo, y actividades de apoyo. Las tareas y procesos que son llevados a cabo dentro de esta metodología engloban tanto las tareas de gestión como de apoyo, además da la posibilidad de utilizar o no una ontología existente anteriormente. Cuenta con una gran gama de herramientas que le dan soporte como son ODE, WebODE, OntoEdit y Protégé³, además de una abundante documentación y ejemplos de ontologías desarrolladas bajo esta metodología. (Corcho, y otros, 2004).

La mayoría de las metodologías existentes están diseñadas a partir de los resultados obtenidos en proyectos particulares de desarrollo de ontologías. Cada una propone un conjunto de pasos que guían la creación de ontologías desde el inicio o a través de una ya creada. Por las características del negocio relacionado con la Clasificación de Mercancías se decide crear una nueva metodología, los procesos que allí ocurren son bien específicos y pueden sufrir algunos cambios, debido a que los métodos empleados para la clasificación y los textos de las mismas pueden variar de acuerdo a resoluciones establecidas por el estado.

Metodologías como The Cyc y Uschold and King's no se comprende una fase de conceptualización ni de diseño, sino que los conocimientos se codifiquen directamente en el lenguaje propio de la metodología. (Lozano Tello, 2002). Otros como Gruninger y Fox y On-To-Knowledge son muy dependientes de la solicitud, de los recursos disponibles para el proyecto para el cual es concebida la ontología. Kactus es una metodología condicionada por el desarrollo de aplicaciones, de forma que, cada vez que se construye una aplicación, se construye la ontología que representa los conocimientos necesarios para la aplicación. Sensus posibilita la construcción de una ontología a partir de una enorme, utilizando términos de esta y agregándole otros necesarios.

³ Protégé :Algunas de estas herramientas serán explicadas más adelante en este documento.

Capítulo 1: Fundamentación Teórica.

En general la mayoría de las metodologías mencionadas proponen una serie de pasos por los cuales regirse a la hora de llevar a cabo la ontología, la mayoría centran su atención en las actividades de desarrollo y no en las de evolución y evaluación. Describen muy abiertamente las actividades a desarrollar sin determinar el cómo hacerlas.

Otra de las metodologías estudiadas es Methontology la cual cuenta con una serie de características que la hacen superior a las otras estudiadas.

- No está inspirada en experiencias concretas, ha sido empleada para modelar ontologías de diversos dominios.
- Propone un ciclo de vida iterativo y basado en prototipos evolutivos.
- Propone actividades específicas a desarrollar.
- Se explican bien los pasos a seguir para el desarrollo de una ontología.
- Se proponen elementos para el cómo desarrollarla y no sólo el qué hacer.
- Varios software existentes para la implementación de ontologías le dan soporte a la misma.
- Cuenta con una amplia documentación que posibilita conocer mejor los procesos que en ella ocurren.
- Independiente de las aplicaciones en las que se vayan a utilizar.

Las metodologías estudiadas han servido de guía para el desarrollo de ontologías, las mismas permiten tener un punto de partida para la determinación de algunos pasos para elaborar la ontología sobre el dominio de la Clasificación de Mercancías. Tomando como base la metodología Methontology por las facilidades que brinda, se proponen un conjunto de pasos específicos para el dominio. Omitiendo algunos pasos y creando otros nuevos propios de la metodología que permitan modelar la ontología dependiendo de las características esenciales del proceso.

1.5 Lenguajes de Presentación.

Las ontologías para ser expresadas requieren de un lenguaje lógico y formal. En la Inteligencia Artificial se han desarrollado numerosos lenguajes para este fin, algunos basados en la lógica de predicados, como

Capítulo 1: Fundamentación Teórica.

Formato de Intercambio de Conocimiento (KIF por sus siglas en inglés) en 1992, para definir términos, metaconocimiento, conjuntos, entre otros y Cycl⁴ que ofrecen poderosas primitivas de modelado, y otros basados en marcos que tienen un mayor poder expresivo, pero menor poder de inferencia; e incluso existen lenguajes orientados al razonamiento. Todos estos lenguajes han servido para desarrollar otros lenguajes.

1.5.1 OIL (Ontology Inference Layer).

Derivado en parte de SHOE⁵, fue impulsado también por el proyecto de la Unión Europea On-To-Knowledge. Utiliza la sintaxis del lenguaje XML y está definido como una extensión de RDFS⁶. Se basa tanto en la lógica descriptiva (declaración de axiomas) como en los sistemas basados en marcos. OIL posee varias capas de sub-lenguajes, entre ellas destaca la capa base que es RDFS, a la que cada una de las subsiguientes añade alguna funcionalidad y mayor complejidad. La principal carencia de este lenguaje es la falta de expresividad para declarar axiomas (Duque Mendez, y otros, 2006).

1.5.2 DAML + OIL (DARPA's Agent Markup Language + OIL).

Este lenguaje nació fruto de la cooperación entre OIL y DARPA y unifica los lenguajes DAML y OIL. Se basa ya en estándares del W3C⁷. El lenguaje DAML se desarrolló como una extensión del lenguaje XML y de Resource Description Framework (RDF) y para extender el nivel de expresividad de RDFS. DAML-OIL hereda muchas de las características de OIL, pero se aleja del modelo basado en clases (marcos) y potencia la lógica descriptiva. Es más potente que RDFS para expresar ontologías. Funciona como formato de intercambio (Corcho, y otros, 2004).

⁴ Cycl: Lenguaje desarrollado para la creación de la ontología Cyc.

⁵ SHOE: Primer lenguaje de etiquetado para diseñar ontologías en la Web.

⁶ RDFS: Extensión semántica de RDF, lenguaje primitivo de ontologías.

⁷ World Wide Web Consortium

Capítulo 1: Fundamentación Teórica.

Este lenguaje representa un avance para los existentes hasta el momento debido a que fusiona las características de OIL y DAML formando un lenguaje más potente, pero tiene algunas carencias como es la complejidad conceptual y de uso, las cuales fueron corregidas con el desarrollo del lenguaje OWL.

1.5.3 OWL (*Web Ontology Language*).

Es un lenguaje de etiquetado semántico para publicar y compartir ontologías en la Web. Se trata de una recomendación del W3C y puede usarse para representar ontologías de forma explícita, permite definir el significado de términos en vocabularios y las relaciones entre aquellos términos. En realidad, OWL es una extensión del lenguaje RDF, aunque es un lenguaje con más poder expresivo que este. Se trata de un lenguaje diseñado para usarse cuando la información contenida en los documentos necesita ser procesada por programas o aplicaciones, en oposición a situaciones donde el contenido solamente necesita ser presentado a los seres humanos. Al igual que OIL, OWL se estructura en capas que difieren en la complejidad y puede ser adaptado a las necesidades de cada usuario, al nivel de expresividad que se precise y a los distintos tipos de aplicaciones existentes (motores de búsqueda, agentes, entre otros) (Gruber, 1993).

Existen 3 sublenguajes de OWL, los cuales van creciendo respecto al nivel de expresión (Duque Mendez, y otros, 2006):

- **OWL Lite:** Útil para la creación de jerarquías y restricciones simples, sólo permite valores de cardinalidad 0 y 1, pierde en expresividad.
- **OWL DL (Description Logic):** es el lenguaje más sencillo e indicado para los usuarios que requieren el máximo de expresividad (todos los cálculos acaban en un tiempo finito). Una clase puede ser a la vez subclase de muchas clases, no puede ser una instancia de otra clase.
- **OWL Full:** Máximo nivel de expresión y la libertad sintáctica de RDF. Permite expresiones de segundo orden, pero sin expresividad. Por ejemplo, una clase puede ser tratada simultáneamente como una colección de individuos y como un individuo por sí mismo. Es el más completo por lo

que se necesita mucho poder computacional para poder hacer inferencias es por eso que se dice que no tiene garantía computacional.

Dentro de los lenguajes de presentación de la información en utilizado en la presente tesis es OWL debido a que este fue concebido como la revisión del lenguaje DAML+OIL siendo mucho más potente que este y evitando la complejidad conceptual y de uso. Está diseñado para usarse cuando la información que se encuentra dentro de los documentos necesita ser procesada por programas o aplicaciones y no por seres humanos, permite definir clases, relaciones y atribuirle ciertas propiedades a dichas relaciones y clases. Además de su facilidad para declarar reglas y axiomas.

1.6 Herramientas para la construcción de ontologías.

El proceso de construir ontologías es complejo y largo, y aún más si los desarrolladores de la ontología tienen que aplicar directamente un lenguaje de ontologías, sin ningún tipo de herramienta de apoyo. Para facilitar esta tarea, a mediados de la década de 1990 fueron creados los primeros ambientes para el desarrollo de ontologías, proporcionando interfaces de usuarios que ayudaron a llevar a cabo algunas de las actividades principales del proceso de desarrollo de ontologías, como la conceptualización, ejecución, comprobación de coherencia y la documentación. En los últimos años, el número de herramientas para el desarrollo de ontologías ha aumentado considerablemente y se han diversificado (Corcho, y otros, 2004).

Para la elaboración de la ontología destinada a la Clasificación de Mercancías se llevó a cabo un profundo estudio sobre algunas de estas herramientas, posibilitando llegar a conclusiones de cuál es la más factible a utilizar, dentro de estas se encuentran:

1.6.1 KAON.

Es una infraestructura de gestión de ontología de código abierto orientada a aplicaciones de negocios. Incluye una completa suite de herramientas que permite la creación de la ontología y administrar con facilidad, proporciona un marco para la creación de aplicaciones basadas en ontologías. Un aspecto importante de KAON es su razonamiento escalable y eficiente. KAON consiste en una serie de módulos

Capítulo 1: Fundamentación Teórica.

diferentes que proporcionan una amplia gama de funcionalidades entorno a la creación, almacenamiento, recuperación, mantenimiento y aplicación de ontologías. Fue desarrollada en un esfuerzo conjunto, principalmente por miembros de Instituto de la Universidad de Karlsruhe (AIFB) y el Centro de Investigación en Tecnologías de la Información. (KAON, 2001).

1.6.2 OntoEdit.

OntoEdit es una herramienta de edición de ontologías que apoya el desarrollo y mantenimiento de las mismas utilizando medios gráficos en un entorno web. Permite la representación semántica de lenguajes conceptuales y estructuras mediante conceptos, jerarquías de conceptos, relaciones y axiomas.

La interfaz es abierta y permite a los usuarios ajustarlo a sus necesidades. Además tiene disponibles una gran variedad de funcionalidades fáciles de usar y adaptables a los diferentes escenarios.

Ofrece una herramienta para representar gráficamente ontologías y que además, pueda almacenarlas y posteriormente manipularlas en una base de datos relacional.

Otros editores de ontologías se centran en ofrecer interfaces gráficas ricas en funcionalidades para editar, integrar y traducir ontologías; pero como estas herramientas no poseen integración en una base de datos, las ontologías no pueden ser guardadas, gestionadas y manipuladas directamente por el editor como si fuera un Sistema Gestor de Base de Datos (Corcho, y otros, 2004).

Aunque OntoEdit está pensado para un entorno web, también puede aplicarse a una Intranet. De esta forma se preserva la seguridad de la información que contiene.

1.6.3 Protégé.

La primera herramienta Protégé fue creada con el principal objetivo de simplificar el proceso de adquisición de conocimiento para sistemas expertos. Para lograr este objetivo, se utilizan los conocimientos adquiridos en etapas anteriores del proceso para generar formularios personalizados para la adquisición de más conocimiento. Desde entonces, Protégé ha pasado por varias versiones y se ha

Capítulo 1: Fundamentación Teórica.

centrado en diferentes aspectos de la adquisición de conocimiento (bases de conocimiento, los métodos de resolución de problemas, ontologías, etc.).

Protégé se basa en los marcos y la lógica de primer orden. Es compatible con Open Knowledge Base Connectivity (OKBC), lo que significa que los componentes principales de modelado son las clases, los slots, las facetas e instancias. Las clases están organizadas en jerarquías de clases donde se permite la herencia múltiple y los slots también pueden ser organizados en jerarquías.

Protégé reconoce marcos, XML Schema, RDF Schema y OWL, que son lenguajes semánticos utilizados en la Web, en contraposición a la rigidez del HTML. (Protégé, 1995)

Dentro de las herramientas que dan soporte a la creación de ontologías la más utilizada y que más se ajusta a las necesidades de la investigación es Protégé debido a la transparencia de su trabajo con OWL. A pesar de tener su propio lenguaje interno para definir ontologías, cuenta con un entorno abierto y fácil de entender, que ha generado en torno suyo toda una comunidad que contribuye activamente a ampliar el entorno con todo tipo de contribuciones en forma de plugins y está haciendo de esta herramienta un entorno sumamente potente. Además es la herramienta más empleada para el trabajo con la metodología Methontology y permite integrarse con otras herramientas que permiten visualizar el contenido de las ontologías y realizar validaciones a estas.

1.7 Razonadores Semánticos.

Una de las herramientas más utilizadas para trabajar con las ontologías son los razonadores, que sirven para realizar inferencia, a través de los conceptos y en algunos casos las instancias, obteniendo nuevo conocimiento. Generalmente difieren en el lenguaje formal en el que se especifica el conocimiento, así como los lenguajes de consulta que puedan utilizar. En la actualidad, son varios los razonadores o sistemas deductivos basados en lógica descriptiva que permiten el razonamiento y la inferencia en las ontologías. Dentro de los principales razonadores se encuentran:

Capítulo 1: Fundamentación Teórica.

1.7.1 Cerebra.

Es un motor de inferencia desarrollado por Network Inference. Parecido al razonador de lógica descriptiva FaCT⁸, sólo que en este caso sí aporta soporte para la inferencia sobre instancias y tipos de datos. Soporta ontologías en los lenguajes RDF, DAML + OIL y OWL, y como lenguaje de consulta para razonar emplea XQuery. Tiene versiones para trabajar conjuntamente con los editores OilEd⁹ y Protégé.

Los razonadores basados en lógica descriptiva deben de poseer la suficiente expresividad para permitir que el conjunto de constructores que forman parte de los lenguajes de ontologías sean soportados, y poder permitir ambos tipos de inferencia, tanto de clasificación como de chequeo de instancias. (Zapater Samper, 2005).

1.7.2 RACER.

RACER (Renamed ABox and Concept Expression Reasoner): desarrollado por Ralf Möller y Volker Haarslev en 1999, pero que ha sido renovado periódicamente. Es un razonador diseñado para la Web Semántica. Permite la inferencia tanto en conceptos como en instancias, soporta ontologías escritas en RDF/RDFS/DAML/OWL posee un lenguaje de consulta sencillo para la inferencia de instancias. Puede ser utilizado por OilEd y Protégé para comprobar la consistencia de la ontología y para hacer consultas sobre el conocimiento. (Zapater Samper, 2005).

El sistema RACER es el razonador basado en lógica de descripciones más representativo. No sólo implementa mecanismos de razonamiento sobre la Tbox¹⁰, sino que además proporciona un lenguaje de consultas, el nRQL (new Racer Query Language) que permite realizar consultas conjuntivas. Además, también implementa, en su última versión, mecanismos de razonamiento sobre la Abox¹¹. Sin embargo, el

⁸ Razonador basado en lógica descriptiva.

⁹ Herramienta para el desarrollo de ontologías.

¹⁰ Contiene conocimiento intencional en la forma de una terminología.

¹¹ Contiene conocimiento extensional en la forma de una terminología.

Capítulo 1: Fundamentación Teórica.

sistema no ofrece persistencia y los mecanismos de razonamiento se implementan mediante complejas técnicas de reducción. Esto implica que ontologías con grandes cantidades de instancias no puedan ser cargadas. Además, cada vez que se inicia el sistema se debe cargar de nuevo la ontología y procesarla. Por otro lado, RACER es en la actualidad una herramienta comercial, y por tanto, el razonador Pellet está ganando poco a poco más popularidad, ya que es una herramienta de libre disposición. (Roldán García, 2007).

1.7.3 FaCT ++.

Es un razonador DL que soporta la lógica TBox y ha sido desarrollado bajo el proyecto europeo WonderWeb⁷. Está implementado en C++ y corresponde a una nueva versión del razonador DL FaCT, que fue originalmente implementado en Lisp¹². Para realizar inferencia utiliza los algoritmos tableaux, e implementa nuevas características y optimizaciones, que permite personalizar para adicionar nuevas tácticas de razonamiento y la capacidad de razonar con lógicas descriptivas más potentes y cercanas a la expresividad de OWL-DL. (Rodríguez A, y otros, 2010).

1.7.4 Pellet.

Es un razonador de OWL-DL basado en Java. Puede ser utilizado conjuntamente con bibliotecas del API de Jena¹³ o del OWL. Mediante su uso es posible validar, comprobar la consistencia de ontologías, clasificar la taxonomía y contestar a un subconjunto de consultas RDQL¹⁴ (conocido como consultas ABox en terminología del DL). Se trata de un razonador DL basado en los algoritmos tableaux¹⁵ desarrollados para DL expresiva. Soporta todas las construcciones del OWL-DL.

¹² Lenguaje de programación de alto nivel.

¹³ Marco de trabajo utilizado para el trabajo complementario de las ontologías.

¹⁴ Lenguaje utilizado para realizar consultas a la ontología.

¹⁵ Método automático de deducción por refutación.

Capítulo 1: Fundamentación Teórica.

Entre las características más relevantes de Pellet se destaca que es capaz de (Lama Penín, y otros, 2008):

- Analizar y reparar ontologías, incorporando un conjunto de heurísticas para detectar y corregir automáticamente las ontologías OWL para asegurar que son OWL DL.
- Razonar con tipos de datos definidos en el estándar XML-S.
- Validar “especies” para garantizar que los individuos se clasifican de modo correcto en la taxonomía de conceptos de la ontología.
- Razonamiento con múltiples ontologías.

Después de realizar un estudio sobre los razonadores utilizados para realizar inferencias y validar sintácticamente ontologías, se ha decidido utilizar en el presente trabajo de tesis Pellet debido a que permite realizar consultas A box y Tbox. Además soluciona problemas de RACER, el cual es muy completo pero no permite cargar ontologías con grandes cantidades de instancias. Cada vez que se inicia el sistema debemos cargar de nuevo la ontología y procesarla. Además Pellet puede integrarse con la herramienta de desarrollo Protégé, para hacer la validación de la ontología se utilizarán Pellet y FaCT ++.

1.8 Conclusiones Parciales.

En capítulo se presentaron los conceptos significativos para la correcta comprensión del trabajo, dentro de estas las ontologías, así como las metodologías, lenguajes y herramientas empleadas en la confección de las mismas. Se realizó un estudio donde se identificaron las distintas formas de representar el conocimiento, concluyendo que de estas la que permite realizar inferencias sobre un dominio determinado son las ontologías, siendo así la forma seleccionada para representar el conocimiento relacionado al proceso de Clasificación de Mercancías. La comparación de los métodos en el desarrollo de ontologías ayudó a identificar los pasos a seguir para el desarrollo de la ontología para el dominio de la Clasificación de Mercancías en la AGR. Se decide utilizar además OWL 2.0 como lenguaje de representación de la ontología, Protégé en su versión 4.1.0 como herramienta de desarrollo de ontologías y Pellet como razonador.

Capítulo 2: Descripción de Metodología Propuesta.

CAPÍTULO 2 : DESCRIPCIÓN DE METODOLOGÍA PROPUESTA.

2.1 Introducción al Capítulo.

En el presente capítulo se proponen los pasos para el desarrollo de la ontología, se establecen las fases, se definen los artefactos a desarrollar durante el desarrollo de la ontología para comprender mejor el ciclo de vida, se definen los encargados de realizar cada una de las actividades.

2.2 Metodología Seleccionada.

El desarrollo de las ontologías es un proceso evolutivo e iterativo a través del cual se van refinando y completando sus detalles. Ellas modelan la realidad del mundo y todos los conceptos relacionados deben estar claramente reflejados en su representación. Para su elaboración no existen estándares por los cuales se pueda regir. La mejor solución es aquella que cubra todas las necesidades de la aplicación teniendo en cuenta la definición más exacta del dominio en cuestión.

A través del estudio de diferentes metodologías para el desarrollo de ontologías, se tomaron las pautas que se consideraron importantes para el desarrollo de una nueva, específica para el dominio relacionado con la Clasificación de Mercancías. La misma cuenta con roles, fases y artefactos.

Los roles describen las funciones y responsabilidades de aquellas personas que intervienen en la confección de la ontología. Las fases explican qué se debe hacer y cómo llevar a cabo el desarrollo de la ontología, están compuestas por una serie de actividades asignadas a un rol. Los artefactos reflejan los resultados tangibles de cada una de estas actividades.

Capítulo 2: Descripción de Metodología Propuesta.

2.3 Proceso de Clasificación de Mercancías.

La metodología propuesta tiene como objetivo el desarrollo de una ontología destinada al proceso de Clasificación de Mercancías que tiene lugar en el departamento Despacho Comercial en la AGR.

La misma se implementa para ayudar a mejorar este proceso.

La clasificación comienza desde la entidad (declarante), que será aquella empresa que importan algún tipo de mercancías al país. El declarante cuenta con el SA propuesto por la Oficina Nacional de Estadísticas e Información (ONEI) que contiene los productos y sus unidades de medida, para a partir del mismo realizar una clasificación previa, la cual irá junto con la descripción del producto en una DM entregada a la AGR.

En el departamento Despacho Comercial se recibe la DM, donde es comparada la información que la misma contiene con la del SA contenido en la Aduana que además de contener los datos de la que propone la ONEI cuenta con las tarifas arancelarias destinadas a cada producto.

En este paso es donde interviene la ontología, haciendo que el proceso de obtener la clasificación de un producto se haga en un menor período de tiempo y con mejores resultados.

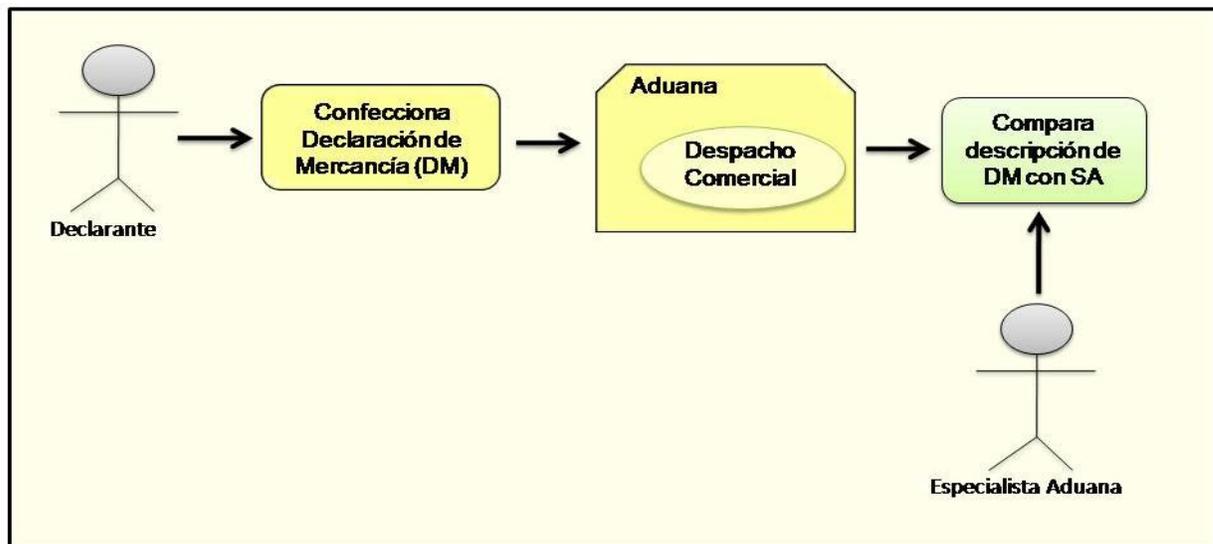


Fig. 1: Proceso para la Clasificación de Mercancías. (Fuente Propia).

Capítulo 2: Descripción de Metodología Propuesta.

2.4 Componentes de una ontología.

Clase (Concepto): descripción formal de una entidad del universo o dominio que se quiere representar. Constituye la pieza básica de estructuración del conocimiento. La decisión sobre qué considerar una clase del dominio no es fácil. Tal decisión debe ser tomada de acuerdo a los objetivos de la ontología (extraídos de las sesiones con los expertos, preguntas relevantes y de estándares existentes en el dominio). Una clase puede tener subclases que representan conceptos que son más específicos que dicha clase.

Clase abstracta: clase que no permite que existan instancias de ella. Se usa para agrupar conceptos, introducir cierto orden en la jerarquía, pero suelen ser demasiado generales para admitir instancias.

Instancia: representan objetos concretos del dominio, pertenecientes a una clase. La colección de instancias constituye la base de hechos (también denominada base de datos o base de conocimiento) del modelo.

Propiedad (Atributo, Slot): característica que permite describir más detalladamente la clase y sus instancias. Establece que la clase o concepto posee una propiedad que se concretará mediante un valor. Los valores de las propiedades o atributos pueden ser tipos básicos como cadenas de caracteres o números, pero también pueden ser otras clases o instancias.

Faceta (Restricción sobre las propiedades): es alguna propiedad de la propiedad. Por ejemplo, la cardinalidad, si la propiedad es obligatoria o no, etc.

Relación: interacción o enlace entre los conceptos o clases del dominio que se modelan. Algunas relaciones semánticas básicas son: subclase de, parte de, parte exhaustiva de, conectado a, es un. Suelen configurar la taxonomía del dominio. Las relaciones más simples se modelan mediante una propiedad de una clase cuyo valor es una instancia de otro concepto. Por ejemplo, dadas las clases AUTOR y OBRA, se puede definir la relación CREACIÓN como una propiedad (de cardinalidad múltiple) de la clase AUTOR cuyos valores sean instancias de la clase OBRA. Si las relaciones son más complejas

Capítulo 2: Descripción de Metodología Propuesta.

(deben tener a su vez propiedades o deben organizarse en jerarquías, por ejemplo) se suelen definir mediante clases (conceptos) que las representen.

Axioma: regla que se añade a la ontología y que permite describir el comportamiento de los conceptos o clases. Se establecen a partir de valores específicos de las propiedades. Por ejemplo: "Para todo A que cumpla la condición C, entonces A es B". Permiten dejar constancia de que ciertos valores de propiedades introducidos son coherentes con las restricciones de la ontología, o bien inferir posteriormente valores de atributos que no se han introducido explícitamente. De esta forma, a través de los axiomas es posible inferir conocimiento no codificado explícitamente en la ontología.

Anotación: es el proceso de relleno de instancias a partir de texto libre. Existen dos maneras de anotar texto: la más usual es la inclusión de etiquetas semánticas dentro del texto que se está procesando. Esto implica que el formato del texto debe ser editable y procesable. En el caso de no disponer de este formato, la segunda manera consiste en rellenar las instancias directamente en el modelo, dejando el texto original sin modificar.

Herencia: propiedad de la relación 'es_un' que permite que las clases relacionadas (heredadas) cuenten con los atributos de la clase con la cual se relacionan (clase padre).

Herencia múltiple: se representa cuando una clase dada hereda o cuenta con las propiedades de dos clases padre con las que establece dos relaciones del tipo 'es un'.

Derivación: organización de las clases de la ontología en un árbol de jerarquía mediante sucesivas relaciones 'es un' (también llamadas tipo de, es un, o herencias) con la propiedad de herencia. Esta organización permite el encadenamiento sucesivo de herencias desde las clases de nivel superior a las clases situadas en niveles inferiores, llamadas clases derivadas. (Comeche, 2007).

Capítulo 2: Descripción de Metodología Propuesta.

2.5 Roles en el desarrollo de la ontología.

Para el desarrollo de la ontología se definen los roles de acuerdo al dominio planteado, en la presente metodología se definen 3 roles.

Analista: Es el encargado de propiciar al desarrollador de la ontología los conocimientos sobre el negocio del problema, describe cuáles son los problemas existentes y las principales elementos a desarrollar en la ontología. Está presente en las dos primeras fases del desarrollo de la ontología, siendo el mismo el encargado además de determinar las clases, propiedades y las relaciones establecidas en el dominio.

Especialista de Aduana: Es el encargado de brindar la información necesaria para la confección de la ontología, es el cliente de la misma, con sus conocimientos y la información que el mismo posee se puede refinar la descripción del dominio y los términos seleccionados de forma tal que la ontología que se obtenga sea con la calidad requerida. Está presente en todo el desarrollo de la ontología, brinda tanto a los analistas como a los desarrolladores información necesaria para la correcta implementación de la ontología.

Desarrollador de la ontología: Es el encargado de la implementación y validación de la ontología. El mismo está presente en las dos últimas fases del desarrollo. Debe establecer cuáles son los axiomas a declarar, las instancias necesarias y validará la ontología a través de herramientas y en acuerdos con el especialista de Aduana.

2.6 Fases para el desarrollo de la ontología.

Tomando como base la metodología *Methontology* debido a que pone énfasis en la determinación clara y concisa de los requerimientos de la ontología, propone un ciclo de vida basado en prototipos evolutivos, la descripción detallada de las actividades a desarrollar y otras de las característica expuestas en el capítulo 1 de la presente investigación, se propone el siguiente proceso de desarrollo.

Capítulo 2: Descripción de Metodología Propuesta.

Se han definido cuatro fases. Las fases delimitan las diferentes actividades que se van a llevar cabo durante la metodología definida para la concepción de la ontología (Véase Fig. 2).

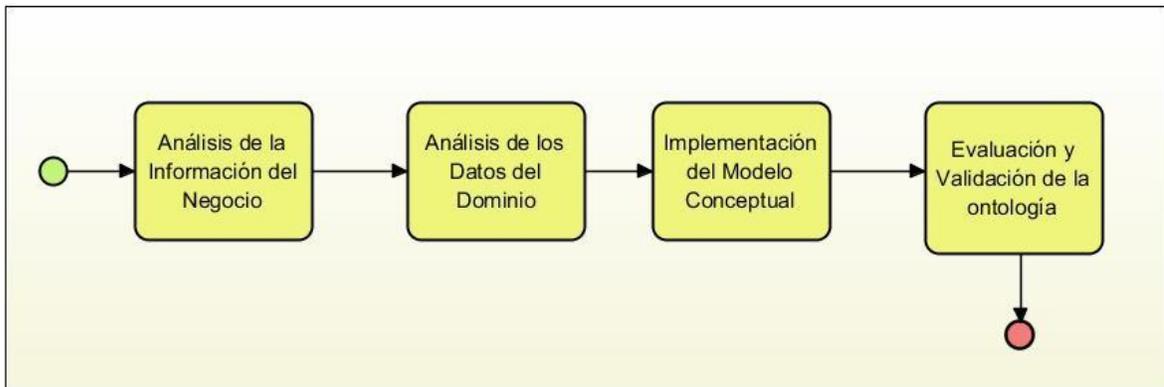


Fig. 2: Fases del Desarrollo de la ontología. (Fuente Propia).

Antes de comenzar el proceso de construcción de la ontología conviene destacar algunas reglas fundamentales según Noy & McGuinness (Peñalvo, 2006):

- No existe una única forma correcta de modelar un dominio; al contrario, siempre hay alternativas posibles. La mejor solución casi siempre depende del propósito u objetivo final de la ontología o de sus aplicaciones.
- El desarrollo de una ontología es necesariamente un proceso iterativo.
- Los conceptos o clases de una ontología deberían ser muy cercanos a objetos o entes (tanto físicos como lógicos) y a las relaciones existentes en el dominio de interés. Estos conceptos o clases serán probablemente nombres (objetos) o verbos (relaciones) que se encuentran en las frases que describen el dominio.

2.6.1 Análisis de la información del negocio.

Comprende las actividades iniciales de entendimiento del negocio y de la planeación de las diferentes acciones que se van a ejecutar durante el proceso de desarrollo (Véase Fig. 3).

Capítulo 2: Descripción de Metodología Propuesta.

Paso 1: Entendimiento del negocio de la Clasificación de Mercancía.

Se debe llevar a cabo un estudio detallado del negocio relacionado con la Clasificación de Mercancía, para poder conocer porqué es necesario el desarrollo de la ontología.

Paso 2: Planificar tareas a desarrollar y responsables.

Plantear las diferentes tareas que se van a llevar a cabo durante el desarrollo de la metodología, los pasos y los hitos más importantes en un documento donde se puedan consultar para poder llevar el control de las actividades (Cronograma de tareas), se limitan los responsables de cada una de estas así como el período en el que se van a llevar a cabo.

Paso 3: Definir partidas, capítulos y secciones a emplear.

Se debe definir bien cuál es el dominio dentro del negocio con el cual se va a trabajar, es decir sobre qué secciones, capítulos y partidas y de esta forma definir su alcance. Una correcta determinación del dominio de aplicación de un problema ayuda a conocer sus fronteras y mejora la comunicación entre todas las partes involucradas. Dentro de este se debe definir cuál es el objetivo de la ontología, para qué se usará y quien será el encargado de usar y mantener dicha ontología.

Las respuestas a las siguientes preguntas pueden ser de gran ayuda a la hora de definir su dominio y alcance:

¿Qué capítulos cubrirá la ontología?

¿Para qué se va a emplear la ontología?

¿Qué preguntas debería contestar la ontología?

¿Quién utilizará y mantendrá la ontología?

Paso 4: Elaborar preguntas de competencias.

Capítulo 2: Descripción de Metodología Propuesta.

Se deben elaborar las preguntas significativas o de competencia con las cuales se puede conocer si al concluir el desarrollo de la ontología esta cumple con los requisitos del dominio. Estas preguntas deben ser significativas dentro del dominio donde se trabaja debido a que al concluir el sistema debe ser capaz de contestar dichas preguntas. Es habitual que estas preguntas de competencia se extraigan, al menos de manera parcial de artículos, libros u opinión de expertos representativos del dominio que se desea modelar.

Ejemplo de esto es el SA con sus notas explicatorias, secciones y capítulos, la opinión de los especialistas de la Aduana relacionados con la Clasificación de Mercancías y artículos propuestos por la Oficina Nacional de Estadísticas e Información. Tanto la definición del dominio, el alcance y las preguntas de competencias que se realicen quedarán guardadas en un documento (Documento Visión).

En el Documento Visión se recogen los elementos más importantes durante el desarrollo de la metodología, el mismo permite conocer a grandes rasgos cuál es el dominio, las clases seleccionadas para la implementación, así como la forma de validación y los resultados obtenidos.

Paso 5: Se selecciona el lenguaje y el editor gráfico con el que se trabajará en la implementación

Se debe seleccionar cuál será el editor gráfico o herramienta de desarrollo empleada para la confección de la ontología y con este, el lenguaje a emplear, el mismo debe ser guardado en el (Documento Visión) para que otras personas que interactúen con la ontología tengan conocimiento sobre cuál herramienta y con qué lenguaje fue elaborada.

Capítulo 2: Descripción de Metodología Propuesta.

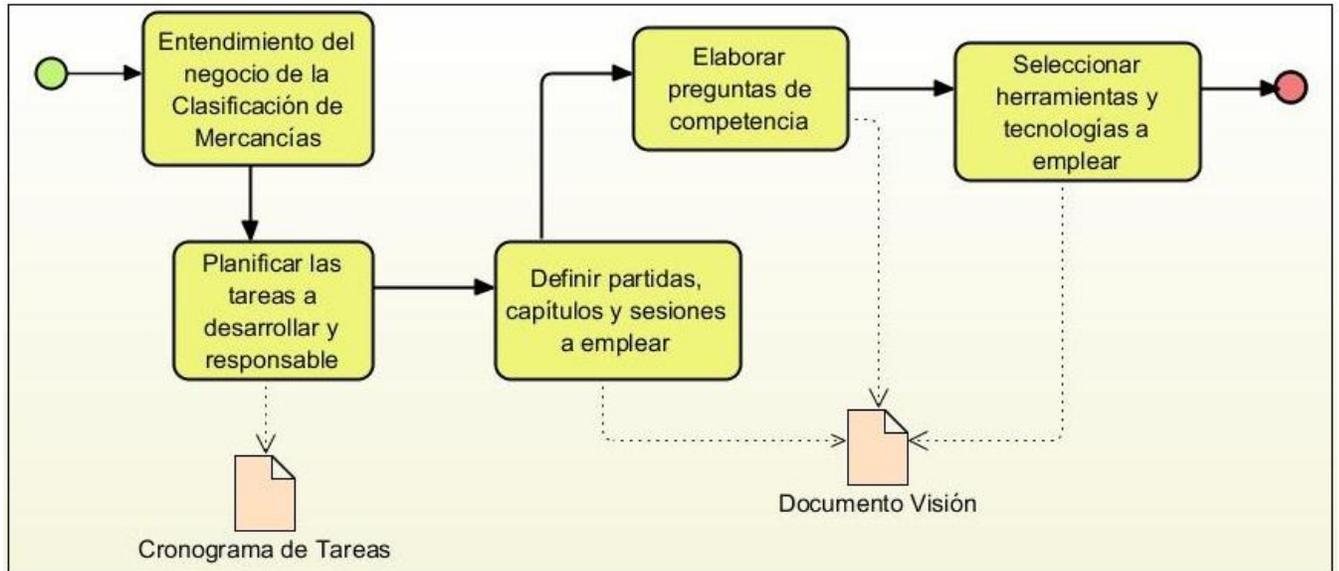


Fig. 3: Fase Análisis de la Información del Negocio. (Fuente Propia).

2.6.2 Análisis de los datos del dominio.

Comprende las actividades de análisis de los datos del dominio, haciendo un enfoque en los principales conceptos y la obtención de un modelo conceptual robusto, claro para la futura implementación (Véase Fig. 5).

Paso 1: Definir términos del capítulo seleccionado.

Deben ser analizados detalladamente los datos de la sección y del capítulo que se va a emplear, con el objetivo de conocer bien a fondo cuáles pueden ser términos importantes en el desarrollo de la ontología. Deben ser cercanos a los objetos y relaciones del dominio, los mismos pueden ser sustantivos (productos) y verbos (relaciones), se debe definir si dichos términos son persistentes o no. La persistencia indica si se registrará la información concerniente a dicho término en el modelo final y los términos que persisten son finalmente modelados mediante concepto del negocio, deben ser los términos más importantes. (Clase o entidad) relación de clases o atributo.

Capítulo 2: Descripción de Metodología Propuesta.

Se definen los términos relacionados con el dominio detallando el nombre de cada concepto y su persistencia (Listado de Términos del Dominio). La lista debe estar escrita en un lenguaje que sea entendible por los clientes.

Paso 2: Definir productos, partidas y subpartidas que pertenecen al capítulo.

Tomando como base dichos términos seleccionados se definen los productos, partidas, subpartidas, propiedades y las relaciones con una breve descripción de cada uno ellos y su clasificación (Ver Glosario de Términos).

Paso 3: Definir jerarquía entre los productos, partidas y subpartidas.

Se establece la jerarquía entre las clases que han sido seleccionadas para crear la ontología. A través de la misma se describen los conceptos y la interacción entre ellos. Este proceso se puede realizar de tres formas, la primera es definiendo desde los conceptos más generales hasta los más específicos dentro del dominio (top-down). La segunda consiste en el proceso inverso, o sea, ir desde lo más específico hasta lo más general (bottom-up) y la tercera vía es la combinación de los dos procesos anteriores, primero se detallan los conceptos más relevantes y luego se generalizan y especifican apropiadamente. Se definen las jerarquías (Diagrama de clases).

Paso 4: Definir propiedades de los productos.

Definir las propiedades o slot de los productos (Listado de clases y propiedades). En estos momentos se tienen reflejadas en el modelo todas las clases y sus relaciones, ahora se deben **definir las propiedades o SLOTS** de las mismas. Los SLOTS son los que describen las cualidades internas de los conceptos y representan las propiedades distintivas de los objetos.

Por ejemplo: una ontología que tenga definido como dominio los animales tendrá como propiedades la edad, el área en donde habitan, tipo de comida de la que se alimentan, con qué otros animales se relacionan, número de patas, etc. (Véase Fig. 4).

Capítulo 2: Descripción de Metodología Propuesta.

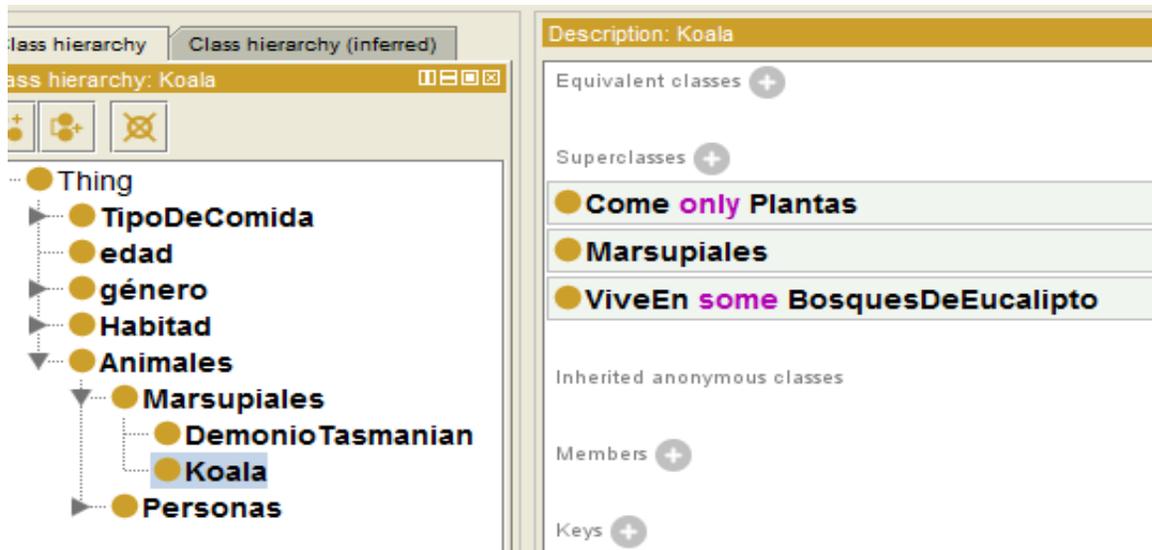


Fig. 4: Ejemplo de propiedades con la herramienta Protégé. (Fuente Propia)

Paso 5: Definir relación entre productos, partidas y subpartidas.

Definir cómo se relacionan los productos, partidas y subpartidas a través de propiedades (Diagrama de relaciones binarias) permitiendo saber el tipo de relación que se establecen entre estos y cuáles son las propiedades que intervienen para un mejor entendimiento del negocio.

Capítulo 2: Descripción de Metodología Propuesta.

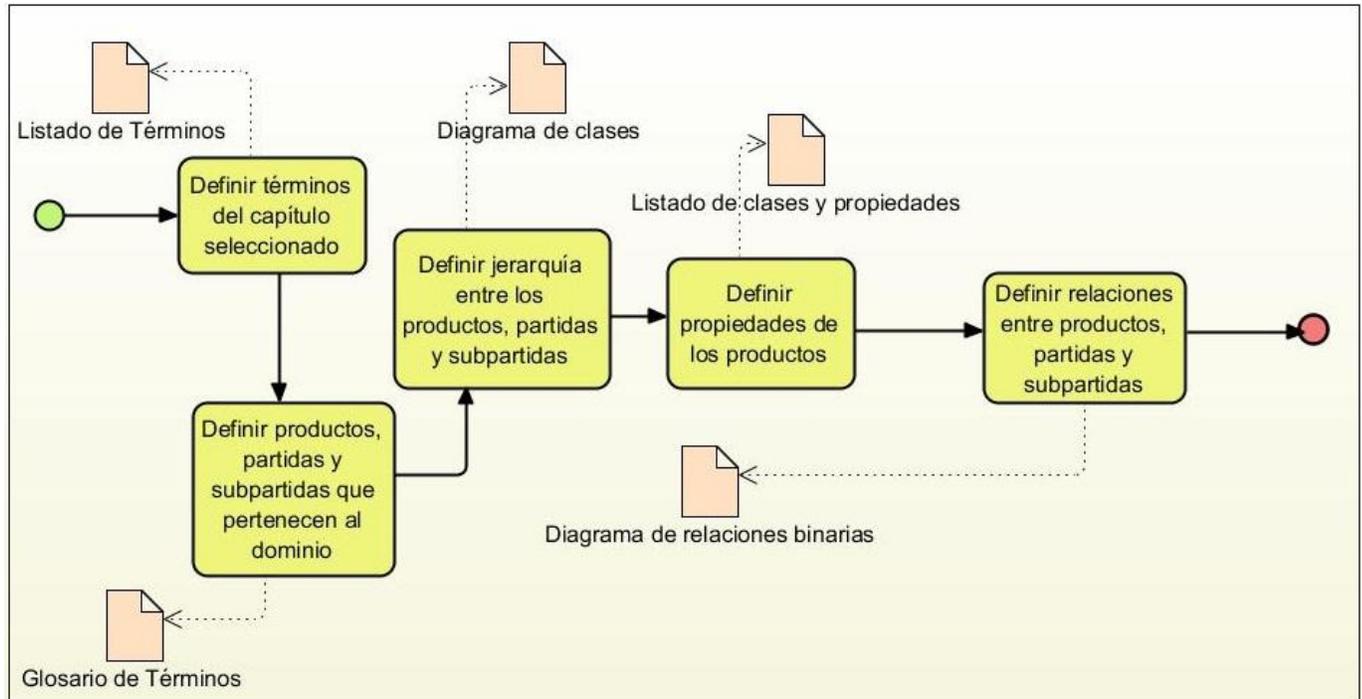


Fig. 5: Fase Análisis de los Datos del Dominio. (Fuente Propia).

2.6.3 Implementación del Modelo Conceptual.

Ya definido el lenguaje y editor gráfico a emplear, se definen las clases y relaciones entre ellas, atributos, axiomas y se declaran las instancias de las clases y consultas (Véase Fig. 6Fig. 6).

Mediante la confección de axiomas y reglas de inferencia se hace posible el razonamiento del futuro sistema. Es importante aclarar que el uso de estas no es necesario para el caso en el que sólo se necesite consultar los datos, a esto se le conoce como ontología de concepto.

Paso 1: Definir Axiomas Formales.

Se definen los axiomas formales (Listado de axiomas). La tabla de axiomas lógicos define las expresiones lógicas que, en la ontología, son siempre verdaderas. La definición de cada axioma incluye el nombre, la descripción de la regla en lenguaje natural, el concepto al que se refiere el axioma y la relación.

Capítulo 2: Descripción de Metodología Propuesta.

Paso 2: Crear Instancias.

Después de determinada las restricciones de los slots y los axiomas formales, se procede a crear las instancias tanto de los productos y subpartidas, las cuales, definiendo los valores de sus propiedades ayuda a inferir nuevos conocimientos en el dominio.

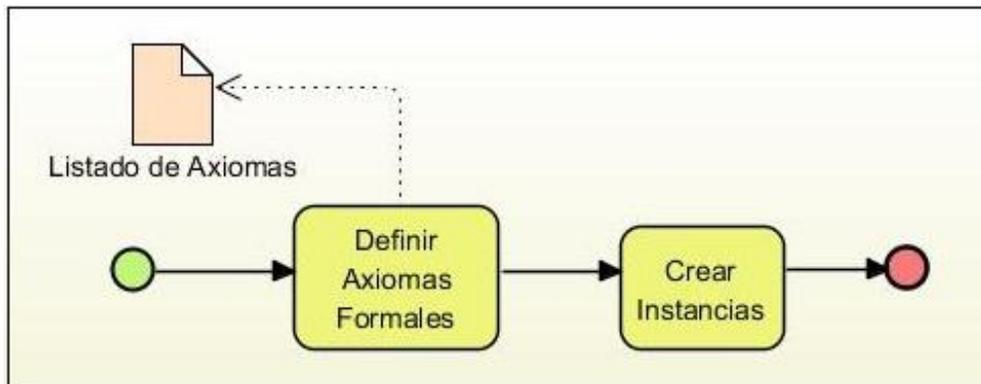


Fig. 6: Fase Implementación del Modelo Conceptual. (Fuente Propia).

2.6.4 Validación de la ontología.

En esta fase se pueden detectar errores de la fase anterior, en caso de detectarlos se tendría que retroceder para corregirlos. Sólo cuando se llegue al final de esta fase, se puede decir que ha concluido la fase anterior (Véase Fig. 7).

Paso 1: Pruebas de Consistencia.

Utilizando algunos sistemas de validación para OWL que existen online como con Validation Service ¹⁶y Pellet OWL Reasoner¹⁷ se evalúa si la ontología cumple con los requisitos de estándar OWL.

¹⁶ <http://www.w3.org/RDF/Validator/>

¹⁷ <http://www.mindswap.org/2003/pellet/demo.shtml>

Capítulo 2: Descripción de Metodología Propuesta.

Los mismos permiten que al ser introducido la URI de la ontología o el código de la misma saber si el mismo es correcto sintácticamente o no, permitiendo así conocer los valores de validación, consistencia y clasificación con los que cuenta la ontología.

Paso 2: Pruebas de Razonamiento.

Un Motor de Razonamiento en base a Lógica Descriptiva, asocia dos mecanismos internos en su entendimiento del conocimiento: ABox y TBox. El primero contiene aserciones sobre objetos, específico para las instancias del dominio de discurso. El segundo contiene el conocimiento normativo en forma de una terminología y es construido a través de declaraciones que describen propiedades generales de conceptos. (Zapater Samper, 2005).

Mediante el uso del razonador Pellet y FaCT++, Protégé permite validar algunos aspectos de la ontología, entre ellos: chequear la consistencia de la ontología, obtener automáticamente la clasificación taxonómica y computar los tipos inferidos.

Comprobación de la consistencia: permite constatar que no existen contradicciones en la ontología. La semántica de OWL define una especificación formal para la definición de la consistencia en una ontología empleando Pellet.

Validación de la taxonomía de clases: observa la relación entre cada clase y comprueba la jerarquía de clases completa.

Verificación de inferencia de clases: encuentra las clases más específicas a las que pertenece una instancia; en otras palabras, determina la clase a la que pertenece cada uno de los individuos.

Capítulo 2: Descripción de Metodología Propuesta.

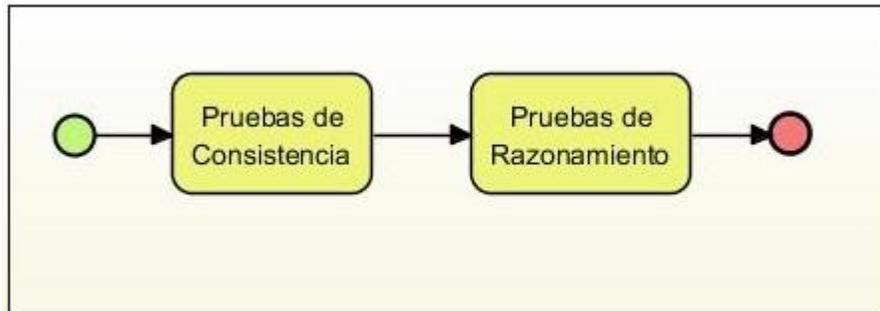


Fig. 7: Fase Validación de la ontología. (Fuente Propia).

2.7 Conclusiones Parciales.

En este capítulo se presentó una propuesta metodológica para el desarrollo de ontologías para el proceso de Clasificación de Mercancías. Se establecieron las fases, permitiendo así determinar cuáles son los pasos y procesos más importantes, se determinaron los principales artefactos que se deben seguir según el dominio para un mejor entendimiento del desarrollo de la ontología. Además se definió la forma de implementación a llevar a cabo y cómo será validada la ontología, probando de esta forma el correcto funcionamiento de la misma.

CAPÍTULO 3 : DESARROLLO DE LA ONTOLOGÍA.

3.1 Introducción al Capítulo.

En el capítulo se expondrán los elementos más importantes en el modelado de la ontología para probar que la metodología propuesta es factible para el desarrollo de la misma. Elementos tales como el dominio y alcance de la misma, las preguntas de competencia, así como los términos y conceptos relevantes en el negocio. También describen las entidades que toman lugar en el negocio y propiedades, así como las instancias que se crean de las mismas y la validación final de la ontología.

3.2 Análisis de la Información del Negocio.

3.2.1 Entendimiento del negocio de la Clasificación de Mercancía.

El Sistema Armonizado de Designación y Codificación de Mercancías (SA), constituye una nomenclatura internacional orientada a los aranceles de Aduana y a las estadísticas del Comercio Exterior. Consta de 21 secciones (que no participan en la codificación), 99 capítulos (dos dígitos), 1234 partidas (4 dígitos) y 5229 subpartidas (6 dígitos).

Todos los países signatarios del Convenio del Sistema Armonizado (que representan la mayoría de los países y mueven la inmensa mayoría del comercio mundial) están en la obligación de respetar sin ningún cambio, las Reglas Generales para la Interpretación del SA, su nomenclatura (secciones, capítulos, partidas y subpartidas) y las Notas explicativas de secciones y capítulos. (ONE, 2006).

Para la realización de la ontología se va a utilizar de la sección XVI destinada a “Las máquinas y aparatos, material eléctrico y sus partes; aparatos de grabación o reproducción de sonido, aparatos de grabación o reproducción de imágenes y sonido en televisión, y las partes y accesorios de estos aparatos”, específicamente el capítulo 85 a que trata “Las máquinas y aparatos, material eléctrico y sus partes”, el mismo contiene máquinas eléctricas con diferentes fines, como motores y generadores.

Capítulo 3: Aplicación de la metodología.

3.2.2 Planificar las tareas a desarrollar y responsables.

Para realizar la planificación se desglosaron las tareas a llevar a cabo en el desarrollo de la ontología y dependiendo de la complejidad de las mismas se estimó un tiempo para ser llevadas a cabo, principalmente por el desarrollador de la ontología y en el período de evaluación con la vinculación de alguno de los clientes (Véase Tabla 1).

Tabla 1: Cronograma de Tareas.

Fecha Inicio	Fecha Fin	Actividad	Responsable
10/10/2010	10/10/2010	Estudio a profundidad del negocio sobre la Clasificación de Mercancías.	Yusney Martínez Estrada
20/01/2011	22/01/2011	Planificar tareas a desarrollar y responsables.	Yusney Martínez Estrada
23/01/2011	10/02/2011	Definir partidas, capítulos y secciones a emplear.	Yusney Martínez Estrada
15/03/2011	18/03/2011	Definir preguntas de competencias.	Yusney Martínez Estrada
20/03/2011	24/03/2011	Selección de las herramientas y tecnologías a emplear.	Yusney Martínez Estrada

3.2.3 Establecer Dominio y Alcance.

¿Qué capítulos cubrirá la ontología?

La ontología Clasificacion_Mercancia permitirá modelar o representar la estructura que presenta el Sistema Armonizado de Designación y Codificación de Mercancías el cual está compuesto por sesiones y capítulos, específicamente del capítulo 85a referido a “Las máquinas y aparatos, material eléctricos y sus partes”. La ontología se utilizará como entrada al motor de inferencias para finalmente permitir la gestión semántica y apoyar el proceso de Clasificación de Mercancías en la AGR.

¿Para qué se va a emplear la ontología?

Capítulo 3: Aplicación de la metodología.

La ontología se empleará con el fin de facilitar la búsqueda de mercancías y su clasificación, además formalizar todo el conocimiento y descubrir nuevos a partir de los ya existentes. Esto permitirá encontrar inconsistencia y razonar sobre lo formalizado.

¿Quién utilizará y mantendrá la ontología?

Los usuarios potenciales de esta ontología son los especialistas de la Aduana encargados de llevar a cabo la Clasificación de Mercancías. Los que se deben hacer responsables de su mantenimiento deben ser expertos en el tema, así como los informáticos de la entidad capacitados en el trabajo con las ontologías, los cuales deben llevar a cabo la administración de la misma.

3.2.4 Preguntas de Competencia.

Las preguntas de competencia permiten conocer cuál es la clasificación de un producto determinado y la respuesta a las mismas debe estar expresada con términos dentro de la ontología. Lo anterior posibilita la evaluación de las necesidades y restricciones que debe cumplir la ontología, la especificación de la ontología mediante la traducción de todo el conocimiento capturado en lenguaje natural, a un lenguaje formal; por ejemplo OWL, la especificación de axiomas y definiciones de los términos en lenguaje formal, a través de la lógica de primer orden y el establecimiento de las condiciones óptimas de la ontología, con el fin de poder determinar cuando esté completa.

En el dominio de la Clasificación de Mercancías, las siguientes cuestiones son relevantes:

- ¿En que subpartida se localiza un motor de corriente continua con potencia inferior a 37 W?
- ¿Cuáles son los motores de corriente continua?
- ¿A qué sección pertenece al capítulo 85a?

Capítulo 3: Aplicación de la metodología.

3.2.5 Selección del Lenguaje y Editor Gráfico.

OWL 2.0 debido a que está diseñado para usarse cuando la información contenida en los documentos necesita ser procesada por programas o aplicaciones, permite definir clases mediante condiciones sobre sus miembros, mediante combinación booleana de clases o por enumeración de las instancias que pertenecen a la clase. Además OWL permite atribuir ciertas propiedades a las relaciones, como cardinalidad, simetría, transitividad, o relaciones inversas.

Protégé 4.1 para construir la ontología debido a la transparencia de su trabajo con OWL a pesar de tener su propio lenguaje interno para definir ontologías, cuenta con un entorno abierto y fácil de entender.

3.3 Análisis de los Datos del Dominio.

3.3.1 Definir términos del capítulo seleccionado.

Los términos de la ontología al ser definidos se convierten en conceptos, los cuales se definen con una palabra cuyo significado identifica un concepto específico. Para definir los términos que conformarían a Clasificacion_Mercancia se hicieron investigaciones sobre el dominio, en este caso los productos y demás términos que componen el SA. Existen diferentes subpartidas y partidas para clasificar los productos comercializables, para definir las mismas se necesitó ayuda de algunos expertos en el tema de la Clasificación de Mercancía provenientes de la Aduana.

Esta investigación desarrollada arrojó una serie de términos que pueden ser utilizados en una primera versión de cuáles pueden ser las posibles clases necesarias para la obtención de la ontología (Véase Tabla 2).

Tabla 2: Listado de Términos.

No	Término	Persistencia
1	Motor Corriente Continua.	S
2	Motor Corriente Alterna.	S

Capítulo 3: Aplicación de la metodología.

3	Generador Corriente Continua.	S
4	Tipo de energía.	N
5	Generador Corriente Alterna.	S
6	Grupo Electrónico de Émbolo Encendido por Chispa.	S
7	Grupo Electrónico de Émbolo Encendido por Compresión.	S
8	Grupo Electrónico de Energía Eólica.	S
9	Transformadores.	S
10	Transformador Dieléctrico Líquido.	S
11	Balastro.	S
12	Bobina Reactancia.	S
13	Convertidor Estático.	S

A partir de la primera versión del listado de términos, con la ayuda de varios especialistas de los que trabajan directamente con la Clasificación de Mercancías se identificaron los principales conceptos que se utilizarían en la ontología.

3.3.2 Definir productos, partidas y subpartidas que pertenecen al capítulo.

Para cada concepto que se decidió que fuera persistente dentro de estos, principalmente los productos, partidas u subpartidas, se realizó una búsqueda de su significado utilizando las definiciones que brinda el Sistema Armonizado de Designación y Codificación de Mercancías, quedaron definidos algunos de los sinónimos y que poseen a dichos conceptos y el tipo al que pertenecen (clases, propiedades, etc.) (Véase Tabla 3).

Tabla 3: Glosario de Términos.

No	Término	Descripción	Tipo
1	Capítulo	Recoge las partidas y subpartidas en las que puede estar un producto determinado.	Entidad

Capítulo 3: Aplicación de la metodología.

2	Sección	Agrupar a los capítulos por las características de los productos que los componen. El conjunto de secciones compone el Sistema Armonizado.	Entidad
3	Partidas	Clasificación en la que puede estar un producto en el Sistema Armonizado.	Entidad
4	Subpartidas	Clasificación en la que puede estar un producto en el Sistema Armonizado.	Entidad
5	Producto	Comprende a los productos que pueden estar contenidos en un capítulo	Entidad
6	Generador CA Inferior a 75 kVa	Generador de corriente alterna con potencia inferior o igual 75 kVa.	Entidad
7	Generador CA entre 75 kVa y 375 kVa	Generador de corriente alterna con potencia que puede estar entre 75 kVa o menos o igual a 375 kVa.	Entidad

3.3.3 Definir jerarquía entre los productos, partidas y subpartidas.

La ontología favorece el proceso de recuperación de la información al representar el conocimiento a través de conceptos y relacionándolos entre sí. Después de haber construido el glosario de términos se procede a la construcción del diagrama de clases, donde se expresan de forma jerárquica las clases del dominio. Para llevar a cabo dicho diagrama existen varios procesos (McGuinness, 2005):

- Un proceso de desarrollo *top-down* comienza con la definición de los conceptos más generales en el dominio, la subsecuente especialización de los conceptos.
- Un proceso de desarrollo *bottom-up* comienza con la definición de las clases más específicas, las hojas de la jerarquía, con el subsecuente agrupamiento de esas clases en conceptos más generales.

Capítulo 3: Aplicación de la metodología.

- Un proceso de desarrollo combinado es el resultado de una combinación de los enfoques *top-down* y *bottom-up*: primero definimos los conceptos más sobresalientes y luego los generalizamos y especializamos apropiadamente.

Para desarrollar el diagrama de clases (Véase Fig. 8) se utilizó el proceso *top-down*, se definieron las clases generales y luego las subclases. De esta forma se abarcan todos los términos declarados.

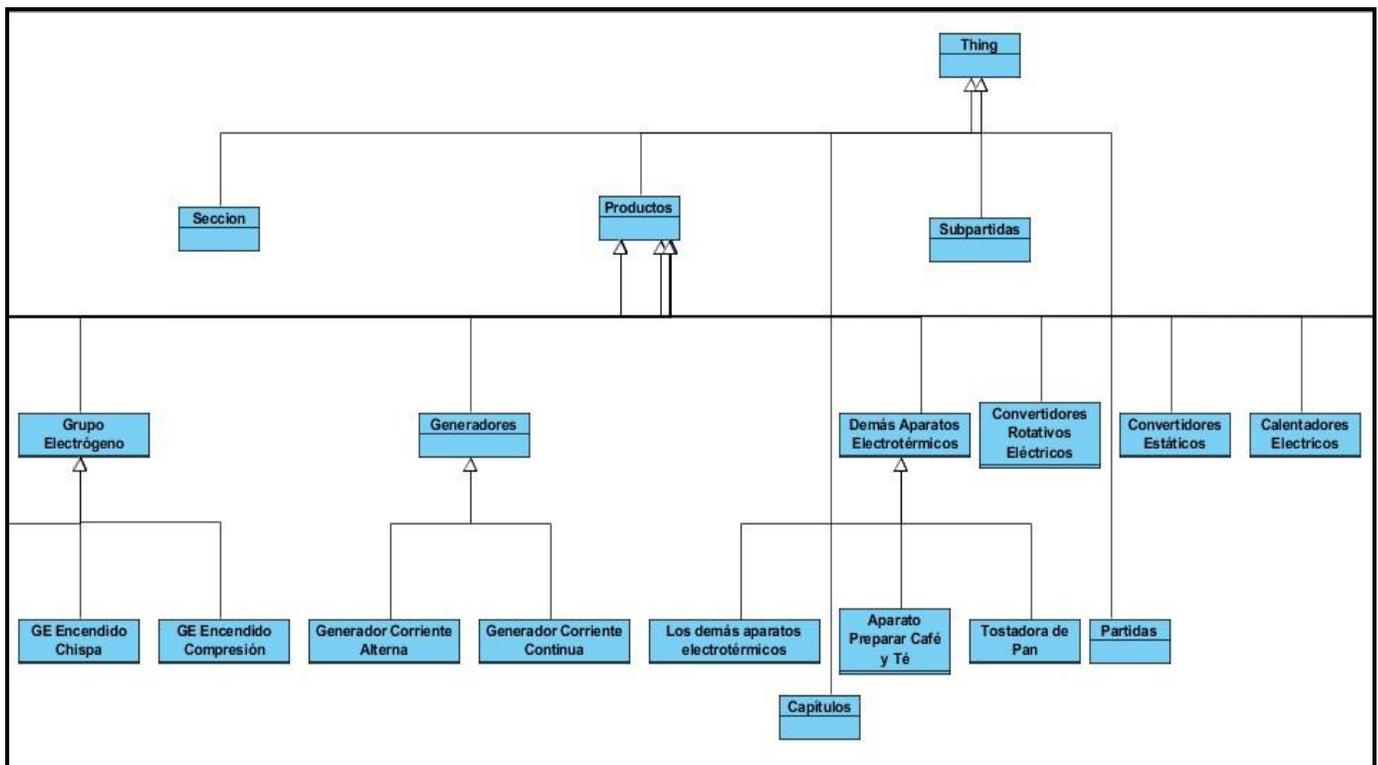


Fig. 8: Diagrama de Clases. (Fuente Propia).

3.3.4 Asociar propiedades a los productos.

Las clases aisladas no proveerán suficiente información para responder las preguntas de competencia. Una vez que se ha definido algunas de las clases, se deben describir las características de los conceptos, así como aquellos conceptos que relacionan a las clases de la ontología.

Capítulo 3: Aplicación de la metodología.

Una propiedad especifica la forma en que se va a relacionar un concepto con otro. Las mismas poseen diferentes características (Enciclo, 2010):

Simétrica: Si se tiene una propiedad simétrica y el par (X, Y) es una instancia de esa propiedad, entonces el par (Y, X) también es una instancia de esa propiedad. Ejemplo: Si tenemos una propiedad simétrica llamada *hermano* y dos instancias de esa propiedad (Pepe, Juan) se está especificando que Pepe es hermano de Juan, por lo tanto Juan es hermano de Pepe.

Inversa: Una propiedad puede ser la inversa de otra propiedad. Si se establece una *propiedad 1* como inversa de la *propiedad 2*, y se relaciona X con Y mediante la *propiedad 2*, entonces Y estaría relacionado con X mediante la *propiedad 1*. Si la propiedad *tieneHijo* es la propiedad opuesta de *tienePadres* y Juan *tieneHijo* Pepe, entonces Pepe *tienePadres* Juan.

Transitiva: Cuando una propiedad es transitiva, si el par (X, Y) es una instancia de la propiedad *transitiva 1*, y el par (Y, Z) es otra instancia de la propiedad *transitiva 1*, entonces el par (X, Z) también es una instancia de la propiedad. Por ejemplo, si se indica que la propiedad *Antepasado* es transitiva, si Juan es un antepasado de Pepe y Pepe es un antepasado de Alberto, entonces Juan es un antepasado de Alberto.

Funcional: Es posible definir propiedades para que tengan un valor único. Si una propiedad es funcional no tendrá más de un valor para cada individuo o ninguno. Esta característica se denomina propiedad única, es una forma de indicar que la cardinalidad mínima de la propiedad es 0 y la cardinalidad máxima es 1. Por ejemplo, *tieneMadre* puede establecerse como una propiedad funcional, cada individuo tiene una sola madre.

Funcional inversa: Si una propiedad es funcional inversa, entonces la inversa de la propiedad será funcional. Por tanto, la inversa de la propiedad tiene como máximo de un valor para cada individuo.

Cada relación es binaria, lo cual indica que cada relación puede ser vista como dos desde el punto de vista de cada uno de los conceptos involucrados. La relación entre Motores Corriente Continua y las

Capítulo 3: Aplicación de la metodología.

Subpartidas es a través de la propiedad “SeLocalizaEn”, entonces existe una que relacione a las Subpartidas con los Motores de Corriente Continua, por tanto existe la propiedad “TieneProductoA”. A este tipo de relación se le denomina relación inversa ya que el dominio de una es el rango de la otra y viceversa.

Estas propiedades explicadas son las propiedades de objeto (Object Properties), ellas relacionan conceptos de la ontología pero existen también las propiedades que no relacionan objetos, que su valor no es una instancia de otra clase si no un dato primitivo, dígame un entero o cadena etc. Estas propiedades son las propiedades de tipo de dato (DataType properties) por ejemplo la propiedad “Potencia” del “Motor” es un Integer o entero, no sería lógico crear un concepto de la ontología llamado “entero” o “Integer” es por esto que potencia no es una Object Property sino una Data Property.

Tabla 4: Ejemplo del Listado de Clases y Propiedades.

Clase:	Motor		
Relaciones:	Se localiza en.		
	Tiene Producto a.		
Atributos:	Tipo de Dato	Descripción	Valor que puede tomar.
Potencia	entero	Valor que toma la potencia en un motor determinado.	Valores enteros.

3.3.5 Definir relación entre productos, partidas y subpartidas.

Las relaciones binarias son propiedades que representan un tipo de asociación entre los conceptos del dominio.

Basándose en la definición de la taxonomía de conceptos y de la descripción de cada concepto por separado, se establecieron algunas propiedades que le permitirán a la ontología lograr un razonamiento para inferir el conocimiento almacenado en ella, creándose así el diagrama de relaciones binarias (Véase Fig. 9).

Capítulo 3: Aplicación de la metodología.

De esta manera se van tomando los datos de la descripción y a través de las relaciones la ontología puede inferir el tipo de producto a clasificar en una partida o subpartida.

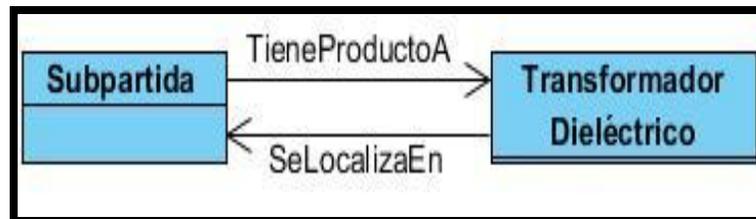


Fig. 9: Ejemplo de Diagrama de Relaciones Binarias. (Fuente Propia).

3.4 Implementación del Modelo Conceptual.

3.4.1 Definir Axiomas Formales.

Se le denominan axiomas formales a expresiones lógicas que son siempre verdaderas y utilizadas normalmente para especificar restricciones en la ontología. Estas restricciones se establecen entre conceptos y propiedades para lograr una inferencia en el conocimiento.

Para probar la metodología propuesta se declararon algunos axiomas formales (Véase Fig. 10) para especificar algunas de las restricciones que deben cumplir los productos para encontrarse en una determinada subpartida.

Capítulo 3: Aplicación de la metodología.

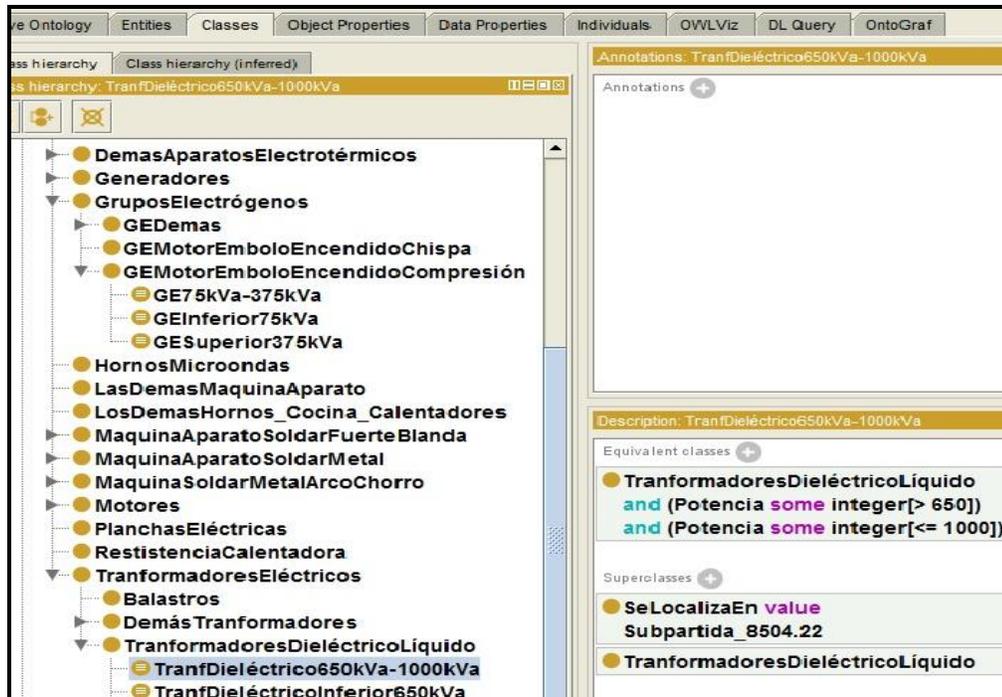


Fig. 10: Axiomas Formales. (Fuente Propia).

3.4.2 Crear Instancias.

Las instancias representan individuos en la ontología, los individuos son objetos de las clases. En la figura (Véase Fig. 11) se muestran algunos de los individuos definidos.

Capítulo 3: Aplicación de la metodología.

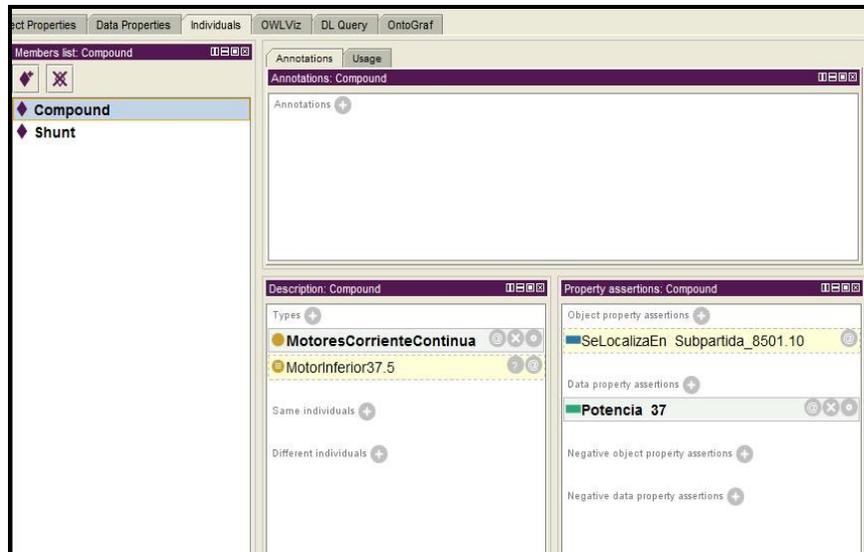


Fig. 11: Ejemplo de instancias en la ontología. (Fuente Propia).

En el ejemplo de la figura anterior se muestran como quedan creadas las instancias en Protégé y como se pueden relacionar con las Object properties y las Data properties, entre otras restricciones. Por ejemplo, la instancia **Compound** de la clase **MotoresCorrienteContinua** es un motor que se relaciona con el individuo **Subpartida_8501.10** de la clase **Subpartidas** a través de la Object property **SeLocalizaEn** y posee la Data property **Potencia**.

3.5 Validación de la ontología.

En la fase de evaluación se realizan pruebas consistencia, de razonamiento y de aceptación para verificar la consistencia de la ontología en cuanto a las relaciones establecidas entre los conceptos (clase equivalentes, propiedades, axiomas, etc.), la realización de inferencias entre clases e individuos y la aceptación por parte del cliente de la ontología. Una vez realizadas las pruebas se podrá verificar si la ontología soluciona el problema planteado, satisfaciendo las necesidades en los procesos de búsqueda y clasificación de los productos en la Clasificación de Mercancías.

Capítulo 3: Aplicación de la metodología.

3.5.1 Pruebas de Consistencia.

Para llevar a cabo las pruebas de consistencia se proponen 2 aplicaciones online capaces de medir la consistencia del lenguaje de representación OWL. Los mismos son Validation Service y Pellet OWL Reasoner.

The image shows two side-by-side screenshots of online validation tools. The left screenshot is from the W3C RDF Validation Service, displaying a 'Validation Results' page with a table of triples. The right screenshot is from the Pellet OWL Reasoner, showing 'Results' including input file details, DL Expressivity, and a classification hierarchy.

Number	Subject
1	http://www.semanticweb.org/ontologies/Ontology1336594304792.owl
2	http://www.semanticweb.org/ontologies/Ontology1336594304792.owl#ContieneA
3	http://www.semanticweb.org/ontologies/Ontology1336594304792.owl#ContieneA
4	http://www.semanticweb.org/ontologies/Ontology1336594304792.owl#PerteneceA
5	http://www.semanticweb.org/ontologies/Ontology1336594304792.owl#seLocalizaEn

Results
Input file: Text area
OWL Species: Full
DL Expressivity: ALCOIF(D)
Consistent: Yes
Time: 5076 ms (Loading: 2070 Species Validation: 661 Consistency: 159 Classification: 2185)

Classification:

- owl:Thing
 - ontology1336594304792:Capitulos
 - ontology1336594304792:Dinamo
 - ontology1336594304792:Seccion
 - ontology1336594304792:Alternadores
 - ontology1336594304792:Partidas
 - ontology1336594304792:SubPartidas

Fig. 12: Resultado obtenido con razonadores online. (Fuente Propia).

Las pruebas realizadas a través de analizadores sintácticos en la web (Véase Fig. 12) arrojaron resultados satisfactorios respecto a la estructura y consistencia de la ontología, mostrando los resultados de los mismos en la Fig. 12, en el Validation Service se muestra como resultado que el documento en RDF se ha validado correctamente y el Pellet OWL Reasoner prueba la consistencia de la ontología, así como el tiempo destinado a evaluar la misma.

3.5.2 Pruebas de Razonamiento.

Los razonadores semánticos se incluyen como plugins adicionales en Protégé para hacer inferencia sobre los datos representados en la ontología. Al aplicarle un razonador a una ontología los datos que se infieren se muestran sombreados en amarillo. De esta forma se comprueba la consistencia de la ontología en cuanto a los conceptos y las relaciones definidas entre ellos.

Capítulo 3: Aplicación de la metodología.

Para la realización de las pruebas de razonamiento a Clasificacion_Mercancia y comprobar el conocimiento que es capaz de inferir, se utilizan los razonadores semánticos *Pellet* y *FaCT ++*, ambos integrados al Protégé. Para comprobar el razonamiento que realiza Clasificacion_Mercancia al inferir los datos, se crearon 52 instancias de la clase **Subpartidas**, las mismas hacen referencia a los productos que se encuentran en el capítulo 85^a de la sección XVI del Sistema Armonizado e instancias de las clases MotoresCorrienteContinua y GeneradoresCorrienteAlterna. A estas instancias se le asociaron propiedades que le permiten a la ontología clasificarlas de acuerdo a sus descripciones. El conocimiento almacenado en Clasificacion_Mercancia será inferido a través de los razonadores *Pellet* y *FaCT++*.

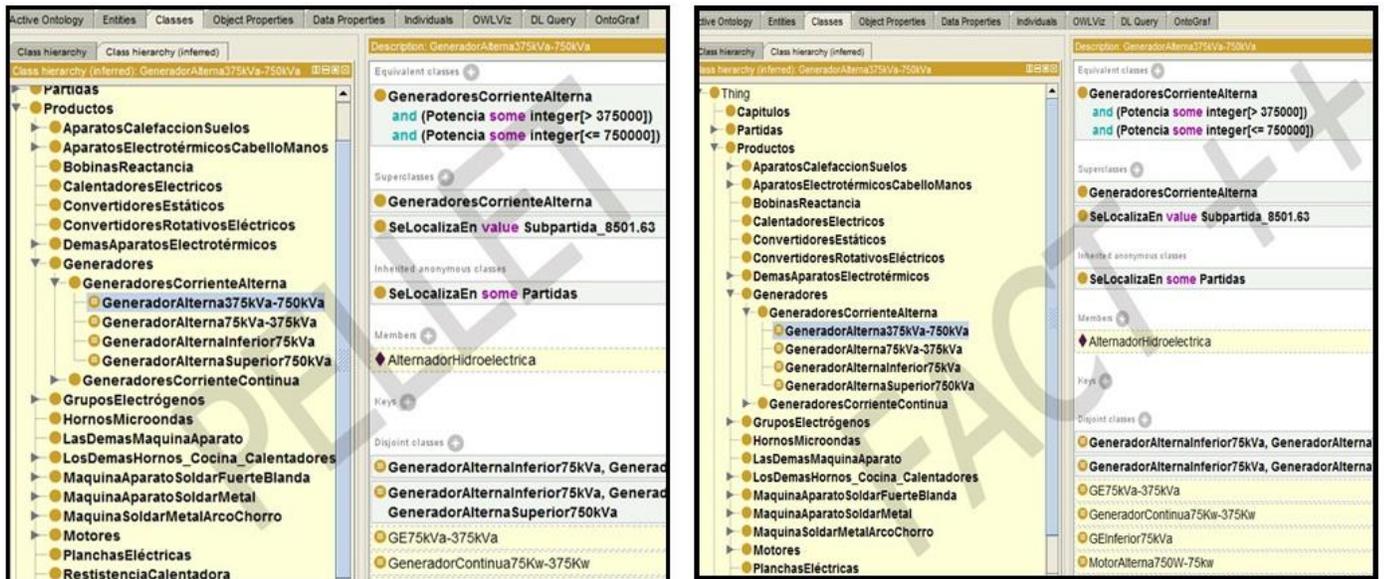


Fig. 13: Inferencia de datos con razonadores semánticos (1). (Fuente Propia).

En la Fig. 13 se muestra la primera prueba de razonamiento realizada a Clasificacion_Mercancia, al aplicársele ambos razonadores por separado se infirieron los mismos resultados. La comparación de los resultados obtenidos demuestra la consistencia de la ontología y la capacidad de razonamiento que posee la misma.

Capítulo 3: Aplicación de la metodología.

En el ejemplo, además se muestra la herencia de clases inferida y los datos que se obtienen sobre la clase seleccionada. Por tanto, se infiere que un **AlternadorHidroelectrico** es un Generador de Corriente Alterna que se encuentra entre 375kVa-750kVa de potencia, también que este es diferente a los demás generadores de corriente alterna y a los de corriente continua, de acuerdo a las características que se establecieron en su descripción.

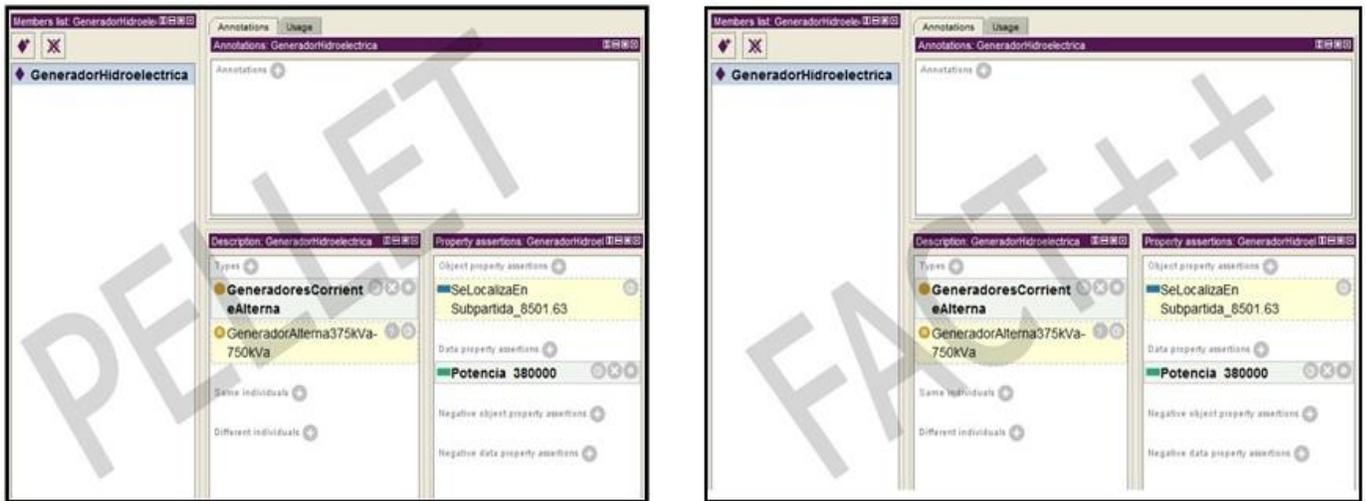


Fig. 14: Inferencia de datos con razonadores semánticos (2). (Fuente Propia).

En la Fig. 14 se muestra otro ejemplo de inferencia donde se obtienen los mismos resultados con los dos tipos de razonadores, en este caso se trata de la instancia **AlternadorHidroelectrica** a la cual se le establece en sus propiedades que tiene Potencia igual a 38 000. En otra propiedad también se relaciona la misma instancia con la clase **Subpartida**, de esta forma los razonadores infieren que la instancia se encuentra en la subpartida 8501.63 que es una instancia de la clase Subpartidas.

Capítulo 3: Aplicación de la metodología.

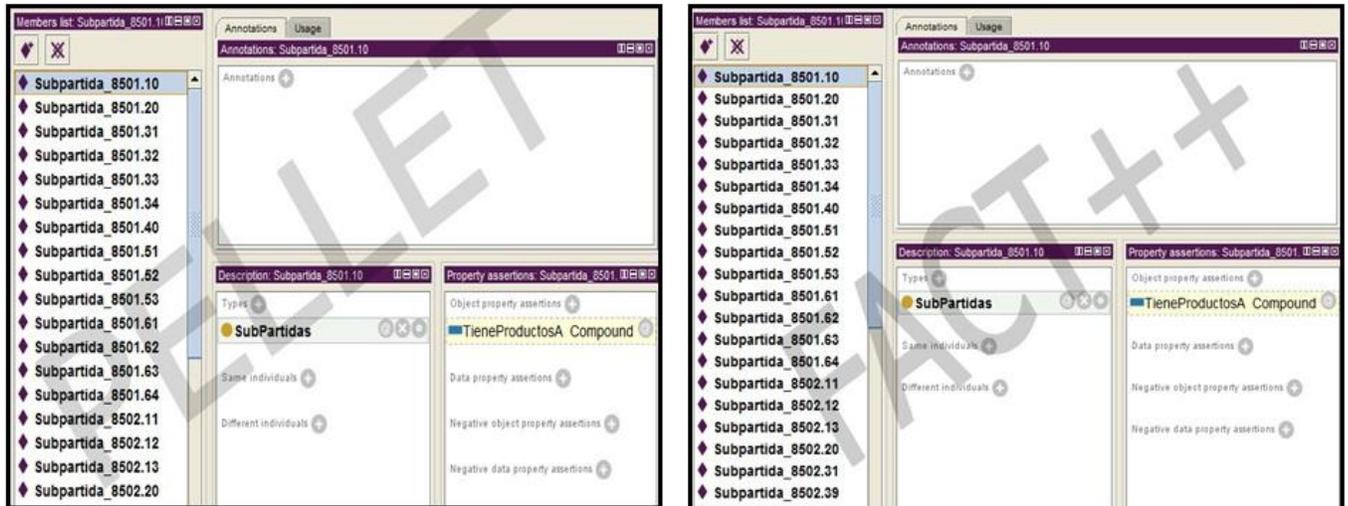


Fig. 15: Inferencia de datos con razonadores semánticos (3). (Fuente Propia).

En este último ejemplo de la figura Fig. 15 se muestran los resultados obtenidos al inferir conocimiento sobre la instancia **subpartida**, donde por ambos razonadores se obtiene que la instancia tiene como productos las instancias de la clase Motor Corriente Continua Inferior a 37 W. De esta forma el especialista obtiene una clasificación del producto perteneciente a esta subpartida.

Después de haberle realizado las pruebas de razonamiento a Clasificacion_Mercancia con los razonadores Pellet y FaCT ++, se llega a la conclusión de que la ontología posee consistencia en cuanto a la representación de los datos. La información inferida por ambos razonadores es prácticamente la misma. Por tanto, la ontología cumple con los objetivos planteados dando respuesta a las preguntas de competencias planteadas, asociando a los productos y su clasificación en subpartidas.

En la siguiente tabla comparativa (Véase Tabla 5) se muestra una comparación de los resultados obtenidos por ambos razonadores al realizar las pruebas de razonamiento.

Tabla 5: Comparación de resultados de los razonadores.

Aspectos	Razonadores Semánticos
----------	------------------------

Capítulo 3: Aplicación de la metodología.

	Pellet	FaCT ++
Clasificación de un producto.	x	x
Reconocimiento de clases equivalentes.	x	x
Reconocimiento de axiomas formales.	x	x
Clasificación de individuos en clases.	x	x
Inconsistencia en la ontología.	Ninguna	Ninguna

3.5.3 Análisis de Resultados.

Las pruebas realizadas demuestran que la ontología es capaz de brindarle al especialista de Aduana una serie de opciones, permitiendo comprobar si un producto determinado se encuentra en su respectiva clasificación, así como a partir de una subpartida conocer cuál o cuáles productos se asocian a ella. Con la realización de las mismas se comprobó la consistencia a través de razonadores sintácticos online, se validó la taxonomía de las clases y se verificó la inferencia de las clases e individuos mediante razonadores integrados al Protégé como plugins, obteniendo resultados satisfactorios.

3.6 Conclusiones Parciales.

En este capítulo se probó la metodología propuesta en el dominio específico para el que fue concebida, la misma se llevó a cabo completamente permitiendo verificar que la misma cuenta con las características necesarias para desarrollar ontologías para la Clasificación de Mercancías.

La ontología desarrollada fue probada y arrojó como resultado que posee una correcta consistencia y semántica. Permite inferir cuál es la clasificación de un producto determinado lo que contribuye a que el proceso de Clasificación de Mercancías ocurra de una forma más eficiente.

CONCLUSIONES.

Con la realización de esta investigación se ha arribado a la era de la gestión del conocimiento, espacio en donde la información recobra un nuevo significado, aproximándose cada vez más al pensamiento humano, contribuyendo al desarrollo de organizaciones como la AGR. La metodología propuesta está encaminada a gestionar el conocimiento en entornos donde la información esté estructurada en secciones y capítulos aprovechando para ello las facilidades que brindan las ontologías. En el presente trabajo se llegó a las siguientes conclusiones:

- La definición del marco teórico permitió identificar los elementos relevantes de la gestión del conocimiento, formas de representar el conocimiento; metodologías empleadas en la elaboración de ontologías.
- La metodología desarrollada se basa en la metodología Methontology, además tiene en cuenta las características propias del dominio de la Clasificación de Mercancías y permitirá clasificar un producto según sus propiedades.
- Siguiendo lo propuesto por la metodología desarrollada se diseñó una ontología que permite clasificar la mercancía según el Sistema Armonizado de Designación y Codificación de Mercancías, además se validó a través de razonadores integrados al Protégé como plugins, razonadores sintácticos online y la opinión de especialistas en el tema.
- Con la elaboración de una metodología menos compleja se hizo más sencillo la aplicación de las ontologías en el proceso de Clasificación de Mercancías.
- Algunas de las ventajas del uso de la ontología son favorecer el entendimiento de los especialistas de la AGR y el Sistema Armonizado de Designación y Codificación de Mercancías, así como el razonamiento automático de las clasificaciones.

RECOMENDACIONES.

Por las experiencias alcanzadas durante la realización de este trabajo y con vista a enriquecer la solución propuesta, el autor sugiere:

- Representar el conocimiento de los demás capítulos del Sistema Armonizado de Designación y Codificación de Mercancías.
- Estandarizar la metodología propuesta para extender la Gestión del Conocimiento a los demás departamentos de la AGR.
- Implementar un agente semántico que permita consultar la información disponible en la ontología Clasificacion_Mercancia.

REFERENCIAS BIBLIOGRÁFICAS.

Castells, P. 2004. *Neptuno: tecnologías de la web semántica para una hemeroteca digital.* España : Ministerio de Ciencia y Tecnología, 2004.

Comeche, Juan Antonio Martínez. 2007. *TUTORIAL ONTOLOGÍAS.* Madrid : Universidad Complutense de Madrid, 2007.

Corcho, Oscar, Fernandez Lopez, Mariano y Gomez Perez, Asuncion. 2002. *Methodologies, tools and languages for building ontologies. Where is their meeting point?* Madrid : Facultad de Informatica, Universidad Politecnica de Madrid, 2002.

Corcho, Oscar, Gómez Pérez, Asunción y Fernández López, Mariano. 2004. *Ontological Engineering with examples from the areas of Knowledge.* Berlín : s.n., 2004. págs. 107-337.

Duque Mendez, Nestor Dario, Chavarro Porras, Julio Cesar y Moreno Laverde, Ricardo. 2006. *Evolución de los lenguajes utilizados en la construcción de la Web Semántica.* Colombia : Scientia Et Technica, 2006. Vol. XII. ISSN.

Enciclo, Grupo. 2010. Aula Digital. *Relaciones Binarias.* [En línea] 2010. [Citado el: 11 de febrero de 2012.]

Figuerola, Liliana y Palavecino, Rosa. 2006. Conocimiento Abierto. Sociedad Libre. [En línea] 20 de noviembre de 2006. [Citado el: 01 de octubre de 2010.] <http://www.cibersociedad.net/congres2006/gts/comunicacio.php?id=618..>

Gruber, Thomas R. 1993. *Toward Principles for the Design of Ontologies.* Italia : Stanford Knowledge Systems Laboratory, 1993.

Guarino, Nicola. 1998. *Formal Ontology and Information Systems.* Italia : s.n., 1998.

Gutiérrez, J. J. 2005. *XQuery.* Departamento de Lenguajes y Sistemas Informáticos. Sevilla : s.n., 2005.

KAON. 2001. KAON tool Suite. [En línea] 2001. [Citado el: 25 de septiembre de 2010.] <http://kaon.semanticweb.org/>.

2003La gestión del conocimiento.<http://www.uoc.edu/dt/20251/index.html#>

Lama Penín, Manuel y Sánchez Vila, Eduardo M. 2008. *Análisis de técnicas de aprendizaje adaptativo con ontologías.* 2008.

López, Francisco Javier Honrubia. 2002. Introducción a las Ontologías. 2002.

Lozano Tello, Adolfo. 2002. *Métrica de Idoneidad de Ontologías.* s.l. : Escuela Politécnica de Cáceres, 2002.

Noy, Natalya F. y McGuinness, Deborah L. 2005. *Desarrollo de Ontologías-101: Guía Para Crear Tu Primera.* s.l. : Stanford University, Stanford, 2005.

Olleta, Javier Echegoyen. 2002. Origen de la Filosofía, Presocráticos, Sofistas y Sócrates. [En línea] 2002. [Citado el: 23 de febrero de 2012.] <http://www.e-torredebabel.com/Historia-de-la-filosofia/Filosofiagriega/Presocraticos/Ontologia.htm>.

ONE. 2006. ONE República de Cuba. [En línea] 2006. [Citado el: 25 de febrero de 2012.] http://www.one.cu/ryc_saclap.htm.

Ortiz, Antonio Moreno. 2008. Esquemas de redes semánticas. [En línea] 2008. ISSN: 1139-8736.

Osorio Galindo, Mauricio y Zepeda Cortés, Claudia. 2009. Revista Iberoamericana de Inteligencia Artificial. [En línea] 2009. [Citado el: 11 de febrero de 2012.] <http://erevista.aepia.org/>. ISSN 1137-3601.

Peñalvo, Francisco José García. 2006. *Web Semántica y Ontologías.* Departamento de Informática y Automática. s.l. : Facultad de Ciencias, 2006.

—. 2006. *Web Semántica y Ontologías.* 2006.

Protégé. 1995. Protégé. [En línea] Stanford Center for Biomedical Informatics Research, 1995. [Citado el: 11 de octubre de 2010.] <http://protege.stanford.edu/doc/dev.html>..

Rodríguez A, Claudia Milena, Cano Montaña, William y Mauricio Martínez, José. 2010. *Razonadores semánticos: Un estado del arte.* Colombia : Revista de la Facultad de Ingeniería, 2010. págs. 26-36.

Rojas Mesa, Lic. Yuniét. 2006. De la gestión de información a la gestión del conocimiento. [En línea] 16 de febrero de 2006. [Citado el: 5 de septiembre de 2010.] http://bvs.sld.cu/revistas/aci/vol14_1_06/aci02106.htm..

Roldán García, Maria del Mar and Aldan Montes, Jose. 2007. *DBOWL: Persistencia y Escalabilidad de Consultas y Razonamientos en la Web Semántica.* España : Universidad de Málaga, 2007.

Salazar, Alejandro Andrés Pavez. 2000. *Modelo de implantación de Gestión del Conocimiento y Tecnología de la Información para la Generación de Ventajas Competitivas .* Departamento de Informática, Alejandro Andrés Pavez Salazar. Valparaíso : s.n., 2000.

Schmitz Barcelo, Mario, Sanchez, Gerardo y Soltero, Alonso Perez. 2006. *Design of an Ontology as a Support to the Knowledge Audit Process in Organisations.* México : s.n., 2006.

Socorro, Raisa. 2009. *Las ontologías en la representación del conocimiento.* Centro de Estudios de Ingeniería de Sistemas (CEIS), Instituto Superior Politécnico. Ciudad Habana : s.n., 2009.

Studer, R., Benjamins, V. R. y Fensel, D. 1998. *Knowledge engineering: principles and methods.* 1998. págs. 161-167.

Tramullas, Jesús. 2003. *Agentes y ontologías para el tratamiento de la información: clasificación y recuperación en Internet.* s.l. : Universidad de Zaragoza, 2003.

Zapater Samper, José Javier. 2005. *Ontologías para Servicios Web Semánticos de Información de Tráfico: Descripción y Herramientas de Explotación.* Valencia : Servei de Publicacions, 2005. I.S.B.N.:84-370-6270-5.

GLOSARIO DE TÉRMINOS.

Aduana General de la República (AGR): Constituye un órgano de control en frontera y en la actividad interna vinculada al comercio exterior, que garantiza la seguridad y protección de la sociedad socialista y de la economía nacional, así como la recaudación fiscal y las estadísticas del comercio exterior, a través del cumplimiento de las políticas estatales de competencia aduanera para el tráfico internacional de viajeros, mercancías y medios de transporte.

Clasificación de Mercancías: Clasificación en el Sistema Armonizado de Designación y Codificación de Mercancías o Sistema Armonizado de un producto en una determinada partida o subpartida.

Declaración de Mercancías (DM): es aquella declaración que se realiza del modo prescrito por la Aduana, por la cual las personas interesadas indican qué régimen aduanero pretenden aplicar a las mercancías y suministran los detalles informativos que la Aduana requiere para la aplicación del régimen elegido.

Gestión del Conocimiento: Es la función que planifica, coordina y controla los flujos de conocimiento que se producen en la empresa en relación con sus actividades y su entorno con el fin de crear unas competencias esenciales.

Metodología: Es un proceso de software detallado que define con precisión los artefactos, roles y actividades involucradas.

Partida o código arancelario: Es un número que clasifica las mercancías.

Sistema Armonizado de Designación y Codificación de Mercancías o Sistema Armonizado (SA): nomenclatura que comprende las partidas, subpartidas y los códigos numéricos correspondientes, las notas de las secciones, de los capítulos y de las subpartidas, así como las Reglas Generales para la interpretación del sistema Armonizado. Se estableció a través del Convenio internacional del sistema Armonizado de Designación y Codificación de Mercancías, que entró en vigencia el 1 de enero de 1998, en sus idiomas oficiales: inglés y francés.