

# **Universidad de las Ciencias Informáticas**



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

**Título: Sistema Basado en Conocimientos para  
apoyar la toma de decisiones de la Subdirección  
de Investigación y Postgrado del CDAE**

**Autor: Jorge Mederos López Del Castillo**

**Tutores: MSc. Maypher Román Duran**

**MSc. Yoisy Pérez Olmos**

Junio del 2010. Año 52 de la Revolución.

# DECLARACIÓN DE AUTORÍA

Declaro ser autor de la presente tesis y reconocemos 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 \_\_\_\_\_.

**<Nombre autor>**

\_\_\_\_\_

Firma del Autor

**<nombre tutor>**

\_\_\_\_\_

Firma del Tutor

## **DEDICATORIA**

*Dedico este trabajo a las personas que más quiero y admiro en el mundo “mis padres”, Jorge Mederos Victoria y Mercedes López del Castillo, a los cuales les debo lo que soy y lo que seré en el futuro. Primeramente por haberme dado el privilegio de la vida, en la que me han entregado, su amor, dedicación y esmero. Por haber influido directamente en mi educación y formación, con sacrificio y apoyo incondicional, por celebrar mis victorias y sufrir mis derrotas. Por esto y tantas razones les doy gracias, pues mi mayor anhelo es estar a la altura del ejemplo que constituyen para mí.*

## **RESUMEN**

En muchas ocasiones resolver un problema conduce al proceso de identificar y seleccionar la acción adecuada para su solución. A este proceso se le denomina toma de decisiones, donde una decisión es la elección de una entre diversas alternativas. En la actualidad la toma de decisiones objetivamente fundamentada, se ha convertido en una tarea fundamental en la dirección de diferentes procesos, con énfasis en el empresarial.

Los Sistemas Basados en el Conocimiento son una de las técnicas de Inteligencia Artificial más empleadas en la actualidad en el proceso de toma decisiones. El presente trabajo surge como una necesidad de la Subdirección de Investigación y Postgrado del Centro de Consultoría y Desarrollo de Arquitecturas Empresariales de contar con un sistema, que además de la gestión clásica de la información de sus recursos, brinde a los directivos elementos para fundamentar sus decisiones estratégicas. En este sentido se implementa la propuesta de un sistema basado en conocimiento que permita extraer datos valiosos a través de un proceso de inferencias, proporcionando un conjunto de elementos favorables al proceso de toma de decisiones. El desarrollo de la solución incluye la modelación de la información del área de Investigación y Postgrado del centro utilizando ontologías como herramienta de representación del conocimiento, lo cual permite recuperar la información modelada.

### **Palabras clave:**

Toma de decisiones, Sistemas Basados en Conocimiento, Ontologías, Gestión de la información.

## ÍNDICE

DECLARACIÓN DE AUTORÍA .....	I
DEDICATORIA .....	I
RESUMEN.....	I
ÍNDICE .....	I
INTRODUCCIÓN.....	1
CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA .....	6
1.1. Gestión del Conocimiento .....	6
1.1.1 Etapas en la gestión del conocimiento. ....	7
1.1.2 Procesos Fundamentales. ....	7
1.2. Ontologías.....	8
1.2.1 Tipos de Ontologías.....	9
1.2.2 Componentes de una Ontología. ....	10
1.2.3 Ventajas del uso de Ontologías. ....	11
1.2.4 Herramientas para la Construcción. ....	13
1.2.5 Metodologías para construir Ontologías. ....	14
1.2.6 Lenguajes de Representación. ....	22
1.3. Sistemas Basados en Conocimientos.....	24
1.3.1 Definición de SBC.....	26
1.3.4 Ventajas del uso de SBC:.....	27
1.3.5 Desventajas del uso de SBC: .....	27
1.3.6 Categorías de aplicación de los SBC. ....	27
1.3.7 Tipos de SBC.....	28
1.3.8 Sistemas Basados en Reglas. ....	29
1.3.9 Ventajas de los SBR.....	31
1.4. Base de Conocimiento.....	32
1.5. Motor de Inferencia. ....	33

1.6. Lógica Descriptiva.....	34
1.6.1 Servicios de inferencia en Lógica Descriptiva.....	35
1.7. Conclusiones Parciales .....	39
CAPÍTULO 2: DISEÑO DE LA SOLUCIÓN.....	40
2.1. Análisis de la información del negocio.....	40
2.2. Análisis de datos.....	42
2.2.1 Preguntas de Competencia.....	42
2.2.2 Reutilización de Ontologías.....	43
2.2.3 Términos Relevantes.....	44
2.2.4 Especificación de la persistencia de los términos relevantes .....	46
2.2.5 Definición de restricciones de rol.....	49
2.2. Conclusiones Parciales .....	52
CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBA.....	53
CONCLUSIONES.....	68
REFERENCIAS BIBLIOGRÁFICAS.....	70

# INTRODUCCIÓN

La sociedad moderna se caracteriza por el cambio acelerado, permanente y continuo. Abrirse paso en la economía global constituye una realidad invariable para los países emergentes o en vías de desarrollo. En el caso específico de Cuba que no cuenta con grandes recursos naturales la fuente esencial de la ventaja competitiva está asociada al conocimiento. La formación de un gran caudal de capital humano, la investigación, el desarrollo y la innovación pasan a ser elementos de vital importancia en el mundo actual en el cual disponer de recursos humanos calificados es esencial, pero el gran desafío es tener la capacidad de administrar el conocimiento que se genera de la organización (1).

El ámbito de la representación y recuperación de la información ha tenido que asumir el impacto de Internet y sus tecnologías asociadas, en especial la Web. Estos cambios en la tecnología están conduciendo a una progresiva digitalización del campo de la representación y recuperación de información que afecta por igual a los recursos de información, las herramientas de representación, recuperación y los requerimientos de los usuarios. En este ambiente de creciente digitalización, son diversas las herramientas de representación y recuperación de la información que pueden ser objeto de estudio. De igual forma son variados los campos de conocimiento donde se aplican estas herramientas: la Lingüística, la Inteligencia Artificial, la Documentación... Así es que en la literatura especializada se analizan como herramientas para la representación y recuperación de información taxonomías: sistemas de clasificación, bases de datos léxicas, tesauros, bases de conocimiento, mapas conceptuales, ontologías, entre otros. Dentro de este amplio espectro de herramientas de representación y recuperación de información, las ontologías constituyen una de las que con mayor frecuencia son objeto de vinculación en la bibliografía. Todas estas herramientas colaboran en la descripción de los diferentes recursos de información y en su posterior recuperación, en términos de efectividad, rapidez y facilidad de acceso a la información (2).

Una ontología es un tipo de instrumento que permite la representación del conocimiento en un área determinada en clara conexión con su recuperación en entornos informáticos, es la descripción explícita de un dominio la cual define un vocabulario común (3). Algunas de las ventajas de su utilización son:

- Optimización de las aplicaciones de sistemas basados en conocimiento, su desarrollo e interoperabilidad.
- Preservación del conocimiento persistente de los expertos en cualquier campo de aplicación.

- La descripción de los conceptos y sus relaciones pueden ser empleadas de forma más general que los actuales tesauros. Estas permiten una gestión rápida, eficaz y permanentemente de la información depositada en Internet permitiendo su recuperación y divulgación.

Al ser las ontologías un nuevo tipo de documento, se debe estudiar su almacenamiento y recuperación para la satisfacción de las necesidades que se puedan plantear (4).

A raíz de la evolución de nuevos paradigmas arquitectónicos y tecnologías de desarrollo de software se crea en la Universidad de las Ciencias Informáticas (UCI), el Centro de Consultoría Tecnológica y desarrollo de Arquitecturas Empresariales (CDAE). Este Centro tiene como misión fundamental ofrecer los servicios de consultoría tecnológica a partir del capital intelectual que se desarrolla en la universidad, como una actividad generadora de recursos financieros para el país. Sus objetivos se basan en:

- Brindar servicios de consultoría en tecnologías informáticas a instituciones nacionales y extranjeras.
- Formar consultores altamente calificados en tecnologías informáticas comprometidos con los principios de la Revolución Cubana.
- Constituir un centro de referencia en la UCI sobre las tecnologías asociadas a las líneas de consultoría propias (Inteligencia Organizacional, SOA-BPM “Service Oriented Architecture - Business Process Management”, JEE “Java Enterprise Edition”).

El CDAE está organizado en líneas de trabajo que se constituyen para ejecutar proyectos de investigación, desarrollo o brindar servicios de consultoría. En las líneas debe existir una estrategia para el desarrollo profesional de sus miembros y cada una constituye en sí misma un observatorio del estado del arte de las tecnologías que trata, de modo que se garantice la actualidad de las temáticas que afronta.

Una de las áreas del CDAE es la de Investigación y Postgrado la cual se encarga de regir las investigaciones científicas que se realizan en el centro así como de organizar los cursos de superación posgraduada para los especialistas del CDAE esta área será el dominio del sistema.

Actualmente en esta área no existe una herramienta que apoye al proceso de toma de decisiones, teniendo en cuenta la retroalimentación con la información previa precedente que se tiene de otros



proyectos, de manera que los administradores de dicha entidad puedan realizar esta actividad de forma rápida e intuitiva, lo cual trae como resultados:

- La gestión es manual, hace un proceso tedioso e ineficiente.
- No se gestiona el conocimiento asociado a la información que se genera en el CDAE.
- No existe un mecanismo que apoye la toma de decisiones teniendo en cuenta la información precedente.

La situación existente en el CDAE actualmente permite definir el siguiente **Problema de investigación**: ¿cómo realizar un mecanismo de gestión de la información de los proyectos del CDAE de manera centralizada, que además apoye la toma de decisiones estratégica a partir de la gestión del conocimiento?

La **hipótesis** plantea que si se desarrolla un Sistema Basado en Conocimientos (SBC) que permita realizar inferencias a partir de una ontología que modele la información del Área de Investigación y postgrado del CDAE de la UCI, entonces se posibilitará una gestión del conocimiento que apoye la toma de decisiones estratégicas de los directivos del centro.

Un Sistema Basado en Conocimientos (SBC) puede ser definido como (30):

*Un sistema computarizado que usa conocimiento sobre un dominio para arribar a una solución de un problema de ese dominio. Esta solución es esencialmente la misma que la obtenida por una persona experimentada en el dominio del problema cuando se enfrenta al mismo problema.*

Por lo que se tiene como **Objeto de Estudio**:

- Las técnicas y herramientas para la representación y gestión del conocimiento.

Especificándose como **Campo de Acción**:

- Los Sistemas Basados en Conocimiento (SBC) para la gestión del conocimiento del Área de Investigación y postgrado del CDAE de la UCI.

El **Objetivo general** de la investigación se enmarca en desarrollar un SBC para la gestión del conocimiento del Área de Investigación y postgrado del CDAE de la UCI.

Para cumplir el objetivo general se proponen algunos **objetivos específicos**:

- Evaluar el estado del arte acerca de la gestión del conocimiento, las ontologías y los sistemas de inferencia.
- Desarrollar una Ontología para modelar la información del Área de Investigación y Postgrado del Centro de Consultoría y Desarrollo de Arquitecturas Empresariales de la UCI.
- Desarrollar un SBC que, a partir de un Modelo Ontológico de la información del Área de Investigación y postgrado del CDAE, favorezca la estrategia de toma de decisiones por los directivos del centro.

Las **Tareas de investigación** que se proponen son las siguientes:

- Estudio del estado del arte de las prácticas, tecnologías y metodologías asociadas al campo de investigación.
- Modelado de la información de los recursos del Área de Investigación y postgrado del Centro de Consultoría y desarrollo de Arquitecturas Empresariales de la UCI.
- Diseño de un motor de inferencia que, a partir del diseño ontológico del Área de Investigación y postgrado del Centro de Consultoría y desarrollo de Arquitecturas Empresariales de la UCI, permita apoyar el proceso de toma de decisiones.
- Diseño de la base de datos con la información del Área de Investigación y postgrado del Centro de Consultoría y desarrollo de Arquitecturas Empresariales de la UCI.
- Implementación del sistema basado en conocimientos.
- Realización de pruebas al sistema basado en conocimientos implementado.

Para el desarrollo de las tareas científicas se utilizan diferentes Métodos de investigación en la búsqueda y procesamiento de la información. Los métodos teóricos son factibles en el estudio de las características poco observables del objeto de investigación.

### **Dentro de este grupo se utiliza:**

El método Análisis Histórico-Lógico, permite desarrollar el estudio del arte previo al desarrollo de la investigación. Es utilizado para identificar si existen otras soluciones al fenómeno en cuestión, y para el estudio crítico de los trabajos anteriores.

El Analítico-Sintético, se utiliza para el estudio a partir de fuentes bibliográficas seguras de los modelos de organización de bases de casos en el proceso de representación y búsqueda de la información. Permite además descomponer el problema de investigación en elementos por separado y profundizar en el estudio de cada uno de ellos, para luego sintetizarlos en la solución de la propuesta.

La tesis está conformada por tres capítulos. En el primero se aborda la gestión del conocimiento en particular las Ontologías, los tipos que existen, sus componentes, las ventajas que trae consigo su uso, las diferentes metodologías para el desarrollo de ontologías y los distintos lenguajes de representación que existen. Se realiza una caracterización de los SBC, componentes fundamentales, su desarrollo y tipos existentes. Se analiza el impacto que trae consigo la utilización de estos sistemas en el proceso de toma de decisiones así como las ventajas que trae su uso frente al modelo del proceso tradicional. Finalmente se describe el Razonamiento Basado en Reglas como una alternativa para la elaboración de los SBC. El segundo capítulo aborda el análisis de los conceptos relevantes del negocio, la designación de su tipo en el modelo ontológico (Clase, Relación, Atributo) así como la definición de los axiomas principales del sistema para finalmente formalizar desarrollo de la Ontología que modela los procesos del Área de Investigación y Postgrado del CDAE. En el tercer capítulo se realizan un conjunto de pruebas para fundamentar la funcionalidad del sistema así como enriquecer algunos elementos de relevancia.

# **CAPÍTULO 1: FUNDAMENTACIÓN TEÓRICA**

En el presente capítulo se describen los elementos y términos más importantes asociados con el desarrollo de un Sistema Basado en Conocimientos. Se aborda además el proceso de Gestión del Conocimiento (GC) y las etapas que lo conforman. Por otra parte se conceptualiza el término de Ontologías, se describen los tipos existentes, sus componentes, ventajas, metodologías de desarrollo y los lenguajes usados en su representación; así como las herramientas que existen para su construcción. Finalmente se tratan los sistemas basados en conocimiento, sus generalidades, ventajas y desventajas, arquitectura y componentes fundamentales.

## **1.1. Gestión del Conocimiento**

El conocimiento está en la mente de las personas y es información combinada con experiencia, contexto, interpretación y reflexión, aplicado a decisiones y acciones. Las organizaciones lo consideran, hoy en día, como activo y riqueza de la nueva economía, dándole la mayor importancia a su creación, transferencia y uso eficiente. El conocimiento es uno de los activos claves en las empresas comerciales actuales, así como en la sociedad moderna, se ha identificado como un elemento clave de las organizaciones para lograr ventajas competitivas. Ante esta realidad, ha surgido un nuevo enfoque dentro de la gestión empresarial: la gestión del conocimiento (1).

Dentro de los conceptos de Gestión del Conocimiento más utilizados se encuentran: “el conocimiento es la capacidad de resolver un determinado conjunto de problemas, con una efectividad determinada”. Otros autores lo definen de la manera siguiente: “Es una mezcla de experiencia, valores, información y saber hacer, que sirve como marco para la incorporación de nuevas experiencias e información y es útil para la acción. Se origina y aplica en la mente de los conocedores” mientras que para otros estudiosos “El conocimiento es un conjunto integrado por información, reglas, experiencia, interpretaciones, relaciones y conexiones en un contexto y una organización, que constituyen la base para la acción y toma de decisiones” (1). De cualquier forma la esencia radica en que el conocimiento es complejo, pues una parte de este es instintiva y difícil de determinar formalmente ya que se establece como un conjunto de elementos mezclados fundamentalmente en las relaciones humanas con los sistemas que se hallan en su medio.

### 1.1.1 Etapas en la gestión del conocimiento.

**Crear conocimiento**, que implica exploración, combinación y el descubrimiento de conocimiento mediante el hacer. Los individuos al interior 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 (5).

**Compartir conocimiento**, que se produce cuando los individuos al interior de una organización transfieren y comparten el conocimiento. Al compartir este 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 (5).

**Aplicar conocimiento**, que 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 (5).

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 UCI. La cual tiene la responsabilidad de estar llamada a convertirse en una importante fuente de ingresos para nuestro país.

### 1.1.2 Procesos Fundamentales.

En la GC hay dos procesos fundamentales, *la creación y la transmisión de conocimiento*.

**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 acopiarlo, 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 (6).

Estos dos procesos que pueden pensarse por separado también están totalmente interrelacionados, pues la creación de conocimiento no es algo que se hace partiendo de la nada, sino que para crear conocimiento se utiliza conocimiento que proviene de otras personas y de otros lugares, por lo tanto, ha habido un proceso de transmisión previo. Son procesos que están muy interrelacionados y que juntos hacen que el conocimiento dentro de las organizaciones mejore y se utilice. Otro elemento importante que ayuda a hacer que el conocimiento funcione es el contexto, este va a permitir interpretar el conocimiento y transmitirlo (6).

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 organización, aportándole nuevos valores a la misma y por ende a la sociedad a la que brinda sus servicios en función de su razón de ser.

### **1.2. Ontologías**

Las Ontologías han sido tradicionalmente usadas como modelo de representación del conocimiento en Inteligencia Artificial. Según la definición ofrecida por Gruber una Ontología es “una especificación explícita y formal de una conceptualización” (9). Más concretamente, una Ontología estará formada por una taxonomía relacional de conceptos y por un conjunto de axiomas o reglas de inferencia mediante los cuales se podrá inferir nuevo conocimiento, por su parte Neches delimita la ontología como un instrumento que define los términos básicos y relaciones a partir del vocabulario de un área, así como las reglas de combinación de estos términos y relaciones para definir extensiones a un vocabulario (10). Sin embargo para Guerrero y Lozano, las ontologías son construcciones que estructuran contenidos explícitos y que son capaces de codificar las reglas implícitas de una parte de la realidad, pese a trabajar con declaraciones explícitas independientes del fin y del dominio de la aplicación (2).

Lo fundamental de todos estos conceptos es que las ontologías brindan la posibilidad de establecer un entendimiento común entre las personas y los agentes de software, posibilita analizar y reutilizar el conocimiento y explicitar las suposiciones de un dominio determinado. El entendimiento entre el agente

y los usuarios es de vital importancia para el desarrollo del presente trabajo de diploma ya que el agente debe ser capaz de generar conocimiento mediante inferencias entrada la información del centro.

Según Noy y McGuinness los objetivos de las Ontologías son (2):

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 en la extracción y recuperación de información, en páginas web, de contenidos conectados temáticamente.

- Permitir la reutilización del conocimiento perteneciente a un dominio.
- Permitir hacer explícitos los supuestos de un dominio.

Separar el conocimiento de un dominio del conocimiento que se puede denominar operacional. Con esto se alude a que, en ocasiones, el conocimiento que se está representando se puede implicar en diferentes áreas al pertenecer más a un conocimiento relacionado con procesos.

Hacer posible el análisis del conocimiento de un campo, por ejemplo en lo que se refiere al estudio de los términos y relaciones que lo configuraran ya sea formalmente o no.

*Las ontologías es la técnica que se ha considerado más óptima para el desarrollo de la base de conocimientos del sistema, debido a que permite además de la información explícita del dominio de aplicación modelar parte de la información implícita lo cual es necesario en el sistema que se desea implementar con el presente trabajo.*

### 1.2.1 Tipos de Ontologías.

Las Ontologías se pueden clasificar de acuerdo a la cantidad y tipo de estructura de la conceptualización en (11):

- *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 el almacenamiento estandarizado de información.

Según su dependencia y relación con una tarea específica desde un punto de vista, las Ontologías tienen la siguiente clasificación (11) :

- *Ontologías de Alto nivel o Genéricas*: Describen conceptos más generales.
- *Ontologías de Dominio*: Describen un vocabulario relacionado con un dominio genérico.
- *Ontologías de Tareas o Técnicas básicas*: Describen una tarea, actividad o artefacto, por ejemplo componentes, procesos o funciones.
- *Ontologías de Aplicación*: Describen conceptos que dependen tanto de un dominio específico como de una tarea específica, y generalmente son una especialización de ambas.

### 1.2.2 Componentes de una Ontología.

Generalizando, se puede afirmar que una Ontología no es más que otra forma de representación del conocimiento para ello consta de las siguientes partes (12):

- **Clases o Subclases** (o conceptos): Descripción formal de una entidad del universo o dominio que se quiere representar. 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.
- **Slots** (en ocasiones llamados roles o propiedades): 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.
- **Relaciones**: Interacción o enlace entre los conceptos o clases del dominio que se modela. 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.
- **Axiomas**: Reglas que se añaden 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.

- **Facetas** (también llamadas restricciones de roles). Estas facetas describen elementos como los tipos de valores, los valores permitidos, el número de valores y cualquier otra característica que un slot puede tomar. Entre los tipos más comunes de facetas están la cardinalidad (permiten definir cuántos valores pueden tener un slot), el tipo de valor del slot (string, número, booleanos, etc.), dominio y rango de un slot, etc.

### 1.2.3 Ventajas del uso de Ontologías.

Las Ontologías se usan para favorecer la comunicación entre personas, organizaciones y aplicaciones, lograr la interoperabilidad entre sistemas informáticos, razonar automáticamente y para la ingeniería de software. 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 (13).

En los últimos años el desarrollo de ontologías ha estado moviéndose del dominio de los laboratorios de Inteligencia Artificial a los escritorios de los expertos de un dominio dado. Las ontologías han llegado a ser comunes en el WWW. Muchas disciplinas desarrollan ahora ontologías estandarizadas que los expertos de ciertos dominios pueden usarlas para compartir y anotar información en sus campos de trabajo. Una ontología define un vocabulario común para investigadores que necesitan compartir información en un dominio. Ella contiene definiciones de conceptos básicos y sus relaciones que pueden ser interpretadas por una máquina (14).

Seguidamente se especifican algunos de los beneficios de su utilización:

- I. **Compartir el entendimiento común de la estructura de información entre personas y agentes de software** es una de las más importantes metas al desarrollar ontologías (15). Por ejemplo, se supone que distintos sitios Web contengan información médica o provean servicios de *comercio electrónico* médico. Si esos sitios Web comparten y publican la misma ontología subyacente de los términos que usan, entonces agentes de software podrían extraer y agregar información de esos sitios diferentes. Los agentes podrían usar esta información agregada para responder solicitudes de los usuarios o servir como datos de entrada a otras aplicaciones.
- II. **Permitir la reutilización de conocimiento de un dominio** fue una de las fuerzas conductoras detrás de recientes trabajos en la investigación sobre ontologías. Por ejemplo, modelos para diferentes dominios necesitan representar la noción de tiempo. Esta representación incluye las nociones de intervalo de tiempos, puntos en el tiempo, medidas relativas de tiempo, y cosas por el estilo. Si un grupo de investigadores desarrollo tal ontología en detalle, otros podrían simplemente rehusarla en sus dominios (15).
- III. **La explicitación de suposiciones de un dominio que subyacen bajo una implementación**, permite cambiar esas suposiciones fácilmente si el conocimiento del dominio cambia. Suposiciones codificadas explícitamente acerca del mundo en algún lenguaje de programación hacen que las suposiciones no solo sean difíciles de hallar sino también difíciles de cambiar, en particular para alguien sin competencias en programación. Además, las especificaciones explícitas del dominio de conocimiento son útiles para nuevos usuarios que deben aprender el significado de los términos del dominio (15).
- IV. **La separación del conocimiento del dominio del conocimiento operacional** es otro uso común de las ontologías. Se puede describir la tarea de configuración de un producto a partir de sus componentes de acuerdo a especificaciones requeridas e implementar un programa que hace independiente esta configuración de los productos y componentes en sí (2). Se puede desarrollar una ontología de componentes de PC y características y aplicar el algoritmo para configurar PCs ordenadas a medida. Se puede usar el mismo algoritmo para configurar elevadores si “alimentamos” la ontología con elevador como componente (15).
- V. **Analizar el conocimiento de un dominio**: es posible una vez que una especificación declarativa de los términos está disponible. El análisis formal de los términos es extremadamente valioso al intentar

rehusar ontologías existentes y al extenderlas. A menudo, desarrollar una ontología de un dominio no es la meta en sí. Desarrollar una ontología es comparable a definir un conjunto de datos y sus estructuras para que otros programas los usen. Métodos que resuelven problemas, aplicaciones independientes del dominio, y agentes de software usan ontologías y bases de conocimiento construidos a partir de ontologías como datos (16).

En resumen el uso de ontologías en diversas esferas de la vida moderna proporciona un protocolo de comunicación entre varios sistemas interrelacionados proporcionando una forma para representar y compartir el conocimiento utilizando un vocabulario común.

### 1.2.4 Herramientas para la Construcción.

Entre las diversas herramientas informáticas empleadas actualmente en el desarrollo de ontologías se destacan las siguientes (6):

#### **Protégé:**

Protégé es un editor para construir ontologías y un marco general para representar el conocimiento. Con la ayuda de este software se pueden construir aplicaciones basadas en el conocimiento que representen la información de forma ontológica. La característica de este tipo de herramientas es que utilizan un lenguaje muy flexible para ser utilizado en la Web, por eso se ha acuñado el término de Web Semántica, es decir, que facilitan la tarea no sólo de ser manejable a través de la Web, sino que hacen una descripción semántica de la información, con lo que todo el sistema es menos rígido y por tanto se hace más flexible y potente.

Protégé reconoce por tanto, Frames, XML Schema, RDF Schema y OWL, que son lenguajes semánticos utilizados en la Web, en contraposición a la rigidez del HTML. Protégé permite realizar programas en OWL con una gran facilidad, debido a su entorno gráfico, tan sólo con el uso de otro programa asociado a él que se denomina Protégé-OWL Plugin. Este es una especie de anexo a Protégé para poder trabajar con el lenguaje OWL (17).

- **KAON:** Desarrollada por la Universidad de Karlsruhe (el nombre de KAON procede de Karlsruhe Ontology). Es una herramienta de libre acceso que puede descargarse desde Sourceforge (16).
- **OilEd:** Basado inicialmente para el desarrollo de Ontologías OIL y DAML + OIL se han ido realizando numerosas actualizaciones para que acepte la mayoría de los lenguajes de especificación actuales. Es un editor bastante utilizado por los investigadores porque aporta la posibilidad de interactuar con un razonador como FACT o RACER los cuales permiten comprobar la consistencia de una Ontología. Una de las desventajas que presenta este editor es la carencia de recursos para soportar Ontologías grandes, migración e integración de otras Ontologías y diferenciación de versiones (18).
- **WebOnto: Desarrollado** por el *Knowledge Media Institute* (KMI de la Open University (Reino Unido). Se puede acceder libremente a ella, aunque los usuarios deben solicitar una cuenta de usuario a los administradores para disfrutar de todas sus capacidades de edición (18).

*Se propone como herramienta de modelado para la construcción de la ontología el editor Protégé, ya que permite realizar programas en OWL con una gran facilidad, debido a su entorno gráfico basado en plugins, además tiene a su favor una gran comunidad de desarrolladores que brindan una bibliografía abundante y actualizada, así como un entorno extensible que se adapta a la construcción de numerosos tipos de ontologías.*

### 1.2.5 Metodologías para construir Ontologías.

Actualmente existe una gran cantidad de metodologías para la construcción de ontologías, pero de todas ellas las más usadas son (1):

**Metodología Cyc** (22): nace como un proyecto de inteligencia artificial que busca la construcción de una ontología comprensible para habilitar el razonamiento humano. Esta metodología propone los siguientes pasos:

- Extracción manual del conocimiento común (de diversas fuentes)
- Utilización de herramientas de procesamiento de lenguaje natural o aprendizaje natural para la adquisición de nuevo conocimiento en la ontología.

- El proyecto Cyc surgió en el año de 1984, por parte de la Corporación de Tecnología en Computación y Microelectrónica. La base de conocimiento de Cyc es propietaria, aunque una pequeña versión fue liberada y está disponible como OpenCyc, la misma que busca definir un vocabulario común para el conocimiento automatizado.
- Actualmente Cyc cuenta con más de un millón de aserciones en su base de conocimientos que han sido definidas por el humano mediante el lenguaje CycL1.

**Metodología de Uschold y King (22):** Permite la creación de ontologías en base a otras ya existentes. Recomienda los siguientes pasos:

- Identificación del propósito para el cual se construye la ontología.
- Capturar los conceptos y las relaciones entre ellos.
- Codificación de la ontología
- Evaluación de la ontología
- Documentación de la ontología

**Metodología de Grüniger y Fox (23):** En esta metodología se proponen los siguientes pasos:

- Definición de los escenarios motivadores, es decir identificación de las posibles aplicaciones en las que la ontología será usada.
- Formulación de preguntas en lenguaje natural a las que se les denomina cuestiones de competencia, esto con la finalidad de determinar el ámbito de la ontología.
- Especificación de la terminología, en base a las preguntas realizadas en el paso anterior se define conceptos principales, relaciones, propiedades, etc.
- Formalización de las interrogantes.
- Especificación de axiomas formales
- Verificación de la ontología.
- Dentro de esta metodología las cuestiones de competencia hacen referencias a consultas a las cuales la ontología debería responder.

---

<sup>1</sup>CycL es un lenguaje ontológico usado por el proyecto de inteligencia artificial de Doug Lenat conocido como Cyc. Es un lenguaje está basado en cálculo de predicados que tiene una sintaxis similar a Lisp.

**Metodología Kactus** (24): Esta metodología centra la construcción de la ontología sobre una base de conocimiento, mediante un proceso de abstracción. Para ello hace uso de:

- Especificación del contexto de la aplicación y el punto de vista de modelado. Lo primero hace referencia a la descripción del dominio que tendrá la aplicación, así como también los objetos de interés y tareas que realizará la ontología, mientras que el punto de vista de modelado se refiere a definir qué tipo de modelado se va a realizar: dinámico-estático, funcional-causal.
- Realizar un diseño preparatorio sobre la base de una ontología existente, lo que implica realizar un estudio de ontologías que se hayan construido (mapeo).
- Refinamiento y estructuración de la ontología.
- Finalmente la documentación y reutilización de la ontología.

**Metodología Methontology** (22): Es una metodología creada en el Laboratorio de Inteligencia Artificial de la Universidad Técnica de Madrid. La creación de la ontología puede empezar desde cero o en base a la reutilización de otras existentes. Sus pasos principales son:

- **Especificación:** Consiste en delimitar los objetivos de su creación (compartir información entre personas o por agentes 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 modelar 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, etc.).
- **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 consiste en convertir el modelo anterior en un modelo formal o semi-computable. Se puede emplear en este paso una herramienta como Protégé On-to-Knowledge.

- **Implementación:** Convierte el modelo formalizado en un modelo computable mediante un lenguaje para construcción de ontologías. Se puede emplear en este paso una herramienta como Protégé.
- **Mantenimiento:** Labor que puede acarrear 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.

**Metodología On-To-Knowledge (2):** Esta metodología incluye la identificación de los objetivos alcanzados por las herramientas de gestión del conocimiento y está basada en el análisis de escenarios de uso. Los pasos propuestos por la metodología son los siguientes:

- Punto de partida: donde los requerimientos de la ontología son especificados y capturados, preguntas de habilidades son identificadas, ontologías potencialmente reusables son estudiadas y se construye una primera versión o primer borrador de la ontología.
- Refinamiento: donde se produce una ontología madura y orientada a aplicación.
- Evaluación: donde los requisitos y las preguntas de habilidades son chequeadas y la ontología es testeada en el entorno de la aplicación.
- Mantenimiento de la ontología.

### **Metodología Interna del CDAE para la Construcción de Ontologías.**

La metodología propuesta por el CDAE consta de varios elementos dentro de los cuales se encuentran: seis fases para la creación de una ontología, tres escenarios en dependencia del entorno y las características de este, varios roles, artefactos, técnicas, herramientas y estándares. Estos se describen puntualmente en las secciones siguientes.

#### Escenarios de aplicación de las ontologías

El desarrollo de una ontología se propone el estudio, especificación y representación de diversos escenarios, entiéndase por escenarios diversos ámbitos que pudieran propiciar las necesidades para el desarrollo de ontologías, para diversos fines y propósitos que pueda dársele.

Durante la ejecución de los proyectos desarrollados en el CDAE fueron identificados diversos entornos o escenarios de aplicación de las ontologías, que implicaron un esquema de trabajo diferente en cada caso, véase Figura. 1.2

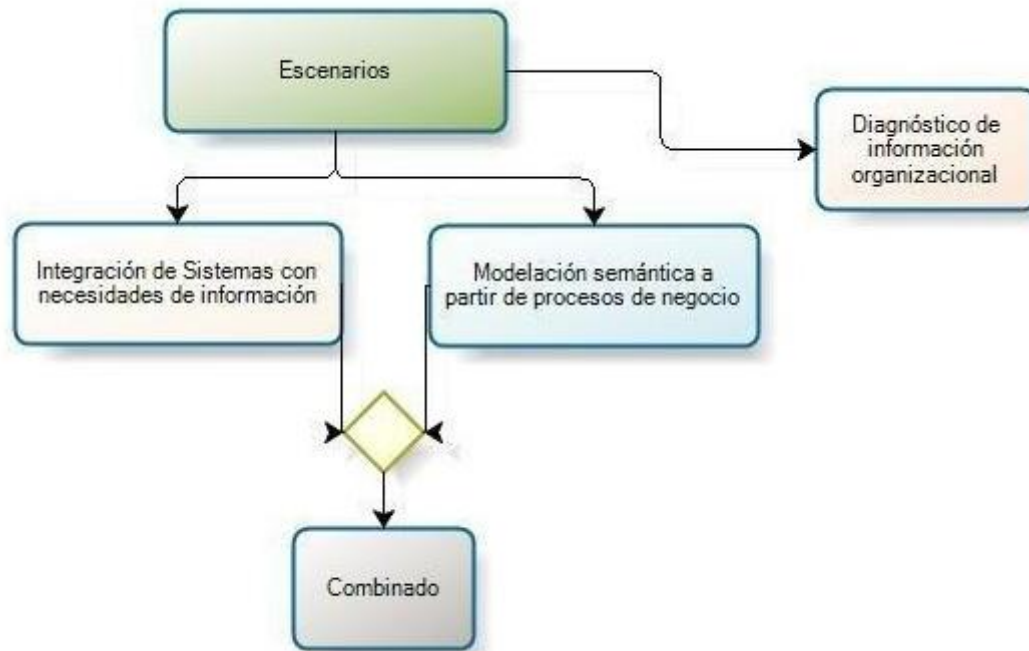
### *Escenario 1: Desarrollo de ontologías partiendo de necesidades de información de la organización*

Dichas necesidades de información pudieran ser requerimientos de un sistema, necesidades de manejo de datos o simples consultas a sistemas de almacenamiento que proporcione parcial o completamente la información requerida.

Las necesidades de información pudieran ser satisfechas mediante la vinculación de varias fuentes de datos o sistemas autónomos, o sea, un mismo requisito podría demandar de partes de información de varios sistemas o fuentes para conformar y dar respuesta a la necesidad en particular. Podría darse el caso que la información y los datos de la organización no estén completamente automatizados, en este caso la fuente será documental y se creará la ontología procesando la información disponible.

Para este escenario en particular gran parte de las actividades para el desarrollo de la ontología estarán destinadas al estudio y análisis de las fuentes relacionadas con las necesidades de información de la organización, en aras de verificar el estado de las fuentes y conformar un modelo ontológico estándar que permita el entendimiento común entre la organización y las fuentes o sistemas que satisfacen las necesidades de información.





**Figura. 1.2** Escenarios para el desarrollo de ontologías

### *Escenario 2: Desarrollo de ontologías partiendo de los procesos de negocio de la organización*

Es un desarrollo muy característico cuando en una organización sus procesos están bien definidos y estructurados, ya sea porque se haya organizado a través de BPM o por la excelencia de la industria a la que pertenezca y el nivel de organización para el control que ello requiera. Partiendo del estudio de los procesos de negocio, y de la información que de su análisis se arroje, se desarrolla una ontología propia del negocio de la organización, teniendo como principal característica que estará completamente alineada al funcionamiento de la organización y a la cadena de valor de sus procesos.

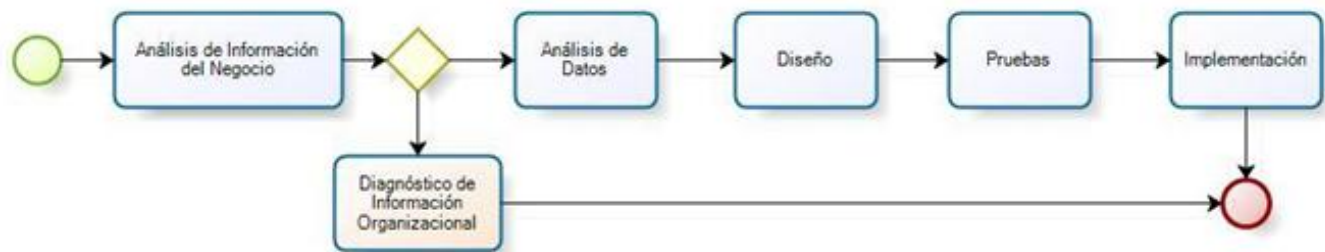
Es aconsejable, en el caso de encontrarse la organización en fase de levantamiento o reestructuración de sus procesos, o ya sea que posea documentación de estudios previos de sus procesos, que las actividades para el desarrollo de ontologías se abastezcan directamente de la información recogida o participe en el levantamiento.

*Escenario 3: Desarrollo de ontologías partiendo de procesos de negocio y necesidades de información de la organización. Combinado*

Es el caso cuando se combinan dos escenarios y se crea un híbrido entre ambos. Se necesitan integrar soluciones dispersas y no conectadas, con la característica que la integración de los sistemas es basado en las necesidades de información que se determine. Este caso de desarrollo de ontologías es propiciado cuando los sistemas son muy complejos, las necesidades de información vinculas a varios sistemas y el estudio de cada una se torna una tarea compleja, por lo que se ajustan los dominios y el alcance solo a satisfacer las necesidades de información que se determinen.

### Fases para el desarrollo de la metodología

Se han descrito un total de cinco fases básicas para el desarrollo de una ontología, con una fase adicional en dependencia de las necesidades del cliente y las características de la organización o ente a la cual se le realice la ontología, véase la Figura. 1.3 Las fases delimitan las acciones que se realizan durante la metodología. Como parte de esta se han diseñado un conjunto de actividades correspondientes a cada fase, así como una lista de artefactos como resultado de estas.



**Figura. 1.3** Fases para el desarrollo de una ontología

*Diagnóstico de la información organizacional:* relaciona un conjunto de actividades que pueden realizarse opcionalmente, en el caso de que al estudiar la organización en la primera fase, se decida diagnosticar la entidad, la información que maneja, las interrelaciones a nivel macro, un estudio global que conlleve aun análisis que produzca una valoración del estado actual de la empresa para asumir un escenario en particular. Es una fase completamente informativa para la organización, puede considerarse en el caso que no se encuentren explícitas ciertas características para la selección de uno y otro escenario, cuando no exista la información del estado de sus fuentes, especificación de sus procesos de negocio o cuando no estén delimitadas y establecidas necesidades de información.

## Capítulo 1: Fundamentación Teórica

---

*Análisis de la información del negocio:* Comprende las actividades iniciales de comprensión del negocio y de la planeación de las acciones a ejecutar durante todo el desarrollo. Se establecen el dominio y alcance del negocio, y durante el proceso de planificación se evalúa cuál escenario es el adecuado en dependencia de las necesidades establecidas por el cliente, luego se seleccionan los recursos y se establece el cronograma de tareas

*Análisis de datos:* Se analizan los datos del negocio, se comienzan a levantar los términos, en cualquiera de los tres escenarios expuestos anteriormente. Se identifican las relaciones y restricciones de los términos del negocio. Es importante señalar que en esta fase se define el dominio y alcance que tendrá la ontología. Se valora además la posible reutilización de ontologías o cualquier otra forma de visualización y representación de la información o los datos de la organización que se estudia.

*Diseño:* Comprende el diseño de la información recogida del negocio, de los términos relevantes de la fase de análisis de datos, obteniendo como resultado un modelo de datos. Adicionalmente se diseña el modelo ontológico en un editor de ontologías, como es el caso del Protégé.

*Pruebas:* Es la fase donde se crean las instancias de las clases definidas en el modelo ontológico. Se diseñan casos de prueba en correspondencia a las preguntas de competencia que se definieron en la fase de análisis de datos. Se generan los esquemas de las clases y los metadatos de la ontología que podrán ser consumidos para el desarrollo de aplicaciones y sistemas.

*Implementación:* Fase en la que se implementa un motor semántico que emplee reglas de inferencia conjuntamente con reglas del negocio establecidas previamente y una base de conocimiento almacenada.

*La metodología interna del CDAE es la seleccionada para el desarrollo del modelo ontológico del sistema a desarrollar ya que conjuga las buenas prácticas de las demás metodologías y su vez le atribuye componentes únicos de acuerdo a las necesidades actuales, que para el desarrollo de sistemas requiere el CDAE.*

### 1.2.6 Lenguajes de Representación.

Las ontologías son representaciones formales de la semántica de los datos y para constituirse deben contar con un mecanismo que le permita esta actividad. Este mecanismo no es más que los lenguajes de representación de las ontologías (4):

**SHOE** (Simple HTML Ontology Extensions) (2): fue el primer lenguaje de etiquetado para diseñar Ontologías en la Web. Este lenguaje nació antes de que se ideara la Web Semántica. Las Ontologías y las etiquetas se incrustaban en archivos HTML. Este lenguaje permite definir clases y reglas de inferencia, pero no negaciones o disyunciones. Con él se desarrollaron muchos editores, buscadores, APIs, etc; el proyecto fue abandonado a medida que se desarrollaron OIL y DAM; aunque también existe una socialización de este lenguaje en XML.

**OIL** (Ontology Inference Layer) (27): este lenguaje, derivado en parte de SHOE, 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 RDFS. Se basa tanto en la lógica descriptiva (declaración de axiomas) como en los sistemas basados en frames (taxonomías de clases y atributos). 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.

**OWL** (Web Ontology Language o Lenguaje de Ontologías para la Web) (27): 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 (Ontologías). En realidad, OWL es una extensión del lenguaje RDF y emplea las tripletas de 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, etc.).

Existen 3 sub lenguajes de OWL, los cuales van creciendo respecto al nivel de expresión:

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 he 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.

**XML SCHEMA**: Es un lenguaje cuyo objetivo principal es definir la estructura en bloques de un documento *XML* pero de una forma mucho más precisa. El propósito de un esquema es definir y describir una clase de documentos XML usando estas construcciones para restringir y documentar el significado, uso y relaciones de las partes constituidas: tipo de datos, elementos y su contenido, atributos y sus valores, entidades y su contenido y anotaciones. Los esquemas documentan su propio significado, uso y función (7).

**RDF (ResourceDescription Framework)**: Es un esquema para definir e intercambiar metadatos basado en XML, que se ocupa de la semántica y permite, mediante relaciones de clasificación, crear una jerarquía de conceptos de un dominio. En pocas palabras, es el modelo que establece el formato de datos que permite al usuario crear lenguajes entendibles para las computadoras de manera que estas puedan procesar y clasificar semánticamente la información (8). Lenguaje para la definición de Ontologías y metadatos en la web. El elemento de construcción básica en RDF es el “triple” o sentencia, que consiste en dos nodos (sujeto y objeto) unidos por un arco (predicado), donde los nodos representan recursos y los arcos propiedades.

**RDF Schema (RDFS):** Se pueden definir jerarquías de clases de recursos, especificando las propiedades y relaciones que se admiten entre ellas (8).

**KIF** (Knowledge Interchange Format) (27): es un lenguaje para representar Ontologías basadas en la lógica de primer orden. KIF está basado en la lógica de predicados con extensiones para definir términos, meta conocimiento, conjuntos, razonamientos no mono tónicos, etc. y pretende ser capaz de representar la mayoría de los conceptos y distinciones actuales de los lenguajes más recientes de representación del conocimiento. Está diseñado para intercambiar conocimiento entre distintos sistemas de computación.

*Se propone como lenguaje de representación de ontologías el OWL en su versión DL ya que es un lenguaje diseñado para usarse en el procesamiento de la información contenida en el modelo ontológico cuando es procesada por programas o aplicaciones de software brindando propiedades avanzadas para el proceso de inferencia, lo cual es sumamente importante para el sistema a desarrollar.*

### 1.3. Sistemas Basados en Conocimientos.

En la década del 70 se reconoció que los métodos de solución de problemas generales eran insuficientes para resolver los problemas orientados a aplicaciones. Se determinó que era necesario conocimiento específico sobre el problema, limitado a los dominios de aplicación de interés, en lugar de conocimiento general aplicable a muchos dominios. Este reconocimiento condujo al desarrollo de SBC (28).

El conocimiento representado en los SBC es el de los expertos en el dominio. Una parte de un conocimiento experto consiste de relaciones de causa y efecto. Estas relaciones o reglas se originan a partir de la experiencia pasada y son llamadas heurísticas. Ellas representan conocimiento informal o atajos, que permiten a un experto encontrar rápidamente una solución a un problema sin tener que realizar un análisis detallado de una situación particular debido a un análisis de un problema similar resuelto de forma exitosa anteriormente o a relaciones que han sido aprendidas como resultado de un intento fallo pasado para resolver un problema similar. El experto puede no

recordar todos los detalles del análisis del problema original, pero puede reconocer el enfoque aplicado a un problema similar (29).

El uso de heurísticas en los SBC es marcadamente diferente de las funciones heurísticas usadas en la búsqueda heurística. Las funciones heurísticas representan conocimiento general usado para guiar la búsqueda, mientras que la heurística en los SBC consiste en el conocimiento heurístico usado por un experto en la solución de un problema, esta es más específica del dominio y más profunda en naturaleza. Un experto es una persona que posee habilidades que le permiten sacar conclusiones de experiencias pasadas y rápidamente focalizar sobre el centro de un problema dado. Las mayores posibilidades del éxito de un experto en la solución de problemas se deben a que ha adquirido un conjunto de relaciones de causa y efecto poderosas basadas en la experiencia. Un experto es capaz de utilizar este conocimiento básico para reconocer rápidamente rasgos sobresalientes del problema, clasificar este de acuerdo a estas características y buscar una solución. Últimamente, se reconoce que en lugar de relaciones de causas y efecto los expertos recuerdan casos similares al problema a resolver y toman de ellos las soluciones dadas como punto de partida para encontrar la solución al nuevo problema (30).

Al experto lo caracterizan:

- Una mayor amplitud en la variedad de problemas que puede resolver que los otros profesionales del dominio.
- Más rapidez en encontrar una solución más calidad en la solución encontrada.

El conocimiento en cualquier especialidad es usualmente de dos tipos: público y privado. El conocimiento público incluye las definiciones, hechos y teorías publicadas. Pero la experticia usualmente incluye más que esta clase de conocimiento. Los expertos humanos generalmente poseen conocimiento privado. Elucidar y reproducir tal conocimiento es la tarea central en la construcción de sistemas basados en el conocimiento; el sujeto de esta acción es el ingeniero de conocimientos (IC). Sobre el conocimiento público hay consenso, el privado puede llevar a controversias entre los expertos (31).

Entre las razones que sustentan enfatizar en el conocimiento en lugar de en los métodos de razonamiento formal están (32):

- Los expertos humanos obtienen su destacado desempeño gracias al conocimiento que poseen.
- La mayoría de los problemas difíciles e interesantes no tienen soluciones algorítmicas tratables.
- El conocimiento es una fuente escasa cuyo refinamiento y reproducción crea riqueza.

A modo de resumen, la potencia de un SBC radica en el conocimiento que posee.

### 1.3.1 Definición de SBC.

En términos generales, un SBC puede ser definido como (30):

*Un sistema computarizado que usa conocimiento sobre un dominio para arribar a una solución de un problema de ese dominio. Esta solución es esencialmente la misma que la obtenida por una persona experimentada en el dominio del problema cuando se enfrenta al mismo problema.*

Esta definición puede ser aplicada a programas que no se consideran SBC. Por ejemplo, un programa en un lenguaje procesual que permita calcular magnitudes usando fórmulas conocidas llega a iguales resultados que los especialistas que manejan esas fórmulas para resolver un problema; y no por eso dicho programa se puede considerar un SBC. Realmente, a los SBC lo caracterizan otros rasgos además de duplicar el conocimiento y experticia de un experto humano para un dominio específico. Tres conceptos fundamentales relativos a los SBC lo distinguen de los programas basados en búsqueda general (33):

- La separación del conocimiento de cómo este es usado (distinción entre conocimiento y estrategia de control).
- El uso de conocimiento muy específico del dominio.
- Naturaleza heurística, en lugar de algorítmica, del conocimiento empleado.

Generalmente los términos SBC y Sistemas Expertos se usan indistintamente, aunque algunas autores limitan el uso del término SE a aplicaciones donde el conocimiento al nivel de experto es requerido.

Los SBC tienen ventajas y desventajas cuando se comparan con otras soluciones como el software convencional o expertos humanos.



### 1.3.4 Ventajas del uso de SBC:

El uso de los SBC como aplicaciones informáticas de uso práctico trae ventajas algunas de las cuales se presentan a continuación (34):

- Consistencia en las respuestas (los expertos humanos pueden diferir en sus explicaciones, incluso un mismo experto puede responder de forma diferente en momentos diferentes).
- Gran accesibilidad (los SBC trabajan las 24 horas todos los días).
- Preservación de la experticia (constituye una memoria institucional y poseen la capacidad para adquirir nuevo conocimiento y perfeccionar el que poseen).
- Solución de problemas que incluyen datos incompletos.
- Explicación de soluciones (justifica sus conclusiones y explica por qué hace una pregunta).
- Permite evaluar el efecto de nuevas estrategias añadiendo o modificando conocimiento.
- Constituye un entrenador en el dominio de aplicación.
- Fácil modificación (conocimiento explícito y accesible).
- Amplia distribución de experticia escasa.

### 1.3.5 Desventajas del uso de SBC:

Al igual que el uso de los SBC trae consigo muchas ventajas también implica algunos riesgos o desventajas algunas de las cuales son (34):

- Las respuestas no siempre son correctas.
- Conocimiento limitado al dominio de experticia.
- Ausencia de sentido común.
- No reconocen el límite de su conocimiento.

### 1.3.6 Categorías de aplicación de los SBC.

Los SBC tienen innumerables usos en los procesos empresariales entre los cuales están (35):

- Interpretación: Son aquellos Sistemas que a partir de datos suministrados por determinados sensores pueden inferir situaciones determinadas.

- Producción: Son aquellos Sistemas que a partir de datos suministrados pueden Inferir consecuencias de situaciones dadas.
- Diagnóstico: A partir de observaciones inferir el mal funcionamiento de un sistema.
- Diseño: Son utilizados para configurar objetos bajo instrucciones especificadas.
- Monitoreo: Comparan observaciones con resultados esperados.
- Instrucción: Se utilizan en la educación para diagnosticar, evaluar y corregir el comportamiento de un estudiante.
- Control: Se utilizan para gobernar o supervisar el comportamiento de otros sistemas.

### 1.3.7 Tipos de SBC.

En la siguiente tabla se realiza una caracterización de los principales tipos de SBC así como la especificación de algunos de sus atributos más relevantes.

Nombre	Forma de Representación del Conocimiento (FRC)	Explicación	Método de solución del Problema (MSP)	Fuentes de Conocimiento
Sistemas basados en reglas (SBR)	Reglas de producción	Reglas activas	Usualmente búsqueda primero en profundidad con dirección backward o forward	Expertos, publicaciones, ejemplos
Sistemas basados en frames (SBF)	Frames		Herencia y procedimientos adjuntos	Expertos, publicaciones, ejemplos

Sistemas basados en casos (SBCasos)	Casos	Casos semejantes	Razonamiento basado en casos (búsqueda por semejanza y adaptación de las soluciones)	Ejemplos
Sistemas basados en probabilidades (SBP)	Probabilidades o frecuencias	Valores de probabilidades condicionales	Teorema de Bayes y otras técnicas de inferencia estadística	Ejemplos
Redes expertas	Pesos y alguna otra FRC	Según el modelo simbólico	Cálculo de niveles de activación de las neuronas	Ejemplos
Sistemas basados en modelos	Modelo del artefacto		Razonamiento basado en modelos	Esquemas estructurales y funcionales del artefacto

**Tabla 1.1** Elementos Importantes de los diferentes tipos de SBC (31).

### 1.3.8 Sistemas Basados en Reglas.

Los SBR son los más conocidos de los SBC. Los sistemas basados en el conocimiento clásico como Dendral, MYCIN y R1<sup>2</sup> también conocido como XCON fueron SBR.

<sup>2</sup>Dendral, MYCIN y R1 son los nombres de SBC clásicos que se utilizaron en propósitos reales demostrando la funcionalidad de estos sistemas, cada uno de ellos marco una pauta en el desarrollo de estos sistemas.

Los SBR son SBC en los que la forma de representación del conocimiento usada son las reglas de producción y como método de inferencia utiliza la regla de modus ponens.

Los SBR son llamados frecuentemente sistemas de producción, una idea de 1943 expresada por E. L. Post.

Las reglas utilizan un formato IF - THEN para representar el conocimiento, la parte IF de una regla es una condición (también llamada premisa o antecedente), y la parte THEN de la regla (también llamada acción, conclusión o consecuente) permite inferir un conjunto de hechos nuevos si se verifican las condiciones establecidas en la parte IF (36).

Las reglas pueden expresar un amplio rango de asociaciones (37):

- Situación – acción  
Si está lloviendo y va a salir, entonces debe buscar una capa.
- Premisa – conclusión  
Si su temperatura es de 40°C, entonces usted tiene fiebre.
- Antecedente – consecuente  
Si x es un gato, entonces x es un animal.

El proceso de solución de problemas en un SBR es crear una cadena de inferencias que constituye un camino entre la definición del problema y su solución. Esta cadena de inferencias puede construirse por dos vías “direcciones de búsqueda” (31):

- Comenzar con todos los datos conocidos y progresar hacia la conclusión (data driven o forward chaining).
- Seleccionar una conclusión posible y tratar de probar su validez buscando evidencias que la soporten (goal driven o backward chaining)
- La dirección forward es apropiada cuando:
  - Hay pocos datos de entrada ó
  - La cantidad de conclusiones posibles es grande.

Entre los campos de aplicación para esta dirección de búsqueda están monitoreo y diagnóstico en sistemas de control en tiempo real, diseño y planificación (29).

La dirección backward es apropiada cuando:

- Hay pocas conclusiones posibles.
- Los valores de entrada no son adquiridos automáticamente.

Los problemas de diagnóstico y clasificación son frecuentemente resueltos con esta dirección de búsqueda.

### 1.3.9 Ventajas de los SBR.

El uso del desarrollo de SBR trae consigo un conjunto de características ventajosas entre ellas se destacan (30):

- **Modularidad:** Los SBR son altamente modulares. Cada regla es una unidad de conocimiento que puede ser añadida, modificada o removida independientemente de las otras reglas existentes (a pesar de los problemas antes señalados). Esto da flexibilidad en el desarrollo de la BC.
- **Uniformidad:** Todo el conocimiento del sistema se expresa en el mismo formato.
- **Naturalidad:** Las reglas son un formato natural para expresar conocimiento en algunos dominios. Los expertos lógicamente piensan en los problemas y sus soluciones usando las situaciones existentes para indicar las conclusiones deseadas.

Comparación entre el proceso de toma de decisiones por un Sistema Basado en Conocimiento y por el sistema tradicional.

Sistema Basado en conocimiento	Sistema Tradicional
No tiene limitación de volumen de datos.	Manipula conocimiento.
Procesa datos.	El control está separado del conocimiento.
El control depende de la aplicación.	Se agrega conocimiento para mayores resultados.
Se perfeccionan la lógica y los cálculos para ampliar los resultados.	Para ser complementado requiere la participación de expertos en el área.

Su complementación depende, en algunos casos, de la capacidad de la herramienta o del programador.	El proceso en general es heurístico.
Provee ayuda para el proceso pero no por los resultados.	Es interactivo.
Pueden ser interactivos o no.	Da explicaciones y ayuda de acuerdo con el problema o proceso.
El proceso es algorítmico.	El usuario debe tener conocimiento explícito del área.
El usuario sólo requiere instrucción de manejo.	El área de conocimiento del problema es reducida.

**Tabla 1.2** Comparación de los modelos de toma de decisiones (30).

La composición de un sistema experto es simple: 1) Bases de conocimiento (hechos, reglas, heurísticas) y; 2), Motor de inferencia (elemento que procesa las reglas en las bases de conocimiento). Además incluye: Interfaz del usuario y módulos explicativo e interrogativo.

### 1.4. Base de Conocimiento.

Las bases de conocimiento surgieron a partir de la investigación en Inteligencia Artificial como respuesta a las necesidades que las aplicaciones de esta disciplina planteaban. Más adelante se hace un muestreo de las aportaciones que la IA ha hecho al estudio del lenguaje natural. Las bases de conocimiento son la evolución lógica de los sistemas de bases de datos tradicionales, en un intento de plasmar no ya cantidades ingentes de datos, sino elementos de conocimiento, así como la manera en que este ha de ser utilizado. También se les trata de dotar de conocimiento sobre sí mismas, es decir, una KB ha de “saber lo que sabe” (28).

Las bases de conocimiento contienen los hechos y las reglas necesarias para hallar solución a los problemas. La estructura de las bases de conocimiento depende de la experiencia del Ingeniero de Conocimiento para extraer conocimiento de los expertos y la capacidad para representarlo (18):

- a. ¿Qué objetos serán definidos?
- b. ¿Qué relaciones existen entre los objetos?
- c. ¿Cómo se procesarán y formularán las reglas?
- d. ¿Cuál debe ser la consistencia de las Bases de Conocimiento?

Las bases de conocimiento se han clasificado en dos grandes tipos (29):

**Bases de conocimiento legibles por máquinas**, diseñadas para almacenar conocimiento en una forma legible por el computador, con el fin de obtener razonamiento deductivo automático aplicado a ellas. Contienen una serie de datos, usualmente en la forma de reglas que describen el conocimiento de manera consistente. Operadores lógicos como *Y* (conjunción), *O* (disyunción), *condición lógica* y *negación* son utilizados para aumentarla desde el conocimiento atómico. En consecuencia, la deducción clásica puede ser utilizada para razonar sobre el conocimiento en la base de conocimiento. Este tipo de bases de conocimiento son utilizadas por la Web semántica.

**Bases de conocimiento legibles por Humanos**, diseñadas para permitir a las personas acceder al conocimiento que ellas contienen, principalmente para propósitos de aprendizaje. Estas son comúnmente usadas para obtener y manejar conocimiento explícito de las organizaciones, incluyen artículos, manuales de usuario y otros. El principal beneficio que proveen las bases de conocimiento es proporcionar medios de descubrir soluciones a problemas ya resueltos, los cuales podrían ser aplicados como base a otros problemas dentro o fuera del área de conocimiento.

### 1.5. Motor de Inferencia.

Una ontología es una opción muy adecuada, ya que es un formalismo muy extendido para modelar y almacenar conocimiento. Sin embargo, sin otras herramientas, de una ontología no es posible extraer más conocimiento que los conceptos almacenados en ella. Aun así, su organización permite un razonamiento sencillo sobre ella utilizando un razonador. Estas herramientas permiten, dada una ontología, razonar nuevas relaciones o conceptos no explícitos en el modelo original (38).

La definición de una ontología involucra tres componentes principales:

- Clases e instancias, que son de forma básica los objetos o elementos que componen la ontología.
- Relaciones, que representan las interacciones entre los anteriores.
- Reglas, que se emplean para modelar el conocimiento que no puede recogerse a través de los elementos anteriores.

Las ontologías que no incluyen o desarrollan las reglas son las llamadas ontologías ligeras, ello no es más que una clasificación de los elementos y trae como consecuencia la pérdida de la capacidad expresiva de la ontología.

En el caso particular de las ontologías, cuanto mayor sea el trabajo de definición de axiomas y relaciones, mayor será su potencial semántico. De esta forma una ontología se convierte en algo más que en un establecimiento de categorías o subcategorías de conceptos. Es en este momento cuando el empleo de razonadores cobra gran interés. Tal y como se perfila en los párrafos anteriores el empleo de razonadores está motivado por el gran poder lógico o de deducción de conocimiento que presentan las ontologías. De esta forma podríamos decir que recogen más conocimiento del que realmente se presenta en un principio y este puede extraerse, haciendo deducciones o respondiendo a consultas lógicas, mediante el empleo de razonadores (24).

### 1.6. Lógica Descriptiva.

Lógica Descriptiva (DL) es un lenguaje de representación de conocimiento adaptado para expresar conocimiento sobre conceptos y jerarquías entre estos, mediante la creación de ontologías formales. Este tipo de lenguajes es apropiado para estructurar la información.

Es un formalismo lógico cuyo origen son los fundamentos de otros formalismos como redes semánticas o frames. Hay una gran variedad de DL, siendo especificada cualquier lógica por tres conjuntos de características (39):

- Operadores para formar conceptos, tales como conceptos primitivos, conjunción y cuantificadores existenciales.
- Operadores de formación de propiedades, tales como primitivas e inversas.
- Axiomas soportados, tales como implicación de conceptos igualdad o implicación en las relaciones.



## 1.6.1 Servicios de inferencia en Lógica Descriptiva

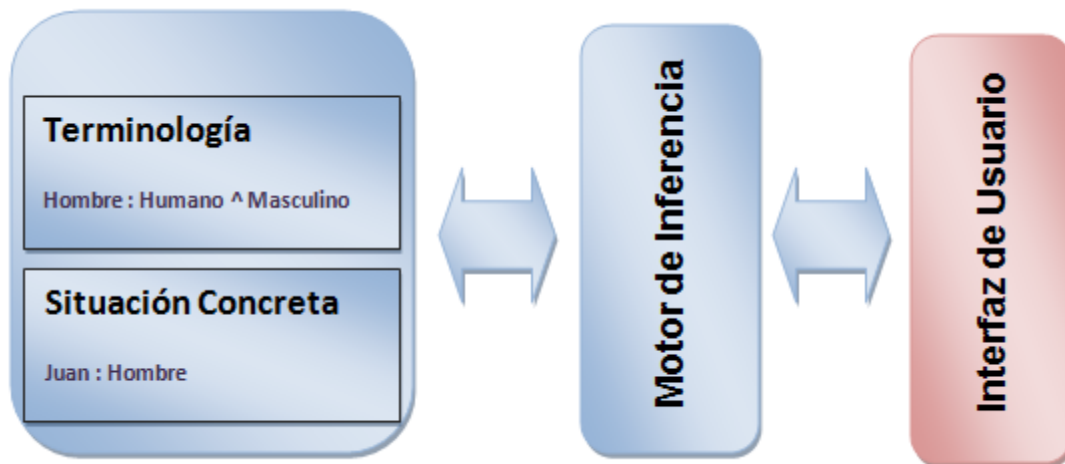


Figura 1.4 Arquitectura de un Sistema Basado en Conocimiento (22)

La arquitectura de un Sistema Basado en Conocimiento está formada por tres subsistemas: la base de conocimiento, un mecanismo o sistema de inferencia y una interfaz. En la base de conocimiento se almacena información sobre el dominio organizada utilizando una jerarquía de clases y relaciones entre ellas. En ella se distinguen claramente dos partes bien diferenciadas: Terminology Box (Tbox) y Assert Box (Abox). La primera de ellas contiene el conocimiento normativo (terminológico) en forma de una terminología y es construido a través de declaraciones que describen propiedades generales de conceptos (definiciones intencionales de conceptos y roles). El Abox contiene conocimiento extensional o factual (aserciones sobre objetos), específico para los individuales del dominio de discurso. En otras palabras, los Tbox contienen las definiciones de los conceptos y roles, mientras que los Abox contienen las definiciones de los individuales (instancias).

El sistema de inferencia se utiliza para poder razonar sobre la base de conocimiento. El conjunto de razonamientos que se pueden realizar está directamente relacionado con la expresividad de la DL que se esté utilizando y finalmente la interfaz que permite la interoperabilidad con el Sistema Terminológico (22).

Un Sistema Terminológico permite razonar sobre el dominio de conocimiento. Los principales problemas de inferencia que se plantean resolver utilizando un Sistema de Razonamiento suelen ser similares en todos los dominios de aplicación e independientes de la DL utilizada.

La representación de conocimiento y el sistema de razonamiento basado en DL provee un número determinado de facilidades de razonamiento que puede ser separado en dos grupos: razonamiento terminológico (sobre el Tbox) y razonamiento sobre las instancias (Abox) (22).

### Razonamiento Terminológico.

Las cuatro tareas siguientes son la base del razonamiento terminológico, implementadas en la mayoría de los sistemas:

- **Subsunción:** Consiste en comprobar la posición que ocupa una clase dentro de la jerarquía de la base de conocimiento. Comprobar si una clase es subclase de otra.
- **Consistencia:** Consiste en chequear si el modelo de conocimiento que se ha construido es consistente, si la estructura jerárquica generada es consistente con el conocimiento almacenado. Comprobar si existen definiciones de clases inconsistentes.
- **Construcción taxonómica:** Consiste en procesar todas las relaciones de subclase (incluyendo aquellas que no son explícitamente especificadas pero que son implicadas por las definiciones dadas).
- **Clasificación:** Determina las clases que inmediatamente son más generales que otras o aquellas que son incluidas (menos generales) que una clase dada (22).

### Razonamiento Instancial.

Las tareas básicas que involucran los individuales o instancias son:

- **Realización:** Dada una descripción parcial de un individual, encontrar el concepto más específico que lo describe.
- **Comprobación de instancias:** Dada una descripción parcial de un individual y una descripción de una clase, averiguar si la clase describe la instancia. Consiste en chequear si un individuo es instancia de una determinada clase cumpliendo las características de esta.

- **Recuperación de instancias:** Encontrar todos los individuales que son descritos por un concepto dado. Consiste en obtener todas las instancias de una determinada clase.
- **Equivalencia:** Hace referencia a comprobar si el conjunto de instancias de una determinada clase es el mismo al de otra (22).

### 1.6.2 Herramienta para la construcción de SBC

#### **Pellet:**

Es un razonador de Lógica Descriptiva del Lenguaje OWL basado en Java. Puede ser utilizado conjuntamente con bibliotecas del API de Jena o del OWL. Mediante su uso es posible validar, comprobar la consistencia de Ontologías, clasificar la taxonomía y contestar a un subconjunto de consultas RDQL (conocido como consultas a Abox en terminología del DL). Algunas de las funcionalidades de esta herramienta son: (40)

- **Comprobación de la consistencia:** Comprobando 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.
- **Corrección de los conceptos:** Verifica si es posible que se definan instancias para una clase; si no es así, definir una instancia para una clase provocaría que toda la ontología fuera inconsistente.
- **Clasificación:** Observa la relación entre cada clase y crea la jerarquía de clases completa.
- **Realización:** 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.

#### **RACER (Renamed Abox and Concept Expression Reasoner):**

RACER fue desarrollado por Ralf Möller y Volker Haarslev en 1999, pero que ha sido renovado periódicamente hasta la fecha. 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 y 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 (40).

#### **BOR:**

Es un razonador desarrollado por el Laboratorio Sirma del proyecto On-To-Knowledge. Tiene soporte para ambos tipos de inferencia, tanto sobre instancias como sobre conceptos: chequeo de la consistencia y del modelo de la Ontología, construcción de la jerarquía de conceptos (40), clasificación de conceptos definidos. El razonador puede ser usado con Ontologías escritas en DAML + OIL, con algunas restricciones y con Ontologías escritas en la especificación OWL Lite. Además se puede incorporar a la aplicación Sesame para dar soporte a Ontologías DAML + OIL en este tipo de repositorios y poder inferir conocimiento o simplemente recuperarlo (13).

### **Cerebra:**

Es un motor de inferencia desarrollado por Network Inference. Parecido al razonador de lógica descriptiva FaCT, sólo que en este caso sí aporta soporte para la inferencia sobre instancias y tipos de datos. Soporta Ontologías RDF, DAML + OIL y OWL, y como lenguaje de consulta para razonar emplea Xquery [Xquery04]. Tiene versiones para trabajar conjuntamente con los editores OilEd y Protégé (40).

### **Jena:**

Entorno para Web Semántica de código libre programado en Java. Proporciona una API para extraer y escribir datos de un grafo RDF. Los modelos pueden ser consultados mediante SPARQL. Se puede utilizar OWL con Jena. Proporciona varios razonadores internos y se pueden añadir otros mediante una interfaz para Descripción Lógica (DIG). El razonador Pellet se puede conectar directamente y así superar las limitaciones de velocidad de DIG (16).

Los requisitos fundamentales del razonador para ofrecer descubrimiento y búsqueda inteligente son los siguientes (16):

- Licenciamiento abierto. Idealmente el razonador elegido debería distribuirse en código abierto, por si la funcionalidad ofrecida no satisface plenamente los requisitos, poder modificar/añadir funcionalidad de manera sencilla.
- Ofrecer una API adecuada. Dado que el razonador no va a ser una herramienta de uso final, sino que será utilizada por una plataforma que posteriormente se desarrollará, se necesita una API de acceso a la funcionalidad del razonador.

*Se propone el empleo de Jena como framework de desarrollo del SBC debido a las múltiples funcionalidades que este brinda, principalmente en los aspectos de almacenamiento persistente, manejo de los elementos del modelo ontológico y razonamiento sobre el mismo. Además cuenta con una amplia documentación actualizada en su sitio oficial.*

### **1.7. Conclusiones Parciales**

A partir del análisis realizado sobre los principales elementos relacionados con el marco teórico de la investigación, se concluye que:

- Se propone el uso de las Ontologías como herramienta para la representación del conocimiento.
- Fue seleccionada la metodología propuesta por el CDAE para la construcción del modelo ontológico.
- Se propone el empleo del lenguaje OWL-RDF para la representación de un modelo ontológico.
- Fueron seleccionadas, respectivamente, las herramientas Protégé y Jena para el modelado e implementación de una Ontología.

## **CAPÍTULO 2: DISEÑO DE LA SOLUCIÓN.**

En este capítulo se describen los principales elementos que conforman la solución propuesta. Se abordan los elementos que conforman el diseño de la ontología: el dominio y alcance, las preguntas de competencia, los términos y conceptos relevantes en el negocio, entidades y sus propiedades y relaciones, así como la especificación de su tipo de persistencia. Por otra parte se hace una introducción al dominio de negocio para el cual se propone la solución en cuestión.

### **2.1. Análisis de la información del negocio**

Durante la 1ra fase de la metodología propuesta para la construcción de ontologías, se propone el análisis de la información del negocio. En este sentido se debe comenzar con la definición del dominio y alcance del negocio que comprende el modelo ontológico que se presente construir.

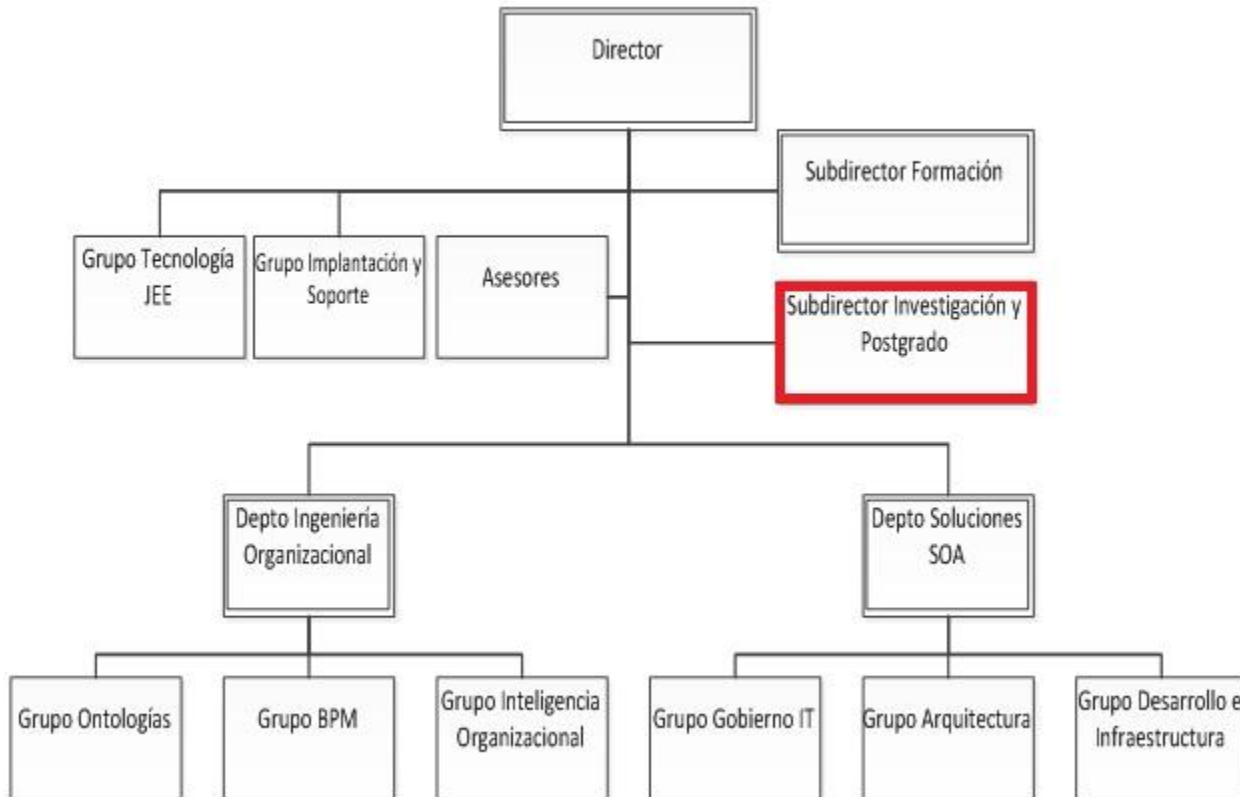
La Ontología propuesta modelar algunos de los procesos que tienen lugar en el Área de Investigación y Postgrado del CDAE, así como la incorporación de algunos elementos de información de otras áreas del centro.

#### **2.1.1 Introducción al Negocio.**

En la Universidad existe el Centro de Consultoría y Desarrollo de Arquitecturas Empresariales (CDAE), enmarcado dentro de la estrategia de la UCI de creación de centros de desarrollo, con el objetivo de elevar el nivel de respuesta frente a las crecientes demandas para la informatización del país y de exportación de productos informáticos y servicios profesionales.

Tiene como Misión brindar servicios de consultoría en tecnologías informáticas y desarrollar soluciones para organizaciones que buscan optimizar sus procesos de negocio y elevar la eficiencia operacional, empleando los modelos de Arquitectura Empresarial y SOA/BPM como paradigmas tecnológicos de referencia.

### Estructura organizativa del centro



**Figura. 2.1** Estructura Organizativa del CDAE.

El Área de Investigación y Postgrado es el dominio de aplicación del presente trabajo. Dicha área gestiona la información referente a las investigaciones científicas del CDAE, así como del proceso de superación posgraduada de los especialistas del centro.

## 2.2. Análisis de datos

Siguiendo la metodología del CDAE para la construcción de ontologías luego de haber analizado la información del negocio se analizan los datos que este arroja, se comienzan a levantar los términos, en el escenario seleccionado para la situación en el caso del sistema se selecciona el *Escenario 1: el cual se basa en el desarrollo de ontologías partiendo de necesidades de información de la organización*. Se identifican las relaciones y restricciones de los términos del negocio. Es importante señalar que en esta fase se define el dominio y alcance que tendrá la ontología. Se valora además la posible reutilización de ontologías o cualquier otra forma de visualización y representación de la información o los datos de la organización que se estudia.

### 2.2.1 Preguntas de Competencia.

En esta fase de la metodología se define el dominio y alcance que tendrá la ontología propuesta a través de las preguntas de competencia.

Las Preguntas de Competencia permiten explicitar lo que la ontología debe describir y sus respuestas deben ser formuladas con términos de la propia 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. También posibilita 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 cuándo esté completa.

La siguiente ontología deberá proveer respuestas a preguntas tales como:

- ¿Cuáles son las categorías tanto docente como científica de un determinado Especialista<sup>3</sup>?

---

<sup>3</sup> Se refiere a los profesores que trabajan en el CDAE, al existir varias denominaciones para ellos se estandariza el término finalmente como *Especialista* del Centro.



- ¿Cuáles son las tareas de un determinado Especialista, así como los resultados de estas?
- ¿Qué asignaturas ha impartido un especialista del CDAE?
- ¿Qué idiomas domina un determinado Especialista?
- ¿En qué eventos ha participado un determinado Especialista?
- ¿Qué cursos ha recibido o impartido un determinado Especialista?
- ¿Qué materiales pertenecen a un curso determinado?
- ¿Qué artículos han sido realizados por determinada Línea de Investigación y que Especialista los ha escrito, así como cuáles de ellos han sido publicados?
- ¿Qué proyectos integran una determinada Línea de Investigación?
- ¿A qué proyecto y Línea de Investigación pertenece un determinado Especialista?
- ¿A qué Línea de Investigación y proyecto pertenece una tesis, que Especialista la tutora y quiénes son sus autores?
- ¿Qué Línea de Investigación, proyectos y Especialista integran un tema de investigación?
- ¿Qué servicios se ofrecen en el CDAE?
- ¿Cuáles eventos se han realizado en determinada región o país?
- ¿Qué tipos de Especialistas (Roles) contiene un determinado proyecto según su tipo?
- ¿Qué habilidades debe tener un Especialista para desempeñar un rol determinado?
- ¿Cómo ha sido el desempeño de determinado grupo o Línea de Investigación en una etapa dada?

### 2.2.2 Reutilización de Ontologías.

Durante el desarrollo de la ontología es importante analizar la posibilidad de reutilización de ontologías o cualquier otra fuente de información existente en el negocio. En el caso específico del negocio del sistema actual, no existían modelos de datos previos o sistemas de gestión de información.

Se aplicaron diferentes técnicas para la captura de información:

- Análisis documental: Las principales fuentes de datos existentes en el contexto del negocio estudiado fueron básicamente documental. En este sentido se hizo necesario hacer un análisis.
- Entrevistas: El conocimiento del negocio estaba concentrado en algunos de los especialistas del centro. En este sentido fue necesario realizar una serie de entrevistas a estos especialistas con el

fin de capturar la mayor cantidad de información posible asociado a los procesos que tienen lugar en dicha área del centro y la información que en los mismos se gestiona. A medida que se avanzaba en el diseño del modelo ontológico los prototipos obtenidos fueron sometidos a la verificación de estos especialistas para su aceptación final.

### 2.2.3 Términos Relevantes

Los términos relevantes son los elementos que se destacan en el levantamiento de la información. 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 (Clase o Entidad) relación de clases o atributo.

No.	Término	Persistencia
T1	Acreditación de Conocimientos	N
T2	Acreditación de Asignaturas	N
T3	Artículos	S
T4	Banco de Problemas	S
T 5	Bitácora de Resultados	N
T 6	Boletín de Ciencia	N
T 7	Comité de tesis	S
T 8	Categoría Docente	S
T 9	Categoría Científica	S
T 10	Cronograma del proyecto (Elemento de Chequeo)	N
T 11	Cronograma de Tesis	N
T 12	Control de las Categorías de los Profesionales	N
T 13	Cursos	S
T 14	Currículo	S
T 15	Diseño teórico de los proyectos	N

## Capítulo 2: Diseño de la Solución

---

T 16	Doctorados	S
T 17	Expediente de Investigación de la Línea de Consultoría	N
T 18	Estrategia de Investigación	N
T 20	Estudiantes	S
T 22	Especialistas	S
T 23	Fuentes de Publicación	S
T 24	Indicadores de Ciencia	S
T 25	Libros	S
T 26	Líneas de Investigación	S
T 27	Maestrías	S
T 28	Marco teórico del proyecto	N
T 29	Materiales Didácticos	S
T 30	Perfil de Tesis	S
T 31	Plan de Formación del Centro	S
T 32	Plan de Desarrollo Investigativo Individual	S
T 33	Procesos de Auditoría	N
T 34	Presentaciones en Eventos	S
T 35	Proyectos de Investigación	S
T 36	Proyectos Productivos	S
T 37	Publicaciones	S
T 38	Registro de Eventos	S
T 39	Servicio	S
T40	Tesis	S
T 41	Tesistas	S
T42	Tribunal de Tesis	S
T43	Vinculación a la Docencia	N

T44	Wiki	N
T45	servicios	S
T46	Actividad	S
T47	Indicador	S
T48	Región	S
T49	País	S
T50	Tarea	S

**Tabla 2.1** Términos relevantes del negocio

## 2.2.4 Especificación de la persistencia de los términos relevantes

De los términos relevantes del negocio se extraen los persistentes y se especifica el tipo de su persistencia (Entidad, Atributo, Relación), así como una breve descripción de estos.

No.	Concepto	Descripción	Tipo
C1	Artículos	Se refiere a los artículos de investigación que realizan los trabajadores del CDAE.	Entidad
C2	Asignatura Impartida	Se refiere a las asignaturas que los especialistas afiliados al CDAE imparten.	Entidad
C3	Categoría Docente	Se refiere al nivel docente con que cuentan los especialistas del CDAE.	Entidad
C4	Categoría Científica	Se refiere al nivel científico con que cuentan los especialistas del CDAE.	Entidad
C5	Comité de tesis	Tiene como objetivo que a los temas de interés de centro y los problemas identificados en el banco de	Relación

## *Capítulo 2: Diseño de la Solución*

		problemas del Centro se les brinde solución mediante las tesis. Aprueba o desaprueba los perfiles de tesis	
C 6	Cursos	Proceso mediante el cual un(os) especialistas imparten un determinado conocimiento.	Entidad
C 7	Currículo	Conjunto de información que avala ciertas competencias de un especialista.	Relación y atributos
C 8	Doctorados	Tipo de curso de postgrado.	Atributo
C 9	Tesista	Estudiante afiliado al CDAE por medio de la autoría de una tesis	Entidad
C 10	Especialistas	Trabajadores del CDAE.	Entidad
C 11	Evento	Se refiere a actividad científica donde se presentan un conjunto de trabajos científicos se realizan intercambios de experiencias etc.	Entidad
C 12	Indicadores	Elementos que reflejan estado o característica determinada.	Entidad
C 13	Idioma	Se refiere a los diferentes idiomas que un determinado especialista puede dominar.	Entidad
C 14	Líneas de Investigación	Grupo de especialistas enfocados en el trabajo con algún tipo de tecnología.	Entidad
C 15	Maestrías	Tipo de curso de postgrado.	Atributo
C 16	Materiales	Artefactos de valor para la enseñanza de un contenido específico.	Entidad
C 17	Publicación	Artículo de autoría de algún especialista del CDAE que ha sido publicado.	atributo
C 18	Problemas	Circunstancia que dificulta la consecución de algún fin productivo del CDAE.	Entidad
C 19	Premio	Sacado de las planillas(Pendiente a análisis)	Entidad

C 20	Plan de Desarrollo Investigativo	Conjunto de tareas planificadas por un especialista para llevar a cabo determinados objetivos de trabajo.	Relación
C 22	Proyectos	Grupo de recursos (tanto RH como tecnológicos) destinados y guiados a alcanzar determinado objetivo.	Entidad
C 23	Tesis	Trabajo investigativo que busca resolver una problemática determinada.	Entidad
C24	Tribunal	Conjunto de especialista encargado de evaluar el avance de los trabajos de tesis.	Relación
C25	Servicios	Se refiere a los servicios que se ofrecen por el CDAE tales como Consultoría Tecnológica, Solución Informática etc.	Entidad
C26	Rol	Labor que desempeña un especialista en su área de trabajo ej. Programador, Diseñador etc.	Entidad
C27	Habilidades	Competencia que los especialistas desarrollan producto de la superación o la experiencia de trabajo.	Entidad
C28	Actividad	Se refiere acciones desde el punto de vista científico pero que no son medibles tales como Tribunal de Evento, Tribunal de Tesis etc.	Entidad
C30	Indicador	Elementos medibles que indican el desempeño de determinado grupo de trabajo o Línea de investigación	Entidad
C31	Región	Área geográfica perteneciente a un país donde se realizan determinados eventos.	Entidad
C32	País	País donde se encuentra la región donde se realizan los eventos.	Entidad

**Tabla 2.2** Conceptos relevantes del negocio

### 2.2.5 Definición de restricciones de rol.

Las clases aisladas no proveerán suficiente información para responder las preguntas de competencia. Una vez que se ha definido algunas de las clases, se debe describir la estructura interna de los conceptos. Las propiedades son los elementos ontológicos que permiten establecer las relaciones, además brindan una descripción detallada de las características de los conceptos (22).

En general, hay varios tipos de propiedades de objeto que pueden llegar a ser slots en una ontología(22):

- Propiedades “intrínsecas” tales como Formación académica de un especialista.
- Propiedades “extrínsecas” tales como los servicios que brindan las Líneas.
- Partes, si el objeto es estructurado; pueden ser “partes” físicas y abstractas (ej., el plan de actividades de un especialista).
- Relaciones con otros individuos; estas son las relaciones entre miembros individuales de una clase y otros ítems (ej. el rol de un especialista representa el papel de este en un proyecto determinado de acuerdo a sus habilidades).

También existen las facetas de los slots (restricciones de rol) o valores permitidos de una propiedad que especifican y restringen el rango de los valores de una propiedad. Por ejemplo, el nombre de un especialista es “String”, por lo tanto si se le pone como nombre “1234” a un especialista la aplicación lanzaría un error porque “1234” no es un valor permitido para la propiedad “Nombre” de la clase “Especialista”. También sucede así con la cardinalidad, un “Especialista” solo tiene un nombre, lo que significa que el concepto “Especialista” solo admite una cadena de caracteres para representar la característica “Nombre” (Ej. “Pedro Pérez”, “Pedro José Pérez” etc.)

En el caso de la ontología desarrollada, cada relación descrita anteriormente en la especificación de las entidades, constituye una propiedad que relaciona dos conceptos. Existe la propiedad “tieneCatCientifica” que relaciona los conceptos “Especialista” y “Cat\_Cientifica”, de esta manera se representa en el modelo ontológico la categoría científica de un especialista pero hay que especificar en la ontología que esta propiedad va a relacionar estos conceptos y no otros; ya que sería absurdo, por ejemplo, relacionar a

“Actividad” con “Cat\_Cientifica” porque el concepto “Actividad” no cuenta con esta relación con “Cat\_Cientifica”.

La manera de especificar los conceptos que relaciona una propiedad dada es mediante la definición del rango y el dominio de la misma. El **dominio** es el conjunto de clases que tendrán la propiedad (Ej. para el caso anterior “Especialista” pues es la clase que tiene inicia la relación lógica: “Un *especialista* tiene una categoría científica”); mientras que el **rango** es la clase que se implica en la relación (Ej. para el caso anterior “Cat\_Cientifica” pues es la clase es implicada en la relación lógica: “Un *especialista* tiene una *categoría científica*”).

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. Siguiendo el ejemplo anterior la relación entre el especialista y la categoría puede tomarse desde el punto de vista de la categoría en sí, o sea, que si existe la propiedad “tieneCatCientifica” que relaciona a “Especialista” y “Cat\_Cientifica” como dominio y rango respectivamente de la misma; entonces existe una que relacione a la categoría científica con el especialista, por tanto existe la propiedad “esCatCientificaDe” que tiene como dominio y rango “Cat\_Cientifica” y “Especialista” respectivamente. A este tipo de relación se le llama *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) ya que 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 “Nombre” del “Especialista” es un string o cadena no sería lógico crear un concepto de la ontología llamado “Cadena” o “String”; es por esto que nombre no es una Object Property sino una DataType Property.

Crear una instancia en la ontología no es más que escoger un concepto o clase determinado y darle valores a sus propiedades (Ej. una instancia para la clase Especialista seria):

<b>Entidad:</b>	<b>Especialista</b>
<b>Relaciones:</b>	Cada una de las relaciones es una instancia de la clase implicada ya que son propiedades de objetos.



Campos	Valor
Id	1
Nombre	Yoisy Pérez
Teléfono	2721
Correo	yoisy@uci.cu

Tabla 2.27 Ejemplo de Instancia de Especialista.

## 2.2.2 Vistas del Modelo Ontológico.

Las reglas son una parte importante de las ontologías. El Protégé brinda una interfaz (Ver Figura 2.32) en la cual se ofrecen un conjunto de elementos para hacer más fácil la edición de reglas en el lenguaje Semantic Web Rule Lenguaje (SWRL).

A continuación se muestran algunos ejemplos de las reglas que fueron definidas en el modelo ontológico, expresadas tanto en lenguaje natural como su equivalente en SWRL:

Lenguaje Natural	SWRL
Si (e es un Especialista que imparte la asignatura a y a es "Practica Profesional") entonces el especialista e tiene experiencia en "Practica Profesional" y está capacitado para ser analista.	$\text{imparteAsignatura}(?e, ?a) \wedge$ $\text{asignatura}(?a, \text{"Practica Profesional"})$ $\rightarrow \text{expPractica\_Profesional}(?e, \text{true}) \wedge$ $\text{Analista}(?e, \text{true})$
Si (e es un <i>Especialista</i> que desempeña el rol r y r es " <i>Lider de Proyecto</i> ") entonces el rol actual del especialista e es " <i>Lider de Proyecto</i> " y está capacitado para ser Líder de Proyecto.	$\text{desempennaRol}(?e, ?r) \wedge \text{rol}(?r, \text{"Lider de Proyecto"})$ $\rightarrow \text{rolActual}(?e, \text{"Lider de Proyecto"})$ $\wedge \text{Lider\_de\_Proyecto}(?e, \text{true})$
Si (l es <i>Linea_Investigacion</i> que tiene	$\text{tieneTemalInvestigativo}(?l, ?t) \wedge$

<i>Tema_Investigacion</i> T y t es Tesis de Linea_Investigacion l ) entonces el tema de Investigación de la tesis es T.	esTesisDeLinea(?t, ?l) ∧ nombre(?t, ?n) → sobreTema(?t, ?n)
---	---

**Tabla 2.28** Ejemplo de Reglas en SWRL.

En total se definieron 28 reglas para el modelo ontológico propuesto.

## 2.2. Conclusiones Parciales

El desarrollo del presente capítulo permitió arribar a las siguientes conclusiones:

- El diseño de la ontología propuesta se obtuvo siguiendo las fases que componen la metodología propuesta por el CDAE para la construcción de ontologías. En este capítulo se llegó hasta la fase de diseño, de dicha metodología.
- El alcance de la ontología fue definido a través de 17 preguntas de competencias. Fue necesario aplicar diferentes técnicas de captura de la información del negocio entre las que se encuentran: análisis documental y la entrevista.
- El modelo ontológico propuesto está compuesto por 48 conceptos, 135 relaciones y 106 propiedades que se dividen en propiedades de tipo de datos y propiedades de concepto, que comprenden la información modelada del Área de Investigación y Postgrado del CDAE.
- Se definieron un total de 28 reglas en el diseño de la ontología.

# CAPÍTULO 3: IMPLEMENTACIÓN Y PRUEBA

En este capítulo se lleva a cabo la implementación del SBC y sus respectivas pruebas. En este sentido se crean las instancias de las clases definidas en el modelo ontológico, se especifican algunos de los patrones arquitectónicos utilizados en la implementación del sistema y se implementa un SBC como solución a la situación problemática dada en el Área de Investigación y Postgrado del CDAE. Se realiza además una descripción de la arquitectura del sistema propuesto. Se diseñan además los casos de prueba en correspondencia a las preguntas de competencia que se definieron en la fase de análisis de datos, con el fin de validar el sistema propuesto.

## 3.1 Creación de las instancias

Las instancias son una parte importante del sistema basado en conocimiento ya que mediante ellas se realiza gran parte del proceso de razonamiento, además de que muestran en la práctica, la funcionalidad del sistema. El término **Razonamiento basado en instancias** hace referencia a una especialización del proceso de razonamiento. Este tipo de razonamiento se basa en llegar al caso que es conocido y se encuentra en la base de conocimiento del sistema. Para ello se utilizan un gran número de instancias para dirigir el razonamiento. La representación de estas instancias es normalmente en forma de vectores. En el caso específico de este sistema se construyeron 5 instancias de cada tipo de entidad para comprobar las funcionalidades del sistema.

## 3.2. Arquitectura del Sistema

Como se detalló en el capítulo 1, el sistema está conformado por subsistemas conformados según las responsabilidades de las clases que los conforman. El subsistema de acceso a datos está compuesto por la clase `.java OntLoader`, la cual tiene la responsabilidad de establecer la conexión con la base de datos y de crear el modelo persistente. Implementa además, el patrón `Data Access Object`, ya que aísla al resto de la aplicación del origen de los datos.

El otro subsistema es el de la lógica del negocio, el cual está compuesto por 3 paquetes: `Inference_Subsystem`, `Inference_Subsystem.impl` y `Ontology_Model`. En el primero se agrupa el conjunto

de interfaces que permitirá manejar el modelo ontológico desde un lenguaje de alto nivel como Java. En el segundo está la implementación de las interfaces antes mencionadas, que van a permitir un conjunto de funcionalidades para la manipulación de objetos relacionados con el modelo ontológico y en el tercero están declarados los principales elementos del modelo ontológico, este constituye una muy útil herramienta a la hora de desarrollar aplicaciones basadas en ontologías. Constituyen las declaraciones de los elementos de la ontología previamente modelada en Protégé esto se ilustra en la Figura 3.1.

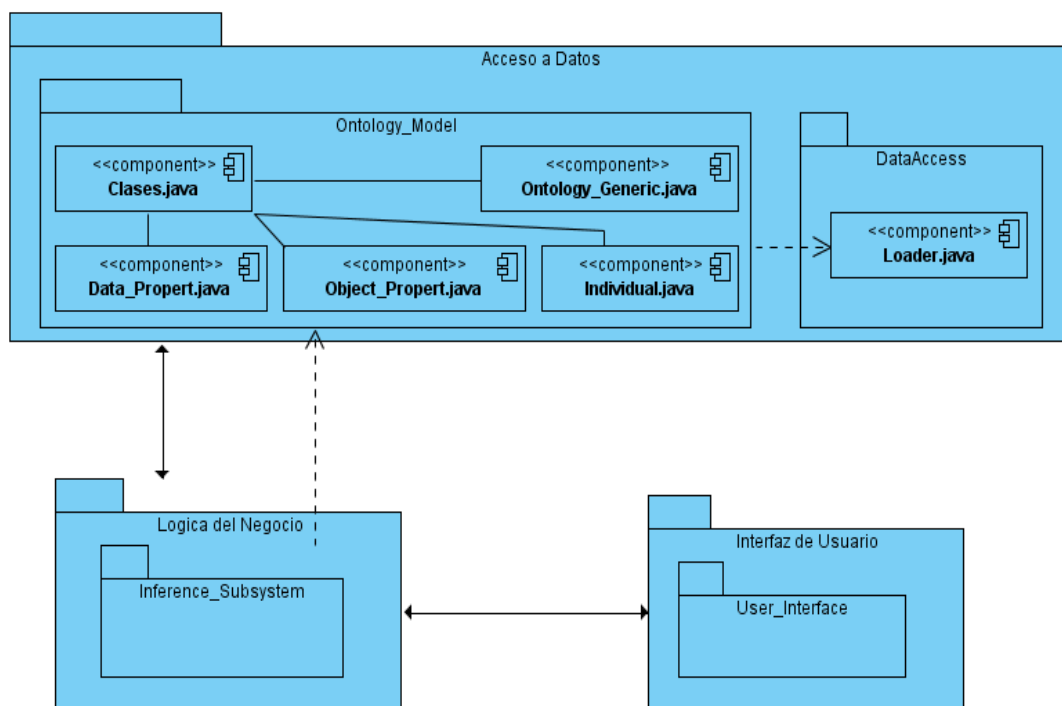


Figura. 3.1 Arquitectura del Sistema.

### 3.3. Patrones Utilizados en la implementación

Los **patrones de diseño** son un conjunto de soluciones a problemas que generalmente se encuentran en el diseño de un programa. Cada patrón explica cómo resolver un determinado problema, bajo determinadas circunstancias y las ventajas y desventajas de su uso.

#### **Data Access Object:**

La utilización del patrón Data Access Object (DAO) permite abstraer y encapsular todos los accesos a la fuente de datos. El DAO maneja la conexión con la fuente de datos para obtener y almacenar datos. El DAO implementa el mecanismo de acceso requerido para trabajar con la fuente de datos. Esta fuente de datos puede ser un almacenamiento persistente como una Base de Datos, un servicio externo como un intercambio B2B o un repositorio LDAP.

El DAO oculta completamente los detalles de implementación de la fuente de datos a sus clientes. Como la interface del DAO no cambia cuando cambia la implementación de la fuente de datos subyacente, este patrón permite al DAO adaptarse a diferentes esquemas de almacenamiento sin que esto afecte a sus clientes o componentes de negocio. Esencialmente, el DAO actúa como un adaptador entre el componente y la fuente de datos.

En el caso específico del sistema se implementa este patrón mediante la Clase OntLoader la cual tiene la responsabilidad de manejar la conexión con el almacenamiento persistente.

#### **Adapter:**

Convierte la interfaz de una clase en otra distinta que es la que esperan los clientes. Permiten que cooperen clases mediante la creación o adaptación de un tipo de dato, que sirve como protocolo de comunicación entre dos funcionalidades. Este patrón es muy utilizado cuando se está trabajando con comparaciones entre diferentes valores de instancias diferentes pues los métodos que brinda el framework para manipular los objetos desde java entregan diversas interfaces.

#### **Objeto de Valores:**

Básicamente consiste en **agrupar** varios valores dentro de un objeto para enviarlo y recibirlo con mayor comodidad/seguridad de esta forma el código es más legible y escalable. Es utilizado cuando la información de entrada o salida de algún método está compuesta por varios valores entonces se crea un objeto compuesto por estos.

En el caso específico del sistema se implementa este patrón con la existencia de las clases “RepIndividual”, “IndivRol”, “IndivPro” etc. Cuyo principal función es la de agrupar un conjunto de datos ya sea para mostrarlos al usuario o para que sirvan de entrada a algunos métodos.

### **Alta Cohesión:**

Este patrón permite que las clases del diseño realicen las funcionalidades necesarias para cumplir con su responsabilidad. Mejora la claridad y facilidad para entender el diseño. Busca soluciones para asignar los métodos a las clases de forma coherente, completa y relacionada, permitiendo el cambio, poniendo toda la información que se necesita controlar a la vista en el mismo fichero. Fomenta la reutilización.

### **Experto:**

Este patrón es muy utilizado para asignar responsabilidades. Una clase contiene la información necesaria para llevar a cabo sus funcionalidades. Permite conservar el encapsulamiento, debido a que los objetos se valen de su propia información para cumplir lo que se le pide. Proporciona que las clases cuenten con la funcionalidad requerida, brindando así una alta cohesión.

### **Creador:**

Este patrón establece la creación de instancias entre clases. Brinda un soporte a bajo acoplamiento. Guía la asignación de responsabilidades relacionada con la creación de objetos, tarea frecuente en la programación orientada a objeto.

### **Bajo acoplamiento:**

Es imprescindible utilizarlo cuando se vaya a diseñar. Permite la reutilización y el diseño de clases más independientes, minimiza el impacto de los cambios.

### **3.4. Almacenamiento Persistente**

Programa de base de datos: Es una herramienta de software para organizar el almacenamiento y la recuperación de esa información.

En el caso específico de este sistema se usa la base de datos del Framework Jena, el cual es brindado por el plugin “ProtégéToJena” el cual permite exportar el modelo Ontológico a formato persistente. Contar con la información de forma persistente es sumamente importante para los SBC ya que de esta manera se puede tener acceso a información anterior y de esta manera se pueden establecer correlaciones entre situaciones que han pasado para ayudar en el proceso de toma de decisiones en una situación concreta del presente. A continuación se muestra una vista del diseño relacional de clases persistentes de la Base de Datos de Jena.

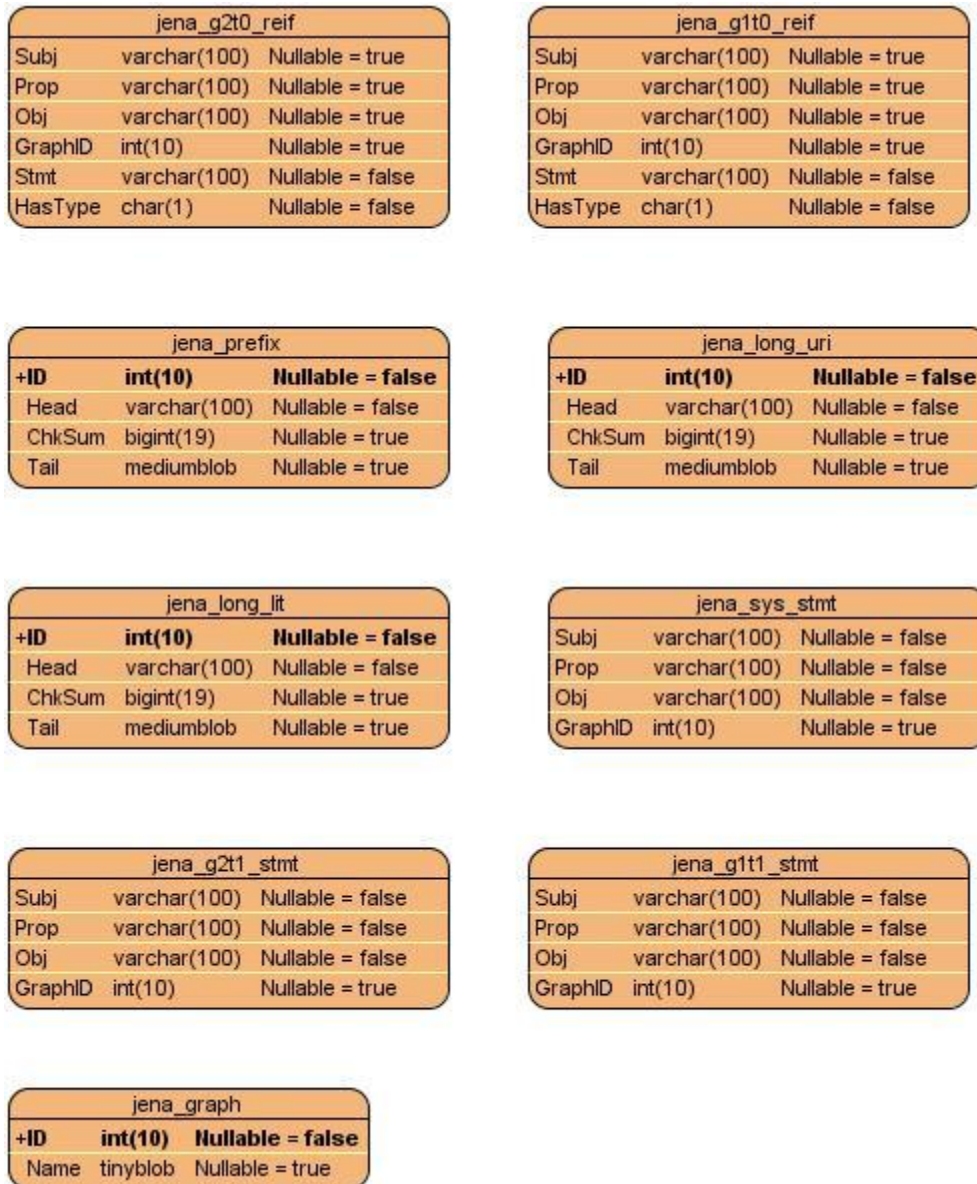


Figura. 3.5 Diseño ORM de la BD del Framework Jena



### **3.5. Pruebas**

Las pruebas de software constituyen un instrumento para determinar el nivel de la calidad de un producto. En este proceso se ejecutan pruebas dirigidas a componentes del software o al sistema de software en su totalidad, con el objetivo de medir el grado en que el software cumple con los requerimientos. En las pruebas se usan casos de prueba, especificados de forma estructurada mediante técnicas de prueba (41).

#### **3.5.1. Pruebas de Caja Negra**

Las pruebas de caja negra se centran en lo que se espera de un módulo, es decir, intentan encontrar casos en que el módulo no se atiene a su especificación. Por ello se denominan pruebas funcionales y el probador se limita a suministrarle datos como entrada y estudiar la salida, sin preocuparse de lo que pueda estar haciendo el módulo por dentro. Los Casos de Prueba son un producto de desarrollo de software que ayudan a validar y verificar las expectativas de los clientes. El propósito de un Caso de Prueba es especificar una forma de probar el sistema, incluyendo las entradas con las que se ha de probar, los resultados esperados y las condiciones bajo las que ha de probarse (42).

En el caso específico del sistema se realizaron un total de 30 pruebas a través de los casos de prueba que describen a continuación posibilitando encontrar algunos errores lógicos introducidos en la programación del subsistema de lógica del negocio.

#### **3.5.2. Descripción del Caso de Prueba**

A continuación primeramente se muestran las Figuras que muestran las relaciones en las que se sustentan la lógica del negocio que es utilizada para dar solución a los requerimientos funcionales que se presentan en forma de casos de pruebas luego se visualizan unas tablas en las que se describe en sí cada caso de prueba.

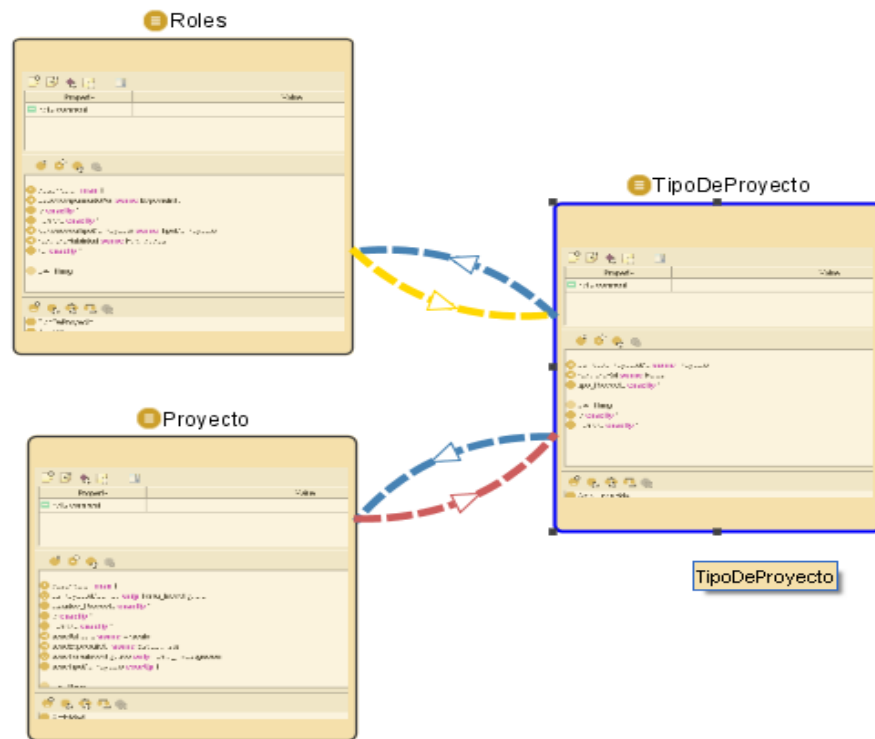


Figura. 3.6 Relaciones entre los conceptos que soportan el requisito funcional.

### Caso de Prueba Mostrar Roles para Tipo de Proyecto

Funcionalidad: Retorna un arreglo de string que indican los roles necesarios en un proyecto de determinado tipo.		
Entradas	Resultado	Condiciones
Al escoger el administrador la opción mostrar rol para proyecto	El Sistema Muestra los tipos de proyectos válidos.	
El administrador introduce un tipo de proyecto	El Sistema muestra los roles que debe llevar un proyecto	El tipo de proyecto debe ser válido.

	de ese tipo.	
El administrador le entra el tipo de proyecto <b>“Desarrollo”</b>	<p>Escribiendo Roles para el Tipo de Proyecto : Desarrollo</p> <pre> -----&gt; Consultor -----&gt; AnalistaDelNegosio -----&gt; ArquitectoDelNegosio -----&gt; ArquitectoSOA -----&gt; DiseñadordeServicios -----&gt; ImplementadorDeServicios -----&gt; ImplementadorDeSolucion -----&gt; Probador -----&gt; LiderDeProyecto -----&gt; AseguradorDeLaCalidad -----&gt; Planificador </pre>	

**Tabla 3.1** Caso de Prueba Mostrar Roles para Tipo de Proyecto

### Caso de Prueba Mostrar Especialistas para Rol

<b>Funcionalidad:</b> Retorna un arreglo de individuales que constituyen los especialistas capacitados para desempeñar el rol pasado por parámetro.		
Entradas	Resultado	Condiciones
Al escoger el administrador la opción	El Sistema Muestra los roles válidos.	

mostrar especialistas para rol.		
El administrador introduce un rol	El Sistema muestra los especialistas capacitados para desempeñar ese rol	El rol debe ser válido.
El administrador le entra el rol <b>“Consultor”</b>	<p>Imprimiendo los especialista propuestos para el desempeño del rol : Consultor</p> <p>----&gt; Juan Manuel Mederos --- Este Especialista Tiene Experiencia Profesional Docente</p> <p>----&gt; Maypher Roman Duran --- Este Especialista Tiene Experiencia en el Rol : Consultor</p> <p>----&gt; Maypher Roman Duran --- Este Especialista Tiene Experiencia Profesional Docente</p>	

**Tabla 3.2** Caso de Prueba Mostrar Especialistas para Rol

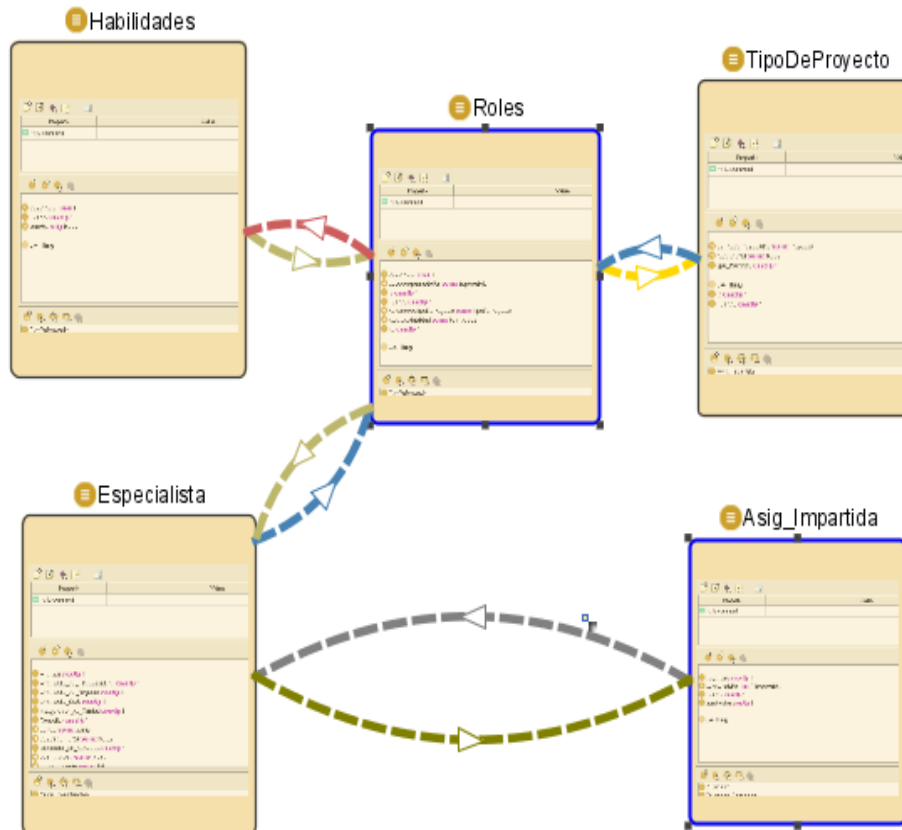


Figura. 3.7 Relaciones entre los conceptos que soportan el requisito funcional.

**Caso de Prueba Mostrar Especialistas para Proyecto de tipo determinado**

**Funcionalidad:** Retorna un arreglo de individuales que constituyen los especialistas capacitados para ingresar en un proyecto de tipo determinado.

Entradas	Resultado	Condiciones
----------	-----------	-------------

Al escoger el administrador la opción mostrar especialistas para proyecto de tipo.	El Sistema Muestra los tipos de proyectos validos	
El administrador introduce un tipo de proyecto	El Sistema muestra los especialistas capacitados para ingresar en un proyecto de ese tipo.	El tipo debe ser válido.
El administrador le entra el tipo "Formacion_de_RRH H"	<pre> Imprimiendo los Especialista Propuestos para Participar en el Proyecto de Tipo : Formacion -----&gt; Pedro Luis Perez --- Tiene Experiencia previa en Proyectos de Formacion_de_RRH                     </pre>	

**Tabla 3.3** Caso de Prueba Mostrar Especialistas para Proyecto de tipo determinado

### Caso de Prueba Mostrar Documentos para Proyecto determinado

Funcionalidad: Retorna un arreglo de individuales que constituyen los documentos que tratan el tema del proyecto en cuestión.		
Entradas	Resultado	Condiciones
Al escoger el administrador la opción mostrar documentos para proyecto	El Sistema Muestra los proyectos	

## Capítulo 3: Implementación y Pruebas

<p>El administrador introduce un proyecto</p>	<p>El Sistema muestra los documentos asociados al tema del proyecto en cuestión.</p>	<p>El proyecto debe estar registrado.</p>
<p>El administrador le entra el proyecto <b>“Sistema para la Automatización del Control Aéreo de Cuba (SACAC)”</b></p>	<p>Imprimiendo los Documento que pueden ser consultado por los especialista que comiencan en un nuevo Pa</p> <p>Tesis del tema de investigacion asociado al proyecto : Sistema para la Automatizacion del Control Aeo</p> <p>----&gt; Tesis - Privacidad y Nuevas Tecnologías.</p> <p>Articulos del tema de investigacion asociado al proyecto : Sistema para la Automatizacion del Control</p> <p>----&gt; Apple explica su política con respecto a Flash, Adobe responde</p>	

**Tabla 3.4** Caso de Prueba Mostrar Documentos para Proyecto determinado

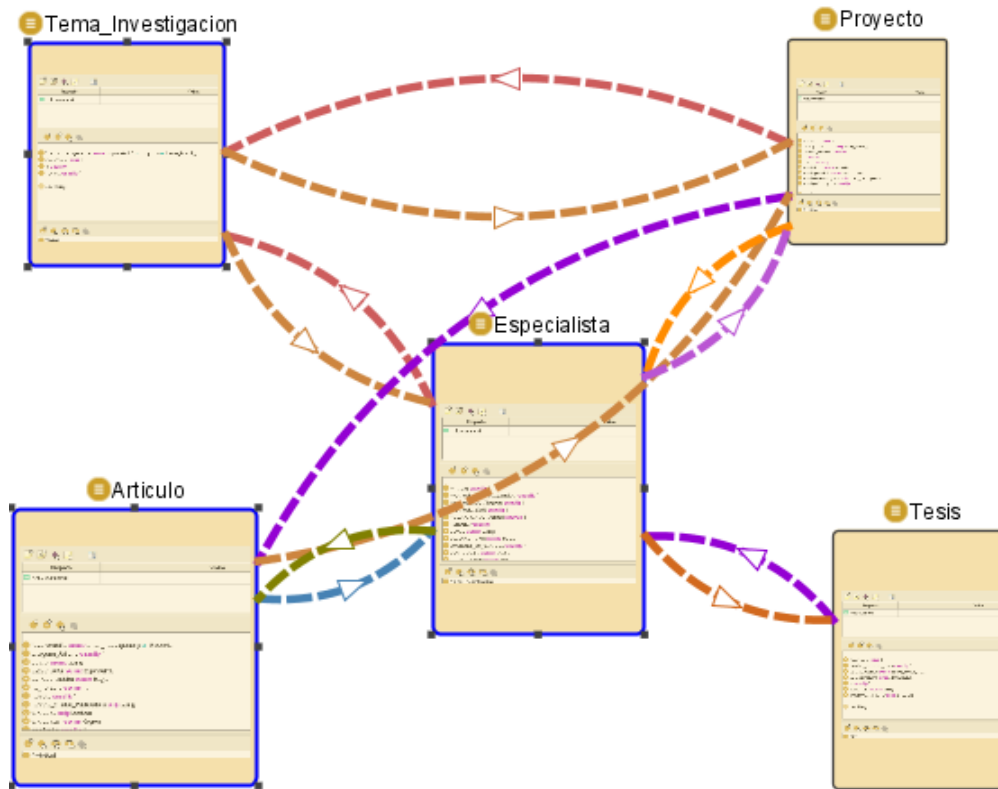


Figura. 3.8 Relaciones entre los conceptos que soportan el requisito funcional.

### Caso de Prueba Mostrar Necesidad de formación para un especialista

<b>Funcionalidad:</b> Retorna un arreglo de individuales que constituyen las líneas de investigación que el especialista no tiene experiencia.		
<b>Entradas</b>	<b>Resultado</b>	<b>Condiciones</b>
Al escoger el administrador la opción mostrar necesidad de formación	El Sistema Muestra los especialistas	



El administrador introduce un especialista	El Sistema muestra las líneas que el especialista no tiene experiencia.	El especialista debe estar registrado.
El administrador le entra el especialista " Yoisy Pérez Olmo	Necesidades de Formación para el especialista : Yoisy Perez Olmo  ----> Problemas Teóricos del Diseño  ----> Modelación de Sistemas y Procesos  ----> LGAC diseño y gestión  ----> SOA_BPM	

**Tabla 3.5** Caso de Prueba Mostrar Necesidad de formación para un especialista

### 3.6. Conclusiones Parciales

El desarrollo del presente capítulo permitió:

- Implementar el sistema basado en conocimientos, incluyéndose el empleo de las buenas prácticas establecidas en los patrones de diseño y la definición del almacenamiento persistente de los datos que almacenarán la información del sistema.
- Probar la solución obtenida a partir de un conjunto de casos de pruebas que fueron definidos y finalmente probadas mediante la técnica de pruebas de caja negra.

## **CONCLUSIONES**

- Se realizó una evaluación del estado del arte acerca de la gestión del conocimiento, sus técnicas y herramientas y los sistemas de inferencia, que permitió la selección de las Ontologías como técnica de representación de conocimiento, la metodología elaborada en el CDAE para su construcción, OWL-RDF como lenguaje de representación de la Ontología y Protégé como herramienta de modelado, además del framework Jena para la implementación del Sistema.
- Se obtuvo una Ontología que modela la información de los procesos involucrados en el Área de Investigación y postgrado del Centro de Consultoría y Desarrollo de Arquitecturas Empresariales de la UCI.
- Se desarrolló un SBC que permite gestionar el conocimiento a partir del modelo ontológico del Área de Investigación y postgrado del CDAE, favoreciendo la toma de decisiones de los directivos del centro.
- Se ejecutaron diferentes casos de pruebas de caja negra en el SBC para la información representada, que permitieron validar satisfactoriamente la calidad del mismo.

- Aplicar el SBC obtenido para gestionar el conocimiento y tomar decisiones en el área de investigación y postgrado del CDAE
- Hacer extensible el modelo ontológico y el SBC obtenido a las demás áreas del CDAE para lograr una mayor eficiencia en la gestión del conocimiento del centro.
- Implementar el Subsistema de Interfaz de usuario propuesto en el diseño arquitectónico del SBC desarrollado.

### Referencias bibliográficas

1. **Álvaro Palma Quiroz, Emilio Rodríguez Ponce.** GESTIÓN DEL CONOCIMIENTO Y GESTIÓN DE CALIDAD EN LA EDUCACIÓN SUPERIOR. 2008, Vol. II, 0718-3429.
2. *Instrumentos de reprecentacion del conocimiento: Tesoros VS Ontologias.* **Jiménez, Antonio Garcia.** Murcia : s.n., 2004.
3. **Keilyn Rodríguez Perojo, Rodrigo Ronda León.***Web semántica: un nuevo enfoque para la organización y recuperación de información en el web.* Ciudad de La Habana : s.n., 2005.
4. **Esther Sánchez Lucas, Rosa Martinez Rubio.***LAS ONTOLOGÍAS Y SU APLICACIÓN EN EL ÁMBITO DE LA DOCUMENTACIÓN.*
5. **Rodríguez-Ponce, Emilio R.***GESTIÓN DEL CONOCIMIENTO Y EFICACIA DE LAS ORGANIZACIONES: UN ESTUDIO EMPÍRICO EN INSTITUCIONES PÚBLICAS.* Caracas : s.n., 2007.
6. **Canals, Agustí.***La gestión del conocimiento.* Barcelona : s.n., 2003.
7. **José A. Senso, Antonio de la Rosa.***Especificaciones XML aplicadas a la Documentación.* 1999.
8. **Julià Minguillón, Enric Mor, Francesc Santanach, Lourdes Guàrdia Ortiz.***PERSONALIZACIÓN DEL PROCESO DE APRENDIZAJE USANDO LEARNING OBJECTS REUTILIZABLES.* 2005.
9. **Peis Redondo Eduardo, Herrera Viedma Enrique, Hassan Montero Yusef, Herrera Torres Juan Carlos.***Ontologías, Metadatos y Agentes: Recuperación 'Semántica' de la Información.* Granada : s.n., 2003.
10. **Sebastián Amengual Galdón, José María Cubero Muñoz, Enrique Estalayo Gutiérrez,.***Motores de búsqueda para contenidos.* Madrid : s.n., 2006.
11. **López, Francisco Javier Honrubia.***Introducción a las Ontologías.* Albacete : s.n., 2007.
12. **Antonio Martín, Sonsoles Celestino, Adela Valdenebro, Julia Mensaque.***ONTOLOGÍAS E INTELIGENCIA ARTIFICIAL PARA LA RECUPERACIÓN.* Sevilla : s.n., 2008.
13. **Flor Nancy Diaz, Luis Joyanes , Victor Hugo Medina.***Taxonomía, Ontología y Folksonomía: Qué son y qué Beneficios u Oportunidades Presentan para los Usuarios de la Web?* Bogot : s.n., 2009.
14. **María de los A. Martín, Luis Olsina.***Fortalezas y Limitaciones de las Tecnologías Actuales para el Procesamiento Semántico en la Web.* La Pampa : s.n., 2007.
15. **Peñalvo, Francisco José García.***Web Semántica y Ontologías.* Salamanca : s.n., 2008.
16. **Contreras, Jesús.***TUTORIAL ONTOLOGÍAS.* Madrid : s.n., 2006.

17. **Javier ESQUIVIAS, Jose Manuel ZURITA, Mercedes GOMEZ, Eduardo VILLAR, Jose Manuel SANCHEZ, Jose.***Representación del Conocimiento mediante Ontologías : Ensayo sobre.* Granada, España : s.n., 2005.
18. **García Sánchez, Francisco.***Sistema basado en tecnologías del conocimiento para entornos de servicios web semánticos.* Murcia : s.n., 2008.
19. **Erdmann, Michael.***OntoEdit: Collaborative Ontology Development for the Semantic Web.* Karlsruhe, Germany : s.n., 2002.
20. **Gruber, Tomas R.***Ontolingua: A Mechanism to support Portable Ontologies.* Stanford : s.n., 1992.
21. **Cottam., Hugh.***Ontologies to Assist Process Oriented Knowledge Acquisition.* Stanford : s.n., 1999.
22. **ZAPATER, JOSÉ JAVIER SAMPER.***ONTOLOGÍAS PARA SERVICIOS WEB SEMÁNTICOS DE INFORMACIÓN DE TRÁFICO: DESCRIPCIÓN Y HERRAMIENTAS DE EXPLOTACIÓN.* Valencia : s.n., 2005.
23. **Sean M. Falconer, Natalya F. Noy, and Margaret-Anne Storey.***Towards understanding the needs of cognitive support.* Stanford : s.n., 2002.
24. **Arano, Silvia.***Los tesauros y las ontologías en la Biblioteconomía y la Documentación.* España : s.n., 2005.
25. **Hernandez, Anislebis Fernandez.***Las Ontologías, Nuevos Retos.* Ciudad de la Habana : s.n., 2009.
26. **Antonio Cuadra, Francisco Garcés, Marcos Reyes, Damaris Fuentes-Lorenzo, Luis Sánchez, José M. Cantera.***Tecnologías Web Avanzadas para el Aseguramiento de la Calidad de Servicio.* Madrid : s.n., 2005.
27. *ALGUNAS REFLEXIONES SOBRE LA GESTIÓN PARA EL CONOCIMIENTO.* **Reyes, Luis del Toro.** 2008.
28. **Carrillo, Francisco Javier.***Desarrollo Basado en el Conocimiento.* Monterrey : s.n., 2007.
29. **Romero, Juan Gomez.***Knowledge Mobilization Architectures, Models and Applications.* Granada : s.n., 2008.
30. **MacGraw, DW Rolston.***Principios de inteligencia artificial y sistemas expertos.* Ciudad Mejico : s.n., 1990.
31. **Giarratano, Joseph y Riley, Gary.***Sistemas expertos principios y programación.* Ciudad Mejico : s.n., 2001.

32. **Lourdes Sáiz Bárcena, Miguel Ángel Manzanedo del Campo, Ana María Lara Palma.***La Gestión del Conocimiento exige una nueva formulación de las Políticas.* Burgos : s.n., 2004.
33. **Claudio Rancán, Patricia Pesado.***Hacia una Propuesta Integradora de Sistemas Basados en Conocimiento y de Descubrimiento.* Buenos Aires : s.n., 2006.
34. **Secundino López García, Jorge Puente Peinador.***INGENIERÍA DEL CONOCIMIENTO.* Oviedo : s.n., 2002.
35. **J.F. Sigut, J.D. Piñeiro, R.L. Marichal, L. Moreno.***INTEGRACIÓN DE TÉCNICAS DE ANÁLISIS Y CLASIFICACIÓN MEDIANTE UN SISTEMA BASADO EN EL CONOCIMIENTO PARA PROBLEMAS DE DIAGNÓSTICO.* Tenerife : s.n.
36. **Emilio Vivancos, Luis Hernández, Vicente Botti.***Construcción y análisis temporal de sistemas basados en reglas para entornos de tiempo real.* Valencia : s.n.
37. **Luís Gilberto Salinas, Roger Saravia Aramayo.***Desarrollo de un Sistema Experto para la Elección de un Método Ágil.* San Andres : s.n., 2007.
38. **Tim Berners-Lee, James Hendler y Ora Lassila.***La Red semántica.* 2003.
39. **E. Peis, E. Herrera-Viedma, Y. Hassan, J. C. Herrera.***Análisis de la web semántica: estado actual y requisitos futuros.* 2003.
40. **Aylen Martínez Gamboa, Katuska Cabrera Pupo.***PROPUESTA METODOLÓGICA PARA LA GESTIÓN DE CONOCIMIENTO BASADA EN ONTOLOGÍAS.* Ciudad de la Habana : s.n., 2007.
41. **Segurado, Teresa Gómez del Castillo.***UN INSTRUMENTO PARA EVALUAR VALORES EN EL SOFTWARE EDUCATIVO.* Sevilla : s.n.
42. **Mañas, José A.***Prueba de Programas.* 1994.
43. *Los tesauros y las ontologías en la Biblioteconomía y la Documentación.* **Arano, Silvia.** 2005.
44. <http://jipl2gaby.blogspot.com/2009/07/ontologias-rdf-y-owl.html>. <http://jipl2gaby.blogspot.com>. [En línea] septiembre de 2009.

