

## ONTOLOGICAL REPRESENTATION OF CATATUMBO FISH

### REPRESENTACIÓN ONTOLÓGICA DE PECES DEL CATATUMBO

**MSc. Torcoroma Velásquez Pérez, Esp. Andrés Mauricio Puentes Velásquez**  
**Ing. Luis Ernesto López Vargas**

**Universidad Francisco de Paula Santander Ocaña**

Sede Algodonal Ocaña, Norte de Santander, Colombia.

Tel.: (577) - 569 0088, Fax: (577) - 569 0088, Ext. 212.

E-mail: tvelasquezp@ufps.edu.co, ampuentesv@ufps.edu.co, lelopezv22@hotmail.com

**Abstract:** This work develop ontology whose contains the highlights of ichthyology in Catatumbo watershed. All the conceptualization was created using the methodology for development of ontologies named Methontology, which were implement a series of tasks that guided the final work. The next step in development was to encode the model generated in the owl ontological language; for this job was selected the editing tool Protégé. In the knowledge base generated at the ontology, users can find information more structured and formalized related to Catatumbo fish.

**Keywords:** Ontologies, Knowledge representation, OWL, Semantic web, Catatumbo fish

**Resumen:** Este trabajo trata sobre la creación de una ontología que contiene los aspectos más sobresalientes de la ictiología de la cuenca del Catatumbo. La conceptualización se desarrollo usando la metodología de desarrollo de ontologías Methontology, en la cual se realizaron una serie de tareas que guiaron el trabajo final. Se codificó el modelo generado en un lenguaje ontológico OWL; para esto se seleccionó la herramienta de edición Protégé. En la base de conocimiento generada en la ontología, los usuarios encontrarán información mejor estructurada y formalizada relacionada con los peces del Catatumbo.

**Palabras clave:** Ontologías, Representación de Conocimiento, OWL, Web Semántica, Peces del Catatumbo

## 1. INTRODUCCIÓN

El Catatumbo es una región muy rica en diversidad ecológica, en sus aguas se encuentran una gran cantidad de peces de diferentes formas, tamaños, colores, hábitos, que enriquecen en gran medida la ictiofauna de los ríos y quebradas. Esto permite que esta región sea una de las más importantes en la actividad pesquera de Colombia, proporcionando un alto grado de desarrollo económico de la zona.

Es de vital importancia ecológica proteger la biodiversidad de estas regiones apoyándose en herramientas como los Sistemas Basados en

conocimiento que permiten almacenar y recuperar eficientemente la información y el conocimiento asociado a dominios de este tipo por parte de los expertos en dichas labores de protección y conservación. Las ontologías deben considerarse como una solución efectiva a este tipo de problemas por ser uno de los componentes fundamentales dentro de la arquitectura de la web semántica. Por definición, se entiende la web semántica como una extensión de la web actual, dotada de mayor significado, en la que cualquier usuario en internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Para construirla se

necesita poder representar el conocimiento de forma que sea legible por los computadores, esté consensuado, y sea reutilizable.

En este proyecto se usó la metodología de desarrollo *Methontology* que permite la construcción de ontologías a nivel de conocimiento e incluye la identificación del proceso de desarrollo de la ontología, un ciclo de vida basado en la evolución de prototipos y técnicas particulares para realizar cada actividad (Gómez-Pérez, 2004). Además se utilizó el editor de ontologías *Protégé* que es de código abierto y que proporciona una arquitectura extensible para la creación de aplicaciones de bases de conocimiento personalizadas; con esta herramienta se logró codificar la ontología en un lenguaje dotado de una expresividad adecuada como es OWL.

### 1.1 Generalidades de los peces del Catatumbo

La región del Catatumbo toma su nombre gracias al río Catatumbo que nace en el municipio de Abrego, donde se forman los ríos Frío y Oroque que en un transcurso más abajo forman el Río Algodonal. Con el levantamiento de los Andes formados en parte por rocas de origen marino y volcánicas ricas en nutrientes, aparecieron los grandes ríos que drenan esos suelos, estos ríos y sus planicies de inundación, ocasionando, sin duda, el cambio más significativo para los peces de agua dulce del continente. La fertilidad de las aguas puso a su disposición una enorme y variada oferta alimenticia y proporcionó una multitud de nichos y ambientes nuevos que dieron lugar a nuevas especies derivadas de la antigua fauna (Galvis, 1997). En Colombia, el levantamiento de la cordillera oriental y su unión con la sierra de Mérida terminaron aislando las cuencas del Catatumbo, Magdalena y todas las que se formaron posteriormente al occidente de esta cordillera; esta barrera determinó que la cuenca del Orinoco dejara de desembocar junto con el río Magdalena en el golfo de Venezuela y desvió su curso hacia el oriente. Los ríos Magdalena y Catatumbo continuaron compartiendo una cuenca común hasta el final del terciario, cuando el levantamiento de la serranía del Perijá dividió en dos este drenaje común. (Maldonado, 2005).

El Catatumbo es una región muy rica en diversidad ecológica, en donde tanto su flora como su fauna son muy abundantes, razón por la cual es considerada una zona de productividad. Las aguas que recorren el territorio son muy caudalosas y apropiadas para la supervivencia de las especies

acuáticas que viven allí. Estas aguas poseen una gran cantidad de peces de diferentes formas, tamaños, colores, hábitos, que enriquecen en gran medida la ictiofauna de los ríos y quebradas de esta zona (Maldonado, 2005). Toda la variedad de peces que se encuentra en la cuenca del Catatumbo permite que esta región sea una de las más importantes en la actividad pesquera de Colombia.

## 2. METODOLOGÍA PARA EL DESARROLLO DE LA ONTOLOGÍA

Como en cualquier campo de estudio dentro de las ciencias computacionales, para construir un proyecto de software de una manera confiable es aconsejable emplear una adecuada metodología, siguiendo los procesos de desarrollo, las actividades y las técnicas necesarias para su uso. El proceso de construir una ontología lo podemos ver como un proyecto de desarrollo de software de alto nivel, donde confluye la ingeniería del software, la ingeniería ontológica y la representación de conocimiento entre otros.

No se ha reconocido hasta el momento por parte de los expertos en ingeniería ontológica una metodología estándar para el desarrollo de ontologías, puesto que a medida que se difunde su uso, van surgiendo metodologías que identifican como construir las mismas, de acuerdo a cada necesidad y a cada época. Sin embargo, se escogió *Methontology* entre todas las metodologías existentes por las siguientes razones:

- Tiene sus raíces en las actividades identificadas por el proceso de desarrollo de software propuesto por la organización IEEE y en otras metodologías de ingenierías de conocimiento (Gómez-Pérez, 2002).
- La creación de la ontología puede empezar desde cero o con base en la reutilización de otras existentes.
- Incluye la identificación del proceso de desarrollo de la ontología (calendario, control, aseguramiento de calidad, adquisición de conocimiento).
- Es la que propone la descripción más ajustada de cada actividad a realizar.
- Es la metodología recomendada para la construcción de Ontologías por la “Fundación para los Agentes Físicos Inteligentes (FIPA), que promueve la interoperabilidad entre aplicaciones basadas en agentes.
- Propone un ciclo de vida de construcción de la ontología basado en prototipos evolutivos, porque esto permite agregar, cambiar y

remover términos en cada nueva versión (prototipo) (Gómez-Pérez, 2004).

- Provee un marco que consiste de: identificación del proceso de desarrollo de la ontología donde las actividades principales son identificadas; un ciclo de vida en prototipos espirales; y la propia metodología, que especifica los pasos para desarrollar cada actividad, las técnicas utilizadas, los productos de salida y cómo deben ser evaluados.

### 3. RESULTADOS

#### 3.1 Jerarquía taxonómica de los peces del Catatumbo.

Los peces se encuentran ubicados en cuatro órdenes, cada uno de ellos se divide en familias y estas a su vez en géneros que es donde se ubican las especies; se relaciona cada nivel de la jerarquía teniendo en cuenta que una especie pertenece solamente a un género; un género pertenece a una familia y una familia pertenece a un orden, obteniéndose de esta forma un árbol de clasificación de conceptos. En la Figura 1 se presenta un esquema diseñado que contiene las familias, géneros y especies del orden *Siluriformes*.

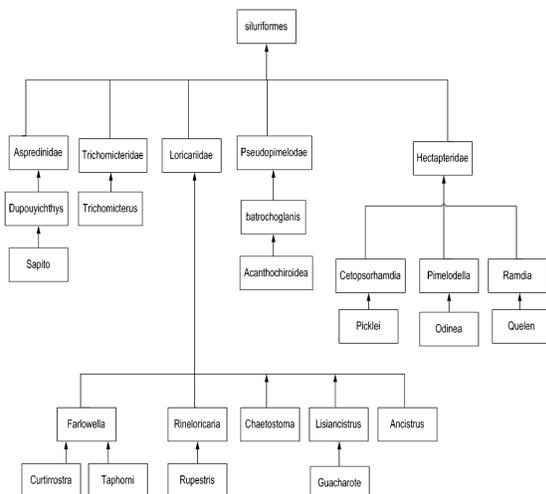


Fig. 1. Taxonomía peces del Catatumbo

**Características de cada pez.** Se realizó un estudio sobre las características generales de los peces para determinar los rasgos distintivos de cada uno de ellos y posteriormente se elaboraron una serie de mapas conceptuales plasmando las características más sobresalientes de cada pez. Las características más importantes que se pueden observar en un pez son Aletas, Arco branquial, Barbicelos, Cabeza, Cuerpo, Dientes, Escamas, Espinas, Hocico,

Mandíbula, Placas y Radios; estas características permiten agrupar una especie en un determinado género.

#### 3.2 Formalización del conocimiento adquirido a través de una técnica de representación del conocimiento

La web semántica se basa principalmente en mecanismos que permiten representar el conocimiento de un modo estandarizado, haciendo posible su tratamiento automático y la existencia por lo tanto, de nuevas aplicaciones que puedan beneficiarse de estas representaciones. La representación del conocimiento como tal, es una materia en la que se lleva trabajando desde hace varias décadas, desde mucho antes de que surgiera la web semántica. Surgió en el ámbito de la Inteligencia Artificial al tratar de crear representaciones de conocimiento que pudieran ser utilizadas por mecanismos que simulen el razonamiento humano (Echarte, 2006).

**Reglas:** esta técnica representa el conocimiento presentando unas premisas o condiciones y las conclusiones o acciones que de ellas se derivan. Se suelen representar de la forma IF – THEN - Las premisas se colocan a continuación del IF en forma normalmente de tripletas O-A-V y utilizando operadores booleanos, mientras que las conclusiones definirían nuevos hechos o realizarían acciones (Echarte, 2006).

En la ontología de peces del Catatumbo se usaron reglas como técnica para representar el conocimiento extraído, previamente organizado y clasificado de acuerdo a las características de las diferentes especies de peces.

#### 3.3 Conceptualización

En esta etapa se estructuraron los conocimientos en un modelo conceptual que describe el problema y su solución. Al desarrollar la ontología se debe tener en cuenta que la mayoría de las tareas aquí descritas se realizan simultáneamente con el proceso de adquisición de conocimientos.

##### 3.3.1. Construcción del glosario de términos.

El glosario de términos debe incluir todos los términos relevantes del dominio (conceptos, instancias, atributos, relaciones entre conceptos, etc.), sus descripciones en lenguaje natural, sus sinónimos y acrónimos (Ramos y Núñez, 2007).

### 3.3.2. Construcción de árboles de clasificación de conceptos.

Cuando el glosario de términos tenga una cantidad importante de elementos, se debe construir una taxonomía que defina la jerarquía entre los conceptos, como la expresada en Hábito (Fig. 2). Se debe evaluar que la taxonomía creada no contenga errores.

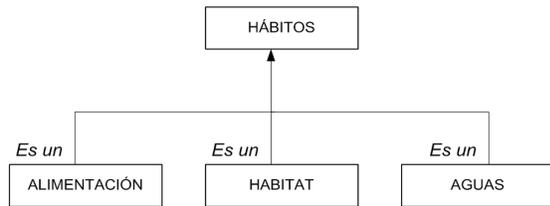


Fig. 2. Árbol de Conceptos: Hábito

### 3.3.3. Construcción del diagrama de relaciones binarias

El objetivo de este diagrama es establecer las relaciones entre los conceptos de una o más taxonomías de conceptos. Se muestra el diagrama de relaciones binarias (Fig. 3), en las cuales se representa el género.

Desde **Género** hacia cada una de las características.

*TieneAletaAdiposa*    *AletaAdiposa*  
*TamañoDeCuerpo*    *TamañoCuerpo*  
*ColorDeCuerpo*      *ColorCuerpo*  
*TipoDeHocico*        *TipoHocico*

Desde **Género** hacia los hábitos.

*TieneAlimentacion*    *Alimentacion*  
*ViveEnAguas*        *Aguas*  
*TieneHabitat*        *Habitat*

### 3.3.4. Construcción del diccionario de conceptos

El diccionario de conceptos contiene los conceptos del dominio, sus relaciones, instancias, atributos de clases y atributos de instancias. Las relaciones, atributos de instancias, y atributos de clases son locales al concepto, lo que significa que sus nombres pueden repetirse en diferentes conceptos.

### 3.3.5. Definición de las relaciones binarias en detalle

Se crea la tabla de relaciones binarias en la que se describe detalladamente todas las relaciones binarias incluidas en el diccionario de conceptos. Para cada relación binaria se debe especificar: nombre, conceptos fuente y destino, cardinalidad y relación inversa.

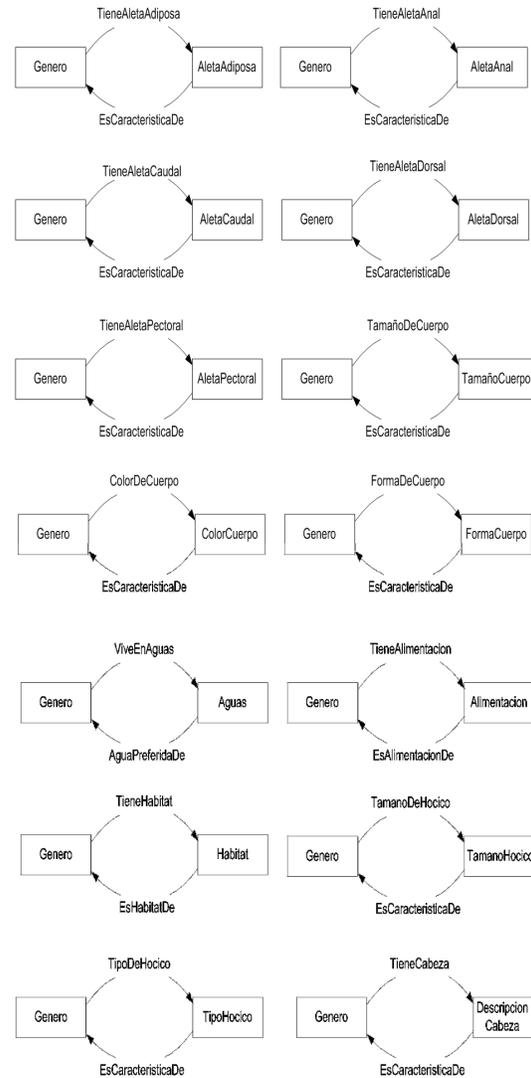


Fig. 3. Diagramas de Relaciones Binarias

### 3.3.6. Tabla de instancias

Todas las instancias que aparecen en el campo 'Instancias' del diccionario de conceptos deben estar presentes en la tabla de instancias. Para cada una de estas instancias se debe indicar: nombre de la instancia, nombre del concepto, atributos y valores que toman los atributos.

### 3.3.7. Tabla de reglas

Para cada fórmula que aparece en las tablas de atributos de instancias. Estas tablas se utilizan para deducir el valor numérico de un atributo de instancia a partir de otros atributos de instancia, atributos de clase y/o constantes. Cada tabla incluye: nombre de la regla o fórmula; descripción,

expresión en lenguaje natural; conceptos, atributos referidos, relaciones referidas y variables.

Estas reglas expresadas son de tipo *IF – THEN*, se representó mediante estas reglas el conocimiento adquirido sobre los peces del Catatumbo utilizando una tabla de reglas por cada especie de peces. Se incluye un ejemplo de la representación de una regla para la especie *RamdiaQuelen* (Tabla 1).

Tabla 1. Regla para la especie *RamdiaQuelen*

Nombre de la Regla	Regla de la especie <i>RamdiaQuelen</i>
Descripción	Cuerpo alargado y cilíndrico; cabeza algo deprimida con boca terminal; ojo con margen libre; aletas pectorales con espina aserrada en el margen posterior; lóbulos de la aleta caudal redondeados, con el inferior un poco mas largo que el superior.
Expresión	<p>if [Genero] (?G) and [Especie] (?E) and [Cuerpo] (?C) and [Aletas] (?A) and [Habit] (?H) and [Cabeza] (?CA) [Reproduccion] (?R) and [ContieneA] (?G,?E) and [tieneCaracteristica] (?G,?C) and [tieneCaracteristica] (?G,?A) and [tieneCaracteristica] (?G,?CA) and [tieneHabit] (?G,?H) and [CaracteristicaReproduccion] (?G,?R) and [FormaCuerpo] (?C,"CuerpoAlargado") and [ColorCuerpo] (?C,"MarronOscuro") and [Alimentacion] (?H,"MaterialVegetal") and [Habitat] (?H,"CorrientesDelPieDeMonte") and [Aguas] (?H,"DeCursoLento") and [DescripcionCabeza] (?CA,"CabezaDeprimidaConBocaTerminal") and</p> <p>then [NombreEspecie] (?E, "RamdiaQuelen")</p>
Conceptos	Cuerpo Aletas Habit Cabeza hábitat Alimentacion ColorCuerpo FormaCuerpo
Atributos referidos	NombreEspecie
Relaciones binarias ad-hoc	tieneCaracteristica
Variables	?G ?A ?C ?H ?CA ?R

#### 4. IMPLEMENTACIÓN DE LA ONTOLOGÍA

El objetivo de esta fase es escribir la ontología en un lenguaje formal que sea computable. Se escogió OWL por ser uno de los lenguajes más importantes para la construcción de ontologías; su entorno de desarrollo se puede soportar en el editor de Ontologías Protégé 3.4 (Puentes, 2010). Las clases definidas en la clasificación taxonómica están agrupadas en Orden, Familia, Género. Las clases concernientes a los conceptos del dominio están clasificadas en Actividad, Características, Hábito y Reproducción.

Para cada relación se define un rango y un dominio determinados, el rango indica los valores que puede tomar la relación y el dominio son las clases a las que se asigna dicha relación. La relación *ColorDelCuerpo* tiene como dominio Género y

como rango *ColorCuerpo*, la clase familia tiene 13 subclases, cada una de ellas esta relacionada con uno o varias subclases de Género; La clase género tiene 25 subclases, cada una de ellas esta relacionada con una única subclase de Familia, por ejemplo, *astyanax* tiene como familia *characidae*, a su vez cada género contiene a una o varias especies. La clase orden tiene 4 subclases, cada una de ellas esta relacionada con una o varias subclases de Familia. Por ejemplo, el orden *siluriformes* tiene como familia *aspredinidae*, *trichomycteridae*, *loricariidae*, *pseudopimelodae*, *Hectapteridae*.

Cada subclase de la clase *Características* también contiene a su vez subclases, cada una de ellas compuesta por instancias, estas últimas representan los datos actuales en la base de conocimiento de la ontología creada. Por ejemplo, la subclase *FormaCuerpo* tiene tres instancias las cuales son: *CuerpoAlargado*, *CuerpoCilindrico*, *CuerpoOvalado*. Las subclases *Actividad*, *Hábito* y *Reproducción*, también poseen instancias definidas, como se ilustra en la vista de instancias para actividad, hábito y reproducción (Fig. 4a y Fig. 4b). A través del OWL Viz se muestra la representación de la ontología de peces (Fig. 5).

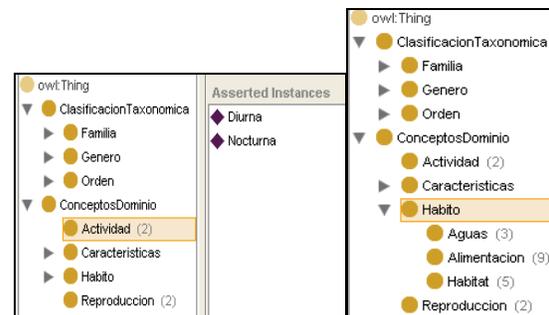


Fig. 4a. Vista de instancias para Actividad, Hábito

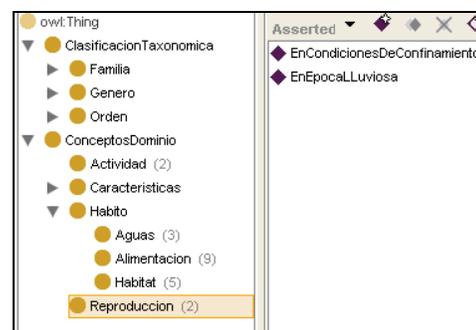


Fig. 4b. Vista de instancias para reproducción

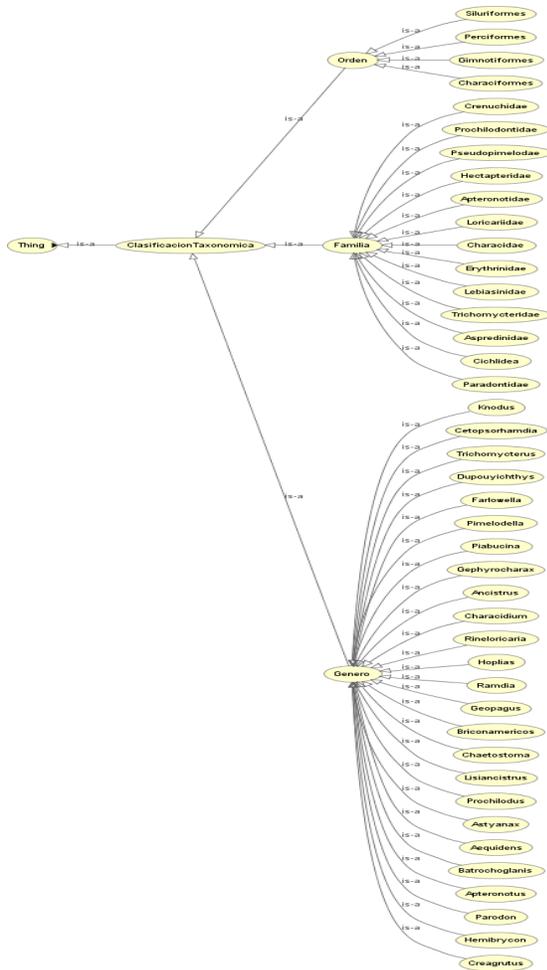


Fig. 5. Representación visual de la ontología

## 5. CONCLUSIONES

Para el desarrollo de una ontología, se debe inicialmente identificar el contexto, conceptualizar el conocimiento y representarlo, además, construir y validar la ontología. La adquisición de conocimiento es uno de los procesos más complejos en el desarrollo de ontologías, en este paso se deben consultar diversas fuentes de información, determinando los principales aspectos referentes al dominio tratado y así obtener un conocimiento efectivo que permite la determinación de las clases, las instancias, las relaciones, los atributos, los cuales representan su conocimiento.

Se logró conceptualizar apropiadamente una especie biológica representativa de la región del Catatumbo; estructurando un conocimiento representativo concerniente a los peces, mediante la información de fuentes de información tanto

primarias como secundarias, apoyadas por expertos en el área así como de libros, internet, etc. Estos datos fueron seleccionados y organizados de acuerdo a los aspectos más representativos como la clasificación taxonómica, sus características, hábitos y su actividad entre otros; este conocimiento debe estar en continuo refinamiento y validación.

Aunque existen varias metodologías para el desarrollo de ontologías, para la representación ontológica de los peces del Catatumbo se usó Methontology, ya que propone la descripción más ajustada de cada actividad a realizar, especificando los pasos para desarrollar cada tarea, las técnicas utilizadas, los productos de salida y su respectiva evaluación. Una vez obtenido y organizado el conocimiento sobre los peces, se desarrolló cada tarea según Methontology. En cada una de las actividades realizadas se crearon unas tablas y diagramas, por medio de las cuales se organizó y modeló de forma fácil y comprensible todo el conocimiento relevante del dominio tratado.

El proceso realizado en cada actividad según la metodología utilizada, se representó en el editor de ontologías protege; en donde se plasmaron los conceptos, propiedades, atributos e instancias y de la cual se obtuvo la respectiva codificación de la ontología en formato OWL.

## 6. DISCUSIÓN

La inteligencia artificial es de gran importancia en el mundo moderno, hace uso de las nuevas tecnologías, inspirándose en nuestra inteligencia para implementar soluciones que logren mejorar sustancialmente la forma en que vivimos (Russell, 2004). La universidad Francisco de Paula Santander Ocaña, propiamente el programa de ingeniería de sistemas, ha venido impulsando el estudio en este campo de conocimiento a través de su grupo de Investigación GITYD y del Semillero de Investigación en Inteligencia Computacional. En la actualidad se realizan trabajos que pretenden abordar problemas como la integración entre sistemas de información tradicionales y sistemas basados en conocimiento; aplicación de métodos y herramientas de razonamiento y de consulta cada vez más eficaces; modelar dominios complejos de manera fácil; diseñar e implementar interfaces de usuario enriquecidas, que faciliten y mejoren la experiencia del mismo.

Se ha logrado verificar con numerosos trabajos realizados en el área la pertinencia y eficacia de emplear Methontology como metodología de desarrollo de ontologías, considerada la mejor metodología más adecuada por todas las ventajas mencionadas previamente; sin embargo, se siguen realizando estudios en otros métodos y herramientas de desarrollo que puedan llevar a la implementación de ontologías empleando el menor esfuerzo en la fase de implementación por parte del desarrollador, con el fin de concentrar esfuerzos en el modelado y formalización del conocimiento.

### REFERENCIAS

- Fernández-López M, Gómez-Pérez A (2002) Overview and analysis of methodologies for building ontologies. *The Knowledge Engineering Review* 17(2):129–156.
- Galvis G., Mojica J., y Camargo M. (1997). *Peces del Catatumbo*. Santafé de Bogotá, D.C., Colombia.
- Gómez-Pérez A, Corcho O (2002). *Ontology Languages for the Semantic Web*. *IEEE Intelligent Systems & their applications* 17(1):54–60.
- Gómez-Pérez, A., Fernández-López, M., & Corcho, O. (2004). *Ontological Engineering: with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*. Springer.
- Maldonado, J.A.; Ortega A.; Usma J., S.; Galvis, G.; Villa, F.A.; Vásquez, L.; Prada, S. y Ardila, C. (2005). *Peces de los andes de Colombia*. Bogotá D.C., Colombia.

Puentes A. (2010). *Ontología para la Búsqueda Semántica de Géneros de Orquídeas de la Flora Colombiana*. Trabajo presentado en el evento Internacional “VII Pan-American Workshop in Applied & Computational Mathematics” en Choroní, Venezuela Junio.

Puerto E (2006). *La Expresiva SHIQ Como Lenguaje Ontológico para la Web Semántica*. Publicado en el volumen 1 Número 7 de La Revista Colombiana de Tecnologías de Avanzada de la Universidad de Pamplona, ISSN 1692-7257, p 78-83.

Russell & Norvig (2004). *Inteligencia Artificial, Un Enfoque Moderno*. Editorial Pearson.

### SITIOS WEB

- Echarte, F. (2006). *Técnicas y lenguajes para la representación del conocimiento*. <http://www.eslomas.com/index.php/archives/2006/12/14/tecnicas-y-lenguajes-para-la-representacion-del-conocimiento/>.
- Ramos, E. y Núñez, H. (2007). *Ontologías: componentes, metodologías, lenguajes, herramientas y aplicaciones*. <http://www.ciens.ucv.ve/escueladecomputacion/documentos/archivo/51>.
- Ramos, L. y Gil, R. (2007). *Hacia un Sistema de Información para Apoyar la Gestión de la Educación a Distancia*. <http://www.geocities.ws/industrialcla/ramos.pdf>.