

# LENGUAJES DE RECUPERACIÓN DE INFORMACIÓN SOBRE LA WEB SEMÁNTICA

Jorge E. Giraldo Plaza<sup>1</sup>, Sandra P. Mateus<sup>2</sup>, Maryem A. Ruiz<sup>3</sup>

<sup>1,2</sup> Ing. de Sistemas, MSc. en Ingeniería de Sistemas de la Universidad Nacional de Colombia, sede Medellín, Docente del área de Programas Informáticos, Facultad de Ingenierías, Politécnico Colombiano Jaime Isaza Cadavid, Medellín - Colombia.

<sup>3</sup> Ing. de Sistemas, MSc. en Informática de la Universidad Eafit. Docente del área de Programas Informáticos, Facultad de Ingenierías, Politécnico Colombiano Jaime Isaza Cadavid, Medellín - Colombia. Integrantes del Grupo de Investigación en Desarrollo de Software, GRINSOFT [jegiraldo, mruiz, spmateus]@elpoli.edu.co

## RESUMEN

La recopilación de información digital en las distintas áreas del saber es de vital importancia ya que facilita el intercambio de información, mejorando así las redes de apoyo académico y científico. Una manera para acceder a este conocimiento es mediante la recuperación de información sobre documentos, es así que existe una variedad de lenguajes de consulta y recuperación de información basados en la sintaxis XML. Este trabajo de investigación busca realizar un estudio comparativo desde el punto de expresividad de los siguientes lenguajes de consulta y recuperación: XQuery, XQL, RQL, SPARQL, SeRQL y OWL-QL.

**Palabras clave:** Web Semántica, Recuperación de Información, Tecnologías Web, Lenguajes de Consulta.

Recibido: 17 de Abril de 2009. Aceptado: 30 de Junio de 2009

*Received: April 17, 2009 Accepted: June 30, 2009*

## INFORMATION RETRIEVAL LANGUAGES ON SEMANTIC WEB

### ABSTRACT

*The collection of digital information from different areas of knowledge is vital as it facilitates the exchange of knowledge, thereby improving support to and academic scientists networks. One way to access this knowledge by retrieving information about documents, so that there is a variety of languages query and XML-Based information retrieval. This research seeks a comparative study in terms of expressiveness of the following query and recovery languages: XQuery, XQL, RQL, SPARQL, and OWL-QL SeRQL.*

**Keywords:** *Semantic Web, Information Retrieval, Web Technologies, Query Languages.*

## 1. INTRODUCCIÓN

Los lenguajes de etiqueta permiten describir información sobre la web; últimamente se intenta representar conocimiento sobre esta, no obstante, la acumulación de conocimiento genera un problema, la recuperación efectiva de información; para ello se proponen distintos lenguajes de consulta y recuperación de información que garanticen su integridad al ser descrita y almacenada en la web mediante los lenguajes de etiquetas.

Con el fin de determinar el lenguaje de recuperación de información apropiado en ambientes de bibliotecas digitales con características semánticas se propone entonces en este documento un estudio comparativo desde el punto de expresividad de los siguientes lenguajes de consulta y recuperación: XPath, XQuery, XQL, RQL, SPARQL, SeRQL y OWL-QL. Los criterios de evaluación se orientan al manejo y manipulación de los datos, el análisis de los grafos de representación del conocimiento y las inferencias que se realicen sobre estos: expresiones condicionales, operadores matemáticos, manejo de clases y objetos, análisis de atributos mediante sus rangos y dominios, determinación de rutas entre nodos que representan los conceptos y criterios orientados al manejo de las expresiones relacionadas con dichos conceptos<sup>1</sup>.

El documento de estructura de la siguiente manera: En la sección 2 se presenta un marco teórico de los lenguajes de consulta y recuperación de información sobre la web. En la sección 3 se exponen los criterios que medirán la expresividad de cada uno de los lenguajes. Luego en la sección 4 se realiza una caracterización de cada uno de los lenguajes desde la óptica de los criterios. En la sección 5 se presentan las conclusiones y trabajo futuro, orientadas a la propuesta de un Modelo de Recuperación Semántica sobre Bibliotecas Digitales. Por último de expone la bibliografía de la literatura consultada.

<sup>1</sup> Esta investigación se presenta como un objetivo específico del proyecto de investigación en curso Modelo de Recuperación Semántica sobre Bibliotecas Digitales asociado al Grupo de Investigación en Desarrollo de Software – GrinSoft – del Politécnico Colombiano Jaime Isaza Cadavid de la ciudad de Medellín, Colombia.

## 2. LENGUAJES DE RECUPERACIÓN DE INFORMACIÓN SOBRE LA WEB

### 2.1 Lenguajes de recuperación para XML

Los lenguajes de consulta basados en XML se presentan como una introducción teórica, esperando que su entendimiento de pie para que el lector pueda comprender directamente los antecedentes de los lenguajes soportados por moledos semánticos.

XML [1] (Extensible Markup Language), es un lenguaje de marcado que puede crear sus propias etiquetas. Fue propuesto por el Consorcio para la World Wide Web (W3C) para superar las limitaciones de HTML. XML como lenguaje estándar de marcado ofrece la posibilidad que la información que se modele ser asequible y administrable. A continuación se presentan los lenguajes de consulta sobre documentos XML más conocidos.

Xquery [2] (XML Query), también conocido como XML Query es un lenguaje de consultas estándar que utiliza la notación XML para definir consultas y manejar los resultados, se centra en encontrar y extraer elementos y atributos de documentos XML. XQuery hace un uso intensivo de XPath [3]. XPath en un principio fue parte de XSL 1.0 [4] y luego se desarrolló como una especificación separada. XQuery está definido en términos de un modelo basado en una jerarquía que fácilmente se representa en un árbol de información, no en términos de texto XML. Cada entrada a una consulta es una instancia de un modelo de datos, y la salida de una consulta también. El núcleo estructural que ofrece Xquery para realizar las expresiones se conoce como FLWOR, que son al XQuery lo que las distintas cláusulas dentro de una sentencia Select (select, from, where, etc) son al SQL. El nombre viene de For, Let, Where, Order by y Return; a continuación se explican cada uno de estos bloques de consulta:

**FOR:** Vincula una o más variables a las expresiones escritas.

**LET:** Vincula una variable al resultado completo de una expresión, pero a diferencia del FOR lo hace sobre el resultado de la expresión sin iteración.

**WHERE:** Filtra variables, quedando solo aquellas que cumplen con la condición.

**ORDER BY:** ordena las variables según una serie de criterios dados.

**RETURN:** construye el resultado de la expresión FLWOR para una variable dada.

XQL [5] (XML Query Language) en este lenguaje la información es vista como un conjunto de nodos, información sobre las relaciones entre nodos, o valores derivados. XQL es una extensión natural del sistema de patrones XSL. XQL puede realizar consultas en infinidad de tipos de documentos como son documentos estructurados, colecciones de documentos, bases de datos, estructuras DOM [6], catálogos, etc.

XQL extiende las posibilidades de XSL, permitiendo identificar clases de nodos añadiendo lógica booleana, filtros, indexado de los nodos, etc. XQL está diseñado para ser usado en diversos contextos: al ser una extensión de XSL, se puede aplicar para asociar nodos, buscar en repositorios, y para muchas otras aplicaciones.

La sintaxis de XQL esta basada en el concepto de contexto. El contexto es el conjunto de nodos sobre el que se recuperará la información en XQL. Se emplea la misma sintaxis de XPath como parte de los elementos que se pueden controlar. Si se desea la recuperación de la información desde la raíz, se emplea el símbolo "/" como prefijo. Si se desea la recuperación de la información desde el nodo actual, se emplea el símbolo "./" como prefijo (normalmente no es necesario).

Para hacer búsquedas de información recursivas descendentes se emplea el símbolo "/" como prefijo. Al igual para hacer búsquedas de información recursivas descendentes a partir del nodo actual se emplea el símbolo "./" como prefijo.

Para seleccionar nodos hijos o descendientes de cierto tipo se emplea el símbolo "/" o "/" con ayuda del elemento "\*", este devuelve todos los hijos del contexto actual. Para recuperar atributos utilizamos "@" delante del nombre del atributo.

XQL también permite realizar recuperaciones de información condicionales, para ello se utiliza la cláusula "[expresion]", los operadores lógicos "\$and\$", "\$or\$" y "\$not\$" y las equivalencias igual "=" (\$eq\$) y distintas "!=" (\$ne\$). Lo anterior se enmarca como la principal diferencia con XPATH.

Aunque toda la información (los elementos y atributos) es String podemos querer realizar comparaciones numéricas, ello también es posible, pero siempre colocando la variable a la izquierda. Para realizar estas comparaciones tenemos "<",

"<=", ">" y ">=". Así mismo permite operaciones sobre conjuntos con los operadores "\$union\$" e "\$intersect\$". Y también algunas funciones de agregación como "count()" que devuelve el numero de nodos de un conjunto.

## 2.2 Lenguajes de recuperación para RDF

RDF [7] (Resource Description Framework) es una base para procesar metadatos y considerado el derrotero para la implementación de la Web Semántica [8]; proporciona interoperabilidad entre aplicaciones que intercambian información legible por máquina en la Web. Uno de los objetivos de RDF es hacer posible especificar la semántica para las bases de datos en XML de una forma normalizada e interoperable. El objetivo general de RDF es definir un mecanismo para describir recursos que no cree ninguna asunción sobre un dominio de aplicación particular, ni defina (a priori) la semántica de algún dominio de aplicación.

El modelo RDF se basa en la idea de convertir las declaraciones de los recursos en expresiones con la forma sujeto-predicado-objeto. El elemento de construcción básica en RDF es el "tripleto" o sentencia, que consiste en dos nodos (sujeto y objeto) unidos por un arco (predicado), donde los nodos representan recursos, y los arcos propiedades. El objeto es el valor de la propiedad o el otro recurso con el que se establece la relación.

Por ejemplo una sentencia podría expresar el hecho de que el autor (predicado) del libro "En la cripta" (sujeto) fue el escritor H.P. Lovecraft (objeto). Encadenando estos tripletes se construyen grafos o redes semánticas para la web. A continuación se presentan los lenguajes de consulta y recuperación sobre documentos RDF.

SPARQL [9] (pronunciado "sparkle") es un lenguaje de recuperación basado en RDF; su nombre es un acrónimo recursivo del inglés SPARQL Protocol and RDF query Language. Se trata de una recomendación para crear un lenguaje de consulta dentro de la Web semántica que está ya implementada en muchos lenguajes y bases de datos.

Con SPARQL los desarrolladores y usuarios finales pueden representar y utilizar los resultados obtenidos en las búsquedas a través de una gran variedad de información como son datos personales, redes sociales y metadatos sobre recursos digitales como música e imágenes.

Similar a otros lenguajes de consultas sobre conjuntos de datos, SPARQL permite a los usuarios declarar específicamente las condiciones requeridas para los datos a ser recuperados más que describir explícitamente los pasos orientados a la descripción de una ruta para recuperarlos.

SPARQL consiste en tres especificaciones separadas, que contienen diferentes partes de su funcionalidad. En total, consiste en un lenguaje de consulta, un formato para las respuestas, y un medio para el transporte de consultas y respuestas, estos son:

SPARQL Query Language: Núcleo de SPARQL. Explica la sintaxis para la composición de sentencias y su concordancia.

SPARQL Protocol: Formato utilizado para la devolución de los resultados de las búsquedas (sentencias SELECT o ASK), a partir de un esquema de XML.

SPARQL Query XML Results Format: Describe el acceso remoto de datos y la transmisión de consultas de los clientes a los procesadores. Utiliza WSDL [10] para definir protocolos remotos para la consulta de bases de datos basadas en RDF.

Para una explotación total del conocimiento almacenado en modelos de datos RDF, se requiere un lenguaje de consulta a nivel semántico. Este lenguaje semántico debe ser sensible a la semántica de las primitivas RDF; RQL [11] es un lenguaje de consulta declarativo para RDF que explícitamente captura esta semántica en su diseño.

RQL fue desarrollado en el instituto ICS-FORTH, y su potencia semántica está basada en la evaluación de caminos de expresiones sobre grafos RDF. Permite el uso de variables tanto para denotar clases y propiedades. Permite consultar esquemas RDF, RDF-S [12], y descripciones RDF en una misma consulta. RQL está definido por medio de un conjunto de consultas básicas, e iteradores que se permiten construir otras consultas a través de una composición funcional con base en la teoría planteada por OQL [13] (Bases de datos orientadas a objetos).

Sin embargo la semántica de RQL no es completamente compatible con la semántica de RDF, existen un número de restricciones adicionales que deben ser puestas sobre los modelos RDF para poder llevar a cabo las consultas sobre los documentos con entera

confiabilidad; un ejemplo de tales restricciones es que cada propiedad debe tener exactamente un dominio y un rango específico.

SeRQL [14] (Sesame RDF Query Language, pronunciado como "circle") es un lenguaje de recuperación para RDF/RDFS desarrollado por Aduna como parte del software Sesame. Combina características de otros lenguajes (principalmente RQL, RDQL [15], N-Triples [16] y N3 [17]) y añade otras propias. Algunas de las características más importantes de SeRQL para la recuperación y organización de la información son:

- Tranformación de grafos.

- Soporte de RDF Schema.

- Soporte de los tipos de datos de XML Schema.

- Emparejado de caminos opcionales.

La sintaxis de SeRQL es similar a la de RQL, añadiendo algunas modificaciones. Al igual que RQL se basa en una interpretación de los grafos modelados con RDF.

### 2.3 Lenguaje de Consulta para OWL

OWL-QL [18] se presenta como el lenguaje de recuperación y consulta sobre documentos descritos en OWL [19] (Web Ontology Language). El OWL está 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. OWL puede usarse para representar explícitamente el significado de términos en vocabularios y las relaciones entre aquellos términos. Esta representación de los términos y sus relaciones se denomina una ontología. En realidad, OWL es una extensión del lenguaje RDF y emplea las tripletas de RDF, aunque es un lenguaje con más poder expresivo que éste.

OWL-QL es un lenguaje formal de consulta y un protocolo de comunicación entre agentes del tipo consulta-respuesta, que usa el conocimiento representado en el lenguaje OWL para realizar las interacciones. OWL-QL es el lenguaje candidato a convertirse en el estándar para la comunicación entre agentes computacionales. Aunque OWL-QL fue especificado para ser usado en OWL, puede ser fácilmente adaptado para utilizar en lenguajes como RDF y RDF-S. OWL-QL ha sido diseñado teniendo en cuenta varias asunciones de la web semántica, en orden de adaptarse a diferentes sistemas. La primera es que los agentes usan un sistema

automático de razonamiento para derivar las respuestas a las preguntas. La segunda, que algunos servidores solo tiene información parcial acerca de un tema dado.

De acuerdo a los aspectos anteriores, este lenguaje de consulta provee un medio para transferir parcialmente las consultas-respuestas y permite que el usuario provea el número máximo de respuestas que espera recibir. Adicionalmente OWL-QL dado que es posible que el usuario desee condicionar los servidores en los que se encuentran las bases de datos de conocimiento para el dialogo consulta-respuesta, la estructura de la consulta OWL-QL esta diseñada en un nivel de abstracción cuya sintaxis no restringe este tipo de situaciones. La especificación solo describe que tipos de objetos son pasados entre servidores y clientes y los componentes opcionales que se manejan en la comunicación. Finalmente OWL-QL requiere que el conocimiento sea plasmado en un lenguaje cuya semántica y teoría lógica este formalmente representada como OWL o DAML [20].

### 3. CRITERIOS DE EVALUACIÓN

Con base en las investigaciones tratadas realizadas [21], [22], [23], los criterios a tener en cuenta para la evaluación de los lenguajes de consulta descritos en la sección 1 están orientados a medir su expresividad desde el punto de vista teórico más que por rendimiento. Los criterios también se orientan a la medición del comportamiento de los lenguajes con base en la consulta y los datos recuperados, del mismo modo en función del grafo que se genera una vez se hace el modelamiento de un dominio específico, así pues, los criterios son los siguientes:

**Expresiones Condicionales:** Hacen referencia a estructuras de control y de filtrado de la información a recuperar, entre estos se encuentran el Si-Entonces y el Where.

**Cuantificadores Existenciales:** Determinan la existencia de los datos dentro de un documento, entre estos se encuentran el Existe, Todos, Alguno y Ninguno.

**Operadores Matemáticos:** Complemento de los dos anteriores criterios, ya que permite el cálculo de valores con base en los datos a recuperar, entre estos se encuentran las Operaciones Aritméticas, Comparadores, Operadores Lógicos, Funciones de Agrupación y Funciones de Cadena de caracteres.

**Clases y Objetos:** Hacen referencia a la recuperación de datos representados por conceptos y sus instancias, así mismo del soporte de características como la herencia y manejo de tipos de datos.

**Rango y Dominio:** Refiere a la información que puede se asocia a atributos de clases en específico, es decir, determinar la pertenencia de un atributo y el rango de valores que estos puedan adquirir.

**Recursos Adyacentes:** Este criterio hace referencia a la capacidad de determinar cuantos y cuales conceptos se encuentran asociados a los datos recuperados en la consulta.

**Predicados sobre recursos:** Significa la capacidad de determinar cuantas expresiones dentro del modelado de la información existe sobre un determinado concepto.

**Distancia entre recursos:** Refiere a la capacidad del lenguaje de determinar la distancia entre dos conceptos, es decir, con base en el grafo que lo representa determinar el número de arcos que existen entre dos conceptos.

**Reificación:** Hace referencia a la capacidad de recuperar datos contenidos en sentencias realizadas sobre otras sentencias.

### 4. CARACTERIZACIÓN DE LOS LENGUAJES

En esta sección se evalúan las características de cada uno de los lenguajes expuestos con respecto a los criterios de evaluación descritos.

#### 4.1 XQuery

Respecto a las Expresiones Condicionales maneja el For, el Where y el If-Then-Else. Por el lado de los Cuantificadores Existenciales, permite el uso de Every y Some.

En cuanto los operadores matemáticos XQuery ofrece buenas prestaciones, ya que permite: operaciones aritméticas como +, -, \*, /, mod, operadores de comparación como =, !=, <, >, <=, >= y not(), por su parte los operadores de agrupación que ofrece son count(), min(), max() y sum(), por último para el manejo de cadena de caracteres ofrece concat(), startswith(), end-with(), substring() y otros.

Aunque no permite obtener instancias de conceptos representados en clases, si permite el manejo de tipos de datos, estos son: reales, fecha, hora, cadena de caracteres y boléanos.

Los criterios relacionados con el análisis de la estructura del grafo de conceptos y sus propiedades representadas en lenguajes como el RDF no los soporta, ya que Xquery se encarga únicamente de recuperar datos bajo condiciones específicas conociendo con anterioridad su existencia, no obstante soporta con confianza aceptable recuperación de datos sobre expresiones hechas sobre expresiones, pues en sí estas se refieren a documentos XML que no requieren inferencias de tipo conceptual.

#### 4.2 XQL

Respecto a las Expresiones Condicionales maneja XQL no los soporta al igual que los Cuantificadores Existenciales.

En cuanto los operadores matemáticos XQL presenta los mismos beneficios que XPath, por tanto ofrece: operaciones aritméticas como +, -, \*, /, mod, operadores de comparación como =, !=, <, >, <=, >=, , así mismo soporta operados lógicos como or, and y not. Por su parte los operadores de agrupación que ofrece son count(), union(), conjunction() y disjunction().

Al igual que Xquery no permite obtener instancias de conceptos representados en clases, pero sí permite el manejo de tipos de datos, estos son: cadena de caracteres, enteros y reales.

Del mismo modo no soporta los criterios relacionados con el análisis de la estructura del grafo pero debido a su compatibilidad con sintaxis XPath permite el análisis de las rutas del grafo, no obstante teniendo un conocimiento previo. En cuanto a la reificación presenta aceptable confianza en la recuperación de datos.

#### 4.3 SPARQL

Las Expresiones Condicionales que soporta este lenguaje son similares a las de SQL, ya que la estructura de las consultas cumplen el mismo formato, se empieza con Select y se incluye el condicional Where para filtrar el resultado, existen algunos auxiliares de ordenamiento, agrupación y modificación, estos son: Distinct, Describe, Order By Desc, Order By Asc, Limit, Union y Offset.

Soporta operadores aritméticos de dos tipos: Binarios (A operador B) +, -, \*, / y los unitarios que se aplican a una sola variable, así +A, -A. Los operadores de comparación que soporta son =, !=,

<, >, <=, >=, , así mismo soporta operados lógicos como or, and y not.

SPARQL soporta consultas sobre la estructura del grafo que se representa ya que pretende encajar los patrones de preguntas al mismo. Permite la descripción del rango y el dominio de una propiedad, pues soporta el concepto de tripleta proveniente de la teoría del lenguaje RDF. Así mismo soporta los siguientes tipos de datos: decimales, enteros, cadena de caracteres, boléanos y colecciones.

Soporta consultas sobre expresiones de reificación. Por su parte soporta el análisis de la estructura del grafo de modelamiento en términos distancias y caminos entre nodos que representan los conceptos. No soporta operadores para contar el número de expresiones que se relacionan con un concepto.

#### 4.4 RQL

Las Expresiones Condicionales que soporta este lenguaje son similares a las de SeRQL, así mismo los operadores aritméticos de dos tipos: Binarios (A operador B) +, -, \*, / y los unitarios que se aplican a una sola variable, así +A, -A. Los operadores de comparación que soporta son =, !=, <, >, <=, >=, , así mismo soporta operados lógicos como or, and y not.

Del mismo modo soporta consultas sobre la estructura del grafo. Permite la descripción del rango y el dominio de una propiedad, pues soporta el concepto de tripleta proveniente de la teoría del lenguaje RDF. No obstante la semántica de RQL no es completamente compatible con la semántica de RDF, existen un número de restricciones adicionales que deben ser puestas sobre los modelos RDF para poder llevar a cabo las consultas sobre los documentos con entera confiabilidad; un ejemplo de tales restricciones es que cada propiedad debe tener exactamente un dominio y un rango específico. En cuanto a los datos que soporta permite el manejo de: decimales, enteros, cadena de caracteres, boléanos y colecciones.

En cuanto a expresiones de tipo reificación presenta un soporte débil, aunque es posible realizarlo pero sobre el mismo sujeto no sobre diferentes. Por su parte el análisis de la estructura del grafo de modelamiento en términos distancias y caminos entre nodos que representan los conceptos. No soporta la especificación de una ruta específica, ya

que la consulta se basa en la descripción de los datos más que en los pasos para llegar a estos como parte del resultado.

#### 4.5 SeRQL

El análisis de SeRQL es similar al de RQL, pues este surgió con base en la sintaxis del anterior. Sin embargo presenta una mejora en el manejo de expresiones de tipo reificación.

#### 4.6 OWL-QL

Las expresiones condicionales que maneja el OWL-QL se refieren a manejo explícito de instancias sobre las clases, cabe recordar que este lenguaje basa su funcionamiento en inferencias realizadas sobre el conocimiento modelado. Adicionalmente se aprovecha de las ventajas ofrecidas en el modelamiento de OWL y RDF.

Acoge las características de los lenguajes de consulta sobre RDF. Sin embargo es débil su propuesta en cuanto al manejo de operadores matemáticos y de comparación. Sin embargo es viable al momento de analizar la estructura del grafo de modelamiento en términos distancias y caminos entre nodos que representan los conceptos. No soporta la especificación de una ruta específica, ya que la consulta se basa en la descripción de los datos más que en los pasos para llegar a estos como parte del resultado.

### 5. CONCLUSIONES Y TRABAJO FUTURO

En este apartado se presentan las conclusiones del estudio realizado sobre los lenguajes desde el punto de vista de criterios de expresividad, adicionalmente se presenta una tabla resumen para facilitar su entendimiento, en donde un “+” significa que cumple esa característica, un “-” que no cumple la característica y un “+” que cumple débilmente la característica indicada.

Del estudio realizado se puede concluir que los lenguajes de consultas sobre documentos XML presentan fortalezas en el manejo y manipulación de los datos recuperados con base en los operadores aritméticos y las expresiones condicionales, así mismo en el trabajo con cadena de caracteres, el ordenamiento y secuenciamiento de los mismos.

De los lenguajes que realizan consultas sobre documentos RDF, su ventaja en que se complementan efectivamente con los modelos

RDF, sin embargo no permiten hacer consultas de tipo semántico, en el sentido de que no tienen porqué basarse necesariamente en elementos (conceptos, atributos y relaciones) de una ontología, sino exclusivamente en el modelo RDF.

A excepción del OWL-QL los lenguajes de consulta no llevan a cabo inferencias para generar nuevo conocimiento, sino que se limitan a generar los individuos o grupos de estos en base al contenido de la fuente de datos. Es importante mencionar que a excepción de OWL-QL, ninguno de los lenguajes descritos lleva a cabo ningún tipo de inferencias para extraer conocimiento. Por último y pensando en la representación sintáctica, OWL-QL es un lenguaje que se basa en la sintaxis XML.

A continuación se presenta la tabla resumen (Tabla 1.), en la primera columna se presentan los lenguajes y en la primera fila los criterios de evaluación haciendo uso de las siguientes convenciones:

- EC = Expresiones Condicionales.
- CE = Cuantificadores Existenciales.
- OM = Operadores Matemáticos.
- CO = Clases y Objetos.
- RD = Rango y Dominio.
- RA = Recursos Adyacentes.
- PR = Predicados sobre recursos.
- DR = Distancia entre recursos.
- RF = Reificación.

Tabla 1. Comparación de Lenguajes

	E C	C E	O M	C O	R D	R A	P R	D R	R F
<b>XQUERY</b>	+	+	+	-	-	-	-	-	+ -
<b>XQL</b>	+	+	+	-	-	-	-	-	+ -
<b>SPARQL</b>	+	+	+	+	+	+	+	+	+
<b>RQL</b>	+	+	+	+	+	+	+	+	+
<b>SeRQL</b>	+	+	+	+	+	+	+	+	+
<b>OWL-QL</b>	-	+	-	+	+	+	+	+	+

Como trabajo futuro se pretende realizar una exploración en la literatura especializada en métodos de búsqueda y recuperación semántica en bibliotecas digitales con el fin de determinar el modelo apropiado para su uso en bibliotecas digitales. Posteriormente se realizará una

caracterización de los componentes que intervienen en una consulta semántica para el modelamiento de búsqueda y recuperación de información en bibliotecas digitales y así aplicar el lenguaje de búsqueda elegido en una validación que se haga sobre una biblioteca digital desarrollada en el Politécnico Colombiano Jaime Isaza Cadavid de la Ciudad de Medellín.

## 6. BIBLIOGRAFÍA

- [1] Bray, T. Extensible Markup Language (XML) 1.0, World Wide Web Consortium Octubre 2000, <http://www.w3.org/TR/REC-xml>.
- [2] Marchiori, M. XML Query (XQuery), World Wide Web Consortium, 23 Septiembre 2003, <http://www.w3.org/XML/Query>.
- [3] World Wide Web Consortium (W3C). XML Path Language (XPath) Version 1.0, Noviembre 1999. <http://www.w3.org/TR/xpath>.
- [4] World Wide Web Consortium (W3C). XSL Transformations (XSLT) 2.0. 2002. <http://www.w3.org/TR/xslt20>.
- [5] Robie, J., Lapp, J., y Schach D. XML Query Language (XQL). World Wide Web Consortium, <http://www.w3.org/TandS/QL/QL98/pp/xql.html>.
- [6] World Wide Web Consortium (W3C). Document Object Model. 2003. <http://www.w3.org/dom>.
- [7] Brickley, D., Interest Group Chair and Core Working Group co-chair. Resource Description Framework (RDF). World Wide Web Consortium (W3C). 5 Agosto 2003. <http://www.w3.org/rdf>.
- [8] Berners-Lee, T., Hendler, J. y Lassila, O. The Semantic Web. Scientific American.com. May 17, 2001.
- [9] Prud'hommeaux, E. y Seaborne, A. SPARQL Query Language for RDF. World Wide Web Consortium (W3C). 2006. <http://www.w3.org/TR/2006/CR-rdf-sparql-query-20060406>.
- [10] Bultan, T., Fu, X., Hull, R. y Su, J. Conversation Specification: A New Approach to Design and Analysis of E-Service Composition, Proceedings of the 12th International World Wide Web Conference (WWW 2003), ACM, 2003.
- [11] Karvounarakis, G., Alexaki, S., Christophides, V., Plexousakis, D. y Schol, M. RQL: A Declarative Query Language for RDF. In Proceedings of the Eleventh International World Wide Web Conference (WWW'02), USA, May 7-11 2002.
- [12] Brickley, D. y Guha, R. RDF vocabulary description language 1.0: RDF Schema. World Wide Web Consortium (W3C), 2003. <http://www.w3.org/TR/rdf-schema>.
- [13] Cattell, G., Barry, D., Berler, M., Eastman, J., Jordan, D., Russell, D., Schadow, O., Stanienda, T. y Velez, F. The Object Database Standard ODMG 3.0. Morgan Kaufmann, January 2000.
- [14] Broekstra, J. y Kampman, A. SeRQL: An RDF Query and Transformation Language. Submitted to the International Semantic Web Conference, ISWC 2004.
- [15] Andy Seaborne, A. Rdfql - a query language for rdf, w3c member submission, January , 2004. <http://www.w3.org/Submission/2004/SUBM-RDQL-20040109>.
- [16] M. Sintek and S. Decker. TRIPLE - an RDF query, inference and transformation language. In Deductive Databases and Knowledge Management (DDLK), 2001.
- [17] Berners-Lee, T. Notation 3. 2001. <http://www.w3.org/DesignIssues/Notation3>
- [18] Fikes, R., Hayes, P. y Horrocks P. OWL-QL: A Language for Deductive Query Answering on the Semantic Web. KSL Technical Report 03-14.
- [19] McGuinness, D. y van Harmelen, F. OWL Web Ontology Language Overview. 2004. <http://www.w3.org/TR/2004/REC-owl-features-20040210>
- [20] DAML. The DARPA Agent Markup Language Homepage. 2001. <http://www.daml.org>
- [21] Angles, R. y Gutierrez, C. Querying RDF Data from a Graph Database Perspective. 2nd European Semantic Web Conference (ESWC), Heraklion, Greece. Lecture Notes in Computer Science, Volume 3532/2005, pp. 346-360. May 2005.
- [22] Angles, R., Gutierrez, C. y Hayes, J. RDF Query Languages Need Support for Graph Properties. Reporte técnico Nro. TR/DCC-2004-3, Departamento. de Ciencias de la Computación de la Universidad de Chile. Junio 2004.
- [23] Haase, P., Broekstra, J., Eberhart, A. y Volz, R. A Comparison of RDF query languages. In: McIlraith, Sheila A., Plexousakis, Dimitris, van Harmelen, Frank (Eds.), Proceedings of the Third International Semantic Web Conference, Springer, Berlin / Heidelberg. pp. 502-517. 2004.