

Generación Semiautomática de una Ontología Geográfica a partir de Fuentes Heterogéneas

Manuel E. Puebla-Martínez¹, José M. Perea-Ortega², and Alfredo Simón-Cuevas³

¹ Universidad de las Ciencias Informáticas, La Habana, Cuba
mpuebla@uci.cu

² Universidad de Extremadura, Badajoz, Spain
jmperea@unex.es

³ Instituto Superior Politécnico José Antonio Echeverría, CUJAE, La Habana, Cuba
asimon@ceis.cujae.edu.cu

Resumen. En contextos donde el conocimiento requiere ser semánticamente formalizado para su procesamiento computacional, el desarrollo y uso de ontologías se ha ido incrementado, como por ejemplo ocurre en los Sistemas de Información Geográfica. En este trabajo se presenta un método semiautomático para construir una ontología geográfica a partir de fuentes de información geográfica heterogéneas, a saber: Geonames, OpenStreetMap y Bases de Datos Espaciales. El método propuesto parte de la construcción manual de una ontología preliminar, la cual es extendida a partir de cada una de las fuentes mencionadas y de la generación automática de relaciones espaciales. La principal contribución de este trabajo radica en la integración en una ontología OWL2 de las fuentes de información ya mencionadas y los mecanismos desarrollados para su posterior gestión. Se realizaron pruebas de los procesos de extensión incluidos en el método, considerando información de una región específica de Cuba, así como pruebas funcionales a través de consultas, algunas de las cuales son ejemplificadas.

Palabras clave: Ingeniería ontológica, Extracción de información, Recuperación de Información Geográfica

Abstract. In contexts where knowledge should be semantically well-defined for computational processing, the development and use of ontologies has been increasing. An example are the Geographical Information Systems. This work presents a semi-automatic method to generate a geographic ontology from heterogeneous geographical information sources such as Geonames, OpenStreetMap and spatial databases. The proposed approach begins with a preliminary ontology, which is enhanced and extended from the prior sources along with the automatic generation of spatial relationships. The main contribution of this work is the integration of several geographic information sources into an OWL2-ontology, together with different mechanisms developed for its management. Several experiments were performed on the specific region of Cuba, as well as functional testing through different queries; some of them are shown as examples.

Keywords: Ontological Engineering, Information Extraction, Geographic Information Retrieval

1. Introducción

Según se ha reportado, aproximadamente el 80 % de la información en formato electrónico actualmente almacenada, se corresponde con información georeferenciada o susceptible de serlo [7]. En este escenario, la Recuperación de Información Geográfica (RIG) resulta un tema de especial interés, el cual se centra en el desarrollo de métodos y sistemas que permitan el acceso oportuno y preciso a dicha información. Teniendo en cuenta la gran cantidad de información que manejan estos sistemas, en muchos casos provenientes de fuentes diversas y en lenguaje natural, así como los mecanismos de consultas actualmente empleados, el tratamiento de la semántica asociada a la información constituye una necesidad para incrementar la precisión de los resultados y la satisfacción de los usuarios. En la solución de esta problemática, el uso de ontologías [5] ha sido una de las alternativas más comúnmente consideradas. El uso de ontologías en RIG, específicamente las ontologías geográficas [17], permiten recuperar información teniendo en cuenta la semántica con la que han sido descritos los datos y el conocimiento representado, dando la posibilidad de explicitar información adicional: tipo, estructura, jerarquía, relaciones y reglas [9]. Sin embargo, la construcción manual de ontologías geográficas representa un trabajo agotador y propenso a errores humanos, debido a la rapidez con que crecen los volúmenes de información y el tiempo a dedicar por un especialista en la captura de la información a ser representada en dicha ontología.

Por otra parte, el desarrollo de Internet y la interconectividad global ha propiciado el surgimiento de grandes fuentes de información generadas por los usuarios, conocidas en el ámbito espacial como Información Geográfica Voluntaria (VGI). Dos de las más abarcadoras en este contexto son Geonames⁴ y OpenStreetMap⁵ (OSM). Geonames es una base de datos geográfica que cubre todos los países y contiene más de ocho millones de nombres de lugares. La información está organizada en nueve clases, conocidas como *featureClass*, donde cada una cuenta con un grupo de códigos para clasificar la información contenida. OSM es un proyecto de colaboración para crear un mapa del mundo que sea editable por los usuarios, inspirado en Wikipedia, pero en lugar de artículos los usuarios editan entidades geográficas (*Nodes*, *Ways* y *Relations*). Por otra parte, también se encuentran las Bases de Datos Espaciales (BDE), las cuales constituyen un tipo de base de datos donde se describen objetos espaciales a través de tres características básicas: atributos, localización y topología [13]. Estas fuentes de información han sido utilizadas como recursos en sistemas de RIG, en su mayoría usando solo una de ellas, y se han desarrollado proyectos con el objetivo de describir la información que almacenan en formato RDF⁶, por tanto con bajo nivel de descripción semántica, quedando mucho por hacer cuando se trata de la representación y el razonamiento de las características espaciales [11].

⁴ <http://www.geonames.org>

⁵ <http://www.openstreetmap.org>

⁶ <http://www.w3.org/RDF>

En este trabajo se propone un método para la construcción semi-automática de una ontología geográfica en la que se integrará información proveniente de Geonames, OSM y BDE. El método parte de la construcción manual de una ontología preliminar del dominio geográfico usando *Protégé*, en la que se representan conceptos, propiedades y relaciones espaciales comunes, e incluye mecanismos de extensión de esta ontología a partir de cada una de las fuentes de información, dividido en tres fases, así como un proceso final de generación automática de relaciones espaciales. Como resultado se obtiene una ontología geográfica codificada en OWL2. Este diseño se ha concebido como un proceso de desarrollo de la ontología de manera iterativa e incremental, y no partiendo de cero, sino tomando de referencia información existente en otras ontologías geográficas, así como de los propios recursos mencionados. La ontología geográfica resultante será utilizada en un sistema de RIG, con el objetivo de mejorar la precisión de sus resultados e incrementar la satisfacción de los usuarios.

2. Trabajo relacionado

En la literatura revisada existen varios trabajos que abordan el problema de la generación semiautomática de ontologías a partir de bases de datos [1,16,8], sin embargo, no se pudo encontrar y adaptar sus resultados de software. De ellos solo en [1] se aborda el ámbito espacial.

La vinculación de datos geoespaciales desde sistemas de información heterogéneos ha sido abordado previamente en el proyecto GeoKnow⁷, siendo uno de los resultados más relevantes la obtención del LinkedGeoData (representación en RDF de OpenStreetMap). En este proyecto los procesos de extracción y administración de información geoespacial se llevan a cabo a partir del tratamiento de tripletas RDF y no a través de la construcción de ontologías geográficas codificadas en lenguaje OWL2, lo cual limita las posibilidades de descripción semántica de esta información y con ello la generación de nuevo conocimiento. Sin embargo, al ser un proyecto genérico, que no se centra en ninguna fuente de datos específica, se ha considerado que existe un margen de mejora si se profundiza en las características particulares de cada conjunto de datos en el momento de la integración.

En [4] se reporta una propuesta de integración entre LinkedGeoData y Geonames, donde cada par de topónimos, es analizado mediante la combinación de información referente a distancia espacial y similitud sintáctica de las etiquetas que los identifica. Sin embargo, el método de integración propuesto solo enlaza una pequeña parte de la información contenida en ambas fuentes de datos, específicamente, la referente a las zonas residenciales. Al igual que en GeoKnow, el resultado de la integración de los datos se describe y almacena como tripletas RDF. Por otra parte, en [4] también se asume que todas las zonas residenciales en Geonames tienen asociada un rectángulo de delimitación (*boundingBox*, en inglés), lo cual no ocurre en todos los casos.

⁷ <http://geoknow.eu>

En [6] se presenta la ontología geográfica GeoNW, construida a partir del análisis de tres fuentes de datos: Geonames, WordNet⁸ y Wikipedia. Los autores señalan que la comparación de nombres hubiera sido más eficaz si se hubiera hecho a través de una función de distancia entre cadenas, para tener en consideración errores tipográficos. Además, teniendo en cuenta que WordNet constituye una base de datos léxica de ámbito general, por lo que aporta muy poco en cuanto a sinónimos de lugares geográficos de regiones específicas. Por ejemplo, de los 812 topónimos almacenados en Geonames de Cuba, solo en dos de ellos se identifican sinónimos en WordNet: 'Cuba' y 'Havana'. En esta propuesta Wikipedia es utilizada para añadir nacionalidades a los lugares administrativos. En este sentido, Wikipedia brinda mayores beneficios como fuente de datos geográfica, debido al crecimiento de proyectos asociados a las coordenadas geográficas de los artículos publicados en este recurso, tales como WikiProject Geographical Coordinates⁹. Finalmente, GeoNW elimina un gran número de inconsistencias encontradas en Geonames, sin especificar la cantidad, ni referencia a la zona en la que se encontraron.

3. Descripción del método propuesto

El método propuesto se inicia con la construcción manual de una ontología preliminar del dominio geográfico en que se representan conceptos, propiedades y relaciones espaciales comunes en este dominio. Luego se ejecutan secuencialmente en tres fases mecanismos de extensión de esta ontología usando Geonames, OSM y BDE. Estos procesos incluyen tareas de extracción de conceptos, relaciones e individuos desde cada fuente. Al final, se lleva a cabo un proceso de generación automática de relaciones espaciales, obteniéndose como resultado la ontología geográfica codificada en OWL2 en la cual se integra información proveniente de las tres fuentes mencionadas. La Figura 1 esquematiza el flujo de trabajo del método.

Teniendo en cuenta el volumen de información a ser integrada en la ontología resultante, y para reducir el consumo de memoria RAM a la hora de procesar esta ontología se diseñó una clase específica llamada *Ontology*, capaz de administrar ontologías con un gran número de individuos y propiedades, y de obtener información de ella sin la necesidad de utilizar razonadores. En esta clase se implementaron funcionalidades que sustituyen a los razonadores en la generación de conocimiento, por ejemplo, para recuperar las subclases e instancias de una clase específica, así como las clases equivalentes, dando la posibilidad de almacenar los valores de las propiedades en una base de datos diseñada para esa tarea y no en el fichero ontológico, y con ello reducir el consumo de memoria en el procesamiento de la ontología. La gestión de la ontología en las diferentes fases se hizo con la versión 4.1.4 del framework OWLAPI, un framework Java diseñado

⁸ <https://wordnet.princeton.edu>

⁹ http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Geographical_coordinates

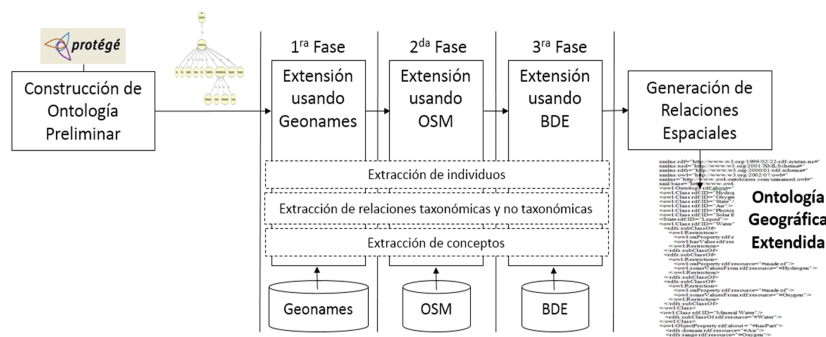


Figura 1. Etapas fundamentales del proceso de generación de la ontología

para gestionar ontologías en OWL2. *Protégé* 4, el editor de ontología de código abierto más popular, por ejemplo, está construido alrededor de OWLAPI.

3.1. Construcción de la ontología preliminar

En este proceso se lleva a cabo la construcción manual de una ontología preliminar del dominio geográfico en la cual se representan conceptos, propiedades y relaciones espaciales comunes dentro de este dominio. En la construcción de la conceptualización de esta ontología se utilizó como referencia fundamental Geonames, la cual posee un elevado nivel de aplicabilidad y constituye una fuente de datos geográficos bien documentada y extensa, cuyo vocabulario básico constituye un buen punto de partida para definir una ontología del dominio geográfico. Específicamente, se utilizó la ontología GeonamesOntology¹⁰, desde donde fueron seleccionados conceptos básicos y propiedades de tipos de datos. También fueron representados en la ontología preliminar conceptos y propiedades identificadas en OSM, así como las propiedades de objetos (ej. 'está al norte de', 'contiene', 'se superpone', 'intersecta con') que son comunes en el estado del arte [2,12,15]. En la Figura 2 se muestra la taxonomía de conceptos representada en esta ontología.

Las relaciones espaciales fueron modeladas de la misma forma que normalmente se hace en la literatura especializada [3,2]. Este aspecto representa una ventaja con respecto a las variantes de integración de datos espaciales basadas en tripletas RDF. Se modelan en la ontología, como propiedades de datos, todas las características de las fuentes de datos propuestas. Se define una nueva propiedad de dato que es la base del futuro proceso de RIG y el mecanismo utilizado en esta investigación para almacenar en una ontología espacial las geometrías de los objetos espaciales recuperados: la cadena WKT¹¹ (*Well Known Text*), la cual es almacenada como un tipo cadena en la ontología. La construcción de

¹⁰ http://www.geonames.org/ontology/ontology_v3.1.rdf

¹¹ <http://www.opengeospatial.org/standards/wkt-crs>

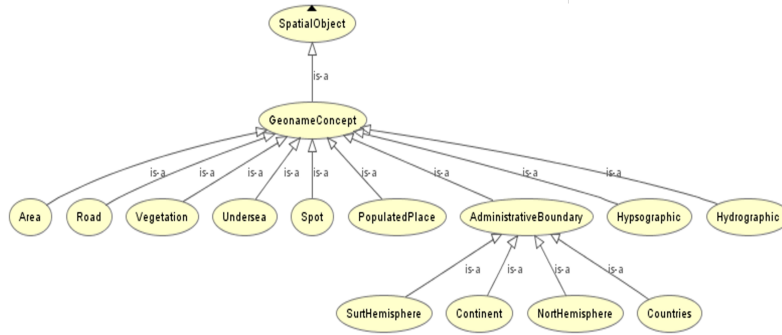


Figura 2. Taxonomía de conceptos de la ontología preliminar

esta ontología preliminar se llevó a cabo usando *Protégé*, siguiendo la propuesta metodológica reportada en [10] y codificada en OWL2.

Como resultado, se obtuvo una ontología formada por 15 clases, 14 relaciones taxonómicas, 15 propiedades de objeto (*owl:ObjectProperty*) organizadas en una taxonomía y que representan relaciones semánticas entre las clases, así como 32 propiedades de tipo de dato (*owl:DatatypeProperty*) también organizadas en una taxonomía, y se han incorporado 25 anotaciones. Tanto en las propiedades objeto como en las de tipo de datos están especificados en su dominio (*rdfs:domain*) todas las clases definidas, brindando así capacidades para alcanzar un nivel alto de descripción de los objetos espaciales que se representen en esta ontología. A partir de esta ontología preliminar el método propone la ejecución de un proceso de extensión a partir de Geonames, OSM y BDE, que se describen a continuación.

3.2. Extensión usando Geonames

En esta primera fase, se lleva a cabo un proceso automático de extensión de la ontología preliminar, a partir de extraer toda la información posible de Geonames. Como parte de este proceso se extraen de la fuente específica nuevos conceptos a ser incorporados a la ontología preliminar, así como propiedades e individuos. Este proceso de extracción de información es guiado por los individuos almacenados en Geonames, es decir, a partir de cada individuo se extraen todos los conceptos con los cuales está relacionado, así como sus relaciones y propiedades. En correspondencia con la información descrita en Geonames, se extiende automáticamente la taxonomía de clases de la ontología preliminar. La extensión incluye la propia descripción (*Annotations* en OWL2) que aparece en Geonames para cada nueva clase. Se extrae información implícita, evitando la utilización de razonadores para inferir conocimiento. Aquí también se identifican y eliminan topónimos ambiguos.

En esta investigación dos topónimos serán considerados ambiguos si la distancia espacial entre sus dos puntos más cercanos es menor que 1000 metros y sus nombres o nombres alternativos son sintácticamente similares según la función

de similitud Damerau-Levenshtein¹². Con el objetivo de considerar la longitud de las cadenas a comparar, la función se utilizó de la siguiente forma: dos cadenas son similares si el resultado de la función es menor o igual que uno, cuando el promedio de las cadenas comparadas es menor o igual a siete, o menor o igual que dos cuando el promedio es mayor que siete. Esta modificación permite ajustar el impacto de las diferencias en las cadenas a comparar en correspondencia con la longitud de las mismas, no es lo mismo una distancia de uno en cadenas de longitud cuatro que en cadenas de longitud 30. La fundamentación del uso de la función de similitud Damerau-Levenshtein está basada en que dos de las fuentes de datos utilizadas en esta investigación son construidas de manera voluntaria por los humanos, por tal motivo los errores tipográficos tienen mayor probabilidad de aparecer en comparación con otros tipos de errores. Cabe señalar que no todos los topónimos de Geonames tienen un *boundingBox* definido, para lo cual se construye la cadena WKT a partir de las coordenadas geográficas del punto que identifica al objeto. Por el contrario, para los casos que tienen definida la propiedad *boundingBox*, la cadena WKT es construida a partir de las coordenadas de los cuatro puntos que definen el propio *boundingBox*. Para acceder a los recursos de Geonames se utilizó la biblioteca de Java geonames-1.1.13¹³, la cual garantiza acceso pleno a los recursos de Geonames a través de servicios web desde el lenguaje de programación Java.

3.3. Extensión usando OSM

Una vez que la ontología preliminar ha sido enriquecida con la información de Geonames, en esta segunda fase se extrae de manera automática toda la información posible de OSM. De forma análoga al proceso anterior, también se identifican en esta fuente conceptos, propiedades e individuos a ser incorporados a la ontología en construcción, igualmente guiado por estos últimos. En correspondencia con la información descrita en [14], se extiende automáticamente la taxonomía de clases de la ontología. La extensión incluye la propia descripción (*Annotations* en OWL2) que aparece en [14] para cada nueva clase. Como parte del proceso de integración de información, también se analizan y reducen posibles ambigüedades tanto en individuos como en clases, usando la misma técnica empleada anteriormente basada en el análisis de similitud sintáctica y distancia espacial. Además, para cada tipo de objeto se extraen todos sus puntos del borde y se genera la cadena WKT a partir de ellos. A diferencia de Geonames, OSM permite generar con exactitud la geometría del objeto espacial, ya que proporciona mecanismos que permiten acceder a los puntos que delimitan el objeto espacial.

El algoritmo de extensión usando OSM que se ha aplicado se muestra a continuación:

¹² https://en.wikipedia.org/wiki/Damerau-Levenshtein_distance

¹³ <http://www.geonames.org/source-code>

```

Ontology ExtensionOSM(Ontology ontology){
    int cantObject= ontology.cantObjetosOSMPertenecientesZona
    Geográfica();
    for i = 1 to cantObject hacer {
        Tnode objetoEspacial = ontology.ObtenerObjetoEspacial(i);
        if (InfoValida(objetoEspacial)){
            string typeClass = identificarClase(objetoEspacial);
            ontology.WriteToponymInOntology(objetoEspacial);
            switch (typeClass){
                case "type1":
                    AnalizarCorrespondenciaConClasesDeOntologíaEn
Construcción(typeClass);
                    BuscaryDefinirIndividuosAmbiguos(objetoEspacial,
listClases_a_Comparar(typeClass));

                    break;
                case "typeN":
                    AnalizarCorrespondenciaConClasesDeOntologíaEn
Construcción(typeClass);
                    BuscaryDefinirIndividuosAmbiguos(objetoEspacial,
listClases_a_Comparar(typeClass));

                    break;
            }
        }
    }
    return ontology;
}

```

InfoEsValida(*objetoEspacial*): define si la información presente en *objetoEspacial* es válida o no.

IdentificarClase(*objetoEspacial*): identifica la clase presente en la ontología a la cual debe pertenecer el objeto espacial *objetoEspacial* o define una nueva clase para él.

AnalizarCorrespondenciaConClasesDeOntologíaEnConstrucción (*typeClass*): identifica la clase *typeClass* en la ontología o construye una nueva clase. En caso de construir una nueva clase, la ubica en la taxonomía de la ontología y define un comentario o *Annotations* para esa nueva clase.

listClases_a_Comparar(*typeClass*): obtiene el listado de clases presentes en la ontología que pudieran tener individuos ambiguos al objeto espacial que se está analizando.

BuscaryDefinirIndividuosAmbiguos: identifica individuos ambiguos en la ontología y los define como similares.

Por último, la gestión de los recursos disponibles en OSM se hizo a través de la biblioteca Java *osm4j*¹⁴, la cual brinda varias ventajas si la comparamos con sus homólogas más conocidas como *Osmosis*¹⁵.

¹⁴ <http://www.topobyte.de/projects/osm4j>

¹⁵ <http://wiki.openstreetmap.org/wiki/Osmosis>

3.4. Extensión usando BDE

El objetivo de esta fase es extraer toda la información posible de una BDE que esté al menos en Primera Forma Normal, e incorporarla a la ontología en construcción, también tratando la ambigüedad de individuos y clases. El mecanismo definido posibilita obtener información e integrarla a la ontología proveniente de cualquier BDE. Se propone para la integración de información proveniente de una BDE la utilización de la cadena WKT como alternativa para representar en OWL2 el campo *Geometry* de los gestores de BDE. Todas las relaciones entre tablas se han modelado como relaciones funcionales del tipo “*HasA*”, definida en la ontología preliminar como una propiedad de objeto funcional. Por último, también se ha llevado a cabo la traducción a OWL2 de restricciones sobre atributos y tablas: condición de Primera Forma Normal, NOT NULL, UNIQUE y PRIMARY KEY.

Toda la estructura organizativa de la BDE es generada automáticamente a partir del modelo físico de la misma. El usuario final solo pudiera intervenir para especificar por cada tabla qué concepto de la ontología es equivalente al concepto asociado a la tabla y el nombre de la columna que representa el nombre de los objetos espaciales. Si este último atributo no se suministra, se asume que es el campo llave y si no se suministra un concepto equivalente se asume que no existe. En este enriquecimiento es posible incorporar nuevas clases e individuos a la ontología, además de equivalencia entre clases, relaciones entre esos individuos, restricciones de datos y de objetos. Una vez ejecuta las reglas de mapeado entre los componentes de la BDE y la ontología en construcción, se analiza la posible ambigüedad de los individuos. Por cada nuevo individuo que esté dentro del área geográfica de interés, se buscan individuos ambiguos en la ontología, incluyendo los de la BDE. El algoritmo utilizado para buscar ambigüedades es similar al utilizado en las fuentes anteriores. Por último, la gestión de la BDE se hace a partir del Java Database Connectivity(JDBC)-9.3-1104 de PostgreSQL¹⁶ y el 1.3.3 de PostGIS¹⁷ para administrar los datos espaciales.

3.5. Generación Automática de Relaciones Espaciales entre Individuos

Una vez integradas en la ontología las tres fuentes de datos, se identifican todos los pares de individuos que cumplen con alguna de las propiedades de objetos definidas en la ontología preliminar, las cuales representan relaciones espaciales. Esta información es almacenada en una base de datos y no en el fichero OWL, usando la clase *Ontology* descrita anteriormente. En este proceso se aprovecha la integración de información realizada en la ontología al culminar las tres fases anteriores, en el cual todos los individuos tienen su geometría en un mismo formato: cadena WKT. Al igual que en la extensión desde Geonames, este módulo es capaz de identificar todos los pares de individuos que cumplen

¹⁶ <http://www.postgresql.org>

¹⁷ <http://postgis.net>

	1ra fase	2da fase	3ra fase	Generación de Rel. Esp.
Clases	667	699	723	-
Clases equivalentes	0	1	11	-
Relaciones taxonómicas o subclases	682	749	956	-
Individuos	47	5.863	8.158	-
Rel. de objeto entre pares de individuos	264	264	264	27.763.887
Prop. de datos con valor en individuos	1.316	48.343	105.054	-
Pares de individuos similares	1	1.216	2.115	12.217
Anotaciones	696	728	728	-

Tabla 1. Resultado de la extensión de la ontología preliminar para el municipio Mariano (La Habana)

con alguna de las propiedades de objeto definidas en la ontología preliminar, sin importar que algunos puedan ser generados posteriormente por los razonadores. Para ello, se apoya en la biblioteca ESRI JAVA Geometry¹⁸.

4. Resultados

A partir de la ontología preliminar construida se realizaron pruebas de los procesos de extensión definidos en el método. Las pruebas se realizaron mediante la ejecución de esos procesos sobre las fuentes seleccionadas, pero considerando solo información relacionada con una región de Cuba, específicamente la del municipio de Mariano de la provincia de La Habana. No obstante, el método está diseñado para procesar información relacionada con cualquier área geográfica sobre la que se almacene información en las fuentes. En la Tabla 1 se muestra como la ontología se va enriqueciendo con la ejecución de los diferentes procesos, obteniéndose al final un recurso ontológico con un gran volumen de información representada; sobre todo muy enriquecido en cuanto a individuos y relaciones semánticas entre ellos. Esto alcanza especial relevancia teniendo en cuenta el propósito de esta ontología (fuente primaria de información en un sistema de RIG) ya que mientras más información sea representada, el sistema de RIG tendrá mayores capacidades para satisfacer las necesidades de información de los usuarios. La inclusión de 728 anotaciones con información descriptiva y complementaria de los elementos representados también es otro aspecto relevante porque facilita el uso de este recurso y, además, esta información también puede ser suministrada a los usuarios como resultado de la RIG.

Adicionalmente, se realizaron pruebas de funcionalidad teniendo en cuenta el propósito de esta ontología. En este sentido, se realizaron 50 consultas sobre la ontología dirigidas a recuperar información referente a objetos espaciales (ej. topónimos), con resultados satisfactorios. En la Tabla 2 se muestran los resultados de cuatro consultas realizadas, las cuales se formalizan mediante la terna:

¹⁸ <http://github.com/Esri/geometry-api-java/wiki>

Id	Requisito de información	Información recuperada	Clases que intervienen
Q1	Recursos hidrográficos en Marianao	Geo, 3547581, Arroyo Marinero; Geo, 3545352, Arroyo Paila; Geo, 3746220, Presa Teresita; OSM, 288952829, Río Orengo; OSM, 223205922, Río Almendares; OSM, 288923290, Río Quibú; ...	Múltiples. La clase <i>Hydrographic</i> y sus descendientes
Q2	Calles que cruzan la localidad de “Los Pocitos”	36 resultados. Aquí solo algunos: OSM, 387758012, Avenida 51; OSM, 369443969, 130; OSM, 366902784, 116; OSM, 38657142, 128; OSM, 351920955, 120; ...	Múltiples. La clase <i>Road</i> , sus descendientes y todas las relacionadas con la propiedad de objeto <i>crosses</i>
Q3	Objetos espaciales en “Los Pocitos”	Geo, 3567880, Arroyo Bañabuey; OSM, 2641153075, Policlínico Docente 27 de Noviembre; OSM, 3268060632, Plaza de Marianao; Geo, 3548916, Los-Pocitos; ...	Múltiples. Todas las relacionadas con el objeto espacial “Los Pocitos” a través de la propiedad <i>inside</i>
Q4	Objetos espaciales en un radio de 3 kms del Hospital Infantil “Juan Manuel Márquez”	628 resultados. Aquí solo algunos: OSM, 258806292, Instituto Tecnológico Militar José Martí ITM; OSM, 367184830, Ciudad Escolar Libertad; Geo, 10114388, San Alejandro; ...	Múltiples. Todas las relacionadas con el Hospital Infantil “Juan Manuel Márquez” a través de la propiedad de objeto <i>isNearOf</i>

Tabla 2. Ejemplos de consultas y resultados

fuente, identificador (en la fuente), nombre del objeto; en el caso de la BDE, los dos últimos elementos de la terna se convierten en uno. En los resultados de cada consulta se puede apreciar como la información recuperada proviene de diferentes fuentes, como intervienen varias clases de la ontología en la obtención de la respuesta, y el requisito de información.

Por último, cabe reseñar que dadas las características de la geontología obtenida donde la información no está íntegramente en el fichero OWL, no fue posible utilizar SPARQL como lenguaje de consultas, utilizándose en su lugar los métodos de razonamiento implementados.

5. Conclusiones

En este trabajo se describe un método computacional para generar, de manera semiautomática, una ontología geográfica que conceptualiza cualquier zona geográfica del mundo con el objetivo de ser utilizada en un sistema de RIG. Los experimentos realizados sobre el municipio de Marianao (La Habana) evidencian que las características de la ontología geográfica resultante son idóneas para el éxito de un futuro sistema de RIG sobre dicho municipio. Aunque solo se hicieron pruebas funcionales utilizando la ontología resultante, los resultados obtenidos demuestran el progreso de características que definen la evaluación cualitativa

de la ontología generada por cada una de las etapas de enriquecimiento y las consultas realizadas confirman el éxito del proceso de RIG.

Agradecimientos

Este trabajo ha sido parcialmente financiado por el Ministerio de Economía y Competitividad del Gobierno de España, proyecto REDES (TIN2015-65136-C2-1-R).

Referencias

1. Baglioni, M., Masserotti, V., Renso, C.: Building geospatial ontologies from geographical databases. In: Proceedings of the Second International Conference on Geospatial Semantics (2007)
2. Egenhofer, M.J.: On the Equivalence of Topological Relations. *International Journal of Geographical Information Systems* 9, 133–152 (1995)
3. Freeman, J.: The modelling of spatial relations. *Computer Graphics and Image Processing* 4(2), 156 – 171 (1975)
4. Hahmann, S., Burghardt, D.: Connecting LinkedGeoData and Geonames in the Spatial Semantic Web. In: 6th International Conference on Geographic Information Science, GIScience 2010 (2010)
5. Hess, G.N., Iochpe, C., Ferrara, A., Castano, S.: Towards effective geographic ontology matching. In: Fonseca, F.T., Rodríguez, M.A., Levashkin, S. (eds.) *GeoS. Lecture Notes in Computer Science*, vol. 4853, pp. 51–65. Springer (2007)
6. Linares Zaila, Y., Montesi, D.: Geographic Information Extraction, Disambiguation and Ranking Techniques. In: Proceedings of the 9th Workshop on Geographic Information Retrieval. pp. 11:1–11:7. GIR '15, ACM (2015)
7. Llano, E.G.: Estado actual de la interpretación semántica de datos espaciales. *Blue Series. Pattern Recognition. Digital version* (2007)
8. Maruf Pasha, Abdul Sattar: Building Domain Ontologies From Relational Database Using Mapping Rules. *International Journal of Intelligent Engineering & Systems* 5(1) (2012)
9. Mustafa, J., Khan, S., Latif, K.: Ontology based semantic information retrieval. In: 4th International IEEE Conference Intelligent Systems. vol. 3, pp. 22–14–22–19 (2008)
10. Natalya F. Noy, Deborah L. McGuinness: Desarrollo de Ontologías-101: Guía Para Crear Tu Primera Ontología (2005), <http://ocw.uc3m.es/ingenieria-informatica/sistemas-avanzados-de-recuperacion-de-informacion/ejercicios/ontology101-es.pdf/view>
11. Patroumpas, K., Giannopoulos, G., Athanasiou, S.: Towards GeoSpatial Semantic Data Management: Strengths, Weaknesses, and Challenges Ahead. In: Proceedings of the 22Nd ACM SIGSPATIAL International Conference on Advances in GIS. pp. 301–310. SIGSPATIAL '14, ACM (2014)
12. Ping, D., Yong, L.: Building Place Name Ontology to Assist in Geographic Information Retrieval. vol. 1, pp. 306–309. IEEE (2009)
13. Sarabia López, G.: Búsqueda y ponderación de información contenida en bases de datos espaciales, utilizando jerarquías. Ph.D. thesis, Instituto Politécnico Nacional de México (2008)

14. Stadler, C., Lehmann, J., Höffner, K., Auer, S.: LinkedGeoData: A core for a web of spatial open data. *Semantic Web* 3(4), 333–354 (2012)
15. Torres Ruiz, M.J.: Representación Ontológica Basada en Descriptores Semánticos Aplicada a Objetos Geográficos. Ph.D. thesis, IPN. CIC (2007)
16. Vega Ramírez, A., Grangel González, I., Sáez Mosquera, I., García Castro, R.: Procedimiento para la obtención de un modelo ontológico para representar la información contenida en bases de datos. In: *Proceedings of the 1st Cuban Workshop on Semantic Web 2014* (2014)
17. Vera Voronisky, F., Garea Llano, E.: Alineamiento de ontologías en el dominio geoespacial. Tech. rep., Centro de Aplicaciones de Tecnologías de Avanzada, CENATAV (2009)