

SISTEMA DE RECUPERACIÓN DE LA INFORMACIÓN EN LA WEB

Integrantes:

- José Augusto Erazo Gómez Joerazogo@unal.edu.co
- Alejandro Gaviria Sepulveda agavirias@unal.edu.co
- Felipe Cortés Norena fcortesn@unal.edu.co

Sistema de recomendación: Búsqueda de alojamiento para turistas

1. Planteamiento del problema.

Un sistema de recomendación es un sistema inteligente que realiza recomendaciones personalizadas a los usuarios sobre un conjunto de elementos. Estos sistemas de recomendación determina, según la característica del usuario y un conjunto masivo de datos con su procesamiento, un subconjunto de elementos que pueden resultar de interés para el usuario. En este caso se tiene la situación de que un turista desea saber qué ciudad escoger en base en unos criterios, como lo es el clima, la condición ambiental, la seguridad, los lugares de entretenimiento, los lugares importantes, etc. Estos criterios son soportados con datos encontrados en datos abiertos en Colombia, cuyo sitio posee registros de cierta variedad de tópicos que cualquier persona con acceso a internet puede observar. El sistema de recomendación de “las mejores ciudades a visitar” es la que nos indica qué ciudades son recomendables para visitar donde posee las calificaciones, los puntajes, las escalas, índices, etc en cada uno de estos criterios. Esto da un indicio de que el turista puede recomendar una ciudad o no según sus expectativas, su decisión y sobre todo su experiencia en dicha ciudad escogida.

Cabe mencionar que estos datos de algún modo hay que estructurarlos de tal forma que, cuando un cliente acceda a ellos por medio de una aplicación, pueda leerlos con facilidad y de este modo familiarizarse con la aplicación, lo cual, la estructura será por medio de ontologías, que es un modo de representar el conocimiento en la que nos permite encontrar información de manera precisa para poder realizar deducciones y toma de decisiones. Además, esta información alimentará tres sitios de publicación de *linked data*, en este caso serán **dbpedia**, **D2R Server** y **Virtuoso** y la información almacenada en estos tres sitios será integrada en un nuevo repositorio. Es un cuarto sitio donde debe recoger e integrar información en tiempo real de estos tres sitios mencionados anteriormente. Dicho sitio de publicación se implementará haciendo uso de JENA y PELLET para llevar a cabo una serie de consultas haciendo uso de sparql (lenguaje utilizado para extraer información de ontologías).

2. Objetivos

2.1 Objetivo general

Diseñar un sistema de recomendación sobre las mejores ciudades por visitar en el país en base a unos criterios, dicho diseño parte de ontologías donde su estructura es mucho más flexible y además sea fácil de acceder para cualquier persona que esté interesada y de esta forma le facilite la toma de decisiones.

2.2 Objetivo específico

- Diseñar un sistema de recomendación sobre las mejores ciudades por visitar en el país en base a unos criterios (lugares, hoteles, clima, etc).
- Representar correctamente los elementos del sistema definiendo sus propiedades e interacción entre ellos mismos y con dicho sistema.

3. Marco Teórico

3.1 Estado del arte y antecedentes

El turismo es un rama de la actividad económica que representa aproximadamente el 10% del PIB mundial y esta participación viene creciendo en los últimos decenios. Las cifras que anualmente registra el ingreso mundial por turismo internacional refleja la importante contribución a la dinámica económica que esta actividad representa para muchas regiones en el mundo. Entre sus principales impactos positivos encontramos los relacionados con el ingreso por divisas, la contribución al ingreso privado y público, la generación de empleo, el incentivo a la creación de tecnologías y a la formación de capital humano y las oportunidades de negocio que puede generar en una economía.

El turismo es una área que ha experimentado un gran cambio en los últimos años, debido principalmente a la aparición de nuevas tecnologías (internet, web social, dispositivos móviles con GPS...) y a la aparición de información geográfica accesible de forma gratuita y al alcance de todos (Openstreetmaps, Google Maps, Bing Maps...). En la última década, muchos sitios web populares de turismo en línea ofrecen servicios a los usuarios, por ejemplo, reservas de hotel, planificación de viajes, etc. A pesar de esto, estos servicios parecen ser deficientes porque llevan a los usuarios a una situación en la que tienen dificultades para elegir los servicios adecuados con su perfil y su preferencia real. Estos usuarios pueden dedicar mucho tiempo y esfuerzo en encontrar lo que buscan. Además, los turistas desean descubrir nuevos sitios y culturas del país o de la ciudad visitada por su patrimonio material e inmaterial. Por lo tanto, la integración de sistemas inteligentes como los sistemas de recomendación es una operación obligatoria en el ámbito turístico.

Los sistemas de recomendación ofrecen una solución eficiente a los problemas de sobrecarga de información (los resultados contiene información demasiado condensada y mal estructurada, por lo tanto, el usuario puede perderse en esta masa de información que debe filtrarse a sí misma), su función principal es seleccionar o filtrar la información en esta masa de

información para el usuario de manera transparente y proporcionar sugerencias para temas de interés con que el sistema proporcionará al usuario una lista de recursos de acuerdo con sus preferencias para satisfacer sus necesidades.

4. Antecedentes

En este punto se realizó un cuadro comparativo entre diferentes trabajos que cuyo tema principal también fue llos sistemas de recomendación en turismo

Trabajo	Hacia la mejora de la creación de rutas turísticas a partir de información semántica	Ontología Turística Geográfica Memoria del proyecto	Ontologías Turísticas Geográficas: Creación de una Ontología sobre Rutas Turísticas (a Pie o en Bicicleta) por Espacios Naturales.
Pregunta problema	Cómo representar la información turística de manera que sea fácilmente interpretable por programas de cálculo de rutas.	Cómo crear, usando tecnologías de web semántica, de una ontología de carácter turístico que permita clasificar y relacionar documentos digitales en internet con puntos de interés turístico, y la creación de un buscador web que permita hacer consultas sobre un grupo de datos que utilicen la ontología creada	Cómo desarrollar un marco de trabajo que permita apoyar al usuario en sus rutas turísticas, guiándolo en función de su posición (proporcionada por su dispositivo móvil), sus preferencias y el momento del día en que se encuentre, lo que se denomina como Location Based Services (LSB), pudiendo recomendar al usuario información personalizada con respecto a sus preferencias o características como viajero.
Población objetivo	Turistas que deseen recorrer Valls d'Aneu	Turistas que deseen recorrer Mallorca	Turistas españoles.
Estrategias de solución	Se presentó el concepto de ontología, sus principales características y ventajas. Luego se enumeraron diferentes	Se creó una ontología turística geográfica. Se usó un buscador de recursos turísticos. Desarrollar una api de web semántica.	Desarrollo de la Ontología. Enlazar con otras ontologías.

	<p>ontologías de carácter general, turísticas y del proyecto Open Linked Data que podrían usarse de soporte en una aplicación SIG del ámbito del turismo. También se presenta la arquitectura de una aplicación SIG que permite ofrecer rutas personalizadas en función de las preferencias de usuario (horario de la visita, dispositivo móvil y lugar a visitar) y el papel que juegan las ontologías en dicha arquitectura.</p>	<p>Para la creación del prototipo de buscador se usó Java-J2EE, utilizando Java Server Faces para la presentación de las páginas en el cliente y la comunicación con el servidor.</p> <p>Para la gestión del modelo semántico se utilizó la API de web semántica Jena, que cumple los requerimientos indicados en el diseño de la aplicación.</p>	
Resultados	<p>Las ontologías de alto nivel son una opción cuando se necesitan ontologías muy formales, pero su magnitud y dificultad de manejo las desaconsejan en gran medida.</p>	<p>Aún siendo un ejemplo muy simple, muestra unas pocas funcionalidades interesantes que sirven como ejemplo de cómo trabajar con una ontología.</p>	<p>se demostró la utilidad de las tecnologías relacionadas con la Web Semántica a la hora de facilitar la integración de información en diferentes formatos, proveniente de diversas fuentes, utilizando ontologías para modelar las entidades fundamentales dentro de un ámbito concreto (en este caso, de las rutas turísticas) e identificar las relaciones existentes entre ellas, facilitando al usuario la tarea de búsqueda de información.</p>

5. Plan de temas y cronograma

En la siguiente tabla se establece la actividad y persona asociada a dicha actividad en un orden cronológico

Personal	Actividad	Descripción	Justificación
Alejandro Gaviria	Realización y búsqueda de artículos	Realización y búsqueda de artículos para apoyo a la sustentación de la problemática de los datos enlazados en Colombia	Investigaciones realizadas por los expertos de la web semántica estableciendo en sus artículos los retos a afrontar
Alejandro Gaviria, José Erazo	Diseño, creación y corrección de la Ontología	Se crea la ontología donde se va a estructurar la información, cuyos datos se consiguen son extraídos en datos abiertos.	Se busca los datos en las páginas de datos abiertos con el fin de estructurarlos por medio de ontologías con el fin de que sea más flexible en su manejo
José Erazo y Alejandro Gaviria	Montar la ontología virtuoso y D2R	Luego de crear la ontología y estructurar correctamente la información se sube a virtuoso y D2R, que es un servidor que nos permite entre muchas opciones, tener nuestro propio endpoint SPARQL para realizar consultas sobre una base de datos semántic	Virtuoso es una base de datos para grafos más completas actualmente.
Alejandro Gaviria	Hacer las consultas en sparql	las consultas son las que nos mostrará que la información que deseemos sacar de la ontologías	Sparql es el lenguaje para base de datos basada en grafos donde nos permitirá determinar la información de interés a través de consultas
Felipe Cortés	Implementación del razonador	Jena interface Api, será una api que realiza en java y que nos permite determinar el razonamiento de la ontología diseñada	Jena interface Api, es una de las herramientas más importante que proporciona Jena en sus librerías, es el

			Inference API o API de razonamiento
Alejandro Gaviria	implementación del sistema de recomendación	Se implementa uno de los sistemas de recomendación explicadas en el curso.	Se implementar un sistema de recomendación con el fin de determinar la mejor manera de evaluar las cosas que se van recomendar o no.
Alejandro Gaviria	Subida de la aplicación en la nube	Se utilizó amazon web service	La parte del servidor es para las personas que quieran acceder a la página web e interactúe con ella con el fin de que se sienta satisfecha por las recomendaciones que se implementan que esté a su gusto.

6. Bibliografía

La bibliografía elegida describe los trabajos realizados basados en la web semántica donde se determina claramente el proceso de organización de la información y sobre todo su flexibilidad en el manejo de la misma, básicamente en la extracción de información y su posterior son las que conducen la toma de decisiones correctas. Además se resalta claramente que al ser más flexible la estructuración de la información, se hace más fácil el acceso a ellas y de hecho el análisis minucioso sobre ellas puede resultar más cómodo.

- [ieeexplore-ieee-org.ezproxy.unal.edu.co](https://ieeexplore-ieee-org.ezproxy.unal.edu.co/stamp/stamp.jsp?tp=&arnumber=8302291). (2018). [online] Available at: <https://ieeexplore-ieee-org.ezproxy.unal.edu.co/stamp/stamp.jsp?tp=&arnumber=8302291> [Accessed 18 Apr. 2018].
- [ieeexplore-ieee-org.ezproxy.unal.edu.co](https://ieeexplore-ieee-org.ezproxy.unal.edu.co/stamp/stamp.jsp?tp=&arnumber=6511683). (2018). [online] Available at: <https://ieeexplore-ieee-org.ezproxy.unal.edu.co/stamp/stamp.jsp?tp=&arnumber=6511683> [Accessed 18 Apr. 2018]
- (2018). [ebook] Available at: http://openaccess.uoc.edu/webapps/o2/bitstream/10609/2284/1/igutierrezl_articulo.pdf [Accessed 1 Jun. 2018].

