

# Buscador facetado en conjuntos de datos de la Linked Open Data Cloud a través del reconocimiento del lenguaje natural



CONACYT

<sup>1</sup>López Ochoa, Betia Lizbeth; <sup>2</sup>Sánchez Cervantes, José Luis

<sup>1</sup>Maestría en Sistemas Computacionales, División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Orizaba

<sup>2</sup>CONACYT- Instituto Tecnológico de Orizaba

## INTRODUCCIÓN

Comúnmente las aplicaciones Web incorporan interfaces limitadas a la captura de datos a través de teclado, ratón o pantallas táctiles y a la visualización de los resultados obtenidos generando una interacción poco intuitiva, por otro lado, existe una brecha entre los sistemas para navegar y explorar datos semánticos disponibles en la LOD (Linked Open Data) *cloud* y los usuarios, ya que los mecanismos actuales para la recuperación de información semántica son orientados a ser utilizados por personas experimentadas.

## OBJETIVO

Desarrollar un buscador facetado que aplique NLP (*Natural Language Processing*, Procesamiento de Lenguaje Natural) para el reconocimiento de comandos de voz que permita la realización de búsquedas y navegación facetada en bases de conocimiento de dominio médico que formen parte de la LOD *cloud*.

## RESULTADOS

**Aplicación Web FASELOD** (Faceted Search in Linked Open Datasets) desarrollada en el lenguaje Java que permite la obtención de información médica desde bases de conocimiento de la LOD *cloud* tales como DrugBank, ClinicalTrials.gov, DailyMed, Orphanet y PubMed.

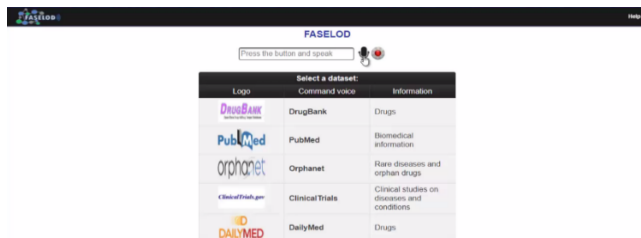


Figura 1 Aplicación Web FASELOD, selección de dataset

El principal medio de interacción con FASELOD es el lenguaje natural a través del reconocimiento de comandos de voz, donde cada comando permite realizar una determinada acción.

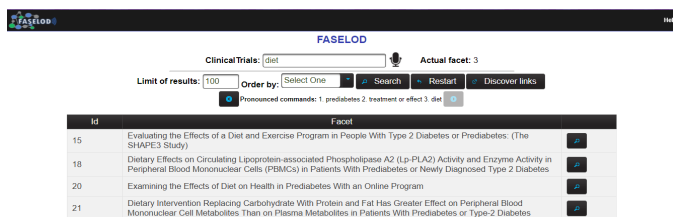


Figura 2 Búsqueda facetada en FASELOD

FASELOD permite guiar al usuario mediante los resultados obtenidos en cada iteración hasta que obtenga el o los resultados esperados, con este tipo de navegación se eliminan resultados ambiguos, repetitivos o que no cumplen con los requisitos de búsqueda. Los usuarios tienen la opción de realizar descubrimiento de información adicional en DBpedia (Figura 2) obteniendo un conjunto de enlaces relacionados con los criterios de búsqueda, dichos enlaces son abiertos de forma externa a la aplicación.

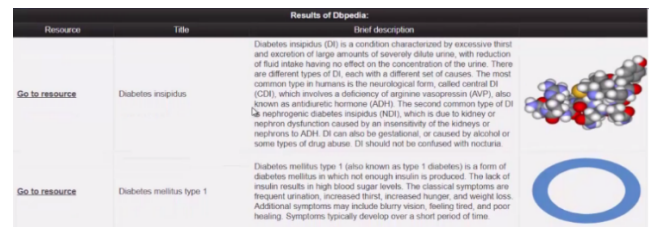


Figura 3 Descubrimiento de enlaces

## CONCLUSIONES

En este trabajo se obtuvieron distintos beneficios como la adquisición de información y navegación de forma sistemática entre grandes cantidades de información que no es posible que se procesen sin apoyo automático, así como la utilización de comandos de voz como principal medio de interacción que permiten una navegación más natural e intuitiva. Finalmente, mediante el uso de facetadas se proporciona un listado de resultados a través de los cuales el usuario navega hacia otros datos relacionados con su tema de búsqueda eliminando resultados ambiguos, repetitivos o que no cumplen con los requisitos de búsqueda.



**Betia L. López Ochoa, José L. Sánchez Cervantes, Giner Alor Hernández, Ma. Antonieta Abud Figueroa, Beatriz A. Olivares Zepahua, Lisbeth Rodríguez Mazahua,** “An Architecture based in Voice Command Recognition for faceted search in Linked Open Datasets”, Springer Press. ISBN: 978-3-319-69341-5.



**Sánchez Cervantes J. L. & López Ochoa B. L.** “FASELOD: Buscador facetado en conjuntos de datos de la Linked Open Data cloud a través del reconocimiento del lenguaje natural”. INDAUTOR.



E. Rodríguez, “Presentan en Zacatecas software innovador”, CONACYT Agencia Informativa, 2017.