

UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

SISTEMAS INFORMÁTICOS Y COMPUTACIÓN

SISTEMAS BASADOS EN CONOCIMIENTOS



Autor:

David Alejandro Burneo Valencia

David Josue Jiménez Jiménez

Karla Lizbeth Ochoa Ludeña

Docente:

Ing. Janneth Chicaiza

Tema :

PROYB1-1 DEFINICIÓN DE DOMINIO Y
APLICACIÓN PARA EL PROYECTO

Abr/2021 – Ago/2021

1. Dominio del trabajo

Como dominio de trabajo para realizar la búsqueda de información se propone el tema de COVID-19, delimitado en las publicaciones referentes a los casos de trombosis presentados como efectos secundarios de las campañas de vacunación.

- Covid-19 & vaccine thrombosis

2. Fuentes de datos

Para el proceso de elección de fuentes de datos se ha procedido a verificar varias páginas web destinadas a la publicación de revistas científicas e investigaciones, con el fin de seleccionar las mejores para la temática escogida, además de que dicha información pueda ser procesada y transformada a RDF.

2.1.Semantic Scholar

2.1.1. Descripción

Semantic Scholar es un motor de búsqueda respaldado por un sistema de inteligencia artificial dedicado a la investigación de publicaciones académicas. Fue desarrollado por el Instituto Allen de Inteligencia Artificial y lanzado al público en noviembre de 2015. Utiliza los últimos avances en el procesamiento del lenguaje natural para proporcionar resúmenes de artículos académicos.

Se trata de un proyecto basado en la aplicación de técnicas de inteligencia artificial en la gestión de búsqueda de publicaciones científicas, con el objetivo de descubrir los artículos científicos más relevantes asociados a un tema en particular. Se basa en la combinación de metodología y técnicas utilizadas en inteligencia artificial como el aprendizaje automático, entre otros, para entender el contexto en el que se dan estas frases para filtrar los resultados obtenidos. (Wikipedia, 2021)

2.1.2. Licencia

Acuerdo de licencia de conjunto de datos y Api de Semantic Scholar
<http://s2-public-api-prod.us-west-2.elasticbeanstalk.com/corpus/legal/>

2.1.3. Características

- DOI
- ID

- CorpusID
- Abstract
- NombrePublicación
- AñoPublicación
- Autor
- RevistaPublicación
- Área
- Referencias
- URL

2.1.4. Resultados

Al aplicar la búsqueda en la presente fuente de información se obtiene un numero considerable de resultados (Ilustración 1), con la cual se espera extraer la información relevante.

The screenshot shows the Semantic Scholar search interface. At the top, the search bar contains the text "covid-19 vaccine thrombosis". Below the search bar, it indicates "About 15,200 results for 'covid-19 vaccine thrombosis'". The results are filtered by "Fields of Study", "Date Range", "Has PDF", "Publication Type", "Author", and "Journals & Conferences". The results are sorted by "Relevance".

The first result is titled "AstraZeneca COVID-19 vaccine: thrombosis with thrombocytopenia" by "anonymous" in "Reactions Weekly" from 1 April 2021. It includes a PDF icon, a "View on Springer" link, and options to "Save", "Alert", "Cite", and "Research Feed".

The second result is titled "Updated Recommendations from the Advisory Committee on Immunization Practices for Use of the Janssen (Johnson & Johnson) COVID-19 Vaccine After Reports of Thrombosis with Thrombocytopenia Syndrome..." by "J. MacNeill, John R Su, +14 authors" in "MMWR. Morbidity and mortality weekly report" from 30 April 2021. It includes a "View on PubMed" link and options to "Save", "Alert", "Cite", and "Research Feed".

The third result is titled "Thrombocytopenia and Intracranial Venous Sinus Thrombosis after 'COVID-19 Vaccine AstraZeneca' Exposure" by "M. Wolf, B. Luz, L. Niehaus, P. Bhogal, H. Bickner, H. Henkes" in "Journal of clinical medicine" from 1 April 2021. It includes a "View PDF" link and options to "Save", "Alert", "Cite", and "Research Feed".

The fourth result is titled "A Rare Case of Cerebral Venous Thrombosis and Disseminated Intravascular Coagulation Temporally Associated to the COVID-19 Vaccine Administration" by "V. d'Agostino, E. Caranci, +9 authors" in "Journal of personalized medicine" from 1 April 2021. It includes a "View PDF" link and options to "Save", "Alert", "Cite", and "Research Feed".

Ilustración 1. Resultados Semantic Scholar

2.2. Crossref

2.2.1. Descripción

Es la agencia más grande de registros de DOI; es una asociación de editoras científicas que promueve el uso de las nuevas tecnologías para mejorar la comunicación y la investigación científica; trabaja en la mejora de la identificación de las publicaciones. (Morcillo, 2018)

2.2.2. Licencia

El contenido de este sitio está sujeto a una licencia Creative Commons Attribution 4.0 International License

2.2.3. Características

- AñoPublicacion
- DOI
- URL
- Autor
- NombrePublicación
- Licencia
- Disciplina
- Referencias
- UltimaEdición
- Abstract

2.2.4. Resultados

De igual manera que en la fuente de información anterior, se aplicó la búsqueda con el fin de obtener el número de resultados obtenidos. Sin embargo, para reducir el numero total de resultados se aplico a filtrar los datos en función del año (2021) y a tipo de información (Journal Article), con lo cual se obtiene el numero de resultados visualizados en la Ilustración 2.

Crossref

Search: covid-19 vaccine thrombosis

Sort by: RELEVANCE PUBLICATION YEAR

TYPE: Journal Article (73,341)

YEAR: 2021 (73,326)

PUBLICATION: SSRN Electronic Journal (1,017)

PUBLISHER: Elsevier BV (16,251)

Results:

- AstraZeneca COVID-19 vaccine: thrombosis with thrombocytopenia**
Journal Article published Apr 2021 in Reactions Weekly volume 1851 issue 1 on pages 1 to 2
<https://doi.org/10.1007/s40278-021-94091-4>
- A Case Report of Acquired Haemophilia following COVID-19 Vaccine**
Journal Article published 30 Mar 2021 in Journal of Thrombosis and Haemostasis
Authors: Mansoor Radwi, Sara Farsi
<https://doi.org/10.1111/jth.15291>
- Driving to Comply: Mind Genomics, Arizona, and the COVID-19 Vaccine**
Journal Article published 13 Jan 2021 in Journal of International Perspectives on COVID-19 volume 1 issue 1
<https://doi.org/10.31038/jipc-2021111>
- Developing standard safety outcomes for COVID-19 vaccines**
Journal Article published Mar 2021 in Vaccine
Authors: Jim P. Buttery
<https://doi.org/10.1016/j.vaccine.2021.03.004>
- Letter to the Editor: Winter is coming and COVID-19 vaccine is available! The role of gastroenterologist in increasing COVID-19 vaccine acceptability among IBD patients.**
Journal Article published Jan 2021 in Gastroenterology
Authors: Alfredo Papa, Antonio Gasbarrini, Loris Riccardo Lopetuso
<https://doi.org/10.1053/j.gastro.2020.12.066>

Ilustración 2. Resultados Crossref

3. Aplicación a construir para la explotación del grafo de conocimiento

3.1. Búsqueda (Elastic Search)

Es un motor de analítica y análisis distribuido, enfocado en todos los tipos de datos tanto estructurados, no estructurados, entre otros. Es de código abierto creado en Apache Lucene y desarrollado en Java. Permite almacenar, buscar y analizar grandes volúmenes de datos de forma rápida y en tiempo real. (Elastic, 2021)

3.1.1. Funcionamiento

Los datos sin procesar fluyen hacia Elasticsearch desde una variedad de fuentes, incluidos logs, métricas de sistema y aplicaciones web. La ingesta de datos es el proceso mediante el cual estos datos son parseados, normalizados y enriquecidos antes de su indexación en Elasticsearch. Una vez indexados en Elasticsearch, los usuarios pueden ejecutar consultas complejas sobre sus datos y usar agregaciones para recuperar resúmenes complejos de sus datos. Desde Kibana, los usuarios crean visualizaciones poderosas de sus datos, comparten dashboards y gestionan el Elastic Stack.

3.1.2. Frameworks/librerías/paquetes

- ElasticSearch
- Logstash

- Kibana
- Spring data
- Curl

3.2.Semantic Scholar

Para la extracción de los metadatos en Semantic Scholar se empleará la librería de Python <semanticscholar>

3.3.Crossref

Para la extracción de los metadatos en Crossref se empleará la librería de Python <crossrefapi> La cual esta enfocada en buscar y extraer la información solicitada a través de iteraciones en los resultados obtenidos y devolver la información

Bibliografía:

- Wikipedia. (2021). Semantic Scholar. 29/04/2021, de Wikipedia Sitio web: https://es.wikipedia.org/wiki/Semantic_Scholar
- Yolanda Morcillo. (2018). CROSSREF: UNA HERRAMIENTA MULTITAREA. 29/04/2021, de Investiga UNED Sitio web: <http://investigauned.uned.es/crossref-una-herramienta-multitarea/>
- Elastic. (2021). ¿Qué es Elasticsearch?. 29/04/2021, de Elastic Sitio web: <https://www.elastic.co/es/what-is/elasticsearch>