

# Actividad Bases de Datos

Santiago Montenegro Woodcock  
Fundación Universitaria Konrad Lorenz

**Abstract**—En la actualidad existen varias herramientas tecnológicas que permiten que una persona tenga acceso a una basta lista de bases de datos que contienen información relacionada con una gran cantidad de temas. Sin embargo, al realizar una búsqueda en una de estas bases de datos, esta nos puede retornar una cantidad excesiva de artículos.

En respuesta a lo anterior, se han creado métodos y herramientas para el análisis de datos con el fin de anexar filtros a la búsqueda a realizar, lo que por consiguiente limitará la cantidad de información ofrecida, volviéndose más específica y efectiva.

## I. INTRODUCCIÓN

Hoy en día las personas tienen acceso a una basta cantidad de fuentes de información gracias a las bases de datos que podemos encontrar en internet, cada una con una gran cantidad de libros, artículos, revistas; que podemos usar para indagar sobre uno o varios temas.

Sin embargo, a la hora de realizar una búsqueda, las bases de datos nos pueden ofrecer una gran cantidad de resultados de los cuales no es fácil saber si nos servirán para dar respuesta al tema de interés por parte del usuario.

En respuesta a eso, se han desarrollado diferentes tipos de análisis de datos, cada uno enfocado en diferentes objetivos para facilitar la filtración de información obtenida a la hora de realizar una investigación.

En este artículo, realizaremos un proceso de análisis de datos con el fin de obtener los mejores documentos relacionados con el tema de “Free Software” o “Software Libre” como concepto contenidos en la base de datos “Scopus” mediante el apoyo de una librería para el lenguaje de programación “Python”.

## II. RESULTADOS

En esta sección podremos encontrar la documentación sobre los procesos que se siguieron para realizar un análisis de datos teniendo como base el tema de “free software”, iniciando con una cantidad de 61229 documentos proporcionados por la base de datos “Scopus” y concluyendo con 3 artículos [1], [2], [3], los cuales se seleccionaron de acuerdo al propósito de la búsqueda.

### A. Flujo del Analisis

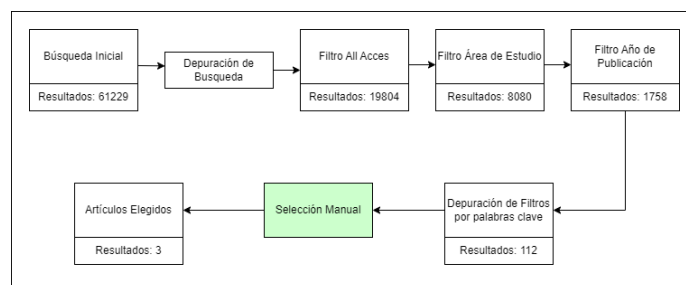


Fig. 1. Query Exploratoria.

### B. Query Exploratoria

Para crear la “Query Exploratoria” tomo en cuenta las palabras principales que deberán encontrarse en los artículos a buscar (free AND software), limitando los resultados a solo los que sean del tipo “open access” para el caso de ser elegidos ciertos artículos pueda tener acceso al contenido completo.

Continuado con el filtrado, limito los resultados a los que estén relacionados con las áreas de estudio sobre informática (computer science), ingeniería (engineering) y matemáticas (mathematics), y concluyendo con un filtro de año de publicación (2023, 2022 y 2021) asegurando que la información que contengan los artículos esté actualizada.

Finalizado con la query representada en la Fig. 2.

TITLE-ABS-KEY ( free AND software ) AND ( LIMIT-TO ( OA,"all" ) ) AND ( LIMIT-TO ( SUBJAREA,"COMP" ) OR LIMIT-TO ( SUBJAREA,"ENGI" ) OR LIMIT-TO ( SUBJAREA,"MATH" ) ) AND ( LIMIT-TO ( PUBYEAR,2023 ) OR LIMIT-TO ( PUBYEAR,2022 ) OR LIMIT-TO ( PUBYEAR,2021 ) )

Fig. 2. Query Exploratoria.

### C. Analisis de Titulo

Al analizar la gráfica de la Fig.3, podemos observar que palabras de peso en la investigación como “analysis”, “data” y “model” están presentes en una buena cantidad de los artículos. Sin embargo, también podemos observar palabras relacionadas al aprendizaje como lo son “study” y “learning” no se quedan atrás estando cerca de la cantidad de apariciones de las palabras anteriormente mencionadas.

En este caso, ya que la investigación se centra en “free software” de un modo conceptual, tomaremos como palabras de filtro a “análisis”, “model” y “learning”.

