

Análisis Bibliografico Scopus

Andrés Salcedo Vera
Fundacion Universitaria Konrad Lorenz

Resumen—En el presente trabajo se documentará el análisis bibliográfico de artículos científicos de la base de datos Scopus, se buscará hacer un análisis profundo en los temas de

Se va a tocar temas como análisis de títulos, abstract, análisis de autores y demas, para poder entender la información suministrada y cuál puede ser la más relevante para el caso del proyecto actual.

I. INTRODUCCIÓN

para poder capturar los mejores resultados en Scopus, se necesita primero tener un tema de referencia, con ello trabajaremos, luego dentro de scopus realizaremos el primer filtro (este tema será tocado más adelante en el artículo en flujo de análisis).

la idea principal es encontrar entre los 1.000 y los 2.000 documentos, el problema si son mas es que Scopus no los descarga directamente, y se supone que demoran 24 horas en dar el link de la descarga, pero no suele ser asi.

II. RESULTADOS

en este trabajo se logró bajar el query de 1750 resultados a 8 lo cual ha sido muy bueno, se puede obtener con una limpieza rápida los documentos más interesantes y más útiles según el prejuicio que se desea encontrar.

Es importante notar cada uno de los análisis hechos, todos son importantes y pueden aplicarse al filtro para obtener menos filtros si llegara a necesitar del caso, cada uno puede obtener diferentes artículos, y cada uno puede ser más útil que otro, en el caso en el que se busque algo más específico también se puede aplicar al filtro para obtener el query más adecuado.

II-A. Flujo del Analisis

El filtro estará determinado en los últimos 5 años, también se hará el filtro sobre ciencias de la computación e ingeniería.

II-B. Query Exploratoria

TITLE-ABS-KEY (intelligent AND assistant) AND (LIMIT-TO (PUBYEAR , 2023) OR LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019)) AND (LIMIT-TO (SUBJAREA , ÇOMP") OR LIMIT-TO (SUBJAREA , .^{EN}GI"))

II-C. Análisis de Título

en los títulos podemos encontrar que using, detection, sytem, based y data son los más repetidos, por lo que sí es de nuestro interés hacer la query con esas palabra clave nos podría funcionar.

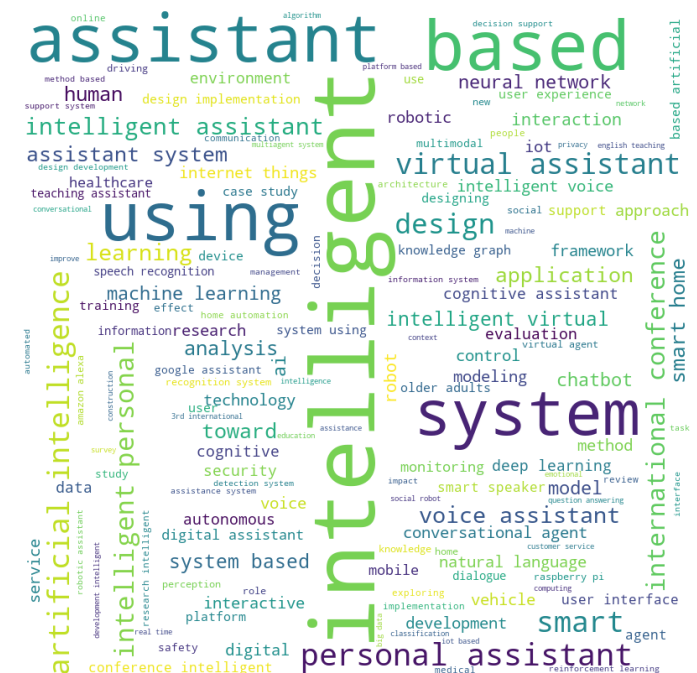


Figura 1. nube de palabras de títulos

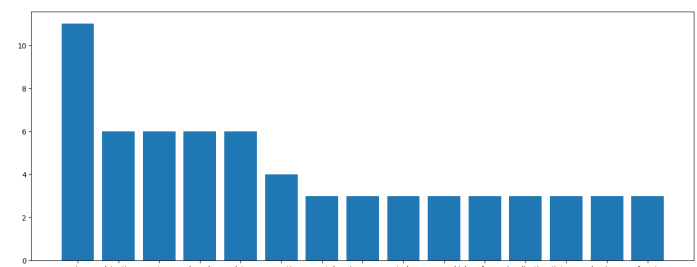


Figura 2. nube de palabras de títulos

II-D. Análisis de Keyword

Las keywords son las palabras más esenciales para el trabajo por hacer, es importante tenerlas muy en cuenta para una query más apropiada, ya que estas son la definición más precisa de que va a tratar el trabajo y de allí podremos sacar más información.

en este caso las keywords más importantes son computing, data, high performance y learning.

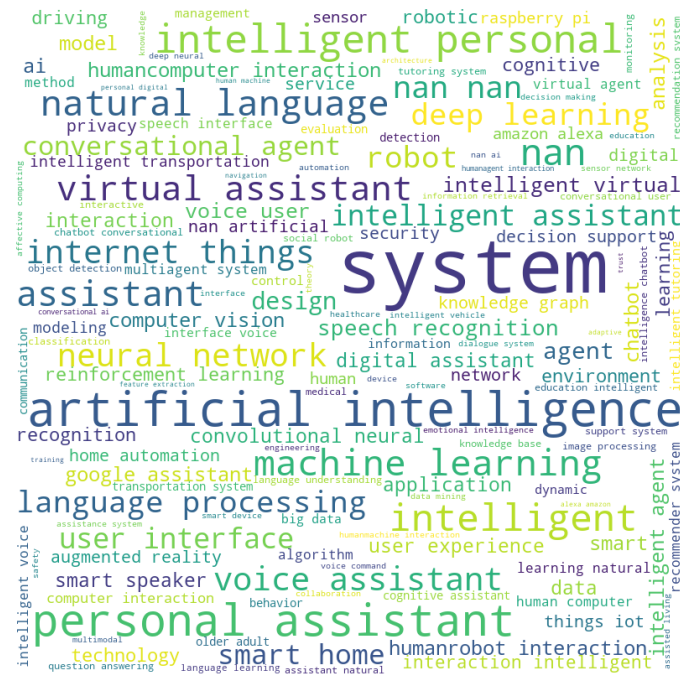


Figura 3. nube de palabras de las keywords

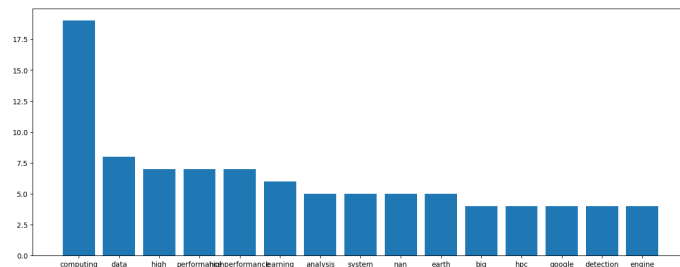


Figura 4. grafico de las keywords

II-E. Análisis de Abstract

El abstract es el siguiente paso más importante, ya que de aquí sacaremos nuestro próximo query, por lo que se usaran las palabras data, computing, high performance, map y results para obtener la query más optimizada, este será el filtro más importante ya que marcara la relevancia del artículo en nuestra propia investigación.

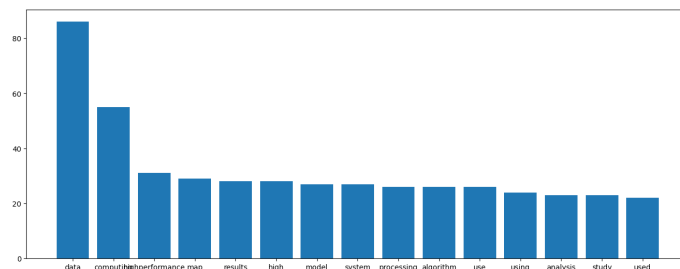


Figura 5. grafico del abstract



Figura 6. nube de palabras del abstract

II-F. Ranking de palabras clave para la construcción del nuevo filtro

Para filtrar mejor nuestra query, se decidió utilizar el top 5 de las palabras del abstract, este da el contexto más profundo del artículo que se desea encontrar.

las palabras clave más importantes fueron:

1. data
2. algorithm
3. performance
4. system
5. detection

II-G. Query Resultante

Este es el query resultante.

TITLE-ABS-KEY (intelligent AND assistant AND algorithm AND data AND performance AND detection AND system) AND (LIMIT-TO (SUBJAREA , ÇOMP") OR LIMIT-TO (SUBJAREA , .ENGI")) AND (LIMIT-TO (PUBYEAR , 2023) OR LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2019))

II-H. Otros Análisis

tambien se decidió hacer análisis de datos más a profundidad del dataset que tenemos, como los autores, los journals con más cantidad de artículos encontrados, y los años con más cantidad de artículos subidos.

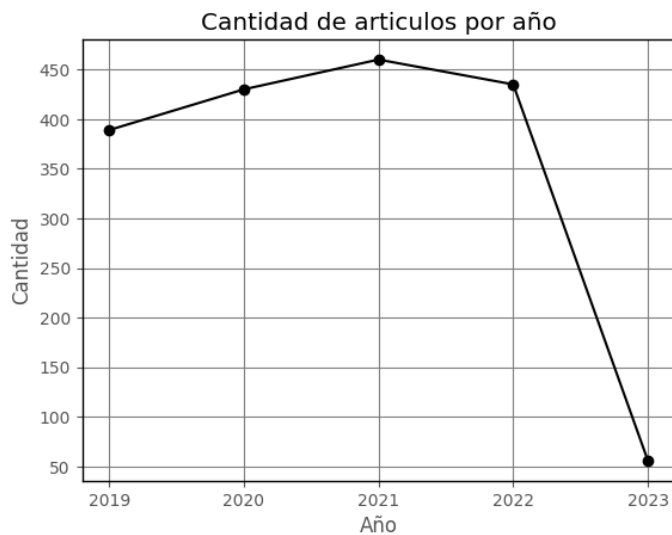


Figura 7. grafico de los años con más publicaciones

el año con más publicaciones sin duda es el 2021, pero la gráfica no debe engañar, aún estamos actualmente a principios del 2023 lo cual aún da mucho tiempo para poder llegar a una cantidad satisfactoria y comparable de artículos como lo ha sido los demás años.

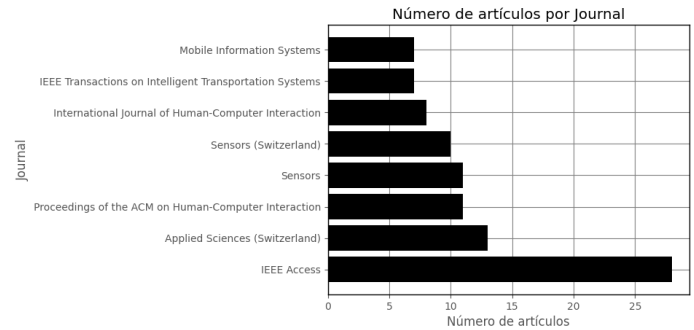


Figura 9. Cantidad de artículos por journal

con los artículos es muy interesante, se puede ver que hay bastantes journal que siguen creyendo en los proyectos de las personas y los publican, y algo más interesante aun es que si hay tantos por journal, eso significa que hay muchos proyectos por terminar o interesantes para diferentes ramas de las ciencias.

III. CONCLUSIONES

Hay muchos artículos interesantes, y muchos temas de investigación y análisis, en este caso el query logro reducirse mas de un 90% por lo que resulta muy util, para mas informacion como los notebooks que se usaron para el proyecto se puede acceder al siguiente link para ver todo el proceso.

<https://github.com/AndresDardex/BasesDeDatos>

IV. BIBLIOGRAFIA

OpenAI. (2021). ChatGPT: A large language model.

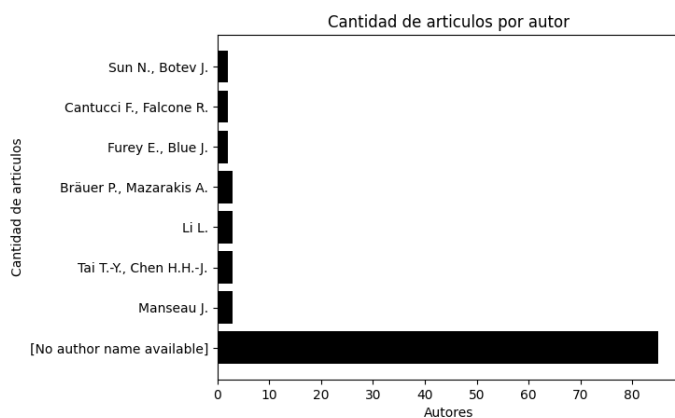


Figura 8. Grafica de los autores

algo muy curioso es que en la base de datos más de 80 de los artículos no tienen registrado un autor, lo cual es realmente triste porque hacer este tipo de artículos requieren de mucho esfuerzo, investigación y dedicación, por lo que no tener al menos su nombre es realmente triste.