

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

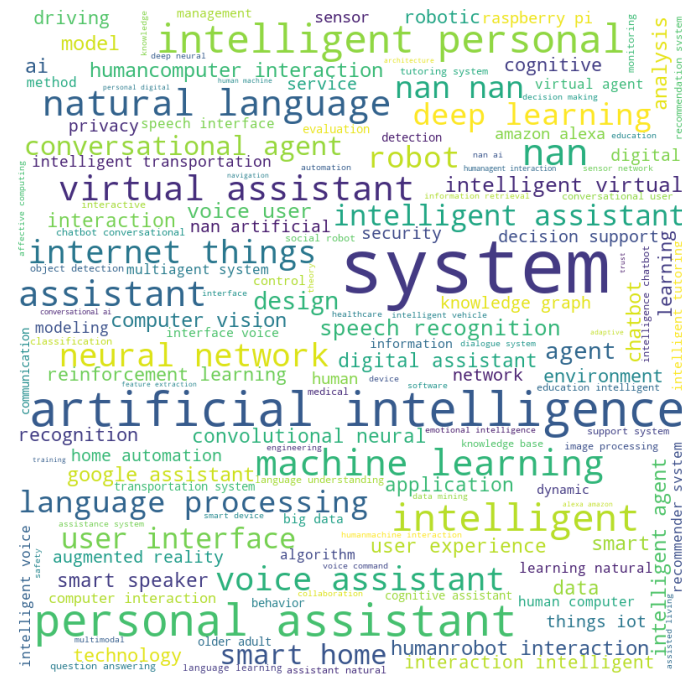


Figura 3. nube de palabras de las keywords

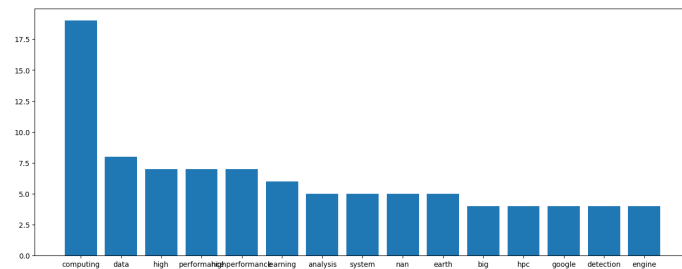


Figura 4. grafico de las keywords

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 marcará la relevancia del artículo en nuestra propia investigación.

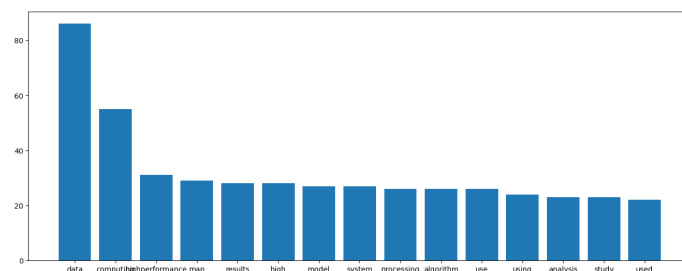


Figura 5. grafico del abstact

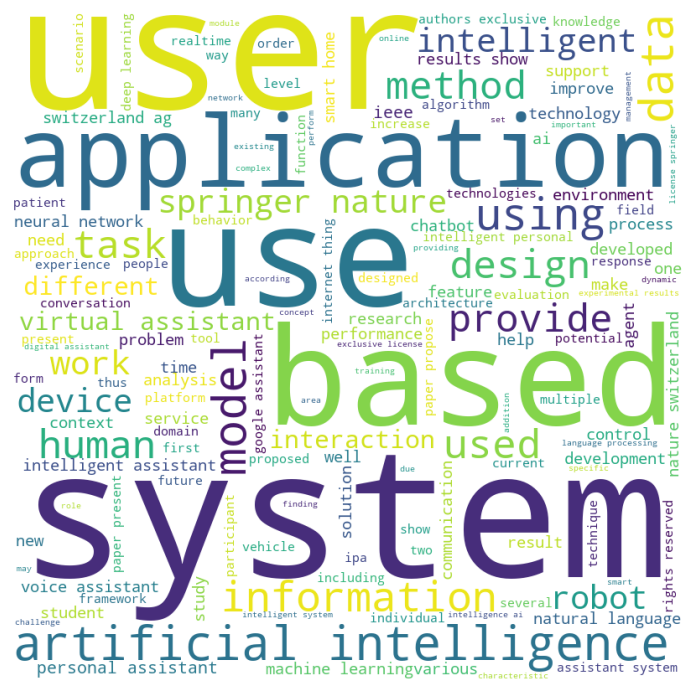


Figura 6. nube de palabras del abstact

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 , ^{EN}GI")) 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.

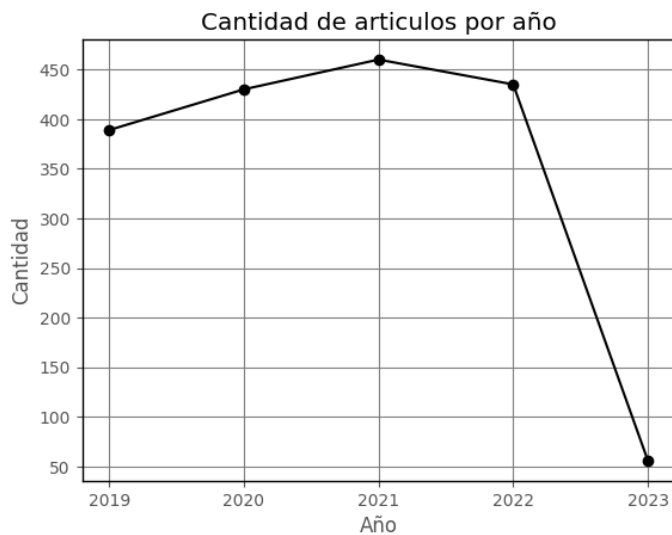


Figura 7. grafico de los años con más publicaciones

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.

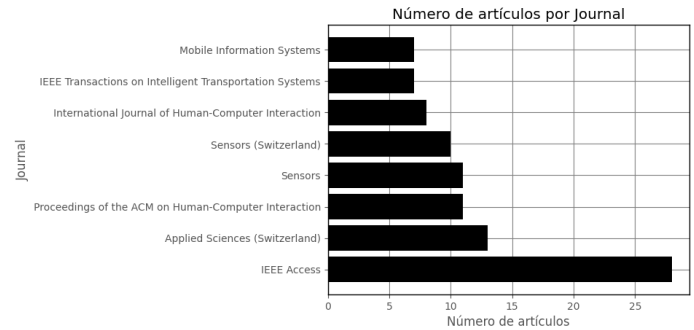


Figura 9. Cantidad de artículos por journal

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 articulos interesantes, y muchos temas de investigacion y analisis, 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>

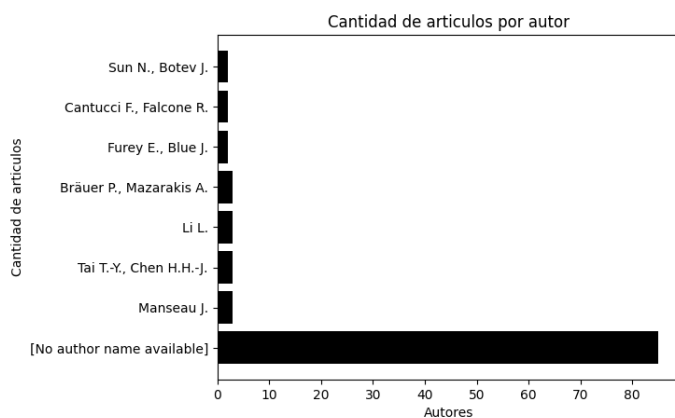


Figura 8. Grafica de los autores

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.