

UNIVERSIDAD POLITÉCNICA SALESIANA
SEDE CUENCA

CARRERA DE COMPUTACIÓN

*Trabajo de titulación previo a la
obtención del título de Ingeniero
en Ciencias de la Computación*

ARTÍCULO ACADÉMICO:

**“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA INTELIGENTE PARA LA
RECOMENDACIÓN DE ARTÍCULOS CIENTÍFICOS CON BASE A
PERFILES DE NIÑOS EN EL ÁREA DE LA TERAPIA DEL LENGUAJE Y LA
COMUNICACIÓN”**

AUTOR:

JOHN HENRY TENESACA CRIOLLO

TUTOR:

ING. DIEGO FERNANDO QUISI PERALTA, MgS

CUENCA - ECUADOR

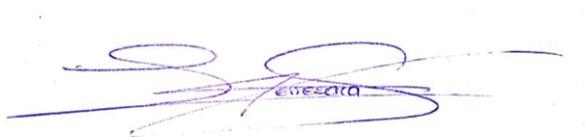
2021

CESIÓN DE DERECHOS DE AUTOR

Yo, John Henry Tenesaca Criollo con documento de identificación N° 0105598254, manifiesto mi voluntad y cedo a la Universidad Politécnica Salesiana la titularidad sobre los derechos patrimoniales en virtud de que soy autor del trabajo de titulación: **“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA INTELIGENTE PARA LA RECOMENDACIÓN DE ARTÍCULOS CIENTÍFICOS CON BASE A PERFILES DE NIÑOS EN EL ÁREA DE LA TERAPIA DEL LENGUAJE Y LA COMUNICACIÓN”**, mismo que ha sido desarrollado para optar por el título de: *Ingeniero en Ciencias de la Computación*, en la Universidad Politécnica Salesiana, quedando la Universidad facultada para ejercer plenamente los derechos cedidos anteriormente.

En aplicación a lo determinado en la Ley de Propiedad Intelectual, en mi condición de autor me reservo los derechos morales de la obra antes citada. En concordancia, suscribo este documento en el momento que hago entrega del trabajo final en formato digital a la Biblioteca de la Universidad Politécnica Salesiana.

Cuenca, octubre de 2021.



John Henry Tenesaca Criollo

C.I. 0105598254

CERTIFICACIÓN

Yo, declaro que bajo mi tutoría fue desarrollado el trabajo de titulación: **“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA INTELIGENTE PARA LA RECOMENDACIÓN DE ARTÍCULOS CIENTÍFICOS CON BASE A PERFILES DE NIÑOS EN EL ÁREA DE LA TERAPIA DEL LENGUAJE Y LA COMUNICACIÓN”**, realizado por John Henry Tenesaca Criollo, obteniendo el *Artículo Académico*, que cumple con todos los requisitos determinados por la Universidad Politécnica Salesiana.

Cuenca, octubre de 2021.

A handwritten signature in blue ink, appearing to be 'Diego Fernando Quisi Peralta', written over a horizontal line.

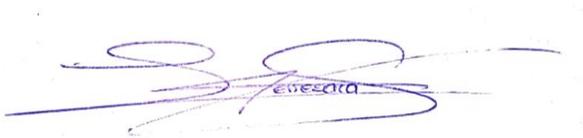
Ing. Diego Fernando Quisi Peralta, MgS

C.I. 0104616461

DECLARATORIA DE RESPONSABILIDAD

Yo, John Henry Tenesaca Criollo con documento de identificación N° 0105598254, autor del trabajo de titulación: **“DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA INTELIGENTE PARA LA RECOMENDACIÓN DE ARTÍCULOS CIENTÍFICOS CON BASE A PERFILES DE NIÑOS EN EL ÁREA DE LA TERAPIA DEL LENGUAJE Y LA COMUNICACIÓN”**, certifico que el total contenido del *Artículo Académico*, es de mi exclusiva responsabilidad y autoría.

Cuenca, octubre de 2021.



John Henry Tenesaca Criollo

C.I. 0105598254

DEDICATORIA

*Dedico este trabajo a mi familia Silvia Patricia, Juan Andres Tenesaca y Ana Orellana que son lo más importante en mi vida y a través de su amor, respeto y honestidad que me inspiraron en una de las etapas más importantes en mi vida.
Por último y no menos importante, dedico esto a mis amigos Pablo Arevalo, William Barzallo, Mateo Mogrovejo, Andres Iñiguez, Sebastian Gomez, Sebastian Velez, Jose Calle, Juan Calderon ,Andy Auquilla y Juan Diego Cabrera, gracias por el apoyo incondicional que siempre he recibido de su parte, los llevo siempre en mi corazón.*

John Henry Tenesaca Criollo

AGRADECIMIENTOS

*Agradezco a mis padres por haberme dado educación y un hogar donde desarrollarme
adquiriendo valores que hoy me definen.*

*Agradezco a la Universidad por todos los conocimientos ofrecidos a lo largo de estos 5
años que me han formado como un buen cristiano y honrado ciudadano.*

*Agradezco a mis maestros que con su dedicación y esfuerzo han hecho de mi un buen
profesional. De manera especial agradezco a mi tutor, maestro y amigo Diego Quisi, por
haberme guiado a lo largo de este trabajo y darme ánimos para seguir adelante con todas
mis metas.*

RESUMEN

El progreso de las tecnologías abre la posibilidad de la creación de nuevas herramientas que facilitan y optimizan la labor dentro de un área específica como el trastorno del lenguaje y la comunicación. El trastorno del lenguaje es un problema de comprensión o al unir palabras a fin de generar una oración y transmitir ideas, la solución son los terapeutas que se especializan en el diagnóstico, tratamiento y rehabilitación de niños o adultos. Los terapeutas al realizar búsquedas de artículos científicos en bibliotecas digitales sobre áreas específicas del lenguaje y la comunicación no son recomendados de la mejor manera, por lo cual se implementa un sistema inteligente para la recomendación de artículos científicos con base a perfiles de niños en el área del lenguaje y la comunicación mediante diferentes algoritmos inteligentes basada en bases de datos ontológicas. Los sistemas de recomendación de artículos científicos se basan en el tipo de persona, las personas menos expertas realizan una lectura general con el objetivo de introducirse en el tema y durante el proceso adquirir conocimiento específico, mientras que las personas expertas realizan búsquedas específicas. Las técnicas de recomendación de artículos se basan en el análisis de contenidos en particular dos aproximaciones: 1) La similitud de tópicos en los artículos, 2) La divergencia de tópicos en los artículos. Al realizar las dos aproximaciones se consigue resolver lo general o específico del artículo científico, mejorando la recomendación del mismo y superando a los sistemas tradicionales de recomendación. La metodología empleada consta de 3 fases importantes: 1) Reingeniería de elementos ontológicos, 2) Integración de elementos ontológicos, 3) Descubrimiento del conocimiento a través de algoritmos inteligentes. Al concluir con las fases se genera una red de conocimiento que presenta los artículos con más relevancia dentro de la red de conocimiento por medio de la detección de comunidades y la similitud de nodos. Los resultados obtenidos demuestran una satisfacción alta por medio del Alfa de Cronbach en las encuestas realizadas con diferentes terapeutas en el área del lenguaje y la comunicación, demostrando que las recomendaciones de los artículos son de utilidad o valor útil.

Palabras Claves: Bases de Datos Ontológica, Terapia del Lenguaje y la Comunicación, Algoritmos Inteligentes, Recomendaciones.

ABSTRACT

The progress of technologies opens the possibility of the creation of new tools that facilitate and optimize the work within a specific area such as language and communication disorder. Language disorder is a problem of understanding or putting words together in order to generate a sentence and convey ideas, the solution is therapists who specialize in the diagnosis, treatment and rehabilitation of children or adults. Therapists when searching for scientific articles in digital libraries on specific areas of language and communication are not recommended in the best way, so an intelligent system is implemented for the recommendation of scientific articles based on profiles of children in the area of language and communication through different intelligent algorithms based on ontological databases. The scientific article recommendation systems are based on the type of person, less expert people perform a general reading with the aim of getting introduced to the topic and during the process acquire specific knowledge, while expert people perform specific searches. The article recommendation techniques are based on content analysis in particular two approaches: 1) The similarity of topics in the articles, 2) The divergence of topics in the articles. By performing the two approaches, the general or specific nature of the scientific article is resolved, improving its recommendation and surpassing the traditional recommendation systems. The methodology used consists of 3 important phases: 1) Re-engineering of ontological elements, 2) Integration of ontological elements, 3) Knowledge discovery through intelligent algorithms. At the end of the phases, a knowledge network is generated that presents the most relevant articles within the knowledge network by means of community detection and node similarity. The results obtained show a high satisfaction by means of Cronbach's Alpha in the surveys carried out with different therapists in the area of language and communication, demonstrating that the recommendations of the articles are of usefulness or useful value.

Keywords: Ontological Databases, Language and Communication Therapy, Intelligent Algorithms, Recommendations.

ÍNDICE DE CONTENIDOS

I. Introducción	10
1.1 Estado Del Arte	10
A. Ontologías	11
B. Sistemas de recomendación	11
C. Trabajos relacionados	12
II. Metodología	15
2.1 Recolección de datos	18
A. Selección información	18
B. Preprocesamiento	22
C. Procesamiento de lenguaje natural del artículo	22
2.2 Descubrimiento del conocimiento de los artículos	23
A. Modelado y agregación de conocimiento	24
B. Reingeniería de elementos no ontológicas	25
C. Integración de elementos ontológicos	28
D. Algoritmos inteligentes neo4j	29
a. Descubrimiento del conocimiento a través del algoritmo similitud de nodo (Node Similarity)	30
b. Descubrimiento del conocimiento a través del algoritmo centralidad de intermediación (Betweenness Centrality)	31
c. Descubrimiento del conocimiento a través del algoritmo centralidad de grado (Degree Centrality)	31
III. Resultados y discusión	32
3.1 Resultados obtenidos por los algoritmos inteligentes	32
IV. Conclusiones	41
V. Referencias	42
VI. Anexos	44
Anexo 1	44

I. Introducción

El terapeuta al realizar búsquedas de artículos científicos en bibliotecas digitales en el área del lenguaje y la comunicación la información presentada no es la que el terapeuta necesita en función al perfil del paciente, ya que no siempre se encuentra el contenido correcto, para mejorar este sistema es necesario la implementación de un sistema inteligente de recomendación sobre los datos ingresados por los terapeutas para realizar una búsquedas especializada en el área.

Este artículo trata sobre el diseño de un sistema inteligente que permite recomendar artículos científicos en base a perfiles de niños en el área del lenguaje y la comunicación por medio de algoritmos inteligentes que permiten recomendar artículos científicos en base a palabras clave, generando una búsquedas especializada que pueda ser de utilidad o valor para el terapeuta. Los artículos serán extraídos de las bases de datos científicas (Proquest, Scopus, IEEEExplore) y recursos científicos (Doaj, USPTO, Google Noticias).

Se recomiendan los artículos por medio de una base de datos orientada a grafos que ofrece un mejor rendimiento, agilidad, flexibilidad y escalabilidad en los sistemas de recomendación. Las BD orientadas a grafos son capaces de conectar un conjunto de “A” con una característica “B” de otro conjunto con la finalidad de descubrir nuevo conocimiento y realizar una recomendación en su área de dominio. Las BD orientadas a grafos permiten realizar el descubrimiento de conocimientos más conocido como KDD (Knowledge Discovery in Databases) que es el proceso de descubrir conocimientos útiles a partir de una colección de datos y su objetivo es extraer conocimientos de alto nivel a partir de datos de bajo nivel. Entre las técnicas de recomendación de artículos científicos existen tres algoritmos inteligentes que ofrece Neo4j que es una BD orientada a grafos, con la finalidad de ofrecer un sistema escalable de conocimiento(KDD) y un rendimiento alto.

1.1 Estado Del Arte

La terapia del lenguaje es la más efectiva, ya que permite mejorar las destrezas de la comunicación (Belsky & G., 2021). Los terapeutas que realizan este tipo de terapia empiezan por identificar qué tipo de problema del habla o el lenguaje tiene un niño luego determina la causa y posteriormente se decide cuál es el mejor tratamiento, ya que la terapia debe adaptarse a las necesidades específicas de cada niño, entre las habilidades más comunes en las que trabajan están (Belsky & G., 2021):

- Conocimiento fonológico: “Los terapeutas consiguen ayudar a los niños a rimar e identificar los sonidos de las palabras”.
- Expresando ideas más complejas: “Los terapeutas pueden enseñar “unir palabras” para ayudar a los niños a combinar sus ideas en oraciones”.
- Construyendo vocabulario: “Los terapeutas logran ayudar a los niños a recordar palabras representándolas o usándolas para contar una historia”.
- Habilidades de conversación: “Los terapeutas pueden representar una conversación y ayudar a los niños a captar las señales sociales” .

El progreso de las tecnologías abre la posibilidad de la creación de nuevas herramientas que facilitan y optimizan la labor dentro de un área específica, por esta razón se propone realizar la implementación de un sistema inteligente para la recomendación de artículos científicos mediante algoritmo inteligentes con el objetivo de presentar información relacionada que será de utilidad o valor para el terapeuta.

A. Ontologías

Una ontología es una descripción formal de propiedades y relaciones entre conjuntos que permiten representar un área de dominio. “En los campos de la Inteligencia Artificial y la Web Semántica se crean ontologías para limitar la complejidad y para organizar la información. Las ontologías pueden proveer los mecanismos para organizar y almacenar los componentes genéricos de sistemas de información que incluyen esquemas de las bases de datos (BD)”. (Alvarez & BARCHINI, 2021)

B. Sistemas de recomendación

- **Sistemas de recomendación basados en contenido**

Los sistemas de recomendación basada en contenido “consideran atributos y características de los ítems de forma que el usuario obtenga de resultado ítems recomendados con características similares a los ítems que prefirió previamente”. (Artigot, 2017)

- **Sistemas de filtrado colaborativo**

Los sistemas de filtrado colaborativo no consideran las características propias de un ítem, “se basan en emplear las preferencias de otros usuarios con rasgos similares para recomendar ítems al usuario actual”. (Artigot, 2017)

- **Sistemas de recomendación semánticos (SRS)**

Las principales barreras de los enfoques tradicionales de recomendación, es ignorar la semántica de los elementos porque se basan en representaciones de palabras claves. Según (Utrera & Cuevas, 2017) “la precisión de las recomendaciones se puede mejorar integrando la estructura del sitio y el contenido del sitio web, extendiéndose con datos semánticos que lo caracterizan”. Los SRS se basan en la interpretación “semántica” y de las necesidades del usuario. Los sistemas de recomendación semánticos basados en ontologías se apoyan en una base de conocimiento para mejorar semánticamente sus recomendaciones haciendo uso de ontologías.

C. Trabajos relacionados

La recomendación de artículos científicos se basa en el tipo de persona, “las personas menos expertas prefieren leer un artículo general con el objetivo de introducirse en el tema mientras que las personas expertas prefieren artículos más específicos” (Mena, 2018). Las técnicas de recomendación se centran en el análisis del contenido en particular dos aproximaciones para recomendar artículos científicos:

1. La similitud que se dan entre estos tópicos de los Artículos
2. La divergencia de tópicos que se dan en los Artículos

Según (grapheverywhere, 2020) “un grafo es una combinación de objetos, conocidos como nodos que se relacionan entre sí mediante conexiones nombradas como aristas permitiendo a los grafos estudiar todas las relaciones que interactúan”.

Además (Antonio et al., 2015) propone que “con ambas medidas se consigue determinar lo general o específico de un artículo para su recomendación, superando a un sistema de recuperación de información tradicional”.

Se planteó (Huaman, 2018) “el uso de un modelo ontológico usando las relaciones con tesauros para la recomendación de libros basándose que todo libro tiene

descriptores o tesauros que los identifican y estos a su vez pueden pertenecer a otros libros del mismo tema de investigación. Es decir, para conocer un término debemos conocer previamente otros”.

Según (Brut & Sedes, 2010) expone una técnica para desarrollar una solución para las recomendaciones personalizadas, adoptando un modelado basado en ontologías de perfiles de usuario y modelos de documentos. Básicamente, “para cualquier técnica de recomendación se analiza la actividad de navegación del usuario y algunos enfoques de nivel de abstracción, se consideraron la navegación basada en conceptos, donde cada concepto y página es un centro de navegación”.

La metodología para integrar ontologías implica tres pasos según (Brut & Sedes, 2010):

1. Preparación de datos: Analizar documentos para generar ontología de dominio.
2. Descubrimiento de patrones: Analizar las elecciones del usuario con el fin de establecer patrones de uso semántico.
3. Recomendaciones: Hacer coincidir el modelo de dominio.

Propone (Xia, 2016) un método de recomendación que incorpora “relaciones de autor comunes entre artículos para ayudar a generar mejores recomendaciones para los investigadores”. Los investigadores buscan artículos publicados por los mismos autores para encontrar artículos que les interesen esto lo llaman patrón de búsquedas basado en autor

Plantea (Sun et al., 2013) el uso de tres tipos de medidas de similitud: similitud de palabras clave, similitud de revista y similitud de autor para medir la relevancia de los artículos para los investigadores. “La similitud de palabras clave se utiliza para generar una lista de artículos candidatos, la similitud de la revista y la similitud del autor se utilizan para seleccionar los artículos más adecuados de la lista de candidatos”. La medida de similitud integrada se utiliza para clasificar los artículos según su relevancia.

En el artículo (Sun & Ma, 2014) se propone un enfoque para recomendar artículos científicos a los investigadores aprovechando el contenido y las conexiones. En este enfoque, “analiza el contenido semántico del artículo mediante el cálculo de similitud

de palabras clave y luego extrae las conexiones de los usuarios en línea para respaldar la votación del artículo y finalmente emplea un proceso de recomendación de dos etapas para sugerir artículos relevantes”. Las palabras clave generadas mediante la consulta de un investigador a menudo tienen ambigüedades semánticas y estas relaciones semánticas también pueden existir en los documentos. La solución planteada por (Sun & Ma, 2014) “utiliza una matriz de similitud de palabras clave para abordar el problema de la falta de coincidencia”.

II. Metodología

En esta sección se detalla la metodología que se va a emplear.

A. Conceptos Previos

Antes de presentar la metodología es necesario señalar la importancia de los sistemas de recomendación basada en BD orientadas a grafos. Las BD orientadas a grafos brinda herramientas para crear aplicaciones inteligentes y flujos de aprendizaje automático de la actualidad. La implementación de BD orientadas a grafos ofrece algoritmos que se dividen en categorías con la finalidad de resolver diferentes clases de problemas:

- **Centralidad:** El algoritmo revela qué nodos son importantes según la topología del gráfico identificando los nodos influyentes en función de su posición en la red.
- **Detección de comunidad:** El Algoritmo agrupa su gráfico en función de las relaciones para encontrar comunidades donde los miembros tienen interacciones más significativas. La detección de comunidades ayuda a predecir comportamientos similares, encontrar entidades duplicadas o preparar los datos para otros análisis.
- **Semejanza:** El Algoritmo emplea comparaciones de conjuntos para puntuar qué tan parecidos son los nodos individuales en función de sus vecinos o propiedades. Las propiedades y atributos de los nodos se utilizan para puntuar la semejanza entre los nodos. El enfoque es utilizado en aplicaciones de recomendaciones personalizadas y desarrollo de jerarquías categóricas.
- **Pathfinding y búsquedas:** Encuentra los caminos óptimos evaluando la disponibilidad y la calidad de la ruta. Las búsquedas de rutas son fundamentales para el análisis de gráficos y encontrar las rutas más eficientes o más cortas para atravesar nodos.

Cada algoritmo inteligente dentro de neo4j contiene una o varias de las categorías presentadas previamente con la finalidad de ofrecer un rendimiento alto y tiempos

bajos de respuesta. Los algoritmos inteligentes implementados son: Similitud de nodo, Centralidad de intermediación y Centralidad de grado.

Se seleccionan los 3 algoritmos inteligentes por causa de que permiten aplicarse a grafos no ponderados estos son aquellos grafos que no tienen un valor o peso asociado. Los algoritmos inteligentes son rápidos y realizan un consumo mediano de recursos por lo que beneficia los tiempos de carga para el usuario. Estos algoritmos son de gran importancia para la generación de conocimiento, puesto que como resultado obtienen los pesos de los artículos más relevantes de la red de conocimiento.

Las BD orientadas a grafos permiten realizar el descubrimiento de conocimientos que es el proceso de descubrir conocimientos útiles a partir de una colección de datos y su objetivo es extraer conocimientos de alto nivel a partir de datos de bajo nivel, por lo tanto se implementa la generación de conocimiento por medio del esquema vivo y mundo.

Las ontologías se relacionan con la finalidad de generar una red de conocimiento sobre un dominio de modo que se implementan 2 ontologías de dominio que ofrecen las relaciones para administrar datos sobre un investigador para los artículos científicos y las relaciones de las palabras en el área del lenguaje y comunicación.

Ontologías de dominio:

- **VIVO:** “Es una ontología para administrar datos sobre investigadores de código abierto que cuando se rellena con los intereses, permite descubrir la investigación y la erudición en todas las disciplinas” (Colaboradores BioPortal, 2018).
- **Mondo Disease Ontology (Bio Portal):** “Es una ontología sobre trastornos que se caracteriza por la incapacidad de un individuo para comprender o compartir ideas o sentimientos debido a una discapacidad en el lenguaje, el habla o la audición” (Colaborador BioPortal, 2021).

Para obtener toda esta información se debe trabajar con RDF (Resource Description Framework) que es “un modelo conceptual que proporciona información descriptiva sobre los recursos que se encuentran en la web que permite el intercambio de información sin que los datos pierdan su significado permitiendo que las aplicaciones

puedan acceder a un esquema público vía web y recuperar la estructura y la semántica sin perder información” (gazteaukera, 2019).

A continuación, se da una explicación más detallada de todos los conceptos revisados previamente, con el objetivo de entender de mejor manera el método propuesto. Las relaciones entre esquemas permiten descubrir el conocimiento por tanto se deben relacionar vivo, mundo y ontoartículo. Ontoartículo será nuestro esquema después de pasar la recolección de datos de las diferentes BD Científicas, preprocesamiento de datos y la reingeniería de elemento no ontológicos, ofrecerá una red de conocimientos de los artículos científicos relacionados con un trastorno específico en el área. Obtiene este conocimiento al relacionar las palabras claves de los artículos con las palabras claves obtenidas por la red de dominio mundo que es la red de conocimiento en el área del lenguaje y comunicación. Las relaciones se generan mediante consultas Cypher que son el lenguaje de consulta de gráficos de neo4j que permite seleccionar, insertar, actualizar o eliminar de la BD orientada a grafos.

El método propuesto en este trabajo consiste en 7 pasos los cuales se pueden apreciar en la figura 1.

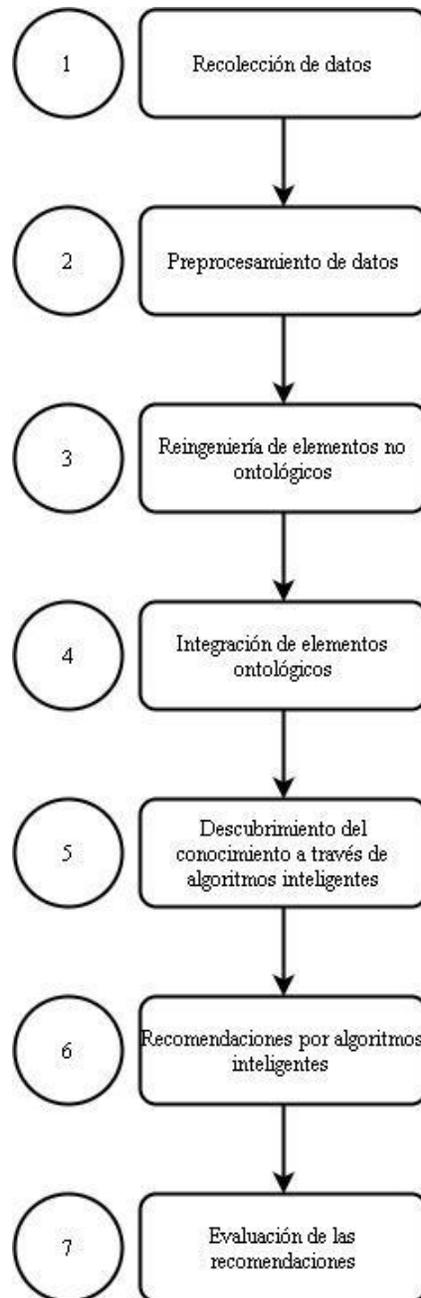


Figura 1:Método propuesto

2.1 Recolección de datos

En esta fase se describen los procesos de recolección en donde se abordan las fases de selección y preprocesamiento de los resultados.

A. Selección información

Como principal proceso se debe seleccionar la información de las BD científicas:

- **Scopus:** Es un repositorio de documentación científica que “basa su modelo de datos en tres grupos: autores, documentos y afiliaciones” (Mena, 2018). Para realizar el consumo del Api de scopus debemos tener un api key y se extraen los campos más relevantes como se observa en la tabla I.

Campo	Descripción
scopus_link	Contiene el link del artículo de Scopus.
title	Contiene el título del artículo de Scopus.
abstract	Contiene el abstract del artículo de Scopus.
authors	Contiene los Id de los autores.

Tabla I:Tabla descriptiva de los campos seleccionados del servicio REST de Scopus.

La API retorna el id de los autores por lo que se debe realizar una siguiente consulta sobre los autores por medio del id obtenido en la primera consulta y solo extraemos los campos más relevantes sobre el autor, como se observa en la tabla II.

Campo	Descripción
name	Contiene el nombre del autor.

Tabla II:Tabla descriptiva de los campos seleccionados del servicio REST de Scopus sobre autores.

- **DOAJ:** Es un repositorio en línea que “registra y proporciona acceso a revistas académicas de acceso abierto” (Doaj, 2018). Doaj ofrece su API de forma gratuita así que extraemos los campos más relevantes dentro del API, como se observa en la tabla III.

Campo	Descripción
url	Contiene el link del artículo de DOAJ.

title	Contiene el título del artículo de DOAJ.
abstract	Contiene el abstract del artículo de DOAJ.
author	Contiene los nombres de los autores del artículo.
keywords	Contiene los keywords del artículo.

Tabla III:Tabla descriptiva de los campos seleccionados del servicio REST de DOAJ.

- **ProQuest:** Suministra archivos científicos de fuentes diversas como revistas, tesis doctorales y bases de datos. Para realizar el consumo del Api de ProQuest debemos tener una autorización y extraemos los campos más relevantes dentro del API de ProQuest, como se observa en la tabla IV.

Campo	Descripción
URL	Contiene el Link del artículo de ProQuest.
field	Contiene el Título del artículo de ProQuest.
Abstract	Contiene el Abstract del artículo de ProQuest.
Author	Contiene los nombres de los autores del artículo.
HeadingTerm	Contiene los keywords del artículo.

Tabla IV:Tabla descriptiva de los campos seleccionados del servicio REST de ProQuest.

- **Uspto:** Es un directorio en línea que gestiona patentes a las empresas para la identificación de productos y de propiedad intelectual. Uspto ofrece su API de forma gratuita así que extraemos los campos más relevantes dentro del API, como se observa en la tabla V.

Campo	Descripción
filelocationURI	Contiene el link del artículo de Uspto.
inventionTitle	Contiene el título del artículo de Uspto.
abstractText	Contiene el abstract del artículo de Uspto.

inventorNameArrayText	Contiene los nombres de los autores del artículo .
-----------------------	--

Tabla V:Tabla descriptiva de los campos seleccionados del servicio REST de Uspto.

- **IEEEExplore:** Es una BD que permite la búsquedas y recuperación de publicaciones de IEEE. Para realizar el consumo del Api de IEEEExplore debemos tener un api key. Extraemos los campos más relevantes como se observa en la tabla VI.

Campo	Descripción
pdf_url	Contiene el link del artículo de IEEEExplore.
title	Contiene el título del artículo de IEEEExplore.
abstract	Contiene el abstract del artículo de IEEEExplore.
authors	Contiene los nombres de los autores del artículo .
terms	Contiene los keywords demediol artículo .

Tabla VI:Tabla descriptiva de los campos seleccionados del servicio REST de Uspto

- **Google Noticias:** “Es un buscador de noticias que rastrea la información de los medios de comunicación en línea” (colaboradores de Wikipedia, 2021). Extraemos los campos más relevantes dentro de Google Noticias, como se observa en la tabla VII.

Campo	Descripción
link	Contiene el Link.
title	Contiene el Título.
desc	Contiene la Descripción.
media	Contiene la Fuente.

Tabla VII:Tabla descriptiva de los campos seleccionados del servicio REST de Google Noticias

Para obtener la información se utilizó el lenguaje Python mediante el consumo de servicios REST publicados por cada BD Científica y se realizó una consulta de 100 artículos

relacionados con una palabra clave obtenida por medio de Mondo. El tesoro principal se presenta mediante las propiedades de la red de dominio de Mondo esta contiene todas las palabras claves relacionadas con las discapacidades en el lenguaje.

Como información extra de cada artículo se agrega los siguientes campos presentados en la tabla VIII que son de gran importancia para realizar una integración de elementos ontológicos.

Campo	Descripción
Fuente	La fuente es el nombre de la BD Científica de donde se recolectó dicho artículo.
Key	Es la palabra ingresada en el servicio de la BD Científica.

Tabla VIII:Tabla descriptiva de los campos de integración ontológica.

B. Preprocesamiento

En este paso se perfeccionan los datos mediante la limpieza de los datos faltantes y eliminado el ruido. El objetivo de este paso es eliminar los datos erróneos y empezar a generar datos de gran calidad.

- **Datos Redundantes:** La eliminación de datos redundantes es por medio de la columna link y abstract. Esta columna permite realizar una limpieza de todos los artículos redundantes dentro de la recaudación.
- **Incompletos:** Los artículos incompletos son aquellos que no cuentan con link, título o abstract estos artículos son eliminados.

C. Procesamiento de lenguaje natural del artículo

Al terminar de realizar la limpieza, encontramos las principales palabras usadas en los artículos por medio de su título o abstract. Para obtener las palabras principales utilizamos spaCy que es una biblioteca para el procesamiento avanzado de lenguaje natural en Python (Fosado, 2019) y agregamos los campos presentados en la tabla IX.

Campo	Descripción
paltitulo	El campo contiene las palabras claves obtenidas por Spacy del título del artículo.
palabstract	El campo contiene las palabras claves obtenidas por Spacy del abstract del artículo.

Tabla IX:Tabla descriptiva de los campos para la recomendación.

La generación de las palabras clave del título y abstract permitirá obtener una mejor recomendación mediante sus relaciones con otros artículos que contengan las mismas palabras clave. Al terminar el preprocesamiento obtiene una dataset con la siguiente información presentada en la tabla X:

Columna	Descripción
link	Contiene los links de los artículos.
title	Contiene los títulos de los artículos.
abstract	Contiene los abstracts de los artículos.
authors	Contiene los autores de los artículos.
keywords	Contiene los keywords de los artículos.
paltitulo	Contiene las palabras claves del título de cada artículo.
palabstract	Contiene las palabras claves del abstract de cada artículo.
key	Contiene la palabra ingresada en el servicio de cada BD Científica

Tabla X:Tabla descriptiva de los campos obtenidos de la recaudación de información.

2.2 Descubrimiento del conocimiento de los artículos

En esta etapa se desarrolla la técnica de KDD que es la creación de conocimiento a partir de esquemas estructurados y no estructurados que permite encontrar patrones en los artículos,

realizando una etapa de agregación de conocimiento por medio de ontologías y algoritmos inteligentes para clasificar los contenidos y descubrir conocimiento que no se presentaba en el dataset.

A. Modelado y agregación de conocimiento

El modelado y agregación de conocimiento “busca construir un modelo que conceptualiza una abstracción de una parte de la realidad basada en algunas partes del marco de desarrollo de ontologías, para la construcción de la red ontológica y base de conocimiento” (Mena, 2018).

En la Figura 2 se observa el modelado de conocimiento que inicia con el proceso de recolección de datos a través del tesoro principal que se obtiene mediante el esquema Mondo que en sus propiedades contiene las palabras relacionadas con los problemas en el área de dominio. Este pasa por un preprocesamiento que realiza una limpieza de la recolección de datos el cual debe pasar por una etapa de reingeniería de elementos no ontológicos con la finalidad de generar la integración de elementos ontológicos formando la base de conocimiento en la herramienta Neo4j que genera una capa de información extra en el área de dominio.

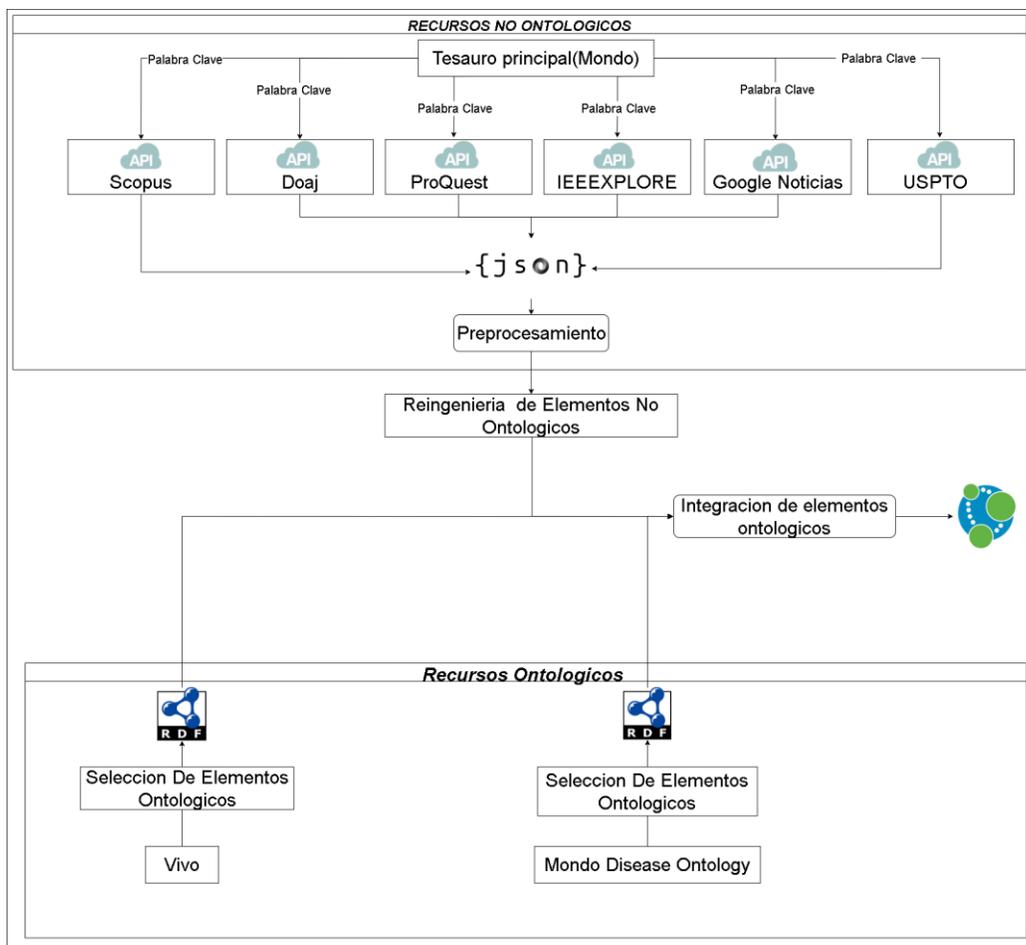


Figura 2:Esquema de proceso de modelado de conocimiento

B. Reingeniería de elementos no ontológicas

Para relacionar los resultados no ontológicos debemos establecer relaciones dentro de la BD Neo4j a través de consultas Cypher. Con el fin de relacionar los datos no ontológicos dentro de la base de datos debemos conocer las relaciones que tiene un artículo.

Los artículos contienen la siguiente información presentada en la tabla XI:

Campo	Prioridad	Descripción
link	Obligatorio	El link del artículo
título	Obligatorio	Titulo del articulo
abstract	Obligatorio	Abstract del artículo
keyword		Palabras claves que contiene el artículo.

autores	Obligatorio	Nombres de los autores del artículo
palabras clave titulo	Obligatorio	Palabras clave del título del artículo.
palabras clave abstract	Obligatorio	Palabras clave del abstract del artículo.
fuelle	Obligatorio	Nombre de la BD Científica de la que recogió esta información.
key	Obligatorio	Palabra ingresada dentro del API.

Tabla XI: Tabla descriptiva de los campos seleccionados para relacionar.

Para generar relaciones dentro de neo4j debemos utilizar “MERGE” que permite coincidir con los nodos existentes y los vincula, o bien crea nuevos datos y los vincula. Las relaciones que se generan se presentan en la tabla XII.

El ARTÍCULO contiene RELACIÓN con un LINK .
El ARTÍCULO contiene RELACIÓN con un TÍTULO .
El ARTÍCULO contiene RELACIÓN con un ABSTRACT .
El ARTÍCULO contiene una o varias RELACIONES con KEYWORDS .
El ARTÍCULO contiene una o varias RELACIONES con AUTORES .
El ARTÍCULO contiene una o varias RELACIONES con las PALABRAS CLAVE TÍTULO .
El ARTÍCULO contiene una o varias RELACIONES con las PALABRAS CLAVE ABSTRACT .
El ARTÍCULO contiene RELACIÓN con una FUENTE
El ARTÍCULO contiene RELACIÓN con una KEY

Tabla XII: Relaciones del Artículo.

Para relacionar un artículo a varios autores, keywords, palabras clave título, palabras clave abstract se debe generar una instancia con un ID para identificar el nodo de un artículo en específico y agregar los autores en un mismo artículo como se observa en la figura 4 .

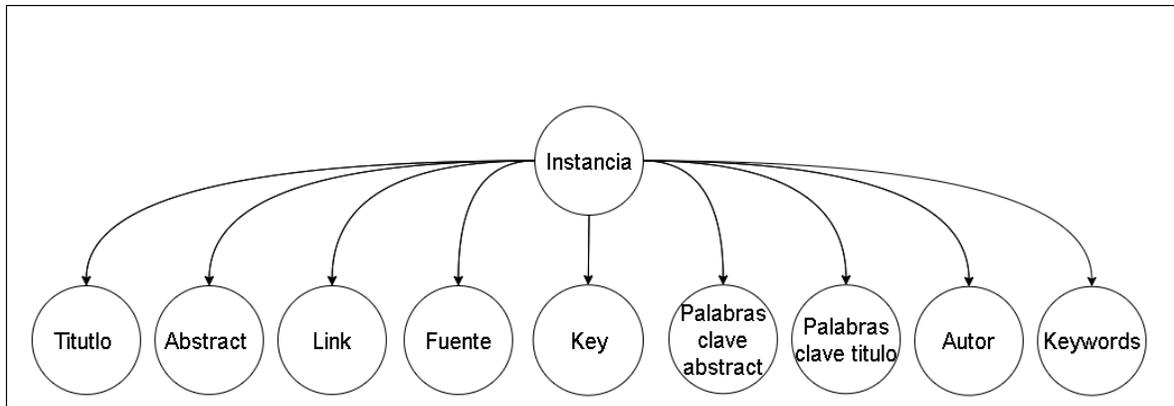


Figura 3:Modelo de grafo generado

Se deben generar 4 instancias que son para los autores, keywords, palabras claves del título, palabras claves del abstract. La generación de las palabras claves del título y abstract se relacionan, ya que contienen las mismas palabras en el título como en el abstract como se puede observar en la figura 4.

Para comprobar que las ontologías de la base de conocimiento estén integrados, se utiliza la herramienta Neo4j para importar las tres ontologías.

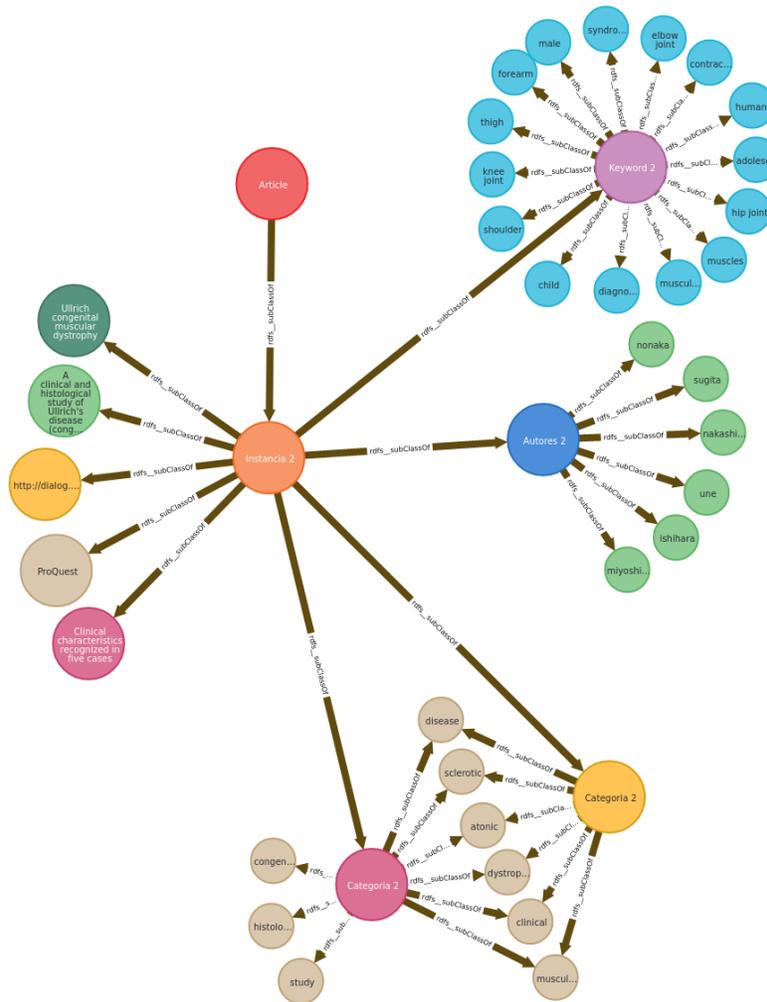


Figura 4: Relaciones en Neo4j

C. Integración de elementos ontológicos

La integración de los elementos ontológicos es el proceso que une los elementos que con anterioridad fueron alineados y se comprueba el funcionamiento en conjunto. La integración inicia en el acoplamiento de los esquemas para formar la red ontológica (Mena, 2018). Las ontologías vivo, mundo y ontoartículo se relacionan por características específicas que contienen dentro de la base de conocimiento.

En esta etapa relacionamos las 3 ontologías (vivo, mundo, ontoartículo) para la generación de KDD.

1. Vivo Y Ontoartículo: Los relacionamos de forma que el nodo “article” de mundo que tiene una relación con la instancia del artículo.

2. Mondo y Ontoarticulo: Los relacionamos con cada uno de los trastornos de la comunicación a través del título.
 - a. Agregamos pesos según las coincidencias de palabras con la categoría de mundo.
 - b. Relacionamos la key del artículo de ontoarticulo con la categoría de mundo.

D. Algoritmos inteligentes neo4j

En neo4j se desarrollan algoritmos inteligentes dentro de Graph Data Science. Cada algoritmo inteligente devuelve como resultado los pesos de cada nodo, esto es a través de la selección de los nodos con mayor relevancia en el caso de las recomendaciones se basan en los nodos con palabras clave como son título y abstract que contiene una relación con otro conjunto de nodos con las mismas características.

- **Graph Data Science(GDS):** “Es la librería que contiene algoritmos de gráficos para el desarrollo de predicciones en las relaciones y las estructuras de red de datos existentes para la generación de predicciones” (Neo4j, 2021). La biblioteca GDS contiene algoritmos de gráficos reforzados para obtener resultados consistentes, con incrustaciones de gráficos y modelos entrenados dentro del espacio de trabajo de análisis.

Los algoritmos de GDS seleccionados son:

- a. Similitud de nodo
- b. Centralidad de intermediación
- c. Centralidad de grado

Los algoritmos inteligentes seleccionados permiten aplicarse a grafos ponderados o no ponderados estos son aquellos grafos que tienen un valor o peso asociado. Estos algoritmos son de gran importancia para la generación de conocimiento que se presentará más adelante a profundidad.

a. Descubrimiento del conocimiento a través del algoritmo similitud de nodo (Node Similarity)

El algoritmo de similitud de nodos “calcula las similitudes entre pares basándose en la métrica Jaccard” (Neo4j, 2019) se observa su relación en la figura 5. Los nodos se consideran similares si comparten varios de los mismos vecinos.

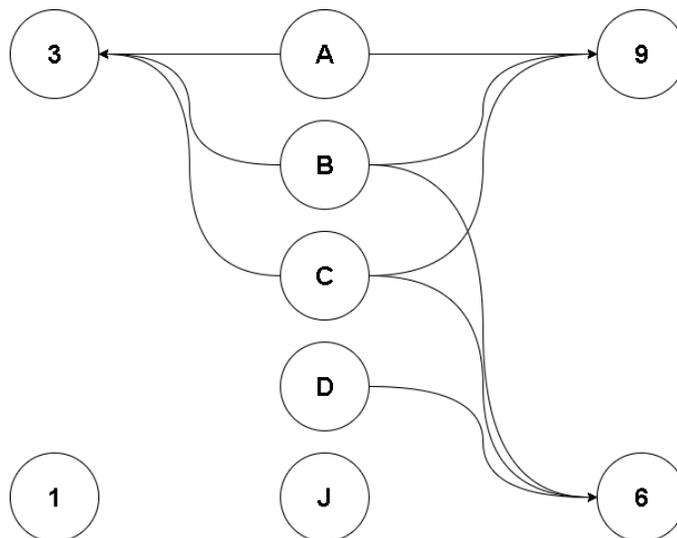


Figura 5:Node Similarity

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

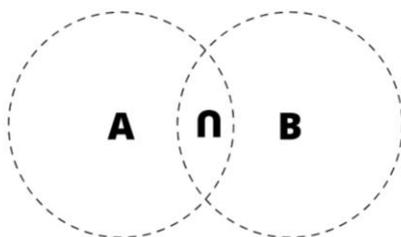


Figura 6:Jaccard

- **Jaccard:**Mide el grado de similitud entre dos conjuntos sin importar el elemento a comparar. Se observa su formulación en la figura 6.

El algoritmo de similitud de nodos compara cada nodo que tiene relaciones salientes con cada uno de esos nodos. “Para cada par ”a”, ”b” el algoritmo calcula la similitud para ese par a través de las similitudes Jaccard” (Colaboradores Neo4j, 2021).

En la figura 5 se denota que los nodos “B” y “C” contienen relaciones similares B(3, 9, 6), C(3, 9, 6) el algoritmo node similarity valora (B,C)=1, entonces la valoración de cada nodo permite saber que tan fuertemente un “Artículo A” está relacionado con un “Artículo B” para realizar una recomendación basado en el peso que ofrece el algoritmo inteligente.

b. Descubrimiento del conocimiento a través del algoritmo centralidad de intermediación (Betweenness Centrality)

La centralidad de intermediación “detecta la cantidad de influencia que tiene un nodo frente al flujo de información en el grafo” (Neo4j, 2019). El algoritmo calcula los caminos más cortos entre todos los pares de nodos de un gráfico y recibe una puntuación basada en el número de caminos más cortos que pasan por el nodo (Neo4j, 2019).

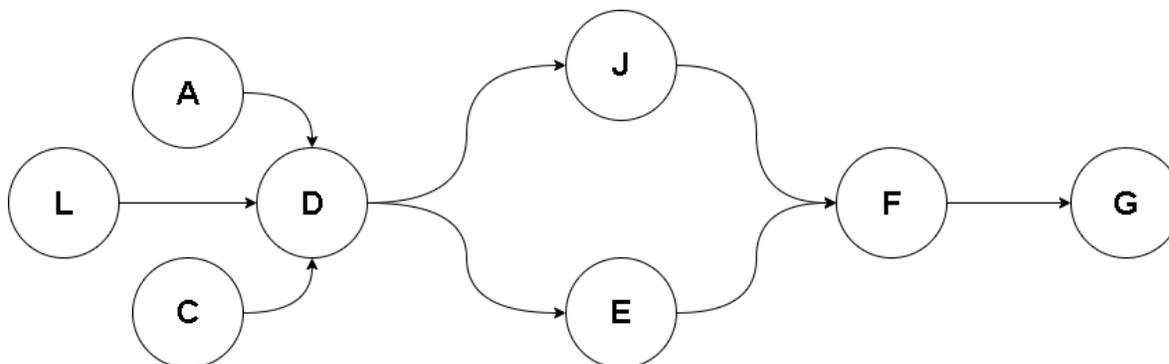


Figura 7:Betweenness Centrality

En la figura 7 se presenta que el nodo “D” está en posición de cuello de botella en el gráfico, el nodo “D” contiene “A, L, C” y todos los demás nodos lo que aumenta su puntuación o peso. Esta valoración permite saber qué tan relacionado está un nodo, obtiene como resultado el nodo con más influencia frente a los demás, para realizar una recomendación basado en la puntuación del nodo que es un artículo.

c. Descubrimiento del conocimiento a través del algoritmo centralidad de grado (Degree Centrality)

El algoritmo de centralidad de grado permite encontrar los nodos populares dentro de un gráfico (Neo4j, 2019). La centralidad de grado mide el número de relaciones entrantes o

salientes de un nodo, dependiendo de la orientación de las relaciones. Degree Centrality determina cuál es el artículo más popular, ejemplo en la figura 7.

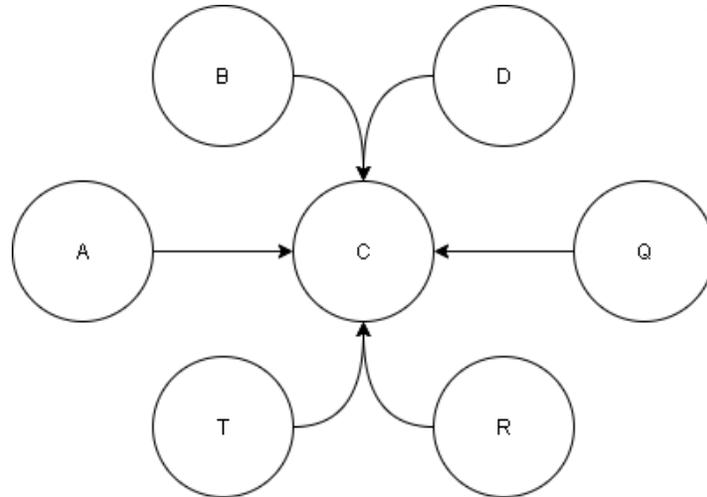


Figura 8: Degree Centrality

En la figura 8 se denota que “C” es el más popular con 6 relaciones entrantes, todos los demás nodos se relacionan con “C” pero estos no contienen relaciones, por lo tanto “C” se considera como centralidad, esto permitiendo realizar una recomendación basada en la popularidad de cada nodo que es un artículo que es valorada entre 0 y 1 que es el peso guía

III. Resultados y discusión

3.1 Resultados obtenidos por los algoritmos inteligentes

Para comparar los resultados obtenidos por los algoritmos inteligentes se generaron diferentes búsquedas de los mismos términos en cada algoritmo inteligente y se analiza la recomendación en los primeros cinco artículos. Los terapistas califican si los artículos son de utilidad o valor la encuesta se presenta en el anexo 1. Los resultados obtenidos a través de las encuestas se presentan en las tablas XIII, XV, XVII esta calificada por los terapistas entre (1 a 5) con resultados por el coeficiente alfa de Cronbach en las tablas XIV, XVI, XVIII se

profundizan los resultados mediante gráficos de barra de cada algoritmo inteligente en las figuras 9, 10, 11.

	Art 1	Art 2	Art 3	Art 4	Art 5	Art 1	Art 2	Art 3	Art 4	Art 5
Terapista 1	4	4	3	3	3	5	3	2	1	5
Terapista 2	4	3	3	3	3	4	3	3	3	3
Terapista 3	5	4	5	5	5	5	3	3	3	4
Terapista 4	3	4	4	3	3	3	3	3	3	3
Terapista 5	3	3	3	2	2	3	3	2	4	4
Terapista 6	3	3	3	3	3	3	3	3	3	3

Tabla XIII:Resultados de los terapeutas encuestados por el algoritmo Node Similarity

Los resultados obtenidos a través del coeficiente alfa de Cronbach del algoritmo Node Similarity los presentamos en la tabla XIV:

	Valor
K(Número de Ítems)	10
Vi(Varianza del Ítem)	5.388888889
Vt(Varianza Total)	18
Resultado	0.8117283951

Tabla XIV:Resultados a través del coeficiente alfa de Cronbach

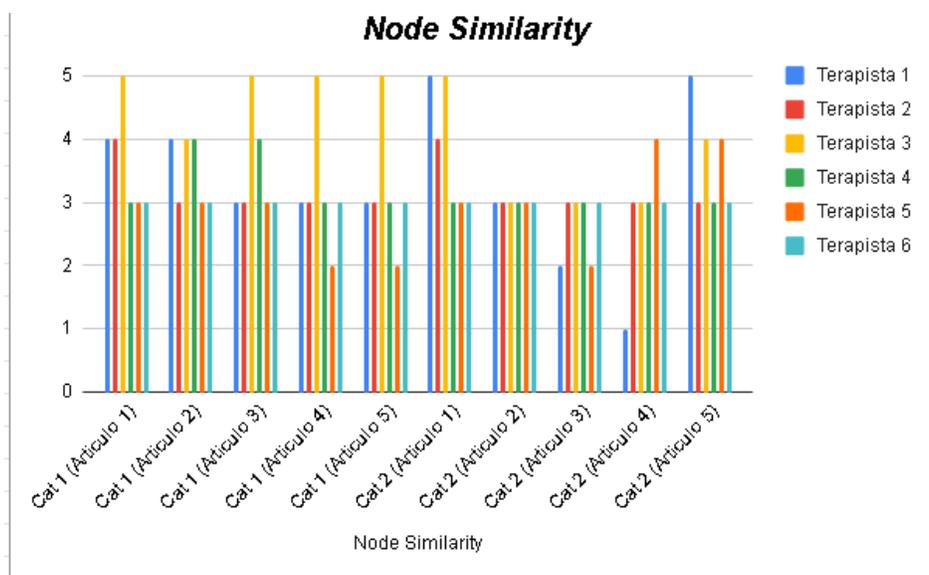


Figura 9: Gráfico de columnas de los resultados de Node Similarity realizado por los terapeutas

	Art 1	Art 2	Art 3	Art 4	Art 5	Art 1	Art 2	Art 3	Art 4	Art 5
Terapista 1	2	1	3	2	1	1	2	2	3	3
Terapista 2	4	1	3	3	2	2	3	3	2	4
Terapista 3	4	4	3	2	4	1	3	3	3	2
Terapista 4	2	2	3	3	3	3	3	3	3	2
Terapista 5	2	3	2	2	3	1	2	2	2	1
Terapista 6	3	4	3	2	3	4	3	4	4	2

Tabla XV: Resultados de los terapeutas encuestados por el algoritmo Degree Centrality

Los resultados obtenidos a través del coeficiente alfa de Cronbach del algoritmo Degree Centrality los presentamos en la tabla 16:

	Valor
K(Número de Ítems)	10
Vi(Varianza del Ítem)	7.027777778
Vt(Varianza Total)	19.80555556
Resultado	0.7562724014

Tabla XVI:Resultados a través del coeficiente alfa de Cronbach del algoritmo Degree Centrality

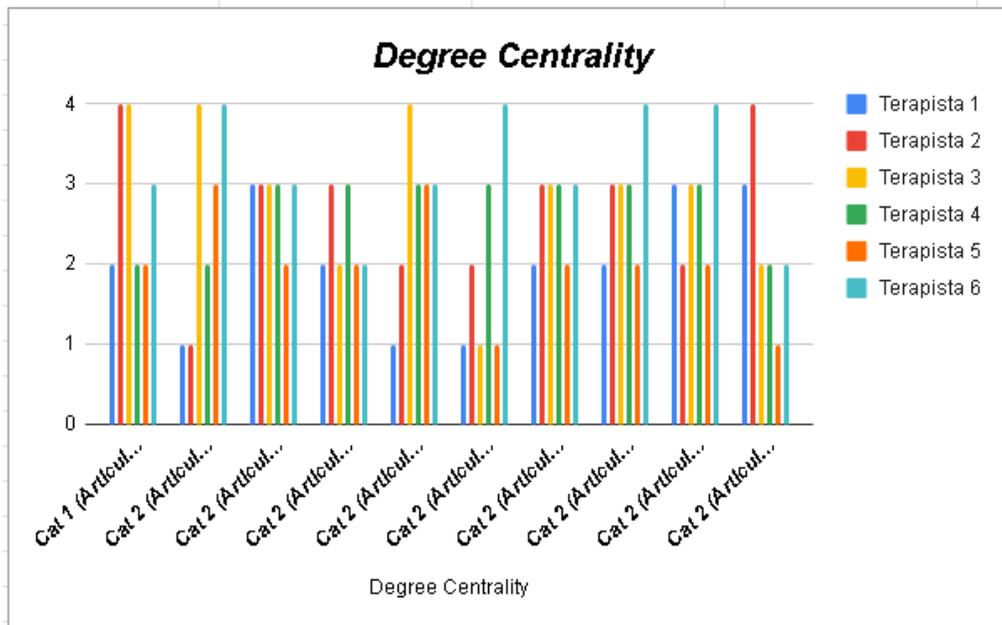


Figura 10: Gráfico de columnas de los resultados de Degree Centrality realizado por los terapeutas

	Art 1	Art 2	Art 3	Art 4	Art 1	Art 2	Art 3	Art 4
Terapista 1	5	5	5	4	4	3	3	3
Terapista 2	1	3	3	3	4	3	4	3
Terapista 3	4	4	4	4	4	5	5	5
Terapista 4	3	3	2	3	3	3	3	4
Terapista 5	3	2	3	3	1	3	3	4
Terapista 6	4	4	3	3	4	4	4	4

Tabla XVII:Resultados de los terapeutas encuestados por el algoritmo Betweenness Centrality

Los resultados obtenidos a través del coeficiente alfa de Cronbach del algoritmo Betweenness Centrality los presentamos en la tabla 18:

	Valor
K(Número de Ítems)	8
Vi(Varianza del Ítem)	6.416666667

Vt(Varianza Total)	22.80555556
Resultado	0.8614929528

Tabla XVIII: Resultados a través del coeficiente alfa de Cronbach del algoritmo Betweenness Centrality

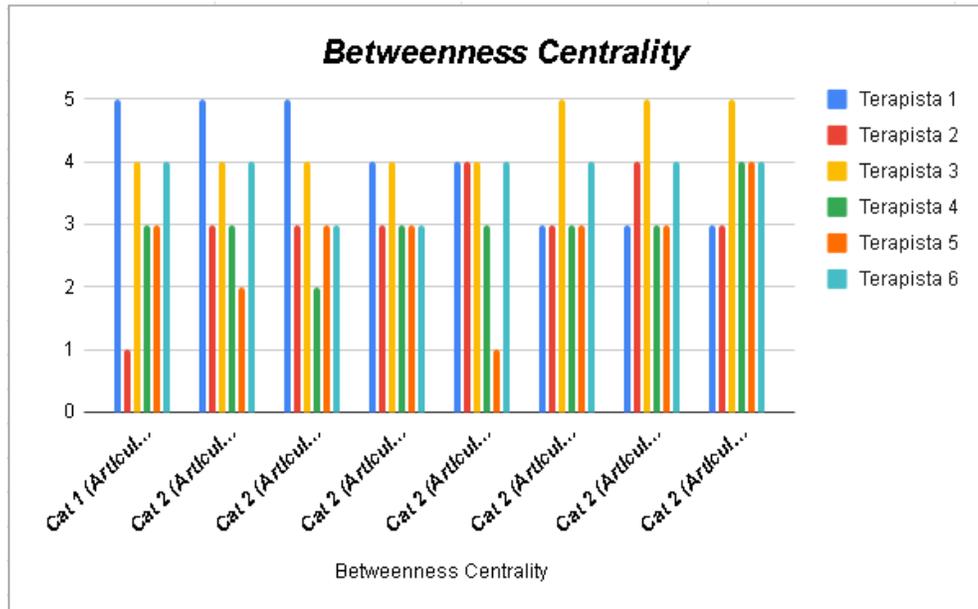


Figura 11: Gráfico de columnas de los resultados de Betweenness Centrality realizado por los terapeutas

Los resultados obtenidos en la encuesta demuestran que los tres primeros artículos son de importancia o de valor para el terapeuta. Las pruebas fueron realizadas al buscar una relación del título del artículo con la palabra buscada.

- **Diseño**

En la figura 12 el diseño del buscador es a través del algoritmo inteligente similitud de nodo que tiene una área de texto para el ingreso de las palabras claves, las palabras deben estar en inglés, puesto que se entrenó con palabras en inglés. Tiene un tiempo de respuesta entre 30 a 60 segundos.

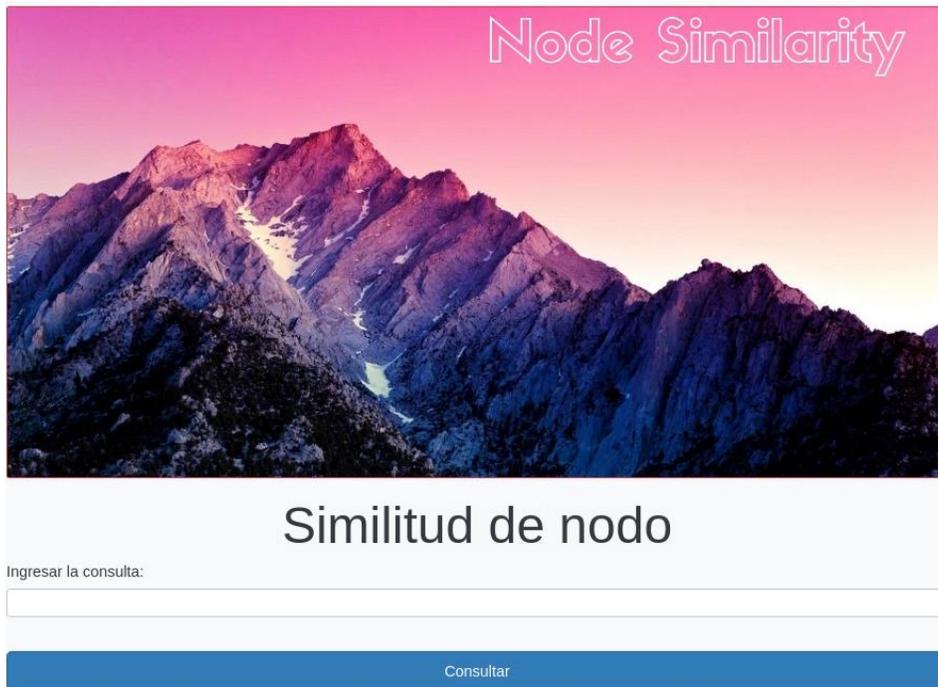


Figura 12: Página búsquedas Similitud de Nodo

En la figura 13 el diseño del buscador es a través del algoritmo inteligente degree similarity que tiene una área de texto para el ingreso de las palabras claves, las palabras deben estar en inglés, puesto que se entrenó con palabras en inglés. Tiene un tiempo de respuesta entre 10 a 30 segundos.

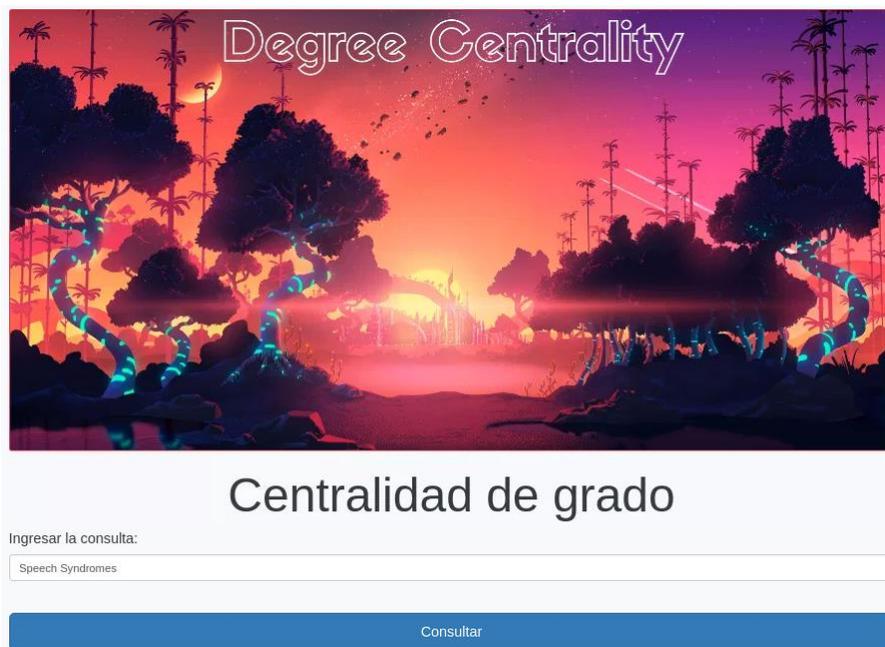


Figura 13: Página búsquedas Centralidad De Grado

En la figura 14 el diseño del buscador es a través del algoritmo inteligente degree similarity que tiene una área de texto para el ingreso de las palabras claves, las palabras deben estar en inglés, puesto que se entrenó con palabras en inglés. Tiene un tiempo de respuesta entre 10 a 30 segundos.

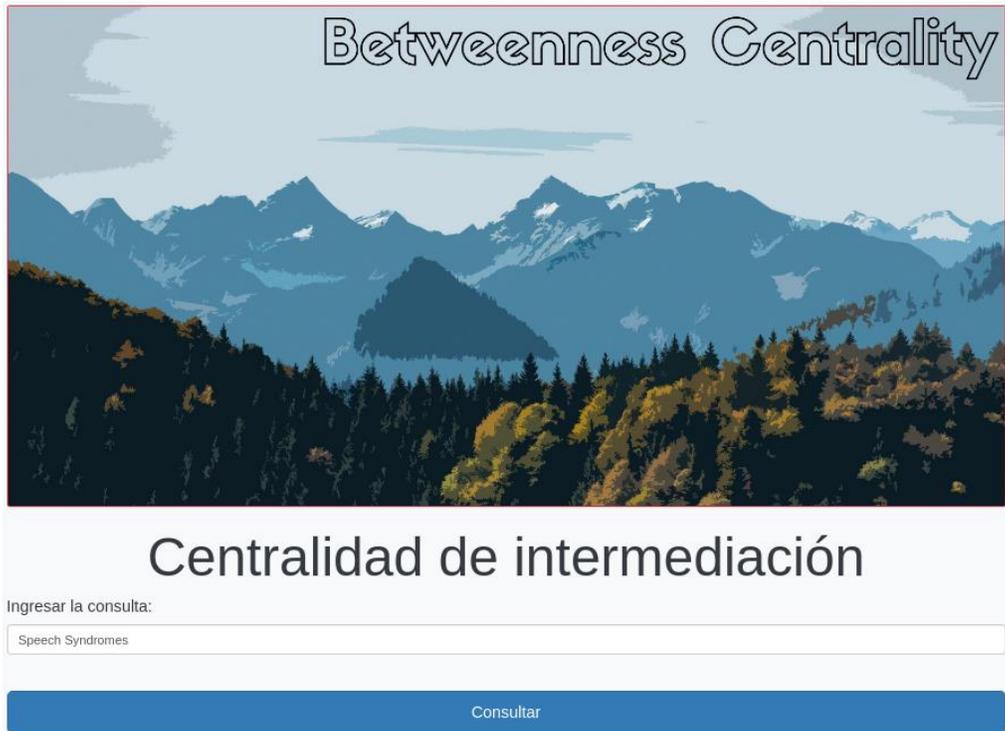


Figura 14: Página búsquedas Centralidad De Intermediación

- **Ejemplo**

El terapeuta recibe un paciente en la edad de 9 años con problemas en la comunicación el representante del paciente le comenta que tiene problemas respiratorios como musculares. El terapeuta al realizar pruebas determina que tiene problemas musculares graves y procede a realizar una investigación sobre trastornos relacionados al lenguaje con problemas musculares. El terapeuta al realizar varias búsquedas en las BD científicas no consigue un artículo científico relacionado con los síntomas del paciente es informado sobre OntoArticulo que es una página de recomendación de artículos científicos a través de ontologías.

Al ingresar en OntoArticulo se presenta brevemente la introducción al funcionamiento de los 3 tipos de algoritmos inteligentes y al revisar la información presentada en la página principal de OntoArticulo que se observa en la figura 15, realiza una búsquedas a través de palabras claves sobre los síntomas que tiene el paciente en los 3 algoritmos inteligentes mediante las palabras clave “niños con

problemas de comunicación debido a problemas musculares” los resultados obtenido presentados al terapeuta se observan en la tabla 16, 17, 18 de los primeros artículos presentados por cada algoritmo inteligente.

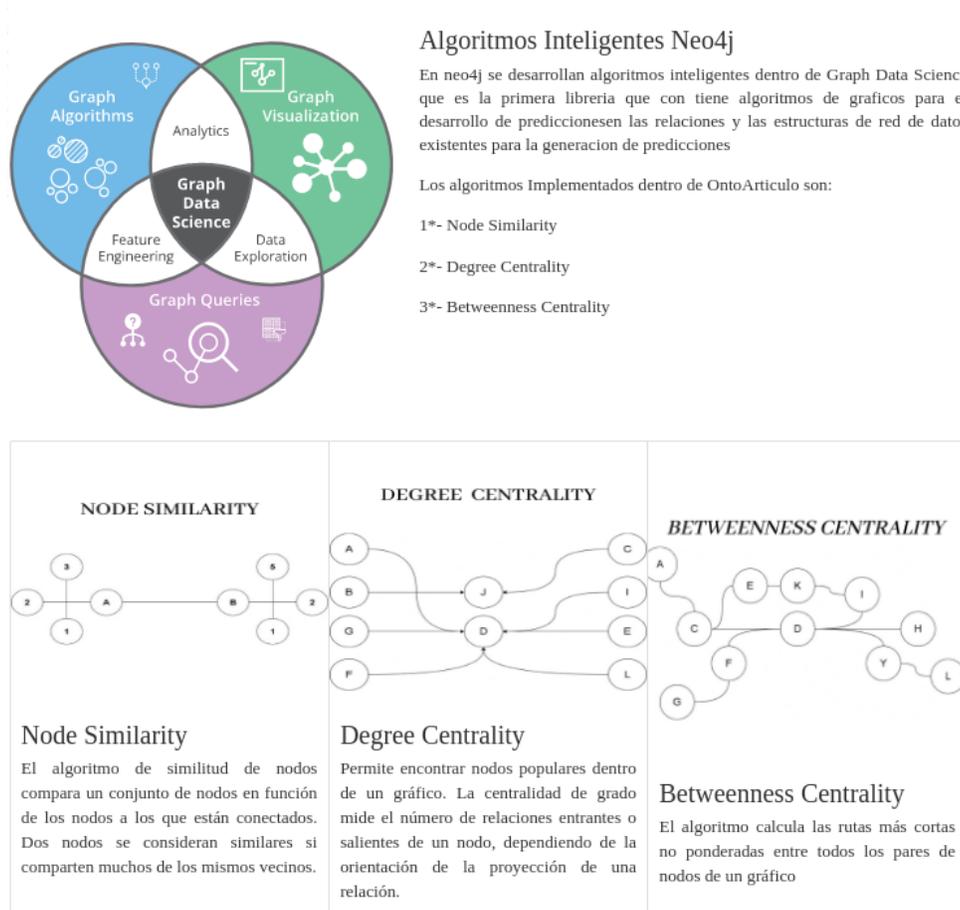


Figura 15: Página Principal de OntoArticulo

Los resultados presentados por el algoritmo inteligente Betweenness Centrality le ofrecen una mediana solución a su problema. Los contenidos son relacionados a su problema, pero no son lo que está buscando para el paciente como se observa en la figura 16.

Ingresar la consulta:

children with communication problems due to muscle problems

Consultar

#	Título	Abstracto	Fuente	Categoría
0	Método de evaluación de la enfermedad relacionada con el hígado graso, aparato de evaluación de la enfermedad relacionada con el hígado graso, método de evaluación de la enfermedad relacionada con el hígado graso, producto del programa de evaluación de la enfermedad relacionada con el hígado graso, sistema de evaluación de la enfermedad relacionada con el hígado graso, aparato terminal de comunicación de información y método de búsqueda para sustancia profiláctica / para mejorar la enfermedad relacionada con el hígado graso	Un método de evaluación de enfermedades relacionadas con el hígado graso incluye (I) un paso de obtención de datos de concentración de aminoácidos en un valor de concentración de aminoácidos en sangre recolectado de un sujeto para ser evaluado y (II) un criterio de valor de concentración que evalúa el paso de evaluación un estado de una enfermedad relacionada con el hígado graso que incluye al menos uno de hígado graso, NAFLD y NASH en el sujeto, en base a los datos de concentración de aminoácidos del sujeto obtenidos en la etapa de obtención.	uspto	enfermedad de la unidad pilosebácea
1	Detección automatizada de embolia pulmonar basada en aprendizaje profundo en	© 2021 Los autores Objetivos: La comunicación rápida de los exámenes de TC positivos para la embolia pulmonar (EP) es crucial para el inicio oportuno de la anticoagulación y el resultado del paciente. Se desconoce si la detección automatizada de aprendizaje profundo de EP en angiogramas pulmonares por tomografía computarizada (CTPA) en combinación con la priorización de la lista de trabajo y un sistema de notificación electrónica (ENS) puede mejorar los tiempos de comunicación y la respuesta del paciente en el Departamento de Emergencias (DE). Método: En 01/2010 se instaló un EHC que permite la comunicación directa entre radiología y CII. A partir de 10/2010, los CTDA se procesaron mediante un algoritmo de aprendizaje profundo en	scopus	tuberculosis abdominal

Figura 16: Resultados Obtenido por el algoritmo Betweenness Centrality

El terapeuta realiza la misma búsquedas en el algoritmo Degree Centrality examina los títulos del los artículos y observa que tienen una mediana correlación con sus búsquedas por lo que pierde el interés de revisar la información de los artículos presentados.

#	Título	Abstracto	Fuente	Categoría
0	Remisión de enterocolitis similar a la enfermedad de Crohn después del tratamiento con empaglifozina en un niño con enfermedad por almacenamiento de glucógeno tipo Ib: reporte de un caso	© 2021. Los autores. Antecedentes: Además de las principales características clínicas / bioquímicas, la neutropenia y la enfermedad inflamatoria intestinal (EII) constituyen complicaciones comunes de la enfermedad por almacenamiento de glucógeno tipo Ib (GSD Ib). Sin embargo, su gestión sigue siendo un desafío. Aunque los informes anteriores han documentado el beneficio de la administración de empaglifozina sobre la neutropenia, no se dispone de datos de seguimiento sobre la morfología intestinal (macro / microscópica). En este documento presentamos por primera vez una evaluación longitudinal de la morfología intestinal en un niño con GSD Ib que padece enterocolitis similar a la enfermedad de Crohn tratado con empaglifozina. Presentación del caso: Un niño de 14 años con GSD Ib y EII grave fue tratado (no indicado en la etiqueta) con empaglifozina (20 mg / día) después de obtener el consentimiento informado oral y escrito de los padres del paciente. No se observaron eventos adversos. Los síntomas clínicos y la frecuencia de las deposiciones mejoraron durante la primera semana de tratamiento. El índice de actividad de la enfermedad de Crohn pediátrica (PCDAI) se normalizó durante el primer mes de tratamiento. La resonancia magnética del abdomen (IRM) realizada 3 meses después del inicio del tratamiento mostró una disminución dramática en la actividad y duración de la enfermedad. Se informaron hallazgos similares en histología a los 5,5 meses. A los 7,5 meses, los niveles de hemoglobina se normalizaron y la calprotectina fecal casi se normalizó. También se observaron mejoras en el recuento de neutrófilos, el control metabólico y la calidad de vida. La dosis de G-CSF se redujo en un 33% y el paciente se retiró parcialmente de la alimentación por sonda. Conclusiones: Este es el primer informe que presenta un seguimiento extenso de la morfología gastrointestinal en un paciente con GSD Ib que recibe empaglifozina. El presente caso sugiere que la empaglifozina puede ser segura y eficaz para inducir la remisión de la EII en pacientes con GSD Ib e incluso puede posponer la cirugía. Se requieren estudios futuros para confirmar su efecto a lo largo del tiempo y evaluar su beneficio en diversas etapas de la enfermedad. Merece la pena el desarrollo de redes de colaboración internacional para la recopilación sistemática de datos.	scopus	enfermedad del sistema endocrino
1	Tumor de músculo liso asociado al virus de Epstein-Barr en un receptor de trasplante de riñón: reporte de un caso y revisión de la literatura	© 2020 Wiley Periodicals LLC Introducción: El virus de Epstein-Barr (EBV) es un herpesvirus vinculado a enfermedades linfoproliferativas premalignas y hasta nueve tumores humanos distintos. Las neoplasias malignas asociadas al VEB más frecuentes son los linfomas y el carcinoma nasofaríngeo. Al promover la proliferación del músculo liso, el VEB puede inducir tumores de músculo liso (SMT) asociados al VEB, que siguen siendo una entidad oncológica muy rara. Este estudio informa de un caso clínico de SMT y tiene como objetivo ofrecer la revisión más amplia de la literatura sobre SMT postrasplante (PT-SMT) en receptores de trasplante de riñón, con un enfoque en el manejo terapéutico y la evolución de la función del injerto. Métodos: Se recogieron reportes de casos y series de casos de PT-SMT en receptores de trasplante renal desde 1996 hasta 2019. Resultados: Se evaluaron un total de 59 PT-SMT. La mediana de tiempo en el momento del diagnóstico fue 74. 6 meses después del trasplante de riñón. Las localizaciones más frecuentes fueron hígado y pulmón. Se notificó la seroconversión del VEB en los seis pacientes con estado previamente negativo. La opción terapéutica preferida fue la cirugía (65,9%), asociada a una reducción de la inmunosupresión (77,2%), que incluye el cambio a inhibidores de mTOR (29,5%) y la suspensión de MMF (32%). En nuestra revisión, el 13% de los pacientes experimentaron rechazo, el 8,7% perdió el injerto y volvió a someterse a hemodiálisis; El 9,8% de los pacientes fallecieron por PT-SMT. Conclusión: PT-SMT es una enfermedad rara pero grave en los receptores de trasplante renal. La seroconversión del VEB después del trasplante aparece como un factor de riesgo en el desarrollo de PT-SMT en receptores de órganos sólidos. En ausencia de guías, el manejo terapéutico para PT-SMT es desafiante y expone al paciente a un alto riesgo de pérdida del injerto. Las localizaciones más frecuentes fueron hígado y pulmón. Se notificó la seroconversión del VEB en los seis pacientes con estado previamente negativo. La opción terapéutica preferida fue la cirugía (65,9%), asociada a una reducción de la inmunosupresión (77,2%), que incluye el cambio a inhibidores de mTOR (29,5%) y la suspensión de MMF (32%). En nuestra revisión, el	scopus	carcinoma de faringe in situ

Figura 17: Resultados Obtenido por el algoritmo Degree Centrality

Al revisar la información del sistema inteligente Node Similarity el terapeuta resuelve la problemática del paciente al realizar una lectura del artículo científico “La evaluación de la discapacidad motriz funcional en niños con parálisis cerebral”.

1	ESTUDIO DESCRIPTIVO SOBRE LA EVALUACIÓN DE LA DISCAPACIDAD MOTRIZ FUNCIONAL EN NIÑOS CON PARÁLISIS CEREBRAL	Antecedentes: la parálisis cerebral (PC) es una de las afecciones más comunes en la niñez que causa discapacidad física severa. La paresia espástica es la forma más común de PC. Según la clasificación topográfica, la PC se divide en hemiplejía espástica, diplejía y cuadriplejía. La distribución de la discapacidad motora funcional es variada en cada tipo de PC. Objetivo: Describir la discapacidad motora funcional en niños con parálisis cerebral utilizando escalas estándar. Método: Este estudio descriptivo transversal incluyó a 93 niños con parálisis cerebral (PC). La discapacidad motora funcional de cada tipo de PC espástica se evaluó mediante escalas estándar. Resultados: El subtipo dominante de parálisis cerebral en el presente estudio fue la diplejía espástica. El músculo más afectado con espasticidad fue el grupo de músculos gastrocnemio-sóleo. El rango de movimiento activo de eversion y dorsiflexión del pie fueron los movimientos más afectados en todos los tipos de PC. En la muestra general, solo el 35% pudo caminar de forma independiente. La mayoría de los sujetos con cuadriplejía se encontraban en los niveles III y IV de la Escala de clasificación funcional motora gruesa que representa una discapacidad grave. Se observó una relación significativa entre el tono muscular y el rango de movimiento de sus articulaciones correspondientes, así como entre el tono muscular del grupo de músculos gastrocnemio-sóleo y los componentes del tablo del Análisis de la marcha por observación. Conclusiones: Los resultados del presente estudio confirman la impresión clínica de los niveles de discapacidad en cada tipo de PC y mostraron que la evaluación de la discapacidad motora funcional en niños con diferentes tipos de PC espástica es útil en la planificación y evaluación de las opciones de tratamiento.	doi	parálisis cerebral	Revisar Artículo	Revisar Relaciones
2	Un abordaje conservador de los tumores testiculares en niños: 12 casos y su manejo	Durante un período de 34 años, se atendieron a 12 niños con tumores testiculares. De los 8 tumores benignos, 7 se trataron mediante orquiectomía y 1 solo mediante escisión focal. Los 4 tumores malignos se trataron mediante orquiectomía y 1 niño también recibió quimioterapia complementaria. Todos los pacientes están vivos y no han tenido tumores recurrentes después de 1 1/2 a 34 años de seguimiento. Debido a que los tumores testiculares suelen ser menos agresivos en los niños prepúberes, la posibilidad de curación mediante orquiectomía sola en este grupo de edad es buena. Sin embargo, el potencial de metástasis de estos tumores hace imperativo que sean observados de cerca.	ProQuest	tumor maligno de células de Leydig	Revisar Artículo	Revisar Relaciones
3	Oftalmoplejía externa progresiva crónica: I estudio histológico cuantitativo de los músculos esqueléticos	Este estudio cuantifica los principales cambios morfológicos y citoquímicos en biopsias de músculos de las extremidades de 37 pacientes con síndrome de oftalmoplejía externa progresiva crónica (CPEO). El objetivo fue evaluar el valor de la biopsia de músculo de la extremidad en el diagnóstico de este síndrome; definir los cambios miopatológicos y determinar si existían correlaciones clínico-patológicas específicas. Los pacientes se dividieron en tres grupos clínicos: 11 pacientes con CPEO con debilidad de los músculos faciales y/o de las extremidades; 10 con CPEO con debilidad muscular facial y/o de las extremidades y antecedentes familiares positivos; 16 con CPEO con uno o más de los siguientes: retinopatía pigmentaria, ataxia cerebelosa, signos piramidales y neuropatía periférica. Se midieron los siguientes parámetros: las proporciones de tipos de fibras histoquímicas, las áreas de fibras musculares y el porcentaje de fibras musculares que muestran una mayor actividad enzimática oxidativa. Se compararon los resultados agrupados para cada una de las categorías clínicas. El análisis estadístico de las áreas de fibras y el porcentaje de fibras con mayor actividad enzimática oxidativa mostró que el grupo 2 difería de los demás ($p < 0.05$). Los pacientes del grupo 2 mostraron la mayor incidencia de hipertrofia de fibras de tipo 1, atrofia de tipo 2A y la menor incidencia de fibras con mayor actividad oxidativa. Se produjeron desproporciones de tipo de fibra en los tres grupos, pero las diferencias no fueron significativas. mostró que el grupo 2 difería de los demás ($p < 0.05$). Los pacientes del grupo 2 mostraron la mayor incidencia de hipertrofia de fibras de tipo 1, atrofia de tipo 2A y la menor incidencia de fibras con mayor actividad oxidativa. Se produjeron desproporciones de tipo de fibra en los tres grupos, pero las diferencias no fueron significativas.	doi	oftalmoplejía externa progresiva	Revisar Artículo	Revisar Relaciones
4	Hipotermia local artificial en el tratamiento de niños con parálisis cerebral	La hipotermia local se utilizó para tratar los trastornos motores en la parálisis cerebral infantil. Se aplicó hielo a los músculos siendo los factores clave en las sinergias musculares patológicas y determinantes del nivel de trastornos motores en la etapa residual de la parálisis cerebral infantil. El método hipotérmico se detalla con respecto a diferentes sinergias musculares patológicas y peculiaridades del síndrome de parálisis cerebral infantil: diplejía espástica de gravedad variable, síndrome hiperkinético. Las investigaciones EMG arrojaron resultados que se correlacionaron con un efecto clínico positivo. Los datos de EMG indicaron que la privación de la aferenciación muscular patológica no solo redujo la espasticidad muscular sino que también condujo a la reconstrucción funcional en el conjunto de neuronas motoras espinales.	ProQuest	parálisis cerebral espástica	Revisar Artículo	Revisar Relaciones

Figura 18: Resultados Obtenido por el algoritmo Node Similarity

IV. Conclusiones

Los resultados presentados en este trabajo demuestran que las recomendaciones de artículos científicos a través de algoritmos inteligentes basados en ontologías son de importancia o de valor para el terapeuta que permite un nuevo enfoque de recomendación de artículos científicos a los enfoques tradicionales.

Cada algoritmo inteligente tiene un enfoque diferente pero cumplen con su cometido de realizar búsquedas de artículos científicos en el área del lenguaje y la comunicación permitiendo la recomendación a través de palabras clave, ofreciendo varios artículos científicos de utilidad para el terapeuta. De igual manera vemos una tendencia por parte de los terapeutas a los algoritmos Node Similarity y Betweenness Centrality que demuestran realizar recomendaciones fuertemente relacionadas con las necesidades de un terapeuta.

Ontoarticulo demuestra que es rápido y escalable para mejorar las recomendaciones de artículos científicos en el área. El trabajo realizado da a conocer nuevas técnicas de recomendación basadas en algoritmos inteligentes a través de una base de datos orientada a grafos con respecto a las nuevas técnicas que se pueden utilizar para mejorar los sistemas de recomendación.

V. Referencias

- Alvarez, M. M., & BARCHINI, G. E. (2008, n/a n/a). *SISTEMAS DE INFORMACIÓN BASADOS EN ONTOLOGÍAS. SISTEMAS DE INFORMACIÓN BASADOS EN ONTOLOGÍAS. UN ÁREA EMERGENTE*. Retrieved 8 5, 2021, from <https://core.ac.uk/download/pdf/15777408.pdf>
- Antonio, H. A., David, T. D., & Borja, N. C. (2015). *Una aproximación a la recomendación de artículos científicos según su grado de especificidad*. An approach to the recommendation of scientific articles according to their degree of specificity. Retrieved 7 29, 2021, from <http://rua.ua.es/dspace/handle/10045/49280>
- Artigot, R. A. (2017). *Recomendación contextualizada usando una ontología de contexto genérica y de gran escala construida semiautomáticamente a partir de DBpedia*. Recomendación contextualizada usando una ontología de contexto genérica y de gran escala construida semiautomáticamente a partir de DBpedia. Retrieved 8 5, 2021, from <https://repositorio.uam.es/handle/10486/681054>
- Belsky, G., & G. (2021, 05 24). *¿Qué es la terapia del habla?* Understood. Retrieved 8 26, 2021, from <https://www.understood.org/articles/es-mx/what-you-need-to-know-about-speech-therapy>
- Brut, M. B., & Sedes, F. S. (2010, 7 5). *Ontology-Based Solution for Personalized Recommendations in E-Learning Systems. Methodological Aspects and Evaluation Criterias*. IEEEExplore. Retrieved 8 5, 2021, from <https://bibliotecas.ups.edu.ec:2095/document/5572209>
- Colaborador BioPortal. (2021). *Mondo Disease Ontology | NCBO BioPortal*. BioPortal. Retrieved 8 26, 2021, from <https://bioportal.bioontology.org/ontologies/MONDO>
- Colaboradores BioPortal. (2018, 07 07). *VIVO Ontology for Researcher Discovery | NCBO BioPortal*. BioPortal. Retrieved 8 26, 2021, from <https://bioportal.bioontology.org/ontologies/VIVO>
- colaboradores de Wikipedia. (2021, 08 14). *Google Noticias*. Wikipedia. Retrieved 08 18, 2021, from https://es.wikipedia.org/wiki/Google_Noticias
- Colaboradores Grapheverywhere. (2019, 12 02). *Sistemas de recomendación | Qué son, tipos y ejemplos*. Graph. Retrieved 8 26, 2021, from <https://www.grapheverywhere.com/sistemas-de-recomendacion-que-son-tipos-y-ejemplos/>
- Colaboradores Neo4j. (2021). *Node Similarity*. Neo4j Docs. Retrieved 8 18, 2021, from <https://neo4j.com/docs/graph-data-science/current/algorithms/node-similarity/>

Doaj. (2018, 11 09). *¿Qué es el DOAJ?* Doaj. Retrieved 8 28, 2021, from <https://blog.doaj.org/2018/11/09/que-es-doaj/>

Fosado, G. (2019, 07 24). *Primeros pasos en NLP con spaCy, un vistazo general*. Medium. Retrieved 8 27, 2021, from <https://yeralway1.medium.com/primeros-pasos-en-nlp-con-spacy-un-vistazo-general-734686843a57>

gazteaukera. (2019). *RDF (Resource Description Framework)*. gazteaukera. Retrieved 8 28, 2021, from https://www.gazteaukera.euskadi.eus/contenidos/informacion/opendata_rdf_euskadi/es_info/adjuntos/RDF.pdf

grapheverywhere. (2020, 03 10). *Qué son los grafos*. grapheverywhere. Retrieved 08 26, 2021, from <https://www.grapheverywhere.com/que-son-los-grafos/>

Huaman, D. G. (2018, 10). *Diseño de un sistema de recomendación de libros y tesis basado en ontologías asociadas a tesauros*. Diseño de un sistema de recomendación de libros y tesis basado en ontologías asociadas a tesauros. Retrieved 8 5, 2021, from https://www.researchgate.net/publication/329957450_Disenio_de_un_sistema_de_recomendacion_de_libros_y_tesis_basado_en_ontologias_asociadas_a_tesauros_el_caso_de_las_bibliotecas_de_la_UNMSM

Mena, G. G. (2018). Comparativa de tecnicas para el descubrimiento de conocimiento aplicadas al comportamiento de produccion cientifica. *Unir*, 1(1), 88. <https://reunir.unir.net/bitstream/handle/123456789/9631/Gal%C3%A1n%20Mena%2C%20Jorge%20Andr%C3%A9s.pdf?sequence=1&isAllowed=y>

Mena, G. M. (2018). *Comparativa de tecnicas para el descubrimiento de conocimiento aplicadas al comportamiento de produccion cientifica*. reunir. Retrieved 8 18, 2021, from <https://reunir.unir.net/bitstream/handle/123456789/9631/Gal%C3%A1n%20Mena%2C%20Jorge%20Andr%C3%A9s.pdf?sequence=1&isAllowed=y>

Neo4j. (2019). *Betweenness Centrality - Neo4j Graph Data Science*. Neo4j. Retrieved 8 26, 2021, from <https://neo4j.com/docs/graph-data-science/current/algorithms/betweenness-centrality/>

Neo4j. (2019). *Degree Centrality - Neo4j Graph Data Science*. Neo4j. Retrieved 8 26, 2021, from <https://neo4j.com/docs/graph-data-science/current/algorithms/degree-centrality/>

Neo4j. (2019). *Node Similarity*. Neo4j. Retrieved 8 26, 2021, from <https://neo4j.com/docs/graph-data-science/current/algorithms/node-similarity/>

Neo4j. (2021, 05 16). *Graph Data Science Algorithms | Graph Data Analysis | Neo4j*. Neo4j. Retrieved 8 26, 2021, from <https://neo4j.com/product/graph-data-science-library/>

Puzo, M. P. (1990). *La Amistad*. Proverbias: Frases y citas célebres. Retrieved 8 14, 2021, from <https://proverbias.net/cita/453059647-la-amistad-lo-es-todo-la-amistad-vale-mas-que-el->

- Sun, J. S., & Ma, J. M. (2014, 10). *Leveraging Content and Connections for Scientific Article Recommendation in Social Computing Contexts*. IEEEExplore. Retrieved 8 8, 2021, from <https://bibliotecas.ups.edu.ec:2095/document/8130565>
- Sun, J. S., Ma, J. M., & Liu, Z. L. (2014, 10). *Aprovechar el contenido y las conexiones para la recomendación de artículos científicos en contextos de informática social*. IEEEExplore. Retrieved 8 8, 2021, from <https://bibliotecas.ups.edu.ec:2095/document/8130565>
- Sun, J. S., Ma, J. M., & Lyu, Z. L. (2013, 10 7). *A Novel Approach for Personalized Article Recommendation in Online Scientific Communities*. IEEEExplore. Retrieved 8 7, 2021, from <https://bibliotecas.ups.edu.ec:2095/document/6480025>
- Utrera, E. B., & Cuevas, A. J. (2017, 04 10). *Sistemas de recomendación semánticos: Una revisión del Estado del Arte*. *Sistemas de recomendación semánticos: Una revisión del Estado del Arte*. Retrieved 07 29, 2021, from http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2227-18992017000200014
- Xia, F. X. (2016). *Scientific Article Recommendation: Exploiting Common Author Relations and Historical Preferences*. IEEEEXPLORE. Retrieved 8 6, 2021, from <https://bibliotecas.ups.edu.ec:2095/document/7458844>

VI. Anexos

Anexo 1

- Encuestas realizadas a los terapeutas
 - Link de la encuesta realizada a todos los terapeutas en el área del lenguaje y la comunicación.
 - https://docs.google.com/forms/d/e/1FAIpQLSeQIpdq26kP8x0eUeYFuz9erWnkByaWzQQCxSDpBnPIQkQm8g/viewform?usp=sf_link
- Preguntas
 - **Pregunta 1**
 - Se presentan 5 artículos científicos buscados en ontoartículo por medio de la palabra clave “oftalmoplejía y síndromes” con el algoritmo Node Similarity, no se presentan los artículos por presencia de plagio.
 - Del texto anterior cómo calificaría la recomendación del algoritmo Node Similarity mediante la palabra clave? (Ophthalmoplegia and Syndromes-Oftalmoplejía y síndromes)
 - **Pregunta 2**

