

Biblans: aplicación para gestionar, compartir representar e interpretar los datos de artículos para estudios bibliométricos.

Biblans: an application for managing, sharing, representing and interpreting article data for bibliometric studies.

Annier Jesús Fajardo Quesada,¹ René Herrero Pacheco,² Rolando Javier Álvarez Pérez,³ Eduardo Antonio Hernández Gonzáles.⁴

¹ Estudiante de 6to año de Medicina. Universidad de Ciencias Médicas de Granma. Facultad de Ciencias Médicas de Bayamo. Granma, Cuba. annierfq01@gmail.com, <https://orcid.org/0000-0002-2071-3716>

² Residente R1 en Oftalmología. Hospital Clínico Quirúrgico General Carlos Manuel de Céspedes. Granma, Cuba. c1101zion@nauta.com.cu, <https://orcid.org/0000-0002-9450-1572>

³ Estudiante de 6to año de Medicina. Universidad de Ciencias Médicas de Granma. Facultad de Ciencias Médicas Celia Sánchez Manduley. Granma, Cuba. rolialvarez19@gmail.com, <https://orcid.org/0000-0003-0606-2993>

⁴ Estudiante de 6to año de Medicina. Universidad de Ciencias Médicas de Pinar del Río. Facultad de Ciencias Médicas Dr. Ernesto Che Guevara de la Serna. Pinar del Río, Cuba. eduardohernandezgonzalez10@gmail.com, <https://orcid.org/0000-0001-7325-6099>

Autor para correspondencia: Annier Jesús Fajardo Quesada annierfq01@gmail.com

RESUMEN

Introducción: Los estudios bibliométricos han ganado popularidad con el aumento de publicaciones científicas. Sin embargo, no existe un software que gestione eficazmente todos los datos y permita su análisis, representación e interpretación. **Objetivo:** diseñar una aplicación para recolectar, compartir, representar e interpretar datos de artículos en estudios bibliométricos. **Método:** se realizó una investigación de desarrollo tecnológico en agosto de 2024. La aplicación se diseñó en Ubuntu 14.04 utilizando Python 3.12, C++17 y MySQL. Se emplearon Pandas y Matplotlib para el tratamiento de datos, FastAPI para el backend, y QtCreator para la interfaz. La aplicación se convirtió en un ejecutable de Windows con auto-py-to-exe. **Resultados:** la aplicación creada a pesar de funcionar completamente offline requiere una conexión inicial para cargar los datos de usuario y licencia, y para habilitar el modo de intercambio colaborativo de usuarios afiliados a los proyectos. Su sistema de detección de partes del artículo con el modelo de Inteligencia artificial Grobid unido al sistema de scraping de GoogleScholar, Mendeley y otros fueron un punto fuerte en la creación de la base de datos del proyecto. Permite el análisis de datos usando las potencialidades de rapidez de C++ y la representación usando la versatilidad de

las librerías de Python. Conclusiones: la aplicación es una herramienta útil para gestionar, compartir, representar e interpretar datos bibliométricos. Integra tecnologías potentes como Grobid, demostrando ser eficaz en su propósito.

Palabras clave: Bibliometría; Diseño de Software; Literatura; Programas Informáticos; Tecnología.

ABSTRACT

Introduction: Bibliometric studies have gained popularity with the increase in scientific publications. However, there is no software that effectively manages all the data and allows its analysis, representation and interpretation. Objective: to design an application to collect, share, represent and interpret article data in bibliometric studies. Method: a technological development research was carried out in August 2024. The application was designed in Ubuntu 14.04 using Python 3.12, C++17 and MySQL. Pandas and Matplotlib were used for data processing, FastAPI for the backend, and QtCreator for the interface. The application was converted into a Windows executable with auto-py-to-exe. Results: the application created despite running completely offline requires an initial connection to load user and license data, and to enable the collaborative sharing mode of users affiliated to the projects. Its article part detection system with the Grobid artificial intelligence model coupled with the scraping system of GoogleScholar, Mendeley and others was a strong point in the creation of the project database. It allows data analysis using the speed potential of C++ and representation using the versatility of Python libraries. Conclusions: The application is a useful tool for managing, sharing, representing and interpreting bibliometric data. It integrates powerful technologies such as Grobid, proving to be effective in its purpose.

Keywords: Bibliometrics; Software Design; Literature; Software; Technology.

INTRODUCCIÓN

La bibliometría, disciplina que emplea herramientas matemáticas y estadísticas, se ha consolidado como un pilar fundamental en la evaluación y categorización de la producción científica. A través del análisis de publicaciones, revistas y autores, permite trazar un mapa detallado de la ciencia y sus interconexiones. La esencia de la bibliometría reside en la capacidad de transformar datos brutos en conocimiento significativo, utilizando para ello índices bibliométricos que cuantifican y cualifican el impacto y la relevancia de la investigación.⁽¹⁾

Históricamente, el análisis bibliométrico se realizaba de forma manual, un proceso laborioso y propenso a errores. Sin embargo, la revolución digital ha transformado radicalmente esta disciplina. La aparición de herramientas informáticas avanzadas, integradas en bases de datos científicas y softwares especializados, ha impulsado una evolución sin precedentes.⁽²⁾ Esta transformación ha democratizado el acceso a la bibliometría, permitiendo a investigadores de diversas áreas explorar y comprender la dinámica de la ciencia con mayor profundidad y precisión.

En la actualidad, existe una proliferación de aplicaciones diseñadas para asistir a los investigadores en sus proyectos bibliométricos. Estas herramientas abarcan desde la recolección de datos esenciales, como palabras clave, resúmenes y autores, hasta la visualización avanzada de datos mediante graficadores especializados. No obstante, a pesar de esta diversidad de herramientas, persiste una carencia fundamental: la ausencia de un software integral que permita recolectar, interpretar y visualizar datos de forma unificada. Esta fragmentación del flujo de trabajo obliga a los investigadores a utilizar múltiples aplicaciones, lo que genera ineficiencias y dificulta la realización de estudios bibliométricos exhaustivos.

Surge entonces la necesidad de crear una aplicación que integre las funcionalidades de diversas herramientas bibliométricas, permitiendo el tratamiento integral de datos de artículos científicos de manera eficiente, por lo que se realiza la presente investigación, con el objetivo de diseñar una aplicación para recolectar, compartir, representar e interpretar datos de artículos en estudios bibliométricos.

MÉTODOS

Se realizó una investigación de desarrollo tecnológico o aplicado entre los meses de agosto del 2023 a mayo del 2024.

El proceso de creación del software se llevó a cabo en 4 grandes pasos, con vistas a garantizar una posterior escalabilidad del proyecto, así como bases sólidas para su proceso de creación.

Conceptualización, planificación y recolección de requisitos

Se realizó una revisión bibliográfica exhaustiva usando de manera automatizada el motor de búsqueda Google Académico incluyeron los términos: indicadores bibliométricos, bibliometría, cienciometría, aplicaciones informáticas y otros términos afines al tema de la revisión tanto en español como sus traducciones al inglés.

Se recolectó información referente a otras aplicaciones con fines similares, así como la disponibilidad de recursos de libre acceso para ser incluidos en la aplicación.

En este paso se hizo el listado de todas las futuras funcionalidades que se intentarían implementar.

Selección de las herramientas de programación y diseño del software

Se seleccionó para la confección de la aplicación el lenguaje de programación Python 3.12, utilizando como editor de código Microsoft Visual Studio Code 1.60. Para crear la interfaz gráfica el módulo externo PyQt6 y para trabajar sobre los ficheros .ui (extensión de los ficheros de apariencia de Qt) con el software QtCreator 5.0.2(Community).

Algunos datos fueron procesados por aplicaciones de línea de comandos creadas con C++, por la velocidad que ofrece el mismo frente a Python. Se uso como compilador del mismo CodeBlocks 20.03 en el estandar cpp-17.

Para el BackEnd se utilizó FastAPI, y para almacenar las bases de datos SQL con el gestor de bases de datos MariaDB.

Se optó por un modelo de arquitectura de software Modelo-Vista-Controlador.

Implementación

Se hizo en tres partes. Primero se creó la interfaz gráfica, luego la API y posteriormente se le dio funcionalidad a la aplicación de escritorio, así como la conexión de la misma con la API por protocolo HTTPS.

Pruebas y aprobación del producto

Se valoró vía informática. Sobre la interfaz gráfica se hizo el testing usando pytest-qt, y sobre el backend se hizo usando el propio sistema de documentación de FastAPI de OpenAPI.

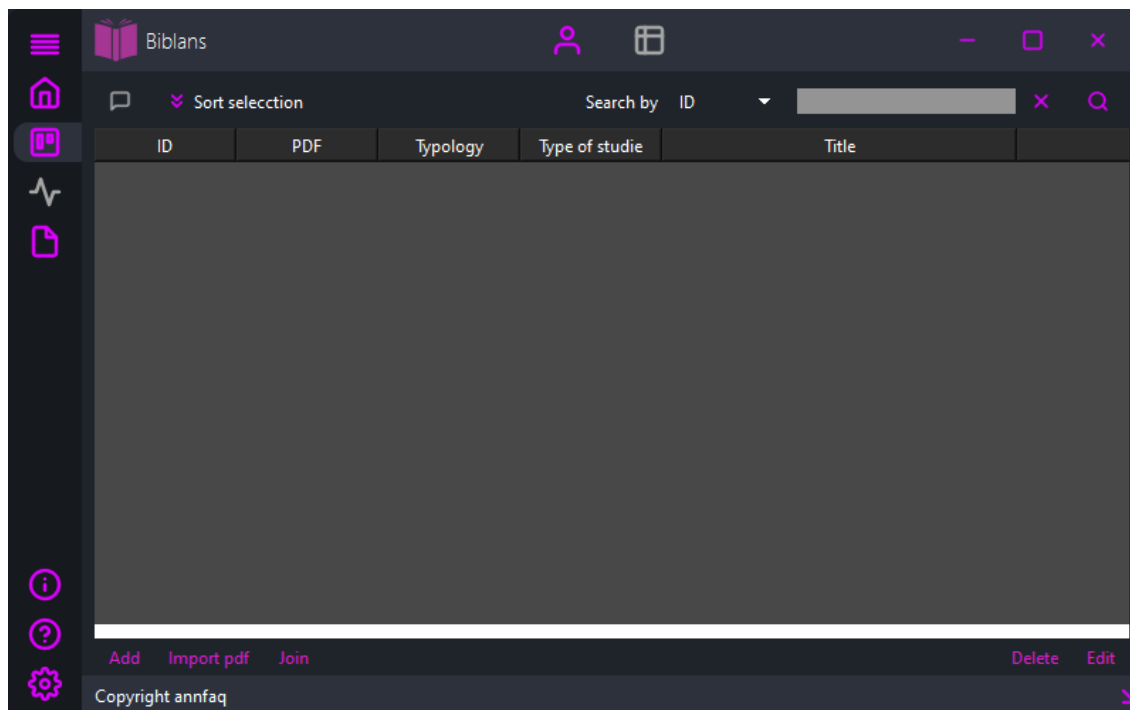
De igual manera se probó en dispositivos con diferentes especificaciones técnicas para evidenciar la existencia de dependencias adicionales u otros problemas técnicos que pudieran surgir en los mismos.

RESULTADOS

La aplicación se construyó teniendo en cuenta que un proyecto de estudio bibliométrico está compuesto por varios artículos o entradas, que son el objeto de análisis de la misma.

Se creó una aplicación para sistemas Windows, con una interfaz gráfica amigable de fácil utilización, con un diseño predeterminado en colores oscuros, esta se muestra en la Figura 1.

Figura 1. Proyecto en blanco en BiblAnS.



Fuente: Elaboración propia.

Cuenta con una barra de título(superior) dividida en tres bloques:

- 1- Logo y título.
- 2- Usuario y proyectos de usuario(Solo se habilita al activarse o registrarse el usuario)
- 3- Control general de ventana(minimizar, maximizar/restaurar, cerrar)

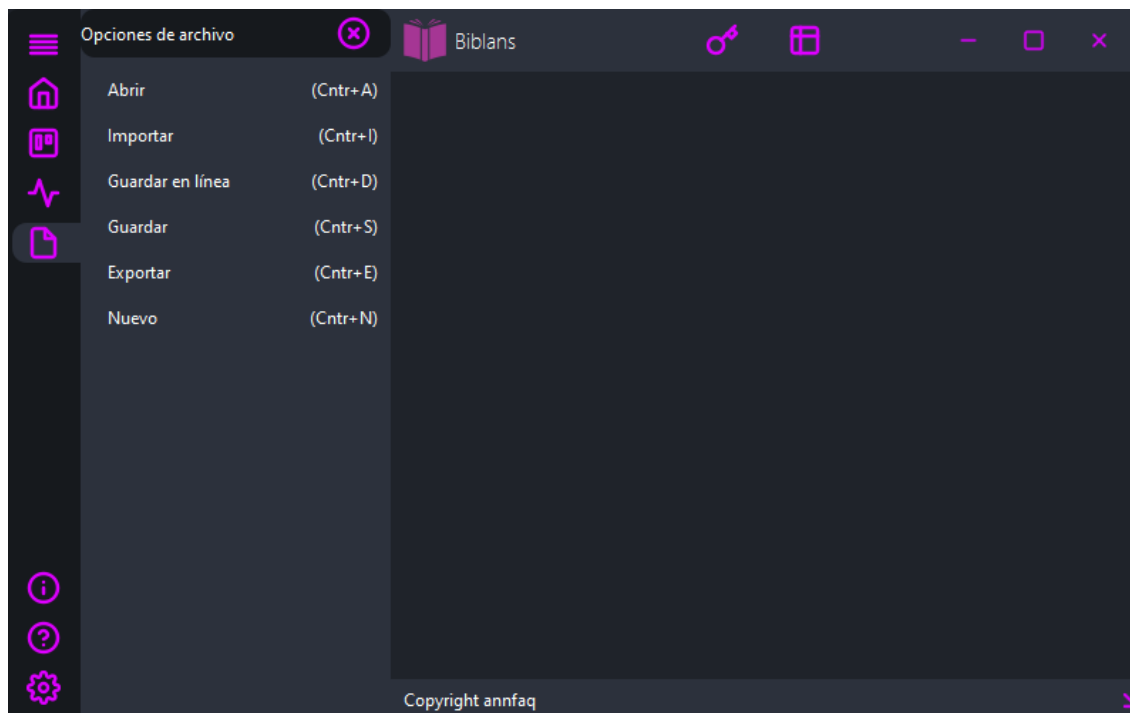
Una barra lateral o de menú dividida en dos bloques:

- 1- Control de proyecto: (inicio, base de datos, representación, opciones del proyecto)
- 2- Información/ayuda y ajustes de la aplicación

Una barra de propiedad (inferior).

En el área central se encuentra la zona de trabajo, donde se mostrara el proyecto, gráficos y otros elementos de configuración de los mismos. En la figura 1 se observa en el área de trabajo un proyecto en blanco, el cual cuenta con opciones de búsqueda y filtrado de artículos (superior) y las opciones de agregar artículos (inferior izquierda) y borrar o editar artículos (inferior derecha). En la figura 2 se observan las opciones generales del archivo de proyecto, con sus teclas de acceso o shortcut.

Figura 2. Opciones del archivo del proyecto.



Fuente: Elaboración propia.

Los artículos del proyecto pueden agregarse a partir de un pdf, donde se extraerán los metadatos del mismo o se hará reconocimiento automático de sus partes con Grobid, o por entrada manual como se observa en la figura 3.

Figura 3. Entrada manual de artículos.

Article datas

Language: English Cites: 0

Typology: Original Descriptive

Title:

Abstract:

Keywords:

Authors:

Institutions:

References:

Journal: Country:

Year: 2022 Volume: 400 Issues: 2022

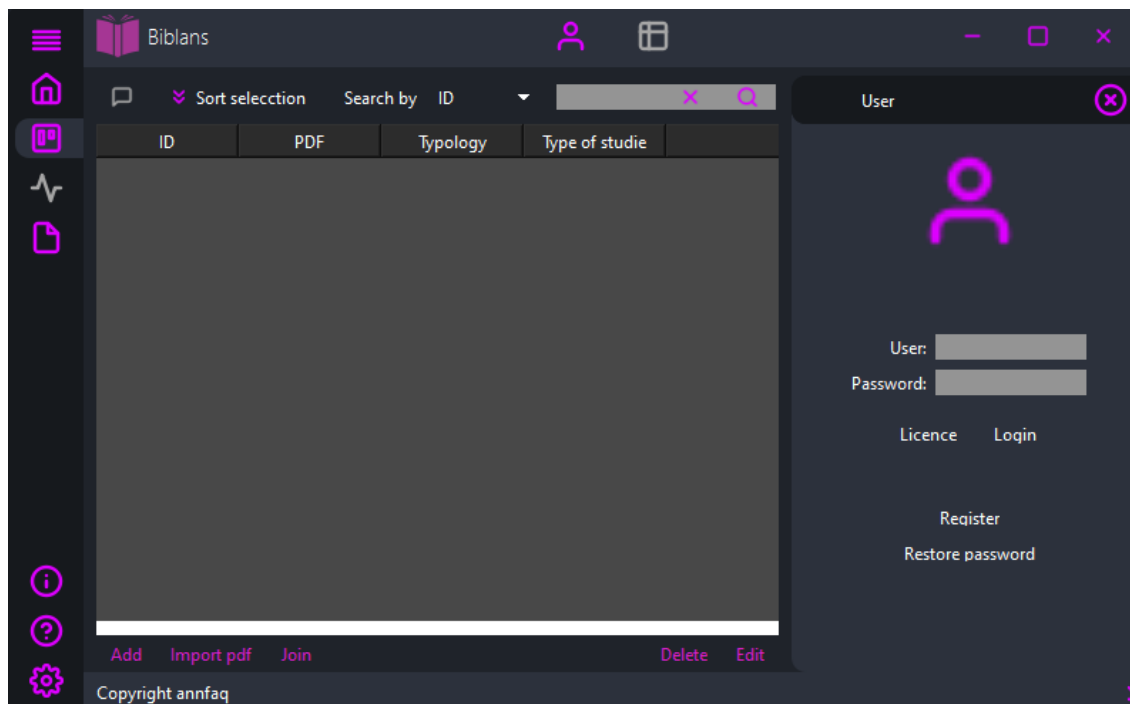
Send: Accepted: Published:

Autofill Verify Cancel Accept

Fuente: Elaboración propia.

La entrada manual cuenta con una opción de auto llenado a partir del título, una vez puesto el mismo la aplicación buscara en algunas bases de dato el mismo y completara el resto de campos. La aplicación cuenta con sistema de autenticación y registro de usuarios como se muestra en la figura 4. Esto posibilita la colaboración de usuarios en un mismo proyecto así como el sistema de licencias y privilegios de la misma.

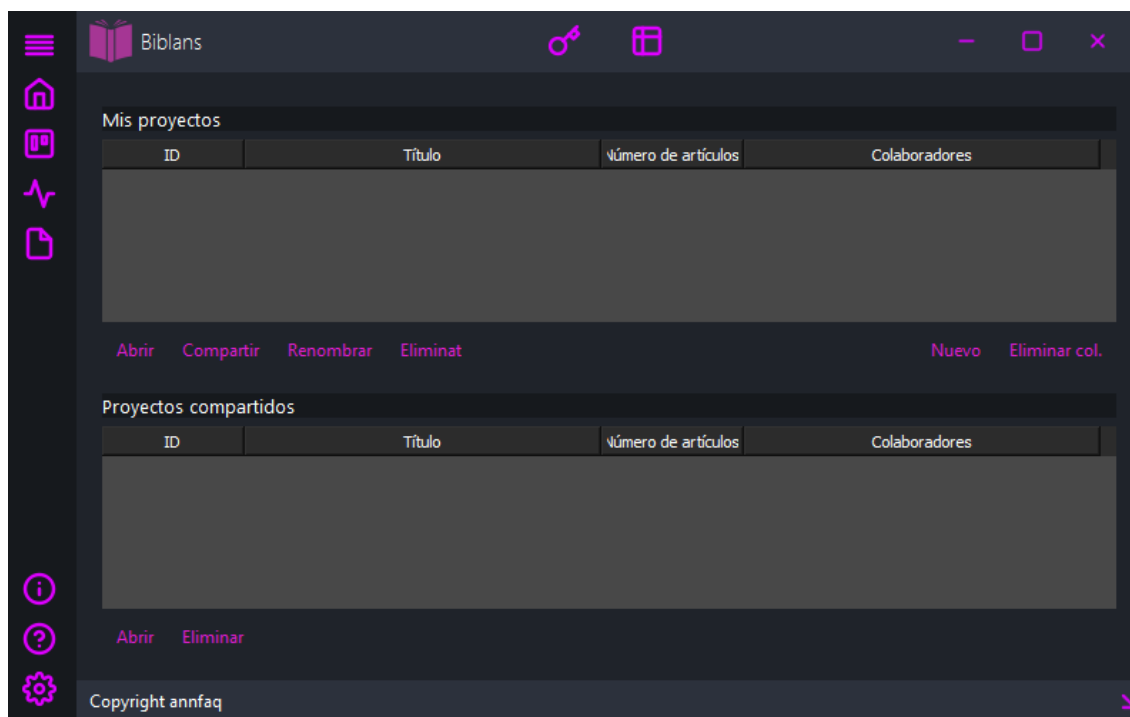
Figura 4. Pestaña de usuario.



Fuente: Elaboración propia.

Una vez registrado y conectado a internet se puede acceder a los proyectos guardados por el usuario, así como aquellos en los que es miembro (figura 5).

Figura 5. Proyectos de usuario.



Fuente: Elaboración propia.

Todos los proyectos cuentan con un chat el cual es compartido entre sus colaboradores. Este se observa en la esquina superior izquierda de la zona de trabajo del proyecto.

Una vez cargados los artículos del proyecto se pueden analizar los resultados en el área representación de datos (figura 6). En esta se pueden seleccionar diferentes tipos de representación de dichos resultados (texto, tablas, gráficos, grafos, mapas e informes).

Figura 6. Área de representación de datos.



Fuente: Elaboración propia.

La aplicación fue probada en sistema operativo Windows 11, Windows 10, Windows 8 y Windows 7. Con características de ordenador mínima Intel Pentium Silver 4Gb de RAM 8 núcleos a 1.10 GHz. Requiere espacio en disco de 1Gb mínimo, sin dependencias externas.

DISCUSIÓN

En la actualidad, existe una amplia gama de aplicaciones diseñadas para facilitar los estudios bibliométricos, abarcando desde la recolección de datos hasta el análisis avanzado y la visualización.⁽²⁾ Sin embargo, muchas de estas herramientas se centran en áreas específicas, lo que obliga a los investigadores a utilizar múltiples plataformas para completar un estudio. Esta fragmentación del flujo de trabajo puede generar inefficiencias y dificultades en la gestión de datos. La aplicación desarrollada busca abordar esta problemática mediante la integración de todas estas funcionalidades en una única plataforma unificada, simplificando así el proceso de investigación y mejorando la eficiencia general.

Uno de los principales desafíos al trabajar con diversas aplicaciones bibliométricas es la potencial incompatibilidad entre los formatos de archivo nativos. Esta falta de interoperabilidad puede llevar a la pérdida de información durante las conversiones de formato. Por ejemplo, la conversión de archivos de PubMed (.txt) a formato RIS puede resultar en la eliminación de resúmenes, un elemento crucial para el análisis de coocurrencia.⁽³⁾ La presente aplicación resuelve este problema al proporcionar un entorno integrado que minimiza la necesidad de conversiones de formato y garantiza la integridad de los datos a lo largo de todo el proceso de investigación.

La recolección y organización de información bibliográfica pueden ser tareas tediosas y consumir mucho tiempo. Afortunadamente, existen herramientas como Grobid, que utilizan modelos de inteligencia artificial para la detección automática de metadatos en artículos científicos. Grobid ha demostrado ser altamente preciso en la identificación de elementos como autores, instituciones y títulos. Biblans aprovecha las capacidades de Grobid a través de una API externa, lo que nos permite ofrecer una extracción de metadatos eficiente y precisa. Sin embargo, teniendo en cuenta el rápido avance de la IA se mantiene la flexibilidad para incorporar modelos más avanzados en el futuro, como Meta-llama-3, a medida que estén disponibles.⁽⁴⁾

La representación visual de los datos es esencial para el análisis bibliométrico. Python, con su popular librería Matplotlib, ofrece una amplia gama de opciones para la visualización de datos.⁽⁵⁾ Sin embargo, para la creación de grafos y mapas, que son fundamentales en los estudios bibliométricos, se optó por utilizar las librerías especializadas Graph y Geopandas. Estas librerías proporcionan funcionalidades avanzadas para la visualización de datos espaciales y de redes, lo que permite generar representaciones visuales más completas y detalladas.

La investigación bibliométrica a menudo implica la colaboración entre múltiples investigadores.⁽⁶⁾ Las redes sociales científicas y los grupos de investigación desempeñan un papel crucial en la facilitación de esta colaboración, permitiendo la distribución de tareas y el intercambio de conocimientos.⁽⁷⁾ Biblans incorpora funcionalidades avanzadas para la creación y gestión de proyectos colaborativos, así como herramientas de comunicación integradas. Esto la convierte en una verdadera red social científica especializada, que fomenta la colaboración y el intercambio de ideas entre investigadores.

El costo de acceso a las aplicaciones y bases de datos bibliométricas puede ser un obstáculo para muchos investigadores. Plataformas como Scopus y Web of Science, así como las APIs de Google Scholar y Mendeley, a menudo requieren suscripciones o pagos por acceso a datos.⁽⁸⁾ La presente aplicación busca democratizar el acceso a la investigación bibliométrica al ofrecer recursos gratuitos y un sistema de licencias asequible. Si bien se reconoce el valor de las bases de datos y APIs más potentes, el enfoque principal de la aplicación es proporcionar una plataforma accesible y de bajo costo para investigadores de todos los niveles.

CONCLUSIONES

Se creó una aplicación de escritorio para Windows con la capacidad de ayudar en el desarrollo de investigaciones bibliométricas, facilitando la recolección, representación y análisis de los datos. Implementa las mejores opciones de las aplicaciones actualmente usadas para esos fines proporcionando un entorno de desarrollo integrado. Posibilita el desarrollo de proyectos en comunidad, mediante el intercambio de información entre sus afiliados.

REFERENCIAS BIBLIOGRÁFICAS

1. Manzo JC, Fabian F, González B, Carlos L, Barba O, Sergio F, et al. Bibliometría del Uso del Blockchain en la Economía Circular: Decisión Hacia la Sostenibilidad. *Cienc Lat Rev Científica Multidiscip* [Internet]. 2024 [citado 2 de marzo de 2025];8(5):9165-80. Disponible en: <https://www.ciencialatina.org/index.php/cienciala/article/view/14304/20476>
2. Iyelolu TV, Agu EE, Idemudia C, Ijomah TI. Driving SME innovation with AI solutions: overcoming adoption barriers and future growth opportunities. *Int J Sci Technol Res Arch* [Internet]. 2024 [citado 2 de marzo de 2025];7(1):036-54. Disponible en: <https://shibata.yubetsu.com/article/3fjdqymC>
3. Bukar UA, Sayeed MS, Razak SFA, Yogarayan S, Amodu OA, Mahmood RAR. A method for analyzing text using VOSviewer. *MethodsX* [Internet]. 2023 [citado 2 de marzo de 2025];11:102339. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2215016123003369>
4. Masalkhi M, Ong J, Waisberg E, Zaman N, Sarker P, Lee AG, et al. A side-by-side evaluation of Llama 2 by meta with ChatGPT and its application in ophthalmology. *Eye* 2024 3810 [Internet]. 2024 [citado 2 de marzo de 2025];38(10):1789-92. Disponible en: <https://www.nature.com/articles/s41433-024-02972-y>
5. Pahira Dongoran N, Cahyati Sitorus Pane A, Amalia Wardani S, William Iskandar Ps J V, Baru K, Percut Sei Tuan K, et al. Pemanfaatan MATLAB dalam Analisis Turunan Parsial : Visualisasi dan Implementasi Fungsi Multivariat. *J Pengabd Masy Sains dan Teknol* [Internet]. 2024 [citado 2 de marzo de 2025];3(4):92-7. Disponible en: <https://ftuncen.com/index.php/JPMsAINTEK/article/view/638>
6. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: An overview and guidelines. *J Bus Res*. 2021;133:285-96.
7. Thornhill-Miller B, Camarda A, Mercier M, Burkhardt JM, Morisseau T, Bourgeois-Bougrine S, et al. Creativity, Critical Thinking, Communication, and Collaboration: Assessment, Certification, and Promotion of 21st Century Skills for the Future of Work and Education. *J Intell* 2023, Vol 11, Page 54 [Internet]. 2023 [citado 2 de marzo de 2025];11(3):54. Disponible en: <https://www.mdpi.com/2079-3200/11/3/54/htm>
8. De Moraes LL, Kafure I. Bibliometrics and data Science: An example search and analysis of scientific information from the Web of Science (WoS). *Rev Digit Bibliotecon e Cienc da Inf*

[Internet]. 24 de julio de 2020 [citado 2 de marzo de 2025];18:e020016. Disponible en:
<https://www.scielo.br/j/rdbci/a/WkSBdJB9zNjc7zhx7CHqBcJ/?lang=en>