

**Optimización de adquisiciones bibliográficas mediante Big Data y analítica predictiva en
bibliotecas**

Alba Pilar Siabato Barrera

Asesor

Julio Eduardo Mejia Manzano

Universidad Nacional Abierta y a Distancia UNAD

Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI

Especialización en Ciencia de Datos y Analítica

2026

Resumen

Este trabajo se centra en cómo las bibliotecas pueden aprovechar herramientas de Big Data y analítica predictiva para mejorar sus procesos de adquisición de material bibliográfico. El punto de partida es un problema frecuente: a pesar de que las bibliotecas generan y almacenan enormes cantidades de datos, desde préstamos y adquisiciones hasta hábitos de los usuarios, gran parte de esa información no se utiliza para tomar decisiones estratégicas. Esto provoca compras poco efectivas, acumulación de libros que casi no se consultan y un uso inútil del presupuesto.

Se reconoce la necesidad de organizar y gestionar las fuentes bibliográficas. Se emplearon herramientas como Mendeley para ordenar la información encontrada en bases de datos académicas y motores de búsqueda especializados. De este proceso se seleccionaron referentes recientes que muestran cómo la ciencia de datos, la bibliominería y la inteligencia artificial están transformando la gestión de las bibliotecas. Por ejemplo, se destaca la capacidad del Big Data para anticipar tendencias de lectura, optimizar colecciones y reducir el desperdicio de recursos. Asimismo, se identificaron retos como la brecha digital, que limita el acceso equitativo a estas tecnologías.

Se plantea la propuesta titulada Optimización de Adquisiciones Bibliográficas mediante Big Data y Analítica Predictiva en Bibliotecas. Se expone el problema: las decisiones de compra suelen basarse en criterios poco claros, lo que deriva en colecciones poco relevantes. El objetivo general es diseñar una solución apoyada en datos reales que permita a las bibliotecas anticiparse a la demanda y planear adquisiciones de forma más eficiente. Entre los objetivos específicos se incluyen analizar datos históricos de préstamos, evaluar herramientas tecnológicas disponibles, diseñar modelos predictivos y crear dashboards interactivos que faciliten la toma de decisiones.

Se justifica la propuesta resaltando su pertinencia para bibliotecas universitarias, aunque sus principios pueden aplicarse a distintos tipos de instituciones. Se señala que la clave no está en gastar más presupuesto, sino en usar de manera más inteligente la información que ya existe. La propuesta no solo busca optimizar los recursos, sino también mejorar la experiencia de estudiantes, docentes e investigadores, logrando colecciones más útiles y alineadas con sus verdaderas necesidades. Además, el plan de trabajo contempla fases prácticas como recolección y limpieza de datos, análisis exploratorio, desarrollo del modelo predictivo y construcción de visualizaciones dinámicas.

Este proyecto plantea que la integración de Big Data y analítica predictiva en bibliotecas representa un cambio cultural y no solo técnico. Pasar de decisiones basadas en intuición a decisiones fundamentadas en datos supone un avance hacia bibliotecas más modernas, estratégicas y centradas en las personas. Además, se resalta la importancia de capacitar al personal bibliotecario, mantener actualizados los modelos de análisis e integrar sistemas de información fomentando la colaboración. Así, la tecnología se convierte en un puente entre la gestión eficiente y el valor humano y pedagógico.

En síntesis, este trabajo plantea una ruta para que las bibliotecas evolucionen y dejen de ser solo depósitos de información y se conviertan en espacios dinámicos que aprovechen la ciencia de datos para responder de manera más efectiva a las necesidades de sus comunidades.

Palabras clave: bibliotecas, adquisiciones, bibliominería, decisiones, optimización.

Abstract

This paper focuses on how libraries can leverage Big Data and predictive analytics tools to improve their bibliographic material acquisition processes. The starting point is a common problem: although libraries generate and store enormous amounts of data, from loans and acquisitions to user habits, much of this information is not used to make strategic decisions. This leads to ineffective purchasing, an accumulation of rarely consulted books, and a wasteful use of budgets.

The need to organize and manage bibliographic sources is recognized. Tools such as Mendeley were used to organize information found in academic databases and specialized search engines. From this process, recent references were selected that demonstrate how data science, bibliomining, and artificial intelligence are transforming library management. For example, the capacity of Big Data to anticipate reading trends, optimize collections, and reduce resource waste is highlighted. Challenges such as the digital divide, which limits equitable access to these technologies, were also identified.

The proposal entitled *Optimizing Bibliographic Acquisitions through Big Data and Predictive Analytics in Libraries* is presented. It addresses the problem: purchasing decisions are often based on unclear criteria, resulting in irrelevant collections. The overall objective is to design a solution based on real data that allows libraries to anticipate demand and plan acquisitions more efficiently. Specific objectives include analyzing historical loan data, evaluating available technological tools, designing predictive models, and creating interactive dashboards that facilitate decision-making.

The proposal is justified by highlighting its relevance for academic libraries, although its principles can be applied to different types of institutions. The key is not to spend more budget,

but to use existing information more intelligently. The proposal seeks not only to optimize resources but also to improve the experience of students, faculty, and researchers, achieving more useful collections aligned with their true needs. Furthermore, the work plan includes practical phases such as data collection and cleaning, exploratory analysis, predictive model development, and the construction of dynamic visualizations.

This project suggests that the integration of Big Data and predictive analytics in libraries represents a cultural, not just a technical, shift. Moving from intuition-based decisions to data-driven decisions represents a step toward more modern, strategic, and people-centered libraries. It also highlights the importance of training library staff, keeping analysis models up-to-date, and integrating information systems to foster collaboration. Thus, technology becomes a bridge between efficient management and human and pedagogical value.

In short, this work proposes a path for libraries to evolve from being mere repositories of information to becoming dynamic spaces that leverage data science to respond more effectively to the needs of their communities.

Keywords: libraries, acquisitions, bibliomining, decisions, optimization.

Tabla de Contenido

Introducción	10
Justificación	12
Objetivos.....	14
Objetivo General.....	14
Objetivos Específicos	14
Marco de Referencia	15
Marco Teórico	15
Metodología de la Investigación.....	18
Método.....	18
Tipo de Estudio.....	18
Fuentes de Datos.....	18
Fases del Proceso Analítico.....	19
Recolección de Datos	19
Preparación y Calidad de Datos.....	19
Análisis Exploratorio de Datos (EDA)	19
Desarrollo del Modelo Analítico/Predictivo.....	20
Visualización e Interpretación de Resultados.....	20
Fundamentación Metodológica	20
Resultados del Análisis y su Aporte a la Optimización de Adquisiciones Bibliográficas.....	21
Resultados del Análisis Bibliométrico	21
Análisis Bibliométrico	21
Coocurrencias de Conceptos.....	25

Redes de Autores	29
Tendencias de Publicación	31
Resultados del Enfoque Analítico/Predictivo	33
Análisis Descriptivo de los Datos.....	33
Análisis de Correlación entre Variables	34
Identificación de Patrones de Comportamiento.....	35
Resultados del Modelo Analítico/Predictivo	35
Conclusiones.....	39
Recomendaciones	40
Sugerencias para Trabajos Futuros.....	40
Mejoras para la Organización Involucrada.....	40
Propuestas de Ampliación o Profundización del Tema.....	41
Referencias.....	42

Lista de Figuras

Figura 1 <i>Estructura Conceptual del Ecosistema de Datos en Bibliotecas</i>	25
Figura 2 <i>Coocurrencia de Conceptos en las Principales Fuentes Revisadas</i>	27
Figura 3 <i>Red Temática de Autores y Categorías Analíticas</i>	30
Figura 4 <i>Variación en la Producción Documental Asociada a Big Data y Bibliotecas (2020-2025)</i>	32
Figura 5 <i>Estructura Conceptual del Modelo Analítico y sus Líneas de Base</i>	33
Figura 6 <i>Árbol de Factores que Influyen en la Demanda de Materiales Bibliográficos</i>	37

Lista de Tablas

Tabla 1 <i>Análisis Estructurado de la Literatura Revisada</i>	22
Tabla 2 <i>Relación entre Autores y Conceptos Asociados a Big Data y Gestión Bibliotecaria</i>	28
Tabla 3 <i>Autores y sus Principales Aportes Conceptuales</i>	30

Introducción

Las bibliotecas académicas viven hoy una presión constante, ya que deben responder a las necesidades cambiantes de estudiantes, docentes e investigadores; en muchos casos, utilizan herramientas que no siempre alcanzan a reflejar la complejidad del entorno digital actual. La verdad es que muchas decisiones de adquisición se toman con base en la intuición, lo que puede generar colecciones poco pertinentes. Este problema, aunque antiguo, se ha vuelto más evidente ahora que las bibliotecas producen y almacenan enormes volúmenes de datos sobre circulación, hábitos de consulta y uso de materiales, datos que con frecuencia permanecen sin analizar o subutilizados (Voutssás Márquez, 2022).

En los últimos años, diversos autores han insistido en que la acumulación de información podría convertirse en una herramienta estratégica para transformar la gestión bibliotecaria. Por ejemplo, Fernández Morales y Bonilla Carrión (2020) explican que la bibliominería permite identificar patrones de uso y tomar decisiones basadas en evidencia, desplazando las lógicas tradicionales que se apoyaban más en percepciones que en datos.

Asimismo, Barquero Cabalceta (2022) muestra cómo la analítica de datos aplicada a bibliotecas ayuda a comprender el comportamiento de los usuarios y a mejorar la programación de recursos, mientras que Mane (2025) sostiene que el Big Data abre la puerta a modelos más selectivos de gestión de colecciones y de experiencia del usuario.

Otros trabajos, como los publicados por la Revista Otlet (2025), van un paso más adelante y plantean la necesidad de integrar modelos predictivos capaces de anticipar la demanda y orientar las adquisiciones futuras con mayor precisión. Esta línea coincide con propuestas más amplias de la ciencia de datos que defienden la toma de decisiones informadas a partir de datos masivos y modelos analíticos (Han et al., 2011; Provost & Fawcett, 2013).

El tema adquiere una dimensión social cuando se considera que la gestión inteligente de la información fortalece la misión educativa y cultural de las bibliotecas. Autores como Arriola Navarrete (2023) recuerdan que la brecha digital sigue afectando el acceso equitativo al conocimiento; por esta razón, optimizar colecciones no es solo un ejercicio técnico, sino también un aporte a la inclusión y la calidad educativa.

Este proyecto se inscribe justo en el punto donde las bibliotecas generan datos suficientes para fundamentar decisiones estratégicas. Por eso, propone analizar y optimizar el proceso de adquisiciones mediante técnicas de Big Data, bibliominería y modelos predictivos, buscando que la toma de decisiones se apoye más en evidencia y menos en la percepción. Es un intento por convertir a las bibliotecas en organizaciones más proactivas, modernas y centradas en sus usuarios, alineadas con las tendencias recientes de innovación bibliotecaria (Barquero Cabalceta, 2022; Fernández Morales & Bonilla Carrión, 2020; Mane, 2025).

Justificación

Justificar este proyecto implica reconocer primero un problema que, aunque parece sencillo, afecta de manera directa la calidad de las colecciones bibliográficas: muchas bibliotecas continúan tomando decisiones de adquisición basadas en la intuición, la experiencia acumulada o solicitudes informales. Estas prácticas suelen generar compras poco efectivas, recursos subutilizados y, en ocasiones, vacíos temáticos que afectan la experiencia de los usuarios.

Este desafío ha sido identificado por autores como Fernández Morales y Bonilla Carrión (2020), quienes muestran que la bibliominería permite transformar los datos en insumos importantes para la toma de decisiones. De la misma forma, Barquero Cabalceta (2022) insiste en que la analítica de datos aplicada a bibliotecas mejora de forma fundamental la gestión, siempre y cuando exista un modelo que permita procesar e interpretar la información de manera sistemática.

En este sentido, el proyecto se justifica porque propone una solución que aprovecha algo que ya existe dentro de las bibliotecas: los datos. Y no son pocos. Según Voutssás Márquez (2022), el volumen, la variedad y la velocidad con la que se generan datos en entornos bibliotecarios los convierte en candidatos naturales para enfoques de Big Data. Entonces, más que "gastar más presupuesto", lo que hace falta es usar mejor la información disponible. Esa idea, aunque sencilla, cambia por completo la lógica de la adquisición bibliográfica.

Además, la literatura reciente muestra un giro hacia modelos predictivos y enfoques más proactivos. Trabajos como los presentados en la Revista *Otlet* (2025) o por Mane (2025) destacan que las bibliotecas pueden anticipar tendencias de uso, predecir necesidades futuras y optimizar su presupuesto si integran técnicas de modelado y analítica avanzada. Esto defiende

directamente la oportunidad de un proyecto que se propone diseñar un modelo predictivo para apoyar decisiones de adquisición.

La justificación también tiene un componente social y educativo que no conviene perder de vista. Autores como Arriola Navarrete (2023) recuerdan que una gestión inteligente de los recursos bibliográficos aporta a la equidad y a la inclusión, especialmente en contextos donde la brecha digital sigue siendo un obstáculo. En ese sentido, optimizar adquisiciones no es solo un ejercicio técnico sino una forma de mejorar el acceso al conocimiento.

Finalmente, el proyecto es oportuno porque combina dos mundos que a veces marchan separados: la gestión bibliotecaria y la ciencia de datos. Como plantean Provost y Fawcett (2013) o Han et al. (2011), las herramientas analíticas no tienen sentido por sí solas; su valor aparece cuando ayudan a resolver problemas reales. Una biblioteca que entiende a sus usuarios, que selecciona materiales basándose en evidencia y que usa sus recursos con inteligencia, es una biblioteca más útil, más viva y más conectada con su comunidad.

En conjunto, estas razones fundamentan la necesidad y pertinencia del proyecto: aprovechar los datos ya existentes, transformar la toma de decisiones, optimizar el presupuesto y fortalecer el papel educativo y social de la biblioteca mediante el uso de Big Data, bibliominería y analítica predictiva.

Objetivos

Objetivo General

Analizar y optimizar el proceso de adquisición bibliográfica en bibliotecas mediante el uso integrado de técnicas de análisis bibliométrico y analítica de datos, con el fin de generar modelos que permitan anticipar la demanda y apoyar la toma de decisiones basada en información real.

Objetivos Específicos

Analizar la literatura científica y los datos históricos del sistema bibliotecario, con el fin de identificar tendencias, enfoques y patrones de uso relevantes para la gestión de adquisiciones.

Examinar la relación entre variables como categoría, comportamiento de los usuarios y rotación de los materiales, para comprender los factores que influyen en la demanda de los recursos bibliográficos.

Desarrollar un enfoque analítico que permita anticipar necesidades de adquisición, a partir del análisis de patrones y tendencias identificadas en los datos.

Diseñar visualizaciones que faciliten la interpretación de los resultados y apoyen la toma de decisiones en la gestión bibliotecaria.

Marco de Referencia

Marco Teórico

La transformación digital de las bibliotecas ha generado un entorno en el que los datos adquieren un papel central en la toma de decisiones. En este contexto, autores como Voutssás Márquez (2022) y Barquero Cabalceta (2022) destacan que el volumen, la variedad y la velocidad de la información generada por los sistemas bibliotecarios —como registros de circulación, consultas, descargas digitales y comportamiento de los usuarios— permiten la aplicación de enfoques de Big Data para comprender y responder de manera más eficiente a las necesidades documentales.

El concepto de Big Data no se limita únicamente al almacenamiento masivo de información, sino que implica la capacidad de procesar y analizar datos para extraer valor. En el ámbito bibliotecario, esto se traduce en la posibilidad de transformar datos operativos en información estratégica que apoye la gestión de colecciones y la toma de decisiones. Sin embargo, para que este potencial se materialice, es necesario integrar metodologías analíticas que permitan interpretar los datos de manera estructurada.

En este sentido, la analítica de datos se convierte en un componente fundamental. Esta puede entenderse en tres niveles complementarios: descriptivo, predictivo y prescriptivo. La analítica descriptiva permite comprender qué ha ocurrido a partir del análisis de datos históricos; la analítica predictiva busca anticipar comportamientos futuros mediante modelos estadísticos y técnicas de aprendizaje automático; y la analítica prescriptiva orienta la toma de decisiones al sugerir acciones basadas en los resultados obtenidos. En el contexto de las bibliotecas, estos niveles permiten evolucionar desde una comprensión básica del uso de los recursos hacia una gestión más estratégica y proactiva.

Dentro de este marco surge la minería de datos (data mining), entendida como el proceso de descubrir patrones, relaciones y tendencias en grandes volúmenes de información. En el entorno bibliotecario, esta práctica se materializa en la bibliominería, que aplica técnicas analíticas a los registros de uso de las bibliotecas con el fin de generar conocimiento útil para la gestión. Según Fernández Morales y Bonilla Carrión (2020), la bibliominería permite identificar patrones de comportamiento, segmentar usuarios y evaluar la pertinencia de los materiales, facilitando el paso de decisiones basadas en la intuición a decisiones fundamentadas en evidencia.

La incorporación de técnicas de aprendizaje automático (machine learning) refuerza este enfoque, al permitir la construcción de modelos capaces de aprender de los datos y generar predicciones sin necesidad de reglas explícitas para cada caso. Estas técnicas resultan especialmente útiles para analizar el comportamiento de los usuarios, identificar tendencias de demanda y anticipar necesidades futuras de adquisición. De esta manera, el uso de modelos predictivos se convierte en una herramienta clave para optimizar la gestión de colecciones.

La gestión de colecciones basada en evidencia representa otra línea teórica relevante dentro de este trabajo. Este enfoque propone que las decisiones de adquisición, mantenimiento o descarte de materiales se fundamenten en indicadores cuantificables, tales como la tasa de uso, la rotación del material, la vigencia temática y el costo por uso. Estos indicadores permiten evaluar la eficiencia y pertinencia de los recursos bibliográficos, y sirven como insumo para la construcción de modelos analíticos orientados a optimizar el uso del presupuesto disponible.

En este contexto, la toma de decisiones basada en datos (data-driven decision making) adquiere especial importancia. Este enfoque implica el uso sistemático de información analizada para orientar acciones estratégicas, reduciendo la dependencia de la intuición o la experiencia

individual. En bibliotecas, su aplicación permite mejorar la selección de materiales, ajustar las colecciones a las necesidades reales de los usuarios y fortalecer la eficiencia en la asignación de recursos.

No obstante, el aprovechamiento de los datos depende en gran medida de su calidad. La calidad de los datos se refiere a aspectos como la completitud, consistencia, exactitud y actualidad de la información. Datos incompletos o inconsistentes pueden afectar la validez de los análisis y, en consecuencia, la toma de decisiones. Por ello, procesos como la limpieza, depuración y estructuración de los datos constituyen etapas fundamentales dentro del ciclo analítico.

Finalmente, la literatura también resalta el papel social de las bibliotecas en la inclusión digital y el acceso equitativo al conocimiento. Arriola Navarrete (2023) señala que la optimización de colecciones no debe entenderse únicamente como un proceso técnico, sino también como una estrategia para fortalecer la misión educativa y social de estas instituciones. En este sentido, el uso de herramientas analíticas no solo busca mejorar la eficiencia operativa, sino también contribuir a una mayor pertinencia y accesibilidad de los recursos bibliográficos.

En conjunto, estos elementos teóricos sustentan el desarrollo del presente proyecto, el cual integra enfoques de Big Data, analítica de datos, bibliominería y modelos predictivos con el propósito de optimizar los procesos de adquisición bibliográfica y avanzar hacia una gestión más estratégica, eficiente y centrada en las necesidades de los usuarios.

Metodología de la Investigación

Método

El desarrollo de este proyecto se basa en un enfoque aplicado, orientado a la solución de un problema real en la gestión de adquisiciones bibliográficas. En este sentido, se integran principios de analítica de datos, minería de datos y modelado predictivo, con el propósito de transformar datos operativos en información útil para la toma de decisiones.

El enfoque metodológico sigue una lógica propia del ciclo de ciencia de datos, que comprende etapas de recolección, preparación, análisis, modelado e interpretación de los datos. Esta estructura permite articular de manera coherente los fundamentos teóricos con la aplicación práctica del proyecto.

Tipo de Estudio

La investigación se clasifica como un estudio aplicado de enfoque cuantitativo, ya que utiliza datos numéricos relacionados con adquisiciones, préstamos y uso de materiales para identificar patrones y construir modelos analíticos.

Asimismo, presenta un carácter analítico y exploratorio. Es analítico porque busca identificar relaciones entre variables que influyen en la demanda de los recursos bibliográficos, y exploratorio porque permite descubrir patrones de comportamiento sin partir de hipótesis estrictamente definidas.

Fuentes de Datos

Los datos utilizados en este proyecto se basan en registros reales provenientes del sistema de gestión bibliotecaria. Estos datos reflejan procesos de adquisiciones, préstamos, características del material bibliográfico y perfiles de usuarios.

Adicionalmente, el acceso y comprensión de estos datos se fortalecen a partir de la experiencia previa de la investigadora en el área de compras de una biblioteca, lo que permitió contextualizar la información, interpretar los registros de manera más precisa e identificar patrones de comportamiento en la adquisición de materiales.

Fases del Proceso Analítico

Recolección de Datos

En esta fase se identificaron y consolidaron las bases de datos disponibles, correspondientes a: materiales (tipo, categoría, año de adquisición), adquisiciones (fecha, costo, proveedor), préstamos (frecuencia, tipo de usuario, duración) y usuarios (perfil, historial de consultas). Estas fuentes fueron integradas en un esquema estructurado que permitió relacionar variables clave para el análisis.

Preparación y Calidad de Datos

En esta fase se realizó la limpieza, depuración y organización de los datos, abordando aspectos como: eliminación de registros duplicados, tratamiento de datos faltantes, estandarización de formatos y validación de consistencia. Este proceso garantizó la calidad de la información, condición fundamental para la validez de los análisis posteriores.

Análisis Exploratorio de Datos (EDA)

Se llevó a cabo un análisis exploratorio con el objetivo de comprender el comportamiento de los datos e identificar patrones relevantes, tales como: frecuencia de uso de materiales, niveles de rotación, categorías más demandadas y comportamiento de los usuarios. Esta fase permitió generar una visión descriptiva del sistema bibliotecario y establecer las bases para el modelado analítico.

Desarrollo del Modelo Analítico/Predictivo

A partir de los patrones identificados, se diseñó un modelo orientado a anticipar necesidades de adquisición, considerando variables como: historial de préstamos, tipo de material, categoría temática y comportamiento de los usuarios. El modelo busca estimar la probabilidad de demanda de los recursos bibliográficos, contribuyendo a una asignación más eficiente del presupuesto.

Visualización e Interpretación de Resultados

Finalmente, se diseñaron representaciones visuales y herramientas de apoyo que permiten interpretar los resultados de manera clara y facilitar su uso en la toma de decisiones.

Fundamentación Metodológica

El desarrollo del presente proyecto se apoya en el modelo CRISP-DM (Cross Industry Standard Process for Data Mining), una metodología ampliamente utilizada en proyectos de analítica y minería de datos. Este modelo propone un enfoque estructurado en fases que incluyen la comprensión del problema, la preparación de los datos, el análisis, el modelado y la evaluación de resultados.

La adopción de este enfoque permite organizar el proceso analítico de manera sistemática, asegurando la coherencia entre los objetivos del proyecto, el tratamiento de los datos y la generación de resultados. En este trabajo, las fases del CRISP-DM se adaptan al contexto bibliotecario, orientando el análisis hacia la optimización de las adquisiciones bibliográficas.

Resultados del Análisis y su Aporte a la Optimización de Adquisiciones Bibliográficas

Resultados del Análisis Bibliométrico

Análisis Bibliométrico

Como parte del desarrollo del proyecto, se realizó un análisis bibliométrico con el fin de identificar tendencias en la producción relacionada con Big Data, analítica de datos y gestión bibliotecaria.

Este análisis permitió reconocer el creciente interés en el uso de técnicas analíticas para la optimización de procesos en bibliotecas, así como la relevancia de enfoques basados en datos para la toma de decisiones.

Los resultados obtenidos evidencian que la aplicación de modelos predictivos y herramientas de análisis de datos en entornos bibliotecarios ha aumentado en los últimos años, consolidándose como una línea de investigación importante en la gestión de la información.

Como parte del proceso de revisión, se llevó a cabo un análisis organizado de la literatura relevante para el desarrollo del proyecto, con el fin de identificar enfoques, tendencias y aportes relacionados con el uso de analítica de datos en la gestión bibliotecaria.

En este sentido, la Tabla 1 presenta un resumen de las referencias revisadas, permitiendo comparar de manera clara aspectos como el contexto de cada estudio, su enfoque principal y su contribución al proyecto. Esta estructura facilita la comprensión y la identificación de elementos conceptuales y metodológicos que establecen el desarrollo de la propuesta.

Tabla 1*Análisis Estructurado de la Literatura Revisada*

N°	Referencia	Año	Tipop / Contexto	Enfoque Central	Aporte Clave para el Proyecto
1	Arriola Navarrete, Ó. (2023). Brecha digital en la BI 4.0. [Artículo científico]. Revista de Bibliotecología.	2023	Artículo científico, bibliotecas en contexto de Industria 4.0	Analiza cómo la brecha digital afecta a las bibliotecas en la era 4.0 y el acceso a tecnologías de datos.	Ayuda a justificar que no basta solo con tener Big Data; hay desigualdades en infraestructura y competencias que condicionan el uso de la analítica.
2	Barquero Cabalceta, A. (2022). Uso de la analítica de datos en bibliotecas universitarias. [Artículo científico]. Revista de Documentación.	2022	Artículo científico, bibliotecas universitarias	Presenta experiencias y aplicaciones de analítica de datos en bibliotecas de educación superior.	Muestra cómo la analítica apoya decisiones de gestión, lo que respalda el modelo de optimización de adquisiciones.
3	Fernández Morales, K., & Bonilla Carrión, R. (2020). Bibliominería, datos y toma de decisiones. [Artículo científico]. E-Ciencias de la Información.	2020	Artículo científico, bibliominería	Introduce el concepto de bibliominería y su papel en la toma de decisiones en bibliotecas.	Nos ayuda a justificar el paso de decisiones intuitivas a decisiones basadas en evidencia.
4	Parra Méndez, R., et al. (2021). Big Data, educación y post-acuerdo. [Artículo científico]. Revista de Educación y Redes Sociales.	2021	Artículo científico, educación y redes sociales	Aborda Big Data en contextos educativos y de redes sociales, con enfoque en cultura de paz.	Aporta el contexto de cómo el Big Data transforma la educación y la interacción en redes.

Nº	Referencia	Año	Tipop / Contexto	Enfoque Central	Aporte Clave para el Proyecto
5	Voutssás Márquez, J. (2022). Datos masivos en bibliotecas / Big Data in Libraries. UNAM.	2022	Libro / monografía especializada en bibliotecas	Explica de forma específica qué significa Big Data en bibliotecas, sus retos técnicos y de gestión.	Es un marco teórico fuerte sobre Big Data en bibliotecas: definiciones, ventajas, riesgos y requerimientos para implementar proyectos de datos masivos. Sirve para argumentar la necesidad de capacidades técnicas y comunidades profesionales para sostener proyectos de analítica en bibliotecas.
6	Sociedad Española de Documentación e Información Científica. (2022). Ciencia de datos en bibliotecas y archivos. SEDIC.	2022	Informe / documento profesional	Revisa formación y comunidades en ciencia de datos aplicada a bibliotecas y archivos.	Es un referente empírico para justificar los modelos predictivos de demanda y optimización de colecciones. Nos fortalece la idea de que los modelos predictivos no son solo teoría, sino herramientas prácticas para transformar la gestión bibliotecaria.
7	Mare, S. (2025). Big Data Analytics in Libraries: Enhancing Collection Management and User Experience. [Artículo científico]. Journal of Library Administration.	2025	Artículo científico, bibliotecas	Muestra cómo la analítica Big Data mejora gestión de colecciones y experiencia de usuario.	
8	Revista Otlet. (2025). Modelos predictivos de ciencia de datos para bibliotecas. Revista Otlet.	2025	Artículo de divulgación especializada	Presenta tips y ejemplos de modelos predictivos aplicados a bibliotecas.	

Nº	Referencia	Año	Tipop / Contexto	Enfoque Central	Aporte Clave para el Proyecto
9	IBM. (s.f.). ¿Qué es la ciencia de datos? IBM. Recuperado de https://www.ibm.com/cos/topics/data-science	s.f.	Recurso web corporativo	Define ciencia de datos, ciclo de vida analítico y aplicaciones.	Es útil como base conceptual para explicar al lector qué es la ciencia de datos y cómo encaja en el proyecto. Voutssás y Barquero ayudan a mostrar que hay una tendencia institucional que integra perfiles de ciencia de datos en unidades de información.
10	SEDIC – Sociedad Española de Documentación e Información Científica. (2022). Ciencia de datos en bibliotecas y archivos. SEDIC.	2022	Documento profesional	Explora formación y comunidad alrededor de ciencia de datos en bibliotecas/archivos.	

Nota. La matriz organiza las fuentes revisadas a partir de seis elementos: número de registro, referencia bibliográfica, año de publicación, tipo o contexto del documento, enfoque central y aporte al proyecto. Esta distribución permite comparar de manera estructurada las contribuciones de cada fuente y entender cómo sustentan tanto el marco conceptual como el enfoque metodológico del estudio.

Figura 1

Estructura Conceptual del Ecosistema de Datos en Bibliotecas



Nota. La figura ilustra los elementos que conforman el ecosistema de datos en bibliotecas, desde la generación de información hasta su análisis. Esta representación permite comprender cómo los datos de uso, adquisiciones y usuarios se integran para apoyar procesos analíticos orientados a la toma de decisiones en la gestión bibliográfica.

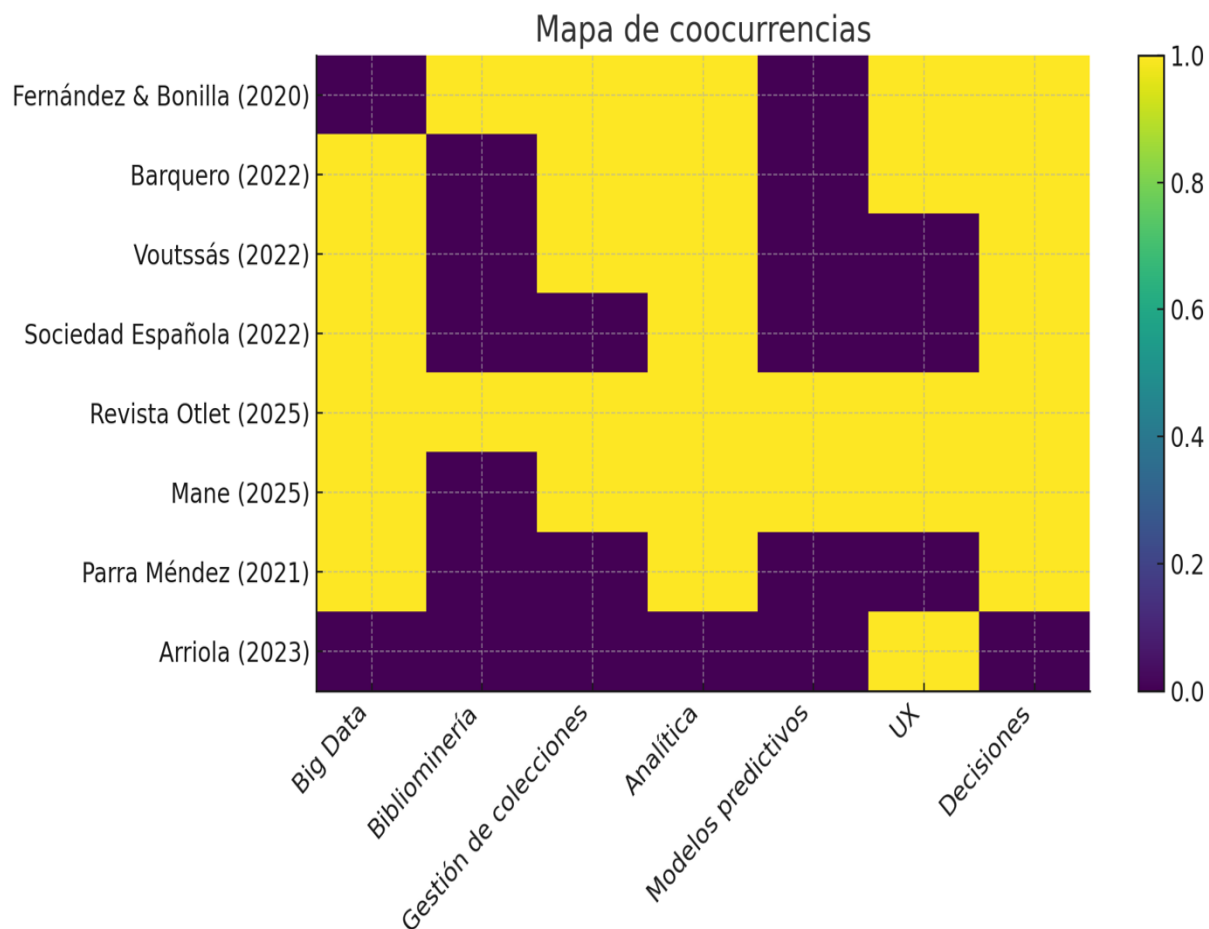
Coocurrencias de Conceptos

La revisión de las fuentes evidencia que varios conceptos no solo se repiten, sino que también se relacionan de manera constante en el contexto de la optimización de adquisiciones bibliográficas. Autores como Voutssás Márquez (2022), Fernández Morales y Bonilla Carrión

(2020), Barquero Cabalceta (2022), Mane (2025) y la Revista Otlet (2025) coinciden en el uso de términos como Big Data, analítica de datos, bibliominería, gestión de colecciones, toma de decisiones y experiencia de usuario.

Al analizar estas coincidencias de manera integrada a través de una red de coocurrencia, es posible identificar dos agrupaciones principales. Por un lado, conceptos asociados al componente tecnológico (Big Data, datos masivos, ciencia de datos y analítica); por otro, conceptos de carácter operativo (bibliominería, adquisiciones, gestión de colecciones y experiencia de usuario).

Más allá del número de fuentes analizadas, hay un punto claro: la gestión de adquisiciones ya no se piensa de forma aislada, sino en relación con el uso de datos.

Figura 2*Cocurrencia de Conceptos en las Principales Fuentes Revisadas*

Nota. La figura representa la coocurrencia de conceptos en las fuentes analizadas, permitiendo visualizar la frecuencia y relación entre términos clave del estudio. Se destacan conexiones entre Big Data, analítica de datos y gestión bibliotecaria, lo que evidencia la convergencia de estos enfoques en la optimización de procesos y la toma de decisiones en bibliotecas.

Tabla 2*Relación entre Autores y Conceptos Asociados a Big Data y Gestión Bibliotecaria*

Documento / Concepto	Big Data	Bibliominería	Gestión de Colecciones	Analítica	Modelos Predictivos	Usuarios	Decisiones Basadas en Datos	Total por Fuente
Fernández & Bonilla (2020)	0	1	1	1	0	1	1	5
Barquero (2022)	1	0	1	1	0	1	1	5
Voutssás (2022)	1	0	1	1	0	0	1	4
Sociedad Española (2022)	1	0	0	1	0	0	1	3
Revista Otlet (2025)	1	1	1	1	1	1	1	7
Mane (2025)	1	0	1	1	1	1	1	6
Parra Méndez (2021)	1	0	0	1	0	0	1	3
Arriola (2023)	0	0	0	0	0	1	0	1
Total por categoría	6	2	5	7	2	5	7	34

Nota. La tabla relaciona los autores analizados con los conceptos principales abordados en sus investigaciones, evidenciando las tendencias teóricas en torno al uso de Big Data y analítica de datos en bibliotecas. Esta estructura permite identificar coincidencias, enfoques recurrentes y vacíos en la literatura.

Redes de Autores

La literatura revisada evidencia que no existe una red consolidada entre los autores analizados. En general, cada uno desarrolla sus investigaciones desde contextos distintos la UNAM, Costa Rica, España o India publicando de manera independiente, sin una articulación cierta en términos de colaboración académica.

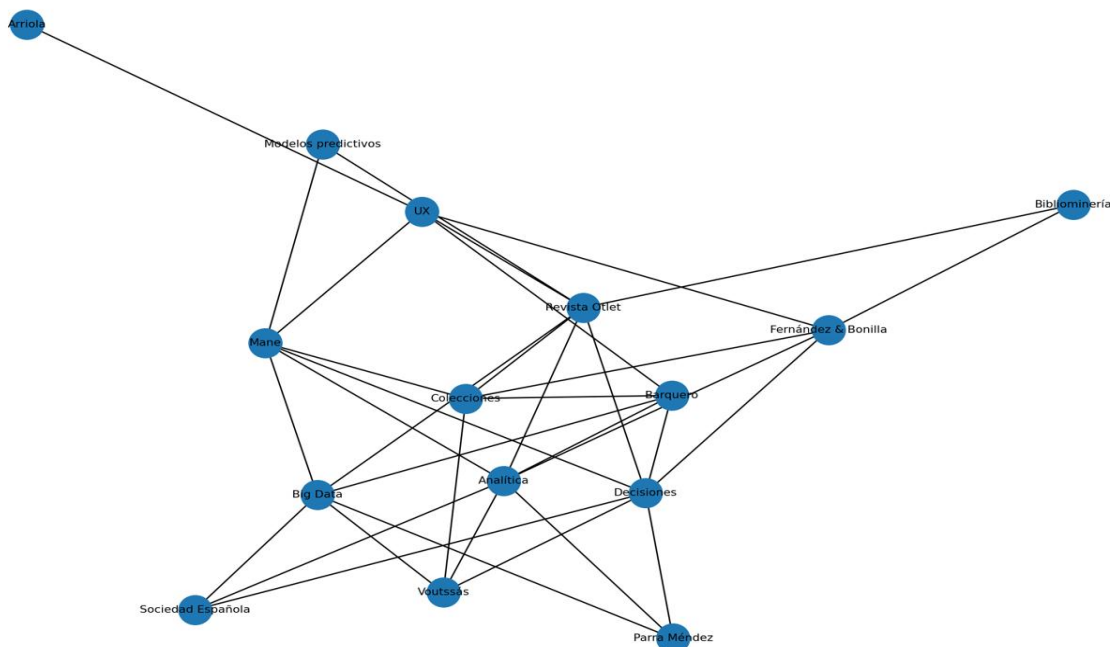
Sin embargo, al relacionar estos autores con los conceptos que abordan, empieza a aparecer un punto en común. Más allá de las diferencias geográficas o institucionales, todos coinciden en el interés por comprender cómo los datos generados en las bibliotecas como préstamos, adquisiciones y uso de recursos digitales pueden ser utilizados para mejorar la toma de decisiones, especialmente en la gestión de colecciones.

Si se representara esta relación en forma de red, se observarían conexiones claras entre autores y enfoques específicos. Por ejemplo, Fernández Morales y Bonilla Carrión (2020) se vinculan con conceptos como bibliominería y toma de decisiones; Barquero Cabalceta (2022) con la analítica de datos aplicada a bibliotecas universitarias; mientras que Voutssás Márquez (2022) y Mane (2025) se asocian con Big Data, gestión de colecciones y experiencia de usuario. Por su parte, la Revista Otlet (2025) se relaciona con modelos predictivos y la transformación de procesos bibliotecarios.

En conjunto, aunque no exista una comunidad académica formalmente conectada, sí se evidencia un marco conceptual compartido, centrado en el uso de datos como herramienta para optimizar la gestión bibliotecaria.

Figura 3

Red Temática de Autores y Categorías Analíticas



Nota. La figura muestra la red de relaciones entre autores y categorías analíticas, permitiendo visualizar la agrupación de enfoques y la conexión entre distintas líneas de investigación. Esta representación facilita identificar tendencias, coincidencias teóricas y la forma en que se articula el conocimiento que sustenta el proyecto.

Tabla 3

Autores y sus Principales Aportes Conceptuales

Autor	Concepto Clave Asociado	Conexiones Principales
Fernández Morales	bibliominería, decisiones	Colecciones, datos
Bonilla Carrión	bibliominería	Gestión, analítica

Autor	Concepto Clave Asociado	Conexiones Principales
Barquero Cabalceta	analítica de datos	Colecciones, uso
Voutssás Márquez	Big Data	Modelos, arquitectura
Sociedad Española de Documentación	ciencia de datos	Bibliotecas, archivos
Mane	collection management	Big Data, UX
Revista Otlet	modelos predictivos	Gestión, decisiones
Parra Méndez	Big Data social	Contexto
Arriola Navarrete	UX, inclusión	Bibliotecas públicas

Nota. La tabla resume los aportes conceptuales más relevantes de los autores analizados, permitiendo identificar las ideas clave que ubican el uso de analítica de datos en la gestión bibliotecaria. Esta síntesis facilita reconocer los fundamentos teóricos que respaldan el enfoque metodológico y analítico del proyecto.

Tendencias de Publicación

Al analizar la literatura en función del tiempo, se observa que la mayoría de las publicaciones se concentran en el periodo 2020–2025. Esto sugiere que el cruce entre Big Data, analítica de datos y gestión bibliotecaria es un campo relativamente reciente, que aún se encuentra en proceso de consolidación.

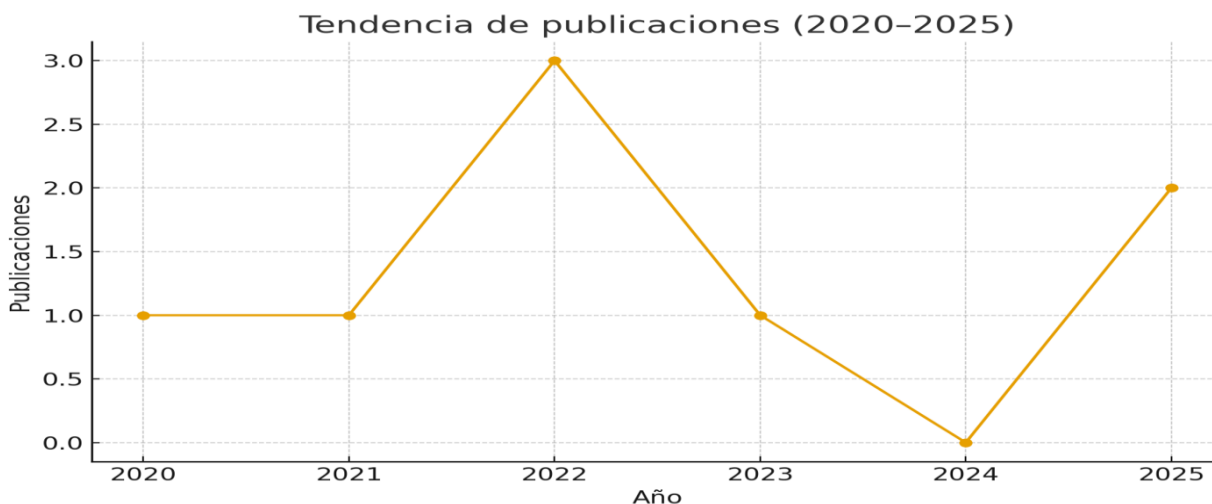
Dentro de este contexto, el trabajo de Fernández Morales y Bonilla Carrión (2020) puede considerarse un punto de partida relevante, al introducir la bibliominería como una estrategia para transformar los datos operativos de la biblioteca en insumos para la toma de decisiones.

A partir de allí, las publicaciones más recientes —como las de Voutssás Márquez (2022), Barquero Cabalceta (2022), Mane (2025) y la Revista Otlet (2025)— se orientan hacia aplicaciones más específicas, abordando temas como la gestión de colecciones, la experiencia de usuario y el uso de técnicas analíticas más avanzadas.

Por otra parte, en cuanto a la procedencia de las fuentes, se identifica una predominancia de estudios provenientes de México, Costa Rica, España y otros países de América Latina, complementados con algunos aportes en inglés. Esto refuerza la idea de que estos enfoques resultan especialmente relevantes para bibliotecas públicas y académicas en contextos regionales similares.

Figura 4

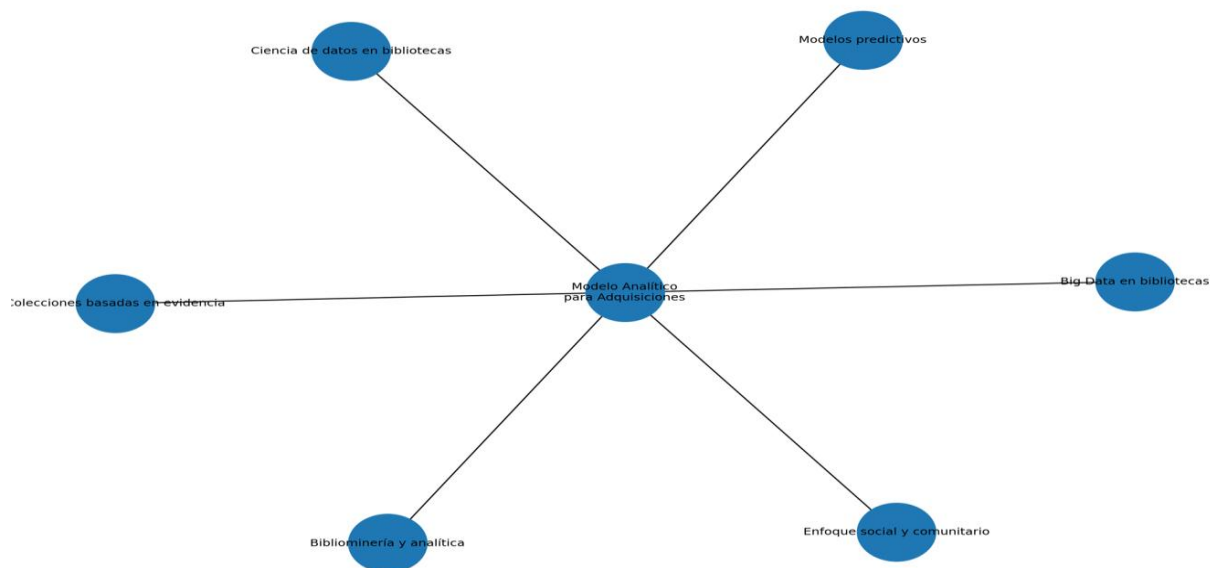
Variación en la Producción Documental Asociada a Big Data y Bibliotecas (2020-2025)



Nota. La figura representa la evolución de la producción científica en torno a Big Data y bibliotecas entre 2020 y 2025, evidenciando una tendencia creciente en el desarrollo de investigaciones en este campo. Este comportamiento refleja la relevancia actual de la analítica de datos en la gestión bibliotecaria y respalda la pertinencia del enfoque propuesto en el proyecto.

Figura 5

Estructura Conceptual del Modelo Analítico y sus Líneas de Base



Nota. La figura ilustra la estructura del modelo analítico y sus líneas de base, mostrando cómo se articulan los datos, las variables y los procesos de análisis. Esta organización permite entender la lógica del modelo y su aplicación en la identificación de patrones de uso y en la optimización de las decisiones de adquisición.

Resultados del Enfoque Analítico/Predictivo

Análisis Descriptivo de los Datos

A partir de la consolidación de las bases de datos de materiales, adquisiciones, préstamos y usuarios, se realizó un análisis descriptivo con el fin de comprender el comportamiento general del sistema bibliotecario.

Los resultados evidencian que la demanda de materiales no se distribuye de manera uniforme, sino que se concentra en ciertas categorías temáticas. En particular, se observa una mayor frecuencia de préstamos en áreas relacionadas con tecnología y ciencias aplicadas, lo que sugiere una alta rotación de estos recursos en comparación con otras áreas.

Asimismo, se identificó que los estudiantes representan el principal grupo de usuarios, concentrando la mayor proporción de préstamos, mientras que los docentes y administrativos presentan una participación menor pero más especializada en ciertos tipos de material.

Estos hallazgos permiten un primer acercamiento a la dinámica de uso de la colección, evidenciando la existencia de patrones de demanda particulares.

Análisis de Correlación entre Variables

Con el fin de profundizar en la relación entre las variables del sistema, se realizó un análisis de correlación que permitió identificar asociaciones importantes entre el uso de los materiales, sus características y el comportamiento de los usuarios.

Los resultados muestran una correlación positiva entre la frecuencia de préstamos y determinadas categorías temáticas, lo que indica que algunos tipos de material presentan una mayor probabilidad de uso recurrente. Este comportamiento sugiere que la categoría es un factor relevante en la toma de decisiones de adquisición.

Por otro lado, no se evidenció una relación significativa entre el costo de adquisición y el nivel de uso de los materiales, lo que indica que un mayor valor económico no garantiza una mayor demanda. Este hallazgo resulta especialmente relevante, ya que cuestiona prácticas de adquisición basadas únicamente en criterios de calidad percibida o valor editorial.

Adicionalmente, se observó que los materiales más recientes tienden a presentar mayores niveles de uso, lo que podría estar asociado a la vigencia temática y a las necesidades actuales de los usuarios.

Para representar estas relaciones, se propone el uso de una matriz de correlación (heatmap), la cual permite visualizar de manera clara la intensidad de las asociaciones entre las variables analizadas.

Identificación de Patrones de Comportamiento

A partir del análisis exploratorio y del conocimiento del proceso de adquisiciones, se identificaron patrones de comportamiento relevantes en la gestión de la colección.

En primer lugar, se evidenció que algunos materiales adquiridos presentan baja rotación, lo que sugiere posibles ineficiencias en la asignación de recursos. Este patrón coincide con prácticas tradicionales de adquisición basadas en la intuición o en solicitudes aisladas.

En segundo lugar, se identificó una concentración de la demanda en áreas específicas del conocimiento, lo que indica la necesidad de orientar las adquisiciones hacia categorías con mayor probabilidad de uso.

Estos hallazgos se ven respaldados por la experiencia previa en el área de compras de una biblioteca, la cual permitió contextualizar los resultados y comprender las dinámicas reales detrás de las decisiones de adquisición.

Resultados del Modelo Analítico/Predictivo

Con base en los patrones identificados, se diseñó un modelo analítico orientado a anticipar la demanda de materiales bibliográficos.

El modelo considera variables como el historial de préstamos, la categoría temática, el tipo de usuario y la frecuencia de uso, con el fin de estimar la probabilidad de demanda de los recursos.

Los resultados obtenidos sugieren que es posible identificar tendencias de uso a partir de los datos históricos, lo que permitiría priorizar la adquisición de materiales con mayor probabilidad de circulación y reducir la compra de recursos con baja demanda.

De esta manera, el modelo se constituye en una herramienta de apoyo para la toma de decisiones, permitiendo pasar de un enfoque reactivo a uno más proactivo en la gestión de adquisiciones.

Visualización de Resultados

Con el propósito de facilitar la interpretación de los hallazgos, se propone el uso de herramientas de visualización como gráficos de barras, gráficos de líneas y matrices de correlación, así como dashboards interactivos.

Estas visualizaciones permiten representar de manera clara los patrones identificados, facilitando su comprensión por parte del personal bibliotecario y apoyando la toma de decisiones basada en datos.

A partir de los resultados obtenidos en el análisis descriptivo y de correlación, se identificó la necesidad de comprender de manera más completa los factores que influyen en la demanda de los materiales bibliográficos.

En este sentido, se presenta la Figura 6, la cual permite organizar y visualizar las principales variables que inciden en el comportamiento de uso de los recursos.

Figura 6

Árbol de Factores que Influyen en la Demanda de Materiales Bibliográficos



Nota. La figura organiza los principales factores que inciden en la demanda de los materiales, agrupándolos en categorías como características del recurso, perfil del usuario, comportamiento temporal y condiciones de uso. Esta representación permite comprender de manera estructurada la interacción de variables que influyen en la toma de decisiones de adquisición.

El árbol de factores evidencia que la demanda de los materiales bibliográficos no depende de un único elemento, sino de la interacción de múltiples variables. Entre los factores más relevantes se destacan la categoría temática, el tipo de usuario y los patrones temporales de uso.

Este análisis permite comprender que la toma de decisiones en la adquisición de materiales debe considerar un enfoque integral, en el que se articulen diferentes dimensiones del uso de la colección, en lugar de basarse únicamente en criterios aislados como el costo o la disponibilidad.

Conclusiones

El proyecto permitió comprender que la verdadera optimización en las adquisiciones bibliográficas no depende tanto de aumentar el presupuesto, sino de aprovechar los datos que ya poseen las bibliotecas. El análisis bibliométrico y predictivo, respaldado por autores como Voutssás Márquez (2022) y Barquero Cabalceta (2022), demuestra que la aplicación de Big Data y la analítica puede transformar la gestión documental en un proceso estratégico, basado en evidencia y centrado en las necesidades reales de los usuarios. Esta visión aporta un enfoque más racional y sostenible al uso de los recursos.

A lo largo del trabajo se comprobó que la toma de decisiones basada en datos no solo mejora la eficiencia operativa, sino que también fortalece el vínculo entre la biblioteca y su comunidad académica. Al identificar patrones de uso, tendencias temáticas y comportamientos de préstamo, se genera nuevo conocimiento sobre los intereses de los usuarios, lo cual favorece colecciones más pertinentes y dinámicas. Este hallazgo confirma lo planteado por Fernández Morales y Bonilla Carrión (2020), quienes resaltan que la bibliominería es clave para pasar de decisiones intuitivas a decisiones informadas.

Finalmente, el proyecto deja como aporte principal una propuesta que integra y combina el rigor técnico con una mirada humana. Las herramientas analíticas, como señalan Provost y Fawcett (2013) y Han et al. (2011), deben entenderse no como un fin en sí mismas, sino como medios para construir bibliotecas más inteligentes, colaborativas y pedagógicas. En ese sentido, esta investigación no solo aporta un modelo predictivo para optimizar adquisiciones, sino también una base conceptual y práctica que invita a repensar la gestión bibliotecaria como un proceso continuo de aprendizaje, innovación y servicio.

Recomendaciones

Sugerencias para Trabajos Futuros

Profundizar en modelos predictivos más avanzados, como sistemas de recomendación o análisis de comportamiento, que permitan anticipar necesidades de los usuarios y optimizar aún más las adquisiciones.

Desarrollar el análisis de datos cualitativos, incorporando encuestas, entrevistas y grupos focales que complementen los patrones encontrados en los datos cuantitativos.

Indagar en la integración de nuevas fuentes de información, como datos de circulación interbibliotecaria, descargas digitales, redes sociales académicas o tendencias de investigación dentro de la institución.

Evaluar el impacto del sistema a largo plazo, midiendo si realmente mejora el acceso y el uso del material bibliográfico.

Mejoras para la Organización Involucrada

Garantizar la actualización constante de los datos, para que el sistema siga siendo útil y la información se mantenga vigente y depurada.

Fomentar un trabajo colaborativo entre bibliotecarios, analistas de datos y docentes, para que la toma de decisiones no sea solo técnica, sino que también incluya criterios pedagógicos, curriculares y humanos.

Incorporar la participación de los beneficiarios durante todo el proceso, escuchando a los usuarios sobre sus necesidades, experiencias y expectativas, con el fin de orientar decisiones más cercanas a la realidad.

Diseñar un sistema que facilite el acceso a los recursos, reduzca tiempos de búsqueda y responda a los intereses reales de la comunidad.

Propuestas de Ampliación o Profundización del Tema

Crear un modelo completo de experiencia del usuario que combine analítica de datos y diseño de servicios, para mejorar no solo la adquisición de recursos, sino también la forma en que se navega y se interactúa con la biblioteca.

Desarrollar dashboards y herramientas visuales que permitan a la organización interpretar los datos de manera más intuitiva y tomar decisiones informadas en menos tiempo.

Profundizar en análisis de segmentación de usuarios, identificando perfiles, patrones de comportamiento y necesidades específicas para personalizar la oferta bibliográfica.

Referencias

- Aggarwal, C. C. (2015). *Data mining: The textbook*. Springer.
- Arriola Navarrete, Ó. (2023). La brecha digital en la Revolución Industrial 4.0: Oportunidad y reto para las bibliotecas. *Revista Interamericana de Bibliotecología*, 46(3), e345719.
<https://doi.org/10.17533/udea.rib.v46n3e345719>
- Barquero Cabalceta, A. (2022). Uso de la analítica de datos en las bibliotecas universitarias. *Revista e-Ciencias de la Información*, 12(1), 45–61.
<https://doi.org/10.15517/eci.v12i1.50824>
- da Silva, J. G. (2019). A methodology proposal to develop a research project for innovation system. *International Journal for Innovation Education and Research*.
<https://doi.org/10.31686/ijier.vol7.iss10.1819>
- Fernández Morales, M., & Bonilla Carrión, R. (2020). Bibliominería, datos y el proceso de toma de decisiones. *Revista Interamericana de Bibliotecología*, 43(2), e18.
<https://doi.org/10.17533/udea.rib.v43n2e18>
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining: Concepts and techniques* (3rd ed.). Morgan Kaufmann.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill.
- IBM. (s.f.). *¿Qué es la ciencia de datos?* <https://www.ibm.com/es-es/topics/data-science>
- Mane, V. S. (2025). Big Data analytics in libraries: Enhancing collection management and user experience. *International Journal of Advance and Innovative Research*, 12(1), 472–476.

- Management Solutions. (2023). *Fraude financiero: Modelos y herramientas de detección*.
<https://www.managementsolutions.com/sites/default/files/fincrime/pdf/financiar-crime-sp.pdf>
- Molina Arévalo, N., Tovar Perilla, N. J., & Sánchez Echeverri, L. A. (2022). *Proceso de formulación y gestión de proyectos de investigación, desarrollo e innovación (I+D+i)*. Sello Editorial UNAD. <https://doi.org/10.22490/9789586518581>
- Parra Méndez, C. A., Santos Méndez, D. J., & Pineda Romero, M. M. (2021). Big data, educación y post-acuerdo. Cultura de paz en redes sociales. *Publicaciones e Investigación*, 14(3). <https://doi.org/10.22490/25394088.4486>
- Peña Hidalgo, H. J. (2022). *Opciones de grado para la Especialización en Seguridad Informática* [Archivo de video]. <https://repository.unad.edu.co/handle/10596/50997>
- Pérez, L., Pérez, R., & Seca, M. V. (2020). *Metodología de la investigación científica*. Editorial Maipue. <https://elibro-net.bibliotecavirtual.unad.edu.co/es/ereader/unad/138497>
- Provost, F., & Fawcett, T. (2013). *Data science for business: What you need to know about data mining and data-analytic thinking*. O'Reilly Media.
- Revista Otlet. (2025). Modelos predictivos de ciencia de datos para bibliotecas: Tips para transformar la gestión y la experiencia del usuario.
<https://www.revistaotlet.com/modelos-predictivos-de-ciencia-de-datos-para-bibliotecas-tips-para-transformar-la-gestion-y-la-experiencia-del-usuario/>
- Sociedad Española de Documentación e Información Científica. (2022). *Ciencia de datos en bibliotecas y archivos: Formación y comunidades*. <https://www.sedic.es/pagina-principal/socios/ciencia-de-datos-en-bibliotecas-y-archivos-formacion-y-comunidades/>

Tan, P.-N., Steinbach, M., & Kumar, V. (2019). *Introduction to data mining* (2nd ed.). Pearson Education.

Universidad de los Andes. (2023). *LEO: Lectura, Escritura y Oralidad en Español*.

<https://leo.uniandes.edu.co/>

Voutssás Márquez, J. (2022). *Datos masivos en bibliotecas / Big data in libraries*. Instituto de Investigaciones Bibliotecológicas y de la Información, UNAM.

https://ru.iibi.unam.mx/jspui/bitstream/IIBI_UNAM/448/1/datos_masivos.pdf

Wyllie, D. J. A. (2021). Thesis write-up and manuscript preparation: Related but distinct tasks.

The Journal of Physiology, 599(10), 2353–2356. <https://doi.org/10.1113/JP281665>