

10, 11 e 12 de novembro de 2025

POLITÉCNICO DO PORTO / ISCAP
PORTO - PORTUGAL

EL IMPACTO DEL BIG DATA EN LAS BIBLIOTECAS: OPORTUNIDADES, RIESGOS Y TRANSFORMACIÓN DE LOS SERVICIOS BIBLIOTECARIOS

Sonia María Puerta Correa, Universidad Carlos III de Madrid, <https://orcid.org/0009-0009-1092-0875>, España, 100547539@alumnos.uc3m.es

1 Introducción

El presente trabajo tiene como finalidad analizar el impacto del Big Data en el ámbito bibliotecario, considerando tanto sus aplicaciones prácticas como los desafíos éticos y sociales que conlleva su implementación. La elección de este tema surge del creciente protagonismo que los datos masivos han adquirido en múltiples sectores, y de la necesidad de comprender cómo esta tecnología puede ser aprovechada en las bibliotecas sin comprometer sus principios fundamentales.

El Big Data se ha consolidado como una herramienta estratégica en ámbitos muy diversos: en el sector sanitario, ha permitido detectar patrones epidemiológicos y mejorar diagnósticos personalizados; en las finanzas, se emplea para prevenir fraudes y analizar riesgos; en el comercio, para optimizar decisiones de marketing y personalizar la experiencia del cliente; en el transporte, para planificar rutas más eficientes; y en la educación, para adaptar contenidos y evaluar procesos de aprendizaje. Asimismo, en el ámbito político, su uso ha suscitado controversias relevantes: un caso paradigmático es el de la empresa Cambridge Analytica, que recopiló datos personales sin consentimiento explícito para influir en procesos electorales como el referéndum del Brexit o la campaña de Donald Trump en 2016. Para esta última, a través de perfiles psicográficos, el equipo de campaña segmentó a votantes indecisos y diseñó mensajes altamente emocionalizados y dirigidos, en ocasiones aprovechando asociaciones simbólicas con productos culturales (De Llano

Neira, 2018; Postedin, s. f.). Este episodio evidenció hasta qué punto el Big Data puede emplearse no solo para personalizar servicios, sino también para manipular decisiones y emociones de forma opaca.

Frente a este panorama, las bibliotecas no permanecen al margen. Como espacios de mediación del conocimiento, generan y gestionan una enorme cantidad de datos sobre el comportamiento informacional de sus usuarios: préstamos, consultas, búsquedas, participación en actividades, etc. Analizar y aprovechar estos datos de forma responsable podría permitirles mejorar sus servicios, anticipar las necesidades y optimizar los recursos. Sin embargo, este potencial debe equilibrarse con el respeto a la privacidad, la autonomía del usuario y la ética profesional. Por ello, este trabajo busca reflexionar críticamente sobre las oportunidades y límites del Big Data en el entorno bibliotecario, en un momento en el que la gestión responsable de la información se vuelve más necesaria que nunca.

El fenómeno del Big Data ha comenzado a tener una presencia significativa también en bibliotecas de referencia a nivel internacional. Por ejemplo, la Biblioteca del Congreso de Estados Unidos ha desarrollado proyectos de minería de datos a partir de colecciones digitales masivas, mientras que la British Library ha puesto en marcha herramientas analíticas para estudiar patrones de uso de sus fondos digitales. En países nórdicos, algunas bibliotecas públicas han explorado el uso de análisis predictivo para planificar adquisiciones según tendencias lectoras emergentes. Estos casos demuestran que el potencial del Big Data

trasciende los límites del marketing o la gestión empresarial y puede ser una herramienta al servicio del acceso al conocimiento, siempre que se gestione bajo principios éticos sólidos.

En el plano normativo y profesional, también se observa una creciente atención hacia los efectos del uso masivo de datos. Organizaciones como la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA) han subrayado la importancia de aplicar marcos éticos sólidos en el tratamiento de información personal en entornos bibliotecarios. Del mismo modo, la legislación europea sobre protección de datos, recogida en el Reglamento General de Protección de Datos (GDPR), establece principios fundamentales como el consentimiento informado, el derecho al olvido o la minimización de datos, que deben ser tenidos en cuenta al implementar sistemas de análisis masivo en bibliotecas. Estas normativas no solo delimitan lo que es legal, sino que reflejan un consenso social sobre la necesidad de proteger la dignidad y la autonomía de los usuarios en la era digital.

Este trabajo adopta una metodología de carácter cualitativo basada en la revisión documental y bibliográfica. Se ha realizado un análisis crítico de fuentes académicas especializadas en Big Data, biblioteconomía y ética de la información, así como de informes institucionales y casos de aplicación en distintos sectores, con especial atención al ámbito bibliotecario. La elección de esta metodología responde al objetivo de comprender en profundidad tanto las potencialidades como los riesgos del uso de datos masivos en entornos informacionales, sin limitarse a una mera descripción técnica, sino incorporando una perspectiva ética, social y profesional.

En definitiva, comprender el impacto del Big Data en las bibliotecas implica ir más allá de su mera utilidad instrumental. Se trata de analizar cómo estas tecnologías reconfiguran los modos de acceso a la información, las relaciones entre usuarios e instituciones, y los marcos de decisión profesional. Este enfoque permite no solo identificar oportunidades, sino también

anticipar dilemas y construir propuestas de futuro que sitúen a las bibliotecas como referentes en el uso responsable de los datos y en la defensa de los derechos informacionales de la ciudadanía.

Desde esta perspectiva, el análisis del Big Data desde el prisma bibliotecario se vincula estrechamente con varios de los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030, especialmente con el ODS 4 (Educación de calidad), el ODS 9 (Industria, innovación e infraestructura) y el ODS 16 (Paz, justicia e instituciones sólidas). Las bibliotecas, como garantes del derecho a la información y agentes clave en la alfabetización digital, tienen el potencial de utilizar los datos masivos no solo para mejorar su eficacia, sino también para fortalecer su función social. La integración crítica del Big Data puede contribuir, por ejemplo, a una gestión más inclusiva de los recursos, a la detección de necesidades formativas emergentes o a la reducción de desigualdades en el acceso a la información.

Por ello, a lo largo del presente trabajo se exploran estos y otros aspectos mediante un enfoque teórico y crítico, que parte de una revisión bibliográfica contrastada y se estructura en torno a cinco grandes ejes: la fundamentación conceptual del Big Data, la metodología adoptada, sus aplicaciones prácticas en el contexto bibliotecario, los principales riesgos asociados a su uso y una reflexión final sobre las oportunidades futuras. Se trata, en definitiva, de una contribución al debate actual sobre el papel que deben desempeñar las bibliotecas en la sociedad del conocimiento, no solo como gestoras de información, sino como instituciones que modelan el acceso, el uso y la ética de los datos.

2 Referencial Teórico

La presente investigación se fundamenta en una selección rigurosa de fuentes académicas y técnicas que permiten comprender el fenómeno del Big Data desde una perspectiva multidisciplinar. Se han consultado artículos científicos, monografías especializadas, capítulos de libros y documentos

institucionales que abordan el impacto de los datos masivos en distintos sectores, con especial atención al ámbito bibliotecario.

El enfoque adoptado ha sido crítico y comparativo, integrando fuentes nacionales e internacionales que aportan tanto marcos teóricos como ejemplos prácticos. Esta diversidad ha permitido construir una visión equilibrada entre las potencialidades del Big Data y los riesgos éticos derivados de su uso indiscriminado. En particular, se han considerado las aportaciones de autores como Zuboff (2020), cuya crítica al capitalismo de la vigilancia resulta esencial para entender los dilemas actuales relacionados con el control y la explotación de datos personales; Pariser (2017), quien alerta sobre los efectos de la personalización algorítmica en la diversidad informativa; y Ríos Hilario y Fraile Sastre (2020), que abordan el uso del Big Data en el entorno bibliotecario desde una perspectiva aplicada.

Además, se han tenido en cuenta marcos normativos internacionales, como el Reglamento General de Protección de Datos (GDPR) de la Unión Europea, así como los principios éticos promovidos por organismos como la IFLA o la UNESCO, que refuerzan la necesidad de un uso responsable de la tecnología en contextos de acceso público a la información. Todas las referencias se presentan al final del trabajo siguiendo el estilo APA (7.ª edición).

2.1 Definición y evolución del Big Data

El término Big Data hace referencia al procesamiento y análisis de grandes volúmenes de datos que, por su tamaño, velocidad de generación y complejidad estructural, superan las capacidades de los sistemas tradicionales. La definición más influyente es la propuesta por Doug Laney (2001), quien estableció el modelo de las “3V”: volumen, variedad y velocidad. A esta clasificación inicial se han ido sumando nuevas dimensiones propuestas por distintos autores y organismos, como la veracidad, el valor, la variabilidad o la validez, lo que ha dado lugar a

una comprensión más rica y crítica del fenómeno (Moreno-Carriles, 2018; Cordero, 2025).

- Volumen: se refiere a la cantidad ingente de datos generados continuamente por sensores, dispositivos móviles, redes sociales, transacciones, servicios digitales, etc.
- Velocidad: indica la rapidez con la que los datos son creados, transmitidos y analizados, en muchos casos en tiempo real.
- Variedad: alude a la heterogeneidad de los datos (estructurados, semiestructurados y no estructurados): textos, imágenes, vídeos, audios, logs de navegación, etc.
- Veracidad: hace referencia a la calidad y fiabilidad de los datos. El ruido informacional y la desinformación son desafíos clave.
- Valor: consiste en la capacidad de transformar los datos en conocimiento útil para la toma de decisiones.
- Variabilidad: describe la inestabilidad contextual de ciertos datos cuyo significado puede cambiar según el entorno de análisis.

El desarrollo del Big Data está estrechamente ligado a los avances tecnológicos de las últimas décadas: la expansión de internet, la computación en la nube, el abaratamiento del almacenamiento digital y la proliferación del Internet de las Cosas (IoT). En un principio, su uso se limitó a entornos empresariales, donde permitió optimizar estrategias de marketing, segmentar clientes y anticipar tendencias. Con el tiempo, se ha extendido a sectores como la medicina personalizada, la predicción de delitos, la investigación científica o la planificación urbana, y hoy constituye uno de los pilares de la transformación digital.

En el ámbito bibliotecario, el Big Data se asocia con herramientas como *dashboards* de análisis, software de minería de datos, sistemas de recomendación y gestores de metadatos enriquecidos. Estos instrumentos permiten a las bibliotecas interpretar el comportamiento informacional de sus

usuarios, gestionar sus colecciones de forma eficiente y justificar políticas de adquisición o renovación de servicios. No obstante, el uso de estos datos plantea también interrogantes sobre el consentimiento, la anonimización, el sesgo algorítmico y la privacidad del usuario, temas que deben abordarse con rigor y transparencia.

2.2 Principales características del Big Data

Vivimos en una sociedad donde los datos se han convertido en un recurso estratégico comparable al capital o la energía. Sin embargo, el acceso a datos no garantiza el acceso al conocimiento. La transformación de información en conocimiento útil exige capacidades cognitivas avanzadas, pensamiento crítico, competencias digitales y una ética sólida en el tratamiento de la información (Colmenarejo Fernández & Pedrajas, 2017). En este contexto, el Big Data actúa como una herramienta ambivalente: puede facilitar la toma de decisiones fundamentadas o, por el contrario, ser instrumento de manipulación, exclusión y vigilancia.

Autores como Han (2018) o Zuboff (2020) denuncian la progresiva sustitución de la esfera pública por lógicas algorítmicas opacas, donde las decisiones son tomadas por sistemas automatizados que escapan al control ciudadano. Este fenómeno, al que algunos llaman “infocracia”, pone en entredicho los principios democráticos de deliberación, transparencia y rendición de cuentas. Frente a ello, las bibliotecas tienen la posibilidad de actuar como espacios de resistencia ética y formación crítica.

Desde esta perspectiva, el papel de las bibliotecas en la sociedad del conocimiento va más allá de la gestión de información: se trata de formar a ciudadanos capaces de comprender cómo se producen, seleccionan y usan los datos en entornos digitales. La alfabetización en datos, la visualización crítica de información, el fomento de la diversidad informativa y la defensa de la privacidad son

hoy competencias clave que las bibliotecas deben promover activamente.

En suma, el Big Data no es solo una tecnología: es una cuestión política, cultural y ética. Su aplicación en bibliotecas debe estar guiada por una reflexión profunda sobre los fines que se persiguen y los valores que se defienden. De este modo, será posible construir una bibliotecología informada, crítica y al servicio de una ciudadanía más libre y empoderada.

2.3 Big Data y su papel en la sociedad del conocimiento

La implementación del Big Data en bibliotecas no se limita a una transformación conceptual, sino que requiere también el uso de tecnologías específicas que permitan capturar, procesar, analizar y visualizar grandes volúmenes de datos. Estas herramientas varían según el propósito del análisis y el nivel de infraestructura disponible, pero todas comparten el objetivo de convertir datos dispersos en conocimiento accionable.

Entre las plataformas más utilizadas en entornos profesionales destaca Apache Hadoop, un marco de software de código abierto diseñado para almacenar y procesar grandes volúmenes de datos distribuidos. Gracias a su arquitectura escalable, Hadoop permite a las instituciones manejar información masiva sin necesidad de servidores centralizados costosos, lo que resulta especialmente útil para bibliotecas con colecciones digitales amplias o datos de uso acumulado durante años.

Otra herramienta destacada es Elasticsearch, un motor de búsqueda basado en texto que permite realizar consultas complejas y análisis en tiempo real. En el ámbito bibliotecario, se utiliza para mejorar la recuperación de información en catálogos, portales de descubrimiento y repositorios institucionales, facilitando búsquedas más precisas e intuitivas para los usuarios.

Para el análisis de datos bibliométricos y la visualización de redes, destacan aplicaciones como VOSviewer y Gephi. Ambas permiten representar gráficamente relaciones entre

autores, instituciones, publicaciones o términos clave, facilitando la interpretación de grandes conjuntos de datos académicos. Estas herramientas han sido empleadas con éxito en bibliotecas universitarias para mapear áreas de investigación, identificar colaboraciones emergentes o evaluar la producción científica de sus comunidades académicas, a menudo utilizando como fuente de datos bases como Web of Science (WoS) o Scopus, que ofrecen indicadores estandarizados y estructuras de metadatos especialmente adecuadas para este tipo de estudios.

Asimismo, OpenRefine se ha consolidado como una solución eficaz para la limpieza, normalización y enriquecimiento de datos. Su utilidad es especialmente visible en proyectos de integración de metadatos procedentes de diferentes fuentes o en la preparación de conjuntos de datos para su análisis posterior.

Además, muchas bibliotecas han comenzado a utilizar tableros de mando o *dashboards* personalizados, que permiten visualizar en tiempo real indicadores clave como la frecuencia de préstamos, las tasas de asistencia a actividades, los patrones de navegación en recursos electrónicos o la satisfacción del usuario. Estos sistemas, basados en herramientas como Power BI o Google Data Studio, ofrecen una visión panorámica del funcionamiento de la biblioteca y permiten tomar decisiones basadas en evidencias.

No obstante, el uso de estas tecnologías también plantea retos significativos. Requiere personal capacitado, inversiones en infraestructura, estrategias claras de protección de datos y una cultura organizacional orientada al uso crítico de la información. En este sentido, la formación del personal bibliotecario en competencias digitales y analíticas se convierte en un factor clave para que el potencial del Big Data se traduzca en mejoras reales y éticamente sostenibles en los servicios ofrecidos.

En resumen, las bibliotecas que incorporan herramientas de análisis de datos no solo están modernizando su gestión interna, sino que

están avanzando hacia un modelo más proactivo, transparente y centrado en las necesidades de sus comunidades. Estas tecnologías no sustituyen la función humanista de la biblioteca, sino que la complementan, siempre que su uso esté guiado por valores éticos y una orientación al bien común.

3 Procedimientos Metodológicos

Este trabajo adopta una metodología de carácter cualitativo basada en la revisión documental y bibliográfica, lo que resulta especialmente pertinente para abordar un fenómeno complejo, emergente y en constante evolución como es el Big Data. El enfoque cualitativo permite explorar en profundidad las implicaciones sociales, éticas y profesionales del uso masivo de datos en bibliotecas, más allá de la cuantificación o medición estadística de variables.

La investigación parte de una revisión exhaustiva de fuentes académicas especializadas en Big Data, ciencia de la información, ética informacional y biblioteconomía, complementadas con informes institucionales de organismos como la IFLA, la UNESCO o la Comisión Europea. Asimismo, se han incorporado estudios de caso y experiencias prácticas documentadas en bibliotecas reales, lo que ha permitido vincular la reflexión teórica con ejemplos aplicados de interés. Esta triangulación de fuentes garantiza una aproximación integral y actualizada al objeto de estudio.

El desarrollo del trabajo ha seguido las etapas propias de una investigación teórica: en primer lugar, se procedió a la delimitación del tema, centrado en la intersección entre el Big Data y el entorno bibliotecario. A partir de ahí, se formularon preguntas orientadoras como: ¿Qué aplicaciones concretas del Big Data se están desarrollando en bibliotecas? ¿Qué riesgos éticos implica su implementación? ¿Cómo pueden las bibliotecas mantener sus principios fundacionales en un contexto de vigilancia digital? Estas cuestiones guiaron la búsqueda y selección de bibliografía relevante, procedente de bases de datos académicas,

catálogos universitarios y repositorios institucionales.

Las fuentes seleccionadas fueron organizadas temáticamente en función de su contribución al marco conceptual, la dimensión práctica o la perspectiva crítica. Para ello se elaboraron fichas de lectura, esquemas comparativos y matrices de contenido que facilitaron la identificación de patrones recurrentes, tensiones entre enfoques o lagunas en la literatura. Este proceso permitió construir un discurso coherente, basado en evidencias textuales contrastadas, que articula los distintos apartados del trabajo.

Aunque el estudio no incluye una fase empírica de recogida de datos mediante encuestas, entrevistas o análisis estadísticos, la metodología adoptada permite generar conocimiento útil y orientado a la práctica profesional. La revisión bibliográfica no se limita a una descripción acumulativa, sino que se aplica con un enfoque analítico, interpretativo y propositivo, que busca problematizar los discursos dominantes sobre la neutralidad tecnológica, el uso de algoritmos en entornos de servicio público o la privacidad informacional de los usuarios.

Un aspecto relevante del procedimiento metodológico es la atención a la dimensión ética de la investigación. Se ha procurado citar todas las fuentes utilizadas de manera transparente, respetar los derechos de autor y contrastar la información procedente de informes o artículos. Además, se ha mantenido un compromiso con la pluralidad de perspectivas, incorporando tanto voces críticas como posturas más optimistas sobre el papel del Big Data en las bibliotecas.

Entre las limitaciones del estudio cabe señalar que, al tratarse de una investigación teórica, los resultados no pueden generalizarse empíricamente ni se derivan de una experimentación directa. Tampoco se han incorporado testimonios de profesionales bibliotecarios o datos propios de usuarios, lo cual podría enriquecer futuros estudios complementarios. No obstante, esta limitación se ve compensada por la profundidad analítica,

la conexión interdisciplinar y la actualidad del objeto de estudio.

En definitiva, este procedimiento metodológico ha permitido construir un marco interpretativo sólido, que fundamenta las reflexiones y propuestas expuestas en el resto del trabajo. El enfoque cualitativo y documental se presenta, por tanto, como una estrategia válida y necesaria para abordar los retos éticos, técnicos y profesionales que plantea el uso del Big Data en las bibliotecas contemporáneas.

4 Aplicaciones del Big Data en el ámbito bibliotecario

La progresiva digitalización de los servicios bibliotecarios ha generado un volumen creciente de datos que, si se analizan adecuadamente, pueden ofrecer un conocimiento valioso sobre el funcionamiento de las bibliotecas y las necesidades de sus comunidades. Este apartado examina algunas de las principales aplicaciones del Big Data en el entorno bibliotecario, con especial atención al análisis del comportamiento de los usuarios, la gestión de colecciones y la visualización de datos para la toma de decisiones.

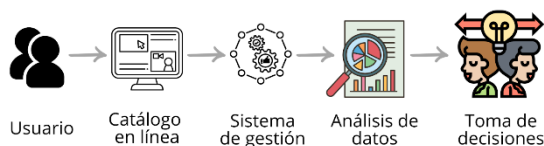
4.1 Análisis de usuarios y mejora de servicios

Una de las aplicaciones más significativas del Big Data en el ámbito bibliotecario es el análisis del comportamiento de los usuarios. Cada acción (como realizar una búsqueda en el catálogo, solicitar un préstamo, participar en una actividad o acceder a recursos electrónicos), genera información que puede recogerse y analizarse para comprender mejor las necesidades reales de la comunidad usuaria. En esta línea, Ríos Hilario y Fraile Sastre (2020) destacan que las bibliotecas universitarias consideran especialmente valiosos los datos internos generados por la propia institución, ya que estos permiten evaluar el funcionamiento del servicio y orientar mejoras basadas en evidencias.

La imagen siguiente representa el recorrido de los datos desde la interacción del usuario con

el catálogo hasta su análisis y utilización para la mejora de los servicios bibliotecarios.

Figura 1: Flujo de datos en una biblioteca digital



Fuente: elaboración propia.

Gracias a este tipo de análisis, es posible identificar patrones de uso, detectar cambios en los intereses temáticos, ajustar horarios o evaluar el impacto de nuevos servicios. Esta información facilita una planificación más eficiente y una toma de decisiones menos basada en intuiciones y más en datos objetivos. Como afirman los autores, uno de los principales beneficios del Big Data en bibliotecas radica en su capacidad para personalizar la atención, adaptar los recursos a las demandas reales y optimizar el uso institucional de forma estratégica.

Además, el uso del Big Data permite avanzar hacia modelos de servicio más personalizados. Por ejemplo, los sistemas de recomendación basados en algoritmos pueden sugerir lecturas según el historial de préstamo, intereses o búsquedas previas. Estas técnicas, ampliamente empleadas en plataformas comerciales como Amazon o Netflix, están comenzando a explorarse también en el ámbito bibliotecario para adaptar la oferta informativa a las preferencias reales de cada usuario.

Sin embargo, esta personalización debe gestionarse con especial cautela. Como advierte Pariser (2017), los algoritmos que filtran la información tienden a construir burbujas invisibles que refuerzan nuestras ideas previas y reducen nuestra exposición a contenidos diversos, lo que puede limitar la capacidad crítica y el acceso a perspectivas alternativas. En el caso de las bibliotecas, esto implica un riesgo: lejos de ser espacios de apertura y pluralidad, podrían transformarse en entornos cerrados si los datos se utilizan sin una reflexión ética. Por ello, la personalización debe estar al servicio del usuario, pero sin comprometer su privacidad, su autonomía ni el

principio de acceso libre y equitativo a la información.

4.2 Gestión de colecciones y adquisiciones

El uso del Big Data también está transformando los procesos internos de las bibliotecas, especialmente en lo que respecta a la gestión de colecciones y a las decisiones de adquisición. Gracias al análisis de datos históricos sobre circulación, reservas, renovaciones, quejas o solicitudes de los usuarios, es posible identificar qué materiales tienen mayor impacto, cuáles presentan baja rotación o qué áreas temáticas muestran una demanda emergente. En este sentido, Ríos Hilario y Fraile Sastre (2020) subrayan que las bibliotecas valoran especialmente los datos internos generados por su propia actividad, ya que permiten conocer mejor su funcionamiento y orientar decisiones estratégicas.

Estos datos permiten fundamentar con mayor rigor las decisiones sobre compra, descarte o renovación de fondos, optimizando los recursos presupuestarios disponibles. Asimismo, ayudan a diseñar colecciones más alineadas con los intereses reales de la comunidad a la que sirve la biblioteca, reduciendo la brecha entre la oferta documental y las necesidades informativas de los usuarios. De hecho, uno de los campos donde las bibliotecas universitarias prevén un mayor potencial de mejora gracias al Big Data es precisamente la adecuación de servicios y recursos a los perfiles y comportamientos reales de su comunidad académica (Ríos Hilario & Fraile Sastre, 2020).

El Big Data también facilita la evaluación continua del rendimiento de recursos electrónicos como bases de datos, revistas digitales o plataformas de préstamo en línea. A través de herramientas estadísticas y cuadros de mando, los profesionales pueden acceder a información detallada sobre tasas de uso, momentos pico de consulta o patrones de abandono, lo que permite negociar licencias más ajustadas a la demanda real o redirigir esfuerzos de formación hacia recursos

infrautilizados. Los datos masivos, analizados en tiempo real o retrospectivamente, ofrecen así una base sólida para tomar decisiones más eficientes y justificadas.

Por otro lado, el análisis predictivo basado en series temporales o algoritmos de aprendizaje automático puede emplearse para anticipar tendencias futuras en la demanda documental, permitiendo una planificación más proactiva. No obstante, es esencial recordar que la interpretación de los datos debe tener en cuenta el contexto social, cultural y educativo, evitando una excesiva dependencia de modelos automatizados que puedan invisibilizar a colectivos minoritarios o reproducir sesgos preexistentes.

4.3 Visualización de datos y toma de decisiones

La visualización de datos es una herramienta fundamental para transformar conjuntos masivos de datos en conocimiento accesible y útil. En el ámbito bibliotecario, permite representar gráficamente patrones de uso, evolución de servicios, impacto de actividades o distribución temática de colecciones, facilitando así la toma de decisiones fundamentadas. Como afirman Adell y Guersenzvaig (2013), los datos no constituyen información por sí mismos; es su organización, procesamiento y visualización lo que les otorga significado y utilidad para el análisis.

Paneles de control (*dashboards*), mapas de calor, gráficos interactivos o informes dinámicos son formatos cada vez más comunes en bibliotecas para comunicar datos tanto a responsables institucionales como a la ciudadanía. Esta visualización de la información permite justificar presupuestos, rediseñar servicios o evaluar programas en base a evidencias. Ríos Hilario y Fraile Sastre (2020) destacan que este tipo de análisis interno es especialmente valorado en el entorno bibliotecario por su utilidad para ajustar los servicios a las necesidades reales de los usuarios.

La visualización también cumple una función clave en la transparencia institucional. Publicar

datos sobre circulación, asistencia o uso de recursos refuerza la rendición de cuentas y puede fomentar la participación si se acompaña de estrategias de alfabetización en datos. Sin embargo, como advierte Zárates-Sánchez (2024), los errores lógicos en la interpretación visual (como la falacia de la falsa causalidad o la paradoja de Simpson) pueden generar lecturas erróneas si no se contextualizan adecuadamente.

Además, la visualización moderna ha dejado de ser estática: hoy se espera que sea interactiva, adaptable e intuitiva. Adell y Guersenzvaig (2013) insisten en que no basta con una buena representación gráfica; es necesario un diseño centrado en el usuario, con atención a la experiencia de navegación, la claridad de los elementos y la posibilidad de manipular los datos para generar nuevas preguntas. Esta interacción potencia el análisis cultural y social a partir de grandes volúmenes de datos.

Por ello, el desarrollo de competencias en visualización crítica y el uso de herramientas especializadas es clave en las bibliotecas actuales. Herramientas como OpenRefine, para la limpieza y normalización de datos, o VOSviewer, para la representación de redes y patrones bibliométricos, permiten tratar grandes volúmenes de información y ofrecer representaciones comprensibles, rigurosas y orientadas a la toma de decisiones. Como resumen Adell y Guersenzvaig (2013), una visualización eficaz requiere la colaboración entre analistas, diseñadores y programadores para convertir los datos en conocimiento compartido.

4.4 Inteligencia artificial y análisis predictivo en bibliotecas

La integración del Big Data con tecnologías de inteligencia artificial (IA) representa una de las aplicaciones más innovadoras y, al mismo tiempo, más controvertidas en el entorno bibliotecario. Gracias al uso de algoritmos de aprendizaje automático, las bibliotecas pueden anticipar necesidades informativas, automatizar procesos rutinarios e incluso ofrecer atención personalizada a través de

sistemas conversacionales, conocidos como chatbots.

El análisis predictivo basado en datos históricos permite a las bibliotecas identificar tendencias futuras en la demanda de recursos o servicios. Por ejemplo, mediante el análisis de series temporales sobre préstamos o reservas, es posible prever picos de demanda y adaptar la oferta documental o la asignación de personal. Esta capacidad de anticipación contribuye a mejorar la planificación operativa y a maximizar la eficiencia de los recursos disponibles.

Además, algunos centros han comenzado a experimentar con sistemas de recomendación automatizados, similares a los que utilizan plataformas comerciales, pero con objetivos formativos o culturales. Estos sistemas analizan los historiales de consulta, los intereses temáticos y las interacciones previas del usuario para sugerir lecturas, recursos académicos o actividades relevantes. Aunque estas herramientas pueden enriquecer la experiencia del usuario y fomentar el descubrimiento de contenidos, también plantean interrogantes éticos similares a los ya señalados en relación con la personalización algorítmica: ¿hasta qué punto se respeta la diversidad informativa? ¿Existe riesgo de reforzar sesgos o de limitar el pensamiento crítico del usuario?

Otra aplicación emergente de la IA en bibliotecas es el desarrollo de chatbots capaces de responder preguntas frecuentes, guiar al usuario en la búsqueda de información o ayudar en la navegación por los recursos electrónicos. Estos asistentes virtuales, cuando están bien diseñados, pueden mejorar significativamente la accesibilidad y la disponibilidad del servicio, especialmente fuera del horario habitual o en contextos de sobrecarga del personal. Sin embargo, su efectividad depende en gran medida de la calidad de los datos con los que han sido entrenados, así como de su capacidad para entender el lenguaje natural y adaptarse a distintos perfiles de usuarios.

Por otro lado, el uso de técnicas de minería de textos (*text mining*) y procesamiento del lenguaje natural (PLN) también está siendo explorado en bibliotecas académicas y de investigación. Estas técnicas permiten analizar grandes corpus documentales, identificar tendencias temáticas, mapear áreas de conocimiento o detectar duplicidades y vacíos en las colecciones. Este tipo de análisis puede contribuir a mejorar las decisiones de adquisición, a optimizar los catálogos y a reforzar el papel de las bibliotecas como agentes activos en la producción y organización del conocimiento.

No obstante, la incorporación de IA en bibliotecas debe ser abordada con una mirada crítica. Tal como advierte Han (2022), la automatización informacional corre el riesgo de desplazar el juicio profesional y de reducir la complejidad del pensamiento a respuestas programadas. La biblioteca no debe renunciar a su función mediadora ni a su papel formativo, sino encontrar un equilibrio entre innovación tecnológica y humanismo informacional.

En este sentido, resulta imprescindible que los desarrollos tecnológicos en bibliotecas estén acompañados de políticas éticas claras, formación del personal, transparencia en los algoritmos utilizados y participación de la comunidad usuaria en las decisiones tecnológicas. Solo así será posible garantizar que la inteligencia artificial y el Big Data se conviertan en aliados al servicio del acceso equitativo, crítico y plural al conocimiento.

5 Riesgos y desafíos

La incorporación del Big Data en bibliotecas, si bien ofrece múltiples beneficios, también plantea una serie de riesgos y desafíos que no pueden ser ignorados. A diferencia de otros sectores, las bibliotecas están estrechamente ligadas a principios como la privacidad, la libertad intelectual o el acceso equitativo, lo que implica una especial sensibilidad frente al uso de tecnologías que pueden afectar a estos valores. Antes de abordar los principales riesgos, se presenta a continuación un análisis DAFO que sintetiza de forma visual las

fortalezas, debilidades, oportunidades y amenazas que implica el uso del Big Data en el entorno bibliotecario.

Figura 2: Análisis DAFO del uso del Big Data en bibliotecas



Fuente: elaboración propia.

5.1 Privacidad y protección de datos

Uno de los principales riesgos del uso del Big Data en bibliotecas es la vulneración de la privacidad de los usuarios. Cada interacción (una búsqueda en el catálogo, un préstamo o la asistencia a una actividad) genera un volumen de datos que, si no se gestionan de forma segura, pueden comprometer derechos fundamentales.

Este riesgo se agrava si los usuarios no son conscientes del alcance del tratamiento de sus datos ni del uso que puede hacerse de ellos (Zuboff, 2020). Este autor define esta dinámica como capitalismo de la vigilancia, un modelo económico que convierte la experiencia humana en materia prima gratuita para ser extraída y comercializada sin consentimiento real.

Desde esta perspectiva, aceptar términos de uso poco claros o políticas de *cookies* invasivas equivale a una expropiación de la autonomía informativa. Por ello, en el ámbito bibliotecario resulta imprescindible establecer políticas claras de privacidad, mecanismos de anonimización robustos y protocolos de consentimiento informado adaptados a un entorno no lucrativo, tal como exigen los principios éticos de la profesión.

5.2 Brechas digitales y exclusión

El Big Data también puede acentuar desigualdades ya existentes en el acceso a los servicios bibliotecarios. Las bibliotecas con mayor dotación tecnológica estarán en mejores condiciones para implementar herramientas avanzadas de análisis, mientras que otras pueden quedar relegadas, reproduciendo desigualdades estructurales.

Además, si los datos utilizados para tomar decisiones provienen únicamente de usuarios activos, es probable que queden invisibilizados aquellos colectivos que interactúan menos con los entornos digitales (Colmenarejo Fernández & Pedrajas, 2017). Esta exclusión no es solo tecnológica, sino también epistémica. Como advierte Zuboff (2020), el capitalismo de la vigilancia actúa mediante una asimetría radical del conocimiento: las plataformas lo saben todo sobre nosotros, mientras que nosotros apenas sabemos nada sobre sus procesos internos. Por tanto, interpretar los datos sin una perspectiva inclusiva puede llevar a decisiones sesgadas. Es necesario complementar el análisis cuantitativo con metodologías participativas y cualitativas que integren la voz de colectivos tradicionalmente subrepresentados en los datos.

5.3 Ética de los algoritmos y sesgos

Los sistemas algorítmicos que gestionan y analizan grandes volúmenes de datos no son neutrales. Están diseñados por seres humanos, entrenados con datos históricos y tienden a reproducir desigualdades sociales preexistentes. En el entorno bibliotecario, esto puede traducirse en recomendaciones automatizadas que favorezcan a determinados grupos, perpetuando sesgos de género, clase o etnia (San Segundo, Martínez-Ávila & Frías Montoya, 2023).

Además, una confianza excesiva en los algoritmos puede disminuir la autonomía profesional del personal bibliotecario. Como recuerdan estos autores, “el usuario es el contenido”: nuestras decisiones, preferencias y comportamientos alimentan sistemas que luego dictan lo que debemos ver, leer o

consumir, generando un peligroso bucle de retroalimentación. Es lo que Eli Pariser (2011) llamó filtro burbuja, algo que limita nuestra exposición a la diversidad informativa. Ante esta realidad, es crucial fomentar la transparencia algorítmica, el pensamiento crítico y la supervisión humana, tal como proponen Colmenarejo Fernández y Pedrajas (2017), quienes insisten en la necesidad de mantener un enfoque ético en la toma de decisiones apoyadas en datos.

5.4 Rendición de cuentas, transparencia y derechos digitales

En el contexto del Big Data, la rendición de cuentas y la transparencia institucional se han convertido en valores clave para preservar la confianza de los usuarios y garantizar una gestión ética de los datos. Las bibliotecas, como espacios de acceso libre y servicio público, deben asumir una responsabilidad activa en la protección de los derechos digitales de sus comunidades.

Uno de los grandes desafíos del entorno digital es la opacidad con la que se gestionan los algoritmos y los sistemas de análisis de datos. En muchas ocasiones, las decisiones automatizadas se toman a partir de procesos que no son comprensibles para los usuarios ni fácilmente auditables por terceros. Esta falta de transparencia puede erosionar la legitimidad de las instituciones, incluso cuando las intenciones son positivas. Tal como señalan Colmenarejo Fernández y Pedrajas (2017), una gobernanza algorítmica sin supervisión externa ni participación ciudadana puede derivar en prácticas tecnocráticas que vulneran los principios democráticos.

Las bibliotecas tienen el potencial de actuar como modelos de transparencia institucional. Esto implica no solo proteger los datos personales, sino también informar a los usuarios sobre qué datos se recogen, con qué finalidad, durante cuánto tiempo se conservan y cómo se toman decisiones automatizadas, si las hubiera. Publicar memorias de actividad con indicadores desglosados, mantener políticas de privacidad accesibles y promover

espacios de participación activa son prácticas que refuerzan la legitimidad institucional y la autonomía de los usuarios.

Además, la defensa de los derechos digitales debe ir más allá del ámbito técnico. El derecho a ser informado, el derecho a la portabilidad de los datos y el derecho al olvido son aspectos clave del Reglamento General de Protección de Datos (RGPD) de la Unión Europea, que también deben aplicarse en contextos bibliotecarios. Por ejemplo, los usuarios deberían poder solicitar la supresión de sus historiales de préstamo, navegación o consulta, especialmente si se han utilizado para generar perfiles o recomendaciones personalizadas.

En este sentido, las bibliotecas pueden desempeñar un papel pedagógico fundamental, actuando como espacios de alfabetización digital crítica. Informar a la ciudadanía sobre los derechos que les asisten en el entorno digital, sobre los riesgos de la sobreexposición de datos y sobre las formas de proteger su identidad informacional contribuye a una sociedad más justa, autónoma y consciente. La transparencia institucional debe ir acompañada, por tanto, de una acción formativa orientada a empoderar a los usuarios en el uso de sus datos y en la toma de decisiones informadas.

También es importante señalar que la rendición de cuentas debe operar en todos los niveles. Las bibliotecas no solo deben rendir cuentas ante las administraciones o los organismos financiadores, sino también ante sus propias comunidades usuarias. Incorporar mecanismos de evaluación participativa, consultas abiertas o auditorías éticas externas puede ayudar a detectar prácticas problemáticas y a fortalecer la legitimidad social del uso del Big Data en estos entornos.

Finalmente, en un escenario donde los datos se han convertido en un recurso estratégico, es urgente reclamar una ética pública del dato que contrarreste las lógicas extractivas propias del sector privado. Las bibliotecas, por su tradición humanista y su vocación de servicio, están en una posición privilegiada para liderar

esta transformación. Adoptar prácticas de gobernanza ética, fomentar la corresponsabilidad informacional y garantizar la transparencia son condiciones imprescindibles para un uso justo, crítico y emancipador del Big Data en el ámbito bibliotecario.

6 Conclusiones: perspectivas de futuro

La irrupción del Big Data en el ámbito bibliotecario ha abierto un abanico de posibilidades que van desde la mejora de la toma de decisiones hasta la personalización de servicios y la optimización de recursos. Las bibliotecas, tradicionalmente defensoras del acceso libre y equitativo a la información, se enfrentan ahora al reto de integrar esta tecnología sin comprometer sus valores fundamentales.

Esta integración exige una transformación profunda en el perfil profesional. Ya no basta con dominar las técnicas clásicas de gestión documental; también se requiere formación en análisis de datos, visualización, pensamiento crítico y ética digital. El bibliotecario del siglo XXI debe actuar como un mediador consciente entre los datos y la comunidad, capaz de interpretar la información de forma responsable y al servicio del bien común.

Asimismo, se impone la necesidad de revisar y actualizar los marcos normativos que regulan el tratamiento de datos en contextos bibliotecarios. Aunque existen leyes de protección de datos como el RGPD, la especificidad del trabajo bibliotecario requiere guías y protocolos propios que traduzcan estos principios generales a las prácticas cotidianas del sector. Esta normativa no debe entenderse como una traba, sino como una herramienta para asegurar la equidad, la transparencia y la justicia en el uso de los datos.

Es previsible que en los próximos años el uso del Big Data en bibliotecas se expanda, especialmente en ámbitos como la evaluación del impacto, la automatización de procesos, la anticipación de demandas y el diseño de servicios inteligentes. La incorporación de inteligencia artificial para tareas de

clasificación documental, chatbots para la atención a usuarios o sistemas predictivos para la gestión de colecciones son ejemplos que ya están comenzando a implantarse. No obstante, esta expansión tecnológica no debe conducir a una pérdida del carácter humano, ético y social del trabajo bibliotecario.

La hibridación entre datos y valores es, por tanto, un terreno fértil para la innovación responsable. Lejos de asumir el Big Data como una moda o como una imposición externa, las bibliotecas pueden apropiarse críticamente de estas herramientas para reforzar su misión cultural, educativa y democrática. En este sentido, la formación continua del personal, el fomento del trabajo colaborativo con perfiles técnicos y la participación activa de los usuarios en la toma de decisiones serán claves para garantizar procesos inclusivos y sostenibles.

En este escenario, las bibliotecas tienen la oportunidad de liderar un modelo de innovación tecnológica con responsabilidad social. Para ello, será necesario:

- Adoptar una visión crítica y reflexiva sobre el uso de los datos, priorizando la transparencia, la privacidad y la equidad.
- Formar al personal en competencias relacionadas con el tratamiento ético y estratégico de grandes volúmenes de información.
- Impulsar la alfabetización en datos de la ciudadanía como parte del compromiso educativo de la institución.
- Fomentar el uso de herramientas tecnológicas que favorezcan la toma de decisiones basadas en evidencias sin perder de vista el contexto humano.
- Reforzar el papel de la biblioteca como garante de una sociedad informada, autónoma y capaz de entender los riesgos y oportunidades de la era digital.

Por otro lado, será fundamental fomentar redes de cooperación entre bibliotecas, universidades, centros tecnológicos y

administraciones públicas. Esta colaboración puede facilitar el desarrollo de proyectos piloto, el intercambio de buenas prácticas y la construcción de estándares comunes para el análisis ético de datos en entornos bibliotecarios.

También cabe destacar el valor estratégico de los datos abiertos y la ciencia abierta como líneas de actuación prioritarias. Las bibliotecas pueden desempeñar un papel clave como mediadoras entre los datos generados por la administración y la ciudadanía, promoviendo su reutilización responsable para la mejora de políticas públicas, investigaciones colaborativas o procesos de participación social.

Integrar el Big Data desde una perspectiva estratégica y humanista no solo hará más eficaces a las bibliotecas, sino que fortalecerá su papel como instituciones clave en una sociedad del conocimiento más justa, crítica y participativa.

Finalmente, a medida que el uso del Big Data se consolida en el ámbito bibliotecario, emergen nuevas líneas de investigación que complementan las prácticas ya iniciadas y abren horizontes para una innovación ética y responsable. Resulta prioritario, por ejemplo, estudiar el impacto real de los sistemas de recomendación sobre la diversidad informativa, comparar modelos de visualización de datos en entornos bibliotecarios o analizar críticamente los algoritmos aplicados a contextos culturales y educativos. Asimismo, sería de gran interés desarrollar indicadores que permitan evaluar no solo la eficiencia de los servicios, sino también su equidad, inclusividad y respeto por los derechos digitales. Estas líneas de trabajo pueden contribuir a consolidar un modelo de biblioteca tecnológicamente avanzada, guiada por principios de justicia informacional y compromiso social.

7 Referencias

7.1 Libros

- Colmenarejo Fernández, R., & Pedrajas, M. (2017). *Una ética para Big Data: Introducción a la Gestión ética de Datos Masivos*. Ed. UOC, Tecnología.
- Pariser, E. (2017). *El filtro burbuja: Cómo la web decide lo que leemos y lo que pensamos*. Taurus.
- Zuboff, S. (2020). *La era del capitalismo de la vigilancia: La lucha por un futuro humano frente a las nuevas fronteras del poder*. Paidós.

7.2 Artículos de Revistas

- Adell Español, F., & Guersenzvaig, A. (2013). Big Data y los nuevos métodos de visualización de la información. *Elisava Temes de disseny*, 29, 78-85. <https://recerca.uoc.edu/documentos/672281261dbb3d0eb5d3e8b4?lang=ca>
- Camargo-Vega, J. J., Camargo-Ortega, J. F., & Joyanes-Aguilar, L. (2015). Conociendo big data. *Revista Facultad de Ingeniería*, 24(38), 63-77. http://www.scielo.org.co/scielo.php?pid=S0121-11292015000100006&script=sci_arttext
- Juárez, H. F. G. (2019). Tendencias tecnológicas en el análisis de datos: Patrones a gran escala o pistas específicas de gran valor, la simbiosis entre Big Data y Small Data. *Revista Académica ECO*, 20(20). <https://revistasguatemala.usac.edu.gt/index.php/race/article/view/1062>
- Moreno-Carriles, R. M. (2018). Big data, ¿pero qué es? *Angiología*, 70(5), 191-194. <https://www.sciencedirect.com/science/article/abs/pii/S000331701830066X>
- Ríos Hilario, A. B., & Fraile Sastre, A. (2020). Aproximación al grado de conocimiento y aplicación de Big Data en las bibliotecas

universitarias españolas. *Anales de Documentación*, 23(1).
<https://doi.org/10.6018/analesdoc.390931>

San Segundo, R., Martínez-Ávila, D., & Frías Montoya, J. A. (2023). Ethical Issues in Control by Algorithms: The User is the Content. *Knowledge Organization*, 50(5), 352–358.

Zárate-Sánchez, R. A. (2024). Explorando los desafíos lógicos en la visualización y análisis de datos en arquitecturas de Big data: un enfoque en falacias, sesgos y paradojas. *Revista Nacional de Administración*. Volumen, 15(1), 103-115.

7.3 Ambiente Web

Cordero, E. (24 de marzo 2025). ¿Qué es Big Data y para qué sirve? ICX Consulting.
<https://blog.icx.co/es/transformacion-digital/big-data/que-es-big-data-y-para-que-sirve>

De Llano Neira, P. (18 de marzo de 2018). Una consultora que trabajó para Trump manipuló datos de 50 millones de usuarios de Facebook. *El País*.
https://elpais.com/internacional/2018/03/17/estados_unidos/1521308795_755101.html

Postedin. (s. f.). ¿Cómo el Big Data fue utilizado en la campaña de Trump? <https://www.postedin.com/blog/como-el-big-data-fue-utilizado-en-la-campana-de-trump/>