

***Desafíos en el entorno
de la información
y la documentación
ante las problemáticas
sociales actuales
Volumen 1***

***Coordinadora
Georgina Araceli
Torres Vargas***



Z669.7 D47 Desafíos en el entorno de la información y la documentación ante las problemáticas sociales actuales / Coordinadora Georgina Araceli Torres Vargas. - México : UNAM. Instituto de Investigaciones Bibliotecológicas y de la Información, 2022.

3 v. - (Sistemas bibliotecarios de información y sociedad)

ISBN: 978-607-30-6416-3 (obra completa)

ISBN: 978-607-30-6417-0 (vol.1)

ISBN: 978-607-30-6418-7 (vol.2)

ISBN: 978-607-30-6419-4 (vol.3)

1. Investigación bibliotecológica - México - Conferencias. 2. Investigación bibliotecológica - España - Conferencias. 3. Documentación - México - Conferencias. 4. Documentación - España - Conferencias. 5. Lectores. 6. Tecnología de la información. i. Torres Vargas, Georgina Araceli, coordinadora. ii. ser.

Diseño de portada: Mario Ocampo Chávez

Primera edición: 25 septiembre 2022

D. R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Instituto de Investigaciones Bibliotecológicas y de la Información

Circuito Interior s/n, Torre II de Humanidades,

pisos 11, 12 y 13, Ciudad Universitaria, C. P. 04510,

Alcaldía Coyoacán, Ciudad de México

ISBN de colección: 978-607-30-6416-3

ISBN volumen 1: 978-607-30-6417-0

Esta edición y sus características son propiedad de la Universidad Nacional Autónoma de México. Prohibida la reproducción total o parcial por cualquier medio sin la autorización escrita del titular de los derechos patrimoniales.

Publicación dictaminada

Impreso y hecho en México

Volumen 1

Contenido

PRESENTACIÓN	ix
GEORGINA ARACELI TORRES VARGAS	

I. TÓPICOS EN DOCUMENTACIÓN, BIBLIOTECOLOGÍA E INFORMACIÓN

Desafíos del campo bibliotecológico ante las problemáticas sociales actuales	3
HÉCTOR GUILLERMO ALFARO LÓPEZ	

La ética del científico en tiempos de guerra. Siete calas cinematográficas	17
JOSÉ LÓPEZ YEPES	

Influencias disciplinarias de los términos utilizados en el lenguaje bibliotecológico	53
CATALINA NAUMIS PEÑA	

Sistemas de organización del conocimiento: retos en la indización temática de recursos derivados de dominios de conocimiento emergentes	69
ADRIANA SUÁREZ SÁNCHEZ	

II. LECTORES, BIBLIOTECAS Y PATRIMONIO DOCUMENTAL

Un nuevo paradigma de formación de jóvenes lectores del presente y el futuro: desafíos para la investigación y la educación bibliotecológica	97
ELSA MARGARITA RAMÍREZ LEYVA	

Tendencias de la innovación abierta en la biblioteca postdigital . . .	117
PEDRO RAZQUIN ZAZPE	

Teoría y praxis de la gestión de los recursos humanos en bibliotecas
en épocas de cambios mundiales 137
FEDERICO HERNÁNDEZ PACHECO

El patrimonio documental en la recuperación de la identidad
de una comunidad española en México: los visontinos de Veracruz
y Xalapa 155
JOSÉ LUIS GONZALO SÁNCHEZ-MOLERO

Estudio de la documentación de archivo del encuadernador
Adrián Marín en la Real Academia de Ciencias Exactas, Físicas
y Naturales: 1933-1945 173
ANTONIO CARPALLO BAUTISTA, ANDRÉS MARTÍN
CARDABA Y ADRIÁN TURLUIANU

Un incunable salmantino con tipos de Monterrey: el *Missale
Compostellanum* de 1945 191
MARÍA EUGENIA LÓPEZ VAREA

III. LAS TIC PARA ENFRENTAR LOS DESAFÍOS ACTUALES

El *podcast* como recurso para documentar la historia contemporánea.
Experiencia de creación de la serie *Voces: memorias del covid* 207
PERLA OLIVIA RODRÍGUEZ RESÉNDIZ

Tendencias de los archivos y los archivistas ante los retos
de la información y las tecnologías en la sociedad en México 223
JUAN MIGUEL CASTILLO FONSECA

Sistemas de recomendación basados en datos abiertos enlazados . . 245
EDER ÁVILA BARRIENTOS

Sistemas de recomendación basados en datos abiertos enlazados

EDER ÁVILA BARRIENTOS
*Instituto de Investigaciones Bibliotecológicas
y de la Información de la UNAM, México*

INTRODUCCIÓN

La sistematización de los datos abiertos enlazados (DAE) es una de las variables que se observan con mayor frecuencia dentro de la literatura especializada. Esto ha llevado a los especialistas en el campo de la organización y sistemas de información a incursionar en el estudio de los sofisticados modelos y métodos tecnológicos que buscan delinear y perfilar la generación de sistemas de nueva generación capaces de interactuar en mayor medida con el perfil informativo del usuario. Los sistemas de recomendación (SR) son herramientas digitales que establecen un conjunto de criterios y valoraciones sobre los datos de los usuarios para realizar predicciones sobre recomendaciones de recursos que puedan ser de utilidad o valor para el usuario, en función de su necesidad informativa. Estos sistemas seleccionan datos proporcionados por el usuario de forma directa o indirecta y proceden a analizar y procesar información del historial del usuario para transformar estos datos en recomendaciones.

Se estima que los sistemas de recomendación basados en DAE pueden proporcionar sugerencias de recursos de información que pueden ser útiles para un usuario. Las sugerencias acerca de los recursos se relacionan con varios procesos de toma de decisiones como qué información

consultar, qué música escuchar o qué libros leer. Bajo esta premisa, son sistemas que intentan personalizar al máximo la oferta para cada usuario, acorde a su perfil informativo. La asignación de tripletas semánticas permite que estos sistemas registren perfiles y preferencias en cuanto al uso de información, pues los datos almacenados en el sistema remiten a recursos, manifestaciones y expresiones que se relacionan entre sí, los cuales pueden ser recomendados al usuario al momento de ejercer un proceso de búsqueda o una determinada consulta de información. Por lo tanto, el objetivo de este trabajo consiste en analizar los atributos y la funcionalidad de los sistemas de recomendación basados en datos abiertos enlazados.

EL PARADIGMA DE LA DATIFICACIÓN

Los datos son piezas elementales que permiten la generación de nuevos y constantes conocimientos. Se puede afirmar que sin datos sería imposible explicar los fenómenos que acontecen en la realidad. Estos datos también son importantes para encontrar respuesta a las grandes problemáticas de las sociedades.

En el ámbito científico y de la investigación los datos son útiles para el desarrollo de metodologías, comprobación de hipótesis y explicación de resultados, esto permite a la investigación fundamentar sus áreas de estudio acorde a la realidad que acontece y en donde se manifiestan los fenómenos que forman parte de sus estudios. Durante muchos años, la actividad científica ha utilizado datos para generar nuevos hallazgos y propiciar la generación de nuevos conocimientos, fomentando con ello la generación de un paradigma basado en el uso constante de los datos mediante diversas tecnologías que faciliten su recolección y procesamiento.

De acuerdo con Goble y De Roure¹ nos encontramos en la época de la investigación científica centrada en datos, en la cual las hipótesis no

1 Carol Goble y David de Roure, "El impacto de las herramientas de flujo de trabajo en la investigación centrada en datos", en *El cuarto paradigma: descubrimiento científico intensivo de datos*, editado por Tony Hey, Stewart Tansley y Kristin Tolle (México: Universidad Autónoma Metropolitana, 2014), 149.

sólo se ponen a prueba mediante la recolección y el análisis dirigidos de datos, sino que también son generadas mediante la combinación y la exploración de la reserva de datos ya disponible.

La reserva de cantidades de datos casi abismales ha generado la necesidad de contar con mayores capacidades de almacenamiento en los equipos de supercómputo, pero también ha propiciado la urgencia de contar con sistemas de información capaces de procesar los datos y mostrar todas las variables y atributos que forman parte de ellos.

Es decir, contar con sistemas que enlacen datos de gran escala con modelos matemáticos y contemplar que estos datos pueden ser de tipología variada, por ejemplo, registros de bases de datos, datos numéricos sin procesar recursos y contenidos visuales que permitan realizar descripciones interpretativas de un determinado fenómeno. Por lo tanto, el paradigma de los datos no se basa en la idea plana de registrar datos a través de un instrumento o una determinada metodología, se trata de dar un manejo efectivo a dichos datos contemplando su naturaleza y las variables que permiten considerar su complejidad.

Ahora bien, conceptos como la inteligencia artificial, el *big data* y el aprendizaje automático han venido a contribuir al paradigma de la datificación en los diversos contextos de la actividad humana, en la actualidad, pareciera un tema de moda hablar de la implementación de la inteligencia artificial en los diversos dispositivos, sistemas y aplicaciones con los que los individuos interactúan diariamente. Esto ha sido posible gracias al amplio desarrollo y socialización de las tecnologías digitales en diversos contextos de la sociedad. Sin embargo, también es cierto que muchos estratos sociales no tienen acceso a este tipo de tecnología, por lo tanto, las bondades del paradigma de la datificación solo son perceptibles para algunos individuos y desapercibidas para otros.

Esta paradoja en los datos ha propiciado en los contextos favorecidos —como el académico, científico e industrial—, la necesidad de establecer sistemas de mayor capacidad analítica para llevar a cabo tareas de descubrimiento intensivo basado en datos. Se trata de desarrollar la posibilidad de tomar decisiones basadas en la gran cantidad de datos

existentes en la actualidad. Para ello, se han adoptado diferentes técnicas, metodologías y algoritmos que tienen el propósito de recabar datos relevantes para la toma de decisiones.

El desafío de analizar cantidades masivas de datos es afrontar abundancia, exhaustividad y variedad, actualidad y dinamismo, desorden e incertidumbre, alta relación, y el hecho de que gran parte de lo que se genera no tiene una pregunta específica en mente o es el subproducto de otra actividad.²

Por lo tanto, los datos son una entidad con características muy particulares y comportamientos que se asemejan a los fenómenos y problemáticas que acontecen en la realidad. Por este motivo, su análisis, recolección e interpretación requiere de habilidades y herramientas que faciliten el desarrollo de estas tareas.

Bajo esta premisa, las herramientas digitales cada día se van adaptando a las características y comportamientos de los datos presentes en diferentes dominios, dando la pauta a la generación de nuevas maneras de obtener información, basadas en las demandas y necesidades informativas de los usuarios.

Aunado a ello, los flujos de trabajo científicos son herramientas clave en muchas áreas de investigación que se basan en operaciones múltiples, diversas y distribuidas sobre varios conjuntos de datos, lo que por lo regular genera una gran complejidad computacional y dependencias de datos. Hoy en día, la creciente cantidad de datos de entrada, datos intermedios e incluso datos de salida relacionados con la ejecución de flujos de trabajo está desplazando estos sistemas, computacionalmente intensivos hacia problemas paralelos intensivos en datos.³

En este contexto, la recomendación de información se ha convertido en una constante sobre todo en sistemas comerciales que buscan cubrir una necesidad de ocio y entretenimiento entre sus usuarios, sin embargo,

2 Rob Kitchin, "Big Data, new epistemologies and paradigm shifts". *Big Data & Society* 1, n.º 12 (2014): 2, <https://doi.org/10.1177/2053951714528481>.

3 Silvina Caíno-Lores *et al.*, "Applying Big Data Paradigms to a Large Scale Scientific Workflow: Lessons Learned and Future Directions", *Future Generation Computer Systems* 110, (2020): 440, <https://doi.org/10.1016/j.future.2018.04.014>.

en los ambientes como el académico y el científico esta modalidad se va convirtiendo cada vez más en una constante para recuperar datos e información.

Este fenómeno deja apreciar la generación de modelos para recomendar información basados en el procesamiento y experimentación de los datos, los cuales buscan establecer una relación entre los usuarios y sus preferencias para utilizar una determinada información, dando con ello lugar a la generación de métodos integrales relacionados con la atención a una demanda informativa.

PRINCIPIOS DE LA RECOMENDACIÓN DE INFORMACIÓN

Las necesidades del usuario expresadas en el ambiente digital pueden ser ampliamente diversas. La mayoría de estas necesidades dejan ver un amplio grado de interés y afinidad por los contenidos y recursos informativos que dicho usuario necesita para cubrir su demanda. La recomendación de información es un principio que se aplica en algunos de los sistemas de información que se utilizan en el contexto industrial y comercial, sin embargo, su uso es cada vez más común dentro de los entornos académico y científico.

Hoy en día, el proceso de recomendación es extremadamente común y se aplica en una gran variedad de herramientas. Quizás uno de los tipos más populares son los sistemas de recomendación de películas que utilizan empresas como Netflix y los recomendantes de música en Pandora o Spotify, así como cualquier tipo de recomendación de producto de Amazon.com.⁴

Sin embargo, la recomendación de información es un aspecto que poco a poco se ve más aterrizado en ambientes sistematizados. En este con-

⁴ Laura Igual y Santi Seguí, "Recommender Systems", en *Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications. Undergraduate Topics in Computer Science*, editado por Laura Igual y Santi Seguí (Cham: Springer International Publishing, 2017), 165, https://doi.org/10.1007/978-3-319-50017-1_9.

texto, la recomendación forma parte de un proceso que es aplicado en una herramienta computacional mediante el establecimiento de un conjunto de criterios y valoraciones sobre los datos que los usuarios de la herramienta utilizan para interactuar con ella.

La recomendación de información parte del supuesto de que los datos representan las preferencias y atributos del uso de información por parte del usuario. Tomando en cuenta estas valoraciones, la herramienta efectúa un filtrado de información que proporciona una serie de resultados compatibles con los usos de información que manifiesta dicho usuario a través de la herramienta.

Por lo tanto, el concepto de recomendación de información está íntimamente ligado al perfil informativo del usuario. Ya que, a través del análisis de las características del perfil, se ofrecen resultados que intentan responder a una pregunta o bien resolver una demanda informativa, “pues las necesidades de información de un individuo son aquellos conjuntos de datos que éste necesita poseer para cubrir un objetivo determinado”.⁵

Estas necesidades se trasformarán acorde con el cumplimiento del objetivo que el usuario desea cubrir. Por ende, la recomendación de información es un proceso que facilita elementos informativos al usuario, como datos, contenidos y recursos con información relacionados con su perfil y necesidad informativa. Ahora bien, la recomendación de información requiere contar con datos perfectamente organizados y fidedignos, acordes a los atributos que representan a la información que forma parte de un contexto determinado, pues esto asegurará que la información tomada en cuenta por el usuario tenga un alto grado de confiabilidad al momento de utilizarla y de establecer un vínculo con la respuesta a su demanda informativa. Debido a ello, la organización de los datos y de la información es un asunto trascendental, toda vez que permite representar de manera uniforme la información que caracteriza los recursos y contenidos que un usuario puede consultar a través de un determinado sistema.

5 Patricia Hernández Salazar, “El perfil del usuario de información”, *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 7, n.º 15 (1993): 17.

De esta manera, las características de los recursos de información pueden provenir de su propio contenido (enfoque basado en el contenido) o del entorno social del usuario en donde interactúa con información de un tipo y tema en particular. El uso de la recomendación de información se está volviendo cada vez más popular en Internet porque es un proceso útil para evaluar y filtrar la gran cantidad de información disponible en la web para ayudar a los usuarios en sus procesos de búsqueda y recuperación.

El crecimiento explosivo en la cantidad de información digital disponible y el número de visitantes de Internet han creado un desafío potencial de sobrecarga de información que dificulta el acceso oportuno a elementos de interés para los usuarios. Los motores de búsqueda como Google, Bing y Baidu han resuelto parcialmente este problema, pero la priorización y la personalización (donde un sistema asigna el contenido disponible a los intereses y preferencias del usuario) de la información estaba ausente. Esto ha aumentado la demanda de sistemas de recomendación más que nunca.

En la figura 1 puede apreciarse la aplicación del concepto de recomendación de información en un sistema comercial de libros. Este concepto básico permite conocer la manera en que las compañías comerciales y editoriales aplican la recomendación de información, como un método para filtrar las preferencias del usuario en función de la adquisición de un determinado producto. Esta analogía del consumo puede aplicarse también a la recomendación de información para satisfacer una demanda de información especializada relacionada con el contexto académico y científico.

Los grandes consorcios de la información científica y académica se han percatado del enorme potencial que la recomendación de información tiene para fomentar el interés del usuario por productos sobre los que no tenía conocimiento de su existencia. A su vez, este proceso ha dado la pauta para la aparición de nuevos comportamientos informativos relacionados con el uso de la información en una determinada plataforma o sistema.

Figura 1. Ejemplo de recomendación de información aplicado a un sistema de comercialización de libros

The screenshot displays the Amazon Audible product page for the audiobook 'Billy Summers' by Stephen King. The page includes the following elements:

- Header:** Amazon México logo, search bar, and navigation links like 'Inicio', 'Ofertas del Día', 'Servicio al Cliente', 'Listas', 'Tarjetas de Regalo', and 'Vender'.
- Product Title:** 'Billy Summers' by Stephen King, narrated by Paul Sparks, published by Simon & Schuster Audio. It has 45,368 ratings.
- Price Comparison:**

Formato	Precio	Uso
Kindle	US\$14.88	Leer con nuestra Aplicación gratuita.
Audiolibro	US\$0.00	Gratuito con tu prueba de Audible.
Pasta dura	US\$10.96	354 Usado de US\$2.30 37 Nuevo de US\$9.03 17 Artículo de colección de US\$11.95
Pasta blanda	US\$10.62	22 Usado de US\$10.62 33 Nuevo de US\$13.35
- Acciones:** 'Leer y escuchar' button, 'Agregar a la Lista', and 'Compartir' options.
- Descripción:** 'The number one New York Times Best Seller', 'An Edgar Best Book of the Year', 'A Wall Street Journal Favorite Book of the Year', and 'A Goodreads Choice Awards Finalist'.
- Metadatos:** Duration: 18 hours 57 minutes; Narrator: Stephen King; Release Date: 8 Agosto 2021; Language: Inglés; Publisher: Simon & Schuster Audio.
- Recomendaciones:** 'Las personas que vieron esto también vieron' section showing related books like 'The Institute', 'Later', 'If It Bleeds', 'The Outsider', 'Sleeping Beauties: A Novel', 'Fairy Tale', and 'Duma Key'.

Fuente: Amazon.⁶

En la siguiente figura se puede apreciar la recomendación de artículos científicos relacionados con la búsqueda y selección de un artículo en específico (vease figura 2). En ocasiones, se ha observado que esta recomendación no responde a ningún interés por parte del usuario y los artículos recomendados, solo se muestran a través de un proceso basado en un *ranking* de consulta relacionado con el conteo de visitas que se realiza a un ítem determinado.

6 Stephen King, *Billy Summers*, Amazon, https://www.amazon.com/-/es/dp/B08V9HZGJP/ref=s9_acsd_obs_hd_bw_bH_c2_x_0_i?pf_rd_m=ATVPDKIKX0DER&pf_rd_s=merchandise-search-11&pf_rd_r=90GGQNNNG51JVANVC375X&pf_rd_t=101&pf_rd_p=c0d78239-d54e-5fea-8a84-f05121de61e3&pf_rd_i=17.

Figura 2. Recomendación de información científica, aplicada en el contexto de un sistema para recuperación de información

The screenshot shows the ScienceDirect interface for the article "Information retrieval models in Neural Networks Framework: A survey" by Komala Anamalamudi, Y.C.A. Padmanabha Reddy. The page includes a navigation bar with "Sign in", "Register", and "Journals & Books" options. The article title is prominently displayed, along with the journal name "materials today: PROCEEDINGS" and the date "Available online 10 May 2021". The abstract is visible, starting with "Information retrieval using neural networks both as shallow or deep neural networks is an emerging field. This paper presents a survey on neural network models used at various stages of information retrieval. Information retrieval models used at various stages of information retrieval. Information retrieval models used at various stages of information retrieval." The page also features a "Recommended articles" section with links to related papers like "GR-RNK: Global-context residual recurrent Pattern Recognition, Volume 117, 2021, Article..." and "Performance analysis of all-optical logic...".

Fuente: ScienceDirect.⁷

7 Komala Anamalamudi e YCA Padmanabha Reddy, "Information retrieval models in Neural Networks Framework: A survey", *Materials Today: Proceedings*, (2021), <https://www.sciencedirect.com/science/article/pii/S221478532102174X>.

De esta manera, la recomendación de un artículo se desarrolla a través de un proceso de similitud entre las palabras que representan el artículo que el usuario ha consultado de manera específica, en concordancia con las palabras que incluyen los artículos recomendados. Este proceso no asegura que la información recomendada sea del agrado e interés para el usuario. Por lo tanto, es necesario implementar técnicas semánticas para recomendar recursos de información con un amplio grado de especificidad. En este sentido, se vislumbra el uso de sistemas de recomendación basados en datos abiertos enlazados.

SISTEMAS DE RECOMENDACIÓN BASADOS EN DATOS ABIERTOS ENLAZADOS

Los sistemas de recomendación engloban a una serie de técnicas computacionales que intentan presentar información sobre productos que probablemente sean de interés para un usuario. Estas técnicas se utilizan principalmente en el comercio electrónico con el fin de proporcionar sugerencias sobre los artículos que, presumiblemente, serán del agrado de un cliente. Sin embargo, existen otras aplicaciones que hacen uso de esta modalidad de sistemas, como las redes sociales, la construcción de comunidades y la generación de procesos informativos, entre otros.⁸

Los SR tienen atributos particulares que los diferencian de cualquier otro sistema para recuperar información, quizás el elemento característico de más realce de este sistema es la capacidad que tiene el usuario para interactuar con los resultados obtenidos a partir de una búsqueda determinada. En este sentido, las interfaces de los SR son cruciales para que el usuario pueda expresar su demanda a manera de consulta.

De acuerdo con Varga⁹ existen tres interfaces principales que pertenecen a los sistemas de recomendación:

8 Luis Fernando Terán Tamayo, "Recommender Systems", en *SmartParticipation: A Fuzzy-Based Recommender System for Political Community-Building. Fuzzy Management Methods*, editado por Luis Fernando Terán Tamayo (Cham: Springer International Publishing, 2014), 21.

9 Ervin Varga, "Recommender Systems", en *Practical Data Science with Python 3: Synthesizing Actionable Insights from Data*, editado por Ervin Varga (Berkeley, CA: Apress, 2019), 321. https://doi.org/10.1007/978-1-4842-4859-1_8.

- Interfaz de filtrado: intenta eliminar temas poco interesantes (según el perfil de usuario) de una transmisión de datos fuente, como noticias, *tweets*, correos electrónicos, etc. Por ejemplo, un motor de sugerencias de noticias personalizado solo puede seleccionar artículos que podrían ser importantes para el usuario.
- Interfaz de recomendación: presenta una selección de elementos que puede gustarle al usuario. Por lo general, los SR modernos son sistemas híbridos que aprovechan diferentes algoritmos de recomendación para producir una mezcla de contenidos. Por ejemplo, pueden ofrecer tanto artículos que coincidan con el gusto del usuario como sugerencias de artículos actualmente populares. Además, el sistema puede que utilizar únicamente su propia recomendación inicial para desencadenar una interacción basada en el diálogo, donde la retroalimentación adicional de un usuario puede ayudar a reducir la selección final.
- Interfaz de predicción: intenta prever con precisión cuánto valora el usuario algunos productos (por ejemplo, la cantidad de estrellas que le asignaría a una película que no ha visto). Por supuesto, estos datos son estimaciones y a veces puede resultar el despliegue inexacto de cantidades o valores.

Aunado a ello, el estudio de los SR ha permitido vislumbrar algunas de las problemáticas que contienen en función de los datos e información que pueden proporcionar de una manera precisa y acorde a los requerimientos del usuario, en este sentido Aggarwal¹⁰ y Terán Tamayo¹¹ resumen los problemas de la siguiente manera:

- Problema de predicción: el valor de la predicción de una combinación de usuario-recurso permite obtener una estimación o pronóstico de la información que el usuario puede consultar

10 Charu C. Aggarwal, "An Introduction to Recommender Systems", en *Recommender Systems: The Textbook*, editado por Charu C. Aggarwal (Cham: Springer International Publishing, 2016), 1-28. https://doi.org/10.1007/978-3-319-29659-3_1.

11 Terán, "Recommender Systems", 22.

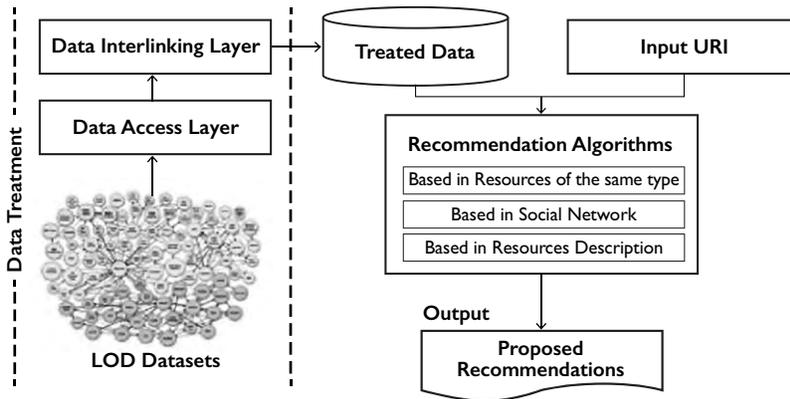
para satisfacer su demanda. Para ello, el algoritmo toma como referencia la huella de comportamiento dejada por el usuario al momento de interactuar con una determinada consulta de información. Sin embargo, en ocasiones estas predicciones poseen un sesgo derivado de la disparidad de los resultados obtenidos en diferentes consultas.

- Problema de calidad en las recomendaciones: la información recibida de un SR debe ser confiable; por esa razón, estos sistemas deben minimizar el número de falsos resultados positivos, es decir, los productos que no le gustan al cliente. Esto permitirá obtener una dispersión menor en el número de recomendaciones

Las problemáticas anteriores han dado la pauta para establecer parámetros basados en el uso de datos abiertos enlazados, los cuales aportan un mayor significado a las relaciones que se establecen entre los datos, los recursos y contenidos que se representan en el universo de la información. De esta manera, los datos abiertos enlazados pueden funcionar como una fuente que permita recomendar piezas de información acorde a los atributos manifestados por el usuario al momento de interactuar con un determinado contenido o recurso de información.

En la figura 3 se puede apreciar una representación que ejemplifica el uso de DAE en un sistema de recomendación. La fuente principal de este sistema es la nube de datos abiertos enlazados (<https://lod-cloud.net/>) mediante la aplicación de algoritmos basados en el tipo específico de recursos; en el uso de redes sociales y en la descripción de dichos recursos se pueden establecer patrones de relaciones entre los datos que se incluyen en los recursos y contenidos disponibles en una determinada fuente, esto permitirá al SR establecer vinculaciones de mayor significado entre los datos que son utilizados para representar los contenidos y recursos, acorde a las características y el perfil de los usuarios que hacen uso de ellos.

Figura 3. Arquitectura de un sistema de recomendación basado en datos abiertos enlazados



Fuente: Oliveira, Delgado y Assaife (2017).¹²

Los datos abiertos enlazados representan una fuente de conocimiento que ha sido generada como parte de la interacción del usuario con diferentes recursos y contenidos. Es preciso recordar que estos datos forman parte de los principios de la web semántica y permiten establecer enlaces de significado entre datos de una manera interoperable a través del uso de normas y estándares desarrollados por el World Wide Web Consortium.

De acuerdo con Wood y Zaidman¹³ los datos abiertos enlazados se basan en los siguientes principios que permiten comprender su aplicación y funcionalidad en los sistemas de información:

1. Los datos están disponibles en la web, en cualquier formato.
2. Los datos están disponibles como datos estructurados legibles por máquina.
3. Los datos están disponibles en un formato no propietario.
4. Los datos se publican utilizando estándares de datos abiertos del World Wide Web Consortium. Por ejemplo, RDF y SPARQL.

12 Jonice Oliveira, Carla Delgado y Ana Carolina Assaife, "A Recommendation Approach for Consuming Linked Open Data", *Expert Systems with Applications* 72, (abril 2017): 409.

13 David Wood et al., *Linked Data: Structured Data on the Web* (Shelter Island, NY: Manning, 2014), 4.

Mediante el uso de la norma para modelar datos RDF y el lenguaje de consulta SPARQL los datos abiertos enlazados permiten crear enlaces entre las preferencias del usuario y un determinado grupo de datos, fomentando la posibilidad de consultar enlaces de significado entre recursos y contenidos disponibles en diferentes fuentes del ambiente digital.

Bajo esta premisa, si los datos de diferentes fuentes son tomados para sugerir las preferencias del usuario respecto al uso de un determinado contenido o recursos, entonces se fomentará en mayor medida la reutilización de los datos y en consecuencia el conocimiento generado por los usuarios formará parte de un ciclo caracterizado por la disponibilidad de los datos y la información que los rodean en un contexto digital común. Este contexto puede ser el sistema de recomendaciones que permitirá reunir diversos perfiles, preferencias y comportamientos para relacionarlos con un recurso o contenido en particular.

Ahora bien, el uso de ontologías dentro de los sistemas también será una etapa trascendental para organizar de una manera uniforme los patrones de comportamiento del usuario con un recurso y contenido específico. De acuerdo con Werner¹⁴ un sistema de recomendación basado en contenido realiza la subtarea de indexación textual de sus elementos, en este escenario, las ontologías son fundamentales ya que se utilizan para representar y gestionar los vocabularios controlados, para describir perfiles y artículos, y finalmente para multclasificar automáticamente a través del proceso de inferencia.

Esto permite generar un modelo de predicción para sustentar la clasificación automática que sirve de base para recomendar los recursos de información dentro del SR. Por lo tanto, el uso de las ontologías y de los datos abiertos enlazados en conjunto permitirá establecer vínculos de mayor significado para la representación de los recursos y contenidos que pueden formar parte de las preferencias del usuario.

14 David Werner *et al.*, "Semantic-Based Recommender System with Human Feeling Relevance Measure", en *Intelligent Systems in Science and Information 2014. Studies in Computational Intelligence* 591, editado por Kohei Arai, Supriya Kapoor y Rahul Bhatia (Cham: Springer, 2015), 178, https://doi.org/10.1007/978-3-319-14654-6_11.

CONSIDERACIONES FINALES

En años recientes y debido principalmente a la sobre carga de información que se tiene en la web han proliferado los sistemas de recomendación, los cuales proporcionan a los usuarios información, productos y servicios que puedan ser del interés del usuario, tras realizar un “estudio” de su perfil, sus gustos e incluso de la forma en la que el usuario navega por la propia web. Se estima que, con mayor frecuencia, estos sistemas puedan desarrollarse en entornos donde los datos y la información son elementos cruciales para la generación de conocimientos.

Se estima que el uso de DAE en los sistemas de recomendación permita obtener resultados de búsqueda con un mayor grado de especificidad y cercanos a las características del perfil informativo del usuario. Esto se lograría a través de una metodología que fuera implementada en la generación del SR y mediante el uso de tecnologías semánticas aplicadas a los datos que describen y representan a los contenidos y recursos.

El procesamiento de datos es una etapa metodológica que proporciona los elementos necesarios para conocer los atributos informativos de los contenidos y recursos informativos que se almacenan y organizan dentro del SR. Este procesamiento permite segmentar semánticamente los datos, mediante el uso de los principios de DAE, aunado a ello, resulta preciso contar con datos que permitan interpretar el comportamiento de los usuarios, para de esta manera, relacionarlos con las preferencias de uso de un determinado recurso o contenido.

Bajo esta premisa, es necesario tener en cuenta los datos que forman parte del perfil del usuario, esto permitirá ofrecer colecciones y recursos de información con un mayor grado de especificidad y, sobre todo, atender de manera oportuna las complejas demandas informativas que se manifiestan en las comunidades.

La generación de SR para atender las demandas de los usuarios de la información es un fenómeno que paulatinamente se va aterrizando en los diferentes contextos informativos. En las herramientas para acceder a la información, como bases de datos especializadas y sistemas de

entretenimiento, se implementan aplicaciones de estas características, por lo tanto, será crucial estudiar y analizar su efectividad para atender la demanda informativa del usuario y considerarlo como parte de las variables que los estudios de usuarios remotos deben contemplar.

BIBLIOGRAFÍA

Aggarwal, Charu C. “An Introduction to Recommender Systems”. En *Recommender Systems: The Textbook*, editado por Charu C. Aggarwal, 1-28. Cham: Springer International Publishing, 2016. https://doi.org/10.1007/978-3-319-29659-3_1.

Caíno-Lores, Silvina, Andrei Lapin, Jesús Carretero y Petter Kropf. “Applying Big Data Paradigms to a Large Scale Scientific Workflow: Lessons Learned and Future Directions”. *Future Generation Computer Systems* 110 (2020): 440-52. <https://doi.org/10.1016/j.future.2018.04.014>.

Goble, Carol y David de Roure. “El impacto de las herramientas de flujo de trabajo en la investigación centrada en datos”. En *El cuarto paradigma: descubrimiento científico intensivo en datos*, editado por Tony Hey, Stewart Tansley y Kristin Tolle, 149. México: Universidad Autónoma Metropolitana, 2014.

Hernández Salazar, Patricia. “El perfil del usuario de información”. *Investigación Bibliotecológica: archivonomía, bibliotecología e información* 7, n.º 15 (1993): 16-22. <https://doi.org/10.22201/ii-bi.0187358xp.1993.15.3816>.

- Igual, Laura y Santi Seguí. “Recommender Systems”. En *Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications. Undergraduate Topics in Computer Science*, editado por Laura Igual y Santi Seguí, con la colaboración de Jordi Vitrà, Eloi Puertas, Petia Radeva, Oriol Pujol, Sergio Escalera, Francesc Dantí y Lluís Garrido, 165-79. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-50017-1_9.
- Kitchin, Rob. “Big Data, new epistemologies and paradigm shifts”. *Big Data & Society* 1, n.º 12 (1 de abril de 2014): 1-12. <https://doi.org/10.1177/2053951714528481>.
- Oliveira, Jonice, Carla Delgado y Ana Carolina Asaife. “A Recommendation Approach for Consuming Linked Open Data”. *Expert Systems with Applications* 72 (abril 2017): 407-20. <https://doi.org/10.1016/j.eswa.2016.10.037>.
- Terán Tamayo, Luis Fernando. “Recommender Systems”. En *SmartParticipation: A Fuzzy-Based Recommender System for Political Community-Building, Fuzzy Management Methods*, editado por Luis Fernando Terán Tamayo, 21-32. Cham: Springer International Publishing, 2014. https://doi.org/10.1007/978-3-319-06551-9_3.
- Varga, Ervin. “Recommender Systems”. En *Practical Data Science with Python 3: Synthesizing Actionable Insights from Data*, editado por Ervin Varga, 317-39. Berkeley, CA: Apress, 2019. https://doi.org/10.1007/978-1-4842-4859-1_8.

Werner, David, Thomas Hassan, Aurélie Bertaux, Christophe Cruz y Nuno Silva. “Semantic-Based Recommender System with Human Feeling Relevance Measure”. En *Intelligent System in Science and Information 2014, Studies in Computational Intelligence* 591, editado por Kohei Arai, Supriya Kapoor y Rahul Bhatia, 177-91. Cham: Springer International Publishing, 2015. https://doi.org/10.1007/978-3-319-14654-6_11.

Wood, David, Marsha Zaidman, Luke Ruth y Michael Hausenblas. *Linked Data: Structured Data on the Web*. Shelter Island, New York: Manning, 2014.

Desafíos en el entorno de la información y la documentación ante las problemáticas sociales actuales, Volumen 1.

Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. La edición consta de 100 ejemplares. Coordinación editorial, Anabel Olivares Chávez; revisión especializada, Angélica Valenzuela, Blanca Vilchis Flores; revisión de pruebas, Anabel Olivares Chávez y Carlos Ceballos Sosa; formación editorial, Oscar Fernando Arcos Casañas. Fue impreso en papel cultural de 90 g en los talleres de Gráfica Premier, 5 de febrero n.º 2309, San Jerónimo Chicahualco, C. P. 52170, Metepec, Estado de México. Se terminó de imprimir en septiembre 2022.