

EL MANEJO DE DATOS

Aproximación desde los estudios
de la información

Georgina Araceli Torres Vargas



Z666.73
L56M3

El manejo de datos. Aproximación desde los estudios de la información / Coordinadora Georgina Araceli Torres Vargas. - México: UNAM. Instituto de Investigaciones Bibliotecológicas y de la Información, 2019.

viii, 116 pp. - Colección: TECNOLOGÍAS DE LA INFORMACIÓN.

ISBN: 978-607-30-2690-1

1. Datos vinculados. 2. Minería de datos. 3. Investigación bibliotecológica.

I. Torres Vargas, Georgina Araceli, coordinadora. II. Ser.

Diseño de portada: Natalia Cristel Gómez Cabral

Primera edición, 2020

D.R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, 04510, Ciudad de México

Impreso y hecho en México

ISBN: 978-607-30-2690-1

Publicación dictaminada

2020

Contenido

Presentación.....	7
GEORGINA ARACELI TORRES VARGAS	

MINERÍA DE TEXTO Y MINERÍA DE DATOS

Identificación de los temas de investigación en los documentos científicos del Colegio de Postgraduados.	11
ÁNGEL BRAVO VINAJA	
SANTIAGO DE JESÚS MÉNDEZ GALLEGOS	
JORGE PALACIO NUÑEZ	

Minería de texto aplicada a un diagnóstico de usuarios en Ciencia y Tecnología: aprendizajes para fortalecer la investigación bibliotecológica.	31
LOURDES FERIA BASURTO	

Minería de Datos, el caso de estudio de la Biblioteca Dr. Valentín Gómez Farías de la Facultad de Medicina de la UNAM.	43
DAVID FLORES MACÍAS	
GUADALUPE VANESA CAROLINA GUTIÉRREZ HERNÁNDEZ	

SISTEMATIZACIÓN DE DATOS Y SERVICIOS DE INFORMACIÓN

Research Data Management and Libraries: Opportunities and Challenges.....	59
KRYSZYNA K. MATUSIAK	

Integración de los principios de <i>linked data</i> en el registro bibliográfico.....	75
---	----

EDER ÁVILA BARRIENTOS

Plan para el Desarrollo de la Ciencia de Datos y Big Data (PDCDBD) en la UNAM para fines académicos y administrativos.....	95
--	----

JAVIER SALAZAR ARGONZA

Presentación

El término *dato* es impreciso; en algunos casos se refiere a la fuente primaria para respaldar una investigación, es decir a la evidencia para validar los resultados de investigación (Johnston, Lisa R. 2017, 2). Sin embargo existe una variedad de datos además de los que derivan de la investigación; el dato puede ser experimental, observacional, operacional, datos de un tercero, del sector público, datos de monitoreo, datos procesados o datos reutilizados (Austin, Claire C. 2016).

Tras la creciente proliferación de dispositivos móviles, transitan grandes cantidades de datos de diversa naturaleza a través de Internet. La coexistencia de esta heterogeneidad de datos es uno de los principales desafíos al momento de su manejo, por lo que surge una amplia diversidad de procesos para su análisis y sistematización, que va desde algoritmos genéticos, procesamiento del lenguaje, aprendizaje automático, redes neuronales, modelos predictivos, análisis de redes sociales, visualización de datos y minería de datos, por mencionar sólo algunos.

Desde los estudios de la información se ha vuelto necesario abordar cómo aprovechar las tecnologías y métodos que existen para efectuar el análisis de datos, con el fin de derivar servicios y productos de información acordes con los requerimientos que se tienen en el ámbito de la investigación, de la empresa, o de cualquier otro ámbito.

Frente a la amplitud de temas que circundan el estudio de los datos, la presente obra tiene por objetivo ofrecer algunas reflexiones en torno al tema del manejo de datos, que por lo general consta de la obtención de datos, su almacenamiento y su tratamiento.

En este sentido, se presentan tres trabajos relacionados con la minería de datos y de texto, que tienen como objetivo explorar el empleo de métodos para interpretar la información cualitativa, así como del análisis diacrónico de la producción científica.

De igual forma se presenta un capítulo relacionado con la gestión de datos de investigación, tema que surge como una nueva área de análisis y de práctica para los estudiosos de la información.

Otro aspecto es el referente a la adopción de los principios de *linked data* (datos enlazados), en la asignación de metadatos, para representar de forma granular los datos bibliográficos y su interrelación con otros datos en el entorno web.

El manejo de datos también requiere de conocimientos especializados, además de la infraestructura tecnológica. En el capítulo "Plan para el Desarrollo de la Ciencia de Datos y Big Data (PDCDBD) en la UNAM para fines académicos y administrativos" se muestran los antecedentes que motivan el desarrollo de este proyecto, así como la problemática que representa su implementación y el logro de sus alcances.

Como el lector podrá notar, un rasgo esencial de esta obra es que a lo largo de sus capítulos se reflexiona sobre las implicaciones sociales que trae consigo el manejo de los datos, así como los usos que y problemas que pueden derivarse. Las pautas de análisis que se dan para los datos, serán de utilidad para los estudiosos del tema y para quienes desean comenzar a adentrarse en la materia.

Austin, Claire C. "Key components of data publishing: using current best practices to develop a reference model for data publishing". En: *International Journal on Digital Libraries*. Junio 2016. Doi:10.1007/s00799-016-0178-2.

Johnston, Lisa R. "Introduction to data curation". En: *Curation research data. Volume One: practical strategies for your digital repository* / edited by Lisa R. Johnston, 2-24. Chicago, Illinois: Association of College and Research Libraries, 2017.

Georgina Araceli Torres Vargas

Manejo de datos. Una aproximación desde los estudios de la información. La edición consta de 100 ejemplares. Coordinación editorial, Israel Chávez Reséndiz; revisión especializada, Francisco Xavier González y Ortiz; revisión de pruebas, Valeria Guzmán González; formación editorial, Natalia Cristel Gómez Cabral. Instituto de Investigaciones Bibliotecológicas y de la Información / UNAM. Fue impreso en papel cultural de 90 gr. en los talleres de Grupo Fogra. Año de Juárez 223. Col. Granjas San Antonio. Alcaldía Iztapalapa. Ciudad de México. Se terminó de imprimir en febrero de 2020.