



UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

COLEGIO DE BIBLIOTECOLOGÍA Y
ARCHIVOLOGÍA

EL BIBLIOTECARIO DE DATOS: PERFIL PROFESIONAL

T E S I S

QUE PARA OBTENER EL TÍTULO DE
LICENCIADA EN BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN

P R E S E N T A:

ELIZABETH NASHIELLI MÁRQUEZ SÁNCHEZ

ASESORA:

DRA. ADRIANA SUÁREZ SÁNCHEZ

Ciudad de México, 2023





Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Índice

Introducción.....	4
Metodología.....	6
Capítulo 1 ¿Qué son los datos?.....	8
1.1 Definición.....	9
1.2 Propósito.....	12
1.3 Tipología	13
1.4 Ciclo de vida.....	23
1.4.1 Ciclo de vida de los datos de acuerdo con Ricardo César Gonçalves Sant’ Ana	23
1.4.2 Ciclo de vida de los datos de acuerdo con Marcos Pérez González.....	27
1.4.2 Ciclo de vida de los datos de acuerdo con el Plan Nacional de Infraestructura de Datos (PNID).....	29
Capítulo 2 <i>Big Data</i>	32
2.1 Definición.....	33
2.2 Propósito.....	35
2.3 Gestión	38
2.4 <i>Big Data</i> en las bibliotecas	40
2.5 El papel de las bibliotecas en la era del <i>Big Data</i>	42
Capítulo 3 Del bibliotecario tradicional al bibliotecario de datos	44
3.1 El bibliotecario tradicional	45
3.1.1 Definición	45
3.1.2 Funciones.....	48
3.2 El bibliotecario de datos	51
3.2.1 Definición	51
3.2.2 Funciones.....	53
3.2.3 Habilidades	62
3.2.4 Educación	67
3.3 Bibliotecario de datos: perfil profesional en la práctica actual	77
3.3.1 ¿Quién requiere y en dónde se requiere el perfil del bibliotecario de datos?.....	77
3.3.2 Otras formas en las que se les conoce a los bibliotecarios de datos	79
3.3.3 Requerimientos del puesto de bibliotecario de datos	79
3.3.4 Actividades que desempeñará el bibliotecario de datos	84
Conclusiones	85
Bibliografía.....	89

Índice de figuras

Figura 1: Nube de datos 1.....	11
Figura 2: Tipología de los datos.	13
Figura 3: Ciclo de vida de los datos para la ciencia de la información (CVD-CI).	26
Figura 4: Ciclo de vida de los datos.	29
Figura 5: Fases del ciclo de vida de los datos.	31
Figura 6: Habilidades básicas para la gestión de datos.	58
Figura 7: Diferentes nombres del bibliotecario de datos.	79
Figura 8: Nube de datos 2.....	83

Índice de gráficas

Gráfica 1: Instituciones que requieren un bibliotecario de datos.	77
Gráfica 2: Países donde están solicitando bibliotecario de datos.....	78
Gráfica 3: Estudios requeridos para el puesto de bibliotecario de datos.	80
Gráfica 4: Requisitos extra para el puesto de bibliotecario de datos.....	81
Gráfica 5: Habilidades.....	84
Gráfica 6: Actividades del bibliotecario de datos.	85

Índice de tablas

Tabla 1: Funciones que los bibliotecarios de datos desempeñan con menor frecuencia.....	60
Tabla 2: Palabras mencionadas con más frecuencia en anuncios de empleo.	82

Anexos

Anexo 1	96
Anexo 2	97
Anexo 3	99
Anexo 4	100
Anexo 5	102
Anexo 6	103

Introducción

Actualmente mantener datos organizados, depurados y sistematizados tiene un impacto significativo en diversas prácticas corporativas y organizacionales. Conforme los datos adquieren relevancia también aumenta su crecimiento y se fortalecen los procesos para su gestión, que representa una actividad fundamental, especialmente para los responsables del análisis y manejo de los datos.

Los usuarios necesitan disponer de datos organizados que les posibiliten la toma de decisiones informadas, el manejo adecuado de situaciones y la identificación de factores de riesgo o factores controlables que permitan el óptimo desarrollo de estrategias dirigidas a la acción. Para que esta gestión pueda darse de forma óptima, en las últimas décadas se habla de un perfil profesional que ha sido denominado bibliotecario de datos. Actualmente, entre la literatura de la especialidad abundan las publicaciones sobre el bibliotecario de datos; sin embargo, sobre tal perfil profesional persisten algunas preguntas de investigación:

- ¿Qué es un bibliotecario de datos?
- ¿Qué hace un bibliotecario de datos?
- ¿Cuáles son las habilidades de un bibliotecario de datos?
- ¿Qué diferencia a un bibliotecario de datos de un bibliotecario común?

El perfil del bibliotecario de datos, si bien no es tan reciente, es impreciso respecto a su lugar en el mundo laboral, académico y científico, debido a las diversas propuestas que existen respecto a sus funciones y características. Derivado de ello, es necesario que exista más investigación al respecto que permita definir, construir y esquematizar dicho perfil. Considerando lo anterior, el objetivo de la presente tesis es:

- Analizar el perfil profesional del bibliotecario de datos, para establecer un vínculo entre la formación bibliotecológica y el contexto laboral, que ayude al bibliotecario a conocer nuevas habilidades y nuevos perfiles profesionales que le permitirán desenvolverse de manera óptima en el campo de los datos.

En aras de alcanzar el objetivo general establecido, se siguió una metodología exploratorio-descriptiva fundamentada en dos técnicas de investigación: revisión bibliográfica y análisis

de casos mediante la revisión de anuncios laborales en los que se explicita el perfil profesional del bibliotecario de datos:

A partir del estudio realizado, la tesis se estructuró en tres capítulos:

- El primer capítulo trata la parte conceptual sobre los datos y su tipología, esto con el fin de tener un contexto que servirá para configurar el perfil del bibliotecario de datos. En tal sentido, el capítulo brinda un contexto de los datos, el *Big data* y todo el conocimiento que conlleva esta profesión. En la sección se consultaron autores interdisciplinarios para promover el entendimiento de los datos, tanto técnicamente como conceptualmente.
- El segundo capítulo trata el *Big data*, con miras a contextualizar el ambiente en el que se desenvuelve el bibliotecario de datos y que significa para las bibliotecas.
- El tercer capítulo se centra en el bibliotecario de datos, haciendo una comparación entre las funciones del bibliotecario tradicional y el bibliotecario de datos y se presenta el análisis de casos a través de anuncios de empleo, con la finalidad de tener una definición más amplia del perfil del bibliotecario de datos.

La importancia de investigar este tema radica en que los bibliotecólogos conozcan este nuevo perfil de la profesión bibliotecaria, con miras a volverse líderes del tema en un futuro cercano. Aunado a ello, se busca la reflexión sobre un nuevo campo de acción que está emergiendo y que será fundamental en el universo documental de los próximos años.

Metodología

El presente estudio es de carácter exploratorio-descriptivo y consiste en la recopilación de información para establecer generalidades y evidenciar el perfil profesional del bibliotecario de datos.

Entre las técnicas de estudio que se aplicaron están:

- Revisión bibliográfica: reside en el hallazgo de significados, interpretación de palabras, textos y discursos en el contexto del que son parte. Su propósito en el estudio es analizar desde los textos la parte teórica que conforma la literatura sobre el perfil profesional del bibliotecario de datos, considerando su definición, funciones, habilidades, educación, entre otros aspectos.
- Análisis de casos mediante la revisión de anuncios laborales, en los que se explicita el perfil profesional del bibliotecario de datos. Se recopiló un conjunto de anuncios obtenidos en varias plataformas de búsqueda de empleo. En la obtención de los casos se consideraron varias etapas:
 - La primera etapa de la búsqueda comenzó con la introducción de la interrogación “Empleo” and “Bibliotecario de datos” dentro del navegador Google. Esta primera búsqueda arrojó pocos resultados, derivando solo en anuncios de empleo sobre bibliotecarios. También se buscó con los términos en idioma inglés “Data Librarian” and “Jobs,” esta segunda consulta produjo más resultados.
 - Para la segunda etapa se escribieron de nuevo los términos de búsqueda “Empleo” and “Bibliotecario de datos” y “Data Librarian” and “Jobs,” pero ahora en las plataformas de búsqueda de empleo populares en México: Indeed, Computrabajo, OCCMundial, Portal del empleo, Glassdoor, LinkedIn. Finalmente, se realizó la misma búsqueda en portales de empleo especializados en Biblioteconomía, como el portal de empleo de la Escuela Nacional de Biblioteconomía y Archivonomía

(ENBA), la *American Library Association* (ALA) y el Colegio Nacional de Bibliotecarios (CNB).

A partir de las búsquedas antes señaladas, fue posible recuperar 35 anuncios que solicitaban bibliotecarios de datos. Entre ellos, 19 fueron detectados a partir de la búsqueda en Google y 16 en plataformas de búsqueda de empleo.

Una vez obtenidos los anuncios, se llevó a cabo su análisis y se propuso una categorización para su estudio, fundamentada en cuatro áreas que eran comunes a la mayoría de los casos obtenidos:

1. ¿Dónde y quién requiere el perfil del bibliotecario de datos?
2. ¿De qué otra forma se le conoce al bibliotecario de datos?
3. ¿Cuáles son las habilidades requeridas para ser bibliotecario de datos?
4. ¿Cuáles son las actividades que desempeña el bibliotecario de datos?

A partir de la metodología que se ha descrito, fue posible cubrir tanto la parte teórica como la parte práctica del perfil profesional del bibliotecario de datos. El procesamiento de la información que da sustento a la tesis se llevó a cabo con ayuda del programa NVivo, para el marcado de etiquetas (definición, educación, habilidades del bibliotecario de datos). También se usó el programa Excel para generar los grafos que se muestran en el capítulo tres.

Capítulo 1 ¿Qué son los datos?

1.1 Definición

El dato tiene una existencia tan antigua como la del hombre mismo, su extensa historia abarca desde las primeras civilizaciones, hasta lo que hoy se conoce como la era de los grandes datos, el dato tiene su origen etimológico en la palabra latina «datum», que significa dado, proporcionado. En esencia se refiere al acto de facilitar algo. Lo que permite que el dato de forma estructurada proporcione la información, permitiendo con ello alcanzar el conocimiento de algo. Algunos autores por su parte describen al dato como:

Hechos que describen el mundo. Las observaciones medidas que pueden transmitirse y procesarse son datos. Si opinamos en una encuesta, esas mediciones son datos. (Johnson, 2018, p. 9)

A medida que fue pasando el tiempo y la tecnología comenzó a progresar, este concepto fue evolucionando hasta conceptualizaciones más técnicas, como la siguiente:

Son la unidad mínima de información que contiene en sí un valor, el cual solo se vuelve útil y expresivo en la medida en que se asocia con otros datos para una determinada finalidad. Así, «1945» es un número que puede revelar una aplicación en la medida en que lo asociamos con otros números (por ejemplo, todos los números divisibles por 5), o con otros datos (por ejemplo, fin de la Segunda Guerra Mundial). Cualquier conjunto de datos organizados se transforma en información. (Barité, 2015, p. 57)

Al desarrollar el concepto de dato, y comparar su relación con otros términos como información o conocimiento Johnson, menciona lo siguiente:

Los datos se pueden procesar para convertirlos en información al ponerlos en contexto. Las colecciones de datos resumidos y contextualizados son información. El conocimiento es lo que la gente sabe y es el resultado de sus experiencias. (Johnson, 2018, p. 9)

Los datos son la unidad mínima con valor que, una vez ya compuestos, adquieren otro significado. El *Online Dictionary of Library and Information Science* (ODLIS) los describe como:

Una cosa que se da o se sabe que es fáctica, sobre la cual se basa un argumento o cálculo razonado. Además, una suposición o premisa a partir de la cual se hacen inferencias posteriormente. (Reitz, 2013, p. 193)

Al ser el dato una unidad mínima de información también representa la base de la realidad, como otros términos lo mencionan definiendo a los datos como eventos o hechos materiales que no tienen un significado relacionado, sin necesidad de interpretación u opinión, y no

están asociados a ninguna documentación que pueda revelar su importancia o relevancia. Algunos ejemplos serían el nombre del cliente, el monto de la compra o el número de transacción bancaria que aparece en la factura.

Los datos constituyen una pequeña parte de todo un ecosistema informativo presente en la vida de las personas, y sus actividades. En este orden de ideas resulta interesante lo que menciona Lagoze, respecto a que los datos, pueden tener sentido solo a través de un intérprete:

Los datos solo pueden existir a los ojos del poseedor: el reconocimiento de que una observación, un artefacto o un registro constituyen datos es en sí mismo un acto académico. Esta perspectiva de los datos es reflexiva; algo (por ejemplo, imágenes, texto y hoja de cálculo de Excel, etc.) son datos porque alguien los usa como datos en un contexto específico y, trascendente, transmite las muchas disciplinas, prácticas y epistemologías de la ciencia. (Lagoze, 2014, p. 1)

De igual forma y de acuerdo con Carrión, los datos por sí mismos no pueden describir todo un contexto “Los datos describen únicamente una parte de lo que pasa en la realidad y no proporcionan juicios de valor o interpretaciones, y por lo tanto no son orientativos para la acción. La toma de decisiones se basará en datos, pero estos nunca dirán que hacer” (Carrión, (s.f.), p. 1). A partir de lo que exponen los autores, se puede afirmar que los datos son los cimientos de una estructura en la cual, se podrán arrojar resultados verdaderos, de acuerdo con el contexto que les corresponda “los datos son los valores que toma una variable cuando es medida”. (Dagnino, 2014, p. 43). Ningún dato por sí solo sin estructura ni contexto podría arrojar algún resultado. Dicho de otra forma, un dato puede variar su significado, dependiendo el contexto en el que se encuentre.

Para complementar, algunos documentos como el escrito por el *Migrant y Seasonal Head Start Technical Assistance Center 2006* aluden, a que: “Al examinar de cerca los datos, podemos encontrar patrones para percibir información, y luego la información se puede utilizar para mejorar el conocimiento” (Migrant y Seasonal Head Start Technical Assistance Center, 2006, p. 5). Es preciso recalcar que, a pesar de que muchos autores mencionan que el dato por sí mismo no puede otorgar un contexto o un significado en particular, no deja de tener la capacidad para comunicar, siempre y cuando estos datos tengan una estructura que los defina, ya que por sí mismos carecen de ella “Se definen como símbolos que representan las propiedades de los objetos, los acontecimientos y su entorno. Son el producto de la

observación. Pero no sirven de nada hasta que se encuentran en una forma utilizable (es decir, pertinente) relevante. La diferencia entre datos e información es funcional, no estructural” (Rowley, 2007, p. 166) Los datos son, en su mayoría, un recurso mínimo que cobra sentido al estar dentro de una estructura. Los datos son independientes de una relación (solamente números o palabras) hasta que se relacionan, entonces se convierten en información que puede llevar al conocimiento. De acuerdo con Sanders: “Los datos suelen requerir un contexto para convertirse en conocimiento” (Sanders, 2016, p. 1). Con fines prácticos para la definición conceptual, en esta tesis se entenderá al dato como la mínima parte de información, que contiene en sí misma un valor, sin ningún significado hasta que es aplicado en un contexto o estructura informativa.

Un claro ejemplo de datos son el número de páginas que contiene esta tesis, así como también la edad de los participantes en una encuesta, el nombre de quien escribió esta tesis y también el número de estrellas que existen en el cielo. En la Figura 1, se recopilan varios ejemplos de lo que es un dato, todos estos ejemplos son algo tan cotidiano y simple, que frecuentemente evidencian que en la vida diaria se convive con datos de diversa naturaleza.

Figura 1: Nube de datos 1.

Fuente: Elaboración propia, 2021.



1.2 Propósito

Los datos como unidad básica no describen el mundo, pero tienen dentro de ellos certezas absolutas, pues una vez trabajados y analizados permiten dejar de lado las corazonadas y suposiciones, dando paso a una toma de decisiones informada.

Gracias a los datos se tiene un conocimiento más cercano a la realidad, sobre los fenómenos que rodean la tierra y la sociedad, su alcance puede incluso llegar a proporcionar predicciones sobre algunos acontecimientos, de tal forma que los datos tienen la capacidad de mostrar fenómenos interesantes, siempre y cuando estos sean trabajados e interpretados de la forma correcta:

Las decisiones no se basan exclusivamente en la información proporcionada por los datos; estos deben ser interpretados por los valores para escoger las mejores decisiones que contribuyan al desarrollo de la persona y el “bien común” (Monfort y Villagra, 2018, p. 1)

La importancia de los datos radica en la forma en que, al ser trabajados, llegan a ser una representación de la realidad, las situaciones o las problemáticas que se viven todos los días, esto con el fin, de controlar y optimizar situaciones.

El análisis de los resultados permite identificar los factores controlables que pueden ser cambiados de tal manera que ello comporte una mejora en la calidad y en el coste/efectividad del servicio prestado. (Monfort y Villagra, 2018, p. 8)

Todo cuanto rodea al mundo está constituido por datos, la mayoría de las cosas se pueden medir, pesar, cuantificar o describir. Es por eso que, saber gestionar y usar los datos es de vital importancia para entender los fenómenos en diferentes circunstancias y procesos.

Como se ha venido diciendo, los datos no logran este propósito por sí mismos, sino, a través de una correcta interpretación y análisis que puede evidenciar significados escondidos a simple vista. Un ejemplo de esto se muestra en el cuento indio del elefante y los tres ciegos. Dicha narración relata que en el valle del río Brahmaputra vivían tres sabios ciegos que un día decidieron conocer lo que era un elefante. El primero de todos ellos tocó su costado, el segundo avanzó, con más precaución tocando sus colmillos y por último el tercero se acercó a su trompa. Aunque los sabios tocaron al animal, ninguno pudo ponerse de acuerdo con su compañero respecto a que forma tenía. Todos habían tocado el mismo animal, pero diferentes partes de él. De manera similar sucede con los datos, todos pueden ver datos en distintas fases

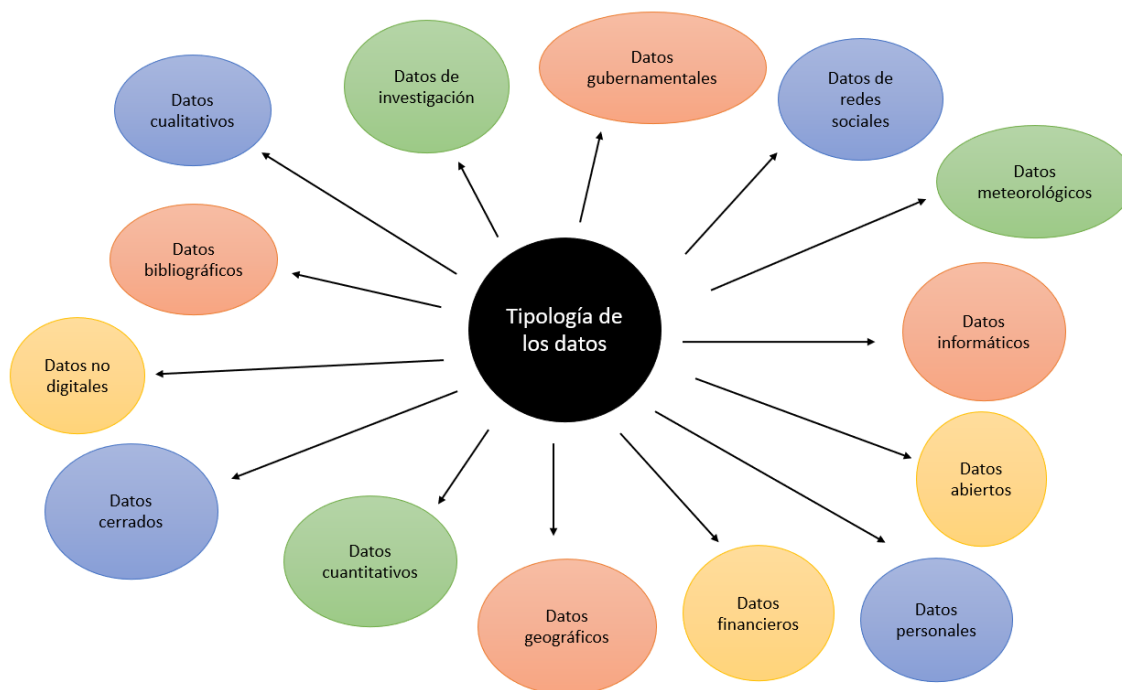
de una situación, contexto o problemática, pero estos, en interacción dentro de varios contextos adquieren significados distintos. De manera similar a las diversas partes del elefante que fueron tocadas por los tres sabios ciegos, los datos son parte de un todo que asoma a la luz de la integridad total.

1.3 Tipología

La sociedad actual, ha generado muchos tipos de datos, volviendo al mundo de los datos más rico y diverso. Hoy en día se cuenta con una tipología de datos amplia, tal como se muestra en la Figura 2. Es importante considerar cuando se habla del trabajo de un bibliotecario de datos, en tanto que estos son su materia prima de interés.

Figura 2: Tipología de los datos.

Fuente: Elaboración propia, 2021.



➤ **Datos cuantitativos**

El dato cuantitativo se considera todo lo que se puede medir, contar o cuantificar, se refiere a la información palpable, la que es obtenida por medio de cualquier procedimiento concreto. Los datos cuantitativos están vinculados con aspectos numéricos y sirven para tomar decisiones.

La siguiente definición indica que los datos cuantitativos:

Se presentan en forma de números, porcentajes y cifras, obtenidos a través del recuento y la medición, por ejemplo, las puntuaciones de un grupo de alumnos en un examen, el orden jerárquico de los mismos de acuerdo con esas puntuaciones, o el número de estudiantes que poseen un atributo común (edad, sexo, clase social). Los datos cuantitativos pueden ser tomados directamente o se pueden desprender de los datos cualitativos. Con el fin de facilitar el recuento y análisis, es posible disponer los datos cuantitativos en matrices, esto es, en forma de tablas de valores. (Palacios et al., 2019, p. 1)

Dentro de las divisiones de estos datos, existe una categorización, dependiendo de sus variables estas son:

Discretas: Este tipo de variables corresponde al conteo de eventos, por lo mismo usa números enteros para caracterizar la magnitud de los datos de las variables en estudio.

Continuas: Corresponde a un valor numérico de espectro real que representa la magnitud del valor al cual representa, por lo mismo puede involucrar valores decimales hasta el infinito si el instrumento de medición es sensible para detectarlos. (Zúñiga y Adasme, 2020, p. 4)

Existen diversos ejemplos de lo que es un dato cuantitativo, y estos van desde la estatura de un individuo, su peso, edad, hasta el número de integrantes de su familia, el sueldo que gana, su presión arterial, entre otros. Al poner una mirada atenta a estos ejemplos, es fácil saber que todos estos son datos concretos, por lo tanto, no existe cabida para la doble interpretación.

➤ **Datos cualitativos**

El dato cualitativo describe como lo dice su nombre, las cualidades, esta clase de información está relacionada con los adjetivos y posee como primordial propiedad que no se pueden medir, ni expresar con números, tienen que ser interpretados, son generalmente descriptivos y más difíciles de examinar ya que, para obtenerlos y examinarlos, se requiere más de tiempo.

El dato cualitativo es un elemento dinámico, resultado de una serie de manipulaciones que transforman la realidad. Cabe destacar, entonces, una serie de

fases, como son el registro de la realidad, su plasmación material en algún tipo de expresión y su transformación mediante un proceso de elaboración conceptual. (Sabiote et al., 2005, p. 137)

Si se analiza una compra y se describe mediante los términos “veloz”, “agradable” o “inútil”, estos adjetivos son datos cualitativos, a semejanza de este ejemplo muchas otras actividades y procesos se caracterizan por elementos cualitativos.

Hay diferentes maneras de obtener datos cualitativos, hay quienes hacen *focus group* para lograrlo, otros se fundamentan con historias, donde se da una averiguación del porqué de sus elecciones, además de analizar otros aspectos. A través de una encuesta en línea, se pueden obtener datos cualitativos indagando aún más en las respuestas abiertas. Una vez obtenidos los datos, se organizan y analizan para lograr interacciones, interpretar la información y sustraer lo importante.

➤ **Datos no digitales**

Los datos no digitales son aquellos que no están siendo analizados, recopilados, almacenados o distribuidos por algún ordenador, celular o máquina. Este tipo de datos no se encuentran en ningún dispositivo electrónico que pueda almacenar datos, llámese celular u ordenador, generalmente este tipo de datos, desde su creación se encuentran en soportes más tradicionales como libros de registro, diarios de investigación o bitácoras, un ejemplo de estos datos son los nombres y edades que se encuentran registrados en un libro de visitas de un museo.

➤ **Datos abiertos**

La era de los grandes datos trajo consigo ciertas ventajas, entre ellas los datos abiertos, la principal característica de estos datos es que son alcanzables, fáciles de consultar y por lo general están disponibles para su uso.

La Carta Internacional de Datos Abiertos define a los datos abiertos como datos digitales que son puestos a disposición del público con las características técnicas y jurídicas necesarias para que puedan ser usados, reutilizados y redistribuidos libremente por cualquier persona, en cualquier momento y en cualquier lugar. (Munte-Kunigami y Serale, 2018, p. 4)

Los datos abiertos hoy en día son un derecho humano, y parte esencial del bienestar colectivo, así mismo son importantes para que el ciudadano se mantenga asertivamente informado.

Son considerados como tales, datos abiertos todos aquellos datos accesibles y reutilizables que no requieren permisos y no presentan exigencias para llegar a ellos. Lo que sucede con los datos abiertos es algo similar a lo que sucede con el software libre, el código abierto, o el acceso libre. Lo que hoy conocemos como Datos Abiertos son en general fuentes de datos que históricamente han sido cerrados y que han estado bajo control de organizaciones -públicas o privadas- y cuyo acceso ha estado restringido mediante diferentes tipos de limitaciones, licencias, derechos de autor y/o patentes. (Bron, 2015, p. 10)

Un ejemplo de lo que pueden llegar a ser los datos abiertos, es la cantidad real de calorías que contienen los productos que se consumen diariamente. Estos datos, a simple vista, podrían parecer sencillos e inofensivos, pero algunas empresas alimenticias por mucho tiempo los negaron y ocultaron, causando a la larga grandes estragos en la salud de la población, así pues:

Los datos abiertos nos deben permitir a todos: *accederlos*, a través de formatos digitales, *utilizarlos* a través de la estructuración en formatos o estándares que faciliten la interoperabilidad, *compartirlos* a través de su certeza jurídica e Integridad. (Munte-Kunigami y Serale, 2018, p. 11)

Este tipo de datos es ahora parte importante de los derechos humanos, gracias a ellos la población puede tomar decisiones informadas que mejoren su calidad de vida.

➤ **Datos cerrados**

Con respecto a estos datos, su nombre deriva de la clasificación según sus restricciones de uso, es decir:

Como datos más restrictivos, nos encontramos con los llamados datos de acceso interno (cerrado), no forman parte del Open Data, son datos privados o personales a los que solamente tienen acceso por ejemplo determinados empleados o miembros de una organización, y cuyo acceso está restringido por contrato y bajo unas condiciones de uso. (Open Data Institute, 2022, p. 1)

Un ejemplo son los informes de ventas de una empresa, o datos a los que solo pueden tener acceso un número limitado de personas bajo contrato.

➤ **Datos de investigación**

De acuerdo con el *Australian National Data Service* (ANDS): “Proporcionar una definición autorizada de los datos de investigación es un desafío, ya que es probable que cualquier definición dependa del contexto en el que se formula la pregunta” (Australian National Data Service, 2017, p. 1). Dado que no existe una definición concisa de lo que son los datos de

investigación debido a que cada institución tendrá una definición acorde con sus necesidades, se tomará en cuenta la definición que varias instituciones le proporcionaron al *Australian National Data Service*, comenzando por la Universidad de Tecnología de Queensland la cual señala lo siguiente:

Los datos de investigación son datos en forma de hechos, observaciones, imágenes, resultados de programas de computadora, grabaciones, mediciones o experiencias en las que se basa un argumento, teoría, prueba o hipótesis, u otro resultado de la investigación. Los datos pueden ser numéricos, descriptivos, visuales o táctiles. Puede ser crudo, limpio o procesado, y puede conservarse en cualquier formato o medio. (Australian National Data Service, 2017, p. 2)

Para la Universidad de Melbourne, los datos de investigación son:

Hechos, observaciones o experiencias en las que se basa un argumento, teoría o prueba. Los datos pueden ser numéricos, descriptivos o visuales. Los datos pueden ser en bruto o analizados, experimentales u observacionales. Los datos incluyen: cuadernos de laboratorio; cuadernos de campo; datos de investigación primarios (incluidos datos de investigación en papel o en formato legible por computadora); cuestionarios; cintas de audio; cintas de video; modelos; fotografías; películas. Las colecciones de investigación pueden incluir diapositivas; artefactos; especímenes; muestras. También se puede incluir información sobre la procedencia de los datos: cómo, cuándo, dónde se recopilaron y con qué (por ejemplo, instrumento). También se puede incluir el código de software utilizado para generar, anotar o analizar los datos. (Australian National Data Service, 2017, p. 2)

Asimismo, los datos de investigación son: “los datos, registros, archivos u otra evidencia, independientemente de su contenido o forma (por ejemplo, en forma impresa, digital, física o de otro tipo), que comprenden observaciones, hallazgos o resultados de la investigación, incluidos materiales primarios y datos analizados” (Australian National Data Service, 2017, p. 2) Es así que, los datos de investigación son todo aquel material que ha sido registrado durante la investigación, reconocido por la comunidad científica y que sirve para certificar los resultados de la investigación que se realiza. En adición a lo anterior: “Además de estas características también se especifica que deben provenir de una fuente única y deben ser difíciles o imposibles de obtener de nuevo por ser propios de un momento o circunstancias irrepetibles de una forma exactamente igual” (Torres et al., 2012, p. 175). En conclusión, los datos de investigación son todos aquellos materiales generados o recolectados durante la investigación. Sin ellos, ninguna investigación puede ser validada o certificar resultados y solo adquieren significado dentro del contexto de dicha investigación. En ellos cae el peso de los resultados, por lo tanto, la importancia de que puedan ser consultados y comprobados,

es importante, no solo para acreditar la investigación, sino también para que futuros investigadores puedan sustentarse a través de ellos y así poder hacer un ecosistema de intercambio informativo más eficiente.

➤ **Datos informáticos**

En informática, los datos son representados como atributos o variables, son la descripción de un código, o una señal que varía de forma continua.

Los datos son, así, la información (valores o referentes) que recibe el computador a través de distintos medios, y que es manipulada mediante el procesamiento de los algoritmos de programación. Su contenido puede ser prácticamente cualquiera: estadísticas, números, descriptores, que por separado no tienen relevancia para los usuarios del sistema, pero que en conjunto pueden ser interpretados para obtener una información completa y específica. (Raffino, 2020, párr. 2)

Entre los ejemplos que se pueden encontrar de este tipo de datos se encuentran los:

- Bits, Es decir, 1 o 0
- Caracteres. Números, letras o caracteres especiales, formados cada uno a partir de una combinación de bits
- Campos. Caracteres agrupados y ordenados (apellidos, palabras)
- Registros. Grupos de campos ordenados (contraseñas)
- Archivos. Conjuntos ordenados de registros (cookies). (Raffino, 2020, párr. 6)

Entre los datos informáticos existe diferentes tipos, dependiendo de cómo son representados:

Caracteres: Dígitos individuales que se pueden representar mediante datos numéricos (0-9), letras (a-z) u otros símbolos.

Caracteres Unicode: Unicode es un estándar de codificación que permite representar más eficazmente los datos, permitiendo así hasta 65535 caracteres diferentes.

Numéricos: Pueden ser números reales o enteros, dependiendo de lo necesario.

Booleanos: Representan valores lógicos. (Raffino, 2020, párr. 6)

En la actualidad los datos en informática son intensamente usados y su estudio es cada vez más necesario.

➤ **Datos geográficos**

Los datos geográficos, también llamados geodatos o datos geoespaciales, de acuerdo con el *Centre for Geographical Information Systems* de la Lund University:

Pueden describirse simplemente como datos de ubicación a los que se han vinculado atributos descriptivos. En formato digital, las características geográficas se pueden representar como una matriz o como coordenadas de puntos, líneas y polígono. (Lund University, Centre for Geographical Information Systems, 2021, párr. 4)

A través de ellos es que, se puede representar la actividad humana en una ubicación geográfica, esto ayuda a entender de mejor manera el impacto de la humanidad sobre el medio ambiente.

Por lo tanto, podemos considerar los datos geográficos como valores, cadenas de caracteres o símbolos que proporcionan a quien los use información sobre la localización geográfica de entidades del mundo real. Es decir, dan respuestas a las preguntas originales que nos habíamos planteado de qué, dónde y cuándo. (Rodríguez y Olivella, 2010, p. 24)

Dentro de los ejemplos de datos geográficos se encuentran, la cantidad de mineros que existen en tal o cual región, cuántas mujeres están embarazadas en cierto estado, entre otras cuestiones. Los datos geográficos presentan tres tipos de componentes:

- Un componente espacial que contiene información asociada sobre la localización.
- Un componente de atributo que contiene información asociada temática.
- Un componente temporal que lleva asociada información del tiempo. (Rodríguez y Olivella, 2010, p. 25)

➤ **Datos financieros**

Los datos financieros son cualquier número o dato personal, utilizado para identificar una cuenta abierta, tarjeta, o instrumento de pago emitido por una institución financiera o por un individuo, estos datos describen la relación con los estados financieros y el historial de crédito. Los datos financieros vienen en muchas formas. Los tipos comunes son series de tiempo, datos de sección transversal y de panel.

Dentro de los tipos de datos financieros se encuentran:

- Los activos, que incluyen todos los bienes.

- Los pasivos, que son las obligaciones financieras (lo que la empresa debe a otros).
- La equidad, que es el valor de la empresa que queda después de que esta ha pagado todos sus pasivos con todos sus activos.

También algunas bases de datos financieros entran dentro de esta categoría, algunos otros ejemplos de datos financieros son: balance general, estado de resultados, estado de cambios en el patrimonio y el estado de cambios en la situación financiera.

➤ **Datos meteorológicos**

Conforme a lo establecido por el *Support Center for Regulatory Atmospheric Modeling* (SCRAM) Los datos meteorológicos son todos aquellos que:

Consisten en parámetros físicos que se miden directamente por instrumentación e incluyen temperatura, punto de rocío, dirección del viento, velocidad del viento, nubosidad, capa (s) de nubes, altura del techo, visibilidad, clima actual y cantidad de precipitación. Estos datos y sus procesadores de datos meteorológicos asociados y programas accesorios suelen ser necesarios para los modelos de dispersión. (Support Center for Regulatory Atmospheric Modeling, 2020, párr. 3)

Esta clase de datos varían con frecuencia, pero normalmente incluyen información sobre la velocidad y dirección del viento, la temperatura y la humedad atmosférica.

➤ **Datos bibliográficos**

Los datos bibliográficos consisten en el estudio, descripción y clasificación de documentos, con el objetivo de compilar fuentes documentales de un trabajo intelectual, o un catálogo de obras sobre un tema específico.

Tradicionalmente, los datos bibliográficos se consideran el contenido de un catálogo de biblioteca, catálogo colectivo o servicio de indexación y resumen, utilizado para conectar a los usuarios de la biblioteca con recursos que cumplirán con su necesidad de información. Sin embargo, la imagen estereotipada de los usuarios de datos bibliográficos está evolucionando junto con el entorno de la información. Además del usuario final de la biblioteca tradicional. (Fallgren, 2007, p. 1)

Algunos ejemplos de datos bibliográficos son: título del libro, nombre del autor, editorial y año de edición. Estos datos son utilizados por el personal de servicios de la biblioteca, los catalogadores, los desarrolladores de metadatos, entre otros actores. Algunos usos más precisos de los datos bibliográficos son:

- El personal de la biblioteca utiliza datos bibliográficos en su trabajo diario para una variedad de propósitos los bibliotecarios interactúan con datos bibliográficos para la investigación y en la creación de guías temáticas.
- Los datos bibliográficos se utilizan en la gestión de colecciones para el control y la colección de inventarios.
- Los catalogadores consultan fuentes externas de bibliografía.
- Los catalogadores pueden navegar o crear informes de los datos bibliográficos en sus propios catálogos con fines de control de calidad. (Fallgren, 2007, p. 1)

A pesar de que los datos bibliográficos, ya se están mudando al ámbito digital, en masa, todavía existen datos albergados de manera física en bibliotecas y centros documentales.

➤ **Datos personales**

De acuerdo con la Ley Federal de Protección de Datos Personales en Posesión de Particulares, estos se describen como “cualquier información concerniente a una persona física identificada o identificable” (Ley Federal de Protección de Datos Personales en Posesión de Particulares, 2010, p. 2). Por lo tanto, los datos personales, son todos aquellos con los que es posible identificar a una persona. Estos datos son únicos y permiten tener un registro irrepetible de un individuo.

Ciertos ejemplos que se encuentran, dentro del ámbito de los datos personales son: nombres y apellidos, domicilio, teléfono, historial laboral y académico, datos patrimoniales y financieros, firma, así como sus características físicas, incluyendo datos biométricos como el iris, o la huella dactilar. Dentro de los datos personales, existe una segunda categoría; llamada datos personales sensibles, los cuales, de acuerdo con la Ley Federal de Protección de Datos Personales en Posesión de Particulares son:

Aquellos datos personales que afecten a la esfera más íntima de su titular, o cuya utilización indebida pueda dar origen a discriminación o conlleve un riesgo grave para este. En particular, se consideran sensibles aquellos que puedan revelar aspectos como origen racial o étnico, estado de salud presente y futuro, información genética, creencias religiosas, filosóficas y morales, afiliación sindical, opiniones políticas, preferencia sexual. (Ley Federal de Protección de Datos Personales en Posesión de Particulares, 2010, p. 2)

➤ **Datos gubernamentales**

Los datos gubernamentales, también llamados datos oficiales, son todos aquellos producidos y recopilados por el gobierno o entidades controladas por el gobierno, muchos de estos datos por lo general no son de acceso abierto, aunque se han llevado a cabo varias políticas para que esto cambie, estos datos son solamente publicados de manera oficial por el gobierno, y son considerados un bien público.

Los datos abiertos gubernamentales que se integran en los sitios web conforman una masa enorme de datos que el Gobierno y sus instituciones producen, gestionan, archivan y distribuyen, los que pueden ser divulgados sin restricciones para uso y distribución en estudios e investigaciones, para la toma de decisiones o para la solución de diversos problemas, a través de una licencia abierta, como Creative Commons. (Quintanilla y Gil-García, 2016, p. 76)

Los datos gubernamentales permiten tener una buena relación ciudadano-gobierno, impulsar el emprendimiento y la innovación, así como también promover y dar seguimiento a programas gubernamentales y rendición de cuentas. Los datos como la estadística, informes o información geográfica pueden mejorar la transparencia, ya que permiten el libre acceso a datos e información que se puede compartir y reutilizar. También pueden crear valor para la sociedad y las empresas, al posibilitar el conocimiento de datos concretos que sean útiles para la realización de alguna actividad de aprendizaje, investigación, innovación o empresa, sobre todo, pueden potenciar la participación y el compromiso de las distintas partes involucradas en los procesos de gobernanza y en toma de decisiones.

Los datos gubernamentales permiten que trabajar con el gobierno sea más sencillo y tener mayor relación con los diferentes actores sociales.

➤ **Datos de redes sociales**

En la actualidad las redes sociales son parte de la cultura diaria. Dentro de todo el espectro, las redes sociales han encontrado la manera de medir y clasificar a la población, a través de gustos y actividades, esta es la razón por la que, los datos de redes sociales, en la actualidad son importantes para muchas empresas.

En resumen, los datos de las redes sociales son la información recopilada de las redes sociales que muestran cómo los usuarios comparten, ven o interactúan con su contenido o perfil. (Barnhart, 2018, párr. 11)

Estos datos, incluso con poco tratamiento, describen aspectos de un sujeto, o incluso una población, algunos ejemplos de estos datos se presentan cuando se comparte una foto con alguna etiqueta, se comenta alguna foto en específico, o se le da “like” a alguna publicación, todos estos datos sirven para saber la opinión de uno o varios grupos en específico, también para tener idea de cuán popular es un tópico a través del uso de etiquetas, pero a pesar de que esto haga parecer a los datos muy imponentes y claros, la realidad es que deben ser analizados para que resulten útiles.

1.4 Ciclo de vida

1.4.1 Ciclo de vida de los datos de acuerdo con Ricardo César Gonçalves Sant’Ana

Para Gonçalves Sant’Ana, la Bibliotecología y Ciencias de la Información, en conjunto con otras disciplinas que se especializan en datos, deben trabajar en equipo para obtener una mayor optimización en datos y para ello es necesario estructurar el análisis del ciclo de vida de los datos. En vista de ello, Gonçalves Sant’Ana propuso utilizar frases o momentos, en los que se hacen presentes diferentes necesidades y competencias del ciclo de vida de los datos, manejando este como una forma de resaltar los diferentes factores que involucra el proceso.

Gonçalves Sant’Ana señala que existen cuatro fases en el ciclo que son: recoger, recuperar, descartar, almacenar, estas representan la base de cualquier ciclo de vida de los datos y son descritas de la siguiente manera:

En la etapa de recoger, se requieren esfuerzos para establecer un plan de acción, que ayude a delimitar el tipo de datos que se recolectarán.

En primer lugar, es necesario obtener los datos que se pueden utilizar para satisfacer una necesidad específica o una demanda esperada de información sobre un contexto determinado. (Gonçalves Sant’Ana, 2016, p. 119)

Posteriormente, en la segunda etapa recuperar, el objetivo se convierte en algo viable. “Así, se inicia una nueva etapa en la que se pueden encontrar, acceder e interpretar las inquietudes y esfuerzos que se dirigen hacia estos datos” (Gonçalves Sant’Ana, 2016, p. 121)

En la etapa descartar se lleva a cabo la depuración del proceso, es aquí donde se escogen los datos que serán de utilidad.

Es posible identificar que datos ya no son necesarios o que deben ser eliminados de la base de datos, lo que lleva a otra fase que se encarga de limpiar o simplemente desactivar parte de los datos. En esta fase, identificada como descartar está la eliminación de parte de los datos. (Gonçalves Sant'Ana, 2016, p. 122)

La última etapa almacenar, se necesitan conocimientos profundos de informática y también expertos en Bibliotecología.

Se procura mantener estos contenidos en un medio determinado. En informática, este proceso denominado persistencia de datos conlleva una serie de inquietudes y aspectos que deben planificarse en detalle. (Gonçalves Sant'Ana, 2016, p. 120)

Junto a los factores anteriormente mencionados como base del ciclo de vida de los datos, existen otros factores que están presentes en todas las demás fases principales, por lo cual pueden considerarse como subetapa la *privacidad, integración, calidad, derechos de autor, difusión y preservación*.

De esta forma podemos hacer explícitas las fases de Recolección, Almacenamiento, Recuperación y Disposición y los factores que están presentes en cada una de estas fases: Privacidad, Integración, Calidad, Copyright, Difusión y Conservación. (Gonçalves Sant'Ana, 2016, p. 119)

A continuación, se dará una breve explicación, de lo que significan las subfases que conforman el ciclo de vida de los datos para Gonçalves Sant'Ana.

La primera subetapa empieza con la privacidad, de esta fase depende, si las demás etapas del ciclo de vida de los datos no están atentando contra los derechos de otras personas o instituciones.

Empezar con Privacidad, en este ámbito adquiere trazos de gran protagonismo. Es necesario identificar, en las fuentes utilizadas, aspectos que puedan constituir una vulneración de la privacidad de las personas o instituciones relacionadas con los datos que se recaben, que pudieran derivar en una responsabilidad futura de la base de datos obtenida, comprometiendo las próximas fases del ciclo de vida. (Gonçalves Sant'Ana, 2016, p. 125)

Una vez superada la primera etapa, se encuentra la integración, que está íntimamente ligada a uno de los momentos base, la etapa de la recolección.

La integración debe ser una preocupación de la fase de recolección a través de la identificación y validación de los atributos que serán responsables de la

identificación unívoca de cada registro (clave candidata o primaria) y sus corresponsales en las otras entidades (claves externas) para que la integración pueda ser garantizado. (Gonçalves Sant'Ana, 2016, p. 125)

En esta etapa es importante que los datos sean de una calidad excelente. Característica esencial de los repositorios de datos es la definición y garantía de elementos que permitan la percepción de la calidad de los datos recolectados, por lo que elementos como el origen, los mecanismos de recolección y las garantías de integridad física y lógica representan solo algunos de los aspectos a considerar. La confiabilidad de los datos es una condición que hace que los datos sean útiles.

Para la cuarta subetapa se muestra que los derechos de autor deben estar presentes en el proceso, esto con el fin de que estén protegidos tanto los responsables, como la información generada a través de estos datos.

Los derechos de autor siempre se deben consultar, la información sobre el derecho de acceso a los datos deseados y sus matices, como la cuestión de los resultados derivados o vinculación para el uso futuro de los resultados producidos a partir de ellos y también la autorización para el cambio y la cita obligatoria de la fuente. El mayor volumen posible de esta información debe estar debidamente registrado en el ámbito del propio repositorio, aumentando su seguridad jurídica y también de los responsables. (Gonçalves Sant'Ana, 2016, p. 126)

Para la quinta subetapa, se menciona la difusión que es también parte de la etapa de almacenamiento. Se espera que los datos almacenados proporcionen acceso en el futuro, entonces la difusión es una de las preocupaciones que también surgen en la fase de almacenamiento. Es necesario proporcionar medios que permitan que estos datos sean accesibles y también incorporar semántica para que puedan ser interpretados, preferiblemente automatizados. Es necesario posibilitar que la base de datos contenga elementos que también permitan y faciliten su localización.

La subetapa de la preservación es para que los datos puedan seguir siendo consultados, utilizados y se pueda acceder a ellos, sin importar los diferentes soportes que manejen.

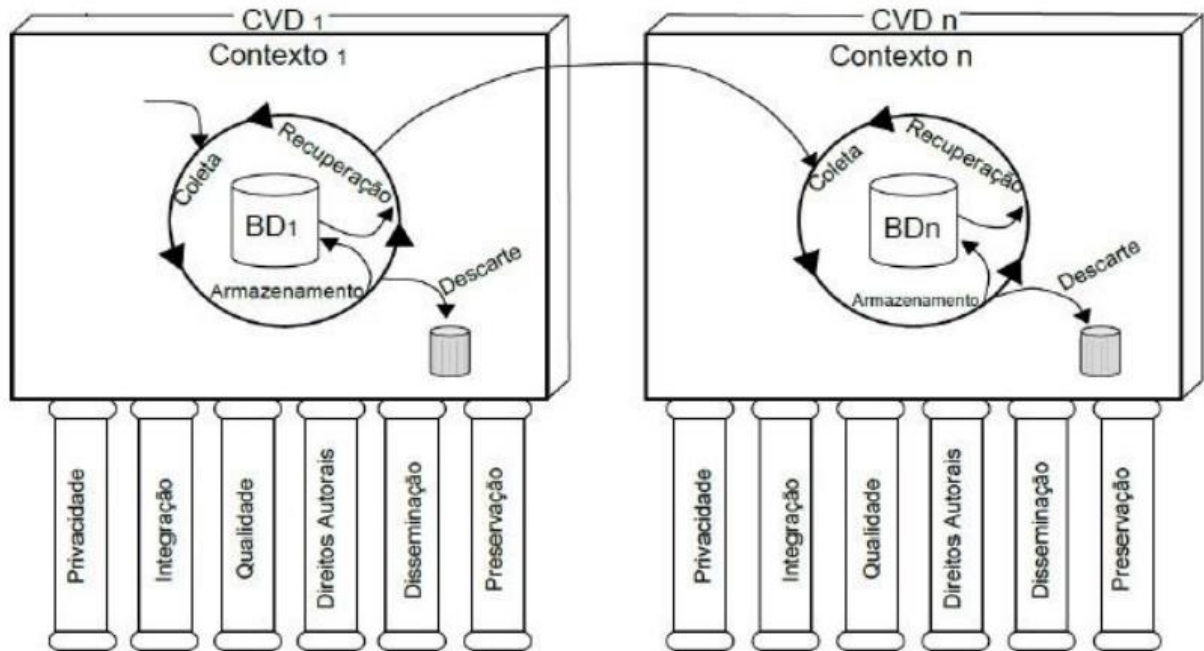
La preservación de los datos recolectados y que eventualmente serán almacenados, también puede requerir que se incluyan datos adicionales en los prerequisites definidos para la recolección, siempre que estos datos puedan ser identificados de manera más amplia e incluso incorporando información sobre posibles características de los dispositivos que han sido utilizados como fuente de los datos, permitiendo que estos datos no solo se conserven sino que también se utilicen, incluso después de cambios inevitables en sus estructuras semánticas y constituciones que surgen de los desarrollos en los dispositivos, como una mayor

exactitud o precisión, con cualquier y muy probablemente, mejoras en los niveles de granularidad de la información. (Gonçalves Sant'Ana, 2016, p. 126)

Con lo anteriormente explicado, el ciclo de vida de los datos quedaría conformado como se muestra en la Figura 3, incluyendo todos los factores que son indispensables para el autor.

Figura 3: Ciclo de vida de los datos para la ciencia de la información (CVD-CI).

Fuente: Gonçalves Sant'Ana, 2016, p. 126.



Cabe destacar que para Gonçalves Sant'Ana tanto las etapas como las subetapas de este ciclo de vida de los datos de investigación, están íntimamente ligadas unas a otras, es decir, aunque las fases principales estén operando, dentro de un ciclo, no pueden de ninguna forma continuar de forma normal, si alguna de las subfases está fallando o no se está llevando a cabo.

1.4.2 Ciclo de vida de los datos de acuerdo con Marcos Pérez González

En el ámbito de las organizaciones, el dato es visto como un activo de valor el cual precisa que se le gestione y se le organice como cualquier otro bien que pudiera llegar a pertenecer a la empresa, en ello radica el ciclo de vida de los datos, para este autor.

Se entiende que el ciclo de vida de los datos:

Se vincula directamente el activo con una actividad o proyecto. Por lo que, en general, el ciclo de vida de un activo puede estar vinculado a las siguientes etapas:

Idea inicial y estudios preliminares (del proyecto).

Evaluación del proyecto respecto a diferentes perspectivas, como viabilidad técnica o económica.

Definición y planeamiento de las etapas del proyecto, y en cuáles de ellas el activo es relevante.

Detalle del proyecto y diseño de los procesos.

Ejecución del proyecto de acuerdo a la planificación.

Puesta en marcha, prueba de todas las instalaciones y aceptación de estas.

Operación de las instalaciones, uso o consumo de los bienes o servicios.

Evaluación del proyecto para identificar posibles mejoras o cambios.

Descarte, reciclaje o venta de la instalación. (Pérez González, 2019, p. 11)

El ciclo de vida de los datos consiste en conocer todo lo que sucede con los datos: desde su creación, integración, proceso, procesamiento final y el reciclaje o la venta de los datos. Entender el ciclo de vida de los datos ayuda a analizar las condiciones de los datos en la organización en relación con el negocio y, por lo tanto, gestionarlo adecuadamente.

Entendemos como gestión del ciclo de vida del dato el enfoque basado en políticas para gestionar el flujo del dato a través de uno o varios sistemas de información a lo largo de su ciclo de vida. (Pérez González, 2019, p. 11)

Para comprender al dato y sus procesos es importante entender su ciclo de vida, a continuación, se describirán sus fases:

Fase 1. Captura del dato. Se crean datos que no existen aún en la organización y que nunca han existido. En esta fase se encuentra la:

1. Captura de datos. Se reciben los datos y se ingresa.

2. Adquisición de datos. La ingesta de datos que han sido creados y existen fuera de la organización.

3. Introducción de datos. La creación de nuevos datos en la organización por personal humano, o generados mediante dispositivos dentro de la propia organización.

4. Recepción de señales. La captura de datos creados por dispositivos, normalmente en sistemas de control, pero cada vez más importantes para los sistemas de información como el Internet de las cosas. (Pérez González, 2019, p. 16)

Fase 2. Mantenimiento del dato. En esta fase se procesan los datos, pero sin que genere un valor para la organización, aquí se maneja la integración, limpieza, enriquecimiento, así como procesos de extracción, transformación y carga.

Fase 3. Síntesis del dato. Aquí se genera la creación de datos de valor aplicando lógica inductiva determinada, usando otros datos para la comparación. En esta fase se usa el modelado de datos, y la lógica inductiva.

Fase 4. Uso del dato. Cuando el dato ha sido capturado y transformado dentro de la organización y se ha usado como fuente, se usa para el beneficio de la propia empresa.

Fase 5. Publicación del dato. En dicha fase se envía el dato, este puede ser interno o externo, una vez el dato ha sido enviado externamente es difícil de recuperar.

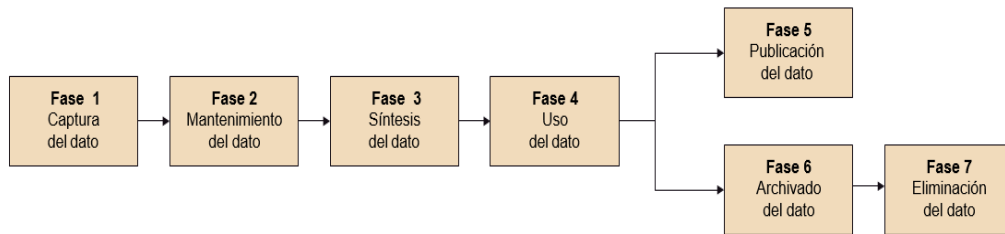
Fase 6. Archivado del dato. Aquí se copian los datos y son almacenados, en caso de que se vuelvan a necesitar en el futuro. Un archivo de datos es simplemente un almacenamiento de datos, pero no de mantenimiento, uso o publicación de datos.

Fase 7. Eliminación del dato. Esta es la fase final del ciclo de vida del dato, y consiste en la destrucción de cualquier copia del dato que exista dentro de la organización. Un reto, en esta fase, sería verificar que la eliminación se ha llevado a cabo de buena manera.

Así es que nace el modelo de ciclo de vida de los datos de acuerdo a las fases anteriormente mencionadas y que se ilustran en la Figura 4.

Figura 4: Ciclo de vida de los datos.

Fuente: Pérez González, 2019, p. 15.



1.4.2 Ciclo de vida de los datos de acuerdo con el Plan Nacional de Infraestructura de Datos (PNID)¹

La Asociación Colombiana de Ingenieros en Sistemas (ACIS) en colaboración con el gobierno colombiano creó el Plan Nacional de Infraestructura de Datos (PNID) con la intención de definir la infraestructura de sus datos y que estos sean aprovechados.

Tal como se describe en la Figura 5, el ciclo de los datos está integrado por varias fases las cuales son:

Crear y obtener: Se pueden obtener datos de diferentes fuentes estructuradas, no estructuradas o fuentes secundarias como las redes sociales. Son diversos los formatos en que los datos pueden aparecer (PDF, Word, Excel, HTML, entre otros). Los datos suelen ser creados por las organizaciones, usualmente de las siguientes formas:

- Adquisición de datos. Adquisición de datos ya existentes que se han producido fuera de la organización.
- Entrada de datos. Entrada manual de nuevos datos por parte del personal dentro de la organización.
- Captura de datos. Captura de datos generados por dispositivos utilizados en varios procesos de la organización. (Asociación Colombiana de Ingenieros en Sistemas, 2021, p. 41)

¹ Con mayor especificidad, en el Anexo 1.2 del Plan Nacional de Infraestructura de Datos.

Procesar. La etapa de procesamiento aporta las siguientes tareas:

Limpieza de datos. En la que un conjunto de datos se limpia y se transforma de su forma sin procesar a algo más accesible y utilizable. Esto también se conoce como corrección de datos.

Compresión de datos. En la que los datos se transforman en un formato que se puede almacenar de manera más eficiente.

Cifrado de datos. En el que los datos se traducen a otra forma de código para protegerlos de problemas de privacidad.

Calidad de los datos. En los procesos de obtención y procesamiento de datos, donde principalmente se aplican las políticas y controles, establecidas en el diseño, con el propósito de determinar cuáles datos son utilizables. (Asociación Colombiana de Ingenieros en Sistemas, 2021, p. 41)

Almacenar. Incluyen el diseño, la implementación y el soporte de los datos almacenados para incrementar su valor, utilizando repositorios que consigan el equilibrio entre disponibilidad y coste de almacenamiento.

Transferir y compartir. Las organizaciones necesitan datos de fuentes secundarias, generados por otros actores del ecosistema de datos, útiles para la toma de decisiones, planificación y mejora operativa, entre otras funciones. Existe mayor presión por compartir los datos generados. Los datos compartidos pueden ayudar a mejorar las decisiones, ya que las organizaciones pueden obtener una visión más completa del impacto de sus decisiones en función de las contribuciones de nuevos conjuntos de datos de más fuentes, tanto internas como externas. Esta capacidad de compartir datos debe ser compatible con la interoperabilidad y mediante el uso de un estándar de lenguaje de intercambio común.

Archivar y preservar. El archivado de datos se refiere a copiar datos en un entorno donde se almacenan en caso de que se necesiten nuevamente en el entorno de producción, así como a eliminar esos datos de todos los entornos de producción activos. Un almacén de datos es simplemente un lugar donde se almacenan los datos, pero no se realiza ningún mantenimiento o uso general. Si es necesario, los datos se pueden restaurar a un entorno en el que se puedan utilizar.

Reutilizar. Significa usar datos que fueron recopilados para otro fin.

Figura 5: Fases del ciclo de vida de los datos.

Fuente: Banco Mundial, 2021, p. 1.



Síntesis de capítulo

Los datos son la unidad mínima con valor que, una vez ya compuestos, adquieren otro significado, de igual forma también representa la base de la realidad y constituyen una pequeña parte de todo un ecosistema informativo presente las actividades cotidianas, pero solo alcanzan sentido a través de un contexto o un análisis. Gracias a los datos se tiene un conocimiento más cercano a la realidad, llegando incluso a proporcionar predicciones sobre algunos acontecimientos, su importancia radica en la capacidad de representación de la realidad, las situaciones o las problemáticas. La sociedad actual, ha generado muchos tipos de datos, volviendo al mundo de los datos más rico y diverso, para ello ha sido necesario estructurar los datos a través de su ciclo de vida el cual tiene como propósito, conocer todo lo que sucede con los datos, desde su creación hasta su almacenamiento.

Capítulo 2 *Big Data*

2.1 Definición

En el año 2009 surgió un artículo que llamó la atención de la comunidad científica, en dicho artículo se evidenciaba cómo un grupo de ingenieros de la empresa Google, mediante el análisis de patrones de búsqueda simples realizadas por los usuarios, pudieron predecir a detalle, el avance de casos de influenza en EUA. A esta herramienta se le nombró *Google Flu Trends*, las predicciones fueron ciertas, en un 97% al ser comparadas con la información oficial del *Centers for Disease Control and Prevention*. es decir, la herramienta, de *Google* pudo recopilar, analizar datos y por consecuente predecir con mucho más rapidez y precisión casos de influenza.

En medio de este trabajo, lo que sorprendió a la mayoría de los testigos del evento, fue la gran ventaja de predicción con la que contaba Google, la cual era de, incluso diez días de anticipación respecto a las cifras oficiales, todo esto operando a un bajo costo. Al paso del tiempo, sentó un precedente sobre lo que se puede lograr, si se trabaja adecuadamente con el *Big Data* o datos masivos.

La cantidad de datos que existen hoy en día nunca se hubieran podido lograr de no ser por la tecnología actual. En consecuencia, para dar paso al tema de *Big Data*, es oportuno comenzar con una breve revisión del auge computacional y de internet.

La primera máquina de procesamiento de datos apareció en 1943 y fue desarrollada por los británicos para descifrar los códigos nazis durante la Segunda Guerra Mundial. Este dispositivo, llamado Colossus buscaba patrones en mensajes interceptados a una velocidad de 5.000 caracteres por segundo. De ese modo, se reduce la tarea de semanas enteras a solo unas pocas horas. (Sánchez, 2018, p. 11)

Fue gracias a la Segunda Guerra Mundial y el gran cambio que trajeron consigo las máquinas y la red, que se empezaron a plantear los centros de procesamientos de datos.

En 1965 el gobierno de los Estados Unidos planea el primer centro de datos del mundo para almacenar 742 millones de declaraciones de impuestos y 175 millones de juegos de huellas dactilares en cinta magnética. (Sánchez, 2018, p. 11)

Al descubrir el gran potencial de los datos, no solamente se empezaron a usar con propósitos bélicos o gubernamentales, sino que se expandieron hacia otras ciencias.

Fue en ciencias como la astronomía y la genética, que experimentaron por primera vez esa explosión en la década de 2000, donde se acuñó el término Big Data, “datos masivos” El concepto está trasladándose ahora hacia todas las áreas de la actividad humana. (Mayer-Schönberger y Cukier, 2013, p. 7)

Las ciencias comenzaban a emplear grandes cantidades de datos, y de acuerdo con Francis X. Diebold fue a mediados de 1990 que el término *Big data* (datos masivos) comenzó a usarse a partir de conversaciones en *Silicon Graphics Inc.* (SGI) y en 2008 saltó a la fama gracias a un artículo de la revista electrónica *Wired*, en dicho artículo:

Se proponía que cantidades masivas de información vuelven obsoleto el método científico o el uso de modelos o teorías para explicar un fenómeno. Es decir, que bastaba una gran cantidad de datos para poder hacer predicciones sin tener que contar con un mayor conocimiento del tema. Sin embargo, existen ejemplos que muestran los grandes riesgos de usar estas predicciones sin tener un modelo o una explicación de fondo. (Oficina de Información Científica y Tecnológica para el Congreso de la Unión, Gobierno de México, 2016, p. 2)

Si bien, la apreciación hecha por los investigadores de *Wired*, fue notablemente riesgosa, no cabe duda de que el hecho de mencionar una nueva forma de investigación causó gran curiosidad a la comunidad científica.

Mientras tanto, el término *Big data* que en Latinoamérica es acuñado como macrodatos, datos masivos o inteligencia de datos, pero es común usar el anglicismo *Big Data*.

Actualmente, hace referencia a una cantidad de datos tal que supera la capacidad del software convencional para ser capturados, administrados y procesados en un tiempo razonable y que, por el momento, es difícil analizarlos utilizando herramientas tradicionales. (Toro y Laniado, 2019, p. 3)

Lagoze propone una visión más compleja del *Big Data* cuando señala que:

Una definición multidimensional popular de Big Data se basa en los llamados 3 V: volumen, velocidad y variedad. El volumen es el factor de tamaño. La velocidad se refiere a la velocidad de acumulación, la naturaleza dinámica resultante de los datos y la capacidad de procesamiento a gran escala necesaria para hacerlos útiles y mantenerlos actualizados. Finalmente, Variedad se refiere a la mezcla o combinación de tipos de datos heterogéneos, modelos y esquemas y la necesidad de resolver estas diferencias. (Lagoze, 2014, p. 2)

Bajo un esquema más amplio, se puede describir al *Big Data* desde tres dimensiones:

- Volumen: cantidad de datos empleada.
- Velocidad: que tan pronto se obtienen o se producen los datos.

- Variedad: que tan complejo es la información que se recibe.

El término «*Big Data*» se refiere a la recopilación de todos los datos y a la capacidad de utilizarlos para el beneficio de las personas o empresas, en una amplia gama de sectores, incluido el de los negocios. Los avances en el almacenamiento y el análisis de datos hacen que se puedan captar, almacenar y trabajar con muchísimos tipos diferentes de datos. Como consecuencia: “los datos hoy abarcan cualquier cosa, desde hojas de cálculo hasta fotos, vídeos, grabaciones de sonido, texto escrito y datos de sensores” (Marr, 2018, p. 3). El *Big Data*, podría parecer imposible de manejar, ya que dentro de sí mismo suele contener distintos tipos de datos, por consiguiente, el *Big Data*, encuentra lugar en muchos negocios, disciplinas, empresas e instituciones distintas. Debido a esto, algunos autores han llegado a la conclusión de que la era de los grandes datos ha comenzado.

Pero esto no es más que el principio. La era de los datos masivos pone en cuestión la forma en que vivimos e interactuamos con el mundo. Y aún más, la sociedad tendrá que desprenderse de parte de su obsesión por la causalidad a cambio de meras correlaciones: ya no sabremos por qué, sino solo qué. Esto da al traste con las prácticas establecidas durante siglos y choca con nuestra comprensión más elemental acerca de cómo tomar decisiones y aprehender la realidad. (Mayer-Schönberger y Cukier, 2013, p. 8)

En algunos casos el *Big Data* hoy en día brinda posibilidades infinitas, si se analiza con mayor profundidad la manera en cómo se desenvuelve en el entorno.

2.2 Propósito

Muchos de los datos masivos que se generan diariamente, permiten a los investigadores y a los científicos entender mejor los fenómenos que acontecen, siempre y cuando estos datos pasen por un proceso adecuado. Pese a que muchos piensan que los datos sustentan a los sistemas, Mayer-Schönberger y Cukier señalan que:

El uso de datos masivos no consiste en intentar “enseñar” a un ordenador a “pensar” como un ser humano. Más bien consiste en aplicar las matemáticas a enormes cantidades de datos para poder inferir probabilidades: la de que un mensaje de correo electrónico sea spam; la de que la combinación de letras “Iso” corresponda a “los”; la de que la trayectoria y velocidad de una persona que cruza sin mirar suponen que le dará tiempo a atravesar la calle, y el coche auto conducido solo necesitará aminorar ligeramente la marcha. (Mayer-Schönberger y Cukier, 2013, p. 11)

El *Big Data* ha establecido parámetros tan precisos de comportamiento, que ahora la cuestión no se enfoca tan solo en predecir movimientos, sino también comprender porque ocurren entre otras funciones. En un mundo de datos masivos, no es necesario concentrarse en la causalidad; por el contrario, se puede descubrir pautas y correlaciones en los datos que ofrezcan perspectivas nuevas y difíciles de ver. Puede que las correlaciones no muestren precisamente por qué está ocurriendo algo, pero alertan de que algo está pasando.

Algo que se debe tener en cuenta es que los datos masivos por sí mismos no van a predecir o describir acontecimientos, los datos necesitan ser gestionados para ofrecer resultados, en tanto que de modo desarticulado no muestran un panorama tan claro.

A menudo, los datos masivos resultan confusos, de calidad variable, y están distribuidos entre innumerables servidores por todo el mundo. Con ellos, muchas veces nos daremos por satisfechos con una idea de la tendencia general, en lugar de conocer un fenómeno hasta el último detalle, céntimo o molécula. No es que renunciemos a la exactitud por entero; solo abandonamos nuestra devoción por ella. Lo que perdemos en exactitud en el nivel micro, lo ganamos en percepción en el nivel macro. (Mayer-Schönberger y Cukier, 2013, p. 12)

Lo prodigioso del *Big Data* reside en lo que se llega a descubrir, esto permite tener un juicio más abierto al momento de analizar los grandes datos. Con todo esto de fondo, se sabe que el *Big Data* cumple variados propósitos, dependiendo el contexto en el que se le acomode, y con que fines se use, un gran ejemplo de ello, son las empresas privadas, que han puesto todo su empeño y recursos en aprovechar hasta la más mínima ventaja competitiva.

Cuando dejamos que hablen los datos, podemos establecer conexiones que nunca hubiésemos sospechado. En consecuencia, algunos fondos de inversión libre analizan Twitter para predecir la evolución del mercado de valores. Amazon y Netflix basan sus recomendaciones, en una miríada de interacciones de los usuarios de sus páginas web. Twitter, LinkedIn y Facebook trazan la “gráfica social” de relaciones de los usuarios para conocer sus preferencias. (Mayer-Schönberger y Cukier, 2013, p. 13)

Esto invita a la reflexión sobre el poder que tienen las grandes empresas, y de qué manera el *Big Data* cumple su propósito en ellas. Cada solución que aporta, dependiendo de la forma en cómo se trabajen los datos, es única, y no se puede abordar un problema de la misma manera que otro. Desde tal perspectiva, las empresas suelen abordar un problema con datos y tecnología, desde el objetivo que tienen en mente. Se comienza analizando el problema, recopilando y explorando los datos necesarios para resolverlo. Así: “El *Big Data* no permite detectar comportamientos individuales o minoritarios que se salgan de la media, sino

tendencias y patrones. Dicho de otra forma, el *Big Data* no es capaz de encontrar una aguja en un pajar, aunque sí podría decirnos qué longitud y tonalidad tienen la mayoría de los trozos de paja” (Pulido, 2016, p. 29). Los propósitos del *Big Data* no son exclusivos del ámbito empresarial, también existen dentro del ámbito académico y científico, muchas veces siendo una herramienta esencial para las investigaciones científicas.

Ya hemos mencionado que el *Big Data* puede sugerir hipótesis e indicar a los investigadores dónde iniciar un análisis, pero para crear modelos siempre será necesario un conocimiento profundo de la materia que solamente poseen los expertos. Ellos y los científicos de datos, de los que ya hemos hablado, son imprescindibles para orientar los análisis e interpretar su resultado. (Pulido, 2016, p. 30)

Una de las limitantes que tiene el *Big Data* es el hecho de que la mayoría de las veces contiene una gran cantidad de ruido, es decir, información inútil, irrelevante o engañosa, que está de manera más presente y crece mucho más rápido que los datos que en verdad son relevantes. Otra limitante es el periodo de validez tan corto de los datos, por esto, es necesario analizarlos para tener en cuenta sus cambios.

Con todo lo descrito anteriormente se concluye que el propósito del *Big Data* dependerá principalmente del contexto en el que se encuentre, cada empresa, institución o profesional de los datos aplicará los análisis correspondientes a sus datos para que respondan a las necesidades de información que desee cubrir.

Pero ¿para qué podemos usar todos estos datos? ¿De qué nos sirve cuantificar? Es imprescindible saber para qué lo estamos haciendo. Las posibilidades son múltiples y pueden ser muy ambiciosas. Una muy poderosa es la de predecir y cambiar comportamientos, que convierte al *Big Data* en una auténtica vía para la persuasión. La técnica requiere de refinamiento. Es necesario realizar un buen procesamiento y análisis de los datos cuantificados. (Paniagua, (s.f.), p. 21)

Una de sus variadas posibilidades, quizá la más emocionante, es revelar aquellos enigmas que impiden tener una perspectiva completamente nueva de algún problema en específico. Por último, y para concluir, uno de los más aterrizados ejemplos de uso del *Big Data* se encuentra en el Neuromarketing:

Se trata de una combinación entre neurociencia y minería de datos, que integra el aprendizaje computacional automático con la intuición humana y el *Big Data*. Esto permite una medición directa de los pensamientos del consumidor acerca de un anuncio para conocer sus efectos persuasivos, teniendo en cuenta que serán mayores cuanto más se alineen emocionalmente con el espectador. (Paniagua, (s.f.), p. 23)

2.3 Gestión

La gestión del *Big Data* es el término con el cual se describe la recopilación, análisis, y empleo de cantidades extraordinariamente grandes de datos, con la finalidad de volver más prácticas ciertas operaciones.

El gran avance hacia la gestión más eficiente de los datos llegó con el advenimiento de la digitalización: hacer que la información analógica fuese legible por los ordenadores, lo que también la vuelve más fácil y barata de almacenar y procesar. Este progreso mejoró drásticamente la eficiencia. La recopilación y el análisis de información, que en tiempos exigía años, podía ahora hacerse en días, o incluso menos. Pero cambió poco más. Los encargados de los datos muy a menudo estaban versados en el paradigma analógico de asumir que los conjuntos de datos tenían propósitos específicos de los que dependía su valor. Nuestros mismos procesos perpetuaron este prejuicio. Por importante que resultase la digitalización para permitir el cambio a los datos masivos, la mera existencia de ordenadores no los hizo aparecer. (Mayer-Schönberger y Cukier, 2013, p. 13)

La gestión de datos permite tener, un mayor control sobre la investigación o el proceso por el que se está atravesando, debido a esto, una empresa privada o pública decide trabajar con grandes datos.

Adicionalmente, entre muchas otras, se pueden incluir propiedades como los cambios en el tiempo (variabilidad), la confiabilidad de las fuentes (veracidad) y qué tan útiles son (valor). Estas propiedades pueden mezclarse dentro del mismo conjunto, pero antes de analizarlos es necesario depurarlos, estandarizarlos y conectarlos o catalogarlos de manera adecuada. A este proceso se le denomina gestión de datos. (Oficina de Información Científica y Tecnológica para el Congreso de la Unión, Gobierno de México, 2016, p. 2)

Los metadatos tienen un papel principal dentro de la gestión de datos, gracias a ellos se obtiene información complementaria como su origen, calidad y confiabilidad, “Si se tiene una colección de fotografías, estas serían los datos y los metadatos pueden ser la fecha de creación, si ha sido editada o no, el autor, la cámara utilizada, etc. La información útil o sensible puede estar tanto en los datos como en los metadatos” (Oficina de Información Científica y Tecnológica para el Congreso de la Unión, Gobierno de México, 2016, p. 2). Esta información extra, ayuda a tener más control sobre los datos, a través de su análisis, permitiendo que los resultados no solo sean más exactos, sino más confiables y comprobables. La gestión de datos implica procesos y tecnologías de apoyo para adquirir, almacenar, preparar y recuperar los datos para su análisis. “El análisis, por su parte, se refiere

a las técnicas utilizadas para adquirir inteligencia a partir de Big Data” (Hernández et al., 2017, p. 19)

El proceso de análisis de datos dependerá en gran medida de las necesidades y requerimientos expresados por la empresa o institución a la cual se ejecuta este proceso, para fines prácticos de la tesis se tomó como ejemplo, el proceso que lleva a cabo la empresa de datos PowerData el cual consta de tres capas, que están íntimamente ligadas con los requisitos tecnológicos actuales, pero también pueden ser adaptadas de acuerdo con las necesidades que presente cada proyecto.

En la primera capa se lleva a cabo el análisis del *Big Data*, este es ejecutado a través de múltiples acciones, de esta manera se visualizan los datos, para posteriormente darles forma a través de la estadística avanzada, análisis predictivo, y el aprendizaje automático.

En la segunda capa entran tres grandes rubros, integración de datos, gobierno y seguridad. “Es necesario incorporar los datos que tengan un mayor rendimiento, procesando de manera optimizada, asegurándose de que sean escalables y por supuesto, que se implementen con la mayor flexibilidad” (PowerData, 2019, párr. 5) También el segundo rubro habla del gobierno de datos, el cual “se va a centrar en la preparación de toda la información que administramos, asegurándose que tiene la máxima calidad. Se les otorgará un linaje a esos datos y habrá que detectar todas las relaciones que presentan entre ellos” (PowerData, 2019, párr. 5) Y finalmente, “la seguridad se utiliza para analizar los riesgos, tener controlados los datos sensibles y protegerlos todo lo posible con una política de protección máxima”. (PowerData, 2019, párr. 5)

Para finalizar la tercera capa dice que, “se relaciona con el procesamiento y el almacenamiento de los datos”. (PowerData, 2019, párr. 5). Este es solo un proceso, de muchos otros, con los cuales el *Big Data* se puede gestionar de manera viable para aprovechar todos sus recursos.

El mantener una correcta gestión del *Big Data*, puede traer grandes beneficios a la empresa, instituto o proyecto a largo y mediano plazo, asegurando un alto nivel de calidad en los datos, también se tiene accesibilidad para otros investigadores, inteligencias de negocios, uso de aplicaciones para conseguir información desconocida incluso viniendo de fuentes diversas,

entre otros beneficios.

La era en la que los datos permanecían guardados, casi escondidos empolvándose, ha terminado, pacientemente los datos esperaron muchos años, hasta que se tuviera la capacidad tecnológica para aprovecharlos. Para finalizar, es importante saber, qué tipo de datos pueden ser útiles para que la gestión sea llevada a cabo de la mejor manera.

Confirmar cuáles son los datos que se conservarán y que serán analizados, con el fin de mejorar los procesos de negocio actuales o proporcionar un negocio con una ventaja competitiva. Garantizar el éxito en la gestión de Big Data depende de llevar a cabo una cuidada clasificación de datos que asegure que, en última instancia, los conjuntos de datos más pequeños pueden ser analizados con la premura suficiente para mantener los niveles de productividad. (Logicalis, 2017 p. 16)

2.4 *Big Data* en las bibliotecas

El *Big Data* es ocupado, en mayor medida por las empresas que desean conocer o predecir los gustos, intereses y movimientos de sus compradores, así como también en el ámbito científico, y en donde se utiliza, para analizar fenómenos, y descubrir patrones que no son obvios a simple vista. No obstante, el *Big Data* como eje fundamental del texto, está presente en múltiples acciones de la vida cotidiana, como las vinculadas a la búsqueda de información, la personalización de información, la publicidad, el reconocimiento de patrones de información, al teclado predictivo y al desarrollo de la investigación de dominios.

Todos estos increíbles usos pueden no parecer de utilidad a simple vista para la biblioteca, pero lo cierto es que, nada puede estar más alejado de la realidad, hoy más que nunca la biblioteca puede entender la relación entre dato, información y conocimiento.

Se ha intentado comprender esta realidad a partir de recopilar y analizar datos en la web en una gran escala en términos de desarrollo de software para descargar datos de la web y de métodos específicos para analizar la información (por ejemplo, análisis de enlaces). A esto se añade el papel de las teorías documentales centradas en la información. Otro de los esfuerzos documentales es la normalización de prácticas para representar e integrar datos, dentro de los que se puede destacar la integración de la web semántica y los Open Data. Por último, y en su vinculación con la social media, sobresalen las herramientas, técnicas y teorías bibliométricas, que también cuentan con nuevas perspectivas. (Valenzuela, 2019, p. 79)

En esta nueva etapa de la biblioteca, se puede decir que el *Big Data* vino a potenciar las labores rutinarias y esfuerzos documentalistas y bibliotecológicos que se han usado por años para mantener la biblioteca activa y funcional. Esto tiene pros y contras, pues la posibilidad de compartir datos supone un nuevo reto en las tareas básicas de selección, análisis y difusión de información que realiza el bibliotecario, pero también significa un nuevo resurgimiento para la profesión y sus labores.

Algunas bibliotecas por ejemplo se han involucrado con el *Big Data* al utilizar los datos que producen, todo el tiempo, esto con el fin de:

Generar mapas de conocimiento a partir de técnicas bibliométricas, permitiendo obtener una gran cantidad de información a partir de los artículos publicados por los investigadores y remitiéndolo directamente a aquellos alumnos interesados por la titulación que cursan o los gustos personales que han establecido previamente a través de algún medio telemático. (Fraile, 2018, p. 17)

Actualmente, la biblioteca ya no está centrada únicamente en los datos de archivo, sino que ahora toma en cuenta el universo de los datos en general.

Desde un plano documental nos encontraríamos con datos de navegación (analítica web) y datos que existen en Internet (cibernetría) alrededor de nichos de mercado, comportamientos de búsqueda, segmentación de clientes o potenciales clientes, relaciones entre marcas, la conexión entre servicios y productos y sus usuarios, las relaciones con competidores, la vigilancia tecnológica. (Valenzuela, 2019, p. 92)

Tenemos entonces que hoy en día existen otras necesidades de información, que ya no pueden ser solamente satisfechas a través de archivos físicos o libros, están surgiendo otras necesidades derivadas de los datos que representan otro recurso, que la biblioteca debe resguardar y gestionar para ser ofrecido a sus usuarios.

La primera reflexión reside en el hecho de que “los datos” son cada vez más un objeto que las bibliotecas de investigación en particular van a tener que recoger y conservar, ya que se trata del producto del trabajo de los investigadores, de igual manera que se hacía antes para su producción documental impresa y posteriormente digitalizada (artículos, tesis, etc.). Ahora bien, estos datos tienen características diferentes de los documentos: se presentan en forma de flujo y no de almacenamiento y requieren herramientas de análisis para poder ser utilizados. (Lagoze, 2014, p. 6)

Una de las ventajas en las que el *Big Data* abrirá muchas puertas a la biblioteca, es el hecho de que por primera vez se podrá analizar el comportamiento del usuario e incluso, se podrán

hacer campañas, a través del *Big Data*, para atraer más usuarios a la biblioteca. Estas campañas podrían estar basadas en estadísticas de consulta de la biblioteca para observar cuáles son los servicios más utilizados y recomendar mejoras.

Dentro del estudio de usuarios, el *Big Data* puede mostrarse como una herramienta útil, a la hora de construir perfiles en lectores que satisfagan de mejor manera sus hábitos de adquisición de ejemplares.

Esta tecnología también se utiliza para gestionar la política de adquisiciones teniendo en cuenta datos como el perfil sociológico de los lectores que frecuentan cada biblioteca de la red, las restricciones del lugar y la tasa de rotación de las colecciones. (Lagoze, 2014, p.7)

De igual manera, se debe tener en cuenta las necesidades de información de los investigadores. En ciertas disciplinas de humanidades, especialmente las relacionadas con la organización documental o las bibliotecas, se están inclinando por el contenido, más que por el soporte, explorando por completo la colección y su organización.

La interrogante ya no radica en definir si el *Big Data* debe o no ser considerado por las bibliotecas y los bibliotecarios, sino en cuánto tiempo estarán activamente trabajando y si los bibliotecarios están listos para enfrentarse a este gran cúmulo de recursos de información. Cito a Lagoze cuando dice que:

La cuestión es saber si las bibliotecas serán capaces de ofrecer a los investigadores un servicio de *Big Data* de las colecciones que conservan: permitirles definir su propio corpus, sus criterios y aplicar herramientas de análisis para extraer nueva información a partir de los datos. (Lagoze, 2014, p. 9)

2.5 El papel de las bibliotecas en la era del *Big Data*

El tiempo de la biblioteca y los grandes datos ha llegado, nunca en toda la historia se había llegado a tener tal cantidad de información en el mundo, la biblioteca, desde siempre y aún hasta la fecha, es el semillero del conocimiento, la cuna y el refugio de todos aquellos que buscan información y actualmente, aunque el *Big Data* ha revolucionado muchas cosas, esta percepción de la biblioteca no debe cambiar muy por el contrario hoy más que nunca, es el tiempo en el que la biblioteca debe tomar la batuta, en este singular acontecimiento. La biblioteca, como ente social, no puede sino adaptarse a las nuevas circunstancias, pero sobre todo al nuevo papel que le toca desempeñar.

Porque la biblioteca y sus profesionales, en definitiva, proponen un servicio al público. Por lo tanto, es necesario conocer a ese público. Pero ese público, así como los datos, también va evolucionando. (Lippenholtz, 2016, párr. 12)

La biblioteca no puede ya ser vista, como un lugar donde solo existen libros o registros, y el usuario no puede ser relegado al lugar de un simple espectador.

Es un camino doble: llega información de los usuarios para que vuelva a ser propuesta a los usuarios en «productos personalizados». Este circuito es el que se ha modificado. Insistimos, los usuarios/lectores antes eran solo receptores. Hoy, son, al mismo tiempo, receptores, pero también productores y fuentes de información. (Lippenholtz, 2016, párr. 13)

En los tiempos actuales, la biblioteca y el bibliotecario deben tomar en serio su vital papel en la participación del *Big Data* para aportar sentido humano y cooperativo en los servicios que se le proporciona a los usuarios.

Nuestras comunidades, nuestras sociedades, nuestras culturas son demasiado importantes para quedarse al margen y simplemente observar o recopilar su producción. Debemos reconocer, sin la neblina de la nostalgia, que somos actores en este mundo y aceptar la responsabilidad de trabajar directamente con comunidades de todo. (Lankes, 2019, p. 2)

El mundo de los datos es un nuevo panorama y en este, sin duda, las nuevas tecnologías, especialmente para muchos bibliotecarios tradicionales, significan un cambio brusco, no obstante, el universo de los datos es una gran oportunidad que tiene la biblioteca para poder aportar y contribuir sobre los recursos de información que brinda a sus usuarios sin embargo se debe considerar que:

La gestión y el almacenamiento, o si se quiere, la cosecha de datos puede ser mecanizado y los algoritmos pueden hacer ciertos cruces, pero aún es el bibliotecario, es el humano quien puede hacer las mejores recomendaciones a través del “descifrado” de la data. (Lippenholtz, 2016, párr. 22)

Síntesis de capítulo

Debido a la tecnología actual, existe una gran cantidad de datos, que, al ser recopilados, crean lo que se conoce como Big Data, dicho termino es acuñado en Latinoamérica como macrodatos. El Big Data a través de la gestión, y el análisis, permite a los investigadores y a los científicos entender los fenómenos que acontecen a través de patrones, con la finalidad de tener un mayor control sobre las investigaciones o los procesos. En esta nueva etapa el Big Data vino a potenciar las labores rutinarias para mantener la biblioteca activa y funcional, significando un nuevo resurgimiento para la profesión y sus labores.

Capítulo 3 Del bibliotecario tradicional al bibliotecario de datos

3.1 El bibliotecario tradicional

3.1.1 Definición

La figura del bibliotecario tradicional es una de las más versátiles y curiosas dentro del ámbito académico, pues en tanto otras profesiones tienen una marcada línea académica, la figura del bibliotecario ha cambiado con los años, una evolución que tiene sentido si analizamos de cerca su razón profesional de ser. En palabras de Ortega y Gasset:

Para determinar la misión del bibliotecario hay que partir, no del hombre que la ejerce, de sus gustos, curiosidades o conveniencias, pero tampoco de un ideal abstracto que pretendiese definir de una vez para siempre lo que es una biblioteca, sino de la necesidad social que vuestra profesión sirve. Y esta necesidad, como todo lo que es propiamente humano, no consiste en una magnitud fija, sino que es por esencia variable, migratoria, evolutiva-en suma, histórica. (Ortega y Gasset, 2005, p. 21)

La biblioteca, al igual que el bibliotecario, tienen una razón social de existir, es decir, residen a través de la sociedad, como una institución con un sentido social e histórico que cambia, dependiendo de los requerimientos de sus usuarios, por consiguiente, es de esperarse que el bibliotecario cambie, acorde a las necesidades de información de su época. La actividad del bibliotecario aparece como respuesta a una necesidad social, no al deseo de indagar en la realidad como la Física o la Biología, o de aplicar conocimientos previos como en Medicina o Arquitectura, por ejemplo: “El "hacer" tuvo prioridad por sobre el "pensar" y este origen situó a la actividad en un marco "no científico” (Roggau, 2006, p. 19). Al ser el usuario, una variante tan cambiante y contextual, el bibliotecario y su profesión son adaptables completamente a él y lo que le haga falta.

El bibliotecario cumple un rol social muy importante que es facilitar la información disponible en su biblioteca a todos los usuarios posibles. Para cumplir con ese rol el bibliotecario debe ser paciente, creativo, ordenado, disciplinado; conocer las características, intereses y necesidades de sus usuarios. Además de estar familiarizado con el contenido temático del acervo; orienta al usuario sobre la consulta interna; dar a conocer y aplicar el reglamento de la biblioteca, revisar diariamente la ordenación de los materiales documentales en la estantería, acomodar los que fueron consultados durante el día y finalmente, elaborar reportes de sus actividades, entre otras. (Universidad Nacional Autónoma de México, Dirección General de Bibliotecas, 2014, p. 40)

Pese a que las definiciones mencionadas con anterioridad dan un panorama amplio del significado social del bibliotecario, vale la pena detenerse a explorar la definición técnica

que ofrece el *Online Dictionary of Library and Information Science*, el cual menciona que el bibliotecario es:

Una persona capacitada profesionalmente responsable del cuidado de una biblioteca y su contenido, incluyendo la selección, procesamiento y organización de materiales y la entrega de servicios de información, instrucción y préstamos para satisfacer las necesidades de sus usuarios. En un entorno online, el papel del bibliotecario es gestionar y mediar el acceso a información que puede existir solo en forma electrónica. (Reitz, 2013, p. 376)

De igual manera se define al bibliotecario por tres aspectos básicos, los cuales son: el administrativo, el procesamiento de su materia prima de trabajo y la circulación del material informativo, conocido como servicio al público.

Existen tres factores determinantes en los aspectos teóricos y prácticos del servicio de información que son indispensables incluir en el uso de la tecnología y la automatización, estos son: el administrativo, la organización bibliográfica o procesos técnicos y el servicio al público. El bibliotecario ejerce sus labores de manera diversa, pero en la mayoría de los casos consiste en:

El trabajo para obtener, organizar, conservar y difundir toda idea registrada por medio de signos gráficos, iconográficos y fonéticos en un soporte llamado papel, plástico, piel, madera, etc.; en forma de libro, revista, cinta, disco, entre otros formatos; del conocimiento generado y construido por el hombre sobre sí mismo y sobre todos los fenómenos que lo rodean ha sido, es y sigue siendo hasta la fecha tarea o trabajo de la disciplina conocida como Bibliotecología. (Gutiérrez, 2013, p. 3)

La figura del bibliotecario goza de un espectro amplio porque, así como lo hace una empresa, la biblioteca, de igual forma, necesita profesionales que desempeñen varias labores, unas completamente diferentes entre sí, pero que, juntas puedan ser parte de una gran maquinaria en funcionamiento. De esta forma es como se llegan a encontrar con muchos tipos de bibliotecarios, cada uno con una función distinta, pero que, con su labor, cubren y enriquecen las funciones de la biblioteca:

- **Bibliotecario administrativo.** Un administrador de biblioteca es responsable del control y supervisión, incluida la planificación, elaboración de presupuestos, formulación de políticas, gestión de personal, relaciones públicas y programas de descarga. También generalmente se refiere a las personas responsables de la administración de la

biblioteca, usualmente el consejo de administración o decanos, el director de la biblioteca y su personal inmediato.

- Bibliotecario catalogador. El catalogador es aquel que se encarga del proceso de creación de metadatos, que representan una fuente de información. La catalogación proporciona información como nombres de autores, títulos y descripción de temas para las fuentes, a menudo mediante la creación de registros bibliográficos. Aunque la catalogación generalmente da como resultado la creación de catálogos de bibliotecas, también produce otros tipos de herramientas de identificación.
- Bibliotecario de referencia. Es aquel bibliotecario que conoce técnicas y servicios bibliotecarios para informar de manera adecuada sobre el uso de la biblioteca, y los productos que ofrece junto con su colección.
- Bibliotecario de servicios. Es el bibliotecario, que se encarga de atender y darle seguimiento a las necesidades de información de los usuarios.

Una vez mostrados los tipos de bibliotecarios, no es de sorprender que el bibliotecario tenga una figura tan experimental y cambiante, sobre todo en las últimas décadas que se han expandido los servicios de información en las bibliotecas. “Es un rol multifacético, multidisciplinario, por qué no solamente se trata de informar al usuario de lo que hay en la biblioteca sino de responsabilizarse de la formación del usuario. Formar y apoyar la investigación” (Infotecarios, 2017, párr. 7). Dentro de todos los conceptos vistos anteriormente, encontramos que muchos coinciden, con base en ello, se puede decir que el bibliotecario es: un profesional, encargado de las funciones específicas de acuerdo con su puesto, mediador experto que conduce al usuario a la satisfacción de sus necesidades de información.

Pese a que los cambios sociales y la nueva forma en como percibimos la información trascienden más allá de la biblioteca, las habilidades requeridas para ordenarla no han cambiado mucho, esto implica que los bibliotecarios poseen las herramientas empíricas necesarias para hacer frente a otro tipo de necesidades profesionales de la información, en tanto que:

Como profesionales de la información tienen las habilidades necesarias para colaborar en la gestión del conocimiento de las organizaciones, participar en procesos de inteligencia competitiva y trabajar en forma interdisciplinaria en el

diseño de sistemas de información, entre otras actividades. Sus competencias y su experiencia les permiten aportar un valor agregado al proceso de toma de decisiones de las empresas. (Gibaja, 2013, p. 51)

Comprender el significado del bibliotecario tradicional, ayuda a entender de forma amplia el espectro de su función como profesional de la información.

3.1.2 Funciones

Anteriormente se señalaron los diferentes tipos de bibliotecarios de modo que sus funciones dependen del puesto que desempeñan. Dentro de las funciones del bibliotecario tradicional encontramos que es el centro mismo de la biblioteca, el puente entre el usuario y la información. Según la Dirección General de Bibliotecas de la Universidad Nacional Autónoma de México:

El bibliotecario desarrollará sus actividades en alguna de las dos áreas de la biblioteca: procesos técnicos o servicios al público. Si se desempeña en el área de procesos técnicos su labor consistirá en elaborar listados, llenar formatos, realizar el proceso menor, actualizar los catálogos mediante los cargos remotos, entre otros procesos; si se le asigna en el área de servicios al público deberá brindar el préstamo en sala y a domicilio, acomodar las colecciones en la estantería, orientará al usuario y colaborar en el inventario, entre otras actividades. Se le indicará su lugar de trabajo y se le dotará del equipo y accesorios de oficina necesarios para desarrollarlo; recibirá instrucciones sobre las tareas, tiempos y movimientos de su puesto de trabajo. (Universidad Nacional Autónoma de México, Dirección General de Bibliotecas, 2014, p. 38)

Al paso del tiempo, las funciones del bibliotecario se han vinculado con tareas específicas de la biblioteca donde la tecnología se hace presente, agregando más actividades al bibliotecario tradicional. Las siguientes son funciones que, de acuerdo con Manzanares, son inherentes y propias del bibliotecario actual:

- Gestión y tratamiento de la información y el conocimiento digital.
- Proceso técnico de todas las publicaciones y todos los soportes electrónicos o digitales.
- Supervisión y mantenimiento de todos los servicios automatizados y robotizados (Manzanares, 2018, p. 30)

Una descripción más amplia señala que el bibliotecario:

- Proporciona, en forma manual o automatizada, los servicios de préstamo, devolución y resello de libros, tesis, revistas, periódicos, mapas y material

documental en cualquier formato; notifica las multas correspondientes por infracciones de los usuarios.

- Proporciona información descriptiva sobre los servicios que ofrece la biblioteca y orienta a los usuarios acerca de los requisitos que debe cubrir para hacer uso de esta.
- Recibe, coteja, coloca y localiza oportunamente los libros, tesis, revistas, periódicos, mapas y material documental en cualquier formato (CD) videocasetes, DVD, disquetes, diapositivas, etcétera así mismo intercala, acomoda y reubica correctamente las colecciones.
- Verifica en el momento de la entrega y/o recepción de los libros y material documental en cualquier formato, el estado en que se encuentra y en su caso, reporta a su jefe inmediato cualquier anomalía detectada.
- Reporta, relaciona y coteja libros, tesis, revistas, periódicos, mapas y material documental en otros formatos que se encuentren deteriorados o mutilados para su envío a restauración o encuadernación. En caso de que proceda, efectuar reparaciones menores del material dañado.
- Elabora tarjetas y credenciales para préstamo de material documental utilizando medios mecánicos, magnéticos y/o electrónicos.
- Resguarda y custodia los libros y demás material documental en cualquier formato.
- Participa y promueve la observancia del reglamento de la biblioteca y reporta a su jefe inmediato las infracciones derivadas de su aplicación.
- Registra la información necesaria, a través de medios mecánicos, magnéticos y/o electrónicos de fascículos de revistas y demás material hemerográfico en cualquier formato.
- Elabora y coloca correctamente en libros y material documental en cualquier formato, etiquetas, marbetes, cintas magnéticas, papeletas de préstamo, sellos, etcétera.
- Colabora en la actualización de catálogos alfabéticos, geográficos, numéricos, onomásticos y otros, de todo tipo de material documental en cualquier formato, así como de usuarios, utilizando medios mecánicos, magnéticos y/o electrónicos.
- Reproduce fichas bibliográficas, hemerográficas y de otros tipos, por medios mecánicos, magnéticos y/o electrónicos.
- Auxilia en la realización periódica de los inventarios de la biblioteca de: libros, tesis, revistas, periódicos, mapas y material documental en cualquier formato.
- Solicita oportunamente y cuida [sic] los materiales y el equipo que se le proporcione para la ejecución de su trabajo.

- Elabora y entrega diariamente el reporte de actividades desarrolladas durante su jornada. (Universidad Nacional Autónoma de México, Dirección General de Bibliotecas, 2014, p. 42)

Pese a que la lista anterior es bastante completa, a ello se agrega las siguientes funciones, indicadas por Gómez:

- Intermediario y filtro: cuando ayuda a acceder a la información sin condicionarla o crearla, evitando sobrecarga irrelevante.
- Almacenador y guardián de cultura.
- Informador y comunicador, independientemente de soportes o formas y canales.
- Asesor y consulta: cuando hace relevante que al informar está orientando, ayudando a resolver los problemas aconsejando como solucionar las necesidades.
- Educador, como consecuencia de la complejidad de las fuentes de información y las tecnologías que canalizan el acceso al saber dan al bibliotecario una misión relacionada con la educación permanente, y la formación para el autoaprendizaje. (Gómez, 1996, p. 78)

En cuanto a la función social, misionera y ética del bibliotecario, el vicepresidente de la *International Federation of Library Associations and Institution*, J. Bowden, recapituló en el último Congreso las siguientes responsabilidades:

Proteger el derecho de acceso a la información: la información es un derecho, un recurso fundamental de los países. Todo individuo tiene derecho a recibir y difundir información, y los bibliotecarios deben favorecer la libertad de información, combatir la censura o la presión sobre el acceso al conocimiento.

Dar a conocer a todos los individuos ese derecho a la información y la importancia de esta.

Dar la información a los usuarios de un modo correcto, exacto y puesto al día.

En relación con el exceso de información, es un deber profesional no solo dar información exhaustiva, sino con la selección, el análisis y la síntesis necesaria para que sea realmente útil.

La excelencia de los servicios es una obligación profesional porque será la calidad el único medio -y no las meras estrategias de promoción de mercado- de hacer avanzar las bibliotecas. Calidad en el trabajo directo con el usuario, en el aprovechamiento de los recursos, etc.

Estar al día, pues tras un corto espacio de tiempo los conocimientos quedan desfasados. (Gómez, 1996, p. 80)

En el presente, el bibliotecario ya no solo es un facilitador de información, sino un individuo social, humanístico y político, de gran importancia, que ayuda de manera profesional al usuario, en aras de favorecer su derecho de acceso a la información.

3.2 El bibliotecario de datos

3.2.1 Definición

Los orígenes del bibliotecario de datos son recientes y su perfil profesional se encuentra en construcción, especialmente en países donde el impacto del análisis de los datos va más lento. No hay certeza de cuándo o cómo nació el término de bibliotecario de datos, solo que empezó en países anglosajones, donde los académicos e investigadores comenzaron a requerir otro tipo de servicios más específicos sobre los datos; servicios que tenían que ver con el cambio que trajeron consigo las tecnologías y el *Big Data*. En este sentido, es importante tomar en cuenta, que no es la primera vez que el bibliotecario se especializa en algún campo que requiera el sector laboral, prueba de ello son los bibliotecarios de referencia, los bibliotecarios de desarrollo de colecciones, entre otro tipo. La profesión del bibliotecario ha tenido distintas evoluciones, durante su existencia, se ha retroalimentado de otras profesiones, e impulsado por la necesidad de sus usuarios, su profesión ha recorrido largos caminos evolutivos, cubriendo los requisitos que solicita su contexto y considerando el ámbito en el que se requiera.

Algunos autores como Steven J. Bell y John Shank han intentado rastrear el origen del bibliotecario de datos, ellos señalan que al principio fue considerado un “bibliotecario mixto”, y unos años más tarde sería nombrado como “bibliotecario incorporado a datos” una figura más presente en los procesos de investigación. Por último, y más recientemente, en un informe de la *Society of College, National and University Libraries* (SCONUL) de 2017 surgió la definición de “bibliotecario borroso” esto “refiriéndose a un bibliotecario tan mezclado con otras profesiones que terminaba desarrollando habilidades asociadas a datos” (Pinfield et al., 2017, p. 16). Posteriormente “en 2013 ya en relación con la gestión de datos, se definió como bibliotecario incorporado una figura que depende de los procesos de investigación para apoyar la gestión de datos” (Morriello, 2020, p. 3) A este bibliotecario se le puede situar gracias a:

Robin Rice y John Southall dos bibliotecarios británicos que reconstruyeron el camino de la profesión del bibliotecario de datos “se sitúan en la década de 1960 cuando se crearon los depósitos de datos debido a la necesidad de conservación e investigación y su progreso ha dado frutos desde entonces” (Morriello, 2020, p. 5)

Si bien aún no se tiene una definición contundente de lo que es el bibliotecario de datos, si se tiene muy en cuenta que pronto se necesitarán para cubrir necesidades emergentes, en vista de que “las bibliotecas han comenzado a contratar bibliotecarios de datos sin que exista un concepto definido” (Khan y Du, 2018, p. 1). Se requieren profesionales enfocados a la gestión de datos, por consiguiente, es de suma importancia que los educadores y las escuelas conozcan las habilidades, bases y conceptos de esta nueva profesión con miras a volverse líderes del tema en un futuro.

De acuerdo con la literatura, un bibliotecario de datos es “aquel que proporciona un medio para almacenar, recuperar, buscar y registrar el acceso a los datos del laboratorio” (Liscouski, 1997, p. 199). Esta definición podría ser válida en el límite de un laboratorio, pero a medida en cómo avanzaron los datos, su función no solamente fue aplicable en laboratorios.

Otros autores afirman que:

La función de los bibliotecarios de datos es actuar como facilitadores en todas las etapas de la investigación científica, contribuyendo con posibles servicios que podrían ser útiles para la gestión y conservación de datos. (Semeler et al., 2017, p. 2)

Dichos autores recomiendan que los bibliotecarios de datos deberían tener las aptitudes para enseñar en las ciencias de datos. Esto se puede entender en el sentido de que hoy en día los datos son un tipo de información muy consultada, no debe ser extraño que los profesionales tengan conocimientos en diversos campos. Así, ser bibliotecario de datos:

No es una especialidad única, sino más bien una colección variada de funciones superpuestas pero distintas que se centran en proporcionar acceso a los datos, documentarlos y conservarlos, de manera muy similar a como lo ha hecho la Bibliotecología tradicional con las fuentes impresas. (Kellam y Thompson, 2016, pp. 5-6).

En cuanto a responsabilidades, su perfil profesional asume que:

Es responsable de investigar y ayudar en la implementación de los servicios de soporte de la biblioteca relacionados con la descripción, el almacenamiento y el intercambio de metadatos de investigación y conjuntos de datos. (Spinak, 2019, párr. 3)

Según una encuesta hecha en *SurveyMonkey* y distribuida a través de la *Journal of the Medical Library Association (JMLA)* y la *Association of Academic Health Sciences Libraries*, se descubrió que los bibliotecarios de datos:

Son una comunidad heterogénea de profesionales de la información de diversa formación educativa y profesional que realizan muchos tipos diferentes de trabajo. (Federer, 2018, p. 299)

El principal problema con esta definición es que los encuestados solamente pertenecían al campo de la Biomedicina y las Ciencias de la Salud, por lo que se muestra un tanto delimitada. El mismo autor reconoció que los trabajos de bibliotecario de datos no solo son cada vez más comunes en otros campos, sino que también conllevan más habilidades y responsabilidades, tal como lo muestra un análisis del mercado de trabajo en las bibliotecas académicas de USA:

El puesto de Bibliotecario de Datos es responsable de investigar y ayudar en la implementación de los servicios de soporte de la biblioteca relacionados con la descripción, el almacenamiento y el intercambio de metadatos de investigación y conjuntos de datos” (Barreto, 2019, párr. 5)

3.2.2 Funciones

El bibliotecario de datos tiene mucho que ver con la representación, organización y difusión de datos y con el uso de tecnologías para la gestión de datos de investigación y los servicios de datos, esto implica que se debe estar comprometido con un constante aprendizaje.

El *Chartered Institute of Library and Information Professionals (CILIP)* identifica la ciencia de los datos, como una de las cuatro macro-áreas de empleo de los bibliotecarios:

Los bibliotecarios de datos se dedican a gestionar datos de investigación, utilizando esos datos como recurso y apoyando a los investigadores en estas actividades. Como bibliotecario de datos, es probable que participe en el desarrollo o la implementación de un plan de gestión de datos de la organización, el almacenamiento y la gestión de datos y la determinación de los períodos de retención y eliminación.

Se espera que involucre y potencialmente capacite a otro personal en temas como derechos de autor, propiedad intelectual, licencias de datos, embargos, ética y reutilización, alfabetización de datos y privacidad. (Chartered Institute of Library and Information Professionals, 2022, p. 3)

En el año 1990 se produjo una aceleración en los datos que contenían los depósitos gracias al nacimiento del Internet. “Rice y Southall creen que el nacimiento de la Bibliotecología de datos es contextual al proceso de convergencia digital” (Morriello, 2020, p. 5) pues justo entonces, hubo un momento en el que el bibliotecario de datos se volvió un puente entre el usuario y el contenido, se trata de una evolución de las actividades tradicionales de los bibliotecarios “en esencia, se trata de actividades que no se desvían mucho de las actividades cotidianas de los bibliotecarios universitarios, por lo que no debería importarles mucho trabajar en publicaciones periódicas, monografías y otras publicaciones o en datos” (Rice y Southall, 2016, p. 19).

Debido al aumento en la cantidad de datos y la necesidad de organizarlos para que estos puedan convertirse en información y conocimiento, han surgido necesidades que ya no pueden ser satisfechas únicamente desde una vista informática, por lo que ya no es asunto exclusivo del dominio de la tecnología de la información sino también del dominio de la documentación, la descripción, el intercambio y la transferencia de conocimiento. Así es como dentro del ámbito de la gestión de datos el bibliotecario de datos encuentra su lugar.

Las nuevas tecnologías han dado paso a varias cuestiones. La primera de ellas es ¿cómo se llevó a cabo un aumento en la cantidad de datos nacidos originalmente en forma digital, en lugar de aquellos que se crearon como datos analógicos? y posteriormente, serían digitalizados. La segunda es ¿cómo podemos manejar grandes cantidades de datos, ya que los costos han disminuido drásticamente? todo este panorama ha dado pie a que surja la gestión de datos.

La gestión de datos es la praxis de organizar y mantener procesos dentro del ciclo de los datos para satisfacer las necesidades en los continuos períodos y cambios de la información. Así pues, la gestión de datos es una de las principales tareas y antecedentes de lo que hoy son las funciones de un bibliotecario de datos.

El ciclo de vida de los datos está íntimamente relacionado con las funciones que el bibliotecario de datos desempeña en sus labores diarias:

Las funciones de gestión de datos dependen de la institución y pueden incluir bibliotecarios científicos, gestores de datos, científicos de datos y otros puestos de trabajo centrados en los datos. Muchos patrocinadores de la investigación solicitan DMPs en las propuestas de financiación. Por lo tanto, hay una mayor

atención a la forma en que los investigadores gestionan los datos de sus proyectos a lo largo del ciclo de vida de los datos. Por lo tanto, para seguir el ritmo de los mandatos de los financiadores y las nuevas expectativas de ciencia abierta, es necesario contar con especialistas dedicados y formados en la curación de datos que garanticen que los datos son utilizables y accesibles en un futuro indefinido. En cualquier punto del ciclo de vida de los datos, "la curación de datos proporciona funciones distintivas separadas de los métodos que los científicos llevan a cabo para la ciencia" y requiere la necesidad de un personal dedicado que puede estar encargado de muchas otras tareas antes y después de otros análisis. A menudo, partes del ciclo de vida de los datos [por ejemplo, la acuñación de un identificador de objeto digital (DOI)] son realizadas por gestores de datos en grandes instituciones después de la creación de los datos. En ese momento, otros gestores de datos normalizan los metadatos para que cumplan las normas institucionales. (Bishop et al., 2021, p. 1387)

Para tener una idea clara de las funciones que puede llegar a desempeñar el bibliotecario de datos, es necesario analizar el proceso que siguen los datos dentro de una biblioteca y cómo los bibliotecarios de datos suelen dar atención a las necesidades de este ciclo.

En primera instancia tenemos que los bibliotecarios de datos deben estar presentes durante todo el proceso de la investigación, guiando a los investigadores en consecuencia, una comunicación que se plantea por ambas partes debe ser lo más concreta posible.

En el primer paso, la importación de datos, los estudiantes e investigadores deben identificar el conjunto de datos adecuado para las preguntas que se plantean. En este paso es esencial tener un conocimiento práctico de los repositorios de datos y otras bases de datos. (Roy y Shollerb, 2018, p. 113)

En segunda instancia se debe poner orden a los datos, así como establecer una estructura de trabajo que ayude a determinar de qué manera se llevará a cabo el orden de los datos en la investigación.

La ordenación de los datos, el analista se encarga de la tarea de limpiar y formatear el conjunto de datos para prepararlo para su exploración y análisis. Los bibliotecarios de referencia pueden trabajar con los investigadores y los estudiantes para identificar los materiales adecuados para aprender a ordenar los datos. La etapa de ordenación se denomina a veces "manipulación de datos" (Roy y Shollerb, 2018, p. 114)

En tercera instancia se tiene la etapa de exploración, donde se deben examinar los datos para encontrar patrones o posibles hallazgos, por ello, es importante que en esta etapa las herramientas de datos sean familiares para el investigador y el bibliotecario de datos, también, deben entenderlas y trabajar cómodamente con ellas.

Las primeras etapas de la exploración suelen incluir la visualización de datos, un tema que ha recibido mucha atención a través de libros de texto y cursos. Indicar a los investigadores y a los estudiantes los tutoriales y las herramientas de visualización de datos les ayudará a descubrir las preguntas que pueden plantear a los datos y les proporcionará información para reformatear, limpiar y transformar de otro modo el conjunto de datos para un análisis específico. (Roy y Shollerb, 2018, p. 114)

En las últimas etapas de la exploración, se incluye un análisis que puede ser estadístico, o de otros tipos y dependerá de la investigación realizada.

En cuarta instancia se comunican los resultados de los análisis, así como los detalles de la investigación.

En este paso, los bibliotecarios de referencia pueden presentar a los investigadores y estudiantes ejemplos de cómo se ilustran normalmente los resultados del análisis de datos en un campo determinado, en una revista o en un formato de informe. Además, las llamadas "habilidades blandas" son una necesidad poco reconocida para los científicos de datos en activo. (Roy y Shollerb, 2018, p. 114)

Los bibliotecarios de datos garantizan que el estudiante, investigador o encargado de analizar los datos conozca las formas más apropiadas de citar los conjuntos de datos.

Por último, documentar el trabajo en ciencia de datos, es tan importante como la ciencia de datos en sí misma, aunque siempre se preste mayor atención a las partes técnicas, ignorando el proceso que se tuvo para obtener los resultados.

Documentar el flujo de trabajo, es decir, todos los pasos de la ciencia de datos, es esencial para garantizar que los análisis resultantes puedan reproducirse o revisarse. Y lo que es más importante, documentar cada paso también ayuda al investigador o al estudiante (a) acelerar futuros análisis, ya que la documentación actúa como una especie de tutorial, (b) identificar posibles errores o equivocaciones en las etapas de importación, ordenación y exploración, y (c) garantizar que la rotación en un equipo de investigación no prohíba el uso futuro de los datos y el software. (Roy y Shollerb, 2018, p.118)

Comprender cómo interactúan los datos entre sí, en su ciclo de vida, ayuda al bibliotecólogo a entender las necesidades de investigadores nuevos en esta área de los datos.

Dada la experiencia de los bibliotecarios en la gestión del conocimiento y su papel de larga data en el ciclo de investigación, los bibliotecarios están bien situados para proporcionar orientación e incluso colaborar con equipos de investigación en varios aspectos de la gestión y la ciencia de datos. (Federer, 2016, p. 36)

Y aunque técnicamente podría estar más o menos claro el papel que desempeña el bibliotecario de datos en todo este contexto se llega a confundir si tal trabajo de datos atañe a informáticos, bibliotecarios de datos o ingenieros de la información.

Muchos científicos han recurrido a la Biblioteconomía y Ciencias de la Información (LIS) en busca de orientación sobre cómo gestionar sus datos. Sin embargo, existe una confusión entre los investigadores principales, los bibliotecarios o los gestores de datos y los profesionales de la tecnología de la información sobre cómo deben distribuirse estas tareas de gestión de datos de investigación (RDM), es decir, quién debe hacer qué. (Bishop et al., 2021, p. 1388)

Estas funciones del bibliotecario de datos y su convivencia con otras profesiones son un asunto complejo que, incluso, puede variar, de acuerdo con la región, los servicios prestados en su biblioteca y los profesionales a los que se les brinda dicha atención. Es imprescindible diferenciar las funciones del bibliotecario de datos de la de los demás profesionistas con los que trabajará en conjunto, es esta razón por la que Camille y Urban consideran que las funciones del bibliotecario de datos son las siguientes:

Curación de datos se refiere a organizar, mantener, integrar y extraer conocimiento de los datos para aplicarlo y reutilizarlo.

Curación digital es más amplia y se define como la gestión activa y la mejora de los activos de información digital para uso actual y futuro.

Alfabetización de datos es la capacidad de leer, crear y comunicar datos como información. La instrucción de alfabetización de datos es a menudo el primer paso para apoyar a los investigadores, fomentar prácticas de datos actualizadas y hacer crecer los servicios de datos.

Preservación digital, que se define como la serie de actividades gestionadas necesarias para garantizar el acceso continuo a la información digital durante el tiempo que sea necesario.

Preservación de datos es una extensión de la preservación digital, ya que los datos digitales se degradan rápidamente y requieren controles de validación frecuentes, copias de seguridad remotas y metadatos antes del almacenamiento. (Camille y Urban, 2018, p. 403)

A esto es necesario agregar que algunos países como Alemania han estado implementando otros perfiles bibliotecarios de trabajo, en los que se observan funciones mayormente relacionadas con las tecnologías, por ejemplo:

Bibliotecario de sistemas es una especialización profesional que tiene tanto tecnologías de la información como tareas bibliotecarias (con un enfoque en sistemas bibliotecarios).

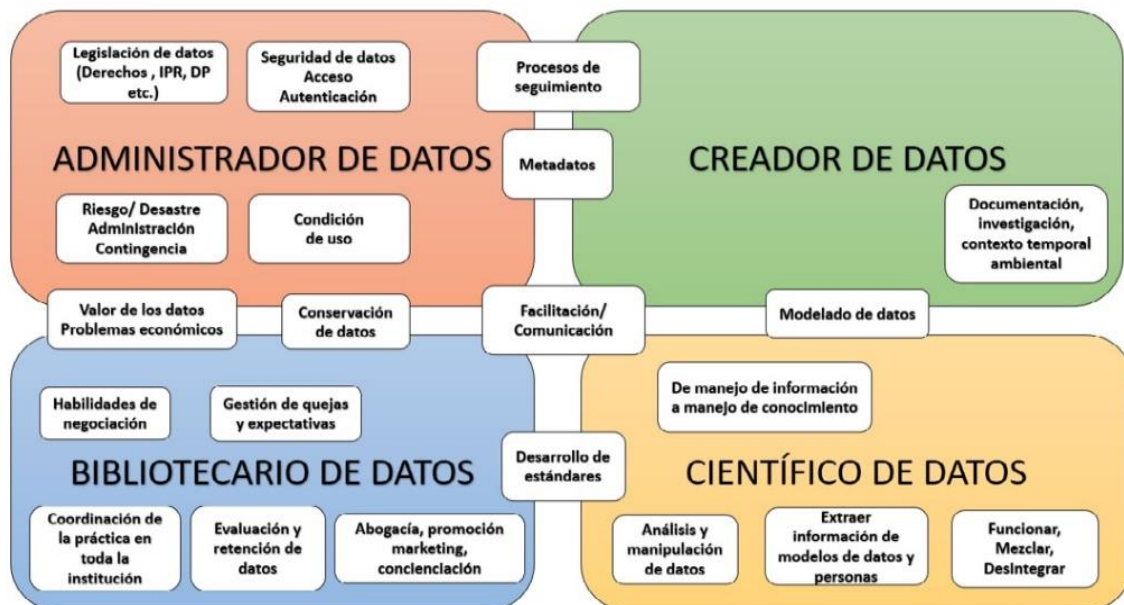
Bibliotecario de datos: es un especialista en datos para el ámbito de los sistemas computacionales.

Especialista en informática bibliotecaria: los especialistas en informática bibliotecaria no son un término alternativo para los sistemas bibliotecarios. Podrían considerarse como un nivel de educación superior para los sistemas de bibliotecarios, ya que sus habilidades de TI son aún más pronunciadas. (Caspers, 2016, p. 12)

Aunque estos puestos de trabajo requeridos en Alemania comparten funciones con los programadores, mantienen un estrecho lazo también con los bibliotecarios, debido a que los dos usan con demasiada frecuencia las Tecnologías de la Información y las Comunicaciones (TIC) para proporcionar sus servicios. Martin Donnelly, describió mejor esta idea y propuso un cuadro en el que se establecen las funciones de distintos profesionales que laboran en el universo de los datos, este se muestra en la Figura 6:

Figura 6: Habilidades básicas para la gestión de datos.

Fuente: Martin Donnelly, 2009, p. 165.



Al observar el cuadro anterior, es evidente que el bibliotecario de datos suele desempeñar más funciones, aunque la mayoría de estas puedan considerarse de carácter blando, dentro de su entorno, esto implica que el bibliotecario de datos actúa como un mediador entre las demás

partes, así como un conector entre sus colegas y compañeros de trabajo, sin embargo, esto no quiere decir que deba prescindir de sus habilidades técnicas.

Desde otro punto de vista, Spinak menciona que las funciones del bibliotecario de datos tendrán que ver con:

- Datos y tecnologías geoespaciales
- Lenguajes de programación
- Paquetes estadísticos (SPSS, Stata, SAS)
- Gestión y preservación de datos
- Redacción de solicitudes de fondos de investigación
- Gestión de proyectos
- Pensamiento crítico y resolución de problemas
- Empaquetado de datos, reutilización de datos, codificación de datos
- Repositorios institucionales y metadatos
- Tecnologías actuales online
- Análisis de datos cualitativos y cuantitativos. (Spinak, 2019, párr. 5)

Al ser partícipe activo del ciclo de vida de los datos, se agrega a esta lista el conocimiento sobre bases de datos y de ser posible bases en ciencias de datos, puesto que los expertos en este tema han aludido que, para poder desempeñar con mayor eficacia sus funciones como bibliotecarios de datos tuvieron que tomar una especialización en ciencia de datos o cursos de ciencias de datos.

Para los trabajos de datos científicos, el futuro está aquí, ya que los trabajadores ayudan a los científicos a encontrar, acceder y hacer interoperables los datos accionables por máquinas en todos los dominios y organizaciones para facilitar la ciencia. (Bishop et al., 2021, p. 1386)

Entre las funciones del bibliotecario de datos, las que más destacan, por diferenciarse de lo que un bibliotecario tradicional suele hacer, son las actividades que tienen que ver con programación o software. Tales actividades, de acuerdo con varios bibliotecarios de datos, son labores que se aprenden rápidamente debido a la experiencia con la gestión de la información, la interacción constante con otros profesionales de tecnologías y el

ordenamiento de datos en catálogos adquirida durante los años de formación en Bibliotecología.

Los conocimientos, habilidades y destrezas necesarios para desempeñar las nuevas funciones se aprenden probablemente en el trabajo y no en la educación formal. Con cualquier especialización emergente, se anticipa el aprendizaje experimental, pero, a medida que estas funciones en los datos científicos se vuelven más comunes, los cursos, certificados y otra educación más formalizada abordarán las nuevas necesidades educativas. (Bishop et al., 2021, p. 1386)

Un estudio interesante, que abarcó esto, fue realizado por Bishop y su equipo de investigación, quienes lograron que los bibliotecarios de datos sintetizaran varias de sus funciones, separándolas en dos grupos: las funciones que más solían realizar y las funciones que menos solían realizar. Algunas de las funciones que los bibliotecarios de datos solían hacer con menos frecuencia se presenta en la Tabla 1:

Tabla 1: Funciones que los bibliotecarios de datos desempeñan con menor frecuencia.

Fuente: Bishop et al., 2021, p. 1389.

Tipo	Herramientas
Lenguajes de programación (7)	Latex, MATLAB, PERL, Python, R, PHP
Suites (7)	Google, Microsoft Office
SIG (3)	ArcGis, GeoServer, QGIS
Bases de datos (5)	Intergraph Metadata, Open Refine, Scholarsphere, SPARQL, SQL, SQLite, editor XML, E-checker, base de datos Postscript
Canales de comunicación (3)	Google Hangouts, Slack, WebEx, Zoom
Otros (7)	Cyberduck, Nvivo, Osprey, VPN, JIRA, WordPress, Drupte, otro software basado en la nube

Entre las tareas que los bibliotecarios realizaron con más frecuencia se encuentran las siguientes:

Comunicación: la tarea laboral más frecuentemente descrita, fue la comunicación, apareció en todas las entrevistas. Aunque la comunicación se da y es importante en todas las profesiones, los participantes destacaron estas "habilidades blandas" por encima de otros conocimientos de dominio o habilidades aplicadas. Las habilidades de comunicación incluyen la comunicación verbal y escrita, el "don de gentes" en general y la capacidad de manejar situaciones de estrés. Los participantes afirmaron que las habilidades blandas son "muy poco valoradas por mucha gente". Porque "la gente suele estar estresada". Seis participantes comentaron la importancia de comprender los métodos científicos para poder comunicarse mejor con los científicos a los que sirven. Al actuar como enlace con comunidades científicas específicas, tanto los gestores de datos como los

bibliotecarios reúnen piezas de información necesarias para completar las tareas específicas de los datos, y las habilidades de comunicación son omnipresentes en todos los aspectos de estos trabajos. Así lo mencionó uno de los participantes en dicho estudio.

Colaboración: los participantes afirmaron que celebran reuniones iniciales al comienzo de un proyecto para que todos estén de acuerdo y planifiquen todo, en lugar de depender constantemente de los correos electrónicos a lo largo del proyecto. Esta comunicación interna permite a los gestores de datos llevar a cabo su trabajo de forma coherente entre los distintos miembros.

Tareas relacionadas con los datos: diez participantes señalaron que las tareas relacionadas con los datos son componentes importantes de su trabajo, y siete de ellos mencionaron las tareas de conservación de datos, que incluyen la limpieza y la comprobación de los datos, el procesamiento de los datos para cargarlos en sitios web y la introducción de datos. (Bishop et al., 2021, p. 1386)

Los autores enfatizaron que muchos de los bibliotecarios de datos eran en realidad profesionales en formación en este nuevo mundo lleno de posibilidades sobre los datos. En este sentido, aseguraron que los profesionales de diversos campos, especialmente bibliotecarios, se formaban mediante cursos académicos que ya están comenzando a surgir, o juntando los conocimientos necesarios que responden a las funciones del puesto. Aunado a ello, lo que más ha ayudado a estos nuevos profesionales en la búsqueda de su formación como bibliotecario de datos, es el afán de seguir aprendiendo, manteniéndose en continua preparación para lo que el nuevo mercado de datos requiere.

La profundidad del aprendizaje que se produce en la formación en el puesto de trabajo es mucho más rica que la obtenida en los cursos optativos, pero cualquier KSA “conocimiento destreza y habilidad” aprendido antes o junto con la experiencia laboral capacita y prepara a los estudiantes para los puestos de trabajo de hoy en día. Las preocupaciones tanto de los instructores como de los que ocupan puestos de trabajo actuales surgen en cualquier momento en que los avances de la tecnología afectan a lo que hay que saber en estas carreras dinámicas. (Bishop et al., 2021, p. 1386)

Independientemente de cómo se logre ser un bibliotecario de datos, ya sea por la experiencia o por medio de la academia, es importante que esta profesión se formalice un poco más y se le dé un carácter más estructurado y homogéneo, de lo contrario se caería en el error de cubrir funciones sin estructura, que, si bien se pueden hacer, no se sabe el propósito ni tampoco el papel que se cumple o se podría cumplir dentro de otros centros de trabajo. Por lo tanto, es pertinente que exista un perfecto balance entre preparación académica y preparación laboral.

3.2.3 Habilidades

El ser un bibliotecario de datos, es una nueva vacante de trabajo para la cual se necesitan de nuevas e interesantes habilidades, que actualmente están requiriéndose dentro del campo laboral. Entonces, si al bibliotecario le interesa este nuevo campo laboral, para el que, por supuesto, está destinado, debe dejar de ser un proveedor pasivo de información o datos y pasar a ser un participante activo en el proceso.

La Bibliotecología de datos requiere habilidades académicas y profesionales, incluidas las áreas tradicionales de la biblioteconomía, como servicio de referencia y apoyo a la investigación académica. Hoy en día, los bibliotecarios de datos necesitan habilidades de datos como transformación de datos, manipulación de datos y análisis de datos. Un bibliotecario de datos necesita combinar habilidades y conocimientos relevantes para la gestión y conservación de datos. (Semeler y Pinto, 2020, p. 2)

Es cierto que en este nuevo contexto se necesitan conocimientos y habilidades en el campo de los datos: curación, representación, organización, recuperación, servicios de datos, entre otros conocimientos, sin embargo, la mayoría de estas habilidades no son nada nuevo para el bibliotecario o al menos no son algo desconocido.

Muchas de las habilidades que los gestores de datos declararon eran habilidades blandas, como el trabajo en equipo, el intercambio de conocimientos y la comunicación. Otro estudio encontró que el conocimiento del dominio de la ciencia es necesario para trabajar con datos científicos, pero igualmente importantes son las habilidades de gestión de datos relacionadas con la comunicación, la gestión de proyectos, la creación de comunidades y las tareas específicas de los datos. (Bishop et al., 2021, p. 1388)

Dentro de las descripciones del puesto dadas por los mismos bibliotecarios de datos, se encuentran que el conocimiento sobre datos es esencial en sus funciones ya que estas se complementan con la gestión de proyectos, búsqueda de información, conocimientos de recuperación, entre otros. Tal argumento fue comprobado por Federer en sus primeros estudios, donde separó a los bibliotecarios de datos en “especialistas en la materia”, que se especializaban en un pequeño número de tareas, y “generalistas de datos” que trabajaban en múltiples campos.

Los bibliotecarios de datos también deben poseer las mismas habilidades de desarrollo estándar que un científico de datos. Las habilidades requeridas de un científico de datos incluyen las de un hacker, científico, analista cuantitativo, asesor de confianza y empresario experto. Estas habilidades no requieren que el

bibliotecario de datos realice un curso de informática, estadística o negocios; sin embargo, él / ella necesita tener al menos un nivel intermedio de conocimiento en estas áreas. Los bibliotecarios de datos deben comprender cómo integrar múltiples tecnologías de datos en los servicios bibliotecarios. Él / ella necesita la capacidad de demostrar habilidades de programación para programar, extraer y estructurar datos. Él / ella también necesita conocer los métodos y herramientas básicos adoptados en los estudios de métricas (bibliométricos e informétricos). (Semeler y Pinto, 2020, p. 5)

Las habilidades requeridas en los bibliotecarios de datos también son definidas a partir de las necesidades de su institución y varían drásticamente, dependiendo de las funciones en las áreas de investigación o la mera gestión de los datos.

Trabajar con datos es una labor de vital importancia al momento de entregar resultados o de hacer nuevos descubrimientos, pero para que esto suceda deben estar varias partes involucradas en la investigación, así como habilidades en conjunto, es esta razón por lo que a pesar de que los puestos de trabajo ofrecen expectativas de formación, experiencia profesional y tareas, estos suelen ser visiones de los candidatos ideales no de los candidatos reales. “Para muchos empleadores, lo ideal sigue siendo una combinación de conocimientos de biblioteconomía y documentación y de conocimientos específicos del campo de los datos, pero sin cursos especializados en el campo o en la biblioteconomía y documentación, seguirá habiendo una escasez de candidatos cualificados” (Bishop et al., 2021, p. 1398).

Se considera que un bibliotecario de datos es un perfil multifacético que debe contar tanto con habilidades técnicas como lo son la programación, la estadística y la ciencia de datos, como con las habilidades blandas, ejemplo la comunicación, el formarse de manera autodidacta, siendo esta última de acuerdo a los encuestados, la que los ayudó a mantenerse en su puesto e incluso sobresalir como bibliotecario de datos, en especial si se adquirían habilidades específicas acordes a los proyectos que surgían.

Según el participante, 005 las herramientas dependían de con quién trabajaba y de lo que hacían. "Muchas de mis herramientas son cosas para averiguar cómo conseguir las herramientas que estaban utilizando y familiarizarme con ellas". Aunque muchos aprendieron estas habilidades técnicas mientras trabajaban, seis participantes mencionaron que desearían haber adquirido esa experiencia antes de comenzar su carrera profesional. (Bishop et al., 2021, p. 1392)

Otras habilidades como la conservación de datos, las capacidades de descripción y documentación de datos son señaladas como útiles, junto con las habilidades legales, políticas y de asesoramiento aunado a una variedad de métodos de investigación.

Esto podría indicar la necesidad de que algunos bibliotecarios académicos se especialicen en conocimientos de datos o se conviertan en bibliotecarios especializados en datos. En otro estudio realizado por Lisa Federer, se encuestó a los bibliotecarios que realizan trabajos relacionados con los datos sobre su trabajo y formación y se les pidió que valoraran la importancia de una serie de habilidades y conocimientos relacionados con los datos para su trabajo. La sorpresa fue que, aunque las competencias tradicionales de los bibliotecarios no fueron muy valoradas por este grupo, las "competencias blandas" sí lo fueron. (Rice, 2019, p. 280)

Las habilidades blandas, como la comunicación, el trabajo en equipo, el aprendizaje autodidáctico, son en casi todos los estudios hechos a bibliotecarios de datos, algunas de las más usadas junto con las habilidades técnicas así es que, de acuerdo con el estudio hecho por Rice de esta manera quedaron valoradas por los profesionistas las siguientes habilidades:

- Los "atributos personales" fueron la categoría más valorada en general (el 70 % de los encuestados los calificaron de muy importantes)
- Las "habilidades bibliotecarias" fueron la categoría peor valorada (40 %). Tal vez estos expertos en datos consideren que simplemente hay una brecha demasiado grande entre cómo los bibliotecarios suelen utilizar sus habilidades (por ejemplo, el trabajo de referencia y la catalogación) y cómo este tipo de habilidades se pueden utilizar en el apoyo a los datos. Sería interesante ver si esta percepción cambia con el tiempo, si los bibliotecarios aprenden las habilidades necesarias para el soporte de datos en las escuelas de biblioteconomía del futuro.
- Los cinco elementos mejor valorados fueron: "Desarrollo de relaciones con los investigadores, el profesorado, etc."; "Habilidades de comunicación oral y de presentación"; "Trabajo en equipo y habilidades interpersonales"; "Habilidades de comunicación escrita"; y "Consulta o instrucción individual" (Rice, 2019, p. 280)

Otros estudios han arrojado algo similar, categorizando como importante en la práctica del bibliotecario de datos.

Las competencias de la categoría "Atributos personales" (como las habilidades interpersonales, escritas y de presentación) fueron calificadas como las más importantes, mientras que las de la categoría "Habilidades bibliotecarias" fueron las menos importantes. Un análisis de conglomerados detectó dos grupos que podrían describirse mejor como especialistas en materias y generalistas de datos. Los especialistas en la materia se centran en un número menor de disciplinas y consideran que un número menor de tareas son importantes para su trabajo en comparación con los generalistas de datos. Además, los generalistas de datos son más propensos a decir que dedican la mayor parte de su tiempo al trabajo relacionado con los datos. (Goldstein, 2018, p. 79)

Las habilidades requeridas para desarrollarse como un buen bibliotecario de datos tienen como eje central el poder investigar, encontrar, facilitar e interpretar la información, habilidades que dentro de la práctica bibliotecológica suelen ser comunes.

Los investigadores, que hace tiempo dejaron de utilizar los servicios de referencia para realizar sus propias búsquedas en Internet, buscan ayuda profesional para la gestión de sus datos de investigación (RDM), y cada vez encuentran más esa ayuda experta en los bibliotecarios. En este sentido, todos los bibliotecarios se están convirtiendo, en cierto modo, en bibliotecarios de datos. (Rice, 2019, p. 265).

Adicional a esto, se encuentran otras características destacables en el bibliotecario de datos:

- Habilidades en comunicación y facilitación.
- Desarrollo de estándares.
- Selección y evaluación de datos.
- Habilidades de negociación.
- Abogacía, promoción y marketing.
- Cuestiones económicas relacionadas con el valor de los datos. (Giannini y Molino, 2018, p. 18)

En cuanto a tecnologías, lo que se busca para que un bibliotecólogo pueda ser considerado un bibliotecario de datos es que tenga entendimiento de estadística enfocada a ciencia de datos, conocimiento en centros de datos existentes, repositorios y mecanismos de descubrimiento de datos, estándares y esquemas de metadatos, los formatos de datos, las ontologías del dominio, las herramientas de descubrimiento, lenguajes de programación especiales para datos como son Python, R y gestores de datos como MySQL.

Quizás parezcan excesivas las habilidades que demanda el mercado laboral para poder ser un bibliotecario de datos, pero inequívocamente esto ya se había anticipado años atrás, cuando el auge de las tecnologías empezó, dando a entender que el bibliotecario del nuevo siglo debía acercarse a un conjunto de roles y saberes, dichos saberes se mencionan a continuación:

Familiarizado con el aprendizaje electrónico: los bibliotecarios deben participar en las prácticas de aprendizaje electrónico; colaborar con miembros de la facultad y enseñar el uso de la biblioteca utilizando la plataforma de aprendizaje electrónico.

Experto en MS Office (Microsoft Office).

Un bloguero: empieza por construir un blog personal.

Un abrazador de las nuevas tecnologías: capaz de considerar nuevas formas de promover la lectura.

Un ciudadano digital: en lugar de un ciudadano analógico.

Un miembro de los foros en línea de bibliotecarios: para la colaboración y la lluvia de ideas.

Un cybrarian: un experto en localizar información en Internet.

Un creador de sitios web: que ayuda a editoriales en la comercialización de los servicios bibliotecarios.

Un catálogo en línea: Por ejemplo, el de la Biblioteca del Congreso. (Eke, 2010, p. 5)

Estas habilidades, anticipadas años atrás, se observan útiles e indispensables para la conformación actual del perfil profesional del bibliotecario de datos.

Finalmente, en cuanto a habilidades más proactivas y personales, el perfil profesional del bibliotecario de datos contempla que sea:

- Autodidacta: como es una profesión que está en formación se debe estar actualizando constantemente, aprendiendo e investigando, un bibliotecario de datos debe aprender por su cuenta, cuestionarse y saber implementar ese conocimiento para ayudar a sus usuarios.
- Innovador: esta profesión necesita de gente creativa que pueda proponer.
- Solucionador: ser bibliotecario de datos significa todo un reto, a nivel profesional, puesto que está cubriendo carencias que apenas están surgiendo y por las cuales se necesitan descubrir sus soluciones.

En el momento en el que la tecnología comenzó a cambiar el mundo, inició también el cambio en los profesionales de la información y otros profesionales. En consecuencia, los nuevos requisitos y habilidades que se piden para el bibliotecario de datos no es algo que deba generar sorpresa o inquietud. Hoy en día, no solamente los profesionales de la información tienen ante sí el reto tecnológico entre sus habilidades curriculares; tal hecho es algo que, a nivel competitivo, se le está pidiendo a profesionistas de toda índole.

En un discurso del 9 de septiembre de 2008, Obama señaló: “Sin una fuerza laboral capacitada en matemáticas, ciencia y tecnología, y otras habilidades del siglo XXI, nuestras empresas innovan menos, nuestra economía crecerá menos y nuestra nación será menos competitiva. Si

queremos competir con el mundo de mañana, debemos educar al mundo de hoy” (Eke, 2010, p. 5)

Esto, lejos de presentar una limitante en el desarrollo del bibliotecólogo, debe suponer un reto que ayude en el crecimiento de la profesión, sin contar también que, desde donde sea que se inicie, no se debe dejar nunca de aprender.

3.2.4 Educación

Actualmente la profesión del bibliotecario de datos se encuentra en construcción y delimitación, por lo tanto, a pesar de su demanda, existen pocos centros educativos que se dediquen a estructurarla como una profesión. A partir de una revisión de los planes de estudio que existen en diversas academias en México, se encontró que no existe una formación como bibliotecario de datos o una especialización, en cambio, sus planes de estudio contienen ciertas materias que podrían relacionarse con el perfil del bibliotecario de datos, así, con el fin de encontrar categorías que pudieran tener relación con esta formación, incluyendo, las habilidades blandas, se analizaron varios planes de estudio, de los cuales a continuación se dará una breve descripción de lo que se encontró en cada uno de ellos.

-Licenciatura en Bibliotecología y Estudios de la Información. Universidad Nacional Autónoma de México. (UNAM)

Este plan de estudios se compone por 51 asignaturas, 360 créditos y una duración de 4 años, al analizar el mapa curricular (Anexo 1) se observa que está relacionado en un 15% con el perfil del bibliotecario de datos, las asignaturas que guardan relación directa con dicho perfil son las siguientes:

- Computación aplicada a la Bibliotecología
- Tecnologías de la información en bibliotecas
- Recursos de información
- Bases de datos
- Servicios bibliotecarios y de información
- Telecomunicaciones en las unidades de información
- Métodos de investigación cuantitativos
- Recursos digitales y multimedia

- Usuarios de la información
- Organización documental
- Servicios al público
- Políticas de información

-Licenciatura en Biblioteconomía. Escuela Nacional de Biblioteconomía y Archivonomía. (ENBA)

Este plan de estudios se compone de 9 semestres, con 2 ciclos de formación del 1° al 5° semestre corresponde a los estudios del Profesional Asociado en Biblioteconomía, el segundo abarca del 6° al 9° complementando la formación de Licenciado en Biblioteconomía. Al analizar el mapa curricular (Anexo 2) se observa que está relacionado en un 25% con el perfil del bibliotecario de datos, las asignaturas que guardan relación con este perfil son las siguientes:

- Servicios al público
- Herramientas prácticas de la computación
- Adquisición de recursos de información
- Usuarios de la información
- Estadística descriptiva
- Conversión y migración de información
- Políticas de información
- Industria de la información
- Introducción a la Bibliometría
- Almacenamiento y recuperación de la información
- Bibliotecas especializadas

-Licenciatura en Bibliotecología y Gestión de Información. Universidad Autónoma de Chiapas. (UNACH)

Este plan de estudios semestral con enfoque humanista está compuesto de 9 semestres y 434 créditos, en su mapa curricular (Anexo 3) se observa que está relacionado en un 26% con el perfil del bibliotecario de datos, las asignaturas que guardan relación son las siguientes:

- Herramientas computacionales
- Introducción a la gestión aplicada a las unidades de información
- Diseño de recursos informáticos
- Bases de datos
- Servicios especializados de información
- Bibliometría
- Métodos de investigación cuantitativos aplicados a la Bibliotecología
- Usuarios de la información
- Automatización de los servicios en las unidades de información
- Organización de unidades y sistemas de información
- Formación y educación de usuarios
- Seminario de integración: gestión, organización y tecnologías

-Licenciatura en Gestión de la Información y Recursos Digitales. Universidad Autónoma de Nuevo León. (UANL)

El siguiente plan de estudios es anual y se compone de 10 semestres con 9 opciones de titulación, en el mapa curricular (Anexo 4) se observa que está relacionado en un 46% con el perfil del bibliotecario de datos, las asignaturas que guardan relación con dicho perfil son las siguientes:

- Fundamentos de la investigación científica
- Competencia comunicativa
- Educación socioemocional
- Información y calidad en las organizaciones
- Ciencia de datos
- Sistemas de gestión de información
- Recursos de información
- Métricas de la información y los datos
- Estándares tecnológicos de información
- Ética y cultura de la legalidad
- Diseño de bases de datos

- Analítica web
- Repositorios
- Didáctica de la ciencia de la información
- Desarrollo de bases de datos
- Minería de datos y textos
- Metadatos y colecciones digitales
- Metodología cuantitativa
- Parámetros y calidad de bases de datos
- Liderazgo, emprendimiento e innovación
- Desarrollo de productos de investigación
- Seminario de tecnología de la información
- Arquitectura de la información en ambientes web

-Licenciatura en Bibliotecología y Gestión del Conocimiento. Universidad de Guadalajara. (UdeG)

Este plan de estudios se compone de 8 semestres, de acuerdo con su mapa curricular (Anexo 5) se observa que se relaciona en un 17% con el perfil del bibliotecario de datos, las asignaturas que guardan relación con dicho perfil son las siguientes:

- Organización de la información
- Diseño de estrategias de servicios de información
- Comunicación y difusión de información y conocimientos
- Recuperación de la información
- Gestión de la información en las organizaciones
- Aplicación de la estadística en el uso de la información
- Toma de decisiones

Licenciatura en Ciencias de la información Documental. Universidad Autónoma del Estado de México. (UAEM)

Este plan de estudios se compone de 10 semestres, con duración de 5 años, de acuerdo con su mapa curricular (Anexo 5) se observa que se relaciona en un 16% con el perfil del

bibliotecario de datos, las asignaturas que guardan relación con dicho perfil son las siguientes:

- Alfabetización informacional
- Metodología para la investigación en ciencias de la información documental
- Normalización de la información
- Elaboración de textos académicos
- Análisis de la información contemporánea
- Mercadotecnia de productos y servicios de la información
- Estadística para documentalistas

Los anteriores planes de estudio contienen materias relacionadas con las bases de datos o incluso a la ciencia de datos, sin embargo, las materias que guardan relación con el perfil del bibliotecario de datos están en porcentajes por debajo del 50% de todo el mapa curricular, es decir, aún no son suficientes para formar a un bibliotecario de datos tal como demanda la época actual.

En cuanto a educación menos formal, existen diversas opciones que contribuyen a la formación del bibliotecario de datos. Una de ellas es el *Research Data Management Librarian Academy* (RDMLA). “El RDMLA es un programa gratuito en línea que fomenta el desarrollo de habilidades profesionales de gestión de datos de investigación para bibliotecarios, profesionales de la información, investigadores y otros profesionales que trabajan en un entorno de investigación intensiva en todo el mundo” (Spinak, 2019, párr. 9).

RDMLA presenta una asociación única entre un programa académico LIS, ciencias académicas de la salud y bibliotecas de investigación, y Elsevier. Las instituciones asociadas incluyen: Harvard Medical School, Harvard Library, Simmons University, Boston University, Brown University, MCPHS, Northeastern University, Tufts University, DataONE, National Taiwan University Library, National Institutes of Health y Indiana University-Purdue University Indianápolis. RDMLA cuenta con el patrocinio financiero de Elsevier. (Research Data Management Librarian Academy, 2019, párr. 2)

Dicho curso está diseñado para que sus alumnos obtengan el conocimiento y las habilidades esenciales para ayudar a los investigadores en la gestión de sus datos, algunas de las materias que imparte son:

Fundamentos de la gestión de datos de investigación (research data management, RDM).

Navegación en la cultura de los datos de investigación.

Abogar y comercializar el valor de RDM en las bibliotecas.

RDMLA Gestión y evaluación de proyectos. (en relación con los servicios RDM)

Descripción general de las herramientas de análisis y visualización de datos.

Descripción general de las herramientas de codificación.

Descripción general de las herramientas de la plataforma. (Open Science Framework, Mendeley Data y otras). (Spinak, 2019, párr. 10)

Otro de los talleres de *Research Data Management Librarian Academy* (RDMLA) es el proyecto RDMRose, este reúne a la Universidad de *Sheffield* con una comunidad de profesionales basada en las bibliotecas del Consorcio de la Universidad *White Rose* en las Universidades de *Leeds*, *Sheffield* y *York*.

RDMRose es un proyecto financiado por Jisc para producir materiales didácticos y de desarrollo profesional continuo (CPD) en RDM adaptados a los profesionales de la información. Está analizando en particular las necesidades específicas de los bibliotecarios de enlace en las bibliotecas universitarias, y los entregables incluyen materiales adecuados para el aprendizaje en múltiples modos, incluido el aprendizaje cara a cara y autodirigido. Todos los materiales pueden ser reutilizados por otros educadores de servicios de información y bibliotecas y hay disponible una versión para el desarrollo profesional continuo autosuficiente. (Guy, 2013, párr. 5)

Otro proyecto es el impartido por la Biblioteca de la Universidad de Exeter.

El proyecto *Holistic Librarian* ha implicado probar un curso autodirigido diseñado para ayudar a los bibliotecarios de materias de Exeter a familiarizarse con los conceptos y aspectos prácticos de RDM. La capacitación toma la forma de "23 cosas para la gestión de datos de investigación". Durante la versión piloto, se le pidió a cada bibliotecario de la asignatura que completara una variedad de "cosas" o tareas específicas o preguntas relacionadas con RDM. (Guy, 2013, párr. 10)

La Universidad de *East London* (UEL) también contribuye a la formación de bibliotecarios de datos mediante el proyecto *Training for Data Management* (TraD). Este proyecto

incorporará buenas prácticas en gestión de datos, mediante el desarrollo de material de capacitación disciplinaria en línea y presencial.

Los cinco módulos de SupportDM se desarrollarán utilizando Xerte y se entregarán a través de Moodle. Se basarán en los módulos MANTRA desarrollados por la Universidad de Edimburgo y el curso de capacitación en inteligencia de datos para personal bibliotecario desarrollado en la Universidad Tecnológica de Delft, Holanda. El curso Data4Librarians es un curso modular basado en competencias, combinando la enseñanza en línea y presencial. SupportDM incluirá recursos en línea y material del curso, notas de conferencias y consejos sobre la implementación y evaluación del curso en las bibliotecas. (Guy, 2013, párr. 12)

A partir de la revisión de la literatura sobre la educación del bibliotecario de datos, se detectó que son pocas las universidades que usan el título de bibliotecario de datos en sus cursos o especialidades, lo más común dentro de los programas de grado de Bibliotecología y Ciencias de la Información es ofrecer el aprendizaje y las herramientas propios, de un bibliotecario de datos, pero sin el título de bibliotecario de datos, ejemplo de ello sería la especialidad que ofrecen programas de certificación en Bibliotecología como (LIS) en Análisis / Diseño de Sistemas de Información:

La mayoría de los programas acreditados por la *American Library Association*, ofrecen "Análisis de datos", "Servicios de datos" o "Gestión del conocimiento e inteligencia competitiva". El plan de estudios de esta especialización incluye los siguientes cursos:

- Análisis y diseño de bases de datos.
- Análisis y diseño de sistemas de información.
- Modelado predictivo.
- Análisis y desarrollo de sistemas.
- Diseño y sistemas de gestión de bases de datos.
- Datos y algoritmos: Ética y Política.
- Planificación y gestión de minería de datos.
- Código y poder.
- Visualización de datos y comunicaciones interesadas.

Tal especialización, si bien no tiene el título de bibliotecario de datos, proporciona herramientas útiles y valoradas para poder cumplir con el puesto de trabajo de un bibliotecario de datos. Como señalan Semeler y Pinto:

El desarrollo de la Bibliotecología de datos está relacionado con los estudios de posgrado en biblioteconomía y ciencias de la información. La certificación en Bibliotecología de datos es interdisciplinaria. En este sentido, para adquirir conocimientos sobre biblioteconomía orientada a datos, existen algunos cursos ofrecidos por instituciones internacionales. La educación académica formal de los bibliotecarios de datos implica estudios a nivel de posgrado; asumimos que un bibliotecario de datos debe tener una amplia experiencia en investigación. La calificación y certificación en áreas relacionadas, por otro lado, se pueden adquirir a través de cursos que enseñan técnicas de aplicación de tecnología de datos, gestión de datos y curación de datos digitales. (Semeler y Pinto, 2020, p. 9)

En Viena existe un curso que ya ofrece el título de bibliotecario de datos y es ofertado por la Universität Innsbruck:

“El curso universitario "*Bibliotecario de datos*" transmite conocimientos que son relevantes para el desarrollo e implementación de servicios en el campo de la gestión de datos de investigación. El enfoque de este curso certificado es la comunicación científica y el apoyo a la investigación, políticas en el manejo de datos de investigación, planes de gestión de datos, metadatos en el campo de los repositorios, análisis de datos, agregación y enlace de datos, estándares de datos, modelado de datos, archivo a largo plazo y recuperación de datos. (Universität Innsbruck, 2020, párr. 2)

El costo de este curso es de 1400 euros y sus requisitos son un examen de elegibilidad o matrícula en Estudios de la Información, en las Universidades de Viena, Graz o Innsbruck.

La Universidad de Viena, en cooperación con la Universidad de Graz y la Universidad de Innsbruck, inició un curso certificado titulado "Bibliotecario de datos" El objetivo es ofrecerlo regularmente como un programa de capacitación en el futuro, ya sea en forma de un curso certificado y/o dentro del programa de estudio interuniversitario "Estudios de Bibliotecología e Información" (LIS), con el objetivo de una vez al año, dependiendo de la demanda y el número de registros. (Hubert, 2020, párr.3)

En los últimos años, la American Library Association, ofertó también el seminario “Datos en bibliotecas: procesamiento y visualización de datos” impartido por el Doctor Ryan Womack, actual bibliotecario de datos en las bibliotecas de la *Universidad de Rutgers*. El seminario estuvo dirigido exclusivamente a bibliotecarios y prometía enseñar sobre los servicios de

datos para bibliotecas, con un enfoque exhaustivo en las técnicas y herramientas de visualización de datos:

Este seminario web es la parte 3 de una serie de cuatro partes.

Sesión 1, Datos en bibliotecas, datos básicos y la entrevista de referencia.

Sesión 2, Datos en Bibliotecas, Datos Abiertos.

Comprensión de la gama y los tipos de servicios que las bibliotecas pueden proporcionar en relación con los datos.

Resultado de aprendizaje 1: Comprensión compartida de las necesidades de los asistentes a los seminarios web (basada en la entrada de encuestas en vivo) para los servicios de datos.

Resultado de aprendizaje 2: Identificar formas en que las bibliotecas pueden responder y satisfacer las necesidades de datos de sus usuarios, incluida una revisión del software estadístico común y sus beneficios y limitaciones.

Resultado de aprendizaje 3: Comprensión del papel de la visualización de datos en el análisis y la presentación de datos, y las herramientas y habilidades que respaldan la visualización de datos. (American Library Association, 2020, p. 3)

De la misma manera, otros cursos adicionales forman como bibliotecarios de datos:

Data Librarianship and Management de Steeves (2020) en Pratt Institute School of Information en los Estados Unidos es una clase académica abierta enfocada en la introducción de conceptos y terminología relacionados con datos y servicios de datos en bibliotecas. (Semeler y Pinto, 2020, p. 10)

Una tercera vía para educarse como bibliotecario de datos, que incluso varios autores han llegado a recomendar, es aprender de manera autodidacta, a través de cursos, ya sea de manera gratuita o con costo. Las opciones aquí incluyen (entre otras) Udacity (categoría de ciencia de datos), Coursera (ciencia de datos, *Big Data*, *Business Analytics*), O'Reilly Medía (datos, programación), Learnquest (datos, *Big Data*), Udemy (datos, análisis, Data Science Career Guide), Code Academy (múltiples opciones) y Lynda.com (*Big Data*, Minería de Datos, Visualización de Datos). Estos proveedores ofrecen opciones populares y de fácil acceso para el desarrollo de habilidades específicas, generalmente a un precio muy razonable.

En México se impartió por parte de la doctora Lourdes Feria Basurto, bibliotecaria mexicana egresada del doctorado en Ciencias de la Documentación de la Universidad Complutense de Madrid y pionera en temas de datos en bibliotecas el taller, "Siete buenas prácticas para la gestión de datos en las bibliotecas"

Por otro lado, también se facilitó por parte de la Dirección General de Bibliotecas, Universidad Nacional Autónoma de México el taller “El bibliotecario de datos: Propósitos y roles en acción” que fue impartido por los profesores Nicholas Cop y José Antonio Yañez. Para dicho taller se tomaron en cuenta los siguientes requisitos:

- Haber tomado el taller previo “*Big Data* para bibliotecarios académicos”, experiencia previa o conocimiento sobre datos de investigación, o tener responsabilidad relacionada con la gestión de datos.
- Identificar previamente a una comunidad de investigadores a entrevistar con el objetivo de tener la base para un Plan de gestión de datos.

Por último, y para concluir el bibliotecólogo podría tomar la licenciatura de Ciencia de Datos en la Universidad Nacional Autónoma de México (UNAM), dicha licenciatura es de ingreso indirecto y se debe aprobar un examen diagnóstico de conocimientos mínimos de matemáticas, computación, español e inglés. Realizar dos entrevistas, una con el Comité Académico y otra con el Coordinador de la Carrera.

También existe el diplomado en Ciencias de Datos impartido por la UNAM, el cual se compone de 15 módulos y una duración de 185 horas.

Con todo lo anterior se concluye que, la preparación académica en la profesión del bibliotecario de datos es una actividad de interés en la que día a día se abren nuevas opciones educativas que buscan dar respuesta a una naciente profesión.

3.3 Bibliotecario de datos: perfil profesional en la práctica actual

3.3.1 ¿Quién requiere y en dónde se requiere el perfil del bibliotecario de datos?

A partir del análisis de los 34 anuncios de empleo, se diseccionó el estudio en cuatro secciones, la primera muestra quién necesita y en dónde se necesita un bibliotecario de datos. Se graficó la clase de institutos que requieren un profesional de esta índole. El resultado mostró que la mayoría eran universidades con el 65%, en segundo lugar, se encuentra el sector privado con un 20% y por último las instituciones de gobierno con un 15% esto parece interesante debido a que en un principio se tenía la teoría de que el sector privado sería el que más requiriera de dicho perfil profesional.

Gráfica 1: Instituciones que requieren un bibliotecario de datos.

Fuente: Elaboración propia, 2021.

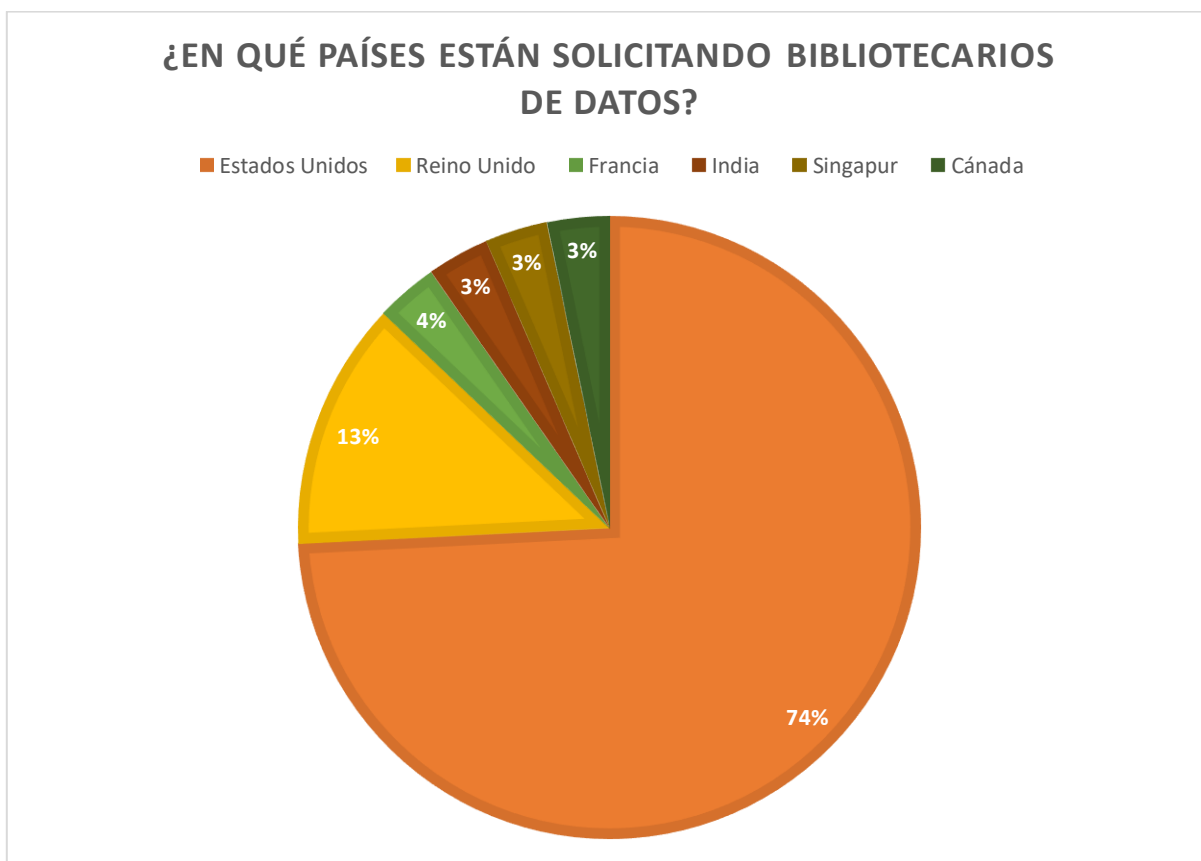


Para hacer frente a la segunda pregunta se consideró que dentro de los países que más solicitan este perfil se encuentra Estados Unidos con un 74%, seguido de Reino Unido con un 13% Francia con un 4% y al finalizar India, Canadá y Singapur con un 3%.

Por último, es pertinente señalar que no se encontró ninguna oferta reciente para bibliotecario de datos en Latinoamérica, esto resulta curioso debido a que bibliotecas de universidades como la Universidad Nacional Autónoma de México (UNAM) y el Instituto Politécnico Nacional (IPN), en sus bibliotecas, brindan servicios de datos a sus usuarios.

Gráfica 2: Países donde están solicitando bibliotecario de datos.

Fuente: Elaboración propia, 2021.

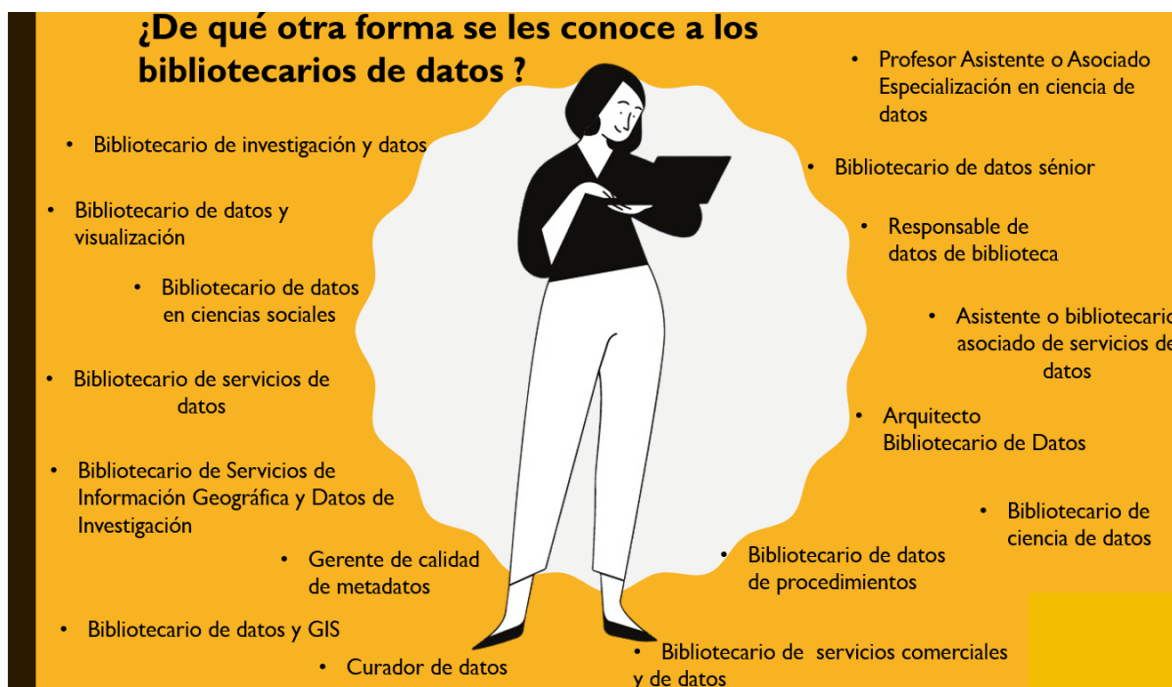


3.3.2 Otras formas en las que se les conoce a los bibliotecarios de datos

A lo largo de la investigación se encontró que el bibliotecario de datos es denominado de múltiples formas, aun cuando las descripciones del puesto y los requisitos eran los mismos. Estas se presentan en la Figura 7.

Figura 7: Diferentes nombres del bibliotecario de datos.

Fuente: Elaboración propia, 2021.



También cabe aclarar que a pesar de que los puestos y perfiles solicitados del bibliotecario de datos tengan distinta denominación para casi todos se solicita el mismo nivel de estudios y el mismo grado académico, como se explica detalladamente en el siguiente apartado.

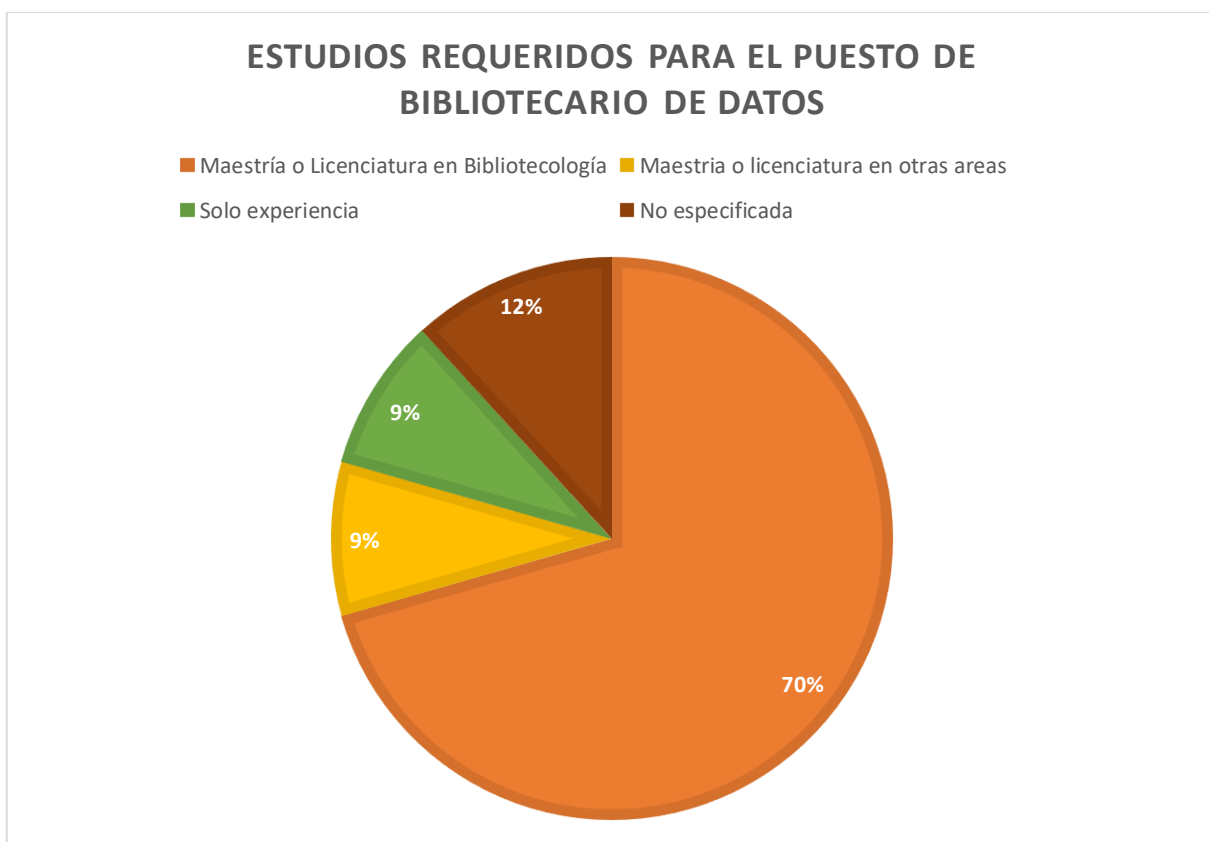
3.3.3 Requerimientos del puesto de bibliotecario de datos

Esta pregunta se contestó a través de tres categorías diferentes; estudios, requisitos extra y habilidades. Para el primer tópico se encontró que el 70% de los anuncios solicitaban un profesional con maestría o licenciatura en Bibliotecología o con alguna certificación acreditada por la *American Library Association*, el 9% de los anuncios podría considerar la postulación si solo se ofrecía experiencia y el otro 9% si se tenía una maestría o licenciatura

en alguna otra área que no fuera la Bibliotecología y por último, el 12% recomendaba contactarse con el anunciante para una entrevista personal, donde se le informaría sobre los requisitos académicos.

Gráfica 3: Estudios requeridos para el puesto de bibliotecario de datos.

Fuente: Elaboración propia, 2021.

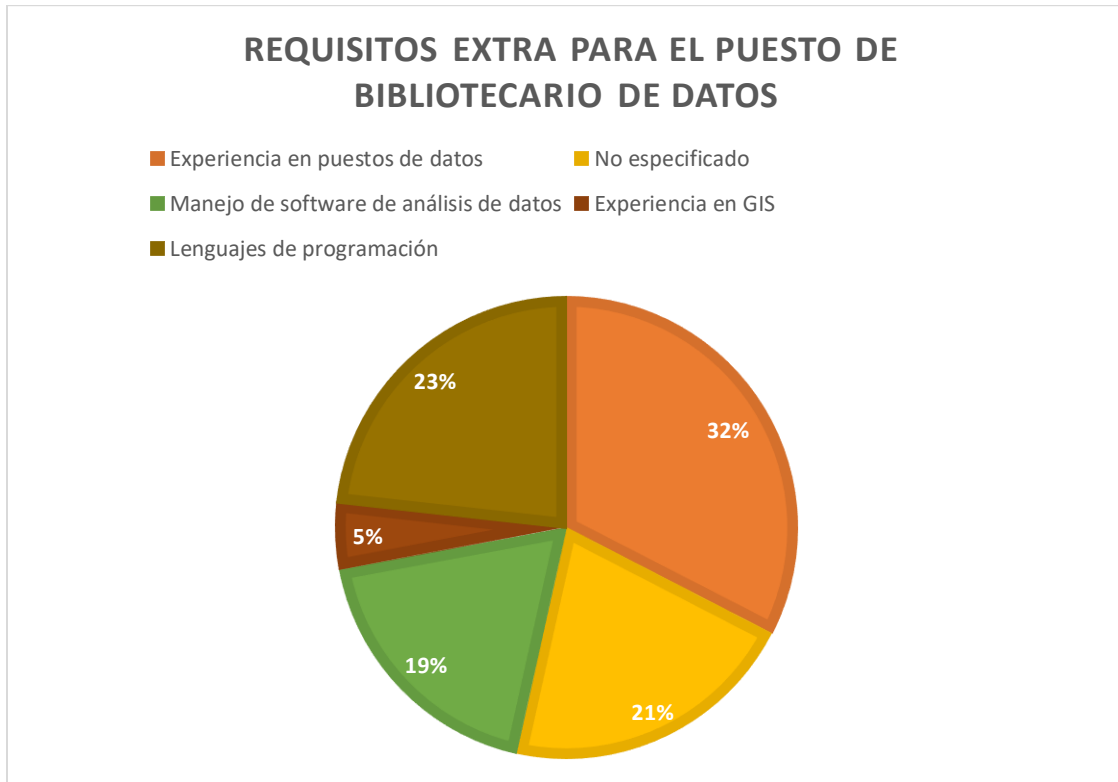


En cuanto a la categoría de requisitos extra, estos son recalcados constantemente en todas las solicitudes. Se encontró que el 32% requerían que el aspirante tuviera experiencia en algún puesto de datos o trabajando con datos, el 23% requería conocimiento en lenguajes de programación un 21% no especificaba los requisitos extra que requería, sino que invitaba a realizar una entrevista de trabajo en persona.

También en un 19% se pedía manejo de software de análisis de datos como lo es NVivo, SAS RapidMiner, *Oracle Business Intelligence Enterprise Edition* (OBIEE), Oracle, entre otros. En última instancia se mostró que en un 5% se necesitan las herramientas de uso geográfico

Gráfica 4: Requisitos extra para el puesto de bibliotecario de datos.

Fuente: Elaboración propia, 2021.



En todos los anuncios se otorga gran peso a la experiencia y al grado académico, incluso, en los anuncios destaca uno cuya intención era encontrar un profesional que tuviera un balance correcto entre experiencia académica y experiencia laboral. Se puede concluir que es esa la principal meta de los empleadores: encontrar un bibliotecario de datos formado a través de la academia y la experiencia.

Tabla 2: Palabras mencionadas con más frecuencia en anuncios de empleo.

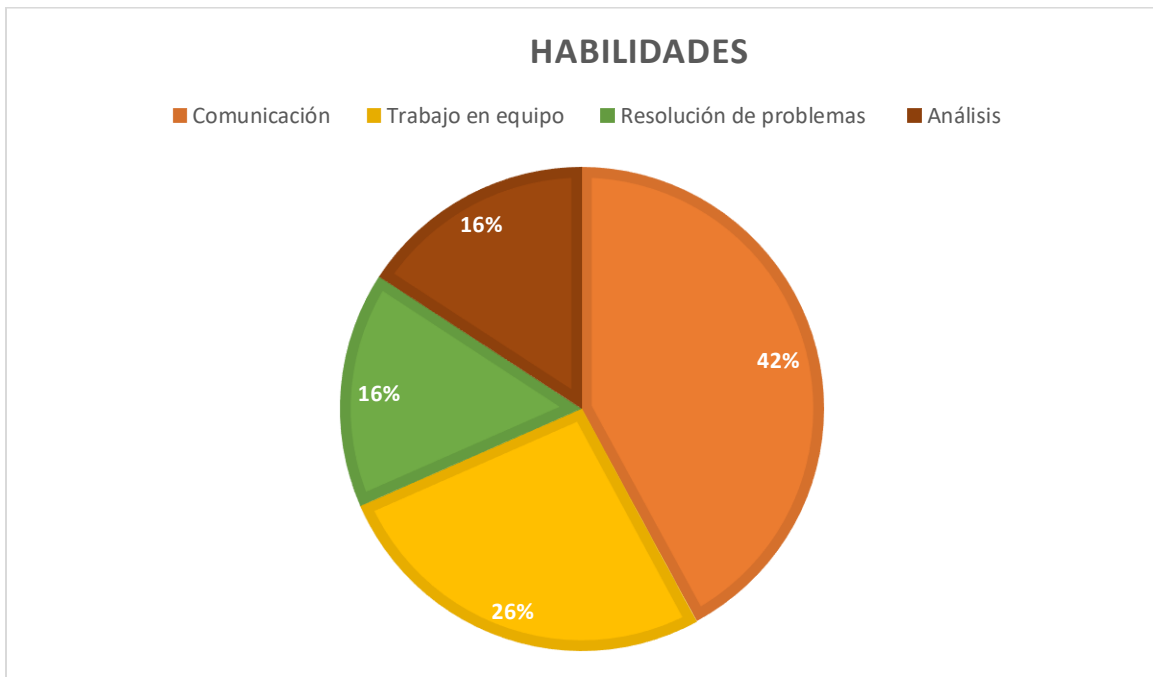
Fuente: Elaboración propia, 2021.

Palabra	Longitud	Conteo	Porcentaje ponderado
Datos	5	78	7.65%
Experiencia	11	37	3.63%
Investigación	13	23	2.26%
Maestría	8	17	1.67%
Capacidad	9	16	1.57%
Información	11	15	1.47%
Gestión	7	14	1.37%
Análisis	8	13	1.28%
Conocimiento	12	12	1.18%
ALA	3	11	1.08%
Servicios	9	11	1.08%
Acreditado	10	10	0.98%
Programa	8	9	0.88%
Título	6	9	0.88%
Bibliotecología	15	7	0.69%
Desarrollo	10	7	0.69%
Trabajo	7	7	0.69%
Visualización	13	7	0.69%
Biblioteconomía	15	6	0.59%
Compromiso	10	6	0.59%
Aplicaciones	12	5	0.49%
Capacitación	12	5	0.49%
Comprensión	11	5	0.49%

La Tabla 2 mostrada con anterioridad, presenta las palabras que fueron mencionadas con más frecuencia en los anuncios de empleo: datos (78 ocurrencias) experiencia (37 ocurrencias), investigación (23 ocurrencias), maestría (17 ocurrencias) capacidad (16 ocurrencias) e información (15 ocurrencias) A continuación en la Figura 14, se presenta la densidad terminológica que se ejemplifica a través de una nube de datos mostrando la importancia de cada característica considerada en el perfil del bibliotecario de datos.

Gráfica 5: Habilidades.

Fuente: Elaboración propia, 2021.

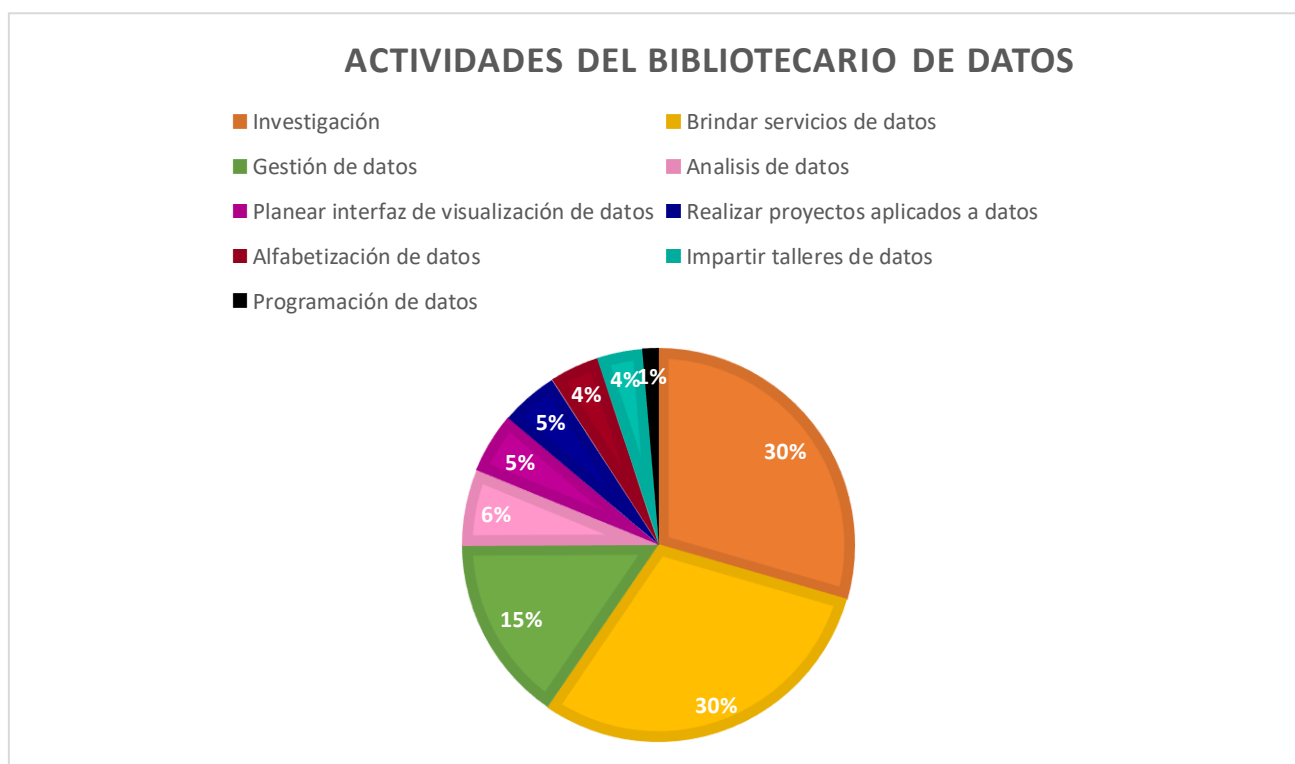


3.3.4 Actividades que desempeñará el bibliotecario de datos

La siguiente categoría fue difícil de analizar, debido a la cantidad excesiva de datos que se encontraban dentro de los anuncios al momento de describir las actividades. No obstante, una vez limpios los datos, se encontró que las actividades que desempeña el bibliotecario de datos son: investigación, esta categoría con un 30% cuenta como una de las más recurrentes dentro de las actividades a desempeñar, seguida de brindar servicios de datos, igual con un 30%, después en tercer lugar se encuentra la gestión de datos con un 15% de recurrencia, posteriormente le sigue la categoría de análisis de datos, con un 6% más adelante en la misma categoría proyectos aplicados a datos y planeación de interfaz de visualización de datos, ambas con un 5% seguida de alfabetización de datos y talleres de datos con un 4% y al finalizar programación con un 1% de recurrencia.

Gráfica 6: Actividades del bibliotecario de datos.

Fuente: Elaboración propia, 2021.



Conclusiones

El dato es la unidad mínima de información que se encuentra asociada con la realidad. Constituye una representación simbólica (numérica, alfabética, algorítmica y espacial) de un atributo o variable cuantitativa o cualitativa. Su propósito es brindar información a la luz del contexto en el que es analizado. Su importancia radica en la capacidad que tiene de representar la realidad de modo certero y brindar información que permita la toma de decisiones informadas. Actualmente, se produce gran cantidad de datos que configuran una variada tipología: datos personales, datos abiertos, datos informáticos, datos geográficos entre otros.

El dato y su creación masiva dio como resultado la recopilación de datos y a la capacidad de utilizarlos para beneficio de una amplia gama de sectores, es decir, se dio paso al *Big Data* dando como resultado, la era de los grandes datos permitiendo a los investigadores y a los científicos entender mejor los fenómenos que acontecen, descubrir pautas y correlaciones en los datos detectando perspectivas nuevas.

Debido al aumento de los datos y la necesidad de analizarlos, gestionarlos y aprovecharlos, en los últimos tiempos se empieza a hablar de un profesional de la información, denominado bibliotecario de datos. Sobre dicho perfil se encontró que el rol del bibliotecario de datos en nuestros tiempos no está conceptualizado de manera precisa; no obstante, se caracteriza por ser un nuevo perfil profesional (multifacético) derivado de la Bibliotecología, que remite a un profesional, encargado de gestionar, difundir, analizar, investigar y programar datos a la vez que atiende las necesidades de usuarios y los alfabetiza sobre el tema.

Entre sus habilidades es posible mencionar dos rubros:

- Habilidades blandas: comunicación, trabajo en equipo, diseño de proyectos, entre otros.
- Habilidades técnicas o habilidades duras: programación, conocimiento en lenguajes especializados en datos como Python y R, conocimiento de estadística, manejo de software para análisis de datos y procesos de datos de investigación.

Ambas habilidades, blandas y duras, deben combinarse de manera equilibrada para poder ser un bibliotecario de datos.

En lo que concierne a la educación, se encontró que la profesionalización del bibliotecario de datos es un asunto incipiente. Un porcentaje de la formación del bibliotecario tradicional contribuye en la definición de su perfil profesional. A partir de la revisión curricular de los programas de estudio de escuelas de Bibliotecología, se detectó que las materias que guardan relación con el perfil del bibliotecario de datos están en porcentajes por debajo del 50%, siendo el mapa curricular de la Licenciatura en Gestión de la Información y Recursos Digitales de la Universidad Autónoma de Nuevo León, la que más se acerca en la formación de dicho perfil. Aunado a ello, en la Universidad Nacional Autónoma de México se cuenta con la Licenciatura en Ciencia de Datos, diplomados, cursos y seminarios que permiten tener

una aproximación al perfil, reforzando las habilidades técnicas: programación, minería de datos, estadística y analítica de datos. Se enfatiza que tanto las bibliotecas como el perfil del bibliotecario a lo largo de la historia se han ido adaptando de acuerdo con la evolución de los recursos de información. Por esto es que hablar de "Bibliotecario de datos", contempla que ahora además de lo tradicional se le suma este nuevo enfoque de la información y por lo tanto de la actividad bibliotecaria.

A partir de la revisión y estudio de los anuncios de empleo -en los que se define el perfil profesional del bibliotecario de datos- se identificó que:

- Las instituciones que más requieren el perfil profesional de bibliotecario de datos son las universidades, seguidas por la iniciativa privada y en tercer lugar el gobierno.
- Entorno a la denominación "bibliotecario de datos" persiste mucha confusión. Se detectaron diecisiete variantes en la manera de identificar al profesional, a pesar de que los requisitos académicos y las funciones enunciadas en el perfil profesional eran bastante similares. Algunas denominaciones encontradas fueron: bibliotecario de investigación y datos, bibliotecario de datos senior, gerente de calidad de metadatos, curador de datos, responsable de datos de biblioteca, entre otros. Este aspecto abre una gran brecha para investigaciones posteriores.
- En cuanto a la formación profesional requerida para el puesto la mayoría de los casos requerían maestría o licenciatura en Bibliotecología.
- Los términos más mencionados en relación con perfil profesional fueron: experiencia, investigación, maestría, capacidad e información. Tales términos evidencian que la experiencia del bibliotecario de datos se relaciona con el acompañamiento durante la investigación, los estudios de maestría y la capacidad de informarse.
- Las habilidades que los empleadores buscan en los bibliotecarios de datos se dividen en dos rubros: habilidades duras o técnicas y habilidades blandas. Las habilidades duras o técnicas son programación, manejo de software de datos, minería de datos, analítica de datos y lenguajes de programación para datos. Las habilidades blandas son: comunicación asertiva tanto interna como externa, trabajo en equipos interdisciplinarios, inteligencia para la resolución de problemas y capacidad de análisis de la información.

- Se encontró que el bibliotecario de datos desempeñará sus funciones en dos sectores: gestión de datos y servicios de datos, ambas funciones fueron descritas como las más relevantes en las descripciones del perfil brindadas por los empleadores.

Actualmente, dado que institutos y bibliotecas de México y América Latina ofrecen servicios de datos, se asumiría que ya hay bibliotecarios especializados en datos. Sin embargo, la formación de los profesionales que laboran en estas áreas no está debidamente definida. La formación profesional del bibliotecario de datos es un asunto en el que se debe trabajar más. Tal responsabilidad concierne a las instituciones educativas que forman al bibliotecario general y la introducción de nuevas materias en el currículo escolar (obligatorias u optativas) que lo formen también como bibliotecario de datos. Se concluye que es muy probable que existan bibliotecarios de datos en México, desempeñándose, en este campo, sin que necesariamente se dé título a las funciones que realizan, este tema queda como asunto de investigación a desarrollar en trabajos futuros.

Los hallazgos del presente trabajo muestran una visión diferente de la profesión bibliotecológica, en la que, como profesionales de la información podemos descubrir un nuevo mundo asociado a los datos. La tendencia actual del perfil profesional del bibliotecario de datos se observa urgente, en vista de que el universo de los datos se expande y el rubro se vislumbra como un espacio laboral prometedor para los bibliotecarios.

Finalmente, a partir de la investigación realizada se descubrió que el tema bibliotecario de datos, es sumamente interesante y relevante de investigar; se observan muchos retos relacionados con el trabajo del bibliotecario de datos en la práctica y mayor investigación sobre su perfil, sus habilidades y la tendencia futura de la profesión. Tal hecho abre un gran panorama de estudios, por ejemplo, realizar grupos focales, con bibliotecarios que ya trabajan con datos, contribuir con artículos en los que se expliciten las tareas reales que los bibliotecarios de datos llevan a cabo en las bibliotecas. También se menciona que una de las limitaciones del presente trabajo fue localizar bibliotecarios de datos en el panorama de las bibliotecas mexicanas, ahí justamente se abre un nuevo rubro de investigación en el ámbito nacional.

Bibliografía

- American Library Association. (2020). *Datos en bibliotecas: procesamiento y visualización de datos*. Recuperado de: <https://www.ala.org/rusa/data-libraries-data-processing-and-visualization>.
- Asociación Colombiana de Ingenieros en Sistemas. (2021). *Plan Nacional de Infraestructura de Datos*. Recuperado de: <https://www.acis.org.co/>.
- Australian National Data Service. (2017). *¿What is research data.?* Recuperado de: https://www.ands.org.au/__data/assets/pdf_file/0006/731823/Whatis-research-data.pdf.
- Banco Mundial (2021). *Data for the Better Lives*. Recuperado de: <https://documentos.bancomundial.org/es/publication/documents-reports/documentdetail/778921588767120094/world-development-report-2021-data-for-better-lives-concept-note>.
- Barité, M. (2015). *Diccionario de Organización del Conocimiento: Clasificación, Indización, Terminología*. Recuperado de: <https://universoabierto.org/2019/12/05/diccionario-de-organizacion-del-conocimiento-clasificacion-indizacion-terminologia/>.
- Barnhart, B. (2018). *How to mine your social media data for a better ROI*. Recuperado de: <https://sproutsocial.com/insights/social-media-data/>.
- Barreto, V. (2019). *Surge nueva especialidad Bibliotecarios de Datos en la Ciencia Abierta*. Recuperado de: <https://umaipur.uprrp.edu/index.php/2019/11/12/surge-nueva-especialidad-bibliotecarios-de-datos-en-la-ciencia-abierta/>.
- Bishop, B.W., Orehek, A.M. y Collier, H.R. (2021). Job Analyses of Earth Science Data Librarians and Data Managers. *American Meteorological Society*, 102 (7), 1384–1393. doi: <https://doi.org/10.1175/BAMS-D-20-0163.1>.
- Bron, M. (2015). *Miradas y perspectivas de los datos abiertos*. La Rioja, Argentina: Universidad Nacional de La Rioja.
- Camillie, V.L. y Urban, J. (2018) What Do Data Librarians Think of the MLIS? Professionals' Perceptions of Knowledge Transfer, Trends, and Challenges. *College y Research Libraries*, 79 (3), 401-423. doi: <https://doi.org/10.5860/crl.79.3.401>.
- Carrión, J. (s.f.) *Diferencia entre dato, información y conocimiento*. México: Instituto de Investigaciones Bibliográficas. Universidad Nacional Autónoma de México. Recuperado de: <http://ibi.unam.mx/voutssasmt/documentos/dato%20informacion%20conocimiento.pdf>.

- Caspers, M. (2016). *Der Systembibliothekar an deutschen Bibliotheken*. (Tesis de licenciatura). Fachhochschule Potsdam. Alemania. Recuperado de: urn:nbn:de:kobv:525-13906.
- Chartered Institute of Library and Information Professionals. (2022). *Bibliotecarios de datos*. Recuperado de: <https://www.cilip.org.uk/page/DataLibrarians>.
- Dagnino, J. (2014). Tipos de datos y escalas de medida. *Revista Chilena de Anestesia*, 43 (2), 109-111. doi: <https://doi.org/10.25237/revchilanestv43n02.06>
- Eke, H. (2010). Quickies for the 21st Century Librarian. *Selected Works*, 18 (8), 1-8. Recuperado de: https://www.researchgate.net/publication/282506805_Quickies_for_the_21st_Century_Librarian.
- Fallgren, N. J. (2007). *Users and uses of bibliographic data background paper for the Working Group on the Future of Bibliographic Control*. Recuperado de: <https://www.loc.gov/bibliographicfuture/meetings/docs/UsersandUsesBackgroundPaper.pdf>.
- Federer, L. (2016). Research data management in the age of Big Data: Roles and opportunities for librarians. *Information Services y Use*, 36 (1-2), 35-43. doi: 10.3233/ISU-160797.
- Federer, L. (2018). Defining data librarianship: a survey of competencies, skills, and training. *Journal of the Medical Library Association JMLA*, 106 (3), 294-304. doi: 10.5195/jmla.2018.306.
- Fraile, A. (2018). *Big Data en las bibliotecas universitarias y científicas* (Tesis de Maestría). Universidad de Salamanca, Salamanca.
- Giannini, S. y Molino, A. (2018). *The data librarian: ¿Myth, reality or utopía?* Italia: Open portal. Recuperado de: https://openportal.isti.cnr.it/data/2019/401532/2018_401532.published.pdf?id=people%3A%3A8203f2747266c27ed34a56ca41d617f1.
- Gibaja, V. (2013). El bibliotecario como profesional de la información. *Información, cultura y sociedad*, (29), 49-76. Recuperado de: <http://revistascientificas.filo.uba.ar/index.php/ICS/article/view/675>
- Goldstein, S. (2018). Data Librarians' Skills and Competencies Are Heterogeneous and Clúster into Those for Generalists and Specialists. *Evidence Based Library and Information Practice*, 3, 294–303. doi: <https://doi.org/10.5195/jmla.2018.306>.
- Gómez, J.A. (1996). *La profesión bibliotecaria*. Madrid, España: Síntesis.
- Gonçalves Sant'Ana, C. R. (2016). Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação. *Informação y Informação*, 21 (2), 116-142. doi: 10.5433/1981-8920.2016v21n2p116.

- Gutiérrez, A. (2013). El currículum de la profesión bibliotecaria; una aproximación. *Investigación Bibliotecológica: archivonomía, bibliotecología e información*, 27 (59), 141-155. doi: 10.1016/S0187-358X (13)72534-6.
- Guy, M. (2013). *RDM Training for Librarians*. Recuperado de: <http://resources/developing-rdm-services>.
- Hernández, E. J., Duque, N. D. y Moreno, J. (2017). Big Data: una exploración de investigaciones, tecnologías y casos de aplicación. *Tecnológicas*, 20 (39), 15-38. doi: <https://doi.org/10.22430/22565337.685>
- Hubert, M. (2020). *Ausbildung zum Datenbibliothekar Wien*. Recuperado de: <https://www.openaire.eu/>.
- Infotecarios. (2017). *Bibliotecarios en el siglo XXI: nuevos retos y desafíos*. México: Recuperado de: <https://www.infotecarios.com/bibliotecarios-siglo-xxi-nuevos-retos-desafios/#.Yu2QRXbMLIU>.
- Johnson, E. O. (2018). *Working as a data librarian a practical guide*. California, EUA: Libraries unlimited.
- Kellam, L. M. y Thompson, K. (2016). *Data librarian ship: The academic data librarian in theory and Practice*. Chicago, EUA: Association of College and Research Libraries, a division of the American Library Association.
- Khan, H.R. y Du, Y. (2018). What is a Data Librarian? A Content Analysis of Job Advertisements for Data Librarians in the United States Academic Libraries. Congreso Mundial de Bibliotecas e Información de la IFLA, 24-30 de agosto de 2018. Kuala Lumpur, Malasia. Recuperado de: <http://ark:/67531/metadc1225772>.
- Lagoze, C. (2014). Big Data, data integrity, and the fracturing of the control zone. *Big Data y Society*, 1 (11), 1-11. Recuperado de: <https://journals.sagepub.com/doi/pdf/10.1177/2053951714558281>.
- Lankes, D. (2019) Defining Terms: *La biblioteconomía en la era de los grandes datos: el toque humano vital*. Conferencia de Bibliotecarios Nacionales Europeos, Mo i Rana, Noruega, 3 de junio de 2019. Recuperado de: https://davidlankes.org/librarianship-in-an-era-of-big-data-the-vital-human-touch/?fbclid=IwAR0Eyog6JVf9iMb-JUP43u9NA95-MPzdLMC_sFbzt9KclAm7IbifdWMMqmo.
- Lippenholtz, B. (2016). *Big Data en bibliotecas*. Recuperado de: <http://documotion.ar/big-data-en-bibliotecas/>.
- Liscouski, J. (1997). The Data Librarian: Introducing the Data Librarian, *Journal of Analytical Methods in Chemistry*, 19 (6), 1-6. doi: <https://doi.org/10.1155/S1463924697000242>.
- Logicalis. (2017). *Diez Claves para Definir tu Estrategia Corporativa de Gestión de Datos*. Recuperado de: <https://cdn2.hubspot.net/hubfs/273205/01->

LOG/ANALYTICS/Ebooks%20y%20CTAs/Estrategia%20corporativa%20gesti%C3%B3n%20de%20datos/90_17_LOG_10%20claves%20para%20definir%20tu%20estrategia%20corporativa%20de%20gestion%20de%20datos.pdf.

Lund University. Centre for Geographical Information Systems (2021). *Datos geográficos*. Recuperado de: <http://www.gis.lu.se/gis-centre/geographical-data>.

Manzanares, G. (2018). Apuntes y reflexiones sobre el papel del bibliotecario en la biblioteca. *Boletín de la Asociación Andaluza de Bibliotecarios*, 33 (115), 24-32. Recuperado de: <https://día.inet.unirioja.es/servlet/artículo?codigo=6554134>.

Marr, B. (2018). *Data Strategy*. Bogotá, Colombia: ECO.

Martin Donnelly, G. P. (2009) Skilling Up to Do Data: Whose Role, Whose Responsibility, ¿Whose Career? *The International Journal of Digital Curation*, 4 (2), 1-13. doi: 10.2218/ijdc.v4i2.10.

Mayer-Schönberger, V. y Cukier, K. (2013). *Big Data La revolución de los datos másivos*. Madrid, España: Turner.

México (2010). Ley Federal de Protección de Datos Personales en Posesión de los Particulares. Recuperado de: http://www.diputados.gob.mx/LeyesBiblio/regley/Reg_LFPDPPP.pdf.

Migrant y Seasonal Head Start Technical Assistance Center (2006) *Introduction to Data Analysis Handbook*. Recuperado de: <https://files.eric.ed.gov/fulltext/ED536788.pdf>

Monfort, A. y Villagra, N. (2018). *Aspectos legales, retos y reflexiones para el desarrollo de la responsabilidad social*. Madrid, España: Comillas.

Morriello, R. (2020). Birth and Development of Data Librarianship. *Italian Journal of Library, Archives and Information Science*, 11 (3), 1-16. doi: 10.4403/jlis.it-12653

Muente-Kunigami, A. y Serale, F. (2018). *Los datos abiertos en América Latina y el Caribe*. Recuperado de: <https://clustertic.org/wp-content/uploads/2018/08/PublicacionDatosAbiertosFinal.pdf>

Oficina de Información Científica y Tecnológica para el Congreso de la Unión, Gobierno de México. (2016). Los datos masivos (Big Data). *NOTA-INCyTU*, 1 (1), 1-6. Recuperado de: https://www.foroconsultivo.org.mx/INCyTU/documentos/Completa/INCyTU_16-001.pdf

Open Data Institute. (2022). *Clasificación de los datos según restricciones de uso*. Recuperado de: [https://cuantrix.com/open-data/clasificacion-datos/#:~:text=Acceso%20Interno%20\(Datos%20cerrados\)&text=Un%20ejemplo%20de%20ello%20pueden,tener%20acceso%20a%20los%20mismos](https://cuantrix.com/open-data/clasificacion-datos/#:~:text=Acceso%20Interno%20(Datos%20cerrados)&text=Un%20ejemplo%20de%20ello%20pueden,tener%20acceso%20a%20los%20mismos).

Ortega y Gasset, J. (2005). *Misión del bibliotecario*. Ciudad de México, México: Fundación José Ortega y Gasset

- Palacios, M. I., Rosa, A. A., Cal, M., Calvo, Y., Polo, F. X., Gómez, L., López, P., Rodríguez, Y. y Varela, J.R. (2019). Diccionario electrónico de enseñanza y aprendizaje de lenguas. Recuperado de: <https://www.dicenlen.eu/es/diccionario/entradas/datos-cuantitativos>.
- Paniagua, E. (s.f.). *Informe público: Big Data El poder de los datos*. Recuperado de: https://www.fundacionbankinter.org/wp-content/uploads/2021/09/Publicacion-PDF-ES-FTF_BigData.pdf.
- Pérez González, M. (2019). *El ciclo de vida del dato*. Recuperado de: https://openaccess.uoc.edu/bitstream/10609/140887/5/Gobierno%20del%20dato_M%C3%B3dulo%201_El%20ciclo%20de%20vida%20del%20dato.pdf.
- Pinfield, S., Cox, A. M. y Rutter, S. (2017). *Mapping the future of academic Libraries*. Recuperado de: <https://sconul.ac.uk/sites/default/files/documents/SCONUL%20Report%20Mapping%20the%20Future%20of%20Academic%20Libraries.pdf>.
- PowerData. (2019). *Los pilares de la gestión del Big Data y arquitectura Big Data*. Recuperado de: <https://blog.powerdata.es/el-valor-de-la-gestion-de-datos/los-pilares-de-la-gestion-y-la-arquitectura-de-datos-de-big-data>.
- Pulido, E. (2016). Big Data: ¿Solución o Problema? *Biblos-e archivo*, 394 (53), 1-12. Recuperado de: <http://hdl.handle.net/10486/677785>.
- Quintanilla, G. y Gil-García, J. R. (2016). Gobierno abierto y datos vinculados: conceptos, experiencias y lecciones con base en el caso mexicano. *Revista del CLAD Reforma y Democracia*, 65 (3), 69-102. Recuperado de: <https://www.redalyc.org/jatsRepo/3575/357546620003/html/index.html>.
- Raffino, M. (2020). *Dato en informática*. Recuperado de: <https://concepto.de/dato-en-informatica/#ixzz6pW0Vi96>.
- Reitz, J. M. (2013). *Online Dictionary of Library and Information Science*. Recuperado de: <https://www.library.ucsb.edu/research/db/1182>.
- Research Data Management Librarian Academy. (2019). *Research Data Management Librarian Academy*. Recuperado de: <https://rdmla.github.io/>.
- Rice, R. (2019). Supporting research data management and open Science in academic Libraries: a data librarian's view. *Open Science*, 72 (2), 1-12. doi: <https://doi.org/10.31263/voebm.v72i2.3303>
- Rice, R., y Southall, J. (2016). *The data librarian's handbook*. London, Inglaterra: Facet Publishing.
- Rodríguez, R., y Olivella, R. (2010). *Introducción a los sistemas de información geográfica: conceptos y operaciones fundamentales*. Recuperado de: <http://repositorio.uasb.edu.bo/handle/54000/1230>.

- Roggau, Z. (2006). Los bibliotecarios, el estereotipo y la comunidad. *Revista del Instituto de investigaciones Bibliotecológicas*, 3 (15), 13-34. Recuperado de: <https://www.redalyc.org/pdf/2630/263019683002.pdf>.
- Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*. 33 (2), 163–180. doi: <https://doi.org/10.1177/0165551506070706>
- Roy, L. y Shollerb, D. (2018). What Reference Librarians Should Know about Data Science. *Taylor y Francis Group*, 60 (2), 109–116. doi: 10.1080/02763877.2018.1560964.
- Sabiote, C., Quiles, O. y Torres, L. (2005). Teoría y práctica del análisis de datos cualitativos: Proceso general y criterios de calidad. *Revista Internacional de Ciencias Sociales y Humanidades*, 15 (2), 133-154. Recuperado de: <https://www.redalyc.org/pdf/654/65415209.pdf>
- Sanders, D. (2016). *Defining Terms: Data, Information and Knowledge*. SAI Computing Conference, Londres, 13-15 de Julio de 2016. Recuperado de: <http://10.1109/SAI.2016.7555986>.
- Sánchez, D. F. (2018). *Manual práctico para el aprendizaje del Big Data*. (Tesis de licenciatura). Universidad Tecnológica de Pereira, Colombia.
- Semeler, R., Pinto, A. L. y Frota Rozados, H. B. (2017). Data Science in data librarianship: Core Competencies of a data librarian. *Journal of Librarianship and Information Science*, 51 (3), 771-780. doi: <https://doi.org/10.1177/0961000617742465>.
- Semeler, R. y Pinto, L. (2020). La biblioteconomía de datos como campo. *Transformación*, 32, (e200034), 10-12. doi: <https://doi.org/10.1590/2318-0889202032e200034>.
- Spinak, E. (2019). *La especialidad Bibliotecarios de Datos en la Ciencia Abierta*. Recuperado de: <https://blog.scielo.org/es/2019/11/01/la-especialidad-bibliotecarios-de-datos-en-la-ciencia-abierta/#.Yib5S3rMLIU>
- Support Center for Regulatory Atmospheric Modeling. (2020). *Meteorological data and processors*. Recuperado de: <https://www.epa.gov/scram/meteorological-data-and-processors#:~:text=Observational%20Meteorological%20Data%20%2D%20Data%20consisting,current%20weather%2C%20and%20precipitation%20amount>.
- Toro, M. y Laniado, H. (2019). Big data: historia, definición, herramientas y aplicaciones en la Industria. *VirtualPro*, (204), 2-5. Recuperado de: www.revistavirtualpro.com.
- Torres, D., Robinson, N. y Cabezas, A. (2012). Compartir los datos de investigación en ciencia: introducción a la data sharing. *El Profesional de la Información*, 21 (2). 1-12. Recuperado de: <http://eprints.rclis.org/16786/1/data%20sharing.pdf>.
- Universidad Nacional Autónoma de México. Dirección General de Bibliotecas (2014). Introducción al trabajo bibliotecario. Ciudad de México, México: Ciudad Universitaria: UNAM

Universität Innsbruck. (2020). *Data Librarian*. Recuperado de:
<https://www.uibk.ac.at/weiterbildung/universitaetskurse/data-librarian/>

Valenzuela, O. (2019). *Bibliotecólogos se refrescan: ahora son expertos en Big Data*. Recuperado de: <https://noticias.utem.cl/2019/08/20/bibliotecologos-se-refrescan-ahora-son-expertos-en-big-data/>.

Zúñiga, J., y Adasme, R. (2020). *Introducción al análisis cuantitativo de datos*. (Tesis de licenciatura). *Universitat de Lleida*. Recuperado de:
<http://10.13140/RG.2.2.25253.93926>.

Anexo 1

Licenciatura en Bibliotecología y Estudios de la Información de la Universidad

Nacional Autónoma de México (UNAM) 2021.

BIBLIOTECOLOGÍA Y ESTUDIOS DE LA INFORMACIÓN (SUA) FACULTAD DE FILOSOFÍA Y LETRAS Total de créditos: 360	
PRIMER SEMESTRE 08 Fundamentos de la Organización Documental 08 Fundamentos de la Bibliotecología 08 Historia del Libro y de las Bibliotecas I 08 Fundamentos de los Servicios de Información 04 Computación Aplicada a la Bibliotecología 08 Introducción a la Investigación	SEGUNDO SEMESTRE 08 Catalogación I 04 Bibliotecología en México 08 Historia del Libro y de las Bibliotecas II 08 Bibliotecología Social 08 Tecnología de la Información en Bibliotecas 08 Métodos de Investigación Cuantitativos
TERCER SEMESTRE 08 Catalogación II 08 Fundamentos de la Administración 04 Recursos de Información 08 Lectura, Lectores y Bibliotecas 08 Bases de Datos 08 Métodos de Investigación Cualitativos 04 Optativa Humanística	CUARTO SEMESTRE 08 Catalogación por Tema 08 Administración de Servicios Bibliotecarios y de Información 08 Teoría y Técnica Bibliográfica 08 Servicios Bibliotecarios y de Información 08 Telecomunicaciones en las Unidades de Información 04 Optativa Humanística 04 Optativa Humanística
QUINTO SEMESTRE 08 Sistema de Clasificación Dewey 08 Administración de Recursos Humanos 08 Industrias Editorial y de la Información 08 Consulta I 08 Sistemas y Programas de Automatización de Bibliotecas 04 Optativa Libre 04 Optativa Humanística	SEXTO SEMESTRE 08 Sistema de Clasificación LC 08 Mercadotecnia de la Información 08 Desarrollo de Colecciones 08 Consulta II 08 Administración de Sistemas Automatizados 04 Optativa Libre
SÉPTIMO SEMESTRE 08 Evaluación de Unidades de Información 08 Bibliografía Mexicana: Siglos XVI-XIX 08 Usuarios de la Información 08 Recursos Digitales y Multimedia 08 Seminario de Titulación I 04 Optativa: Temas Selectos de Bibliotecología I o II	OCTAVO SEMESTRE 08 Indización 08 Bibliografía Mexicana Contemporánea 08 Didáctica de la Bibliotecología 08 Seminario de Titulación II 04 Optativa: Temas Selectos de Bibliotecología III o IV 04 Optativa Libre

Anexo 2

Licenciatura en Biblioteconomía de la Escuela Nacional de Biblioteconomía y Archivonomía (ENBA) 2021.

No.	CLAVE	NOMBRE DE LA ASIGNATURA	CRED.	SERIACIÓN
1er. Semestre				
1	BL00101	Fundamentos de la Organización Técnica	8	---
2	BL00102	Servicios al Público	7	---
3	BL00103	Introducción a la Biblioteconomía	7	---
4	BL00104	Historia del Libro y las Bibliotecas I	8	---
5	BL00105	Introducción a la Organización del Conocimiento	7	---
6	BL00106	Investigación Documental	7	---
7	BL00107	Herramientas Prácticas de la Computación	7	---
2° Semestre				
8	BL00208	Códigos de Catalogación I	7	---
9	BL00209	Fuentes de Información	12	---
10	BL00210	Publicaciones Seriadas	7	---
11	BL00211	Bibliografía	7	---
12	BL00212	Historia del Libro y las Bibliotecas II	7	BL00104
13	BL00213	Redes en el Entorno de la Información	7	---
3er. Semestre				
14	BL00314	Sistema de Clasificación Decimal Dewey	8	---
15	BL00315	Encabezamientos de Materia	7	---
16	BL00316	Uso y Manejo de Fuentes de Información	12	BL00209
17	BL00317	Proceso Administrativo	7	---
18	BL00318	Adquisición de Recursos de Información	7	---
19	BL00319	Automatización de Unidades de Información	7	---

4° Semestre				
20	BL00420	Códigos de Catalogación II	7	BL00208
21	BL00421	Sistema de Clasificación L. C.	8	---
22	BL00422	Usuarios de la Información	8	---
23	BL00423	Áreas Funcionales de la Administración	7	---
24	BL00424	Conservación Preventiva	7	---
5° Semestre				
25	BL00525	Estancia Profesional	31	---
26	BL00526	Informe de la Estancia Profesional	5	---
6° Semestre				
27	BL00627	Indización	6	---
28	BL00628	Planeación Bibliotecaria	7	---
29	BL00629	Bibliotecas Escolares y Universitarias	7	---
30	BL00630	Bibliografía Mexicana	7	---
31	BL00631	Investigación de Campo	6	---
32	BL00632	Estadística Descriptiva	6	---
33	BL00633	Conversión y Migración de Información	6	---
7° Semestre				
34	BL00734	Catalogación Automatizada	7	---
35	BL00735	Promoción de los Servicios	6	---
36	BL00736	Bibliotecas Públicas e Infantiles	7	---
37	BL00737	Políticas de Información	7	---
38	BL00738	Industria de la Información	7	---
39	BL00739	Introducción a la Bibliometría	6	---
40	BL00740	Almacenamiento y Recuperación de la Información	6	---

8° Semestre				
41	BL00841	Bibliotecas Especializadas	7	---
42	BL00842	Desarrollo de Colecciones	7	---
43	BL00843	Seminario de Prospectiva de la Profesión	7	---
44	BL00844	Didáctica	6	---
45	BL00845	Seminario de Investigación I	7	---
46	BL00846	Seminario de Automatización de Bibliotecas	6	---
9° Semestre				
47	BL00947	Seminario de Investigación II	9	BL00845
48	BL00948	Servicio Social	31	---

Anexo 3

Licenciatura en Bibliotecología y Gestión de Información de la Universidad Autónoma de Chiapas (UNACH) 2022.

Licenciatura: Bibliotecología y Gestión de Información **Modalidad:** Presencial, **Plan:** Semestral, **Duración:** 09 semestres, **Créditos:** 434

Primer Semestre

- Fundamento de la ciencia bibliotecológica y gestión de información
- Desarrollo histórico de la bibliotecología
- Herramientas computacionales
- Introducción a la gestión aplicada a las unidades de información
- Catalogación de material bibliográfico
- Naturaleza del conocimiento
- Diseño de recursos informáticos

Tercer Semestre

- Fuentes de información
- Bases de datos
- Sociedad actual
- Gestión de procesos e insumos aplicada a las unidades de información
- Catalogación de materiales especiales
- Archivista
- Deportes
- Actividades culturales

Quinto Semestre

- Desarrollo de colecciones
- Diseño y evaluación de programas en la formación bibliotecológica
- Digitalización y multimedia
- Clasificación de la biblioteca del congreso
- Métodos de investigación cuantitativos aplicados a la bibliotecología
- Usuarios de la información
- Optativa

Séptimo Semestre

- Formación y educación de usuarios
- Cooperación bibliotecaria
- Tesoros
- Seminario de investigación bibliotecológica
- Optativa

Noveno Semestre

- Taller de titulación

Segundo Semestre

- Filosofía y praxis de los servicios y sistemas de información
- Servicios y sistemas de unidades de información
- Sistemas de información automatizados y su aplicación a la bibliotecología
- Gestión de recursos humanos aplicada a las unidades de información
- Catalogación de materia bibliográfico
- Archivista
- Deportes
- Actividades culturales

Cuarto Semestre

- Servicios especializados de información
- Redes y telecomunicaciones
- Mercadotecnia aplicada a las unidades de información
- Encabezamiento de materia
- Clasificación decimal de Dewey y CDU
- Teoría y técnica bibliográfica
- Bibliometría

Sexto Semestre

- Automatización de los servicios en las unidades de información
- Conservación del patrimonio documental
- Organización de unidades y sistemas de información
- Indización y resúmenes
- Métodos de investigación cualitativos aplicados a la bibliotecología
- Optativa

Octavo Semestre

- Seminario de integración: gestión, organización y tecnologías
- Seminario de integración: servicios, usuarios y tecnologías
- Optativa
- Optativa
- Servicio social

Anexo 4

Licenciatura en Gestión de la Información y Recursos Digitales de la Universidad Autónoma de Nuevo León (UANL) 2022.



UANL

UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN

FACULTAD DE FILOSOFÍA Y LETRAS

Plan de estudios: Licenciatura en Gestión de la Información y Recursos Digitales

AC	Primer semestre	C	H/S
ACFGU	Responsabilidad social y desarrollo sustentable	2	2
ACFP-F	Fundamentos de la ciencia de la información	4	4
ACFB	Fundamentos de la investigación científica	4	4
ACFB	Competencia comunicativa	4	4
ACFB	Tradición intelectual de occidente	4	4
ACFB	Historia de la cultura	4	4
	Total	22	22
AC	Segundo semestre	C	H/S
ACFB	Educación socioemocional	2	2
ACFP-F	Teoría de las organizaciones	4	4
ACFB	Investigación documental	4	4
ACFB	Cultura de género	4	4
ACFB	Tradición intelectual hispanoamericana	4	4
ACFB	Diálogos interculturales	4	4
	Total	22	22
AC	Tercer semestre	C	H/S
ACFB	Filosofía del diálogo	2	2
ACFP-F	Información y calidad en las organizaciones	4	4
ACFP-F	Fundamentos de la organización de la información	4	4
ACFP-F	Ciencia de datos	4	4
ACFP-F	Sistemas de gestión de información	4	4
ACFP-F	Archivología	4	4
	Total	22	22

AC	Cuarto semestre	C	H/S
ACFGU	Cultura de paz	2	2
ACFP-F	Recursos de información	4	4
ACFP-F	Descripción documental	4	4
ACFP-F	Métricas de la información y los datos	4	4
ACFP-F	Estándares tecnológicos de información	4	4
ACFP-F	Descripción archivística	4	4
	Total	22	22
AC	Quinto semestre	C	H/S
ACFGU	Ética y cultura de la legalidad	2	2
ACFP-F	Diseño de bases de datos	4	4
ACFP-F	Organización y representación del conocimiento	4	4
ACFP-F	Análítica web	4	4
ACFP-F	Repositorios	4	4
ACFP-F	Evaluación y disposición documental	4	4
	Total	22	22
AC	Sexto semestre	C	H/S
ACFB	Didáctica de la ciencia de la información	2	2
ACFP-F	Desarrollo de bases de datos	4	4
ACFP-F	Sistemas de clasificación	4	4
ACFP-F	Minería de datos y textos	4	4
ACFP-F	Metadatos y colecciones digitales	4	4
ACFP-F	Gestión archivística	4	4
	Total	22	22
AC	Séptimo semestre	C	H/S
ACFB	Práctica docente	2	2

ACFP-F	Sistemas de recuperación de información	4	4
ACFB	Metodología cualitativa	4	4
ACFB	Redacción de textos académicos	4	4
ACFP-F	Accesibilidad de contenidos digitales	4	4
ACFP-F	Optativa del área curricular de formación profesional fundamental	4	4
	Total	22	22
AC	Octavo semestre	C	H/S
ACFP-I	Servicio social	16	2
ACFB	Metodología cuantitativa	4	4
ACFP-F	Equidad y género en la gestión de la información	2	2
	Total	22	8
AC	Noveno semestre	C	H/S
ACFP-F	Parámetros y calidad de bases de datos	4	4
ACFP-I	Optativa del área curricular de formación profesional integradora**	18	18*
	Total	22	22
AC	Décimo semestre	C	H/S
ACFGU	Liderazgo, emprendimiento e innovación	2	2
ACFP-I	Seminario de calidad y gestión de información	6	6
ACFP-I	Desarrollo de productos de investigación	4	4
ACFP-I	Seminario de tecnología de la información	6	6
ACFP-I	Arquitectura de la información en ambientes web	4	4
	Total	22	22
	Total del plan de estudios	220	206

Simbología

ACFGU: Área curricular de formación general universitaria
ACFB: Área curricular de formación básica
ACFP-F: Área curricular de formación profesional fundamental
ACFP-I: Área curricular de formación profesional integradora
H/S: Horas/semana
C: Créditos
AC: Área curricular

Anexo 5

Licenciatura en Bibliotecología y Gestión del Conocimiento de la Universidad de Guadalajara (UdeG) 2021.

	1er Semestre	2do Semestre	3er Semestre	4to Semestre	5to Semestre	6to Semestre	7mo Semestre	8vo Semestre
M A T E R I A S	Organización de la Información 10	Catalogación y manejo de lenguajes controlados 10	Manejo de técnicas de indización 10	Diseño de estrategias de servicios de información 10	Control de autoridades 10	Fomento a la lectura 10	Cuantificación bibliométrica 10	Cooperación e intercambio de información y conocimientos 10
	Apropiación y prospectiva del profesional de la información 10	Desarrollo de colecciones 10	Desarrollo de taxonomías y ontologías 10	Diseminación de información 10	Servicios informativos multi e interculturales 10	Comunicación y difusión de información y conocimientos 10	Preservación y conservación del patrimonio documental 10	Gestión del conocimiento 10
	Fundamentación epistemológica de las competencias informativas e informáticas (C12) 10	Recuperación de la información 10	Evaluación y análisis de fuentes informativas 10	Promoción de centros de información 10	Análisis e integración de la normatividad de la información 10	Elaboración de programas en competencias informativas y objetos educativos digitales 10	Gestión de contenidos web para la gestión del conocimiento 10	Proyecto VIII 12
	Proyecto I 12	Gestión de información en las organizaciones 10	Aplicación de la estadística en el uso de la información 10	Diseño de sistemas y servicios de calidad 10	Toma de decisiones 10	Proyecto VI 12	Gestión de comunidades para la innovación 10	
		Proyecto II 12	Inglés 8	Proyecto IV 12	Proyecto V 12	Laboratorio III 7	Proyecto VII 12	
			Proyecto III 12	Seminario I 7	Taller I 7	Laboratorio IV 7		

Área de formación

 Formación básica común	48
 Formación básica particular	130
 Formación especializante	206
 Formación optativa abierta	28

412 créditos para optar por el título

Anexo 6

Licenciatura en Ciencias de la Información Documental de la Universidad Autónoma del Estado de México (UAEM) 2021.



LICENCIATURA EN CIENCIAS DE LA INFORMACIÓN DOCUMENTAL

MAPA CURRICULAR

PERIODO 1	PERIODO 2	PERIODO 3	PERIODO 4	PERIODO 5	PERIODO 6	PERIODO 7	PERIODO 8	PERIODO 9	PERIODO 10
Inglés 5 2 2 4 6	Inglés 6 2 2 4 6	Inglés 7 2 2 4 6	Inglés 8 2 2 4 6		Digitalización, metadatos y lenguajes de mercado 2 2 4 6				
Teoría de la archivística 4 0 4 8	Paleografía y diplomática I 2 2 4 6	Paleografía y diplomática II 2 2 4 6	Gestión de documentos activos 2 2 4 6	Gestión de documentos semiactivos 2 2 4 6	Gestión de documentos inactivos 2 2 4 6	Descripción documental 2 2 4 6	Gestión electrónica de documentos de archivo 2 2 4 6	Diseño de edificios para unidades documentales 2 2 4 6	
Teoría de la bibliotecología 4 0 4 8	Catalogación I 0 4 4	Catalogación II 0 4 4	Sistemas de clasificación bibliográfica 0 4 4	Hemerografía 2 2 4 6	Gestión de colecciones y fuentes de información 2 2 4 6	Metodología para la investigación en ciencias de la información documental 4 0 4 8	Investigación I 3 3 6 9	Investigación II 3 3 6 9	
Teoría de la documentación 4 0 4 8	Historia de los soportes documentales 2 2 4 6	Patrimonio documental 2 2 4 6	Conservación y deterioro de documentos 2 2 4 6	Encuadernación y restauración de documentos 2 2 4 6	Análisis documental 2 2 4 6	Normalización de la información 2 2 4 6	Bibliografía 2 2 4 6	Legislación documental 2 2 4 6	
Teoría de las humanidades 4 0 4 8	Historia de la ciencia 4 0 4 8	Epistemología de las Ciencias de la Información Documental 4 0 4 8	Ética 4 0 4 8	Didáctica general 2 2 4 6	Sociología de la información 4 0 4 8	Elaboración de textos académicos 2 2 4 6	Lectura, lectores y sociedad 2 2 4 6		
Teoría de la administración y teoría general de sistemas 4 0 4 8	Instituciones políticas novohispanas 2 2 4 6	Administración de unidades documentales 2 2 4 6	Alfabetización informacional 2 2 4 6	Mercadotecnia de productos y servicios de la información 2 2 4 6	Análisis de la información contemporánea 0 4 4 4	Estadística para documentalistas 2 2 4 6	Metría de la información 2 2 4 6		
		Sistema político y administrativo de México 2 2 4 6						Evaluación y visualización de la ciencia 2 2 4 6	

Práctica Profesional

30