

## Hemeroteca Digital Nacional de México, su historia y alcance como fuente para proyectos digitales

### Mexico's National Digital Newspaper Library, its History and Scope as a Source for Digital Projects

Isabel Galina Russell

Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas  
[igalina@unam.mx](mailto:igalina@unam.mx)

Ernesto Priani Saisó

Universidad Nacional Autónoma de México, Facultad de Filosofía y Letras  
[epriani@gmail.com](mailto:epriani@gmail.com)

Miriam Peña Pimentel

Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas  
[miriampp@unam.mx](mailto:miriampp@unam.mx)

Laura Martínez Domínguez

Universidad Pedagógica Nacional  
[lauramartinezdominguez@gmail.com](mailto:lauramartinezdominguez@gmail.com)

Rocío Castellanos Rueda

Universidad Autónoma Metropolitana  
[doroci2011@gmail.com](mailto:doroci2011@gmail.com)



[DOI: 10.24901/rehs.v4i1i63.707](https://doi.org/10.24901/rehs.v4i1i63.707)



Hemeroteca Digital Nacional de México, su historia y alcance como fuente para proyectos digitales por [Isabel Galina Russell](#), [Ernesto Priani Saisó](#), [Miriam Peña Pimentel](#), [Laura Martínez Domínguez](#) y [Rocío Castellanos Rueda](#) se distribuye bajo una [Licencia Creative Commons Atribución-NoComercial 4.0 Internacional](#).

Fecha de recepción: 9 de mayo de 2019

Fecha de aprobación: 6 de abril de 2020

## **RESUMEN:**

La Hemeroteca Nacional Digital de México (HNDM) es una de las colecciones digitales pioneras en México. En este artículo hacemos una reconstrucción de su historia a partir de testimonios orales y escritos con la intención de entender su origen; su relación con la colección impresa; el conjunto de decisiones tanto archivísticas como técnicas que determinaron su formación y su operación actual; y, a partir de todo esto, comprender su naturaleza y los límites que tiene como fuente de proyectos digitales actuales y futuros. Existe poco trabajo que analice críticamente cómo está constituido el archivo digitalizado y generalmente la versión digital es utilizada como un simple sustituto del archivo físico original. En este artículo, siguiendo a Cordell (2017), argumentamos que es necesario reflexionar sobre cómo la computadora, las bases de datos, las interfaces, los algoritmos, por mencionar algunos, agregan capas interpretativas generalmente imperceptibles que median nuestra relación con el objeto estudiado.

## **Palabras clave:**

Digitalización, humanidades digitales, hemerografías digitales, fuentes hemerográficas, Hemeroteca Nacional de México.

## **ABSTRACT:**

Mexico's National Digital Newspaper Library (*Hemeroteca Nacional Digital de México*, HNDM) is a pioneering digital collection in the country. This article presents a reconstruction of the history of this institution based on oral and written testimonies in order to elucidate its origins, its relation to published collections, the series of decisions -technical and archival- that determined its current composition and functioning and, on this basis, better understand the nature of the Library and its limitations as a source for current and future digital projects. Few studies have critically analyzed how digitalized archives are constituted. Moreover, generally speaking, the digital version of documents is used simply as a substitute for the original, physical version. Following Cordell (2017), however, we argue that it is necessary to reflect upon how computers, databases, interfaces, and algorithms, to mention only a few key aspects, add imperceptible interpretative layers that mediate our relation to the object of study.

## **Keywords:**

Digitalization, digital humanities, digital hemerographies, hemerographic sources, *Hemeroteca Nacional de México*.

## **Introducción**

En este artículo nos proponemos documentar y analizar críticamente cómo fue conformada la Hemeroteca Nacional Digital de México (HNDM) desde una perspectiva integral. No sólo se busca reconstruir los procesos técnicos, sino los diversos factores que influyeron en la constitución del acervo digital, buscando destacar qué implicaciones tienen las decisiones tomadas y sus resultados para los investigadores que utilizan la plataforma de la HDNM como fuente para sus investigaciones.

El artículo forma parte de los resultados del proyecto “Intercambios Oceánicos-Oceanic Exchanges (OcEx)”, un proyecto de investigación conformado por once instituciones en seis países que tiene como objetivo utilizar herramientas digitales para estudiar el flujo de información entre periódicos históricos de 1840 a 1914. Con ese fin, dentro del proyecto se integraron los datos de ocho hemerotecas digitalizadas que serían utilizados para el desarrollo de herramientas de minería de datos<sup>1</sup> y como fuente para la investigación de casos de estudio sobre el flujo de noticias de eventos históricos. Uno de los primeros retos al tratar de unir estas colecciones hemerográficas digitalizadas y crear las ontologías<sup>2</sup> que nos permitieran analizar los flujos de noticias fue describir adecuadamente cómo estaban conformadas las diversas

colecciones, con el objetivo de entender cómo la lógica detrás de su conformación podía determinar los resultados que obtuvieramos al investigarlas.

En una primera etapa fue evidente que si bien podíamos identificar y contabilizar hasta cierto punto el número de páginas digitalizadas, la cantidad de títulos y algunos otros datos adicionales como el origen geográfico o la fecha de publicación de cada documento, no era posible para ninguna de las colecciones de hemerotecas digitales que se incorporaron al proyecto explicar con precisión cómo se había conformado la colección ni cómo la versión digital se relacionaba con la colección impresa.

Para estudiar esta cuestión, se decidió investigar la conformación de cada una de las colecciones. Los resultados globales de esta investigación se encuentran en proceso<sup>3</sup> y este artículo presenta los resultados del equipo mexicano que estudió la colección digitalizada de la Hemeroteca Nacional de México (HNM).

Un primer acercamiento al problema nos permitió identificar que uno de los retos de la investigación era la escasa información disponible sobre la creación de la HNDM. El documento de mayor valor y nuestra fuente principal fue el artículo “Diez años de la Hemeroteca Nacional Digital de México (HNDM). Breve reseña de una larga gestión de preservación y acceso, 2002-2012” ([Curiel y Jiménez 2015](#)). Para suplementar esta información decidimos completarla con una historia oral, resultado de entrevistas con personas involucradas en el desarrollo de la HNDM.

El guion que preparamos para las entrevistas tenía como principal objetivo “llenar los huecos” en torno al origen, desarrollo y estado actual de la HNDM con el fin de comprender cómo se configuró la colección digital y su estructura actual incluyendo su alcance; los procesos de toma de decisiones; creación u origen de los metadatos, catalogación y control de calidad. Aunque no fue posible entrevistar a todos los involucrados en la creación de la HNDM, pues, no todos fueron localizados, se entrevistó a personas clave del proceso que ocupan los siguientes puestos: coordinador de la HNM; jefe del Departamento de Unidad de Microfilmación y Fotografía; personal del Departamento de Sistematización Hemerográfica y anteriormente coordinadora de la HN; jefe del Departamento de Catalogación; y jefe del Departamento de Informática. Los comentarios han sido anonimizados e integrados a los resultados e identificadas mediante claves: E1, E2, E3, etcétera.

Las entrevistas tenían un formato semiestructurado y se buscó identificar lo siguiente:

- Profundizar sobre los criterios para la microfilmación y entender el alcance de la colección de microfilmes en relación con la colección general y la digitalizada.
- Identificar con mayor claridad las diferentes etapas por las cuales pasó el proyecto, y la relación entre los distintos procesos: digitalización, creación del sistema, registro y control, por mencionar algunos. Identificar documentación existente que no habíamos localizado.

- Aclarar los cuadros clasificatorios de los periódicos y su relación con los métodos de catalogación en los periódicos, así como información acerca de la organización documental de las tres colecciones: la impresa, la microfilmada y la digitalizada.
- Ampliar la información acerca de las decisiones sobre formas de búsqueda y la interfase con el usuario.
- Información sobre el estado actual del proyecto: nuevas colecciones, procesos de digitalización, políticas de ingreso de nuevos periódicos, por mencionar algunos.
- Cualquier información adicional que nos permitiera tener un cuadro más completo.

Las encuestas fueron transcritas y analizadas para detectar los principales temas, las respuestas anonimadas, y las hemos integrado con la información obtenida por otras fuentes, para ofrecer una historia más completa del proceso de creación del HNDM. El cuestionario está disponible como Anexo I. Este trabajo se realizó durante el 2018.

### **Problemas de la digitalización y la creación de herramientas de investigación**

Los periódicos pueden considerarse ya sea “por separado o en conjunto como objetos de estudio científico” ([Fernández 2010](#)) y las colecciones históricas, en particular, una fuente de información rica e única de manera que es pertinente conocer y comprender la historia y composición de estas colecciones. Para ello es preciso interrogarnos ¿cómo llegó a formarse tal colección? ¿Qué títulos de periódicos existen y qué tan “completa” está tal colección? ¿Qué mecanismos bibliohemerográficos hay que describen la colección y las publicaciones que contiene?

La creación de colecciones que posteriormente se denominan “patrimonio hemerográfico” constituyen un acto de atribución de “relevancia” e “importancia” para un determinado periodo histórico. Los historiadores, al utilizar estas colecciones para entender fenómenos históricos que describen una época delimitada, están circunscritos necesariamente a lo que un coleccionista o una biblioteca determinó en un momento como “digno de resguardo”, ya sea de forma deliberada o por circunstancias azarosas. Al utilizar los periódicos históricos como fuente para realizar algún estudio es imprescindible conocer la conformación de la colección y su representatividad, de manera que estos datos sirvan para contextualizar los límites o los alcances del estudio.

En años recientes, gracias a la creciente disponibilidad de acervos hemerográficos digitalizados se ha renovado el interés en su estudio ([Latham y Scholes 2006](#)). Esto se debe en buena medida a que es cada vez más común que la fuente de estudio no sean los objetos físicos, sino la versión digitalizada de la colección hemerográfica tanto por razones de acceso -en muchos casos los originales pueden no estar disponibles para consulta, por diversos motivos-, como por el uso de herramientas de descubrimiento que ayudan a identificar documentos relevantes, agruparlos y explorarlos. De modo que el conocimiento sobre el origen de la colección digital, sus características y procesos de digitalización, adquiere mucha importancia.

El uso de colecciones digitales permite al usuario acceder a la información de forma remota y, dependiendo de la plataforma de acceso, es posible utilizar diferentes herramientas de

localización y descubrimiento de las colecciones. Por ejemplo, muchas plataformas ofrecen al investigador la opción de realizar búsquedas sobre el texto completo localizando la ubicación de una palabra o palabras en particular en diversos ejemplares de periódicos, algo que no es posible hacer en la colección impresa. Si bien esto permite al usuario nuevas formas de explorar una colección es necesario, al igual que con la colección impresa, tomar una postura crítica ante los resultados ofrecidos a partir de los cuales se llegará a interpretaciones históricas sobre determinados fenómenos. ¿Qué partes de la colección fueron digitalizadas y cómo se llegó a determinar esto? ¿Qué controles se utilizaron para revisar los procesos de digitalización y asegurar su representatividad en comparación con el “original”? ¿Qué información está disponible para los investigadores sobre la toma de decisiones y qué mecanismos de control de calidad existen? ¿Cómo se conformó el archivo digitalizado y cuáles mecanismos de descubrimiento existen? ¿Qué herramientas existen para poder describir adecuadamente la relación entre los procesos de cómputo y los procesos de recuperación y visualización de los periódicos?

Para atender estas consideraciones [Cordell \(2017\)](#) aboga por estudiar estas colecciones digitales como un conjunto de nuevas ediciones y no pensar en la digitalización como una sustitución o reemplazo del original. Considera que se requiere trabajo de análisis y descripción bibliográfica de las colecciones digitales con métodos robustos para describir estos artefactos digitales para: “dar cuenta de las fuentes, tecnologías y realidades sociales de su creación de manera que sus posibilidades y limitaciones sean más fácilmente visibles y estén disponibles para la crítica”<sup>4</sup> ([Cordell 2017, 191](#)).

Existen, sin embargo, pocos estudios sobre la conformación de las colecciones digitales y las implicaciones que tienen para las interpretaciones y estudios. Como comenta McGann: “Existe toda otra problemática sobre cómo integrar el archivo digital emergente con el archivo impreso heredado. Realmente no sabemos cómo hacerlo todavía. Pero es imperativo debido a que las bibliotecas se organizan digitalmente”<sup>5</sup> ([Fraser y McGann 2015](#)). Es de gran importancia desarrollar métodos bibliográficos que permitan entender las colecciones digitalizadas y su relación con la colección impresa de tal forma que posibilite crear marcos interpretativos más robustos.

Una pequeña parte de ese esfuerzo es el que realizaremos en este texto al reconstruir la historia de la Hemeroteca Nacional Digital de México e identificar cómo su conformación establece un marco de interpretación a los objetos que contiene.

### **El origen: la Hemeroteca Nacional de México**

La HNM fue creada en 1944 cuando se trasladó el acervo hemerográfico al templo de San Pedro y San Pablo del exconvento de San Agustín, de acuerdo con la página de la Hemeroteca Nacional ([Hemeroteca Nacional de México s.f.](#)). Sin embargo, como apunta [Vieyra Sánchez \(2015\)](#), hay muchas fechas anteriores a ésta que pueden considerarse importantes en el proceso de su creación y de la formación de las colecciones hemerográficas que la componen. Se menciona, por ejemplo, que desde 1912 el entonces director de la Biblioteca Nacional, Rogelio Fernández

de Güell indicó en su informe que había instalado un “departamento de periódicos y publicaciones nacionales y extranjeras, independiente del resto de la Biblioteca” ([Vieyra Sánchez 2015, 9](#)). Después, en el periodo comprendido entre 1912 y 1944, en que se dotó a la Hemeroteca con su propio edificio, ocurrieron una serie de cambios progresivos que permitieron identificar y reconocer las colecciones hemerográficas como distintas al resto del acervo de la Biblioteca Nacional, con retos bibliográficos de orden, catalogación y resguardo de otra naturaleza.

El trabajo que se ha realizado a través de los años en la HNM es muy importante y gracias a él la colección es hoy una fuente extraordinaria para la memoria documental de México. Precisamente por ello es fundamental identificar adecuadamente los retos enfrentados para custodiar, organizar, sistematizar y administrar esta colección, pues, la colección digitalizada depende, en gran medida, del trabajo de control bibliográfico que se ha hecho a lo largo de estos años.

Comencemos por señalar que, desde su origen, las colecciones periódicas han tenido múltiples sedes. La mayoría de los cambios han sido motivados por falta de espacio de almacenamiento, un problema frecuente con las colecciones hemerográficas. Sin embargo, por las difíciles situaciones por las que ha atravesado el país durante el siglo XX ha significado que muchas veces el rezago para atender necesidades se mida por años. Por ejemplo, ya en 1942, el rector Rodolfo Brito organizó una visita “con objeto de darse cuenta del estado lamentable en que se hallaba el material bibliográfico, particularmente los periódicos, que en grandes cantidades se encontraban hacinados en el piso por falta de muebles donde colocarlos, desde el año de 1925” ([López 1998, 17](#)).

Finalmente, en 1979, se traslada la Biblioteca y Hemeroteca nacionales al Centro Cultural Universitario, en Ciudad Universitaria, y en 1992 se inaugura el Fondo Reservado ([Hemeroteca Nacional de México s.f.](#)). Desde entonces, la Hemeroteca cuenta con una bodega separada para el resguardo de periódicos que, sin embargo, ya desde 2013 “están saturados” ([Olivares Alonso 2013](#)). En 2016 se anunció la creación de otra sede en Querétaro, el Centro de Almacenamiento Documental que “pretende resolver los problemas urgentes de falta de espacio en el edificio y los almacenes” ([Gaceta Digital UNAM 2016](#)). Pero, al inicio del 2020 este proyecto aún no se empieza y si bien pueden existir numerosas razones administrativas y presupuestales, nuevamente pareciera que una situación “urgente” puede demorar años en resolverse. Aunado a esto también están las importantes y difíciles tareas de crear los registros de los periódicos que permitan a los usuarios de la HNM consultarlos y a los bibliotecarios administrar y conocer su acervo para su uso adecuado y su preservación.

A través de los años, la Biblioteca y Hemeroteca Nacionales han utilizado diversos sistemas para el registro de sus colecciones. El trabajo de Gloria Escamilla ([Escamilla 1965](#)), en particular, da cuenta de propuestas para la catalogación en la Biblioteca Nacional adoptados hasta esa fecha. Para los fines de este artículo, sin embargo, resultan particularmente relevantes los procesos de automatización de la Hemeroteca y, sobre todo, la transición de tarjetas a sistemas automatizados, desde bases de datos a catálogos en línea.

Encontramos poca información de los procesos iniciales realizado en los años ochenta, con el sistema MINISIS. En los noventa con la adquisición de Dynix, un software para la administración de bibliotecas, tanto la Biblioteca como la Hemeroteca dieron seguimiento a la automatización de sus procesos. En 1998, la HNM contaba con un catálogo electrónico llamado Ephemeric que podía consultarse en línea vía Telnet ([Villalba Roldán 2000](#)). En 2004, ambas instituciones se movieron a Aleph, sistema para automatización de bibliotecas, que aún se utiliza actualmente. La introducción del software para los procesos de catalogación necesariamente está acompañada de cambios en las formas de trabajo que no siempre son adoptados uniformemente, lo que puede llevar a divergencias en los registros. Como veremos más adelante, esto conlleva a que las fuentes utilizadas para la plataforma digital posean ciertas características que al final impactarán en la construcción del sistema y de su uso.

En los noventa se creó el catálogo electrónico ([Suaste Lugo 2017](#); [Villalba Roldán 2000](#)) y en el 2004 se puso en línea con el nombre de Nautilo, en el cual se destaca “más de un millón de 250 mil libros y documentos con los que cuenta este recinto [Biblioteca Nacional], más la Hemeroteca Nacional” ([El Universal 2004](#)). En su momento, la directora de la Biblioteca indicó que se encontraba automatizado el 40 % del acervo, aunque no está claro si esto incluía también el acervo de la Hemeroteca. Hoy, el catálogo Nautilo para la Hemeroteca se encuentra disponible y ofrece al usuario la posibilidad de descubrir periódicos utilizando búsquedas básicas y avanzadas, así como con índices. Sin embargo, al momento de escribir estas líneas aún es difícil localizar información precisa en relación con cuánto del acervo está catalogado y, en particular, a qué grado de detalle ha llegado la descripción de los objetos.

A la par, la HNM también ha llevado a cabo importantes trabajos de microfilmación. Al crear una copia del impreso en un microfilm, un soporte mucho más robusto que el papel, se busca asegurar la preservación de la colección. En la HNM existen dos copias de microfilm: el original cuya emulsión es de haluro de plata y una copia en película diazónica que se presta para consulta. La copia de consulta en microfilm ayuda a la preservación del original ya que reduce su uso. La microfilmación en la Hemeroteca se realizó en los sesenta, momento que coincide con el traslado de la Hemeroteca Nacional a su nueva sede en Ciudad Universitaria. De acuerdo con [Curiel y Jiménez \(2015\)](#): “la HNM inició un ambicioso programa para la selección y microfilmación de las colecciones antiguas más relevantes de su acervo” que continúa vigente y que actualmente consta de aproximadamente 28 millones de páginas microfilmadas e incluye microfilmes de preservación y copias para consulta. De especial importancia para nosotros es que los materiales por microfilmarse se determinan a partir de una selección que está basada en diversos criterios acerca del valor del documento: la microfilmación de conservación practicada en la HNM es una actividad que comienza a partir de una cuidadosa selección de los materiales por microfilmarse, considerando los factores de importancia histórica, patrimonial, periodística y documental, el estado de conservación y el grado de demanda de los usuarios, así como los asuntos inherentes a los derechos de autor ([Curiel y Jiménez 2015, 4](#)). Sin embargo, no logramos encontrar información más detallada respecto a estos criterios y cómo se han aplicado.

Aunado al trabajo bibliotecario es importante notar que en 1967 se creó el Instituto de Investigaciones Bibliográficas (IIB), entidad académica adscrita a la Universidad Nacional

Autónoma de México, a la cual se le confiere la administración de la Biblioteca y Hemeroteca Nacionales así como realizar e impulsar la investigación bibliográfica, hemerográfica, bibliotecnológica, bibliológica, de archivos y de estudio de fuentes y contribuir al rescate de la bibliografía y la hemerografía mexicanas. El trabajo del IIB es de vital importancia para el conocimiento de los acervos. En el caso de los periódicos históricos de la HN cabe destacar el trabajo del Seminario de Biblioteca Mexicana del Siglo XIX que tiene el objetivo de “buscar, identificar, registrar y estudiar la producción periodística nacional” (Curiel y Castro 2003) y que entre sus productos de investigación destaca la publicación de catálogos impresos de los periódicos en el Fondo Reservado (Curiel y Castro 2003; Curiel, Castro y Quirarte 2019). La investigación bibliohermerográfica que realiza el IIB aporta información especializada sobre los acervos, y puede contribuir enormemente al conocimiento de los mismos para complementar la información de descripción catalográfica. Por ejemplo, explicaciones detalladas de los periódicos por su relevancia histórica. En la versión actual de la HNMD, sin embargo, esta información no está integrada a los sistemas de búsqueda, sino que los libros han sido digitalizados y se encuentran en formato PDF en la sección “Herramientas de consulta”. Se pueden realizar búsquedas sencillas y arroja la descripción detallada, pero esto no está vinculado directamente con los periódicos digitalizados.

Hasta el día de hoy la HNM, continúa resolviendo asuntos relacionados con falta de espacio, identificación y catalogación así como la preservación de su colección. Para una historia completa y detallada véase Lilia Vieyra (Vieyra Sánchez 2015).

### **El proceso: la creación de la Hemeroteca Nacional Digital de México (HNMD)**

En el año 2000 se empezó a gestar la idea de digitalizar los periódicos de la HNM. Guadalupe Curiel fue la principal promotora del proyecto HNMD y durante su periodo como coordinadora de la HNM y, después, como directora del IIB se realizó la mayor parte del trabajo de creación de la HNMD entre los años 2000 y 2015. En el documento, mencionado anteriormente, “Diez años de la Hemeroteca Nacional Digital de México (HNMD). Breve reseña de una larga gestión de preservación y acceso, 2002-2012” (Curiel y Jiménez 2015) encontramos tratados los temas fundamentales de la gestación y desarrollo de la HNMD, en los que profundizamos a través de las entrevistas.

### ***La microfilmación como fuente***

Es relativamente común entre bibliotecas el uso del microfilm para apoyar las tareas de acceso y preservación de las colecciones. En un principio se pensó que la digitalización podría reemplazar la microfilmación. Sin embargo, como argumentan Curiel y Jiménez, la HNM consideró, al igual que otros proyectos alrededor del mundo, que se “podría recurrir al escaneo digital de sus colecciones a partir del microfilm, sin la necesidad de repetir el largo y arduo procedimiento de selección, cotejo y preparación de los ejemplares, acciones ya implícitas en la producción de los microfilmes” (Curiel y Jiménez 2015). La microfilmación además tenía las características técnicas necesarias para considerarse de calidad ideal para el proceso de digitalización. Sin embargo, los autores reconocen que si bien esto facilitaba el trabajo, las

“siguientes actividades en secuencia, tales como enviar, controlar y supervisar el estado de conservación antes y después de la transformación digital, tampoco eran tareas sencillas” ([Curiel y Jiménez 2015, 9](#)).

Esta tarea fue encomendada al Programa de Estabilización y Diagnóstico de la Colección de Microfilmes Originales de la HNDM, que debía estar a cargo “entre otras cosas de la selección, suministro y control de las colecciones micrográficas que alimentan a la HNDM; [...] revisión puntual del estado de conservación de los artefactos fotográficos, antes y después de la transformación digital” ([Curiel y Jiménez 2015, 10](#)).

Sin embargo, hubo ciertas complicaciones “más cuando no se tenía un sistema de control micrográfico confiable que conjuntara las características técnicas micrográficas de los artefactos analógicos, con su condición de conservación y los datos hemerográficos de sus contenidos” ([Curiel y Jiménez 2015, 9](#)). Es decir, no existía una relación entre la información que se tenía de los microfilmes y de los objetos analógicos microfilmados.

Además, había microfilmes que ya contenían imágenes de mala calidad: “De hecho se digitalizó todo, pero si entras a la HNDM no todo se ve bien, porque algunos periódicos están mal encuadrados y se pierden palabras en el centro, cosas así” (E3). A esto hay que añadir que la colección de microfilmes también tuvo distintas etapas y no siempre se manejaron los mismos criterios. Esto contribuye también a que no sea una colección uniforme. “A través de los años la microfilmación pasó por varios directores y cada uno establecía ciertas condiciones” (E1).

Un punto que resaltar dentro del proceso es la dificultad para identificar una empresa que pudiera hacer la tarea de digitalización de los microfilmes. La HNDM fue de los primeros trabajos de digitalización a gran escala en el país y la oferta de empresas con experiencia en digitalización era entonces muy limitada. Las empresas no contaban con escáneres para la digitalización de microfilmes ni el conocimiento para su manejo. Tampoco tenían flujos de trabajo estandarizados ni conocimiento empírico para la digitalización de este tipo de materiales. [Curiel y Jiménez \(2015\)](#) indican que solamente se encontró una empresa que pudiera hacer el trabajo, localizada fuera de la Ciudad de México, que tenía antecedentes de haber trabajado con una empresa canadiense en un proyecto de digitalización de periódicos.

Vale la pena mencionar que de la misma forma que fue difícil encontrar una organización para que se realizara la digitalización de los microfilmes, en su momento, también fue complicado encontrar empresas que realizarán el trabajo de microfilmación, ya que no existían empresas que hicieran esa labor de conservación en México en los sesenta: “Entonces muchos de los microfilmes que se mandaron hacer al principio, resultaban muy malos en cuanto a su calidad técnica” (E1). El rezago en la oferta de servicios de este tipo es de notarse en ambos casos y afecta de manera significativa el proceso completo de microfilmación y su posterior digitalización.

Para el prototipo de la HNDM se evaluaron los microfilmes y se enviaron los que cumplieran con ciertas características combinadas con la buena calidad del microfilm: “La idea fue empezar el prototipo con lo mejor para que la empresa aprendiera cómo manejar los microfilmes y

conforme avanzáramos proporcionaríamos los microfilmes de menor calidad que los primeros pero ya el sistema habría aprendido a trabajar con ellos” (E1).

Ahora bien, a pesar de las dificultades, el trabajo de digitalización resultó ser más rápido de lo esperado. Conforme se aceleraron los ritmos de trabajo:

Llegaban los proveedores y se llevaban todo lo que había en una estantería, sin fijarse si estaban completos, si estaban duplicados, microfilmaban parejo. El resultado es que teníamos una colección de microfilmes en la que había repeticiones: había un volumen microfilmado y más adelante aparecía ese mismo volumen, de la misma fecha microfilmado, sólo que uno tenía faltantes y el otro no (E1).

Otro problema importante se presentó cuando se migró el sistema de la empresa a los servidores en Ciudad Universitaria. El problema central es que resultó difícil empatar los controles de registro. Se encontraron divergencias en los registros entre los números reportados, los microfilmes entregados y recuperados y el número de imágenes digitalizadas, así como su correspondiente OCR<sup>6</sup> en XML<sup>7</sup>: “cuando hicimos el traslado no todo llegó, la empresa tuvo problemas [...] DGTIC y el análisis que también hicimos de la base de datos para hacer el traslado fue lo que nos permitió realmente identificar con mucha precisión, qué es lo que al menos podemos decir que sí hace falta. [...] sí hubo una cantidad de imágenes que no llegaron” (E2). Una vez trasladado e identificadas las divergencias, una tarea importante ha sido asegurar que los pendientes identificados se resolvieran.

Si bien en sus inicios la HNDM se formó con la digitalización de los microfilmes, con el paso del tiempo también hubo posibilidad de incorporar materiales de otra forma. Cuando la HNDM se trasladó a Ciudad Universitaria fue posible que la UNAM adquiriera escáneres para digitalizar directamente los periódicos. En los años que transcurrieron desde el inicio del proyecto, la tecnología disponible cambió mucho y se empezaron a digitalizar los periódicos directamente y ya no los microfilmes. En la actualidad se están implementando sistemas de digitalización internos en la Hemeroteca, pero aún se están afinando procedimientos y la producción es modesta. Estos procesos de digitalización aún no se han articulado con la HNDM y, por lo tanto, las nuevas digitalizaciones no se han integrado a la colección digital. Cabe destacar dos aspectos: la falta de un módulo de gestión de la HNDM y que la HNDM no se ha integrado a la administración de la HNM. Ambos aspectos se discuten más adelante.

### *El diseño de sistema*

Si bien hoy en día pareciera una obviedad, uno de los asuntos con los proyectos de digitalización tempranos es la falta de claridad entre el simple proceso de escaneo y las formas de entrega de los archivos digitalizados. Los procesos de digitalización generan una serie de imágenes que se almacenan en carpetas. Lo crucial es encontrar los mecanismos para que estas imágenes digitalizadas se relacionen con la información metatextual acerca de la imagen, por ejemplo, el título del periódico, la fecha, el lugar de publicación. En síntesis, la información necesaria para

vincular la imagen digitalizada con el original impreso, así como toda la información metatextual relacionada con el nuevo objeto digital, es decir, los metadatos. Esto resulta indispensable para la administración de la colección digital, identificando qué imágenes corresponden a qué periódicos y a la larga para crear plataformas de búsqueda y acceso a la colección digital. En el ámbito bibliotecológico esto se conoce como organización documental.

El proceso de digitalización resultó demasiado rápido, como se señaló antes, lo que causó problemas con los procesos previos (el del seguimiento y control de microfilmes) y secundarios (el portal para tener acceso a lo digitalizado): “Simplemente el tiempo de selección y preparación de los microfilmes resultaba mayor que el tiempo que la empresa requería para su digitalización, y de igual manera resultaba con la programación” (Curiel y Jiménez 2015, 11). Si bien en el artículo referido discuten algunas de las acciones que se realizaron, queda claro que hacían falta suficientes recursos humanos y un tema de “revisar lo nuevo sin descuidar lo pendiente”.

Curiel y Jiménez refieren que aun cuando el sistema de la empresa realizaba las funciones técnicas requeridas, era necesario desarrollar los servicios virtuales. Se requería que el sistema

debería ajustarse a toda clase de publicaciones periódicas [...] además de prestar servicios equivalentes a los tradicionales, tales como la eventualidad de buscar y hojear libremente un fascículo. Pero, además, las nuevas posibilidades de acceso [...] búsqueda por palabras sobre el texto completo [...] bajo el principio de una interfaz intuitiva de fácil operación y alejada de los sistemas aplicados en bibliotecas y hemerotecas tradicionales (Curiel y Jiménez 2015, 11).

Jiménez y Curiel se detienen en las diferentes versiones del sistema de la HNDM que se fueron desarrollando. Entre el 2005 y 2007 se evaluó una tercera versión del portal, que finalmente fue liberada en 2007. Si bien se refieren como “interfaces” mejoradas para los usuarios de la HNDM, la interfaz no es solamente una ventana al archivo físico, la interface es un sistema integrado de remediación del archivo (Cordell 2017, Hayles 2002). Como comenta Cordell: “Los investigadores frecuentemente utilizan la evidencia que encuentran en grandes colecciones digitales como si la hubieran descubierto utilizando métodos análogos [...] suprimiendo o por lo menos menospreciando que las interfaces web, el trabajo humano, y las tecnologías de búsqueda basadas en algoritmos, determinan qué se descubre (y qué no)”<sup>8</sup> (Cordell 2017, 192). Consideramos importante también tratar de entender con mayor detalle qué ocurría en cada nueva “interface” a nivel organización documental, base de datos, metadatos, posibilidades de búsqueda, por mencionar algunos ejemplos. Existe una noción de “mejoras”, pero es importante indagar cómo está compuesto lo que está detrás de la interface, no como construcción tecnológica, sino cómo impacta lo que el usuario encuentra o no.

Al no estar conectados de forma transparente ni el inventario o catálogo de los microfilmes con el catálogo o inventario de los periódicos originales impresos, la creación del sistema para la organización documental del sistema digital aparentemente se tuvo que desarrollar independientemente de los otros sistemas existentes, y el desarrollo que realizó la empresa de

digitalización no estaba integrado con las herramientas de organización documental y, por lo tanto, de acceso y recuperación que ya funcionaban en la HNDM.

Por lo tanto, parece ser que el primer sistema para la administración documental, así como la recuperación y acceso de la HNDM, estuvo enfocado y desarrollado, principalmente, canalizado al usuario y cómo utilizaría el sistema, y no tanto en la administración y gestión documental de la colección. Por lo tanto, hubo mucho énfasis en el desarrollo de la interfaz de usuario. Además, desde el prototipo y las posteriores versiones que existieron de la HNDM pudimos observar que fue cambiando el usuario al que estaba dirigido el sistema. En un principio se pensó en el sistema principalmente para investigadores, por lo tanto, se asumió que ellos ya tendrían conocimiento de la colección y que tendrían conocimiento del título de la publicación. Entonces se organizaron de tal forma para que: “El Ahuizote de mil ochocientos y tal, y no por colecciones. Título y fecha, lo principal” (E2). Sin embargo, unos años después cuando se trasladó la HNDM a los servidores de la UNAM ya se había detectado que había interés en la colección digital por un público más amplio y no tan especializado: “Y ellos requieren otras herramientas de búsqueda, por tema todo lo relacionado con el movimiento del 68, por ejemplo” (E2). Si bien se quiere ofrecer este tipo de recuperación esto actualmente no se puede hacer, no porque el sistema no lo permite, sino porque requiere de trabajo humano: “El sistema ya está preparado para manejar colecciones, pero no las maneja porque los datos no tienen ninguna asignación de este tipo, por eso no existe [...] no hubo personal de catalogación quien pudiera hacer ese trabajo” (E2). Además, si bien el sistema de búsqueda y recuperación quedó completo, los módulos de gestión necesarios para incorporar nuevas digitalizaciones así como otros aspectos de gestión no han sido terminados: “La última versión quedó un poco desprotegida [...]. Había planes de hacer cosas cuando se vino con Toda la UNAM en línea pero hubieron cambios (en la administración universitaria) y esto se vino abajo” (E1). Hay gran coincidencia entre los entrevistados de que es necesario personal para que el sistema pueda ser mejor aprovechado: “había que dejar el sistema preparado para el nuevo director y enriquecerlo con personal para darle forma y vida” (E1).

Uno de los principales atributos de las plataformas digitales en línea fue aumentar los usuarios potenciales ofreciendo la consulta a más personas. Sin embargo, al pensar en un sistema en línea, amplio y abierto, el tema de derechos de autor se agudizó. A esta problemática, Curiel y Jiménez le dedican considerable espacio en su artículo. Y aunque no se mencionan de forma explícita estas inquietudes, también evidenciaron otra problemática, la del “control” sobre el sistema y su contenido (las imágenes digitalizadas y su información metatextual). Aunado a esto encontramos un problema técnico de infraestructura que implicaba que el tiempo de respuesta de la HNDM fuera demasiado lento. Por lo tanto, se tomó la decisión de migrar el sistema -que se encontraba alojado con la misma empresa que realizaba la digitalización- a servidores de la UNAM, apoyado por la Dirección General de Tecnologías de Información y Comunicación (DGTIC). Esta migración no fue sencilla e implicó no sólo procurar la infraestructura necesaria (servidores, ancho de banda), sino también la creación de un nuevo “sistema” a donde se migró la anterior.

Fue en el 2011 que se empezó el trabajo para migrar el sistema con recursos específicos asignados para esta monumental tarea. En agosto se anunció el lanzamiento de Toda la UNAM

en línea<sup>9</sup> para noviembre y hubo interés expreso del rector de incluir a la HNDM en su lanzamiento. Esta premura significó que para la migración “se aceleran los trabajos” y “se revitalizan los grupos de trabajo” para que en noviembre se lograra salir con “una versión reducida de la HNDM”, la cuarta, en un servidor temporal. El artículo fechado en el 2012 indica que “en un corto plazo [...] la HNDM se encontrarán funcionando en plena capacidad de servidores propios” así como una revisión y mejoramiento del sistema mismo.

### *La HNDM como proyecto*

La HNDM es un proyecto de más de quince años, lo que representa en realidad mucho tiempo para plataformas digitales. En su momento uno de los primeros retos fue, como hemos visto, la infraestructura. Para un proyecto tan ambicioso y pionero “en ese momento la UNAM no tenía la infraestructura para hacer eso, obviamente teníamos a la gente pensante, los técnicos, pero no había recursos para empezar con esto en la realidad” (E1). En este sentido “curiosamente era más fácil conseguir el dinero que los recursos para hacerlo dentro” (E1). Además, en aquel entonces no existía un software ya desarrollado para la administración de colecciones digitales. La empresa mexicana que se contrató para el desarrollo del sistema había estado trabajando con una empresa canadiense. Después, las empresas se separaron y el trabajo de la HNM se quedó con la empresa mexicana: “Ellos (los canadienses) querían incorporar todo a un gran portal que ellos ya tenían, pero nosotros queríamos marcar algunas cosas diferentes, darle un poco de aspecto patrimonial a nuestro software, y entonces cambiamos de proveedor” (E1).

En un principio, las consultas a la HNDM que se encontraba en servidores de la empresa se hacían desde puntos estratégicos en Ciudad Universitaria: el 4º piso de la HNM, pero también la Dirección General de Bibliotecas y la Coordinación de Humanidades. Como se mencionó previamente, al principio se pensaba que los usuarios serían principalmente investigadores, pero con el tiempo empezó a incrementar mucho la demanda. La conectividad no era suficientemente buena y los usuarios se empezaron a quejar. Ya para ese momento, “la Universidad en cuanto a telecomunicaciones, tenía una mayor infraestructura que cualquier empresa en México” (E2), y se tomó la decisión de migrar todo a los servidores de la UNAM. Fue durante la migración que “identificamos problemas internos en la Hemeroteca, principalmente de rendimiento del sistema. Los problemas de imágenes y de metadatos eran también algo evidente, pero lo más evidente era el rendimiento del sistema” (E2).

El interés en traer la HNDM a la UNAM estaba también motivada y apoyada por la administración central que tenía el interés de incorporar la hemeroteca digital al proyecto Toda la UNAM en línea. Como la lentitud era mucha, se tomó la decisión técnica de hacer un nuevo sistema y migrar todo ahí. Desde el 2011 hasta el 2015, hubo apoyo de recursos humanos para hacer el desarrollo. Sin embargo, uno de los grandes problemas fue que la empresa no trabajó con base en estándares y eso dificultó los procesos. Se decidió migrar el sistema a una base de datos relacional y esto “implicó hacer un traslado de datos. Se cambió la estructura, no usaba base de datos [...] fue un cambio completamente radical” (E2). Aunado a esto se detectó que se requería un módulo de gestión, para que personal de la HNM pudiera ingresar nuevas digitalizaciones al sistema, y administrar, corregir, organizar, crear colecciones, etcétera. Esto

no era algo que el sistema de la empresa tenía. Pero entonces se acabaron los recursos y el tiempo, y quedó inconcluso el sistema: “Después de haber liberado en el 2015 la versión de consulta, se comenzó a trabajar en esta versión de gestión. Se trabajó del 2015 al 2016 sin embargo, no pudo concluirse y en el 2016 quedó suspendido” (E2).

Hasta el día de hoy el sistema de la HDNM continúa siendo un sistema que no está terminado. Puede argumentarse que con la funcionalidad con la que cuenta ahora es más una colección digital, que una Hemeroteca Digital: “Una biblioteca digital tiene una estructura de organización, una de procesos técnicos y de servicios y van de la mano [...] tienen que ver sus políticas de adquisición, políticas de rescate, de conservación de servicio, de colecciones, etc.” (E4). Si bien se pueden cargar nuevas imágenes digitalizadas a través del *back end*, desde finales del 2015 o principios del 2016 no se ha cargado ninguna imagen nueva, pese a que se llevan a cabo procesos de digitalización en la HNM: “Hay un problema entre administrativo y tecnológico que no nos ha permitido hacer esta actualización de la información” (E4).

La HNDM en su versión actual no es algo que esté terminado: “Algo que quiero que ustedes tengan bien claro: el portal que tenemos no es portal acabado, no hace todo lo que hacía antes, tiene más estabilidad y capacidad, pero hace menos cosas que antes [...] Se acabó el tiempo y el dinero” (E1). También es importante considerar que la migración del sistema, debió haber significado cambios en los procesos de la HNM, que aún no han sucedido. Desde el principio la HDNM se armó como un proyecto. El financiamiento para la digitalización y la migración del sistema fueron recursos extraordinarios. Actualmente no está definido cuál es el estado del proyecto. Si bien la consulta de la HNDM está en el 4° piso de la Hemeroteca Nacional con personal de la HNM y la administración del sistema en el Departamento de Informática, no hay designado un responsable, ni siquiera el coordinador de la HNM, que podría ser la figura pertinente: “no quedó vinculado a los trabajos ordinarios de la Hemeroteca Nacional” (E4). El proceso de incorporar a la HNDM a los procesos administrativos y sus flujos de trabajo diarios es de vital importancia.

## Conclusiones

La HNDM es un proyecto pionero y ambicioso que en sus más de quince años de trayectoria ha pasado por diversas etapas y momentos, impulsados no sólo por la tecnología disponible sino también por intereses administrativos y políticos. La labor de la creación de la HNDM es sin duda loable y el objetivo de esta investigación no es menospreciar el trabajo de más de una década de muchas personas. Como comenta uno de los entrevistados: “No intentamos poner la idea de que es bueno o malo, sino que, con el tiempo, vas viendo que se toman decisiones que en un momento se toman como coyunturales o necesarias [...] y ahora no podemos decir que fueron malas etapas, porque lo vemos en un contexto diferente” (E5). El propósito de este trabajo fue documentar y analizar críticamente desde una perspectiva integral cómo fue constituido el acervo digital haciendo un trabajo de análisis y descripción hemerográfica de la colección digital.

La HNDM se destaca por ser uno de los proyectos de digitalización más ambiciosos hasta la fecha y pionero, no sólo en México, sino en Latinoamérica y de las primeras naciones en el mundo. Con más de 7 millones de imágenes que representan sobre 900 títulos, la HNDM sin duda ha facilitado el trabajo de muchos investigadores, así como también acercar a nuevos usuarios a una rica e importante colección. Como se comentó al principio, el objetivo de esta investigación fue complementar la información disponible acerca de cómo está constituida la colección para así dar los primeros pasos para crear un marco interpretativo más robusto. También da cuenta de cuál es la situación actual de la HNDM y su proyección a futuro.

Es evidente que el tema de recursos humanos ha sido una problemática recurrente de las nacionales, tanto de la Biblioteca como de la Hemeroteca. Los procesos bibliotecarios requieren de recursos humanos para la organización documental y la automatización de las bibliotecas está basada en los trabajos anteriores, desde el control del inventario hasta la catalogación. No hay quien catalogue los periódicos, no hay quien catalogue los microfilmes, no hay quien catalogue lo digital, y se va acumulando el rezago. Los proyectos de digitalización deben plantearse de la mano con la organización documental, que permite no sólo llevar el control, pero también, posteriormente, ofrecer herramientas de descubrimiento mucho más eficaces para los usuarios. A su vez, existen trabajos de investigación hemerográficos descriptivos, como el Seminario de Biblioteca Mexicana del Siglo XIX que podrían integrarse dando un valor agregado importante a la colección digital: “lo que no se previó -aparte de la preservación- fue la organización documental. Entonces, lo que se controla a nivel de documentos digitales es muy frágil, prácticamente en los listados de títulos, los novecientos y tantos títulos, pero ya cerca de las existencias reales de imágenes, no lo sabemos” (E4). Quedó prácticamente desvinculado el control documental de los físicos, los microfilmes y los digitalizados.

Si bien la HNM ha realizado un trabajo arduo para cubrir los requerimientos de los procesos técnicos, la consulta y la preservación de sus colecciones, el presupuesto y la falta de recursos humanos son retos constantes para esta entidad nacional. El tema de la automatización de la Hemeroteca, en particular, la creación de los catálogos, resulta fundamental para comprender la relación entre el acervo físico, el microfilmado, el digitalizado y las bases de datos que registran la información acerca de estos materiales y si existe o no relación entre ellos. Los proyectos de automatización y digitalización de bibliotecas suelen basarse en trabajos previos y generalmente omisiones o deficiencias previas, no son resueltas, sino que continúan impactando en las nuevas plataformas. De la misma forma, el trabajo realizado adecuadamente y con estándares tiene un efecto beneficioso en futuros sistemas. En la HNDM vemos una combinación de ambos.

En su momento la creación de la HNDM como proyecto le dio mucha flexibilidad y la suficiente cantidad de recursos para emprender tareas de gran alcance y también tener una estructura de trabajo responsiva y dinámica. Hubo muchos cambios y ajustes en el trayecto que seguramente se pudieron hacer porque había flexibilidad administrativa. Sin embargo, esa misma independencia ahora ha significado que el proyecto esté “en el aire” y se encuentra en un momento crítico para integrarlo dentro de la estructura administrativa de la institución. Asimismo, la importancia de que se concluya el sistema y que entonces puedan atenderse temas

de control bibliográfico y catalogación y, en su momento, integrar de mejor forma la relación entre las distintas colecciones: microfilm, impresa y digital.

De la misma forma que la Hemeroteca Nacional está históricamente presionada por conseguir espacio para almacenar, catalogar y encuadernar su acervo que crece continuamente, la HDNM también conlleva responsabilidades y necesidades. Si bien el sistema aún no se termina, en su momento, será necesario actualizar o migrar al actual sistema conforme la tecnología vaya avanzando, sólo para mantener las facilidades que ahora ofrece. Se querrá también ir incorporando nuevas opciones de herramientas. Otro reto, por supuesto, es continuar incrementando el acervo y creando los mecanismos para que la HNDM forme parte de las colecciones y servicios que otorga la HNM. Por último, el correcto resguardo y preservación de los archivos digitales que conforman la HNDM se visualiza como apremiante. Con más de 7 millones de imágenes, más los XML correspondientes, el acervo digital, y la inversión millonaria que se realizó, corre el riesgo de perderse si no se establece un plan de preservación digital.

Para los investigadores interesados en la construcción de la bibliografía y la hemerografía mexicana es importante comprender las colecciones digitales no como substitutos o suplentes de la versión impresa. Es necesario entender que la computadora, las bases de datos, las interfaces, los algoritmos, etcétera, agregan capas interpretativas, generalmente, tácitas e invisibles que median nuestra relación con el objeto o los objetos estudiados. Es decir, existe una remediación ([Cordell 2017, 192](#)). Es importante entender la relación entre los materiales digitales y los documentos que representan y explorar el papel que tienen en la producción del conocimiento los materiales digitalizados y construir una historia social de la digitalización ([Mak 2014](#)). Es decir, no sólo cómo manejarlos en términos de acceso y preservación dentro de la biblioteca y el archivo, también explorar cómo contribuyen a la producción del conocimiento. Este artículo no pretende ser exhaustivo ni definitivo, al contrario, es de los primeros ejercicios que busca explicar, desde una perspectiva bibliográfica, la construcción de una colección digital.

## Anexo 1

### CUESTIONARIO PARA OCEANIC EXCHANGES (OCEX)

Historia de la Hemeroteca Nacional Digital de México (HNDM)

Objetivo:

El cuestionario no debe ser aplicado de forma literal, dependerá de la persona entrevistada, su grado de participación en el proyecto y creación de la colección digital, el acceso a documentación y la apertura del personal en detallar aspectos relevantes para el texto.

Algunas preguntas son:

1. Nombre:
2. Cargo que desempeña en la actualidad:
3. ¿De qué forma participó en la creación de la HNDM?

4. ¿Sabe usted si existe un documento inicial sobre los objetivos y alcances de la primera etapa de digitalización de los microfilmes? (Tratar de tener acceso a los archivos de la época, solamente interesa identificar datos generales, como cantidad, calidad y proceso de digitalización).
5. ¿Sabe o tiene conocimiento de cuántos proyectos se han elaborado a lo largo de todo el proceso de digitalización de los fondos microfilmados? (Guarda relación con la forma en que han sido agregados los periódicos al momento de creación de la HNDM y organización de las colecciones digitales).
6. ¿Conoce usted los criterios de microfilmación tenidos en cuenta para seleccionar los documentos hemerográficos a ser digitalizados? (Tanto en la fase inicial del proyecto como en la actualidad. Se debe entre otros aspectos, al uso de tecnología y elementos de archivística aplicados en distintos momentos).
7. ¿Tiene usted idea de cuál es el porcentaje digitalizado correspondiente a microfilmes de publicaciones periódicas del siglo XIX?
8. ¿Sabe quién tomó la decisión de digitalizar los fondos con las características que se hicieron?
9. ¿Sabe usted si la HNDM posee su propio cuadro clasificatorio de fondos?, ¿cuál es?
10. ¿Cuáles son las normativas para ingresar un periódico a la actual colección digital?
11. ¿Cuántos géneros o formas de texto/publicación tiene clasificadas la HNDM para el Fondo Reservado y qué se entiende por cada uno de ellos?
12. ¿Existe una relación de periódicos digitalizados por fecha de publicación?
13. ¿Qué idea de usuario se pensó para la elaboración de la página digital de la HNDM?
14. ¿Cuáles fueron los aspectos tenidos en cuenta para la forma en que el usuario podría intervenir al acceder a la página de búsqueda?

## Bibliografía

- CORDELL, Ryan. 2017. "Q i-Jtb the Raven': Taking Dirty OCR Seriously". *Book History* 20(1): 188-225. <https://doi.org/10.1353/bh.2017.0006>.
- CURIEL DEFOSSÉ, Guadalupe, Miguel Ángel CASTRO y Vicente QUIRARTE. 2019. *Publicaciones periódicas mexicanas del siglo XIX: 1856-1876 (Parte II)*. Vol. 2. Ciudad de México: Universidad Nacional Autónoma de México.
- CURIEL DEFOSSÉ, Guadalupe y Ricardo Javier JIMÉNEZ RIVERA. 2015. "Diez años de la Hemeroteca Nacional Digital de México (HNDM). Breve reseña de una larga gestión de preservación y acceso, 2002-2012". En *Textos, píxeles y bits: Reflexiones sobre la publicación digital*. Primera edición. Isabel Galina Russell: Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas, Biblioteca Nacional/Hemeroteca Nacional.
- CURIEL, Guadalupe y Miguel Ángel CASTRO, eds. 2003. *Publicaciones periódicas mexicanas del siglo XIX: 1856-1876 (Parte I)*. I. ed. Al siglo XIX, ida y regreso. México: Universidad Nacional Autónoma de México.

- El Universal*. 2004. "En marcha catálogo Nautilo en Biblioteca Nacional". 13 de septiembre de 2004. <http://archivo.eluniversal.com.mx/notas/244818.html>.
- ESCAMILLA, Gloria. 1965. "Los servicios de fotorreproducción en las bibliotecas". *Boletín de la Biblioteca Nacional* XVI(1-2): 43-50.
- FERNÁNDEZ FERNÁNDEZ, Íñigo. 2010. "Un recorrido por la historia de la prensa en México. De sus orígenes al año 1857". *Documentación de las Ciencias de la Información* (33): 69-89.
- FRASER, Hilary y Jerome MCGANN. 2015. "Nineteenth-Century Digital Worlds: Hilary Fraser Interviews Jerome McGann". 19: *Interdisciplinary Studies in the Long Nineteenth Century* (21). <https://doi.org/10.16995/ntn.758>
- Gaceta Digital UNAM*. 2016. "Tendrá la UNAM un Centro de Almacenamiento Documental" (4803) (11 de agosto de 2016). <http://www.gaceta.unam.mx/20160811/tendra-la-unam-un-centro-de-almacenamiento-documental/>
- HAYLES, N. Katherine. 2002. *Writing Machines*. Mediawork Pamphlet. Cambridge: MIT Press.
- HEMEROTECA NACIONAL DE MÉXICO. s.f. "Antecedentes". México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Bibliográficas. <http://www.hnm.unam.mx/index.php/quienes-somos/antecedentes> (Fecha de consulta: 13 de febrero de 2019).
- LATHAM, Sean y Robert SCHOLE. 2006. "The Rise of Periodical Studies". *PMLA* 121 (2): 517-31. <https://doi.org/10.1632/003081206X129693>
- LÓPEZ LANDEROS, Víctor Armando. 1998. "Servicios de prensa mexicana (SERPREMEX). Un nuevo servicio de de información en la Hemeroteca Nacional". Tesis de Licenciatura en Ciencias de la Comunicación, Universidad Nacional Autónoma de México. <http://132.248.9.195/pdbis/258190/258190.pdf>
- MAK, Bonnie. 2014. "Archaeology of a Digitization". *Journal of the Association for Information Science and Technology* 65(8): 1515-1526. <https://doi.org/10.1002/asi.23061>
- OLIVARES ALONSO, Emir. 2013. "El impacto de la Hemeroteca Nacional trasciende fronteras". *La Jornada*, 1 de octubre de 2013. <https://www.jornada.com.mx/2013/10/01/cultura/ao4nicul#>
- SUASTE LUGO, María del Rosario. 2017. "El desarrollo de los servicios en la Hemeroteca Nacional de México". *Biblioteca Universitaria* 20(1): 47-55.
- VIEYRA SÁNCHEZ, Lilia. 2015. "La Hemeroteca Nacional de México, 1912-2012: revisión historiográfica". *Boletín del Instituto de Investigaciones Bibliográficas* XX(1-2). <http://publicaciones.iib.unam.mx/publicaciones/index.php/boletin/article/view/777/748>

VILLALBA ROLDÁN, Ángel. 2000. “La catalogación de las publicaciones periódicas en la Hemeroteca Nacional de México informe académico: 1994-1998”. México: Universidad Nacional Autónoma de México.

## Notas

1 La minería de datos es el proceso que intenta descubrir patrones en grandes volúmenes de conjuntos de datos mediante herramientas computacionales.

2 En cómputo, una *ontología* es una definición formal de tipos, propiedades y relaciones entre entidades que se utilizan para organizar la información.

3 Un primer resultado es el “The Atlas of Digitised Newspapers and Metadata: Reports from Oceanic Exchanges”, Beals, M. H. y Emily Bell, con contribuciones de Ryan Cordell, Paul Fyfe, Isabel Galina Russell, Tessa Hauswedell, Clemens Neudecker, Julianne Nyhan, Mila Oiva, Sebastian Padó, Miriam Peña Pimentel, Lara Rose, Hannu Salmi, Melissa Terras y Lorella Viola. 2020. DOI:[10.6084/m9.figshare.11560059](https://doi.org/10.6084/m9.figshare.11560059)

4 “accounting for the sources, technologies, and social realities of their creation in ways that make their affordances and limitations more readily available for critique” (Todas las traducciones son de los autores).

5 “There is a whole other problem about trying to integrate the emerging digital archive with the inherited paper archive. We don’t really know how to do that yet. But it is imperative - because libraries will be organized digitally”.

6 OCR, siglas de reconocimiento óptico de caracteres. Se trata del archivo resultante del proceso de digitalización que “extrae” el texto contenido en una imagen.

7 XML, siglas de Extensible Markup Language. Es la extensión de los archivos que denotan un objeto con el lenguaje estandarizado y flexible. La HNMD ocupa archivos OCR XML en su funcionamiento.

8 “Scholars often use evidence discovered in large-scale digital archives as if they were discovered through analog means [...] eliding or at least downplaying the web interfaces, human labor, and algorithmic search technologies that shaped what is discovered (and what is not)”.

9 Toda la UNAM en línea es un programa institucional que se formalizó en 2011, y tiene como objetivo poner todos los contenidos digitales de la Universidad a disposición de los usuarios en todo el mundo, bajo una política de acceso abierto.