

Perla Olivia Rodríguez Reséndiz
María Teresa Fernández Bajón

Coordinadoras

Conectando los saberes de bibliotecas, archivos y museos (BAM) en torno a la preservación de documentos analógicos y de origen digital



**Z701.3
C65C65**

Conectando los saberes de bibliotecas, archivos y museos (BAM) en torno a la preservación de documentos analógicos y de origen digital / coordinadoras Perla Olivia Rodríguez Reséndiz, María Teresa Fernández Bajón. -- Ciudad de México : UNAM, Instituto de Investigaciones Bibliotecológicas y de la Información, 2018.

344 p.

ISBN: 978-607-30-1543-1

1. Preservación digital. 2. Bibliotecas. 3. Archivos. 4. Museos. I. Rodríguez Reséndiz, Perla Olivia, coordinadora. II. Fernández Bajón, María Teresa, Coordinadora.

Diseño de portada: Improitalia S.L.

Primera edición, 2019

D.R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, 04510, Ciudad de México

Impreso y hecho en México

ISBN: 978-607-30-1543-1

Publicación dictaminada

La publicación de este libro es resultado del proyecto de investigación IN402016 Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica PAPIIT de la Dirección General de Asuntos del Personal Académico de la Universidad Nacional Autónoma de México.

Tabla de contenido

Breve introducción a la preservación digital de los BAM	XI
Perla Olivia Rodríguez Reséndiz	

Prefacio	XIX
Jaime Ríos Ortega	

INICIATIVAS DE COLABORACIÓN

POLÍTICAS EUROPEAS SOBRE LA PRESERVACIÓN DE MATERIALES CULTURALES DIGITALES.....	3
María Teresa Fernández Bajón	

RETOS Y CONTEXTOS PARA LA CONSTRUCCIÓN DEL REPOSITORIO DIGITAL DEL PATRIMONIO CULTURAL DE MÉXICO	25
Ernesto Miranda Trigueros y Vania Ramírez Islas	

LA BIODIVERSITY HERITAGE LIBRARY (BHL), UN PROYECTO COLABORATIVO CON EL INSTITUTO DE BIOLOGÍA PARA LA PRESERVACIÓN DIGITAL DEL ACERVO HISTÓRICO.....	47
María del Socorro Tapia Tinajero y Rosa María Guzmán Vera	

COLABORACIONES PLURIDISCIPLINARIAS PARA LA CREACIÓN DE UN ARCHIVO DIGITAL. EL CASO DEL ARCHIVO DIGITAL DE LENGUAS Y CULTURAS ORIGINARIAS DE MÉXICO.....	63
Margarita Valdovinos y Perla Olivia Rodríguez	

EXPERIENCIAS DE PRESERVACIÓN

LA CUSTODIA DEI DOCUMENTI INFORMATICI IN ITALIA: IL MANUALE DI CONSERVAZIONE.....	77
Antonella Bongarzone	
ARCHIVAR Y DIFUNDIR EL PATRIMONIO ORAL DE LAS CIENCIAS SOCIALES Y HUMANIDADES EN UNA DINÁMICA DE COLABORACIÓN DE ARCHIVOS SONOROS	93
Véronique Ginouvès	
PROPUESTA DE MODELO DE PRESERVACIÓN DE FONDOS AUDIOVISUALES CON DSPACE PARA LA IMPLEMENTACIÓN DE REPOSITORIOS DIGITALES.....	109
Luis Rivera, Julio Rivera y Guadalupe Ramos	
REFLEXIONES EN TORNO A LA PRESERVACIÓN DE LOS DOCUMENTOS SONOROS DIGITALES.....	136
Mariela Salazar Hernández	
SISTEMATIZACIÓN DEL ACERVO FÍLMICO DE LA FILMOTECA DE LA UNAM.....	153
Gerardo León Lastra, Manuel Comi Xolot, Gustavo Lucio José y Luis Felipe Maciel Mercado	
UN ARCHIVO COMO MEDIO PARA TRANSFORMAR LAS PRÁCTICAS MUSEOLÓGICAS.....	173
Héctor Valverde Martínez y Alejandro Sabido Sánchez-Juárez	
EL ORBE DIGITAL, UN RETO PARA LA CAPACIDAD DE ESPACIO EN LOS SISTEMAS DE ALMACENAMIENTO DE INFORMACIÓN	193
Adriana Patricia Vélez García y Salvador Enrique Vázquez Moctezuma	

PATRIMONIO AUDIOVISUAL, UNIVERSIDAD COMPLUTENSE: SERVICIO DE INFORMACIÓN MULTIMEDIA-RED AUDIOVISUAL UNIVERSITARIA HISPANOMEXICANA	210
Alfonso López Yepes	

PRESERVACIÓN DE LA FOTOGRAFÍA DIGITAL A TRAVÉS DE LAS REDES SOCIALES: EL ESTUDIO DEL CASO CRISTERO.....	224
Sandra Peña Haro	

PROBLEMAS COMUNES

LA GESTIÓN DE DOCUMENTOS SONOROS Y AUDIOVISUALES. RETOS Y SOLUCIONES EN TIEMPOS DE CRISIS.....	244
Margarida Ullate i Estanyol	

ESCUCHAR Y ESPERAR. EL DESAFÍO DE LA DIGITALIZACIÓN Y PRESERVACIÓN DEL ACERVO SONORO DE RADIODIFUSIÓN NACIONAL DEL URUGUAY.	266
Fabricia Daniela Malán Carrera	

UNA MIRADA CRÍTICA A LAS DEPENDENCIAS DE GOBIERNO RESPECTO A LA CONSERVACIÓN Y PRESERVACIÓN DIGITAL EN EL ESTADO DE VERACRUZ	285
Adriana Martínez Cadena	

DIGITALIZAR PARA PRESERVAR LA DOCUMENTACIÓN DEL ARCHIVO HISTÓRICO MUNICIPAL DE MORELIA (AHMM)	303
Magali Zavala García	

El orbe digital, un reto para la capacidad de espacio en los sistemas de almacenamiento de información

ADRIANA PATRICIA VÉLEZ GARCÍA
SALVADOR ENRIQUE VÁZQUEZ MOCTEZUMA

Posgrado en Bibliotecología y Estudios de la Información, UNAM

INTRODUCCIÓN

Desde la antigüedad, la información ha formado parte del desarrollo de la humanidad, y por su importancia, se ha resguardado para el uso actual y de las futuras generaciones. Esta labor de preservación se ha realizado en las diversas unidades de información como las bibliotecas, los archivos, museos, entre otros. Sin embargo, los cambios que se han dado en los soportes de información debido al desarrollo tecnológico han llevado a que el resguardo y recuperación sean diferentes.

Actualmente, el mundo digital, que está presente en la vida de muchas personas, ya sea de forma directa o indirecta, alberga una innumerable cantidad de información que necesita ser almacenada para su posterior consulta. Esta tarea se ha enfrentado a múltiples desafíos, como la preservación, seguridad, espacio, organización y permanencia tanto a corto como a largo plazo. Por lo tanto, es indispensable

contar con sistemas de almacenamiento adecuados que faciliten la búsqueda, selección y recuperación.

Para lograr este propósito, las tecnologías de almacenamiento de datos masivos han evolucionado conforme a las necesidades que han surgido por la alta producción de información dentro del mundo digital. Esto ha llevado a que las unidades de información requieran de diversos sistemas que se adapten a sus requerimientos y que cuenten con la capacidad de almacenamiento precisa para resguardar sus datos. Además, el profesional de la información debe conocer la gama de sistemas de almacenamiento para evaluar sus opciones y elegir la más adecuada para su unidad.

Dentro de los sistemas de almacenamiento de datos masivos podemos encontrar del tipo tradicional, distribuido y en la nube. Tienen diversas características que se adaptan a los presupuestos, la cantidad de información y los datos que se deseen resguardar.

La preservación de los datos es una tarea indispensable para las unidades de información que resguardan la memoria de la humanidad. Por ello, la seguridad es un tema imprescindible, ya que la pérdida de bases de datos y otros recursos electrónicos obstaculizarían los servicios. Es ineludible establecer políticas para tener una seguridad adecuada, así como determinar qué tecnologías son apropiadas para almacenar los datos. Estas acciones se complementan con políticas de información y empleo de metadatos para la recuperación de la información. Ante este panorama, nos cuestionamos lo siguiente, ¿cuáles son las opciones de sistemas de almacenamiento con las que cuenta el profesional de la información para la preservación de la información digital? Ésta es una de las muchas preguntas que formulan los bibliotecarios, museólogos, archivistas, entre otros, sobre el tema.

Así, el objetivo principal de este trabajo es analizar los sistemas de almacenamiento de datos masivos para la preservación digital, que incluye el almacenamiento del tipo DAS (Direct Attached Storage), NAS (Network Attached Storage) y SAN (Storage Area Network); almacenamiento distribuido con características ACID y BASE, así como el teorema de CAP y la nube. Empleamos el método de investigación documental para realizar la búsqueda, recolección y análisis de la información.

1. PRESERVACIÓN Y SEGURIDAD DIGITAL

El incremento exponencial de la información que se ha dado dentro del mundo digital desde hace algunos años ha llevado a que la preservación de documentos y datos sea cada vez más compleja. El informe del International Data Corporation (IDC) (Gantz y Reinsel 2011) expone que, para el año de 2015, el total aproximado de datos creados y copiados en el mundo de manera general supera los ocho mil *exabytes*;¹ también estima que esta cantidad será rebasada por lo menos cada dos años. Esto nos da una idea del aumento de datos e información en el ambiente digital y nos esclarece que almacenar un acervo documental digital no necesariamente implica su preservación.

Desde finales del siglo XX, las tecnologías han cambiado de forma vertiginosa, por lo que existe el riesgo de llegar a un estado de obsolescencia que impediría la recuperación y consulta de información. Voutssás (2013) menciona que los registros y metadatos que son incorporados a los acervos digitales son factores importantes, pues no es suficiente contar con un sinnúmero de documentos digitales si posteriormente

1 Un *exabyte* equivale a 1 073 741 824 *gigabytes* o 10^{18} *bytes*.

no podrán ser encontrados. Por ello, el contenido de los documentos guardados debe ser cuidadosamente seleccionado, y debe haber una construcción ordenada y precisa de los metadatos que se usarán para la creación de los acervos digitales y con los que se recuperará la información almacenada.

En sus Directrices para la Preservación del Patrimonio Digital, la UNESCO menciona que: “La preservación digital designa los procesos a que se recurre con objeto de conservar información y cualquier otro tipo de patrimonio existente en forma digital” (UNESCO 2003, s.p.). Por su parte, Voutssás señala que la preservación digital son:

Acciones específicas cuyo fin ulterior y a largo plazo es el de asegurar la permanencia y acceso del contenido de documentos digitales a lo largo del tiempo y las tecnologías, independientemente de su soporte, formato o sistema. Para ello, debemos mantenerlos, esto es, protegerlos y resguardarlos anticipada y permanentemente; en caso de deterioro o daño debemos tratar de restaurarlos (Voutssás 2013, 20).

Naturalmente, para preservar documentos digitales se necesitan varias tareas que permitan acceder a la información y documentos a largo plazo para no caer en algún tipo de obsolescencia.

Una de las acciones importantes es establecer políticas de preservación pertinentes para cada unidad de información, que consideren diversos puntos como saber qué se quiere preservar, conocer el tipo de documentos digitales que se tienen, determinar la tecnología que se utilizará a partir de los fines buscados, crear respaldos de los acervos, tener un plan de mantenimiento y contra imprevistos, entre otros estándares que fortalecen tener documentos digitales recuperables (Jiménez 2009).

Las tecnologías utilizadas para el resguardo de los documentos digitales también son esenciales y se debe contar con ellas desde la creación de los acervos. Entre las características que deben poseer están la confiabilidad, el uso común y la adaptabilidad al tiempo, a fin de evitar un desfase en cuanto a la consulta y recuperación de información.

Aunado a esto se encuentra la seguridad de los datos, la cual tiene que ser tratada desde dos enfoques. El primero se relaciona con el acceso a los documentos y software, mientras que el segundo es la parte física y el hardware donde se almacena la información. De esta forma, la seguridad afecta en los sistemas de almacenamiento de forma íntegra, al incidir en la parte tangible e intangible.

Sin embargo, debe haber ciertas medidas para asegurar el acceso al contenido digital. A este respecto, Voutssás (2010) señala que dentro del ámbito de las unidades de información hay ciertas características que los documentos deben tener para ayudar a que la información se encuentre en un espacio confiable; éstas son permanencia, accesibilidad, disponibilidad, confidencialidad (privacidad), autenticidad (integridad) y aceptabilidad (no repudio). Además, recomienda que las unidades construyan programas, políticas y normas de preservación y seguridad desde la creación de los acervos. La movilidad de la información a través del tiempo se puede asegurar por medio de migración, candados y marcas digitales, que son tácticas de seguridad en las unidades.

Con relación a lo anterior, notamos que el almacenamiento es parte fundamental en las unidades de información para la preservación de la información digital.

2. SISTEMAS DE ALMACENAMIENTO

El almacenamiento de datos digitales empezó de forma masiva a mediados del siglo xx con la aparición de las primeras computadoras, las cuales eran capaces de guardar una gran cantidad de datos, funcionar a la par con bases de datos y atender a diversos usuarios de forma simultánea.

Actualmente existen varias formas de almacenar datos. Se usa por lo general la tecnología de unidades de disco duro o unidades de estado sólido que proporcionan mayor velocidad en la transferencia de información. Por otra parte, se están implementando nuevas formas de almacenamiento como la Memoria de Cambio de Fase (PCM), que provee mayor estabilidad, control, velocidad, almacenamiento y durabilidad de los datos (Gopalakrishnan *et al.* 2010). Conforme pasa el tiempo, los formatos de almacenamiento han ido evolucionando para optimizar el funcionamiento en la búsqueda, recuperación y acceso. Ante esta realidad, uno de los mayores problemas en el manejo de datos son las cantidades exorbitantes que se crean a cada momento, lo cual ha hecho que su administración se complique en las unidades de información que cuentan con información digital.

3. DAS, NAS Y SAN

Direct Attached Storage (DAS) es uno de los sistemas de almacenamiento más sencillos y cotidianos dentro de las conexiones directas. Usualmente, las unidades de disco son conectadas de forma directa a los servidores por medio de una interfaz de datos SCSI o IDE. Dentro de los beneficios de este tipo de conexiones está la facilidad de instalación, necesitan un software sencillo, sus costes de mantenimiento y la compatibilidad con

otros sistemas son bajos y tienen una tecnología probada. Empero, debido a su naturaleza, presenta algunas desventajas técnicas, por ejemplo, su dependencia directa al servidor donde se encuentran conectadas otras terminales puede bajar la eficiencia en la transferencia de datos de forma masiva y el intercambio de información. Este tipo de conexión es funcional en pequeñas unidades de información donde se utilice una conexión sencilla y de bajo costo para su almacenamiento y distribución de datos.

Por otro lado, dentro de los sistemas de almacenamiento, existen los de red, los cuales se caracterizan por tener una capacidad de expansión que permite tener un mejor intercambio de datos de forma masiva. Aquí se encuentran los NAS (Network Attached Storage) y SAN (Storage Area Network).

NAS permite tener varios servidores conectados por medio de una red IP. Este tipo de almacenamiento proporciona una red conectada a un nodo por medio de un servidor de archivos, pero también puede estar inmerso en la red. Este servidor de archivos administra los datos tanto de entrada como de salida en el disco y normaliza el acceso entre varios usuarios de la red.

NAS tiene varias ventajas como la disponibilidad de comunicación entre una computadora y el propio sistema debido a su conexión en red, lo que proporciona un intercambio y recuperación de datos con menos errores, además de que las copias de seguridad se construyen con mayor facilidad y hay mejor exactitud en las búsquedas de información. En contraste, algunas desventajas son el grado de dificultad de la configuración, una capacidad limitada si existe una gran cantidad de dispositivos conectados y el nivel de rendimiento baja cuando hay demasiada transferencia de datos o muchos usuarios conectados a la vez; no es recomendable para aplicaciones de bases de datos a gran escala (Tian 2006).

Se recomienda utilizar el sistema tipo NAS en infraestructuras pequeñas o medianas, con un número moderado de usuarios y cuando se busca una forma sencilla y de costo moderado para tener un acceso a los datos e información.

El SAN es un tipo de almacenamiento que usa una red flexible. Tiene una conexión de fibra óptica que optimiza la velocidad en la transferencia de datos, el almacenamiento y las comunicaciones entre diversos dispositivos son más seguros y se pueden trasladar grandes volúmenes de datos entre largas distancias de forma rápida y confiable. Esta red tiene numerosas ventajas, pero es indispensable contar con conocimientos adecuados, infraestructuras y presupuestos para tener un alto rendimiento. Se recomienda para las grandes unidades de información, con numerosos usuarios y alta transmisión de datos.

Los tipos de almacenamiento DAS, NAS y SAN se adaptan y resuelven diversas necesidades a partir de los datos que se deseen almacenar, el tipo de información y las necesidades que tenga la unidad de información en cuestión.

3.1. ALMACENAMIENTO DISTRIBUIDO

El sistema distribuido consiste en una colección de computadoras autónomas, unidas por una red de cómputo y equipadas con un software específico; ante los usuarios aparenta ser una única computadora (Colouris *et al.* 2012; Tanenbaum 1996). Así, la diferencia entre un sistema distribuido y una red de computadoras estriba precisamente en que el software constituye el sistema distribuido como un solo equipo virtual.

El primer desafío ante las grandes cantidades de datos es el desarrollo de un sistema de almacenamiento distribuido a gran escala para la conservación estratégica, el procesamiento

eficiente y el análisis de los datos (Chen 2014). Aunque se pueda asegurar la cantidad de espacio de almacenamiento, surge el inconveniente con la naturaleza del dato, debido a que es más complicado administrar un número mayor de datos que pocos. Para utilizar un sistema distribuido en almacenamiento de datos masivos deben tomarse en cuenta los factores de consistencia, disponibilidad y tolerancia de partición.

- **Consistencia:** Este tipo de sistema requiere múltiples servidores para la cooperación en el almacenamiento de datos. Es necesario asegurar que múltiples copias de los mismos datos sean idénticas sin importar en qué lugar de la red se encuentren.
- **Disponibilidad:** Un sistema distribuido opera en varios conjuntos de servidores. Es deseable que en caso de fallas moderadas, todo el sistema tenga pocas afectaciones con respecto al servicio que brinda a los usuarios.
- **Tolerancia de partición:** En un sistema de almacenamiento distribuido los servidores son conconectados por una red. Esto podría ocasionar problemas en los enlaces o nodos. Por lo tanto, el sistema debe tener un cierto nivel de tolerancia a problemas causados por divisiones en la red.

Para conseguir un sistema de almacenamiento distribuido eficiente es necesario que los servidores tengan consistencia entre sus copias de datos, que sirvan a la lectura y escritura de datos, además de dividir la red entre los servidores. Hay que considerar que existen tres diseños para la creación de los sistemas de almacenamiento distribuido que permitirán un mejor control de los datos: ACID, BASE y el teorema CAP.

ACID

Es un acrónimo de Atomicity, Consistency, Isolation and Durability. Se caracteriza por la integridad de datos en operaciones que son únicas. Este modelo fue generado principalmente en los sistemas de bancos (Al-Khanjari *et al.* 2013; Morabito 2014). En las aplicaciones, las transacciones de datos individuales se agregan a las transacciones de la aplicación. Las transacciones de datos básicos típicos son crear, leer, actualizar y eliminar; por su parte, las transacciones de aplicaciones siguen el principio ACID. Las partes que componen este sistema son:

- Atomicidad: Es la propiedad que asegura que la operación se ha realizado o no.
- Consistencia: Certifica que sólo se empieza aquello que se puede terminar. Sostiene que cualquier transacción llevará a la base de datos desde un estado válido a otro también válido.
- Aislamiento: Las transacciones aisladas nunca interfieren entre sí. Esto asegura que la realización de dos transacciones sobre la misma información sea independiente y no generen ningún error.
- Durabilidad: Garantiza que, una vez realizada alguna operación, ésta persistirá y no se podrá deshacer, aunque falle el sistema. De esta forma, los datos sobreviven hasta que se cambie por otra transacción comprometida.

ACID ha sido el estándar de algunos sistemas de datos distribuidos, pero, en aplicaciones en la nube, el estándar ACID es difícil, si no imposible, para satisfacer las operaciones a causa de la dispersión de datos.

BASE

Pritchett (2008) menciona que una alternativa para ACID es

BASE (Basically Available replicated Soft state with Eventual consistency). Es una metodología de desarrollo en la que los servicios se ejecutan en un solo centro de datos con una red confiable y están diseñados deliberadamente para usar potencialmente datos obsoletos o incorrectos, rechazan la sincronización a favor de una respuesta más rápida, aunque se corre el riesgo de tener inconsistencias.

BASE es opuesto a ACID, pues esta última brinda consistencia al final de cada operación; BASE tolera que la coherencia de la base de datos esté en un estado de flujo. Aunque lo anterior parece disfuncional e imposible, en realidad es manejable y conduce a niveles de escalabilidad que no se pueden obtener con ACID, lo que hace que la disponibilidad de BASE se logre mediante el apoyo a los fracasos parciales sin falla total del sistema.

TEOREMA CAP

El teorema CAP de Eric Brewer (2000) está formado conceptualmente por Consistency, Availability y Partition tolerance. De forma general, el teorema indica que un sistema distribuido no podía cumplir simultáneamente los requisitos de consistencia, disponibilidad y tolerancia de partición. Como máximo, dos de los tres requisitos pueden ser satisfechos a la par, es decir, se puede tener un sistema CA sin la tolerancia de partición o un sistema CP sin disponibilidad o un sistema AP que ignora la consistencia (Gilbert y Lynch 2002). En un sistema distribuido, la consistencia implica que todos los nodos vean la misma información al mismo tiempo, mientras que la disponibilidad garantiza que cada petición a un nodo reciba una confirmación de si ha sido o no resuelta satisfactoriamente y la tolerancia de partición significa que el sistema continuará trabajando a menos que haya un fallo total en la red.

El teorema CAP advierte que un sistema consistente nunca romperá reglas establecidas. Si un dato se establece con valor determinado para una transacción completa, aquel elemento de datos debe tener un valor de ajuste, sin importar de donde se acceda al elemento. No obstante, si el elemento se replica, no debe ser accesible hasta que el valor de ajuste esté disponible en todos los sistemas replicados. Por otra parte, la disponibilidad se describe generalmente como la relación del tiempo en la que un sistema es servible en el tiempo. La tolerancia de partición se refiere a las consecuencias de fallas en la comunicación entre los nodos de un sistema. Cuando se produce un lapso de comunicaciones, que es generalmente un fallo de red, el sistema se divide o reparte en grupos de nodos que pueden o no comunicarse entre sí.

3.2. ALMACENAMIENTO EN LA NUBE (CLOUD STORAGE)

El *cloud computing* es un equipo virtual que parece invisible al administrador pero tangible debido a que se puede especificar la cantidad de espacio de almacenamiento (Furht y Escalante, 2011). Esta elasticidad en la infraestructura es una de las técnicas utilizadas para el almacenamiento de grandes cantidades de datos (Sakr *et al.* 2011; Schadt *et al.* 2010). Las tecnologías de virtualización han hecho que la computación sea accesible, asequible y rentable (Vázquez 2015).

El almacenamiento en la nube es el espacio para copiar datos, información, objetos digitales, etcétera, a los cuales se accede por internet a través de un servicio web. Incluye el acceso a la red mediante un navegador provisto de recursos informáticos bajo demanda, con control variable para el usuario y neutrales ante sistemas operativos (Sosinsky 2011).

Estas características lo hacen único. El almacenamiento puede ser brindado por un proveedor de servicios (nube pública) o una versión privada (nube privada) en donde una organización particular la crea para uso interno, además de contar con un completo control de los recursos en tecnologías de información.

El servicio de almacenamiento en la nube significa que un proveedor renta espacio en su centro de almacenamiento a usuarios finales que carecen de almacenamiento propio. También se usa cuando no se dispone de personal técnico especializado en la administración de sistemas informáticos, o cuando se adolece de conocimiento para implementar y mantener infraestructura de almacenamiento.

El *cloud computing* ofrecen principalmente tres modelos de servicio, de acuerdo con el NIST: el primero es la infraestructura como servicio (IaaS), el segundo es la plataforma como servicio (Paas) y, por último, el software como servicio (SaaS). Pero para fines de almacenamiento en la nube el más adecuado es la modalidad IaaS, debido a que el proveedor ofrece al usuario recursos como capacidad de procesamiento, almacenamiento o comunicaciones, que el usuario puede utilizar para ejecutar cualquier tipo de software, desde sistemas operativos hasta aplicaciones.

Una de las grandes ventajas del almacenamiento en la nube es el ahorro de recursos económicos. El usuario únicamente paga por la cantidad de datos que transfiere y aloja en los servidores del proveedor. Cada *gigabyte* de alojamiento en promedio cuesta 0.30 centavos de dólar y con capacidad de renta de los 50 a 5000 *terabyte* (Vázquez 2015). Además, permite extensibilidad y escalabilidad en el almacenamiento de la información, que son necesarios cuando se manejan grandes cantidades de datos (Joyanes Aguilar 2012).

Existen varios proveedores como Microsoft, SUN, IBM, Google y Amazon, los cuales usan una interfaz web para almacenar grandes cantidades de datos y tener acceso en tiempo real a la información, y además tiene la ventaja de poder escalar la cantidad de espacio, cuenta con estándares de seguridad en Internet y su velocidad de transferencia de información dependerá del ancho de banda de cada unidad de información.

CONCLUSIONES

Los acervos digitales existentes en las unidades de información requieren de preservación y seguridad; al ser labores complicadas se necesitan políticas de preservación para saber qué es indispensable resguardar. Asimismo, es esencial disponer de un sistema de almacenamiento que permita la organización, búsqueda y recuperación de los documentos a largo plazo.

La seguridad se debe abordar de manera adecuada, ya que de ella depende el acceso a los documentos almacenados, lo cual nos llevará a tener un control de la información que se desea compartir. Por otro lado, las unidades de información que deseen preservar sus documentos digitales requerirán de un sistema de almacenamiento, así como de servidores adecuados que satisfagan las necesidades de información requeridas.

Los profesionales de la información deben considerar que un sistema óptimo dependerá del tipo de unidad de información, sus necesidades de almacenamiento, usuarios, infraestructura y presupuestos. Para ello, existen diversos sistemas de almacenamiento de datos masivos como son DAS, NAS, SAN, el almacenamiento distribuido y en la nube.

Contar con este tipo de conocimientos proporcionará a los profesionales de la información las herramientas adecuadas para hacer una buena selección en cuanto a la preservación, seguridad y almacenamiento de sus acervos digitales.

OBRAS CONSULTADAS

- Al-Khanjari, Z. A. *et al.* "DBSoft: a toolkit for testing database transactions". *Journal of Emerging Technologies in Web Intelligence*, vol. 5, núm. 3 (2013): 205-212.
- Brewer, E. A. "Towards robust distributed systems". Conferencia presentada en Proceedings of the nineteenth annual ACM symposium on Principles of Distributed Computing (PODC), Portland, Oregon, 2000.
- Chen, M. *et al.* *Big data: related technologies, challenges and future prospects*. London: Springer, 2014.
- Colouris, G. *et al.* *Distributed Systems: Concepts and Design*. Nueva York: Addison-Wesley, 2012.
- Furht, B. y A. Escalante. *Handbook of Cloud Computing*. Nueva York: Springer, 2011.
- Gantz, J. y D. Reinsel. "Extracting value from chaos". *IDC iView*, 2011. Disponible en mayo de 2017 en <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>
- Gilbert, S. y N. Lynch. "Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services". *ACM SIGACT News*, vol. 33, núm. 2 (2002): 51-59.
- Gopalakrishnan, K. *et al.* "Highly-scalable novel access device based on mixed ionic electronic conduction (MIEC) materials for high density phase change memory (PCM) arrays". En *Sympo-*

- sium on VLSI Technology*, 205-206. Hawaii: IEEE, 2010.
- Jiménez León, A. “Los metadatos, una herramienta para preservar al conocimiento digital”. En *Memoria del III Encuentro de Catalogación y Metadatos, 29-31 de octubre de 2008*. México: Centro Universitario de Investigaciones Bibliotecológicas, Instituto de Investigaciones Bibliográficas, 2009.
- Joyanes Águilar, L. *Computación en la nube. Estrategias de cloud computing en las empresas*. México: Alfaomega, 2012.
- Morabito, V. *Trends and challenges in digital business innovation*. Nueva York: Springer, 2014.
- Organización de las Naciones Unidad para la Educación, la Ciencia y la Cultura (UNESCO). *Directrices para la preservación del patrimonio digital*, 2003. Disponible el 8 de junio de 2017 en <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>
- Pritchett, D. “BASE: An Acid Alternative”. *ACM Queue*, May/June (2008): 48-55.
- Sakr, S. *et al.* “A survey of large scale data management approaches in cloud environments”. *IEEE Communications Surveys and Tutorials*, vol. 13, núm. 3 (2011): 311-336.
- Schadt, E. E. *et al.* “Computational solutions to large-scale data management and analysis”. *Nature Reviews Genetics*, vol. 11, núm. 9 (2010): 647-657.
- Sosinsky, B. *Cloud Computing Bible*. Nueva York: Wiley, 2011.
- Tanenbaum, A. *Sistemas operativos distribuidos*. México: Prentice Hall, 1996.
- Tian, L. “The selection and application of library storage technology programs”. *Library Work and Study*, vol. 5 (2006): 28-30.
- Vázquez Moctezuma, S. “Tecnologías de almacenamiento de información en el ambiente digital”. *e-Ciencias de la Información*, vol. 5, núm. 2 (2015): 1-18.

Voutssás Márquez, J. *Preservación del patrimonio documental digital en México*. México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 2013.

_____. "Preservación documental digital y seguridad informática". *Investigación Bibliotecológica*, vol. 24, núm. 50 (2010): 127-155.

Conectando los saberes de bibliotecas, archivos y museos (BAM) en torno a la preservación de documentos analógicos y de origen digital.

La edición consta de 100 ejemplares. Coordinación editorial, Carlos Ceballos Sosa e Israel Chávez Reséndiz; revisión especializada, formación editorial y revisión de pruebas, Paola Sanabria López. Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. Fue impreso en papel cultural de 90 gr en la Universidad Complutense de Madrid, España. Se terminó de imprimir el mes de abril de 2019.