IFLA WLIC 201
SINGAPORE

Presentado el 19 de julio de 2013

Dificultades jurídicas y técnicas del Archivo Web en Singapur[1]

Traducción al español del documento original en inglés: "Legal and Technical Difficulties of Web Archival in Singapore".
Traducido por: Dirección de Traducciones de la Biblioteca del Congreso de la Nación Argentina (traducciones@bcn.gob.ar)
Buenos Aires, República Argentina

Jhonny Antonio Pabón Cadavid[2]
Doctorando Universidad Victoria de Wellington, Nueva Zelanda. Dirección de e-mail: pabonjhon@myvuw.ac.nz

Johnkhan Sathik Basha[3]
Máster en Gestión del Conocimiento, Universidad Tecnológica de Nanyang, Singapur. Dirección de e-mail: sathikba001@e.ntu.edu.sg

Gandhimani Kaleeswaran[4]
Máster en Gestión del Conocimiento, Universidad Tecnológica de Nanyang, Singapur. Dirección de e-mail: kaleeswa001@e.ntu.edu.sg

This is a Spanish translation of "Legal and Technical Difficulties of Web Archival in Singapore". Copyright © 2013 by **Jhonny Antonio Pabón Cadavid, Johnkhan Sathik Basha** and **Gandhimani Kaleeswaran**. This work is made available under the terms of the Creative Commons Attribution 3.0 Unported License: http://creativecommons.org/licenses/by/3.0/

URI: http://library.ifla.org/id/eprint/217

Resumen:

El Archivo Web en la Biblioteca Nacional de Singapur es una de las acciones fundamentales para garantizar la preservación del patrimonio digital nacional. Para el año 2006, la Biblioteca Nacional de Singapur (NLS, por sus siglas en inglés) había lanzado una iniciativa para crear un Archivo Web llamado Archivo Web de Singapur (WAS, por sus siglas en inglés). Esta ponencia analiza los aspectos jurídicos y técnicos que actúan como obstáculos para la recopilación y gestión de las colecciones digitales creadas a través de los Archivos

---

[1] Este artículo es una versión resumida de nuestra ponencia final de investigación crítica presentada en la Universidad Tecnológica de Nanyang en 2013. Agradecemos a Ang Peng Hwa (Singapur, NTU) por su asesoramiento, a Peter Lor (Sudáfrica) y a Brian Opie (Nueva Zelanda) por sus comentarios y sugerencias.
[2] Doctorando, Universidad Victoria de Wellington, Nueva Zelanda. Máster (MSc) en Gestión del Conocimiento, Universidad Tecnológica de Nanyang, Singapur. Máster en Historia (2008), PUJ, Colombia. Ha trabajado como investigadora en Propiedad Intelectual en Colombia y Nueva Zelanda. E-mail: pabonjhon@myvuw. ac. nz
[3] Máster en Gestión del Conocimiento, Universidad Tecnológica de Nanyang, Singapur. Graduado en Ingeniería de la Producción, PSG College of Technology, Coimbatore, India (2008). Ha trabajado para IBM en Consultoría Informática en Chennai, India. E-mail: SATHIKBA001@e. ntu. edu. sg
[4] Máster en Gestión del Conocimiento, Universidad Tecnológica de Nanyang, Singapur. Graduado en Ingeniería en Tecnología Informática, Universidad Anna, Chennai, India (2008). Ha trabajado en la industria del software como ingeniero de aplicaciones. E-mail: KALEESWA001@e. ntu. edu. sg

Web en Singapur. WAS carece de un marco adecuado de protección de derechos de autor para la gestión de sus archivos digitales, incluidos el acceso y la preservación a largo plazo. Desde una perspectiva técnica analizamos los últimos avances y desafíos internacionales del archivo web. Tanto en el análisis jurídico como en el técnico nos concentramos en los usuarios de los archivos como principal leitmotiv cuando de las soluciones. Nuestras recomendaciones y conclusiones finales subrayan el hecho de que la NLS debe fomentar la investigación y la creación de conocimiento a través del desarrollo de las humanidades digitales y la implementación de soluciones técnicas y jurídicas que fomenten el Archivo Web.

Palabras clave: Archivo Web, WAS, Derechos de Autor, Depósito Legal, Singapur, Computación en la Nube, Computación Grid, Evolución Terminológica.

1 INTRODUCCIÓN

El archivo web es un área de estudio reciente que surgió a fines del siglo 20 con el propósito de evitar un agujero negro digital en la historia mundial. La publicación cada vez más frecuente a través de la World Wide Web de información concebida originalmente en formato digital y el reconocimiento internacional de dichas publicaciones como parte del patrimonio cultural digital han alertado a los gobiernos acerca del significativo valor nacional de preservar las publicaciones web.

En Asia solo Corea del Sur, Japón, China y Singapur están desarrollando archivos web nacionales en sus respectivas bibliotecas nacionales. El archivo web es una preocupación internacional. Su desarrollo depende de los valores culturales, jurídicos, económicos y políticos de cada país, los cuales definen el alcance, los objetivos, las funcionalidades y la importancia del archivo web.

La complejidad del archivo web conlleva nuevos desafíos jurídicos para la gestión de colecciones digitales en el plano nacional e internacional, especialmente en lo relativo a la legislación sobre depósito legal y sobre derechos de autor. Estas dificultades cubren el espectro que va desde la obtención de contenidos hasta la utilización del Archivo Web por parte de los usuarios.

2 ASPECTOS ÉTICO-POLÍTICOS DEL ARCHIVO WEB

El Archivo Web presenta varias cuestiones éticas en las diferentes áreas de su ciclo de gestión de la información. ¿Qué es lo que debe preservarse? ¿Quién asume la responsabilidad de la preservación? ¿Quién decide quién tiene acceso, en qué circunstancias y a qué materiales? Estas preguntas ponen de relieve las prácticas de archivo anteriores al Archivo Web, pero algunas características de la información digital y de los Archivos Web plantean cuestiones éticas específicas.

Una gran diferencia entre las publicaciones impresas y las digitales es el sentido de permanencia, que es un "principio esencial del patrimonio cultural: la *raison d'être* de la recopilación es mantener la identidad cultural y consolidar los recursos—las colecciones culturales y de investigación—que permiten el enriquecimiento cultural, facilitan la investigación y generan mayores beneficios sociales y económicos para la sociedad que apoya y financia la actividad de recopilación" (Mason, 2007). Las publicaciones impresas temporarias se han archivado en pequeña escala, pero en el Archivo Web una gran cantidad de publicaciones digitales que por su propósito tienden a ser temporarias podrían recopilarse y preservarse en el largo plazo (Lor & Britz, 2012): por ejemplo, los mensajes de twitter; los comentarios hechos en foros; las opiniones de artículos; los *posts* de Facebook y otros contenidos generados por los usuarios cuyos creadores no esperan que sean preservados por instituciones de archivo con fines intergeneracionales.

Otra cuestión que puede surgir del archivo web es la censura, a través de la cual se ejerce control sobre la producción intelectual de la Web. Por ejemplo, un caricaturista y *blogger* singapurense recibió una notificación de la Biblioteca Nacional respecto del archivo de su sitio web y su blog y expresó su temor a ser vigilado a través del archivo de un modo cómico: (El archivo es) "malo en el sentido de que tengo que ser cuidadoso con lo que publico. Jajaja..." (Portrait Workshop, 2007). Además, la censura a través del archivo web puede realizarse mediante la limitación de acceso a los investigadores; la eliminación

intencional de sitios web o la negligencia al recopilar y preservar ciertos sitios web por motivos políticos o de otra índole también es una forma de censura que afecta a los investigadores existentes y futuros. Por ejemplo, la legislación sobre derechos de autor podría utilizarse con motivaciones políticas para ordenar la eliminación y remoción de material archivado.

En algunas situaciones, los archivos web son proyectos de investigación que involucran cuestiones extremadamente sensibles, como la creación de archivos web de la web oscura (*dark web*) para el análisis del terrorismo (Chen et al., 2008). Los archivos web nacionales y temáticos deben ajustarse a pautas éticas que garanticen la minimización de posibles daños.

3 LA IDENTIDAD NACIONAL Y LA BIBLIOTECA NACIONAL DE SINGAPUR

En los últimos años, el desarrollo de los servicios de la Biblioteca Nacional de Singapur (NLS) ha estado en sintonía con el propósito de "nutrir una sociedad de aprendices vitalicios que puedan acelerar la creación de capital intelectual y un nuevo ciclo de innovación nacional" (Seet, 2005, p. 151). El archivo web como nueva herramienta para la Biblioteca Nacional se orienta a la concreción de este objetivo.

El objetivo del Archivo Web administrado por la Comisión de Bibliotecas Nacionales (NLB, por sus siglas en inglés) es crear una colección de información digital originada en Singapur que ha sido publicada en Internet, que registra diferentes facetas de la cultura y el patrimonio de Singapur. Además, se espera que el Archivo Web contribuya a:

generar un sentido de comunidad, identidad nacional y arraigo entre los singapurenses. Esto puede lograrse mediante el archivo de información que da forma a la identidad nacional. La web se utiliza cada vez con más frecuencia como herramienta para la comunicación y la interacción social. Con el tiempo, formaría un registro de eventos que capturaría el ámbito de una nación y que describiría el desarrollo y la evolución de la identidad nacional de Singapur. El archivo de este registro constituye una invaluable fuente de patrimonio documentado para las generaciones singapurenses presentes y futuras. Estos conocimientos crearán un sentido de comunidad y pertenencia, sentimientos comunitarios fomentados en forma conjunta por un archivo coherente y de buena calidad" (Comisión de Bibliotecas Nacionales, 2012).

El Archivo Web pone de relieve los nuevos desafíos de la información digital, especialmente en lo que respecta a nuevos campos de estudio, como el patrimonio digital y las humanidades digitales. Las Humanidades Digitales constituyen una nueva disciplina que aplica técnicas computacionales a la investigación y la comunicación en las humanidades (Hayles, 2012).

Los archivos por sí mismos no crean un sentido de comunidad y nación. Es el acceso y el uso a los archivos lo que puede lograr este resultado. Un ejemplo de investigación sobre identidad nacional en Singapur en el que se utiliza información web es el artículo publicado recientemente con el título "Singapur: La Política de Inventar la Identidad Nacional", que muestra con claridad el modo en que Internet ha sido utilizada por lo singapurenses para repensar y expresar sus ideas de nación (Ortmann, 2009).

El hecho de cuestionar el modo en que la NLB define el patrimonio digital y el modo en que se nutren las humanidades digitales en Singapur contribuirá a identificar las principales características que necesitan los servicios de la Biblioteca Nacional de Singapur para producir los beneficios ofrecidos por el Archivo Web. Sin políticas claras para el archivo web y para acceder a los contenidos del Archivo Web, el propósito de WAS referido a la identidad nacional no podrá lograrse.

4 EL ARCHIVO WEB EN SINGAPUR WAS)

4.1 Alcance Actual del Depósito Legal en Singapur

Con la creación de la Comisión de Bibliotecas Nacionales (NLB, por sus siglas en inglés) en 1995, las funciones de depósito legal en Singapur fueron transferidas a la NLB, en las instalaciones de la Biblioteca Nacional de Singapur (NLS, por sus siglas en inglés). Una de las funciones de la NLB es "adquirir y mantener una colección completa de los materiales bibliográficos relativos a Singapur y su gente" (Artículo 6, Ley de la NLB, Capítulo 197, 1995). El material bibliográfico se define en sentido amplio en el Artículo 2, e incluye

cualquier material o dato que pueda ser reproducido (Artículo 2). La interpretación de este artículo es que el depósito legal en Singapur abarca la información digital que se distribuye a través de medios físicos (como DVD, CD, etc.) pero no la información disponible en línea, que está fuera del alcance de la legislación vigente.

En 2004, la NLB creó un Grupo de Trabajo sobre el Depósito Legal con el objetivo de revisar la legislación y desarrollar una estrategia para ampliar la colección del patrimonio documental, incluida la información disponible en línea (Foo, 2005). El Grupo de Trabajo elaboró un estudio comparativo de la legislación relativa a depósito legal, en el que incluyó cuestiones relacionadas con el depósito electrónico (*e-deposit*) y el archivo web. Recomendó la introducción inmediata del depósito electrónico voluntario y, en el largo plazo, de un nuevo marco de depósito legal que modifique la Ley de la NLB para que incluya el material disponible en línea como parte del patrimonio sujeto a depósito legal (Informe de Singapur a la CDNL, 2005-2006). Después de ocho años, esta reforma aún continúa pendiente.

4.2 Desarrollo del Archivo Web en la Biblioteca Nacional de Singapur

En 2005, el Centro de Investigación sobre Internet de Singapur (SIRC, por sus siglas en inglés) comenzó a generar el Archivo Web del Tsunami Asiático en conjunto con el Archivo de Internet y WebArchivist.org. Algunos miembros del Grupo de Trabajo, especialmente académicos de la Universidad Tecnológica de Nanyang (NTU, por sus siglas en inglés), participaron en este proyecto. El proyecto duró dos meses y archivó aproximadamente 1.600 sitios web de 40 países diferentes en 13 idiomas distintos. En la segunda etapa de preparación para el proyecto se decidió, como política de derechos de autor, implementar un procedimiento de notificación y remoción (Wu & Heok, 2005). Actualmente, el sitio del Archivo Web del Tsunami Asiático http://tsunami.archive.org/ no está disponible y muestra el siguiente mensaje: *HTTP error 404 - The page cannot be found* [HTTP error 404 – La página no puede encontrarse.] Además, este archivo web no se ha utilizado con fines de investigación ni con ningún otro objetivo.

En 2006, gracias a la experiencia adquirida a partir del Proyecto del Tsunami Asiático y de las recomendaciones del Grupo de Trabajo, la NLS creó el Archivo Web de Singapur (WAS). El WAS comparte el objetivo de la NLS tendiente a reunir materiales relacionados con la identidad nacional de los habitantes de Singapur. El criterio de valoración del WAS consiste en reunir sitios web de importancia nacional que podrían ser útiles para futuros investigadores (Chellapandi & Lian San, 2009). El WAS adoptó un modelo de archivo selectivo y temático que ha seleccionado alrededor de 1000 sitios web, especialmente de organismos gubernamentales y organizaciones oficiales. Además, para eventos tales como las Elecciones Generales de Singapur de 2006, se recopilaron más sitios web (Chellapandi & Lian San, 2009).

En 2007 la NLB comenzó a recopilar los sitios de dominio .SG, también selectivamente (en 2009 la NLS tenía capacidad para archivar 500 sitios web por mes). El Centro de Red de Información de Singapur (SGNIC, por sus siglas en inglés) es la organización que se encarga de la administración del dominio nacional .SG. Existen más de 100 000 dominios .SG registrados. En 2009, la NLB celebró un Memorándum de Entendimiento con el SGNIC para facilitar el acceso al listado de dominios .SG (Chellapandi & Lian San, 2009). Gracias a la utilización de la Plataforma Nacional Grid se han archivado más de 20 000 sitios web.

Puesto que no existe una ley de depósito legal que permita a la NLS llevar a cabo la Recopilación Web, la NLS ha adoptado una política de remoción y en algunos casos envía notificaciones (a través de cartas o correos electrónicos) a los administradores de sitios web sobre el archivo de, y el acceso a, los materiales (Chellapandi & Lian San, 2009).

El Grupo de Trabajo tomó como punto de referencia el congreso del Archivo Web llevado a cabo en Australia en 2004 (Foo et al., 2005). El desarrollo del WAS replicó algunas características del proyecto PANDORA de Australia, por ejemplo, el enfoque selectivo, y la interfaz de taxonomía para navegar un archivo web. Además, el WAS opera sin un marco legal adecuado que facilite la selección, el rastreo y la indexación de todos los sitios de dominio nacional y la copia de cualquier tipo de material de interés para la Biblioteca Nacional.

A partir de la publicación del Informe Biblioteca 2010 (Comisión de Bibliotecas Nacionales, 2005) la NLS comenzó una serie de iniciativas para ofrecer un acceso más amplio a los recursos electrónicos, entre las cuales se encuentran las siguientes: la creación de una

Biblioteca Nacional Digital (que incluye la digitalización del patrimonio documental que forma parte de sus colecciones de depósito legal existentes), la enciclopedia Infopedia de Singapur, NewspaperSG (en 2007 la NLS firmó un acuerdo de derechos de autor con Singapur Press Holdings para digitalizar "Straits Times" y poner el periódico a disposición del público a través de internet) y el proyecto en desarrollo "Singapore Memory" (Chellapandi, Han, & Boon, 2010). En los últimos años, otras instituciones como el Archivo Nacional de Singapur han cumplido un rol fundamental en la difusión del patrimonio documental (Beasley & Kail, 2009).

La misión de la NLB indicada en el Informe Biblioteca 2010 "consiste en difundir los conocimientos mundiales en Singapur para crear un impacto social y económico positivo" (Comisión de Bibliotecas Nacionales, 2005). No obstante, el Archivo Web y los materiales concebidos originalmente en formato digital – partes fundamentales del patrimonio digital de Singapur – no estaban mencionados en ese informe y no habían sido incluidos en los planes estratégicos desarrollados por la NLB en 2005.

4.3 Desafíos planteados por los derechos de autor del Archivo Web de Singapur

La gestión de registros digitales del archivo web conlleva muchas cuestiones relacionadas con los derechos de autor de la recopilación, la preservación, el acceso y la utilización de esas colecciones.

Los tratados y acuerdos internacionales sobre propiedad intelectual que Singapur ha suscripto – tales como la Convención de Berna, los acuerdos sobre los ADPIC y el Acuerdo de Libre Comercio con Estados Unidos (USSFTA, por sus siglas en inglés)– dan un marco a la ley sobre derechos de autor que dicho país ha establecido. Sin embargo, últimamente la NLB manifestó que Singapur está trabajando en pos de la flexibilidad de la ley sobre derechos de autor:

"La NLB está trabajando con la Federación Internacional de Asociaciones de Bibliotecarios y Bibliotecas (IFLA), la Organización Mundial de la Propiedad Intelectual (OMPI) y la Oficina de Propiedad Intelectual de Singapur (IPOS, por sus siglas en inglés) para obtener cambios en la legislación sobre derechos de autor con el propósito de educar y sensibilizar a los legisladores acerca de la necesidad de excepciones y exenciones para las bibliotecas en la ley sobre Derechos de Propiedad Intelectual, tanto a nivel internacional como a nivel nacional en Singapur (Ley de Derechos de Autor). La NLB está revisando su actual Ley para incorporar los depósitos legales digitales." (NLB Singapur, 2010).

No está claro cómo la NLB y la IPOS ejercen presión internacionalmente, si así fuera, pero a nivel nacional las modificaciones respecto de los depósitos legales han estado pendientes durante los últimos 10 años.

La Ley sobre Derechos de Autor de Singapur establece que cualquier reproducción de un trabajo protegido es una infracción a menos que exista una autorización por parte del titular de los derechos o una limitación legal a los derechos de autor. Esta Ley no establece ninguna excepción o limitación específica para el Archivo Web ni para la preservación digital. No existe una certeza jurídica sobre la legalidad de un archivo web en Singapur conforme a la cláusula de uso legítimo (*fair dealing clause*).

El Artículo 35 de la Ley sobre Derechos de Autor de Singapur establece una cláusula general de uso legítimo que permite criterios flexibles para evaluar una ley sobre no violación a los derechos de autor dependiendo de los siguientes factores (Artículo 35 (2)): el objeto y la naturaleza del uso, ya sea con fines comerciales o educativos sin fines de lucro; (b) la naturaleza de la obra o adaptación; (c) la extensión e importancia de la parte copiada con relación a toda la obra o adaptación; (d) el efecto del uso sobre el mercado potencial o el valor de la obra o adaptación; y (e) la posibilidad de obtener la obra o adaptación dentro de un plazo razonable y a un precio comercial regular.

4.3.1 1. Colección

Cuestiones relacionadas con los Derechos de Autor

Como fue mencionado anteriormente, el WAS ha atravesado muchas etapas de desarrollo, desde un enfoque muy selectivo de sitios hasta un abordaje más abarcativo de recopilación de sitios de dominio nacional, en su totalidad o en parte. La recopilación de

sitios web conlleva la reproducción de los contenidos. Durante la selección y el rastreo e indexación de sitios web, la NLS adoptó una política de remoción, capturó sitios web y envió a los administradores de sitios web una "Notificación de Archivo Web" que estableció algunas ventajas: "Como administrador de un sitio web, usted cuenta con algunos beneficios si la NLB archiva su sitio web y sus publicaciones online. Podrá, asimismo, acceder a versiones anteriores de sus sitios en el archivo web. La NLB también catalogará su publicación, y aumentará así la difusión de su sitio web entre los investigadores a través de la red de nuestra Biblioteca." Además la notificación indica lo siguiente: "La NLB se reserva el derecho de remover cualquier tipo de material del archivo web que considere una violación a los Derechos de Autor." En consecuencia, cuando se confecciona un archivo de materiales web no se evalúa ninguna cuestión relacionada con los derechos de autor. Sin embargo, posteriormente, a través de una política de remoción, el material puede ser retirado del archivo a pedido de alguna persona que así lo requiera o de la NLB. Los Países Bajos siguen una política de remoción similar (Glanville, 2010).

La política de remoción es la implementada por el Archivo de Internet, y ha sido propuesta como una política beneficiosa para los archivos web de Estados Unidos (Patel, 2007). Sin embargo, existe el riesgo de que pueda presentarse un litigio debido a esta incertidumbre legal que se puede evitar por completo con una ley de depósito legal adecuada, como las que se han sancionado en Nueva Zelanda, en Inglaterra o en Francia. Es importante destacar que el depósito legal no puede ser una solución absoluta porque tiene limitaciones territoriales y tecnológicas; por ejemplo, en Nueva Zelanda, la Biblioteca Nacional solicitó autorización para recopilar contenido del exterior y contenido local protegidos con contraseñas (Paynter & Mason, 2006).

Con respecto al sitio web del WAS, su política actual consiste en excluir del archivo el material no autorizado para su recopilación en materia de derechos de autor: "Asimismo, observe que las imágenes y el contenido que pertenecen a un tercero que es titular de los derechos de autor no pueden ser capturados como parte del archivo debido a cuestiones relacionadas con este tipo de derechos" y "la NLB no captura ni archiva información incluida en otros dominios o servidores" (WAS, NLB, 2013a).

4.3.2 2. Retención y Preservación

Cuestiones relacionadas con los Derechos de Autor

El WAS adoptó una política de remoción en la que los titulares de derechos de autor pueden solicitar a la NLB que retire el material archivado. La política de remoción es muy similar a la que ha sido adoptada por el Archivo de Internet. La NLB estableció un procedimiento de notificación de acceso limitado por parte de los titulares de los derechos de autor y, luego de la investigación, cuando "los fundamentos de la demanda son admisibles, el material será retirado definitivamente del repositorio" (WAS, NLB, 2013b). Sin una legislación que regule los depósitos legales, la retención permanente del material recopilado no está garantizada y depende del consentimiento de los titulares de los derechos de autor.

Los Archivos Web de las Bibliotecas Nacionales son parte del patrimonio cultural del país y, por eso, las colecciones digitales deben ser protegidas a través de una estrategia de preservación a largo plazo. La preservación digital conlleva determinados usos de los contenidos digitales que afectan diferentes derechos de autor exclusivos. Las restricciones a los derechos de reproducción, transformación y, en algunos casos, "difusión" son necesarias si la preservación digital intenta alcanzar los objetivos que se le han propuesto [Programa Nacional de Infraestructura y Preservación de Información Digital (Estados Unidos), Comisión Conjunta de Sistemas de Información, y el Proyecto de Ley de Acceso Abierto al Conocimiento (OAK, por sus siglas en inglés)].

Las disposiciones de la Ley sobre Derechos de Autor de Singapur relacionadas con la preservación son obsoletas e insuficientes en el reino digital. Los Artículos 48 y 113 de la ley mencionada anteriormente autorizan a la Biblioteca Nacional a reproducir una obra con fines de preservación. Sin embargo, esta excepción parece insuficiente si consideramos que la preservación digital a largo plazo requiere el cambio de formato y la emulación; ambas actividades incluyen la adaptación-transformación de la obra que, sin una excepción legal, puede provocar una violación de los derechos de autor. Tal como lo ha indicado un académico

de Singapur, "un Archivo no garantiza definitivamente que sus procesos de ingreso y almacenamiento de datos sean de uso legítimo" (Seng, 2008).

4.3.3 3. Acceso y uso de los Archivos Web

Aunque la recopilación y la preservación de información de los archivos web nacionales planteen cuestiones administrativas y técnicas complejas, el acceso no es diferente. Una de las áreas particularmente sensible a este tema es la de los derechos de autor.

El WAS brinda acceso a una pequeña parte de su colección. De acuerdo con una reciente entrevista realizada a la NLB, el personal manifestó que está a la espera de una modificación a la Ley de la NLB que le permita poner a disposición del público una amplia gama de materiales archivados a través de internet (Yong Fu, 2013). Las disposiciones relacionadas con la preservación en la ley sobre derechos de autor de Singapur son insuficientes para el acceso a largo plazo, permiten una amplia interpretación para autorizar solo las reproducciones de la obra, e impiden la difusión de las colecciones del archivo web (Seng, 2008).

Los sitios a los que se puede acceder en el sitio web del WAS son los que han autorizado a la Biblioteca Nacional o pertenecen al gobierno. En la notificación de Archivo la NLB manifiesta a los administradores del sito web que: "La NLB se ha embarcado en un proyecto de archivo web cuyo objetivo a largo plazo es crear una colección amplia de sitios web y publicaciones relacionadas con Singapur para garantizar el acceso por parte de los singapurenses a su patrimonio cultural presente y futuro. La NLB considera que su sitio web xxx es una parte importante del patrimonio documental de Singapur y desea que siga estando disponible para los investigadores y las futuras generaciones de Singapurenses" (Apéndice 1). Sin embargo, este acceso podría interrumpirse por el uso de la política de remoción.

La única biblioteca nacional que permite el acceso abierto a su Archivo Web es la Biblioteca Universitaria Nacional de Islandia que permite acceder al Archivo Web de Islandia (www.Vefsafn.is) y sólo restringe el acceso al material pago e implementa una política de acceso limitado.

El motivo de la restricción relativa al acceso a los Archivos Web es que la ley sobre derechos de autor no permite el acceso al material sin la previa autorización del titular del derecho de autor y que el uso puede afectar el interés económico de los tenedores de derechos de autor. Sin embargo, no existen pruebas empíricas de que el acceso a los Archivos Web tenga un impacto económico adverso (Programa Nacional de Preservación Digital e Infraestructura de la Información (Estados Unidos et al, 2008).

Los países con Archivos Web que brindan acceso in situ, lo autorizan con fines de investigación o estudio, como excepciones al derecho de autor para educación e investigación. En estas circunstancias el uso del Archivo Web se ha limitado porque la falta de flexibilidad de de la ley sobre derecho de autor impide el uso de herramientas que las humanidades digitales han adoptado en los últimos años. Los investigadores de la Biblioteca Nacional de Francia, por ejemplo, tienen restricciones con respecto al uso del Archivo Web francés para el análisis de minería de datos (Stirling, Illien, Sanz, & Sepetjan, 2012).

Los archivos web se han utilizado en muchos proyectos de investigación como análisis de acciones sociales, historiografía web, e impactos éticos de los mismos (Dougherty et al., 2010). El uso de los archivos web dependerá de la usabilidad, integridad y publicidad de los archivos.

El WAS y la NLB no han brindado a los ciudadanos información sobre el proyecto. Una comunidad singapurense de usuarios de Wikipedia pone de manifiesto esta situación de la NLB al describir su experiencia: "Realizamos varios llamados a la NLB para averiguar sobre esta cuestión del patrimonio, nadie sabe… Si se supone que es nuestro patrimonio, debería ser conocido por la mayor cantidad de personas posible, y fui transferido toda la mañana de un departamento a otro, y nadie sabía de qué se trataba todo esto" (SG Wiki Editor, 2007).

4.3.4 Obras Huérfanas

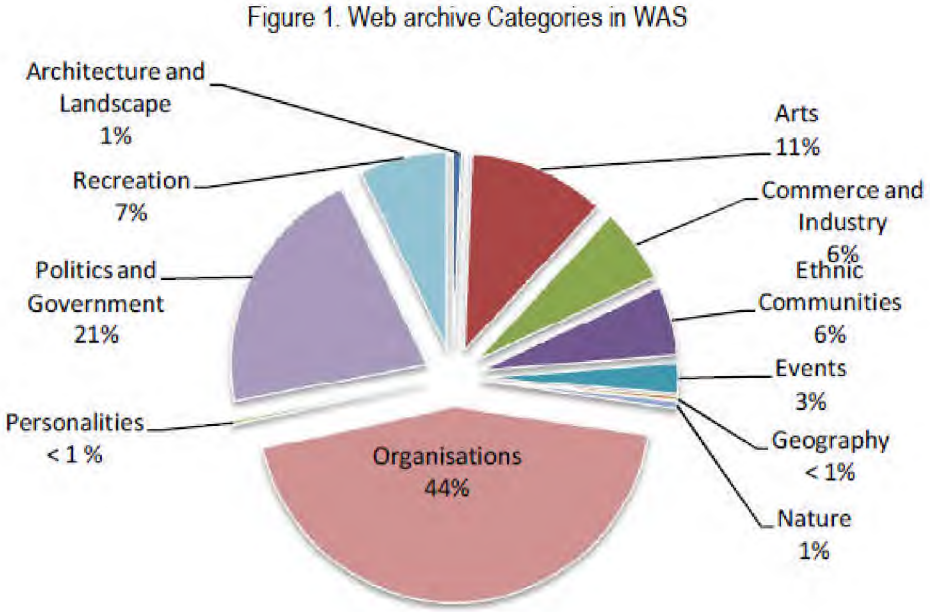
Las obras huérfanas son aquellas para las cuales no puede identificarse o ubicarse al tenedor del derecho de autor (Copyright Office USA, 2005). Esta situación es problemática porque crea una barrera legal para el uso de estas obras. La ley de derechos de autor establece que el uso de obras protegidas requiere la autorización del titular del derecho de autor, y el uso de derechos de autor sin autorización sólo se permite en casos excepcionales.

Actualmente no existe ninguna solución para este problema de las obras huérfanas en Singapur. La relación entre las obras huérfanas y los archivos web es que el contenido no puede ser utilizado en el futuro si no existe una identificación clara del autor y el derecho de autor.

5 TAXONOMÍA DEL ARCHIVO WEB

El Archivo Web de Singapur clasificó 1174 sitios web en 11 categorías amplias. La Figura 2 muestra su distribución. Las organizaciones forman el grupo más grande con el 44% (514), seguido por Políticas y Gobierno (21%) y Artes (11%). Las categorías restantes tienen menos del 10% en cada grupo. De estas categorías, 4 (Arquitectura, Geografía, Naturaleza, Personalidades) tienen 1% o <1% del total de sitios web archivados. La categoría "Personalidades" tiene sólo 4 sitios web y la mayoría de los vínculos web de singapurenses conocidos (Lee Kuan Yew, Lee Hsien Long, etc.) y sus sitios web son ignorados; se debería recopilar un amplio espectro de sitios web de personalidades de Singapur para que la cultura singapurense esté lo más representada posible. No existe ningún criterio que explique por qué el contenido archivado específico de los sitios posee características relevantes para el patrimonio cultural.

Figura 1. Categorías de Archivos Web en el WAS



La categoría "Artes" fue elegida para ser analizada con el propósito de identificar los errores comunes encontrados en los sitios web archivados. Debido a que el sitio incluía un comentario que decía "Algunos vínculos, formularios y cuadros de búsqueda pueden no funcionar adecuadamente", sólo se consideró como contenido erróneo al contenido archivado que no incluía ninguna información. Los errores tales como sitios web con problemas de compatibilidad de buscador y sitios web con sólo un vínculo a una página de inicio fueron agrupados en "Otros".

La Figura 2 muestra la distribución de errores de la categoría Artes. El 39% de los sitios web archivados en esta categoría tiene algunos errores en sus dos últimas versiones. Esta evaluación es indicativa de la baja calidad del contenido de los archivos del WAS y hace hincapié en la necesidad de mejoras técnicas en el proceso de archivo.

Tabla 1. Lista de errores encontrados al analizar la categoría ARTES

No.	Defectos/Errores

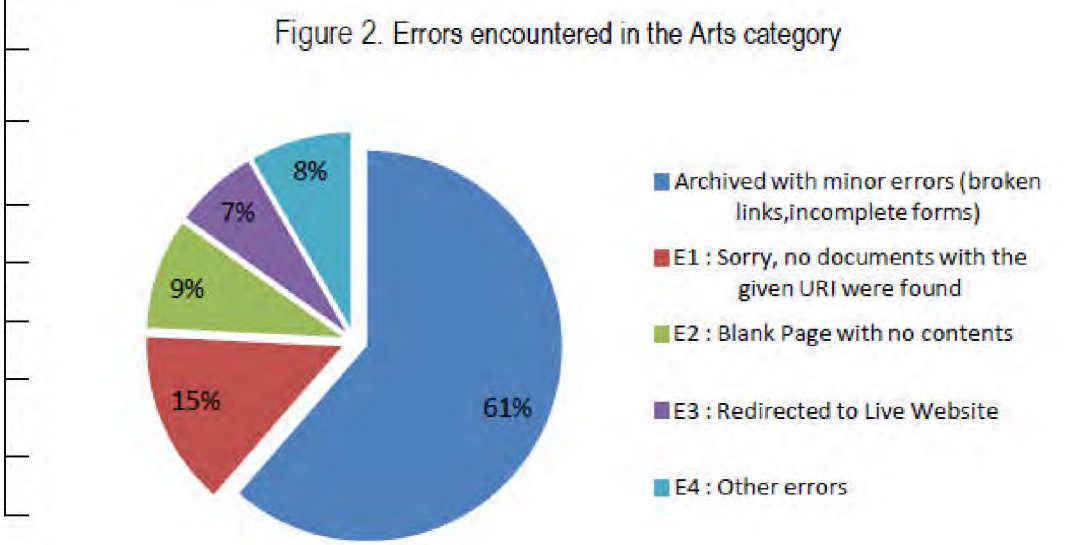
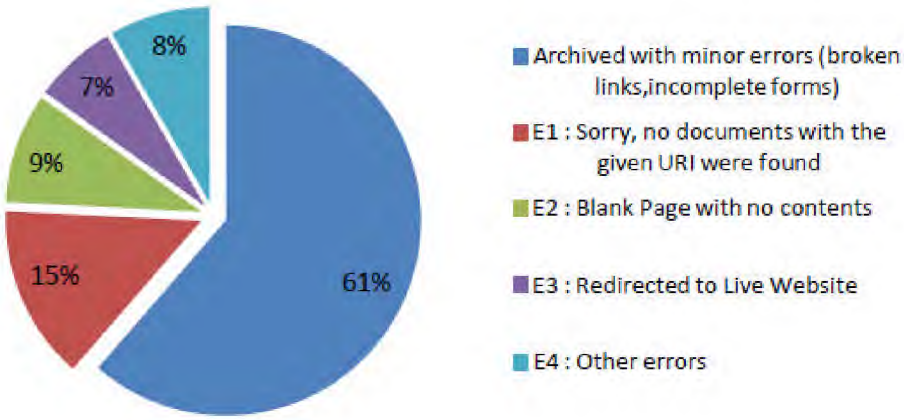


Figura 2. Errores encontrados en la categoría Artes

Figure 2. Errors encountered in the Arts category



5.1 Búsqueda y API's

Las expectativas y el interés de los usuarios experimentados de contenidos de archivos difieren entre si. Los académicos e investigadores podrían requerir documentos web para sus investigaciones y referirse a estos documentos en su trabajo. Para usuarios genéricos, que usan Internet como medio de comunicación, el archivo podría no ser muy valioso. Actualmente la mayoría del contenido archivado se preserva con fines de investigación.

En todo caso, los usuarios pueden esperar búsquedas similares a las que están disponibles en la web en vivo. Existe una necesidad urgente de desarrollar mejores maneras de acceder al contenido archivado. La mayoría de las interfaces de archivo web se restringen actualmente a una búsqueda basada en un URL con una función de búsqueda básica que utiliza categorías de URL. Las herramientas Wayback Machine, WERA (Acceso al Archivo Web) y Hanzo-WARC sólo cuentan con funciones básicas de búsqueda de contenido (Costa.M, Silva.J.M, 2009). En el desarrollo actual de los archivos web, sólo es posible indexar el texto y no los archivos de audio/video o los archivos de imagen. Esta indexación de textos se realiza a partir de los metadatos generados por una araña web presentes en los encabezamientos de archivos. Algunas interfaces de archivo web utilizan Google para las funciones de indexación y búsqueda. El uso de buscadores permite la indexación automática, lo que permite administrar y buscar miles de millones de documentos. Sin embargo, se requieren herramientas de búsqueda de archivos web como NutchWax[5], una extensión del buscador web Nutch para la búsqueda de archivos WARC/ARC, para satisfacer las necesidades específicas de búsqueda de contenidos de archivos web. Si bien los buscadores de gran escala (Google, Yahoo) cuentan con búsquedas de imagen y video, todavía no logran ofrecer resultados precisos en el caso de colecciones grandes (Li, & Wang, 2008).

La catalogación es un factor importante que tiene un impacto en la función de búsqueda en archivos web. Dado que las bibliotecas realizan archivos web de gran escala, tienden a implementar los procedimientos de catalogación tradicionales en la catalogación de documentos web, lo que limita la efectividad de las funciones de búsqueda. La catalogación tradicional no es una opción adecuada para todas las prácticas de archivo sino sólo para determinados archivos, que generalmente son de pequeña escala (Hallgrimsson, 2006). Esto se debe al hecho de que solo un pequeño porcentaje de documentos web contiene metadatos descriptivos confiables. De hecho, algunos estudios demuestran que la catalogación no es una opción viable para ofrecer acceso al contenido archivado; los informes finales del Proyecto Minerva revelan que "La Biblioteca debe basarse en la indexación automática como principal medio de descubrimiento de información de sitios web y del contenido dentro de los sitios. El texto completo de todo el material escrito debería ser indexado periódicamente y susceptible de ser buscado por los usuarios," y "La Biblioteca no debería invertir en extender la catalogación MARC u otros métodos manuales de catalogación a los sitios web, exceptuando los sitios de particular importancia para la Biblioteca y sus usuarios" (Masans, 2006).

5.2 Estudio sobre la evolución de la terminología

Full name (LiWA) es un proyecto financiado por la Comisión Europea que se centra en extender el estado actual de las prácticas de archivo web concentrándose en métodos innovadores de captura de contenido, mejorando la coherencia abordando el tema de la evolución semántica. Uno de los objetivos del proyecto LiWA es encontrar la evolución de la terminología sobre archivos web, lo que ocurre en la preservación a largo plazo. A medida que la sociedad cambia, el uso del lenguaje también se va modificando y esto debe ser abordado en el archivo web (Risse T, 2011). Por ejemplo, en una búsqueda de reproductores de música, el usuario generalmente escribirá la palabra "ipod", pero anteriormente hubiera escrito "discman" y "walkman". Un investigador que desee realizar una investigación sobre un tema específico debe tener conocimiento de todos los términos que se utilizaban anteriormente para encontrar la información deseada. Lo mismo ocurre con los nombres de lugares y los nombres de cargos (presidente, director, etc.) ya que tanto la persona que ocupa el cargo como su título cambia con el tiempo. La evolución de la terminología puede tener un impacto en la recuperación y el uso del archivo web en futuras investigaciones. Para mantener un archivo web semánticamente accesible, LiWA ha propuesto como solución agrupar las palabras comunes en un archivo. Se deberán considerar dos nuevas capas, una capa arquitectónica y una capa lingüística y semántica que vincularía los términos utilizados con su significado previsto. Este tipo de estudio es importante para el archivo de medios sociales ya que evolucionan muy rápidamente,

[5] http://nutch.apache.org/

y los usuarios de medios sociales modifican las frases nominales y crean palabras nuevas con mucha frecuencia.

La NLS podría promover el desarrollo de interfaces de servicios web basados en SOAP o u otros servicios web comerciales que aborden los cambios de la web en vivo. Además, la NLS debe promover el desarrollo de herramientas visuales para la búsqueda de metadatos.

5.3 Análisis Web

Más allá de los temas técnicos relacionados con el archivo, las instituciones y organizaciones que participan de las prácticas de archivo web deben responder las siguientes cuestiones con el fin de enriquecer la experiencia de los usuarios en materia de contenidos de archivos.

- ¿Cómo hacen los usuarios para buscar un contenido archivado importante (palabras clave utilizadas, rutas de navegación del usuario, etc.)?
- ¿Cómo puede filtrarse un contenido archivado de poca precisión y no indexado?
- ¿A qué tipo de contenido archivado desean tener acceso los usuarios?
- ¿Cómo puede ser utilizado más eficazmente un contenido archivado en términos de búsqueda por índice?
- ¿Cómo pueden descubrirse los esquemas de utilización a través del contenido web archivado?

Otras cuestiones similares podrían ser respondidas con técnicas de minería web implementadas en la Web en vivo. La minería web es el proceso de utilizar técnicas de minería de datos con el fin de extraer conocimientos de los datos archivados tales como archivos de registro que contienen versiones, velocidades de conexión de redes, servidores web, sistemas operativos e hipervínculos (Cooley, Srivastava y Mobasher, 1997). Esta visión centrada en datos de la minería web resulta más aceptable entre los investigadores debido a su enfoque y se adapta mejor a las necesidades de la minería de archivos web.

La mayoría de las instituciones que implementan iniciativas de archivos web no demuestra interés en preservar los archivos de registro de uso web y se centraliza únicamente en la recopilación web (web harvesting) y en los procesos externos de archivo. Pero un punto esencial es la preservación de los archivos de registro que puedan ser utilizados para la gestión de registros internos y la minería web.

Los procesos de minería web podrían ayudar a los usuarios a obtener un contenido relevante durante una búsqueda. Abarcan diferentes técnicas, como la minería de contenidos web, la minería de uso web y la minería de estructura web. Además, las técnicas de minería web podrían ofrecer una mejor experiencia de usuario al proporcionar una clasificación basada en la calidad y un contenido de archivo web relacionado con la imagen.

La minería web podría obtener información en base a los datos recolectados durante el ciclo de vida del archivo web. Los datos de la minería web pueden ser categorizados en tres tipos de acuerdo con su origen:

1. Metadatos: Los metadatos pueden clasificarse a su vez en dos tipos: metadatos objeto y metadatos técnicos. Los metadatos objeto son datos extraídos del objeto web, como el tamaño del objeto y fecha de creación mientras que los metadatos técnicos tienen información sobre el tipo de servidor web y sistema operativo utilizado. El examen de estos datos contribuiría a mejorar la autenticidad y procedencia del contenido archivado.
2. Datos de uso: una fuente importante de los datos de uso es la de los archivos del servidor web. Si un usuario envía una solicitud a un servidor web sobre el acceso a un contenido archivado, el mismo podrá almacenar información como: fecha de la solicitud, número de bytes transmitidos a ese usuario y lugar de origen, la dirección IP del usuario, etc. Esta información ayuda al proceso de minería web a analizar las clasificaciones del contenido archivado, clasificaciones de tráfico, etc.
3. Datos de infraestructura: Contienen la información principal del sistema de infraestructura de archivo como las direcciones IP, los números de sistemas autónomos, cantidad de ciclos CPU conectados con routers, ancho de banda utilizada por routers, etc. A través de un proceso de minería de datos, resulta posible analizar y medir las actualizaciones

del tráfico de routers, el tráfico de Internet, etc. y así calcular el tiempo que tardan los paquetes de datos archivados para alcanzar el host de destino.

En una visión holística, todo este conocimiento podría mejorar el rendimiento del proceso de archivo web y ofrecer una mejor categorización de contenidos y una futura predicción de esquemas de uso de datos archivados.

6. DESAFIOS EN MATERIA DE GESTION DE INFRAESTRUCTURA

6.1 Computación de Nube y Computación Grid

Una infraestructura de archivos web debería ser suficientemente sólida como para almacenar grandes cantidades de datos y también debe ser escalable para entregar estos datos a los usuarios durante su extracción. Los sistemas de archivo web apuntan a mantener los datos para una preservación a largo plazo y deben estar en condiciones de retener el contenido archivado y otros datos relacionados en caso de falla o apagado del sistema.

Brewster Kahle (2002), el fundador del Archivo de Internet, ha definido algunos requisitos y operaciones básicos para sistemas de archivos que siguen teniendo importancia:

- El sistema debería utilizar únicamente equipos básicos.
- El sistema no debería basarse en el software comercial.
- El sistema no debería exigir un título de posgrado para la implementación o mantenimiento.
- El sistema debe ser lo más simple posible.

El proceso de infraestructura de archivos web abarca las etapas de almacenamiento, importación, búsqueda, indexación y acceso. La infraestructura debe ser compatible con los requisitos de este proceso y también debe proporcionar seguridad de datos y soluciones de gestión de recursos. Además, debe gestionar los recursos T I y ofrecer soluciones de bajo costo en términos de software, hardware y otros requisitos operativos. Las principales áreas de interés para archivar el contenido de Internet son: interoperabilidad, escalabilidad y seguridad (Jaffe E.,& Kirkpatrick, S, 2009).

Para satisfacer todos estos requisitos sobre archivos web las instituciones emplean diferentes tipos de sistemas de computación en función de sus presupuestos. En la actualidad, los sistemas de computación distribuida o "cluster" y los sistemas de computación grid son ampliamente utilizados por las instituciones de archivo.

La computación grid proporciona un marco para la identificación de servers ociosos y otros activos TI , así como los requisitos del procesador para el proceso de archivo. A su vez, esto incrementa la eficiencia en el uso de los recursos y permite una utilización de recursos equilibrada. Durante un pico impredecible de trabajo, una tarea en particular puede ser enviada a un procesador ocioso del sistema grid y si el recurso grid es utilizado en toda su capacidad, va a suspender o a cancelar la tarea de menor prioridad y determinar la forma de archivar tareas de mayor importancia, como la de capturar la página índice de un sitio web. El tema de la tarea de menor prioridad será desarrollado más adelante de acuerdo con las especificaciones sobre recursos del sistema grid. También calcula los espacios de los discos de almacenamiento no utilizados que son configurados con la arquitectura grid (denominados específicamente grid de datos) y crea un mayor almacenamiento de datos virtuales capaz de archivar contenidos en forma más eficiente a través de una máquina de control remoto como un recurso de almacenamiento compartido.

La computación grid ha sido fundamental para el Proyecto WAS; el mismo facilita la distribución del espacio de almacenamiento y otros recursos, como algunos dispositivos especializados, software, servicios, licencias, etc. Por ejemplo, en un proceso de archivo, si un usuario necesita incrementar la velocidad de Internet para capturar un snapshot de un sitio web multipágina, el proceso puede ser dividido virtualmente en la arquitectura grid. La naturaleza del sistema de computación grid permite el procesamiento paralelo, una ventaja para las iniciativas de archivo web en términos de perfeccionamiento de la infraestructura TI

con gestión de recursos, confiabilidad y administración. También reduce costo y tiempo, necesita menos mantenimiento y mano de obra.

Un sistema de computación grid es compatible con el archivo de Internet pero –en alguna medida- enfrenta dificultades en ejecutar la programación multinivel (Foster I, Zhao, Y, Shiyong Lu, 2008) Por ejemplo, si un trabajo está listo para ser archivado y necesita la capacidad de cien procesadores con un tiempo disponible de 60 minutos cada uno, el trabajo debería esperar hasta que el administrador de recursos (LRM) esté en condiciones de asignar 100 procesadores con la disponibilidad de tiempo estipulada. En el caso del modelo de computación cloud, el caso es totalmente diferente. La tarea particular puede ser compartida por todos los usuarios al mismo tiempo y procesada por el sistema de espera programada permitiendo las aplicaciones para operar de manera nativa en los sistemas cloud de almacenamiento centralizado. La mayor dificultad de la computación grid es el fracaso del sistema; si un procesador del sistema grid falla, toda la red distribuida se verá afectada y la preparación y programación de los datos se detendrán en un determinado momento.

6.2 Posición de preeminencia de la Computación Cloud

El sistema de computación cloud podría ser una solución para las dificultades con las que la computación grid pudiera encontrarse en el futuro. Si bien la computación grid es el segmento principal de la computación cloud, la misma está diseñada para actuar como una infraestructura completa (infraestructura como un servicio-laaS) y es automáticamente intercambiable de manera tal que el sistema no se detiene cuando una red ha dejado de funcionar. Los sistemas cloud proporcionan la mayor seguridad y compatibilidad de datos en términos de almacenamiento. En la computación grid, una tarea importante se divide en varias subtareas y es ejecutada en múltiples procesadores en base al tiempo y los recursos disponibles, mientras que en la arquitectura de la computación cloud, los usuarios pueden recoger sitios web sin involucrar los recursos y ejecutar un rastreo multiprogramado. El mismo puede acceder a servers compartidos de acuerdo con los recursos, el software y los datos.

Para acceder a un servicio de computación cloud, no es necesario contar con software, plataformas de aplicación o una infraestructura individual. Se reducen significativamente los gastos por adelantado, los gastos operativos y las inversiones de capital. Una operación de computación cloud evita el mantenimiento de servers y redes y dado que permite el acceso a múltiples servers desde cualquier lugar a través de un emulador común, representa un método "ecofriendly" de computación. Lo más importante es que en la computación cloud los datos pueden ser restaurados. Cuenta con tecnología de recuperación ante desastres y backup de datos que ayudarían a conservar los datos archivados ante una falla del sistema como, por ejemplo, un "crash" del server o de los datos.

Un ejemplo de computación cloud en Web Archival es la Iniciativa de Archivos de Documentos Gubernamentales de Latinoamérica (LAGDA), que ha implementado un sistema de computación cloud para el archivo de documentos web ministeriales y presidenciales de dieciocho países latinoamericanos y de la región del Caribe (LAWAP[6], 2005). Es un proyecto conjunto con la *University of Texas Libraries*, y el Centro de Computación Cloud Avanzada de Texas da asistencia a los servicios de computación cloud. El archivo web de LAGDA es de aproximadamente 6 TB en la actualidad (Owens. T, 2012).

7 RECOMENDACIONES

Nuestras recomendaciones se dividen en tres categorías:

<u>Etico-políticas:</u>

a) La Biblioteca Nacional debe adoptar un marco transparente y abierto sobre archivos web, difundiendo sus políticas y compartiendo el conocimiento sobre la experiencia WAS.

[6] El objetivo del Proyecto de Archivos Web de América Latina (LAWAP) es la recopilación y preservación de los recursos web de América Latina. El Archivo de Internet ofrece sus servicios a LAWAP.

b) La NLB debe establecer un marco ético o una comisión de ética en temas referidos a WAS.

c) La NLB –en consonancia con el objetivo de apoyar una economía del conocimiento en Singapur- debe ser reformulada como una biblioteca de investigación para proporcionar apoyo a la investigación en humanidades digitales, donde WAS pueda cumplir un rol importante como experimentada fuente de investigación.

Derechos de Autor:
La legislación sobre derechos de autor en materia de archivos web requiere la introducción de enmiendas que proporcionen:

a) seguridad jurídica para Recopilación Web, Preservación Digital de Largo Plazo, Acceso de Largo Plazo y Utilización y reutilización de la información archivada con fines de investigación y estudio; y

b) mecanismos que faciliten la autorización de derechos de autor para la reutilización y transformación de las obras protegidas por los derechos de autor.

Técnicos:

a) Para un mejor desarrollo de las colecciones, WAS debe desarrollar herramientas para mejorar la utilización del acceso de metadatos en sitios web únicos (revisión de herramientas de acceso al mapa del sitio).

b) Implementación de herramientas (ej. Taverna) o desarrollo de nuevas herramientas para un mejor proceso de flujo de trabajo.

c) WAS debería iniciar el proceso de minería web en sus archivos preservados con el fin de mejorar la calidad de la recopilación web.

d) WAS debería apuntar a la creación de nuevas herramientas que permitan mejorar la visualización de los metadatos existentes.

e) WAS debería analizar la evolución de la tecnología y procurar la inclusión de un elemento lingüístico en su arquitectura de archivos web.

f) En el futuro, para el archivo selectivo de determinados hechos, WAS podría intentar con la utilización de la infraestructura cloud.

Bibliografía

Beasley, S., & Kail, C. (2009). a2o: *Access to Archives from the National Archives of Singapore* (Acceso a los Archivos provenientes de los Archivos Nacionales de Singapur). Journal of Web Librarianship, 3(2), 149–155. doi:10.1080/19322900902896531

Chellapandi, S., Han, C. W., & Boon, T. C. (2010). *The National Library of Singapore experience: harnessing technology to deliver content and broaden access* (La experiencia de la Biblioteca Nacional de Singapur: uso de tecnología para entregar contenidos y ampliar el acceso). Interpréstamos y Suministro de Documentos, 38(1), 40-48.

Chellapandi, S., Han, C. W., & Lian San, S. (noviembre, 2009). *Web Archiving Programme at National Library Singapore.* (Programa de Archivo en la web de la Biblioteca Nacional de Singapur) CDNLAO Newsletter. Tokio, Japón. Consultado en http: //www. ndl. go. jp/en/cdnlao/newsletter/066/664. HTML.

Chen, H., Chung, W., Qin, J., Reid, E., Sageman, M., & Weimann, G. (2008). *Uncovering the dark Web: A case study of Jihad on the Web* (Descubriendo la *dark web*: Un estudio del Yihad en la web. Publicación de la Sociedad Americana para la Ciencia de la Información y la Tecnología, 59(8), 1347-1359. doi: 10. 1002/asi. 20838.

Cloud Computing vs. Grid Computing (Computación en la nube c/ Computación grid)« Guía CCSK. (n.d.). Guía CCSK. Consultado el 5 de abril de 2013, en http: //ccskguide.org/cloud-computing-vs-grid-computing/.

Oficina de Derechos de Autor de EEUU. (2005). *Report on Orphan Works* (Informe sobre Obras Huérfanas).

Datta, R., Joshi, D., Li, J., & Wang, J. Z. (2008). *Image Retrieval: Ideas, influences, and trends of the new age* (Recuperación de Imágenes: Ideas, influencias, y tendencias de la nueva era). ACM Computing Surveys, 40(2), 1-60

Diferencia entre la computación en la nube y la computación grid. (n. d.). Cloudways | Managed Cloud Hosting & App Deployment Solutions. Consultado el 3 de abril de 2013, en http: //www. cloudways. com/blog/cloud-computing-vs-grid-computing-differentiated/

Dougherty, M., Meyer, E., Madsen, C., Van den Heuvel, C., Thomas, A., & Wyatt, S. (2010). *Researcher engagement with web archives: State of the art* (Compromiso de los investigadores con los archivos web: tecnología de última generación). Informe de la Comisión Conjunta de Sistemas de Información. Consultado en http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1714997

Foo, S., Wu, H. J. ., Lim, S. K., Haliza, J., Noryati, A. S., Mohamed, S., ... Tim, Y. F. (2005). *Legal Deposit Development in Singapore: Future Challenges and Issues* (Desarrollo del Depósito Legal en Singapur: Asuntos y Desafíos Futuros). En la Conferencia Internacional sobre Bibliotecas Nacionales en la Sociedad basada en el Conocimiento. Bangkok.

Glanville, L. (2010). *Web archiving: ethical and legal issues affecting programmes in Australia and the Netherlands* (Archivos en la web: asuntos éticos y legales que afectan a los programas en Australia y los Países Bajos). *THE AUSTRALIAN LIBRARY JOURNAL*, (Publicación de la Biblioteca Australiana) 59(3), 128.

Hayles, K. (2012). *How we think: digital media and contemporary technogenesis* (Cómo pensamos: los medios digitales y la tecnogénesis contemporánea). Chicago; Londres: *The University of Chicago Press*.

Jaffe. E., & Kirkpatrick. S. (2009). *Architecture of the Internet Archive* (Arquitectura del archivo de internet), *the Hebrew University of Jerusalem* (Universidad Hebrea de Jerusalén), Israel.

LiWA. (2013). *Living Web archive web page* (Archivo de página web "*Living Web*"). Página web consultada de 3 de abril de 2013, en http://liwa-project.eu/

Lor, P. J., & Britz, J. J. (2012). *An ethical perspective on political-economic issues in the long-term preservation of digital heritage* (Una perspectiva ética sobre temas político-económicos en la preservación a largo plazo del patrimonio digital). Publicación de la Sociedad Americana de Ciencia y Tecnología de la Información, 63(11), 2153–2164. doi:10.1002/asi.22725

M. Costa y M. J. Silva. *Towards information retrieval evaluation over web archives* (Evaluación de la recuperación de información de los archivos web). En S. Geva, J. Kamps, C. Peters, T. S. A. Trotman, y E. Voorhees, editores, *Proceedings of the SIGIR 2009 Workshop on the Future of IR Evaluation* [Procedimientos del Taller 2009 del SIGIR, por sus siglas en inglés, (Grupo de Interés Especial sobre Recuperación de Información) sobre el Futuro de la Evaluación de la Recuperación de Información], Boston, Massachusetts, julio de 2009. *IR Publications*, Ámsterdam.

Masanš , J. (2006). *Access and Finding Aids* (Asistencia para el Acceso y la Búsqueda).*Web Archiving* (pp. 131-151). New York: SpringerVerlag Berlin Heidelberg.

Mason, I. (2007). *Virtual Preservation: How Has Digital Culture Influenced Our Ideas About Permanence?* (Preservación Virtual: ¿Cómo ha influido la cultura digital en

nuestros conceptos de permanencia?) *Changing Practice in a National Legal Deposit Library* (Cambio de Prácticas en una Biblioteca Nacional de Depósito Legal). *Library Trends*, 56(1), 198–215.

Programa Nacional de Infraestructura y Preservación de la Información Digital (EEUU), Comisión Conjunta de Sistemas de Información, y el Proyecto de Ley de Acceso Abierto al Conocimiento (OAK, por sus siglas en inglés). (2008). *International study on the impact of copyright law on digital preservation* (Estudio internacional sobre el impacto de la ley de derechos de autor en la preservación digital). Washington, D.C.: Biblioteca del Congreso. Consultado en http://www.digitalpreservation.gov/partners/resources/pubs/wipódigital_preservation_final_report2008.pdf

Comisión de Bibliotecas Nacionales. (2005). *Library 2010: libraries for Life, Knowledge for Sucess* (Biblioteca 2010: bibliotecas para la Vida, Conocimiento para el Éxito). Comisión de Bibliotecas Nacionales.

Comisión de Bibliotecas Nacionales. (2012). Informe Nacional de Singapur para la Reunión 2012 de CDNL-AO. NLB Singapur. (2010). Informe Nacional de Singapur. En la 15° Junta Ejecutiva del CONSAL (Congreso de Bibliotecarios del Sudeste Asiático). Bandung, West Java, Indonesia.

Ortmann, S. (2009). Singapur: *The Politics of Inventing National Identity* (La Política de Invención de la Identidad Nacional). *Journal of Current Southeast Asian Affairs* (Publicación del Sudeste Asiático), 28(4), 23–46.

Owens, T. (6 de junio de 2012). *Web Archiving and Mainstreaming Special Collections: The Case of the Latin American Government Documents Archive* (Integración y Archivo en la Web de las Colecciones Especiales: El Caso del Archivo de los Documentos de los Gobiernos Latinoamericanos) *The Signal: Digital Preservation.* Blogs de la Biblioteca del Congreso. Consultado el 23 de marzo de 2013, en http://blogs.loc.gov/digitalpreservation/2012/06/web-archiving-and-mainstreaming-special-collections-the-case-of-the-latin-american-government-documents-archive/

Patel, K. (2007). *Authors v. Internet Archives: The Copyright Infringement Battle over WEB Pages* (Autores c/ Archivos de internet: La Batalla sobre la Violación de los Derechos de Autor en las páginas web). J. Pat. & Trademark Off. Soc'y, 89, 410.

Paynter, G. W., & Mason, I. B. (2006). *Building a Web Curator Tool for The National Library of New Zealand* (Creación de una herramienta de conservación de la web para la Biblioteca Nacional de Nueva Zelanda). En la Conferencia de la Asociación de Bibliotecas de Nueva Zelanda.

Portrait Workshop (Taller de Retratos). (26 de septiembre de 2007). *Notification of Website Archiving - Portrait Workshop* (Notificación de Archivos web – Taller de Retratos). Caricaturas Singapur. Consultado en http://www.caricature.com.sg/2007/09/notification-of-website-archiving.html

Potter, A (2012) *A Vision of the Role and Future of Web Archives: Research Use* (Una Visión sobre el Rol y el Futuro de los Archivos Web: Uso en Investigaciones) *The Signal: Digital Preservation.* (n. d.). Blogs de la Biblioteca del Congreso. Consultado el 20 de marzo de 2013, en http://blogs.loc.gov/digitalpreservation/2012/05/a-vision-of-the-role-and-future-of-webarchives-research-use/

Seet, K. K. (2005). *Singapore's transformative library: knowledge, imagination, possibility* (Biblioteca transformadora de Singapur: conocimiento, imaginación, posibilidad). Singapur: Ediciones SNP.

Seng, D. (2008). *International Study on the Impact of Copyright Law on Digital Preservation. Singapore's Legal Position* (Estudio Internacional sobre el Impacto de la Ley de Derechos de Autor sobre la Preservación Digital. La Posición Legal de Singapur). Presentado

en el Taller Internacional de la OMPI sobre Preservación Digital y Derechos de Autor, Ginebra, Suiza.

SG Wiki Editor. (4 de julio de 2007). *Discussion about Pacific Centennial Group Wikipedia Page* (Discusión sobre la página de Wikipedia respecto del *Pacific Centennial Group*) Consultado en http://www.zaped.info/Wikipedia:Articles_for_deletion/Log/2007_July_2

Spaniol, M., Mazeika, A., Denev, D., & Weikum, G. (2009). *'Catch me if you can': Visual Analysis of Coherence Defects in Web Archiving* ('Atrápame si puedes': Análisis Visual de los Errores de Coherencia en los Archivos web). En los Procedimientos del 9º Taller Internacional de Archivo en la Web (IWAW 2009, por sus siglas en inglés), Corfú, Grecia. Consultado en http://www.iwaw.net/09/IWAW2009.pdf

Stirling, P., Illien, G., Sanz, P., & Sepetjan, S. (2012). *The state of e-legal deposit in France: Looking back at five years of putting new legislation into practice and envisioning the future* (El estado de los depósitos legales electrónicos en Francia: Teniendo en cuenta los 5 años de vigencia de la nueva legislación y pensando en el futuro). Publicación de IFLA, 38(1), 5–24 doi:10.1177/0340035211435323

Sundaram, K. (n.d.). *Cloud Computing vs. Grid Computing - What is the Difference?* (Computación en la nube c/ Computación Grid - ¿Cuál es la diferencia?) Artículos de Ciencia & Tecnología, Planes Educativos, Consejos Tecnológicos, Revisiones de Hardware y Software, Noticias y Más en *Bright Hub*. Consultado el 29 de marzo de 2013, en http://www.brighthub.com/environment/green-computing/articles/68785.aspx

WAS, NLB. (2013a). Preguntas frecuentes. Consultado en http://was.nl.sg/faq.html

WAS, NLB. (2013b). Política de Remoción. Consultado en http://www.microsite.nl.sg/TakedownPolicy.html

Wu, P. H. J., & Heok, A. K. H. (2005). *Documenting Online Collaboration Between Researchers and Information Professionals: The Case of the Asian Tsunami Web Archive* (Documentación Online de la Colaboración entre Investigadores y Profesionales de la Información: el Caso del Archivo web sobre el Tsunami de Asia). *Singapore Journal of Library & Information Management* (Publicación sobre Administración de Bibliotecas e Información de Singapur), 34, 75–85.

Yong Fu, J. L. (2013). *One Versus The Mob: A Case for the Wisdom of Crowds in Archival Appraisal 2. 0* ("Uno contra La Multitud": Un Caso sobre la Sabiduría del Público en la Valoración de los Archivos 2.0). Tésis. Universidad Tecnológica Nanyang, Singapur.