

**NOTA TÉCNICA No.2  
DESARROLLO DE COLECCIONES DIGITALES  
Elaborado por David Ramírez**

Introducción .....	2
1. Legislación .....	3
1.1. Actualización de la ley de depósito legal digital.....	3
1.2. El depósito voluntario de obras digitales .....	3
1.3. Actualización de otras leyes relacionadas al depósito de obras digitales .....	4
2. Depositante .....	5
3. Obras susceptibles de depósito .....	5
3.1 Depósito de obras en línea .....	6
3.1.1 Mecanismos de adquisición de sitios web .....	6
3.1.2 Lineamientos para la captura de sitios web .....	7
3.2 Mecanismos de adquisición de obras fuera de línea .....	8
3.2.1 Lineamientos para el depósito de obras fuera de línea.....	8
3.2.2 Criterios de aprobación del depósito de obras digitales.....	11
4. Política de Acceso.....	12
4.1 Definiciones .....	12
4.1.1 Archivo blanco .....	12
4.1.2 Archivo gris .....	12
4.1.3 Archivo oscuro .....	13
5. Comité de Selección y Adquisiciones.....	14
6. Visualización y consulta del material recopilado en depósito .....	14
7. Anexos .....	15
7.1 Glosario .....	15
8. Bibliografía .....	18

## **Introducción**

Este documento presenta la política para el depósito de obras digitales<sup>1</sup> que tiene como fin permitir, tanto a la Biblioteca Nacional como a los depositantes, contar con lineamientos para llevar a cabo un depósito de la producción digital colombiana. Esta política será un documento en permanente actualización atendiendo a las fases de implementación de este depósito, entre ellas recepción por tipo de material, tipos de archivos, herramientas tecnológicas de gestión, desarrollo normativo, acuerdos con sectores interesados, entre otros.

La política comprende:

- Marco legal
- Obras susceptibles de depósito en línea y fuera de línea
- Política de acceso
- Comité de Selección y Adquisiciones
- Visualización y consulta de lo depositado.

---

<sup>1</sup> Se usará el término “depósito de obras digitales” para referirse al depósito voluntario mientras no se cuenta con una legislación que permita hablar de depósito legal digital.

## **1. Legislación**

El depósito de obras digitales debe regirse por las leyes colombianas. Los procesos de la Biblioteca se adaptarán de acuerdo a los cambios que se produzcan en la legislación y se tendrá en cuenta el derecho de autor, pero de igual manera el derecho de los usuarios de la información, según las recomendaciones de IFLA (IFLA, 2011a).

Para que la Biblioteca pueda encontrar puntos de equilibrio entre el acceso a sus usuarios y las restricciones del derecho de autor, la Biblioteca Nacional deberá estar al tanto de todas las modificaciones a la legislación sobre depósito legal o derecho de autor. Deberá informarse sobre las propuestas de ley y realizar sugerencias para que se realice legislación que proteja a los autores, pero también a los usuarios.

### **1.1. Actualización de la ley de depósito legal digital**

Mientras no se cuente con una ley actualizada, la Biblioteca deberá realizar todos los trámites requeridos para no infringir el derecho de los autores o titulares. Se deberá seguir entonces líneas de trabajo como las de Estados Unidos o Gran Bretaña para recopilar recursos web (British Library, 2012), donde estos países piden autorización por cada sitio web para incluirlo en el depósito legal. En estos países, en caso que un autor o titular no responda o se niegue, el depósito no se realiza. La Biblioteca revisará periódicamente si la ley de depósito legal es pertinente o si requiere modificaciones, debido a que la gran variedad de implicados en este tema puede dejar por fuera opiniones en las primeras versiones de la legislación (Gömpel & Svensson, 2011). Como mínimo la Biblioteca cada año hará este seguimiento.

### **1.2. El depósito voluntario de obras digitales**

La Biblioteca trabajará en el depósito voluntario hasta que exista una ley que permita la recopilación de documentos o sitios web y se reglamente el depósito legal de documentos digitales. Dicho depósito está orientado por esta política como mecanismo que busca realizar un trabajo concertado con los productores de esta información con el fin de evitar por diferentes medios la pérdida de la memoria digital colombiana.

Como parte de la implementación del depósito de documentos digitales la Biblioteca deberá realizar copias de sitios web o recopilación de documentos a manera de prueba. Si el trámite finaliza con el permiso de los autores, la biblioteca podrá almacenar esta información según las políticas de acceso definidas en esta política.

### **1.3. Actualización de otras leyes relacionadas al depósito de obras digitales**

Los cambios en la legislación apuntan a una tendencia global, debido a que se requiere una actualización de las leyes que regulan el derecho de autor en entornos digitales. La Biblioteca, además de trabajar en la actualización de la ley referente al depósito legal, revisará los debates en temas relacionados con derecho de autor como excepciones y limitaciones para bibliotecas (IFLA, 2012a), obras huérfanas (IFLA, 2011b) préstamo de documentos electrónicos (IFLA, 2012b), copias para preservación y cualquier otro tema afín. Cada uno de estos temas afecta el desarrollo de las labores de la biblioteca y los flujos de trabajo. Por ejemplo el registro que se realiza en el depósito legal puede relacionarse con la identificación de obras huérfanas. Se trabajará en una actualización conjunta de la ley de derecho de autor y depósito legal que abarque todos estos temas para que la Biblioteca tenga más herramientas para proponer una agenda positiva sobre dominio público colombiano (International Communia Association, 2012).

Igualmente la Biblioteca participará en los foros y debates sobre la legislación colombiana en el ámbito de los derechos de autor y bibliotecas ya que estos afectan la política definida. Para ello deberá contar con un comité con profesionales en ciencias de la información, derecho y otras áreas del conocimiento, responsable de velar por el avance en este tema.

En este documento se plantea una definición de obras digitales susceptibles de depósito para que funcione como política interna de la Biblioteca. Entre los temas a actualizar y debatir, se debe incluir un capítulo para definir qué es un documento digital colombiano. Si bien en este documento se plantea una primera definición, se deben realizar debates para revisar los casos donde un sitio web se encuentre alojado en servidores en el extranjero, el caso de obras en colaboración o coautorías o cuando una obra sea escrita por extranjeros, sin domicilio en el país y que hable de temas de interés para la biblioteca. Inicialmente se propone un criterio de nacionalidad, que se revisará más adelante.

## 2. Depositante

El depósito podrá ser realizado por la persona u organización responsable de publicar, producir y poner a disposición del público un documento digital.

## 3. Obras susceptibles de depósito

Serán susceptibles de depósito las obras fuera de línea y en línea ya sean contenidos estáticos o dinámicos<sup>2</sup> que cumplen los criterios de formato, publicación, origen, temática o dominio, así:

- **Criterio de formato:** Es una obra digital. Ésta puede ser una obra nacida digital o el archivo digital fuente de una obra análoga. En este último caso la biblioteca recibirá los dos formatos, excepto cuando no existan copias físicas de la obra.(Stirling, Illien, Sanz, & Sepetjan, 2012).
- **Criterio de publicación:** Es una obra publicada. No se recomienda usar el término “obra terminada” debido a que puede sufrir modificaciones en el transcurso del tiempo. Será susceptible de depósito legal las diferentes versiones de una obra. Se entiende como obra publicada aquellas obras que son accesibles incluso a través de medidas de protección tecnológica como contraseñas.
- **Criterio de nacionalidad:** Es una obra colombiana. Se entienden por obras colombianas:
  - Obras de autores de nacionalidad colombiana.
  - Obras de autores, editores o productores cuyo domicilio principal está en Colombia.
- **Criterio de temática:** Son obras cuya temática principal sea Colombia, incluso si los autores, editores o productores no cumplen el criterio de nacionalidad o el dominio del sitio web o el domicilio del autor no es colombiano. Para el caso de sitios web alojados en otros países la Biblioteca adelantará las consultas jurídicas correspondientes.
- **Criterio de dominio:** Son las obras cuyo dominio web haga referencia a Colombia (.co) y cumpla el criterio de temática o nacionalidad.

Las obras deben cumplir obligatoriamente los criterios de formato y de publicación y al menos uno de los criterios de nacionalidad, temática o dominio.

---

<sup>2</sup> Ver glosario.

Si bien en este documento se plantea una primera definición de las obras susceptibles de depósito, es importante mantener un debate abierto sobre qué es un documento digital colombiano, donde se analicen casos como los relacionados con sitio web que se encuentran alojados en servidores en el extranjero, el caso de obras en colaboración o coautorías o cuando una obra sea escrita por extranjeros, sin domicilio en el país y que hable de temas de interés para la biblioteca. Los detalles jurídicos sobre este punto deben documentarse y presentarse en una posterior fase de ajuste a la política, debido a que exceden el alcance de este documento.

### **3.1 Depósito de obras en línea**

Debe tenerse en cuenta que es imposible que la Biblioteca recupere la totalidad de la web que es objeto de depósito según lo definido en esta política. Sin embargo la Biblioteca buscará obtener los recursos que considere pertinentes y prioritarios para nutrir la memoria colombiana sin ningún tipo de sesgo o creencia en particular.

#### **3.1.1 Mecanismos de adquisición de sitios web**

Para el depósito de obras en línea debe considerarse que el rol de la biblioteca cambia de “obligación de depositar a obligación de copiar” (Elliott, 2010). Esto quiere decir que la biblioteca no sólo recibe las obras de autores que quieren realizar el depósito, sino que debe buscar la información que le interesa a la biblioteca salvaguardar.

Por tal motivo se usan tres métodos para realizar el depósito:

- **Selección:** El comité de Selección y Adquisiciones de la Biblioteca Nacional, con el apoyo de expertos, selecciona los sitios web que deben depositarse.
- **Muestreo:** La biblioteca decide qué sitios se depositan, sin necesidad de pasar por el comité de selección.
- **Depósito voluntario por parte de los usuarios:** El usuario, de manera voluntaria hace el depósito de la web de su autoría.

En todos los casos la Biblioteca concertará con el productor, autor o titular de los derechos de las obras la posibilidad de consulta de estos recursos, bien sea total (archivo blanco) o parcial (archivo gris).

### **3.1.2 Lineamientos para la captura de sitios web**

Una vez identificado el sitio o fragmentos de la web a recuperar se debe decidir el tipo de captura a realizar. Dependiendo el tipo de recurso web se debe seguir un mecanismo diferente. Las variables a considerar son tamaño y frecuencia.

#### **3.1.2.1 Tamaño**

Dependiendo el tamaño del sitio web se debe decidir cuánto se recupera de él. Para sitios que superen un gigabyte de espacio de almacenamiento, se considerará la recuperación tan solo de la primera página u otros mecanismos de captura parcial de información. Esta restricción se debe a las limitaciones técnicas de almacenamiento y ancho de banda disponibles. Si es menor a un gigabyte se capturará toda la página.

Si hay una publicación más relevante que otra de acuerdo al formato, dentro de un sitio web obtenido por selección, se deberá optar por la captura de este tipo de documentos. Por ejemplo la captura de artículos publicados en PDF o únicamente los videos de un sitio web, teniendo en cuenta que la información más relevante de estos sitios se encuentra en estos formatos. La decisión del tipo de captura recae en el comité de selección y adquisiciones de la Biblioteca Nacional, que contará para este caso con invitados conocedores de este tema. Para los sitios obtenidos por muestreo el área de Selección y Adquisiciones podrá decidir, sin necesidad del comité de selección, si se capturan publicaciones, el sitio inicial o la totalidad del sitio.

#### **3.1.2.2 Frecuencia**

Para los sitios obtenidos por el comité, éste debe dar guía sobre la periodicidad con que se capturan los sitios. Los sitios con alta frecuencia de actualización deberán capturarse con mayor frecuencia, a menos que el comité de selección determine lo contrario. Se deben fijar fechas de captura de la totalidad de sitios web seleccionados, para tener hitos de captura web. Esto puede estar sujeto a restricciones técnicas.

Para el caso de sitios entregados por depósito voluntario, el comité decidirá su periodicidad de almacenamiento. Ésta podrá ser ajustada por la Biblioteca atendiendo a aspectos como la capacidad de almacenamiento y preservación. Los autores indicarán la frecuencia de actualización, que servirá como insumo para que el comité defina la periodicidad de recuperación, sin embargo las

indicaciones de los autores sobre la frecuencia de actualización no implican que ésta sea la que se realizará en el depósito de sus obras.

### **3.1.2.3 Sitios o fragmentos web de sucesos**

El comité de selección puede sugerir la captura de sitios web o fragmentos de un evento de especial interés para el país. Pueden ser elecciones, sucesos históricos, deportivos o desastres naturales. En este caso el comité de selección determinará el periodo de tiempo por el cual se hará la captura.

### **3.1.2.4 Sobre la implementación técnica de captura de sitios web**

La Biblioteca contratará con el servicio de terceros para la realización del harvesting mediante la suscripción al servicio de Archive-it, sin embargo la Biblioteca debe aprender a hacerlo con miras a desarrollar paulatinamente la infraestructura suficiente para realizarlo por sus propios medios y recibir la información de Archive-it. Con este fin la Biblioteca probará el uso de herramientas como Web Curator Tool, Heritrix, Way Back, Net Archive Suite, HTTrack, realizará un diagnóstico de las necesidades de la Biblioteca frente a los servicios de estas herramientas y hará pruebas de exportación e importación de la información recopilada a través del servicio de Archive-it. Depósito de obras fuera de línea

## **3.2 Mecanismos de adquisición de obras fuera de línea**

La Biblioteca obtendrá las obras fuera de línea a través de un formulario web, disponible en el sitio principal de la Biblioteca, o a través de la entrega del material en las instalaciones de la Biblioteca. Con el área de Sistemas se evaluará periódicamente la capacidad de envío de documentos a través del formulario, bien sea por tamaño o por mecanismos de envío de los archivos. El formulario debe ser parte del sistema integrado de la biblioteca (del inglés ILS) o estar articulado con éste.

### **3.2.1 Lineamientos para el depósito de obras fuera de línea**

Formatos de archivo

La Biblioteca implementará por fases la recepción de cierto tipo de obras y formatos teniendo en cuenta la política y el plan de preservación definido desde el área de Conservación, así como la infraestructura tecnológica. La Biblioteca promueve la recepción de formatos libres, estándar, no comerciales, con el mejor estándar de digitalización y sin pérdida de información. Teniendo en cuenta lo anterior los formatos definidos para la primera etapa de implementación son:

**Tabla 1: Formatos de archivo fuera de línea para depósito<sup>3</sup>**

<b>Recurso</b>	<b>Formato recomendado</b>
Texto	Texto plano, TXT, RTF, PDF, EPUB, FictionBook, ComicBook, HTML, XML, OpenDocument, PostScript (PS), DVI, LaTeX.
Imágenes	JPG, PNG, SVG, TIFF
Videos	Matroska (MKV, MKS y similares), MPG, OGG Theora
Audio	FLAC, OGG Vorbis, MPC, Speex, MKA, WAV

Se debe tener en cuenta que los formatos que podrían recibirse en depósito legal son diversos. La Biblioteca clasificará los tipos de formato recibidos, para apoyar el proceso de desarrollo de colecciones e identificar qué formatos no se han recibido en la Biblioteca y que podrían requerirse. Los usuarios pueden sugerir a la Biblioteca el depósito de obras en otros formatos, pero el depósito de estos tipos de obras está sujeto a la capacidad de la Biblioteca para conservarlos. Si la Biblioteca no puede garantizar la preservación de las obras sugeridas deberá rechazarse este tipo de depósito. La Biblioteca debe considerar la adquisición de medios para acceder a la información en formatos que no maneje, teniendo en cuenta los formatos especiales que se recibieron en depósito y que no pudieron recibirse.

El depositante debe especificar el software productor, el medio de visualización, así como el hardware u otros requerimientos técnicos específicos para su visualización.

Sin embargo, en caso de que la Biblioteca no pueda acceder a la información de los archivos o a su información relacionada, solicitará al depositante una nueva copia del documento electrónico.

Para etapas posteriores la Biblioteca ampliará paulatinamente los tipos de material y/o formatos a recibir como:

**Tabla 2: Clasificación de tipos de formatos<sup>4</sup>**

<b>Clasificación</b>	<b>Observaciones</b>
----------------------	----------------------

<sup>3</sup> Esta tabla se construyó basados en <http://www.openformats.org/es6> <http://www.ausgoal.gov.au/open-formats> [http://www.erevistas.csic.es/especial\\_revistas/revistas3.htm](http://www.erevistas.csic.es/especial_revistas/revistas3.htm) y la presentación del Comité de Tecnología de la Cámara Colombiana del Libro, durante la primera reunión del 2012. También contó con la verificación de la coordinadora de Conservación de la Biblioteca Nacional de Colombia.

<sup>4</sup> Tabla basada en dotwhat.net y fileinfo.com

3D / CAD	<ul style="list-style-type: none"> <li>• Imágenes en 3D</li> <li>• CAD</li> </ul>
Audio	
Archivos de Backups	
Archivos comprimidos	
Archivos de configuración	
Archivos de datos	<ul style="list-style-type: none"> <li>• Archivos de datos</li> <li>• Archivos de diseño de página (Page layout files)</li> <li>• Hojas de cálculo</li> </ul>
Archivos de bases de datos	
Archivos de desarrollo	
Imágenes de disco	
Archivos encriptados	<ul style="list-style-type: none"> <li>• Archivos de codificación (Encoded file type)</li> </ul>
Ejecutables	
Mapas	<ul style="list-style-type: none"> <li>• TIFF sin compresión, KML, WMS, WFS</li> </ul>
Fuentes	
Juegos	
Imágenes	<ul style="list-style-type: none"> <li>• Imagen de mapa de bits (raster images)</li> <li>• Imágenes vectoriales</li> <li>• Formato de imágenes de cámara sin modificaciones (Camera raw files)</li> </ul>
<b>Clasificación</b>	<b>Observaciones</b>
Archivos misceláneos	<ul style="list-style-type: none"> <li>• Sistemas de información geográfica (del inglés GIS)</li> </ul>
Plugins	
Archivos de sistema	
Archivos de texto	<ul style="list-style-type: none"> <li>• Texto</li> <li>• eBook</li> </ul>

Video	
Web	

### **3.2.2 Criterios de aprobación del depósito de obras digitales**

El depósito se aprobará por personal de la Biblioteca Nacional de Colombia. Este proceso verificará que:

- La información del depositante es correcta
- Cumple con los datos solicitados del depositante en el formulario de carga de documentos. Los objetos digitales han sido verificados por funcionarios de la Biblioteca y se han cargado correctamente.
- Los objetos digitales tienen los formatos establecidos por la Biblioteca.
- Los objetos digitales no se encuentran corruptos o en mal estado
- Los objetos digitales tienen metadatos completos
  - Incluye información sobre derechos de autor, software y hardware de visualización, cumpliendo con los metadatos PREMIS mínimos requeridos en la política de Preservación y Conservación.

Otras fuentes del depósito de obras fuera de línea

Para documentos digitales fijados y publicados en USB u otros dispositivos como CD, DVD, entre otros, se entrega la información en el soporte de producción original. Cuando la información fijada en estos soportes esté en formatos no definidos para la recepción de la primera fase, la Biblioteca recibirá el material y el área de Procesos lo reportará al Grupo de Conservación para que se realice la transferencia correspondiente.

No serán objeto de depósito las obras publicadas en intranets o grupos sociales con acceso restringido que tengan fundamentalmente información personal.

Superadas las primeras fases de la implementación del depósito de obras digitales, la Biblioteca trabajará por la recepción de radio y televisión en el depósito digital.

#### **4. Política de Acceso**

Se implementará un archivo de tres niveles de acceso; blanco, gris u oscuro; siendo el archivo blanco el que permite acceso sin restricciones a la información y el archivo oscuro el de uso exclusivo para funcionarios de la biblioteca. La razón para que existan estos tres niveles de acceso es porque los autores o titulares de derechos de obras digitales pueden solicitar que sus obras estén dentro de la Biblioteca, pero a través de un determinado nivel de restricción, que puede ir de archivo blanco a gris. Cada autor o titular del derecho es autónomo en la elección de uno de estos dos niveles de restricción para sus obras. El archivo oscuro se dará para aquellos casos especiales en que se almacenen obras y éstas no tengan el permiso del autor para su uso.

##### **4.1 Definiciones**

###### **4.1.1 Archivo blanco**

Espacio de almacenamiento de la biblioteca accesible sin restricciones. Los documentos en archivo blanco pueden consultarse a través de internet.

###### **Características:**

- Las obras se incluyen en este archivo por decisión del autor o titular del derecho de autor.
- La obra aparece disponible a través de internet.
- La obra aparece disponible en el catálogo de la Biblioteca.
- La obra aparece disponible para consulta en salas de la Biblioteca.
- No tiene restricciones de acceso.

###### **4.1.2 Archivo gris**

Espacio de almacenamiento de la biblioteca con algún tipo de restricción. Las restricciones pueden ser contraseñas, consulta únicamente en salas de la biblioteca o acceso a la información a través de un permiso especial por la biblioteca o para un público determinado de usuarios, por ejemplo únicamente investigadores de posgrado. Los documentos del archivo gris aparecen en el catálogo de la biblioteca.

También para conservar la privacidad de las personas mencionadas en las obras digitales, en caso de difamación u obras que atenten al buen nombre de alguien u obras que vayan en contra de la moral, se crea el archivo gris. Esto permite el equilibrio entre el acceso a la información a la que tienen derecho los usuarios y a la solicitud de los autores o titulares del derecho de autor que sientan

que el usufructo de su obra pueda verse afectado, al igual que las personas que sientan que su buen nombre o privacidad se encuentre en entredicho. En caso de que un usuario solicite que una obra pase a archivo gris, porque su privacidad o buen nombre se ven afectados, la Biblioteca responderá a la solicitud previa revisión jurídica de la misma.

La mecánica del archivo gris es que si un autor o titular de derecho de autor no quiere permitir que su obra esté disponible en el archivo blanco, puede acogerse al archivo gris, donde la Biblioteca implementará algún tipo de restricción de acceso.

El depositante puede seleccionar cualquiera de las restricciones de publicación en internet, pero la Biblioteca siempre brindará acceso a la obra desde sus salas. La Biblioteca definirá los mecanismos para brindar la contraseña para que se realice la consulta del archivo restringido a través de internet.

**Características:**

- Las obras se incluyen en este archivo por decisión del autor o titular del derecho de autor.
- Tiene restricciones de acceso a usuarios.
- La obra no aparece disponible a través de internet o puede consultarse en internet a través de una contraseña.
- La obra aparece disponible en el catálogo de la Biblioteca.
- La obra estará disponible para consulta en salas de la Biblioteca.

**4.1.3 Archivo oscuro**

Espacio de almacenamiento de la Biblioteca con la máxima restricción de acceso. Los documentos de este archivo son únicamente accesibles por los funcionarios de la biblioteca y no aparecen en el catálogo.

Los autores o titulares que no autoricen el uso de sus obras por parte de la Biblioteca Nacional permanecerán en el archivo oscuro, con máximas restricciones de acceso, hasta que dichas obras pasen a dominio público y puedan usarse sin infringir el derecho patrimonial o hasta que se cuente con una ley que permita a la Biblioteca usar la información digital.

Los autores o titulares que quieran restringir el acceso a su obra al ser depositada y obtener el certificado de depósito, únicamente podrán seleccionar el archivo gris como mecanismo de restricción y no el archivo oscuro.

**Características:**

- Las obras se incluyen en este archivo por decisión del personal de la Biblioteca.
- La obra no aparece disponible a través de internet.
- La obra no aparece disponible en el catálogo de la Biblioteca.
- La obra no aparece disponible para consulta en salas de la Biblioteca.
- Tiene restricciones de acceso.
- Puede consultarse únicamente por funcionarios de la Biblioteca.

**5. Comité de Selección y Adquisiciones**

Los procesos de selección que demanda el depósito digital estarán a cargo del Comité de Selección y Adquisiciones de la Biblioteca Nacional. Para este caso el comité, además de sus miembros permanentes, contará con académicos e investigadores expertos en diferentes áreas del conocimiento, representantes del gobierno y personal de la Biblioteca Nacional de Colombia cuyos conocimientos y experiencia les permita identificar fuentes de información web y otros recursos electrónicos relevantes para la Biblioteca. Por ejemplo: Representantes del Ministerio de las Tecnologías de Información y Comunicación, Colciencias, Biblioteca Luis Ángel Arango, RTVC, Patrimonio Filmico Colombiano, investigadores adscritos a universidades, representantes del comité de la Biblioteca Digital de la Biblioteca Nacional, entre otros.

**6. Visualización y consulta del material recopilado en depósito**

Los documentos digitales fuera de línea serán catalogados en el módulo de control bibliográfico del sistema integrado de la Biblioteca Nacional (ILS), sin embargo, una vez se una el componente de documentos digitales al ILS, serán migrados a éste. Por cada recurso digital fuera de línea recibido se elaborará un registro en el sistema. Para los sitios web se elaborará la catalogación del recurso como un todo representándolo en un solo registro, absteniéndose de elaborar registros independientes por los documentos que estos contienen. Es este caso, tanto la cosecha de los sitios como la captura de metadatos del registro, se harán a través del Sistema de Archive-it, sistema independiente del ILS de la Biblioteca, lo cual significa que en una primera fase no se integrarán, en tanto entre en funcionamiento y se realicen las pruebas de articulación entre dichos sistemas.

## **7. Anexos**

### **7.1 Glosario**

#### **Archivo blanco**

Espacio de almacenamiento de la biblioteca accesible sin restricciones. Los documentos en archivo blanco pueden consultarse incluso a través de internet.

#### **Archivos ejecutables**

Entendemos ejecutables a los archivos que permiten la visualización de la información. Para la edición de los ejecutables se requieren las fuentes.

#### **Archivos fuentes**

Las fuentes son los archivos que permiten la creación o edición de un objeto digital. Una fuente genera un ejecutable, que es el archivo donde se visualiza la información. En algunos casos las fuentes y los ejecutables pertenecen a un mismo archivo. Por ejemplo: Un documento de Word es ejecutable y fuente a la vez. Sin embargo, un archivo PDF puede ser el ejecutable y su fuente un archivo de Word.

#### **Archivo gris**

Espacio de almacenamiento de la biblioteca con algún tipo de restricción. Las restricciones pueden ser contraseñas, consulta únicamente en salas de la biblioteca o acceso a la información a través de un permiso especial por la biblioteca o para un público determinado de usuarios, por ejemplo únicamente investigadores de posgrado. Los documentos del archivo gris aparecen en el catálogo de la biblioteca.

#### **Archivo oscuro**

Espacio de almacenamiento de la biblioteca con la máxima restricción de acceso. Los documentos de este archivo son únicamente accesibles por los funcionarios de la biblioteca y no aparecen en el catálogo.

#### **Carpeta compartida**

Es una carpeta que se duplica en los computadores que tengan acceso a ella. Ejemplo: Dropbox, Ubuntu One o similares.

#### **Contenidos dinámicos**

Es un archivo digital publicado en línea que puede variar con el tiempo y por lo tanto su depósito debería realizarse periódicamente.

### **Contenidos estáticos**

Es un archivo digital publicado en línea que no varía con el tiempo y su depósito puede realizarse una única vez.

### **Formulario web**

Mecanismo para capturar los datos de quien realiza el depósito legal. Se pueden incluir archivos adjuntos (objetos digitales) en el formulario.

### **FTP**

Protocolo de transferencia de archivos. Una forma de subir documentos de gran tamaño o cantidad a un sitio web.

### **Harvesting**

También conocido como web harvesting o web scraping. Mecanismo para realizar copias locales de sitios web. Extracción de datos de sitios web.

### **Metadatos**

Datos que describen un documento. Por ejemplo: Autor, título, palabras clave, dirección web, derechos de autor.

### **Objeto digital o material digital**

Un archivo que contiene información. Un documento susceptible de ser depositado. Un PDF, DOC, MP3 o AVI son ejemplos de objetos digitales. Los objetos digitales pueden ser obras publicadas y creados en línea o fuera de línea.

### **Patrimonio digital**

Conjunto de materiales digitales que poseen el suficiente valor para ser conservados para que se puedan consultar y utilizar en el futuro (Biblioteca Nacional de Australia, 2003).

### **Patrimonio digital colombiano**

Conjunto de obras o documentos digitales y recursos web que conforman una colección nacional que incluye toda obra considerada herencia y memoria nacional o que contribuya a la construcción de identidad de la nación.

Incluye textos, bases de datos, imágenes fijas o en movimiento, grabaciones sonoras, material gráfico, programas informáticos o páginas Web, entre otros.

**Semillas (Seeds)**

Los sitios web que se programan para ser recopilados mediante crawling.

**Unidad lógica (un único archivo a adjuntar)**

Un documento se puede conformar por múltiples archivos. Por ejemplo un sitio web puede conformarse por texto (archivo HTML), imágenes (archivos JPG) y sonidos (archivos MP3) entre otros. Una unidad lógica puede ser un único archivo comprimido (ZIP, por ejemplo) que los contiene a todos los demás.

**Web crawler**

Software que busca en internet sitios web. Imagine una araña en búsqueda de alimento en su red. Así funcionan los web crawlers para rastrear y enviar la nueva información publicada en la web para completar el proceso de harvesting.

## 8. Bibliografía

- Biblioteca Nacional de Australia. (2003). *Directrices para la preservación del patrimonio digital* (p. 186). UNESCO. Retrieved from <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>
- British Library. (2012, April). Legal Deposit in the British Library. Retrieved May 28, 2012, from <http://www.bl.uk/aboutus/stratpolprog/legaldep/>
- Elliott, A. (2010). Electronic legal deposit: The New Zeland experience. Retrieved from <http://conference.ifla.org/past/ifla77/193-elliott-en.pdf>
- Gömpel, R., & Svensson, L. (2011). La gestión del depósito legal de las publicaciones en línea en Alemania. Puerto Rico. Retrieved from <http://conference.ifla.org/past/ifla77/193-goempel-es.pdf>
- IFLA. (2011a). IFLA Statement on Legal Deposit. Retrieved May 30, 2012, from <http://www.ifla.org/en/publications/ifla-statement-on-legal-deposit>
- IFLA. (2011b). IFLA Statement on Orphan Works. Retrieved October 24, 2012, from <http://www.ifla.org/publications/ifla-statement-on-orphan-works>
- IFLA. (2012a). Treaty proposal on limitaions and exceptions for libraries and archives. Retrieved from [http://www.ifla.org/files/assets/hq/topics/exceptions-limitations/documents/TLIB\\_v4.3\\_050712.pdf](http://www.ifla.org/files/assets/hq/topics/exceptions-limitations/documents/TLIB_v4.3_050712.pdf)
- IFLA. (2012b). IFLA E-Lending Background Paper. Retrieved from <http://www.ifla.org/files/clm/publications/ifla-background-paper-e-lending-en.pdf>
- International Communia Association. (2012). Positive Agenda for the Public Domain. *COMMUNIA*. Retrieved from <http://www.communia-association.org/2012/12/05/communia-positive-agenda-for-the-public-domain/>
- 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. *IFLA Journal*, 38(1), 5–24. doi:10.1177/0340035211435323

Política de depósito de documentos digitales

2012

David Ramírez-Ordóñez  
Biblioteca Nacional de Colombia  
10/11/2012