

Library of Congress Network Development and MARC Standards Office

Entender PREMIS

Autora: Priscilla Caplan

Fecha: 1 de febrero de 2009

Copyright © 2009 The Library of Congress, excepto en U.S.A. Deben consignarse los créditos en los extractos de esta publicación.

Traducción: María Luisa Martínez-Conde
Subdirección General de Coordinación Bibliotecaria
Dirección General del Libro, Archivos y Bibliotecas
Gobierno de España. Ministerio de Cultura

NIPO: 551-09-042-5

ENTENDER PREMIS

CONTENIDOS

1. PREMIS EN CONTEXTO	3
1.1. ¿Qué son los metadatos de preservación?	3
1.2. ¿Qué es PREMIS?	4
1.3. ¿Qué incluye el Diccionario de datos PREMIS?	4
1.4. ¿Cómo debe utilizarse PREMIS?	6
1.5. ¿Debe utilizarse PREMIS?	7
2. CONVENCIONES DEL DICCIONARIO DE DATOS	8
2.1. Unidades semánticas	8
2.2. Contenedores y subunidades	8
2.3. Contenedores de extensión	9
3. MODELO DE DATOS PREMIS	10
3.1. Entidad Intelectual	10
3.2. Entidad Objeto	11
3.3. Acontecimientos	12
3.4. Agentes	13
3.5. Derechos	13
4. EL DICCIONARIO DE DATOS	14
4.1. Muestra de una entrada del Diccionario de Datos para una unidad semántica simple	14
4.2. Muestra de una entrada del Diccionario de datos para una unidad contenedora	16
5. PREMIS EN USO	17
5.1. PREMIS en XML	17
5.2. Conformidad con PREMIS	18
6. MÁS INFORMACIÓN	19
APÉNDICES	
Apéndice A: Lista de todas las unidades semánticas de PREMIS	20
Apéndice B: Ejemplo de objeto	25
Apéndice C: Glosario de términos	29

ENTENDER PREMIS

Esta guía constituye una breve visión general del estándar PREMIS de metadatos de preservación. No proporciona la información suficiente para implementar PREMIS, pero da una idea de todo lo que incluye PREMIS. Para muchos lectores será suficiente. Para quienes necesiten dominar el *Diccionario de datos PREMIS* de metadatos de preservación, esta guía puede servir como una sencilla introducción para familiarizarse con el documento más amplio.

En esta guía los términos que aparecen en Verdana son los nombres de las unidades semánticas de PREMIS. La primera referencia de los términos que aparecen en el Glosario se ha impreso en *cursiva*.

1. PREMIS EN CONTEXTO

1.1. ¿Qué son los metadatos de preservación?

Si se trabaja en una biblioteca o en un archivo, es bastante posible que se tengan ciertos conocimientos sobre metadatos y descripción de recursos. Probablemente se sepa que los metadatos se clasifican en distintas categorías de acuerdo con las funciones que cumplen: los metadatos descriptivos ayudan a identificar y recuperar los recursos, los metadatos administrativos ayudan a gestionarlos y rastrearlos y los metadatos estructurales indican cómo reunir objetos digitales complejos para que se puedan visualizar o utilizar de algún modo. De manera análoga, los *metadatos de preservación* soportan las actividades cuyo objetivo es asegurar la utilización a largo plazo de un recurso digital.

El Diccionario de datos PREMIS define los metadatos de preservación como "la información que utiliza un repositorio para soportar el proceso de preservación digital". Son ejemplos de actividades de preservación y de cómo pueden soportarlas los metadatos:

- Un recurso debe almacenarse de manera segura para que nadie pueda modificarlo inadvertidamente (o malintencionadamente). La información de la suma de verificación almacenada como metadatos puede utilizarse para expresar si el archivo almacenado ha variado entre dos momentos en el tiempo.
- Los ficheros deben almacenarse en soportes que puedan leer los ordenadores actuales. Si el soporte está dañado u obsoleto (como los discos blandos de 8" utilizados en los años setenta) puede resultar difícil o imposible recuperar los datos. Los metadatos pueden apoyar la gestión de los distintos soportes mediante el registro del tipo y la edad del soporte de almacenamiento y las fechas en las que se refrescaron los ficheros por última vez.
- Transcurrido un período largo de tiempo, incluso los formatos de fichero más comunes pueden convertirse en obsoletos, lo que significa que las aplicaciones actuales no pueden reproducirlos. Los gestores de la preservación deben emplear *estrategias de preservación* que garanticen que los recursos se

puedan seguir utilizando. Esto puede significar la transformación de los antiguos formatos en otros nuevos equivalentes (*migración*), o la imitación del antiguo entorno de reproducción en el nuevo hardware y software (*emulación*). Tanto las estrategias de emulación como las de migración requieren metadatos sobre los formatos de los ficheros originales y los entornos de hardware y software que los soportan.

- Las acciones de preservación pueden implicar modificaciones de los recursos originales o cambios en su modo de reproducción. Esto puede poner en duda la autenticidad del recurso. Los metadatos pueden ayudar a soportar la autenticidad del recurso mediante la documentación de la *procedencia digital* de dicho recurso – su cadena de custodia y el historial de cambios autorizados.

1.2. ¿Qué es PREMIS?

PREMIS son las siglas de "PREservation Metadata: Implementation Strategies" (Metadatos de preservación: estrategias de implementación) que es el nombre de un grupo de trabajo internacional patrocinado por OCLC y RLG desde 2003-2005. Este grupo de trabajo elaboró un informe denominado *PREMIS Data Dictionary for Preservation Metadata* (Diccionario de datos PREMIS de metadatos de preservación) que incluye un diccionario de datos e información sobre los metadatos de preservación. En marzo de 2008 se publicó la segunda versión actualizada. La Library of Congress mantiene un esquema de representación de PREMIS en XML.

Existe una eficaz *PREMIS Maintenance Activity* (Actividad de Mantenimiento de PREMIS) patrocinada por la Library of Congress. Incluye un sitio web con enlaces a todo tipo de información oficial y no oficial sobre PREMIS, una lista de discusión y una wiki para los implementadores de PREMIS y un Comité Editorial responsable de las revisiones del Diccionario de datos y el esquema. La Actividad de Mantenimiento también promueve el conocimiento de PREMIS, patrocina tutoriales para su utilización y encarga estudios y publicaciones relacionados con PREMIS, como esta guía.

Por lo general, cuando alguien se refiere a "PREMIS", se refiere al Diccionario de datos. A veces, puede referirse al esquema XML, al grupo de trabajo, o a todo, incluida la Actividad de Mantenimiento.

Diccionario de Datos PREMIS: www.loc.gov/standards/premis/v2/premis-2-0.pdf

Sitio web de PREMIS: www.loc.gov/standards/premis/

Lista de discusión del Grupo de Implementadores de PREMIS: listserv.loc.gov/listarch/pig.html

1.3. ¿Qué incluye el Diccionario de Datos PREMIS?

El Diccionario de datos PREMIS define un conjunto de *unidades semánticas* fundamentales (véase sección 2.1) que deben entender los repositorios para llevar a cabo sus funciones de preservación. Las funciones de preservación pueden variar de un repositorio a otro, pero, por lo general, incluirán las acciones que garantizan que los objetos digitales permanezcan viables (es decir, que los soportes se puedan leer) y

recuperables (es decir, que puedan visualizarse, ejecutarse o representarse de alguna manera por una aplicación de software) y que aseguren que los objetos digitales del repositorio no se han alterado inadvertidamente y que se han documentado los cambios legítimos de los objetos.

El Diccionario de datos no tiene como objetivo definir todos los elementos posibles de los metadatos de preservación sino únicamente los que necesitan entender la mayoría de los repositorios. Se excluyen algunas categorías de metadatos por considerarse fuera del alcance del Diccionario. Entre ellas se incluyen:

- Los metadatos de un formato específico, es decir, los metadatos que pertenecen solo a un formato de fichero o a una clase de formato como audio, video o gráficos de vectores.
- La implementación de metadatos específicos y reglas de negocio, es decir, los metadatos que describen las políticas o prácticas de un repositorio en particular, por ejemplo, cómo proporciona dicho repositorio el acceso a los materiales.
- Los metadatos descriptivos. Aunque la descripción de los recursos es, obviamente, relevante para la preservación, pueden utilizarse varios estándares independientes para este objetivo, como MARC21, MODS y Dublin Core.
- La información detallada sobre el soporte o el hardware. De nuevo, aunque asimismo está claro que son relevantes para la preservación, estos metadatos deben ser definidos por otras comunidades.
- La información sobre agentes (personas, organizaciones o software) distintos de los mínimos necesarios para la identificación.
- La información sobre derechos y permisos, excepto los que afectan directamente a las funciones de preservación.

Si se tienen presentes todos los metadatos que necesita una organización que gestiona un repositorio de preservación, PREMIS puede considerarse el subconjunto que se define en el centro, que no está relacionado con la recuperación y el acceso ni se propone definir los metadatos detallados de un formato específico. Solamente define los metadatos que se necesitan, por lo general, para llevar a cabo las funciones de preservación de todos los materiales.

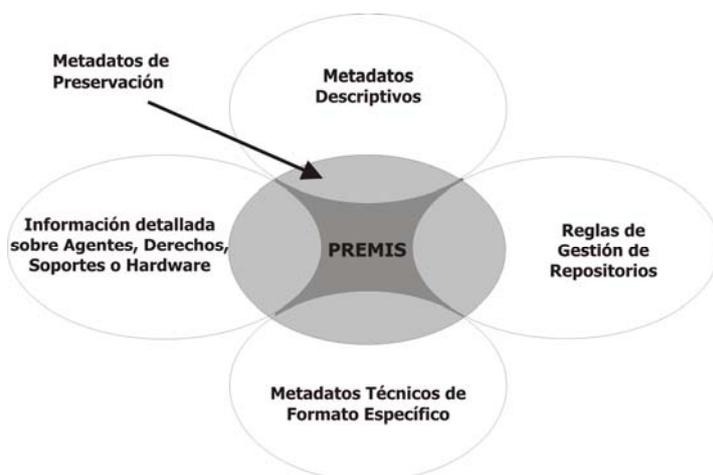


Gráfico 1. PREMIS como subconjunto de todos los metadatos de preservación

El gráfico 1 muestra todos los metadatos relevantes para la preservación en el círculo sombreado en el centro del diagrama. El círculo incluye metadatos descriptivos, reglas de negocio, metadatos técnicos detallados e información detallada sobre agentes, derechos, soportes y hardware. PREMIS es el pequeño núcleo del corazón de los metadatos de preservación, que se muestra en gris oscuro, y que excluye todos los demás tipos.

1.4 ¿Cómo debe utilizarse PREMIS?

El Diccionario de datos PREMIS define lo que necesita entender un repositorio de preservación. Es importante observar que se centra en el sistema de repositorio y en su gestión y no en los autores del contenido digital, las personas que escanean o las que de alguna manera convierten el contenido analógico en digital o en las que evalúan y licencian los recursos electrónicos digitales. PREMIS se utiliza fundamentalmente para el diseño de los repositorios, para su evaluación y para el intercambio de los paquetes de información archivada entre los repositorios de preservación.

Los diseñadores y/o desarrolladores de aplicaciones de software de repositorios de preservación deben utilizar PREMIS como una guía sobre la información que dichas aplicaciones deben obtener y registrar o que de alguna manera debe tenerse en cuenta para la gestión del repositorio.

Quienes tengan previsto implementar un repositorio de preservación, deben utilizar PREMIS como una lista de control para evaluar los posibles softwares. Los sistemas que soporten el Diccionario de datos de PREMIS serán más capaces de preservar los recursos de información a largo plazo.

Un repositorio activo necesitará en algún momento exportar los paquetes de información almacenada para su ingesta en otro repositorio. Por ejemplo, una organización de custodia puede migrar de un sistema de repositorio a otro o un cliente puede querer cambiar de un servicio de terceros a otro. PREMIS proporciona un conjunto común de elementos de datos que pueden entender tanto los repositorios de exportación como los de importación, especialmente si se utiliza el esquema PREMIS XML.

1.5. ¿Debe utilizarse PREMIS?

Depende. La mayor parte del personal de las bibliotecas, archivos, museos y otras instituciones del patrimonio cultural no tienen ninguna implicación directa en la preservación digital. En este caso, es suficiente saber lo que es PREMIS: un diccionario de datos de los metadatos de preservación. Si su trabajo implica alguna responsabilidad en relación con cualquier aspecto de la preservación digital, probablemente le resultará útil estar familiarizado con PREMIS. Si está implicado en la evaluación o implementación de un repositorio institucional o un sistema de preservación, es necesario un buen conocimiento de PREMIS. Considere la posibilidad de realizar un tutorial de PREMIS si es posible; la Actividad de Mantenimiento de PREMIS los ofrece periódicamente.

Si trabaja en proyectos de digitalización, es posible que se pregunte si debe crear los metadatos de PREMIS para utilizarlos posteriormente. La mayor parte de los elementos de PREMIS se han diseñado para que los proporcione automáticamente la aplicación del repositorio de preservación. (Desde luego, esto no significa que las aplicaciones actualmente disponibles los proporcionen). Sin embargo, si es posible, hay cierta información que se debe registrar:

Inhibidores. Los *Inhibidores* se definen como las características de un objeto al que se pretende inhibir el acceso, uso, o migración. Los inhibidores incluyen la palabra de paso de protección y encriptación. Es difícil describir los inhibidores por programa porque este puede no ser capaz de analizar el objeto, por lo que si se sabe que un fichero tiene inhibidores, es importante anotarlo. PREMIS define unidades semánticas para el tipo de inhibidor, el objetivo (las acciones que se inhiben) y la clave (la palabra de paso u otro mecanismo para eludir el inhibidor)

Procedencia. La *procedencia digital* es el registro de la cadena de custodia y la historia de los cambios de un objeto digital. Si una institución ha creado un objeto, las circunstancias de su creación son, obviamente, una parte importante de su procedencia. El nombre y la versión de la aplicación creadora y la fecha de creación pueden extraerse con frecuencia de la cabecera del fichero, pero no siempre, por lo que se recomienda registrar esta información. PREMIS permite que la historia de los cambios se consigne como Información sobre el acontecimiento, que se describe más abajo. Sin embargo, los tipos de acontecimiento de PREMIS se han diseñado fundamentalmente para describir las acciones que ocurren después de enviar algo a un repositorio para su ingesta. Para registrar los acontecimientos que suceden antes de la ingesta, como la adquisición y el acceso, puede ser necesario idear tipos de acontecimiento propios.

Propiedades significativas. Las *propiedades significativas* son las características de un objeto que deben mantenerse a través de las acciones de preservación. Por ejemplo, si se posee un documento, ¿solo son críticos las imágenes y el texto o son, asimismo, importantes los fuentes, el fondo, el formateado y otras características relacionadas con el aspecto? La idea de las propiedades significativas es uno de los conceptos más importantes de la preservación digital y uno de los menos conocidos. Existe un cierto número de iniciativas dedicadas a mejorar el modelado y descripción de las propiedades significativas, pero de momento se encuentran en su fase inicial. No

obstante, cualquier institución que cree o adquiera materiales digitales para una comunidad de usuarios debe tener muy claro que las características de dichos materiales son importantes para dicha comunidad e intentar registrar esta información para utilizarla en el futuro.

Derechos. La información sobre los derechos no es necesaria únicamente para la preservación, pero es muy importante para el proceso de preservación saber lo que se puede hacer con un objeto. Debe registrarse cuidadosamente cualquier información conocida sobre derechos, incluyendo el estado del copyright, los términos de la licencia y los permisos especiales.

2. CONVENCIONES DEL DICCIONARIO DE DATOS

2.1. Unidades semánticas

El Diccionario de datos PREMIS define unidades semánticas, no elementos de metadatos. Esta distinción es sutil, pero es importante. Una unidad semántica es una pieza de información o de conocimiento. Un elemento de metadatos es una manera definida de representar esa información en un registro de metadatos, en un esquema o en una base de datos. PREMIS no especifica cómo deben representarse los metadatos en un sistema, únicamente define lo que necesita entender el sistema y lo que debe poder exportarse a otros sistemas. Así, para ser un purista de PREMIS, hay que pensar en términos de unidades semánticas bastante abstractas. Las unidades semánticas de PREMIS presentan un mapeo directo con los elementos de metadatos definidos en el esquema PREMIS XML, y pueden tener un mapeo menos directo con los metadatos de otro esquema.

Los nombres de las unidades semánticas de PREMIS son cadenas "mayúsculas/minúsculas camello". Es decir, palabras que no se separan por espacios sino por letras mayúsculas: `objectIdentifier` (identificador del objeto), `relatedEventIdentification` (identificación del acontecimiento relacionado). En este documento se han impreso en Verdana.

2.2. Contenedores y subunidades

Algunas unidades semánticas se definen como *contenedores*, lo que significa que no tienen un valor en sí mismas sino que existen para agrupar unidades semánticas relacionadas. Por ejemplo, donde quiera que se registre un identificador en PREMIS debe especificarse de qué tipo de identificador se trata (por ejemplo, "DOI", "ISBN", "asignado por el sistema local"). El contenedor `objectIdentifier` (identificador del objeto) se utiliza para agrupar las dos subunidades `objectIdentifierType` (tipo de identificador del objeto) y `objectIdentifierValue` (valor del identificador del objeto).

Los contenedores proporcionan una estructura jerárquica al Diccionario de datos que en la versión 2.0 se refleja en la numeración de las unidades semánticas:

1.1 objectIdentifier (identificador del objeto) (O, R)

1.1.1 objectIdentifierType (tipo de identificador del objeto) (O, NR)

1.1.2 objectIdentifierValue (valor del identificador del objeto) (O, NR)

Este extracto del Diccionario de datos muestra a simple vista que la unidad semántica objectIdentifier (identificador del objeto) es obligatoria (O) y repetible (R). Puesto que existen unidades semánticas definidas bajo ella, se puede deducir que el objectIdentifier (identificador del objeto) no tiene un valor en si mismo sino que sirve como contenedor de los elementos componentes objectIdentifierType (tipo de identificador del objeto) y objectIdentifierValue (valor del identificador del objeto). Puesto que el objectIdentifierType (tipo de identificador del objeto) y el objectIdentifierValue (valor del identificador del objeto) son no repetibles (NR) dentro del contenedor, hay que repetir toda la estructura del contenedor para registrar dos identificadores distintos.

2.3. Contenedores de extensión

Un *contenedor de extensión* es un tipo especial de contenedor que no tiene subunidades definidas bajo si mismo. Se ha diseñado para disponer de un lugar donde registrar los metadatos no-PREMIS. En este sentido, PREMIS puede extenderse para incluir metadatos que se encuentran fuera de su alcance u otros no incluidos en el Diccionario de datos. Los contenedores de extensión llevan "Extension" (Extensión) en la última parte de su nombre.

Por ejemplo, los metadatos técnicos de un formato específico no se incluyen en PREMIS, pero es una información muy importante para la preservación digital. El contenedor de extensión 'extensión de las características del objeto' (objectCharacteristicsExtension) proporciona un lugar en el que registrar los metadatos técnicos definidos por otros diccionarios de datos, por ejemplo el estándar Z39.87 para describir el mapa de bits de las imágenes.

Si se está familiarizado con XML, será obvio que el Diccionario de datos PREMIS se ha diseñado de manera que sea compatible con XML. Las unidades semánticas de PREMIS pueden implementarse como elementos XML; las unidades contenedoras son elementos que únicamente toman otros elementos como contenido y las unidades de extensión son contenedores de los elementos definidos por un esquema externo. Más adelante, en la sección 5.1, se encuentra más información sobre PREMIS y XML.

3. EL MODELO DE DATOS PREMIS

Gráfico 2. El modelo de datos PREMIS

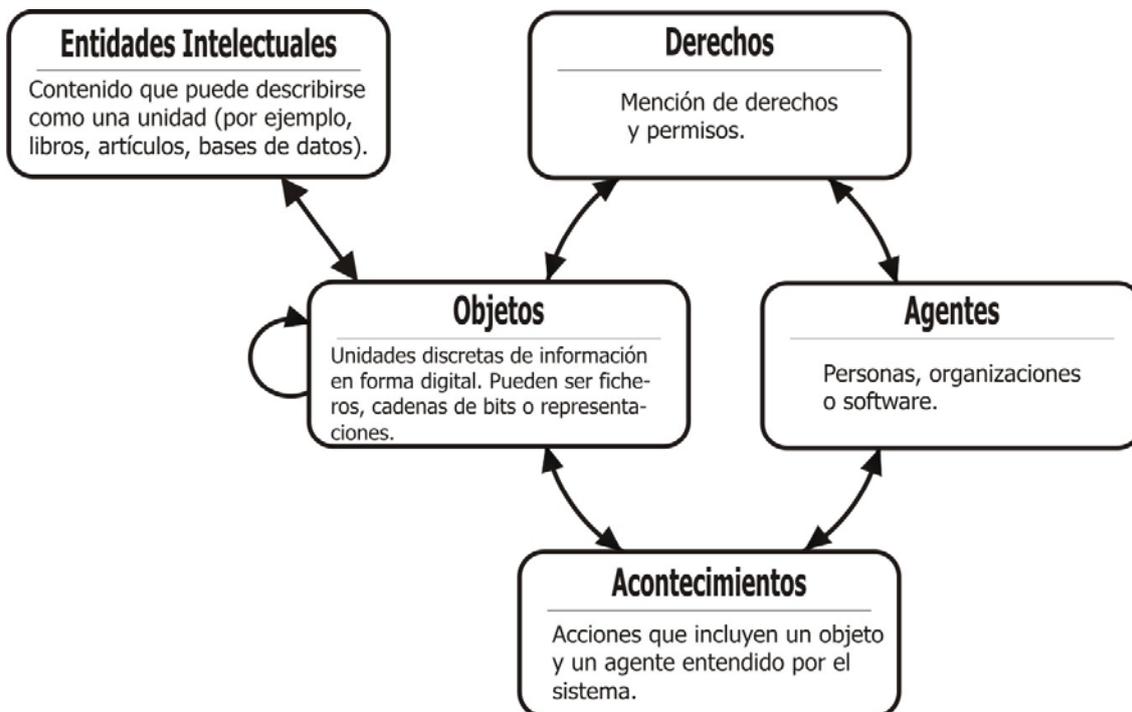


Gráfico 2. Modelo de datos PREMIS

Uno de los principios más importantes de PREMIS es que es necesario tener muy claro lo que se describe. PREMIS define cinco tipos de cosas (llamadas *entidades*) de las que se puede hablar: Entidades Intelectuales, Objetos, Agentes, Acontecimientos y Derechos. Esto es lo que se denomina modelo de datos PREMIS que se muestra más arriba, en el gráfico 2.

3.1 Entidad Intelectual

Las *Entidades Intelectuales* son conceptuales y pueden denominarse “entidades bibliográficas”. PREMIS define la Entidad Intelectual como “un conjunto de contenido que se considera como una sola unidad intelectual para los propósitos de gestión y descripción: por ejemplo, un libro, un mapa, una fotografía o una base de datos”. En realidad, PREMIS no define los metadatos correspondientes a las Entidades Intelectuales porque existen muchos estándares de metadatos descriptivos entre los que se puede elegir.

PREMIS establece que en un sistema de preservación un objeto debe estar asociado a la entidad intelectual que representa mediante la inclusión de un identificador de dicha entidad en los metadatos del objeto. Así, por ejemplo, si estamos preservando un ejemplar de *Buddishm: The Ebook: an Online Introduction* podemos utilizar el ISBN como enlace a la Entidad Intelectual en la descripción PREMIS del e-libro.

3.1 Entidad Objeto

Los *Objetos* son lo que realmente se almacena y gestiona en un repositorio de preservación. La mayor parte de PREMIS se dedica a describir objetos digitales. La información que se puede registrar incluye:

- el identificador único del objeto (tipo y valor),
- fijeza de la información, como la suma de verificación (mensaje cifrado) y el algoritmo utilizado para obtenerla,
- el tamaño del objeto,
- el formato del objeto, que puede especificarse directamente o mediante un enlace a un registro de formatos,
- el nombre original del objeto,
- información sobre su creación,
- información sobre los inhibidores,
- información sobre sus propiedades significativas,
- información sobre su entorno (véase más abajo),
- dónde y en qué soporte se almacena,
- información sobre la firma digital,
- relación con otros objetos y otros tipos de entidades.

Para registrar el *entorno* de un objeto se han definido varias unidades semánticas, es decir, qué hardware y software se necesita para recuperarlo y cuáles son las dependencias con otros objetos. Por ejemplo, un fichero PDF puede visualizarse con distintas versiones de Adobe Acrobat y Adobe Reader así como con otros programas tanto comerciales como de código abierto. Cada una de ellas, a su vez, es soportada por distintos sistemas operativos y requiere ciertas especificaciones mínimas de hardware (velocidad del procesador, memoria y disco). Debido a que Adobe Reader no es una aplicación autónoma sino un navegador plug-in, es dependiente asimismo de determinadas versiones de determinados navegadores; por ejemplo, Reader 9 para Mac OS requiere el navegador Safari versión 2.0.4 o posterior.

La información sobre el entorno es crítica para determinadas estrategias de preservación. Cierta información simple sobre el entorno puede registrarse fácilmente en PREMIS, pero otra información más compleja sobre el entorno puede ser difícil de determinar y llevar mucho tiempo. En el caso de que la información detallada se aplique a clases de objetos más que a instancias locales de objetos, como es el caso de la información sobre el entorno, puede resultar útil agregarla a los registros y que la compartan todos los repositorios. PREMIS permite enlazar los repositorios a información almacenada en registros externos si se prefiere esta opción en lugar de almacenarla a nivel local.

En realidad, PREMIS define tres tipos diferentes de objetos y requiere que los implementadores establezcan una distinción entre ellos. Dichas clases son: objetos *fichero*, objetos *representación* y objetos *cadena de bits*.

El objeto fichero es exactamente como suena: un fichero de ordenador, por ejemplo un fichero PDF o un JPEG.

El objeto representación es el conjunto de todos los objetos fichero que se necesitan para reproducir una Entidad Intelectual. Por ejemplo, supongamos que se quiere

preservar una página web, quizá la página de inicio de su institución en una determinada fecha. Hay bastantes posibilidades de que la página de inicio que se ve en el navegador esté formada por diferentes tipos de ficheros –uno o más ficheros HTML, unas cuantas imágenes TIFF o JPEG, quizá un pequeño audio o una animación Flash. Probablemente también utiliza una hoja de estilo para su visualización. Un navegador reúne todos estos ficheros para reproducir la página de inicio y visualizarla de manera que si un repositorio quiere preservar una página web que se pueda reproducir tiene que entender todos estos ficheros y saber cómo reunirlos. El objeto representación permite al repositorio no solo identificar el conjunto de ficheros relacionados sino también describir las características de la totalidad (e.g., la página web como un todo) que pueden ser diferentes de las de sus partes.

Los objetos cadena de bits son subconjuntos de ficheros. Un objeto cadena de bits se define como datos (bits) dentro de un fichero que a) presenta propiedades comunes para los propósitos de preservación, y b) no puede ser autónomo sin añadir un fichero cabecera u otra estructura. Así, por ejemplo, si se dispone de un fichero en formato AVI (audio-video intercalado) puede que se quiera diferenciar la cadena de bits del audio de la cadena de bits del vídeo y describirlos como objetos cadena de bits independientes.

Algunas unidades semánticas definidas en el Diccionario de datos PREMIS se pueden aplicar a los tres tipos de objetos, mientras que otras solo se aplican a uno o dos tipos de objetos. El hecho de tener distintos tipos de objetos obliga a pensar en lo que se describe y a ser lo más preciso posible, lo que es importante para el proceso automático.

3.3. Acontecimientos

La *entidad Acontecimiento* agrega información sobre acciones que afectan a los objetos del repositorio. Un registro preciso y fiable de los acontecimientos es crítico para el mantenimiento de la procedencia digital de un objeto lo que, a su vez, es importante para demostrar la autenticidad del objeto.

La información sobre los acontecimientos que se puede registrar incluye:

- el identificador único del acontecimiento (tipo y valor),
- el tipo de acontecimiento (creación, ingesta, migración, etc.),
- la fecha y hora en la que ocurrió el acontecimiento,
- la descripción detallada del acontecimiento,
- el resultado codificado del acontecimiento,
- una descripción más detallada del resultado,
- los agentes implicados en el acontecimiento y sus funciones,
- los objetos implicados en el acontecimiento y sus funciones.

Cada sistema de repositorio puede tomar sus propias decisiones sobre los acontecimientos que se registran como parte permanente de la historia de un objeto. PREMIS recomienda que se registren siempre las acciones que modifican un objeto y la entrada para el tipo de acontecimiento (eventType) del Diccionario de Datos proporciona una “lista de entrada” de tipos de acontecimientos importantes con el

objetivo de promover que los repositorios registren estos acontecimientos de manera consistente.

3.4. Agentes

Los *Agentes* son actores con funciones en los acontecimientos y en las menciones de derechos (véase 3.5 Derechos). Los agentes pueden ser personas, organizaciones o aplicaciones de software. PREMIS solo define el número mínimo de unidades semánticas necesarias para identificar los agentes puesto que existen varios estándares externos que se pueden utilizar para registrar información más detallada. (Como muestra de ellos véase “Metadata standards and specifications for describing people and their interests” en www.ukoln.ac.uk/metadata/resources/people/.) Un repositorio puede elegir entre utilizar un estándar independiente para registrar información adicional sobre los agentes o utilizar el identificador del agente para apuntar a la información registrada externamente.

El Diccionario de datos incluye:

- un identificador único para el agente (tipo y valor),
- el nombre del agente,
- la designación del tipo de agente (persona, organización, software).

Siempre que se haga referencia a un agente en relación con un acontecimiento o con una mención de derechos, debe registrarse también la función del agente. Cualquier agente puede tener varias funciones. Por ejemplo, yo podría ser el autor y el derechohabiente de una obra, el autor (pero no el derechohabiente) de una segunda obra y el depositario de una tercera. En el modelo PREMIS un repositorio debe asignarme un identificador único y debe consignar el identificador del registro de cualquier acontecimiento o mención de derechos en el que soy un agente, junto con mi función en ese contexto particular. La función de un agente en relación con un acontecimiento o una mención de derechos se considera una propiedad de la entidad acontecimiento o de la entidad derechos y no del propio agente.

3.5. Derechos

La mayor parte de las estrategias de preservación implican la creación de copias idénticas y versiones derivadas de los objetos digitales, acciones que están restringidas por la ley del copyright a los derechohabientes. La *entidad Derechos* agrega información sobre los derechos y permisos que son directamente relevantes para preservar los objetos del repositorio. Cada una de las menciones de derechos de PREMIS constata dos cosas: las acciones a las que tiene derecho el repositorio y las bases para la reclamación de ese derecho.

Por ejemplo, un repositorio puede albergar la versión escaneada de un libro que se publicó en 1848 y se encuentra, por lo tanto, en el dominio público. El repositorio puede actuar sobre su versión digital sobre la base del estado del copyright del ítem. Otro repositorio alberga un objeto copiado de un CD publicado en el que la licencia de uso individual permite hacer copias de seguridad, pero restringe el acceso y el uso.

La información que puede registrarse en una mención de derechos incluye:

- el identificador único de la mención de derechos (tipo y valor),
- si la base para la reclamación de los derechos es el copyright, una licencia o una ley,
- información más detallada sobre el estado del copyright, los términos de la licencia o la ley, si es aplicable,
- la(s) acción (es) que permite la mención de derechos,
- cualquier restricción sobre la(s) acción(es),
- los derechos otorgados o el período de tiempo durante el que se aplica la mención,
- el (los) objeto(s) a los que se aplica la mención,
- los agentes implicados en la mención de derechos y sus funciones.

La mayor parte de la información está diseñada para ser *accionable* (es decir, registrada de una forma controlada que pueda ser ejecutada por un programa de ordenador). La mención de derechos de PREMIS es una consignación de derechos, no un registro de información a partir del que puedan determinarse los derechos. Esto es, PREMIS no define el tipo de información detallada sobre los autores, fecha y lugar de publicación y la notificación de los derechos de reproducción que se define en la especificación `copyrightMD` de la Biblioteca Digital de California (www.cdli.org/inside/projects/rights/schema/). El objetivo del `copyrightMD` es ayudar a los seres humanos a establecer las determinaciones de los derechos sobre bases actuales, mientras que el objetivo de la entidad Derechos de PREMIS es proporcionar información activable para los sistemas de los repositorios de preservación.

4. EL DICCIONARIO DE DATOS

4.1. Muestra de una entrada del Diccionario de datos para una unidad semántica simple

El gráfico 3 muestra la entrada del Diccionario de datos para la unidad semántica `size` (tamaño) que es un componente o subunidad del contenedor denominado `objectCharacteristics` (características del objeto). El tamaño en sí mismo carece de subunidades. La entrada del Diccionario de datos incluye la definición del elemento y la razón (fundamentos) por la que se incluye entre los metadatos fundamentales de PREMIS así como ejemplos y notas sobre cómo obtener y utilizar el valor. Se pretende que todos ellos ayuden a los implementadores a utilizar adecuadamente el elemento.

Las dos filas “Categoría del objeto” y “Aplicabilidad” se utilizan conjuntamente para mostrar si la unidad semántica es adecuada para describir representaciones, ficheros y/o cadenas de bits. Aquí se presenta el tamaño como perteneciente únicamente a ficheros y cadenas de bits. Finalmente, hay un conjunto de reglas de uso: “Restricción de los datos”, “Repetibilidad”, “Obligatoriedad”.

Las restricciones de los datos especifican restricciones sobre los valores que puede tener una unidad semántica. En este ejemplo, el valor del tamaño puede ser un entero.

Otra restricción común de los datos es que el valor puede tomarse de un vocabulario controlado. A veces los términos del vocabulario se especifican en el Diccionario de datos y a veces no, pero, en cualquier caso, debe registrarse el nombre del vocabulario utilizado. En el Diccionario de datos no se definen unidades semánticas para los nombres de vocabularios, pero la Actividad de Mantenimiento de PREMIS está desarrollando un esquema XML para solucionarlo.

La repetibilidad indica si se puede repetir la unidad semántica.

La obligatoriedad indica si un valor es obligatorio (requerido) u opcional para la unidad semántica. La obligatoriedad es potencialmente confusa, porque PREMIS establece claramente que no requiere un repositorio para almacenar una determinada información. Una unidad semántica obligatoria no debe registrarse y almacenarse en el repositorio. Sin embargo, el repositorio debe ser capaz de generar el valor de la unidad semántica cuando sea necesario, por ejemplo para el intercambio con otro repositorio. Por ejemplo, en el caso improbable de que un repositorio solo almacene imágenes TIFF 6.0, no tendrá que registrar el formato de la información de cada uno de los objetos. No obstante, el gestor del repositorio debe entender que sus objetos fichero son imágenes TIFF 6.0 y debe ser capaz de proporcionar esta información en caso de que sea necesaria. (Véase sección 5.2 Conformidad con PREMIS).

Unidad semántica	1.5.3 size (tamaño)		
Componentes semánticos	ninguno		
Definición	El tamaño en bytes del fichero o cadena de bits almacenados en el repositorio		
Fundamentos	El tamaño es útil para asegurar el número correcto de bytes de almacenamiento que se han recuperado y que una aplicación tiene espacio suficiente para mover o procesar los ficheros. También puede utilizarse cuando se factura por el almacenamiento.		
Restricciones de los datos	entero		
Categoría del objeto	Representaciones	Fichero	Cadena de bits
Aplicabilidad	No aplicable	Aplicable	Aplicable
Ejemplos		2038937	
Repetibilidad		No repetible	No repetible
Obligatoriedad		Opcional	Opcional
Notas de creación/mantenimiento	Obtenidas automáticamente del repositorio		
Notas de uso	La definición de esta unidad semántica como tamaño en bytes hace innecesario registrar una unidad de medida. Sin embargo, para el propósito de intercambio de datos la unidad de medida debe ser establecida o entendida por ambas partes.		

Gráfico 3. Extracto del Diccionario de datos para la unidad semántica size (tamaño)

4.2. Muestra de entrada del Diccionario de datos para una unidad contenedora

El gráfico 4 muestra el principio de la entrada del Diccionario de datos para las características del objeto [objectCharacteristics], la unidad contenedora del tamaño. Se puede decir que se trata de un contenedor porque tiene componentes semánticos y la restricción de los datos es “contenedor”. Obsérvese que los componentes semánticos incluidos pueden ser unitarios, como el tamaño, o contenedores en si mismos, como el formato.

Unidad semántica	1.5 objectCharacteristics (características del objeto)		
Componentes semánticos	1.5.1 compositionLevel (nivel de composición) 1.5.2 fixity (fijeza) 1.5.3 size (tamaño) 1.5.4 format (formato) 1.5.5 creatingApplication (aplicación creadora) 1.5.6 inhibitors (inhibidores) 1.5.7 objectCharacteristicsExtension (extensión de las características del objeto)		
Definición	Propiedades técnicas de un fichero o cadena de bits aplicables a todos o a la mayoría de los formatos.		
Fundamentos	Hay algunas propiedades técnicas importantes que se aplican a objetos en cualquier formato. La definición detallada de las propiedades de un formato específico queda fuera del alcance de este Diccionario de datos aunque dichas propiedades pueden incluirse en <i>objectsCharacteristicsExtension</i> (extensión de las características del objeto).		
Restricciones de los datos	Contenedor		
Categoría del objeto	Representación	Fichero	Cadena de bits
Aplicabilidad	No aplicable	Aplicable	Aplicable
Repetibilidad		Repetible	Repetible
Obligatoriedad		Obligatorio	Obligatorio
Notas de uso	Las unidades semánticas incluidas en <i>objectsCharacteristics</i> (características del objeto) deben tratarse como un conjunto de información que pertenece a un único objeto a un único nivel de composición [<i>compositionLevel</i>]. Las características del objeto pueden repetirse.		

Gráfico 4: Extracto del Diccionario de datos para la unidad semántica objectCharacteristics (características del objeto).

5. PREMIS en uso

5.1. PREMIS en XML

Existe la expectativa (aunque no es un requisito) de que si se utiliza PREMIS para intercambio se representará en XML. La Actividad de Mantenimiento de PREMIS proporciona un esquema XML que corresponde directamente al Diccionario de datos para proporcionar una sencilla descripción de los objetos, acontecimientos, agentes y derechos. El gráfico 5 muestra un fragmento de los metadatos PREMIS utilizando el esquema PREMIS XML.

En la práctica, la mayor parte de los sistemas de preservación ya utilizan formatos XML para la importación y exportación de los datos. Muchos utilizan METS (Metadata Encoding for Transmission Standard), otro estándar mantenido por la Library of Congress, como contenedor XML para reunir distintos tipos de metadatos. Es posible utilizar PREMIS dentro de METS, pero no es muy sencillo por dos razones. En primer lugar, METS divide la información en secciones distintas dependiendo de que se trate de metadatos técnicos, metadatos de derechos o metadatos de procedencia. El esquema PREMIS, de acuerdo con el Diccionario de datos, tiene secciones para objetos, derechos, acontecimientos y agentes. Hay cierta correspondencia entre las dos estructuras, pero no es perfecta, especialmente por lo que se refiere a la información sobre el agente. En segundo lugar, PREMIS y METS presentan algún solapamiento; por ejemplo, ambos definen una etiqueta para almacenar las sumas de verificación. Si se encapsulan los metadatos PREMIS dentro de METS, hay que decidir si los elementos que se solapan se registran en las secciones de PREMIS, en las secciones de METS o en ambas.

Si cada repositorio de preservación hubiera de tomar sus propias decisiones, es evidente que podrían producirse grandes variaciones en la manera de representar los datos que impedirían la interoperabilidad. Por ello, se han puesto en marcha varias iniciativas que ayuden a definir las mejores prácticas de utilización conjunta de METS y PREMIS.

```

<event>
  <eventIdentifier>
    <eventIdentifierType>DAITSS</eventIdentifierType>
    <eventIdentifierValue>10012</eventIdentifierValue>
  </eventIdentifier>
  <eventType>Validation</eventType>
  <eventDateTime>2008-05-06T01:40:22-04:00</eventDateTime>
  <eventOutcomeInformation>
    <eventOutcome>Invalid</eventOutcome>
    <eventOutcomeDetail>ill-formed DateTime value</eventOutcomeDetail>
  </eventOutcomeInformation>
</event>

```

Gráfico 5: fragmento de PREMIS en XML

5.2 Conformidad con PREMIS

El Diccionario de datos PREMIS contiene una sección sobre lo que significa para un repositorio ser conforme a PREMIS. Los requisitos fundamentales son tres:

1) Si el repositorio implementa (almacena o exporta) un elemento de datos que dé a entender que es una unidad semántica de PREMIS, los elementos de datos deben tener la misma definición, las mismas restricciones de datos y la misma aplicabilidad que la unidad semántica definida en PREMIS.

2) Si el repositorio implementa una unidad semántica de PREMIS, su repetibilidad y obligatoriedad puede ser más restrictiva de lo que requiere PREMIS, pero no más libre. Es decir, una unidad semántica repetible puede implementarse como no repetible, pero no al contrario, y un elemento obligatorio no puede convertirse en opcional.

3) Si un repositorio exporta información para que sea utilizada por otro repositorio, debe proporcionar los valores de todas las unidades semánticas que son obligatorias en el Diccionario de datos. Si embargo, existe una cierta flexibilidad en este aspecto debido a que no es necesario que un repositorio soporte unidades semánticas obligatorias para los tipos de entidades que no soporta. Dicho de otra manera, un repositorio puede libremente soportar o no soportar los Agentes de PREMIS, pero en el caso de que los soporte, es obligatorio el identificador del agente (`agentIdentifier`). De la misma manera, un determinado repositorio puede no soportar objetos cadena de bits en cuyo caso no tiene que proporcionar el identificador de la cadena de bits que sería obligatorio en caso de que los soportara.

Estos requisitos deben considerarse en el contexto de ciertos aspectos que NO son requisitos de conformidad. Como se observaba más arriba, no es obligatorio que un repositorio soporte todos los tipos de entidad definidos en el modelo de datos PREMIS. Tampoco es necesario almacenar internamente los metadatos utilizando los nombres de las unidades semánticas de PREMIS o los valores que implican las restricciones de los datos de PREMIS. En otros términos, no importa como “entienda” un repositorio un valor de PREMIS –mediante su almacenamiento con el mismo nombre o con un nombre diferente, mediante el mapeo a otro valor, mediante el apuntamiento a un registro, por deducción, por defecto o por cualquier otro medio-. Un repositorio será conforme en la medida en que pueda proporcionar un buen valor de PREMIS cuando se requiera.

La Actividad de Mantenimiento de PREMIS podría reconsiderar en el futuro los requisitos de conformidad y hacerlos más restrictivos, pero de acuerdo con la versión 2.0 es bastante fácil que un repositorio sea conforme a PREMIS. Por otra parte, cuantas más unidades semánticas soporte un repositorio, mayor es el valor que añade si utiliza PREMIS. El Diccionario de datos PREMIS se desarrolló para identificar la información “esencial” que necesitarán la mayoría de los repositorios para preservar los contenidos digitales a largo plazo. Un repositorio de preservación responsable debe estudiar cuidadosamente PREMIS y tener un buen argumento para no ser capaz de implementar cualquier parte del Diccionario de datos.

6. MÁS INFORMACION

La web de la Actividad de Mantenimiento de PREMIS (www.loc.gov/standards/premis) ofrece algo para todos, incluyendo enlaces al Grupo de Implementadores de PREMIS (PIG), tutoriales de PREMIS, esquemas, herramientas y noticias. Incluye también la sección 'Recursos: artículos y presentaciones' que enlaza a la literatura sobre PREMIS y temas relacionados. Se relacionan aquí algunos de los artículos más útiles para los lectores en general:

Sobre metadatos de preservación en general:

"Preservation Metadata" (PDF:209K/21pp.)

Brian Lavoie (OCLC) and Richard Gartner (Oxford)

Joint report of OCLC, Oxford Library Services, and the Digital Preservation Coalition. Published as DPC Technology Watch Report No. 05-01: September 2005. <http://www.dpconline.org/docs/reports/dpctw05-01.pdf>

Sobre las diferencias entre PREMIS versión 1.0 y 2.0

"PREMIS with a Fresh Coat of Paint: Highlights from the revision of the PREMIS Data Dictionary for Preservation Metadata"

Brian Lavoie, *D-Lib Magazine*, May/June 2008

<http://www.dlib.org/dlib/may08/lavoie/05lavoie.html>

Sobre el uso conjunto de PREMIS y METS:

"Battle of the Buzzwords: Flexibility versus Interoperability when Implementing PREMIS with METS"

Rebecca Guenther, *D-Lib Magazine*, July/August 2008

<http://www.dlib.org/dlib/july08/guenther/07guenther.html>

Apéndice A: Lista de todas las unidades semánticas de PREMIS

Del *Diccionario de datos PREMIS de metadatos de preservación* versión 2.0

Cómo leer esta lista

1.2 objectCategory (categoría del objeto) (O, NR)

La unidad semántica objectCategory (categoría del objeto) es obligatoria y no repetible. Se aplica a todos los tipos de objetos (archivo, representación, cadena de bits).

1.3 preservationLevel (nivel de preservación) (O, R) [representación, archivo]

1.3.1 preservationLevelValue (O, NR) (valor del nivel de preservación) [representación, archivo]

La unidad semántica preservationLevel es opcional y repetible. Solo puede utilizarse para objetos representación y objetos archivo. Es una unidad contenedora porque hay al menos una unidad semántica subordinada a ella, preservationLevelValue (valor del nivel de preservación) que también se aplica únicamente a representaciones y archivos.

Unidades semánticas Objeto Entidad

1.1 objectIdentifier (identificador del objeto) (O, R)

1.1.1 objectIdentifierType (tipo del identificador del objeto) (O, NR)

1.1.2 objectIdentifierValue (Valor del identificador del objeto) (O, NR)

1.2 objectCategory (categoría del objeto) (O, NR)

1.3 preservationLevel (nivel de preservación) (NO, R) [representación, archivo]

1.3.1 preservationLevelValue (valor del nivel de preservación) (O, NR) [representación, archivo]

1.3.2 preservationLevelRole (función del nivel de preservación) (NO, NR) [representación, archivo]

1.3.3 preservationLevelRationale (fundamentos del nivel de preservación) (NO, R) [representación, archivo]

1.3.4 preservationLevelDateAssigned (fecha asignada al nivel de preservación) (NO, NR) [representación, archivo]

1.4 significantProperties (propiedades significativas) (NO, R)

1.4.1 significantPropertiesType (tipo de propiedades significativas) (NO, NR)

1.4.2 significantPropertiesValue (valor de las propiedades significativas) (NO, NR)

1.4.3 significantPropertiesExtension (extensión de las propiedades significativas) (NO, R)

1.5 objectCharacteristics (características del objeto) (O, R) (archivo, cadena de bits)

1.5.1 compositionLevel (nivel de composición) (O, NR) (archivo, cadena de bits)

1.5.2. fixity (fijeza) (NO, R) (archivo, cadena de bits)

1.5.2.1 messageDigestAlgorithm (algoritmo del mensaje cifrado) (O, NR) (archivo, cadena de bits)

1.5.2.2 messageDigest (mensaje cifrado) (O, NR) [archivo, cadena de bits]

1.5.2.3 messageDigestOriginator (creador del mensaje cifrado) (NO, NR) [archivo, cadena de bits]

- 1.5.3 size (tamaño) (NO, NR) [fichero, cadena de bits]
- 1.5.4 format (formato) (O, R) [fichero, cadena de bits]
 - 1.5.4.1 formatDesignation (designación del formato) (NO, NR) [fichero, cadena de bits]
 - 1.5.4.1.1 formatName (nombre del formato) (O, NR) [fichero, cadena de bits]
 - 1.5.4.1.2 formatVersion (versión del formato) (NO, NR) [fichero, cadena de bits]
 - 1.5.4.2 formatRegistry (registro del formato) (NO, NR) [fichero, cadena de bits]
 - 1.5.4.2.1 formatRegistryName (nombre del registro del formato) (O, NR) [fichero, cadena de bits]
 - 1.5.4.2.2 formatRegistryKey (clave del registro del formato) (O, NR) [fichero, cadena de bits]
 - 1.5.4.2.3 formatRegistryRole (función del registro del formato) (NO, NR) [fichero, cadena de bits]
 - 1.5.4.3 formatNote (nota sobre el formato) (NO, R) [fichero, cadena de bits]
- 1.5.5 creatingApplication (aplicación creadora) (NO, R) [fichero, cadena de bits]
 - 1.5.5.1 creatingApplicationName (nombre de la aplicación creadora) (NO, NR) [fichero, cadena de bits]
 - 1.5.5.2 creatingApplicationVersion (versión de la aplicación creadora) (NO, NR) [fichero, cadena de bits]
 - 1.5.5.3 dateCreatedByApplication (fecha creada por la aplicación) (NO, NR) [fichero, cadena de bits]
 - 1.5.5.4 creatingApplicationExtension (extensión de la aplicación creadora) (NO, R) [fichero, cadena de bits]
- 1.5.6 inhibitors (inhibidores) (NO, R) [fichero, cadena de bits]
 - 1.5.6.1 inhibitorType (tipo de inhibidor) (O, NR) [fichero, cadena de bits]
 - 1.5.6.2 inhibitorTarget (objetivo del inhibidor) (NO, R) [fichero, cadena de bits]
 - 1.5.6.3 inhibitorKey (clave del inhibidor) (NO, NR) [fichero, cadena de bits]
- 1.5.7 objectCharacteristicsExtension (extensión de las características del objeto) (NO, R) [fichero, cadena de bits]
- 1.6 originalName (nombre original) (NO, NR) [representación, fichero]
- 1.7 storage (almacenamiento) (O, R) [fichero, cadena de bits]
 - 1.7.1 contentLocation (localización del contenido) (NO, NR) [fichero, cadena de bits]
 - 1.7.1.1 contentLocationType (tipo de localización del contenido) (O, NR) [fichero, cadena de bits]
 - 1.7.1.2 contentLocationValue (valor de la localización del contenido) (O, NR) [fichero, cadena de bits]
 - 1.7.2 storageMedium (soporte del almacenamiento) (NO, NR) [fichero, cadena de bits]
- 1.8 environment (entorno) (NO, R)
 - 1.8.1 environmentCharacteristics (características del entorno) (NO, NR)
 - 1.8.2 environmentPurpose (propósito del entorno) (NO, R)
 - 1.8.3 environmentNote (nota sobre el entorno) (NO, R)
 - 1.8.4 dependency (dependencia) (NO, R)
 - 1.8.4.1 dependencyName (nombre de la dependencia) (NO, R)
 - 1.8.4.2 dependencyIdentifier (identificador de la dependencia) (NO, R)
 - 1.8.4.2.1 dependencyIdentifierType (tipo de identificador de dependencia) (O, NR)

- 1. 8.4.2.2 dependencyIdentifierValue (valor del identificador de dependencia (O, NR)
- 1.8.5 software (NO, R)
 - 1.8.5.1 swName (nombre del software) (O, NR)
 - 1.8.5.2 swVersion (versión del software) (NO, NR)
 - 1.8.5.3 swType (tipo de software) (O, NR)
 - 1.8.5.4 swOtherInformation (otra información sobre el software) (NO, R)
 - 1.8.5.5. swDependency (dependencia del software) (NO, R)
- 1.8.6 hardware (NO, R)
 - 1.8.6.1 hwName (nombre del hardware) (O, NR)
 - 1.8.6.2 hwType (tipo de hardware) (O, NR)
 - 1.8.6.3 hwOtherInformation (otra información sobre el hardware) (NO, R)
- 1.8.7 environmentExtension (extensión del entorno) (NO, R)
- 1.9 signatureInformation (información sobre la firma) (NO, R) [fichero, cadena de bits]
 - 1.9.1 signature (firma) (NO, R)
 - 1.9.1.1 signatureEncoding (codificación de la firma) (O, NR) [fichero, cadena de bits]
 - 1.9.1.2 signer (firmante) (NO, NR) [fichero, cadena de bits]
 - 1.9.1.3 signatureMethod (método de la firma) (O, NR) [fichero, cadena de bits]
 - 1.9.1.4 signatureValue (valor de la firma) (O, NR) [fichero, cadena de bits]
 - 1.9.1.5 signatureValidationRules (reglas de validación de la firma) (O, NR) [fichero, cadena de bits]
 - 1.9.1.6 signatureProperties (propiedades de la firma) (NO, R) [fichero, cadena de bits]
 - 1.9.1.7 keyInformation (información sobre la clave) (NO, NR) [fichero, cadena de bits]
 - 1.9.2 signatureInformationExtension (extensión de la información sobre la signatura) (NO, R) [fichero, cadena de bits]
- 1.10 relationship (relaciones) (NO, R)
 - 1.10.1 relationshipType (tipo de relaciones) (O, NR)
 - 1.10.2 relationshipSubType (subtipo de relaciones) (O, NR)
 - 1.10.3 relatedObjectIdentification (identificación del objeto relacionado) (O, R)
 - 1.10.3.1 relatedObjectIdentifierType (tipo de identificador del objeto relacionado) (O, NR)
 - 1.10.3.2 relatedObjectIdentifierValue (valor del identificador del objeto relacionado) (O, NR)
 - 1.10.3.3 relatedObjectSequence (secuencia del objeto relacionado) (NO, NR)
 - 1.10.4 relatedEventIdentification (identificación del acontecimiento relacionado) (NO, R)
 - 1.10.4.1 relatedEventIdentifierType (tipo de identificador del acontecimiento relacionado) (O, NR)
 - 1.10.4.2 relatedEventIdentifierValue (valor del identificador del acontecimiento relacionado) (O, NR)
 - 1.10.4.3 relatedEventSequence (secuencia del acontecimiento relacionado) (NO, NR)
- 1.11 linkingEventIdentifier (identificador del acontecimiento vinculado) (NO, R)
 - 1.11.1 linkingEventIdentifierType (tipo de identificador del acontecimiento vinculado) (O, NR)
 - 1.11.2 linkingEventIdentifierValue (valor del identificador del acontecimiento vinculado) (O, NR)

- 1.12 linkingIntellectualEntityIdentifier (identificador de la entidad intelectual vinculada) (NO, R)
 - 1.12.1 linkingIntellectualEntityIdentifierType (tipo de identificador de la entidad intelectual vinculada) (O, NR)
 - 1.12.2 linkingIntellectualEntityIdentifierValue (valor del identificador de la entidad intelectual vinculada) (O, NR)
- 1.13 linkingRightsStatementIdentifier (identificador de la mención de derechos vinculada) (NO, R)
 - 1.13.1 linkingRightsStatementIdentifierType (tipo de identificador de la mención de derechos vinculada) (O, NR)
 - 1.13.2 linkingRightsStatementIdentifierValue (valor del identificador de la mención de derechos vinculada) (O, NR)

Unidades semánticas de la Entidad Acontecimiento

- 2.1 eventIdentifier (identificador del acontecimiento) (O, NR)
 - 2.1.1 eventIdentifierType (tipo de identificador del acontecimiento) (O, NR)
 - 2.1.2 eventIdentifierValue (valor del identificador del acontecimiento) (O, NR)
- 2.2. eventType (tipo de acontecimiento) (O, NR)
- 2.3 eventDateTime (tiempo de la fecha del acontecimiento) (O, NR)
- 2.4 eventDetail (detalle del acontecimiento) (NO, NR)
- 2.5 eventOutcomeInformation (información sobre las consecuencias del acontecimiento) (NO, R)
 - 2.5.1 eventOutcome (consecuencias del acontecimiento) (NO, NR)
 - 2.5.2 eventOutcomeDetail (detalles sobre las consecuencias del acontecimiento) (NO, R)
 - 2.5.2.1 eventOutcomeDetailNote (nota sobre los detalles sobre las consecuencias del acontecimiento) (NO, NR)
 - 2.5.2.2 eventOutcomeDetailExtension (Extensión de los detalles sobre las consecuencias del acontecimiento) (NO, R)
- 2.6 linkingAgentIdentifier (identificador del agente vinculado) (NO, R)
 - 2.6.1 linkingAgentIdentifierType (tipo de identificador del agente vinculado) (O, NR)
 - 2.6.2 linkingAgentIdentifierValue (valor del identificador del agente vinculado) (O, NR)
 - 2.6.3 linkingAgentRole (función del agente vinculado) (NO, R)
- 2.7 linkingObjectIdentifier (identificador del objeto vinculado) (NO, R)
 - 2.7.1 linkingObjectIdentifierType (tipo de identificador del objeto vinculado) (O, NR)
 - 2.7.2 linkingObjectIdentifierValue (valor del identificador del objeto vinculado) (O, NR)
 - 2.7.3 linkingObjectRole (función del objeto vinculado) (NO, R)

Unidades semánticas de la Entidad Agente

- 3.1 agentIdentifier (identificador del agente) (O, R)
 - 3.1.1 agentIdentifierType (tipo de identificador del agente)
 - 3.1.2 agentIdentifierValue (valor del identificador del agente)
- 3.2 agentName (nombre del agente) (NO, R)
- 3.3 agentType (tipo de agente) (NO, NR)

Unidades semánticas de la Entidad Derechos

- 4.1 rightsStatement (mención de derechos) (NO, R)
 - 4.1.1 rightsStatementIdentifier (identificador de la mención de derechos) (O, NR)
 - 4.1.1.1 rightsStatementIdentifierType (tipo de identificador de la mención de derechos) (O, NR)
 - 4.1.1.2 rightsStatementIdentifierValue (valor del identificador de la mención de derechos) (O, NR)
 - 4.1.2 rightsBasis (bases de los derechos) (O, NR)
 - 4.1.3 copyrightInformation (información sobre el copyright) (NO, NR)
 - 4.1.3.1 copyrightStatus (estado del copyright) (O, NR)
 - 4.1.3.2 copyrightJurisdiction (jurisdicción del copyright) (O, NR)
 - 4.1.3.3 copyrightStatusDeterminationDate (determinación de la fecha del estado del copyright) (NO, NR)
 - 4.1.3.4 copyrightNote (nota sobre el copyright) (NO, R)
 - 4.1.4 licenseInformation (información sobre la licencia) (NO, NR)
 - 4.1.4.1 licenseIdentifier (identificador de la licencia) (NO, NR)
 - 4.1.4.1.1 licenseIdentifierType (tipo del identificador de la licencia) (O, NR)
 - 4.1.4.1.2 licenseIdentifierValue (valor del identificador de la licencia) (O, NR)
 - 4.1.4.2 licenseTerms (términos de la licencia) (NO, NR)
 - 4.1.4.3 licenseNote (nota sobre la licencia) (NO, R)
 - 4.1.5 statuteInformation (información sobre la legislación) (NO, R)
 - 4.1.5.1 statuteJurisdiction (jurisdicción de la legislación) (O, NR)
 - 4.1.5.2 statuteCitation (cita de la legislación) (O, NR)
 - 4.1.5.3 statuteInformationDeterminationDate (fecha de determinación de la información sobre la legislación) (NO, NR)
 - 4.1.5.4 statuteNote (nota sobre la legislación) (NO, R)
 - 4.1.6 rightsGranted (derechos otorgados) (NO, R)
 - 4.1.6.1 act (ley) (O, NR)
 - 4.1.6.2 restriction (restricción) (NO, R)
 - 4.1.6.3 termOfGrant (período por el que se otorgan los derechos) (O, NR)
 - 4.1.6.3.1 startDate (fecha de inicio) (O, NR)
 - 4.1.6.3.2 endDate (fecha final) (NO, NR)
 - 4.1.6.4 rightsGrantedNote (nota sobre los derechos otorgados) (NO, R)
 - 4.1.7 linkingObjectIdentifier (identificador del objeto vinculado) (NO, R)
 - 4.1.7.1 linkingObjectIdentifierType (tipo de identificador del objeto vinculado) (O, NR)
 - 4.1.7.2 linkingObjectIdentifierValue (valor del identificador del objeto vinculado) (O, NR)
 - 4.1.8 linkingAgentIdentifier (identificador del agente vinculado) (NO, R)
 - 4.1.8.1 linkingAgentIdentifierType (tipo de identificador del agente vinculado) (O, NR)
 - 4.1.8.2 linkingAgentIdentifierValue (valor del identificador del agente vinculado) (O, NR)
 - 4.1.8.3 linkingAgentRole (función del agente vinculado) (O, NR)
- 4.2 rightsExtension (extensión de los derechos) (NO, R)

Apéndice B: ejemplo de objeto

Este ejemplo es una versión simplificada de la que preparó la Library of Congress para un tutorial de PREMIS. Muestra las unidades semánticas de PREMIS y los valores utilizados para describir una imagen TIFF.

La LC utiliza handles (identificadores creados y gestionados por una aplicación denominada Handle System) para los objetos fichero almacenados en el repositorio (véase 1.1. `objectIdentifier` y 1.10.3 `relatedObjectIdentification`). El repositorio entiende el objeto que se describe mediante su identificador único (1.1 `objectIdentifier`), pero su nombre original antes de la ingesta era "001h.tif" (1.6 `originalName`). Sabemos que es un fichero por la categoría del objeto (1.2 `objectCategory`).

El fichero tiene una suma de verificación MD5 computada por el sistema de gestión de contenido digital local de la LC (1.5.2 `fixity`). La información sobre su formato se proporciona de dos maneras: como tipo y versión MIME (1.5.4.1 `formatDesignation`) y apuntando al registro de formatos PRONOM que alberga una copia de la especificación TIFF.

La LC considera que hay dos aplicaciones creadoras. Algo (probablemente un fichero RAW no nombrado) se creó en el escáner al utilizar la aplicación ScandAll y se convirtió en TIFF utilizando Adobe Photoshop (1.5.5 `creatingApplication`). No hay encriptación ni otro tipo de inhibidores lo que la LC considera que merece la pena registrar explícitamente (1.5.6 `inhibidores`).

El fichero se almacena en disco en una localización identificada por la ruta del directorio (1.7.1 `contentLocation`); el dispositivo del disco se identifica muy específicamente, presumiblemente de manera que los gestores del repositorio entiendan en qué momento los datos requieren una migración a un dispositivo más actualizado (1.7.2 `storageMedium`).

La LC se ha preocupado de registrar el entorno recomendado para reproducir el archivo: utilizando la versión 5.0 de Adobe Acrobat en un ordenador Intel x86 que corre bajo Windows XP. La velocidad del procesador debe ser de al menos 60 MHz y debe tener al menos 32 Mb de memoria, y se recomiendan 64 Mb (1.8 `environment`).

El fichero TIFF forma parte de una representación con el identificador *R200802948* asignado internamente. Esta información se proporciona en el tercer bloque de información relacionada (1.10 `relationship`). En esta representación hay otros dos archivos identificados por handles. Sabemos que son objetos fichero porque el tipo de relación se describe como "hermano". Para obtener más información sobre los mismos, como los tipos de ficheros de que se trata, debe verse la descripción del objeto en *loc.music/gottlieb.09602* y *loc.music/gottlieb.mets*.

No disponemos de ninguna información descriptiva sobre la entidad intelectual representada por esta representación. Sin embargo, existe un enlace a su identificador, un LCCN (1.12 `linkingIntellectualEntityIdentifier`). Presumiblemente, si miramos el LCCN encontraremos la descripción bibliográfica de la entidad intelectual (aunque no en este ejemplo porque existe el LCCN).

Unidad semántica	Valor
1.1 objectIdentifier (identificador del objeto)	
1.1.1 objectIdentifierType (tipo de identificador del objeto)	hdl
1.1.2 objectIdentifierValue (valor del identificador del objeto)	loc.music/gottlieb.09601
1.2 objectCategory (categoría del objeto)	file
1.3 preservationLevel (nivel de preservación)	
1.3.1 preservationLevelValue (valor del nivel de preservación)	full
1.5 objectCharacteristics (características del objeto)	
1.5.1 compositionLevel (nivel de composición)	0
1.5.2 fixity (fijeza)	
1.5.2.1 messageDigestAlgorithm (algoritmo del mensaje cifrado)	MD5
1.5.2.2 messageDigest (mensaje cifrado)	36b03197ad066cd719906c55eb68ab8d
1.5.2.3 messageDigestOriginator (creador del mensaje cifrado)	LocalDCMS
1.5.3 size (tamaño)	20800896
1.5.4 format (formato)	
1.5.4.1 formatDesignation (designación del formato)	
1.5.4.1.1 formatName (nombre del formato)	image/tiff
1.5.4.1.2 formatVersion (versión del formato)	6.0
1.5.4.2 formatRegistry (registro de formatos)	
1.5.4.2.1 formatRegistryName (nombre del registro de formatos)	PRONOM
1.5.4.2.2 formatRegistryKey (clave del registro de formatos)	fmt/10
1.5.4.2.3 formatRegistryRole (función del registro de formatos)	especificación
1.5.5 creatingApplication (aplicación creadora)	
1.5.5.1 creatingApplicationName (nombre de la aplicación creadora)	ScandAll 21
1.5.5.2 creatingApplicationVersion (versión de la aplicación creadora)	4.1.4

1.5.5.3 dateCreatedByApplication (fecha creada por la aplicación)	1998-10-30
1.5.5 creatingApplication (aplicación creadora)	
1.5.5.1 creatingApplicationName (nombre de la aplicación creadora)	Adobe Photoshop
1.5.5.2 creatingApplicationVersion (versión de la aplicación creadora)	CS2
1.5.5.3 dateCreatedByApplication (fecha creada por la aplicación)	1998-10-30T08:29:02
1.5.6 inhibitors (inhibidores)	
1.5.6.1 inhibitorType (tipo de inhibidor)	ninguno
1.6 originalName (nombre original)	001h.tif
1.7 storage (almacenamiento)	
1.7.1 contentLocation (localización del contenido)	
1.7.1.1 contentLocationType (tipo de localización del contenido)	Ruta del fichero
1.7.1.2 contentLocationValue (valor de la localización del contenido)	amserver/
1.7.2 storageMedium (soporte de almacenamiento)	IBM DS4000 System 1740-5208
1.8 environment (entorno)	
1.8.1 environmentCharacteristic (características del entorno)	recomendado
1.8.2 environmentPurpose (propósito del entorno)	recuperar
1.8.2 environmentPurpose (propósito del entorno)	editar
1.8.5 software	
1.8.5.1 swName (nombre del software)	Windows
1.8.5.2 swVersion (versión del software)	XP
1.8.5.3 swType (tipo de software)	sistema operativo
1.8.6 hardware	
1.8.6.1 hwName (nombre del hardware)	Intel x86
1.8.6.2 hwType (tipo de hardware)	procesador
1.8.6.3 hwOtherInformation (otra información sobre el hardware)	mínimo 60 Mhz
1.8.6 hardware	
1.8.6.1 hwName (nombre del hardware)	64MB RAM

1.8.6.2 hwType (tipo de hardware)	memoria
1.8.6.3 hwOtherInformation (otra información sobre el hardware)	mínimo 32 MB
1.10 relationship (relaciones)	
1.10.1 relationshipType (tipo de relaciones)	estructural
1.10.2 relationshipSubType (subtipo de relaciones)	tiene "hermano"
1.10.3 relatedObjectIdentification (identificación del objeto relacionado)	
1.10.3.1 relatedObjectIdentifierType (tipo del identificador del objeto relacionado)	hdl
1.10.3.2 relatedObjectIdentifierValue (valor del identificador del objeto relacionado)	loc.music/gottlieb.09602
1.10.3.3 relatedObjectSequence (secuencia del objeto relacionado)	0
1.10 relationship (relaciones)	
1.10.1 relationshipType (tipo de relaciones)	estructural
1.10.2 relationshipSubType (subtipo de relaciones)	está incluido en:
1.10.3 relatedObjectIdentification (identificación del objeto relacionado)	
1.10.3.1 relatedObjectIdentifierType (tipo del identificador del objeto relacionado)	repositorio local
1.10.3.2 relatedObjectIdentifierValue (valor del identificador del objeto relacionado)	R200802948
1.10.3.3 relatedObjectSequence (secuencia del objeto relacionado)	0
1.12. linkingIntellectualEntityIdentifier (identificador de la entidad intelectual relacionada)	
1.12.1 linkingIntellectualEntityIdentifierType (tipo de identificador de la entidad intelectual relacionada)	LCCN
1.12.2 linkingIntellectualEntityIdentifierValue (valor de identificador de la entidad intelectual relacionada)	2007-86121

Apéndice C: Glosario de términos

Este glosario reúne definiciones que aparecían anteriormente en el texto de esta guía. Las definiciones pueden ser menos formales que las que aparecen en el *Diccionario de Datos PREMIS de Metadatos de Preservación*.

accionable: la calidad de registrarse de una forma controlada que puede activarse por medio de un programa de ordenador

agente: persona, organización o programa de ordenador con una función relacionada con un *acontecimiento* o una mención de *derechos*.

contenedor de extensión: un tipo especial de *unidad contenedora* de PREMIS que no tiene definidas subunidades dependientes de sí misma sino que se ha diseñado como lugar de alojamiento de los metadatos no-PREMIS

emulación: *estrategia de preservación* que implica la reproducción de un antiguo entorno de reproducción en un hardware y/o software más nuevo

entidad: en PREMIS, un tipo de cosa de la que se puede hablar. Los tipos de entidades de PREMIS son: *entidades intelectuales, objetos, agentes, acontecimientos y derechos*.

entidad acontecimiento: la entidad de PREMIS que agrega información sobre las acciones que afectan a los objetos del repositorio.

entidad derechos: *entidad* de PREMIS que agrega información sobre los derechos y permisos pertenecientes a los objetos de un repositorio digital.

entidad intelectual: conjunto del contenido que se trata como una unidad para propósitos de gestión y descripción; es similar a la “entidad bibliográfica” en biblioteconomía.

entorno: hardware, software y otros objetos necesarios para reproducir un objeto.

estrategias de preservación: técnicas utilizadas para garantizar que los recursos digitales se puedan utilizar a lo largo del tiempo; dos estrategias comunes son la *emulación* y la *migración*.

inhibidores: las características de un objeto digital cuyo objetivo es restringir el acceso, uso o *migración*

metadatos de preservación: los metadatos que soportan las actividades cuyo objetivo es garantizar la usabilidad a largo plazo de un recursos digital.

migración: estrategia de preservación que implicar elaborar una versión de un archivo digital en un formato de fichero más nuevo.

objeto cadena de bits: un tipo de *objeto* de PREMIS; datos de un fichero que tienen propiedades comunes en relación con los objetivos de preservación y que no son autónomos

objeto fichero: un tipo de *objeto* de PREMIS; un archivo de ordenador, como un PDF o un JPEG.

objetos: ítems digitales que se almacenan y gestionan en un repositorio de preservación digital. PREMIS define tres tipos de objetos: *ficheros*, *cadena de bits* y *representaciones*.

procedencia digital: documentación relativa a la cadena de custodia e historia de los cambios de un recurso digital

propiedades significativas: características de un objeto que deben mantenerse a través de las acciones de preservación.

representación: un tipo de *objeto* de PREMIS; conjunto de todos los *objetos fichero* necesarios para reproducir una *entidad intelectual*.

unidad semántica: parte de información o conocimiento

unidades contenedoras: *Unidades semánticas* que no tienen valor por si mismas pero que permiten agrupar subunidades relacionadas.