

MÓDULO 7: Gestión y preservación de documentos digitales

SECCIÓN 9

METADATOS

MÓDULO 7

Gestión y preservación de documentos digitales

SECCIÓN 9

Metadatos

Adaptación del Archivo Nacional de Costa Rica

Versión 1, 2024

Este curso fue traducido y adaptado por la Dirección General del Archivo Nacional de Costa Rica en colaboración con la Sección de Archivística de la Universidad de Costa Rica a partir del material original del año 2011 de la Asociación Internacional de Archivos Francófonos disponible en línea en el Portal Internacional Archivístico Francófono. Se aclara que pueden existir variaciones respecto al contenido original. Para acceder al material en francés, visite <https://www.piaf-archives.org/se-former/module-7-gestion-et-archivage-des-documents-numeriques>.



ARCHIVO NACIONAL
COSTA RICA



UNIVERSIDAD DE
COSTA RICA

Contenido

Capítulo 1. Introducción	5
1.1. Objetivo de la sección	5
1.2. Primeros elementos de los metadatos.....	5
Capítulo 2. El modelo OAIS y los metadatos	6
2.1. El modelo OAIS y los metadatos.	6
Capítulo 3. Una tipología de metadatos	7
3.1. Clasificación de los metadatos.....	8
3.2. Ejemplos de formatos de metadatos.....	9
3.3. ¿Cómo elegir un formato de metadatos?.....	9
Capítulo 4. Los metadatos descriptivos.....	10
4.1. Definición.....	10
4.2. Un ejemplo de descripción archivística: el estándar EAD	10
Capítulo 5. Los metadatos técnicos.....	13
5.1. Definiciones.....	13
5.2. Información que debe de recopilarse sobre un formato	15

Capítulo 6. Los metadatos administrativos	16
6.1. Los metadatos de identificación	16
6.2. Los metadatos de procedencia y contexto	17
6.3. Los metadatos de integridad	17
6.4. Los metadatos de derecho	17
6.5. Los metadatos orientados a la gestión de documentos.....	18
Capítulo 7. Los metadatos estructurales.....	19
7.1. Definición y ejemplos de formatos de metadatos de estructura.....	19
Capítulo 8. Caso de Estudio: España- Esquema de metadatos para la gestión del documento electrónico.....	22
8.1. Objetivos y descripción general	22
8.2. Esquema e-EMGDE.....	23
8.2.1. Modelo entidad- relación	23
8.2.2. Características.....	25
8.2.3. Estructura del Esquema	25
8.2.4. Metadatos mínimos obligatorios.....	30
Bibliografía	34

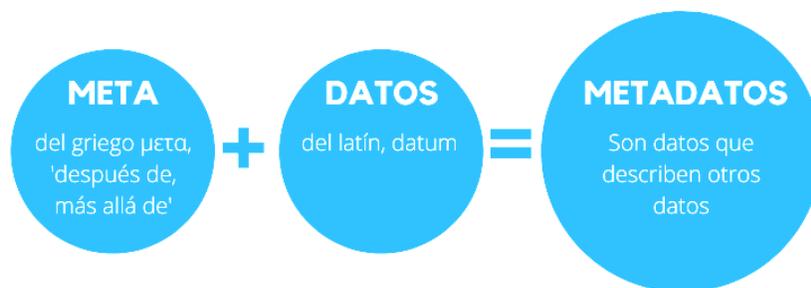
Capítulo 1. Introducción

1.1. Objetivo de la sección

En este apartado se definirán los metadatos; además se verá su papel dentro del modelo OAIS; finalmente se mencionarán todos los diferentes tipos de metadatos que existen, englobándolos en cuatro grandes familias: metadatos descriptivos, metadatos técnicos, metadatos estructurales y metadatos administrativos.

1.2. Primeros elementos de los metadatos

Los metadatos son un tema amplio que guarda una estrecha relación con el papel que desempeña en el proceso de preservación. Intentemos dar una definición: etimológicamente “meta” proviene del griego que significa “después, más allá de, con”, “más allá de los datos”, “que incluye datos”.



Metadatos. Elaboración propia

Por lo anterior, se puede indicar que los metadatos son: “datos que describen datos”, o “datos sobre datos”. Es decir, se trata de datos sobre los datos, en propósito de los datos, que definen, describen datos. Este término es reciente, sin embargo, siempre han existido los metadatos. Según la actividad, se llama catalogar, indexar, clasificar, describir, elaborar un instrumento de investigación, ya sea bibliotecario, documentalista, archivista, científico. Se trata, a partir de los archivos procedentes de un productor, de poder comunicar estos archivos mucho más, a nuevas comunidades de usuarios para las que los metadatos «empresariales» elaborados por los productores deben ser enriquecidos y explicitados para comunidades más amplias.

Los metadatos son una herramienta de ayuda para la búsqueda. Con los datos digitales apareció la necesidad de contar los datos que proporcionen información objetiva sobre las formas de emplear y de manipular los datos (metadatos técnicos). Asimismo, los metadatos relativos a derechos de acceso se convertirán en más complejos de gestionar, en un contexto de difusión en las redes de internet.

Los metadatos son datos, con los mismos desafíos de producción, gestión y conservación que los datos, por supuesto.



COMPLEMENTO

Video: “Microaprendizaje: ¿Qué son los metadatos?”

Fuente: <https://www.youtube.com/watch?v=W9df022FHIE&t=33s>



GLOSARIO

Esquema de metadatos: Instrumento que define la incorporación y gestión de los metadatos de contenido, contexto y estructura de los documentos electrónicos a lo largo de su ciclo de vida. (Archivo Nacional de Costa Rica, 2020, p.3)

Esquema de codificación de metadatos: vocabulario controlado para valores de elementos de metadatos así como sus estructuras de codificación. Proporciona información contextual o reglas de análisis útiles en la interpretación del valor de un término. Esa información contextual puede tomar la forma de vocabularios controlados, notaciones formales o reglas de análisis. Existen dos tipos de esquemas de codificación: esquemas de codificación de vocabularios (en donde el valor del elemento se toma de un vocabulario controlado) y esquemas de codificación de sintaxis (en donde el valor de una cadena de información se formatea de acuerdo con una notación formal). (Barnard, A y Voutssas, J, 2014, p. 105).

Metadato: Dato que define y describe otros datos. Existen diferentes tipos de metadatos según su aplicación. (Archivo Nacional de Costa Rica, 2020, p.5)

Capítulo 2. El modelo OAIS y los metadatos

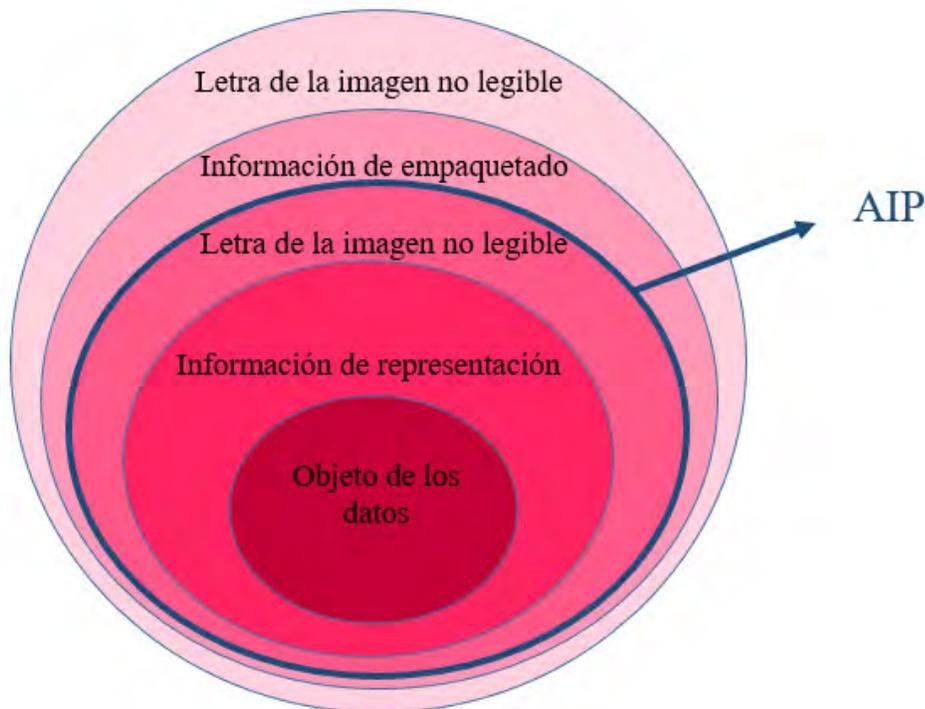
2.1. El modelo OAIS y los metadatos.

El modelo OAIS, es un modelo conceptual destinado a la gestión, el archivo y la preservación a largo plazo de los documentos. Por lo anterior, este modelo muestra claramente que debemos de conservar una cantidad importante de información para garantizar la mejor conservación de la información registrada en los objetos de datos.

El modelo OAIS define también la estructura de los paquetes de información necesarios para la gestión de datos de acuerdo con su posición en el ciclo de vida digital. Estos paquetes son: Paquete de Información de Transferencia (SIP, Submission Information Package), Paquete de Información de Consulta (DIP, Dissemination Information Package) y Paquete de Información de Archivo (AIP, Archival Information Package).

Por lo anterior, el AIP contiene:

- La información de preservación (procedencia, contexto, identificación, integridad, derechos)
- El objeto información, propiamente contiene la información a perpetuidad, en forma de datos de información de representación correspondiente, que permita interpretar los datos en forma de información comprensible.



Los metadatos en el modelo OAIS. Elaboración propia a partir de PIAF

Leyenda: esta imagen representa de manera simbólica la suma de informaciones complementarias que deben añadirse al objeto digital, propiamente el objeto "datos" con el fin de asegurar su preservación y la reutilización a futuro.

Capítulo 3. Una tipología de metadatos

Los calificativos para caracterizar los metadatos son numerosos: descriptivos, administrativos, contextuales, estructurales, técnicos, de preservación, estáticos, evolutivos, externos de contenidos, entre otros.

Algunos de estos datos sobre los datos se utilizan para organizar el conocimiento y otros para mantenerlos accesibles en el tiempo y para garantizar y controlar su acceso.

Se pretende clasificar los metadatos necesarios para la preservación, a partir del modelo OAIS, como sea posible.

3.1. Clasificación de los metadatos

Podemos clasificar los diferentes tipos de metadatos en función de:

- Lo que describen (su contenido)
- La manera en la que se crearon (su proveniencia)
- Momento en el que se crean (su historia)
- Ambiente donde se les encuentra (su localización)
- Aspecto que tienen (su forma)
- Uso que se les da (su objetivo)

Estas categorías no están separadas.

La gestión de un objeto digital complejo implica:

- Una gestión compleja del documento (uno o más documentos, descripciones multinivel)
- Una gestión compleja de los aspectos técnicos (si existen diferentes formatos de archivo, por ejemplo)
- Una gestión compleja de aspectos administrativos (derechos, seguimientos de la evolución del documento, seguimiento de la evolución de los metadatos)
- Una gestión compleja de su estructura física (organización de ficheros entre ellos)
- Una gestión compleja de su estructura lógica (organización de partes del contenido entre ellas)

Este es el papel de los metadatos.



COMPLEMENTO: LO QUE PERMITE DEFINIR UN FORMATO DE METADATOS

En práctica, la definición de un formato de metadatos incluye generalmente los siguientes elementos:

- Un diccionario, que dé cuenta de la semántica del esquema de metadatos. Por ejemplo, Premis https://www.loc.gov/standards/premis/PREMIS_es.pdf
- Un tesoro, instrumento de dominio del vocabulario: ¿Qué sentido preciso dar a cada término?
- Un esquema que traduce la estructuración de conceptos aplicables a el objeto digital. Este esquema se especifica más a menudo en forma de DTD, esquema XML o modelización UML.

3.2. Ejemplos de formatos de metadatos

En los últimos años hemos sido testigos del nacimiento de un gran número de definiciones de metadatos en la iniciativa de organizaciones internacionales, universidades, bibliotecas, cada una tratando de responder a un problema general o específico de su comunidad.

En el campo bibliográfico podemos enumerar:

- **Dublín Core** es un modelo de metadatos elaborado y auspiciado por la DCMI. Este modelo tiene como objetivo definir un conjunto básico de atributos que sirvan para describir todos los recursos existentes en la red (internet). Muchos expertos mencionan que este formato ayuda a los motores de búsqueda en la recuperación global de información en la red. Este formato puede ser enriquecido y especializado en función de los oficios y dominios.
- **Marc 21** es un formato de metadatos para la catalogación, muy utilizado por EAD (Encoded Archival Description). Se utiliza la descripción de los documentos de archivo, TEF. Versión 2.0. Su objetivo es organizar de manera coherente metadatos de tesis ricas y normalizadas, para facilitar su intercambio y difusión, a nivel nacional o internacional. Estos metadatos son heterogéneos. Abarcan metadatos descriptivos (bibliográficos) y metadatos de gestión (administración, derechos, conservación).
- **MODS (Metadata Object Description Schema)** es un esquema de descripción bibliográfica basado en XML; el MODS posee una complejidad intermedia, ya que está más orientado a bibliotecas, por lo cual es más complejo que Dublín Core, sin embargo, es más sencillo que Marc 21.
- **TEI (Text Encoding Initiative)** se ha desarrollado como un estándar internacional e interdisciplinario; debido a que fue creado para representar los materiales textuales en formato electrónico de forma tal que permita a los usuarios, principalmente investigadores, intercambiar y reutilizar los recursos sin importar el hardware y software que utilicen.

3.3. ¿Cómo elegir un formato de metadatos?

Para elegir adecuadamente el formato de metadatos a utilizar, se debe realizar primeramente dos verificaciones. En primer lugar, la reflexión sobre el tema está madurando, por lo cual los formatos están normalizados y utilizados por la comunidad; debido a lo anterior en esta sección se mencionarán los más importantes.

En segundo lugar, el lenguaje XML se ha convertido en la norma para formalizar los metadatos y garantizar que los metadatos sean legibles y comprensibles tanto por las personas como por los ordenadores.

En cualquier proyecto de archivo digital, habrá que examinar el contexto, las diferentes limitaciones, las prácticas y necesidades de la comunidad de usuarios antes de elegir un formato de metadatos y un perfil de aplicación para este formato.

Capítulo 4. Los metadatos descriptivos

4.1. Definición

Los **metadatos descriptivos** son los metadatos que sirven para organizar el conocimiento; estos permitirán identificar, clasificar, jerarquizar la información contenida en el objeto digital. Por esto, se trata típicamente de un título o de un nombre, de autores, de fechas, de términos que permiten la clasificación.

Los metadatos descriptivos se denominan Información descriptiva en el modelo OAIS. Se trata de una información que se construye a partir de los metadatos presentes en el Paquete de Información de Archivo (AIP) a una comunidad de usuarios específica. Las comunidades y sus áreas de actividad pueden ser muy variadas; por esto la naturaleza de la Información descriptiva está extremadamente relacionada con las especificidades del público al que se dirige.

Para la descripción de los documentos de archivo, la EAD (Descripción Archivística Codificada) es un formato de metadatos descriptivos.

4.2. Un ejemplo de descripción archivística: el estándar EAD

Como ya se mencionó, El estándar EAD, que significa Encoded Archival Description es un conjunto de reglas y formatos que se utiliza para describir archivos y colecciones de archivos de una manera estructurada y coherente.

EAD utiliza etiquetas y elementos predefinidos para describir atributos como el título del documento, el autor, la fecha, el asunto y otros detalles importantes. Estas descripciones siguen un formato que puede ser entendido tanto por humanos como por computadoras (bases de datos).

Por ejemplo, en el Archivo Histórico del Archivo Nacional de Costa Rica, se cuenta con una gran parte del fondo documental que se custodia, descrito según los campos de la Norma ISAD G, esto en la plataforma de Atom, por lo que la misma herramienta permite realizar la descarga de la descripción en un XML de EAD.

De manera que este es un ejemplo de la descripción en EAD:

```
<ead>
  <eadheader langencoding="iso639-2b" countryencoding="iso3166-1" dateencoding="iso8601" repositoryencoding="iso15511" scriptencoding="iso15924"
    relatedencoding="DC">
    <eadid identifier="reportaje-sobre-el-libro-de-jose-maria-figueroa" countrycode="CR" mainagencycode="AN"
      url="https://www.archivodigital.go.cr/index.php/reportaje-sobre-el-libro-de-jose-maria-figueroa" encodinganalog="identifier">CR-AN-AH-JMFIGOR-
      ALBUM-000001-001-1-067f</eadid>
    <filedesc>
      <titlestmt>
        <titleproper encodinganalog="title">Reportaje sobre el libro de José María Figueroa</titleproper>
      </titlestmt>
      <publicationstmt>
        <publisher encodinganalog="publisher">Archivo Nacional</publisher>
        <address>
          <addressline>
            Curridabat, San José. Del Centro Comercial Plaza del Sol, 900 m sur y 150 m oeste
          </addressline>
          <addressline>San José</addressline>
          <addressline>Costa Rica</addressline>
          <addressline>Teléfono:2283-1400</addressline>
          <addressline>Correo electrónico:consultasala@dgan.go.cr</addressline>
          <addressline>https://www.archivonacional.go.cr/</addressline>
        </address>
        <date normal="2020-12-15" encodinganalog="date">2020-12-15</date>
      </publicationstmt>
    </filedesc>
    <profiledesc>
      <creation>
        Generado por Access to Memory (AtoM) 2.7.2\n
        <date normal="2024-07-16">2024-07-16 14:06 UTC</date>
      </creation>
      <language>
        <language langcode="spa">español</language>
      </language>
    </profiledesc>
  </eadheader>
```

```

<archdesc level="item" relatedencoding="ISAD(G)v2">
  <did>
    <unittitle encodinganalog="3.1.2">Reportaje sobre el libro de José María Figueroa</unittitle>
    <unitid encodinganalog="3.1.1" countrycode="CR" repositorycode="AN">CR-AN-AH-JMFIGOR-ALBUM-000001-001-1-067f</unitid>
    <unitdate normal="1896-10-16/1896-10-16" encodinganalog="3.1.3">1896-10-16</unitdate>
    <physdesc encodinganalog="3.1.5"> Papel;Digital - 1 unidad </physdesc>
  </repository>
  <corpname>Archivo Nacional</corpname>
  <address>
    <addressline>
      Curridabat, San José. Del Centro Comercial Plaza del Sol, 900 m sur y 150 m oeste
    </addressline>
    <addressline>
      https://www.archivonacional.go.cr
    </addressline>
    <addressline>San José</addressline>
    <addressline>Costa Rica</addressline>
    <addressline>Teléfono:2283-1400</addressline>
    <addressline>Correo electrónico:consultasala@dgan.go.cr</addressline>
    <addressline>https://www.archivonacional.go.cr/</addressline>
  </address>
  </repository>
  <note type="generalNote">
    <p>Notas: Se advierte que la fecha indicada es aproximada</p>
  </note>
  <dao linktype="simple" href="https://archivodigital.go.cr/uploads/r/archivonacional/d/8/f/d8f3bd76ac28be64bbc0ea6fcc181f55ac2c5cda92a71d755363ee0e5177e130/000001-001-1-067f.jpg" role="master" actuate="onrequest" show="embed"/>
  <origination encodinganalog="3.2.1">
    <persname id="atom_4057510_actor">José María Figueroa Oreamuno</persname>
  </origination>
  </did>
  <odd type="publicationStatus">
    <p>Publicado</p>
  </odd>
  <controlaccess>
    <persname role="Creador" id="atom_4057510_actor">José María Figueroa Oreamuno</persname>
  </controlaccess>
  <acqinfo encodinganalog="3.2.4">
    <p>Transferencia</p>
  </acqinfo>
  <processinfo>
    <p>
      <date>11/09/2012</date>
    </p>
    <p>Esteban Cabezas Bolaños, Jorge Emilio Jiménez Espinoza</p>
  </processinfo>
  <dsc type="combined"> </dsc>
</archdesc>
</ead>

```

La misma descripción en la interfaz de consulta se muestra de la siguiente manera:

Área de identidad	
Código de referencia	CR AN CR-AN-AH-JMFIGOR-ALBUM-000001-001-1-067f
Título	Reportaje sobre el libro de José María Figueroa
Fecha(s)	<ul style="list-style-type: none"> 1896-10-16 (Creación)
Nivel de descripción	Unidad documental simple
Volumen y soporte	Papel;Digital - 1 unidad
Área de contexto	
Nombre del productor	José María Figueroa Oreamuno
Institución archivística	Archivo Nacional
Origen del ingreso o transferencia	Transferencia
Área de notas	
Notas	Notas: Se advierte que la fecha indicada es aproximada
Área de control de la descripción	
Fechas de creación revisión eliminación	11/09/2012
Nota del archivista	Esteban Cabezas Bolaños, Jorge Emilio Jiménez Espinoza

Objeto digital metadatos

Access Copies >>	
Master file	<p>Nombre del archivo 000001-001-1-067f.jpg</p> <p>Tipo de soporte Imagen</p> <p>Tipo de documento MIME image/jpeg</p> <p>Tamaño del archivo 1.8 MiB</p> <p>Cargado 13 de noviembre de 2023 16:42</p>
Reference copy	<p>Nombre del archivo 000001-001-1-067f_141.jpg</p> <p>Tipo de soporte Imagen</p> <p>Tipo de documento MIME image/jpeg</p> <p>Tamaño del archivo 28 KiB</p> <p>Cargado 13 de noviembre de 2023 16:42</p>
Thumbnail copy	<p>Nombre del archivo 000001-001-1-067f_142.jpg</p> <p>Tipo de soporte Imagen</p> <p>Tipo de documento MIME image/jpeg</p> <p>Tamaño del archivo 13.5 KiB</p> <p>Cargado 13 de noviembre de 2023 16:42</p>

Al ser un estándar, el uso de etiquetas por cada dato es el mismo, por lo que en caso de generar otro XML de alguna otra entrada descriptiva, tendrá la misma estructura de campos y lo que cambia es la información de cada atributo.

A nivel de bases de datos, esta característica es fundamental, para garantizar la interoperabilidad y comunicación entre las bases, para que en caso de migración o transferencia de un sistema a otro la información se mantenga relacionada a los diferentes objetos digitales que describe.



COMPLEMENTO: ESTÁNDAR EAD

La sociedad de archivistas estadounidenses (SAA) cuenta con una plataforma en la que expone buenas prácticas sobre el estándar EAD y EAC <https://saa-sdt.github.io/EAS-Best-Practices/> donde encontrará material de referencia reciente, en idioma inglés.



COMPLEMENTO: RECORDS IN CONTEXTS

El Consejo Internacional de Archivos (ICA) se encuentra trabajando en la actualización de la concepción de los metadatos, con la propuesta de “Records in Contexts: Modelo Conceptual (RiC-CM)” es un modelo conceptual de alto nivel centrado en la identificación y descripción intelectual de los recursos documentales, de los agentes que los crean, utilizan o están documentados en ellos, y de las actividades llevadas a cabo por los agentes que los recursos documentales facilitan y testimonian. Actualmente se encuentra en elaboración la cuarta y última parte de la Norma, llamada “Records in Contexts: Directrices de aplicación (RiC-AG)”.

Según indica el sitio web oficial del ICA RiC-CM reconcilia, integra, desarrolla y sustituye las cuatro normas actualmente existentes: La Norma Internacional General de Descripción Archivística (ISAD(G)); la Norma Internacional sobre los Registros de Autoridad de Archivos relativos a Instituciones, Personas y Familias (ISAAR(CPF)); la Norma Internacional para la descripción de funciones (ISDF); y la Norma Internacional para describir instituciones que custodian fondos de archivo (ISDIAH).

Capítulo 5. Los metadatos técnicos

Ellos constituyen una de las más grandes novedades con respecto a los metadatos necesarios para los archivos en soporte de papel. Son absolutamente necesarios para permitir la conservación a largo plazo de archivos, sin embargo, también para la restitución, con el fin de saber cómo visualizar lo que se ha conservado.

5.1. Definiciones

Los metadatos técnicos: son los metadatos que sirven para identificar, caracterizar, definir el ambiente técnico de los objetos digitales. En el modelo OAIS, los metadatos técnicos corresponden a la información de representación OAIS que sirve para definir cómo transformar un conjunto de bits en información inteligible



Representación gráfica de la información. Elaboración propia a partir de PIAF

Con el fin de reducir los riesgos, es necesario conservar los datos técnicos periféricos a la información a conservar. Los datos describen lo que puede contener un formato de representación y cómo explotarlo.

La identificación puede limitarse al reconocimiento del tipo de archivo ya sea a partir de su extensión, de su "número mágico/ magic number" o de su tipo "mime":

- *El número mágico o magic number es una técnica que consiste en referenciar las cabeceras de archivo en cada formato para determinar rápidamente el tipo de archivo. Introducir en el sistema UNIX, corresponde al comando "file"*
- *El tipo de MINE (multipurpose Internet Mail Extension) son formatos estándares registrados por el IANA (Internet Assigned Number Authority)*

En ambos casos, el nivel de información obtenido es muy pobre: por ejemplo, vamos a saber que tenemos un archivo de imagen tipo TIFF, pero sin conocimiento de la versión en cuestión.

Para la preservación, este simple nivel de información no es suficiente, por lo anterior, conviene identificar el tipo de formato de la manera más precisa posible y, en particular, obtener la versión exacta del formato del objeto digital.

Para ir más allá, puede ser necesario caracterizar completamente un objeto digital; se trata de no sólo identificar con precisión el tipo de formato, sino también definir las opciones técnicas que se han seleccionado para la aplicación de este formato, por ejemplo: tipo de compresión, tipo de codificación. El formato puede referirse a una norma, pero habíamos visto que las normas sobre los formatos se presentaban como muñecas rusas y que el conocimiento de la muñeca más grande no permitió que se conozcan las características de las más pequeñas.

Con una definición precisa y completa del formato de la información, es posible validar el formato, es decir:

- Asegurarse de que las características técnicas definidas por las especificaciones están bien verificadas (conformidad por relación con la norma o estándar)
- Comprobar que estas características cumplen una especificación relacionada con una aplicación específica (conformidad con las normas de uso o condiciones restrictivas que se han decidido para el archivo).

Además, la definición de las características es a menudo necesaria para desarrollar las herramientas de transformación cuando se vaya a realizar una migración de formato.

Los metadatos técnicos no se limitan a la definición de los formatos. En previsión de operaciones de migración de formato o de utilización de herramientas de emulación, es necesario definir el entorno técnico tanto de software como de material, de creación o de restitución del objeto digital.



COMPLEMENTO: METADATOS TÉCNICOS DE FORMATOS DE IMAGEN

Existen varios formatos de **metadatos internos**:

- EXIF (Exchangeable Image File): conjunto de metadatos esencialmente técnicos relativos a la toma de imágenes y suministrados automáticamente por el aparato digital (fabricante y modelo del aparato, altura y anchura de la imagen, fecha y hora de la toma, orientación, resolución, tiempo de exposición, apertura, presencia de un flash, etc.) que es posible integrar en imágenes JPEG/JFIF en particular. El formato EXIF fue desarrollado en 1995 por la JEIDA (Japan Electronic Industry Development Association); en julio de 2016 la *Japan Electronics and Information Technology Industries Association* (JEITA) entidad que se formó tras la desaparición de JEIDA publicó la versión 2.31, que es la más reciente.
- XMP (Extensible Metadata Platform): formato de metadatos basado en XML, creado por Adobe en 2001, utilizado en el interior de los archivos de imagen (JPEG/JFIF, TIFF, GIF, PNG, PDF, SVG...). Incluso si predetermina la forma de almacenar una serie de informaciones más corrientes, en particular las de Dublin Core y EXIF, XMP está abierto a todo tipo de metadatos XML. Es posible explotar los metadatos XMP incluso en ausencia de las aplicaciones originales.

Y un formato normalizado de metadatos externos de las aplicaciones originales

Existe un formato de metadatos específico para la caracterización técnica de las imágenes fijas digitales: Data Dictionary - Technical Metadata for Digital Still Images, norma ANSI/NISO Z39.87 (última versión 2017). Este diccionario de datos tiene una declinación en forma de esquema XML: MIX (Metadata for Images in XML).

Los 200 elementos previstos por el diccionario de datos se dividen en cinco familias: información básica sobre el objeto digital (identificador, tamaño, formato, compresión, fijación), información básica sobre la imagen (dimensiones, color), metadatos de captura de la imagen (tamaño de la fuente, fecha de captura, información sobre el escáner, información sobre la cámara digital, coordenadas geográficas), metadatos de evaluación de la imagen (muestreo de captura, muestreo colorimétrico), historial de las modificaciones originales.

5.2. Información que debe de recopilarse sobre un formato

<p>Selección</p> <p>-Tengo un contenido ¿en qué formato lo debo presentar?</p>	PRODUCCIÓN
<p>Identificación</p> <p>-Tengo un objeto digital ¿en qué formato está hecho?</p> <p>Validación</p> <p>-Tengo un objeto digital, entra en formato X, ¿es correcto?</p> <p>Caracterización</p> <p>-Tengo un objeto en formato X, ¿cuáles son sus propiedades?</p>	INGESTIÓN
<p>Evaluación</p> <p>-Tengo un objeto en formato X con propiedades y, ¿cuál es el riesgo de obsolescencia?</p> <p>Tratamiento</p> <p>-Tengo un objeto en formato X con propiedades y, ¿cómo realizo la operación Z en este formato?</p>	PRESERVACIÓN

Información que debe recopilarse sobre un formato (según Emmanuelle Bermes). Elaboración propia a partir de PIAF



COMPLEMENTO: DROID Y JHOVE - HERRAMIENTAS PARA IDENTIFICAR, CARACTERIZAR Y VALIDAR FORMATOS

LDROID (Digital Record Object Identification) es una herramienta de código abierto con licencia BSD proporcionada por los Archivos Nacionales del Reino Unido. Se basa en el registro de formato PRONOM también soportado por los Archivos Nacionales del Reino Unido. PRONOM busca proporcionar características y referencias sobre la documentación de cada versión de formato. Define un identificador único para cada una de estas versiones, llamado PUID (Pronombre Único Identificar). DROID identifica el formato de un objeto digital proporcionando su PUID.

http://www.nationalarchives.gov.uk/aboutapps/fileformat/pdf/automatic_format_identification.pdf

JHOVE (JSTOR/Harvard Object Validation Environment) es una herramienta de validación de formato público general reducido GNU de código abierto proporcionada por JSTOR y la Universidad de Harvard.

<https://jhove.openpreservation.org/>

El CINES (Centro Informático Nacional de la Enseñanza Superior) Montpellier en Francia, ofrece un servicio a distancia de validación de los formatos, basada en la herramienta JHOVE: <http://facile.cines.fr/>

Capítulo 6. Los metadatos administrativos

Los metadatos administrativos son los metadatos que se utilizan para gestionar la vida del objeto digital. En el modelo OAIS, los metadatos administrativos representan la Información de preservación que sirve para identificar, autenticar, definir la procedencia y el contexto de la información que debe mantenerse, así como los derechos vinculados a esta información del OAIS.

Naturalmente, parte de estos metadatos administrativos también se utilizarán en el proceso de búsqueda de información.

Agrupan los metadatos de identificación, contexto, procedencia, integridad y gestión de derechos.

6.1. Los metadatos de identificación

Los metadatos de identificación permiten la identificación inequívoca de los objetos archivados. La elección de un tipo de identificador debe ser estratégica en el marco de la preservación. Los tipos de identificación son numerosos. El objetivo es adoptar un identificador quién resistirá la prueba del tiempo, las evoluciones de clasificación intelectual de los contenidos, los cambios en la organización física de los datos. Debe ser adaptable. También debe soportar la evolución del tamaño del Archivo. ¿Imagínese las consecuencias que podría tener una elección de ID que encontraría un límite después de algunas décadas? Debe ser extensible. Debe ser capaz de permitir identificar varios niveles de información de un objeto o colección. Debe ser granular.

6.2. Los metadatos de procedencia y contexto

Los metadatos de procedencia y de contexto informan sobre la vida del objeto digital. Éste es el cómo y el por qué: ¿De dónde viene el objeto digital? ¿Cómo se recogió la información? ¿Por qué medios técnicos (sensor del satélite, recepción de la señal, digitalización)? ¿Qué tratamientos fueron realizados? ¿Por quién? ¿Por qué? ¿Cuándo? ¿Cuáles son las razones que motivaron su creación, su captura, su recepción? ¿Qué confianza podemos tener en esta fuente?

6.3. Los metadatos de integridad

Los metadatos de integridad permiten disponer de información relativa al respeto de la integridad de los objetos. Esta información es para uso interno (supervisión y control de los objetos) y/o para uso externo (prevención de litigios, litigios). Se trata, por ejemplo, de huellas dactilares de los ficheros que deben conservarse, obtenidas mediante procedimientos criptográficos.

6.4. Los metadatos de derecho

Permiten gestionar el estatuto legal del objeto digital. No se trata aquí de los DRM (Digital Right Management) en el sentido de Medios Técnicos de Protección (MTP) que pretenden controlar, coaccionar, impedir usos que pueden ser hechas por los usuarios. Entran en esta categoría, para los archivos públicos, los plazos legales de comunicabilidad establecidos por la normativa. Estos mecanismos se aplicarán bien a través del sistema de difusión, bien por el propio Archivo en el momento del acceso.

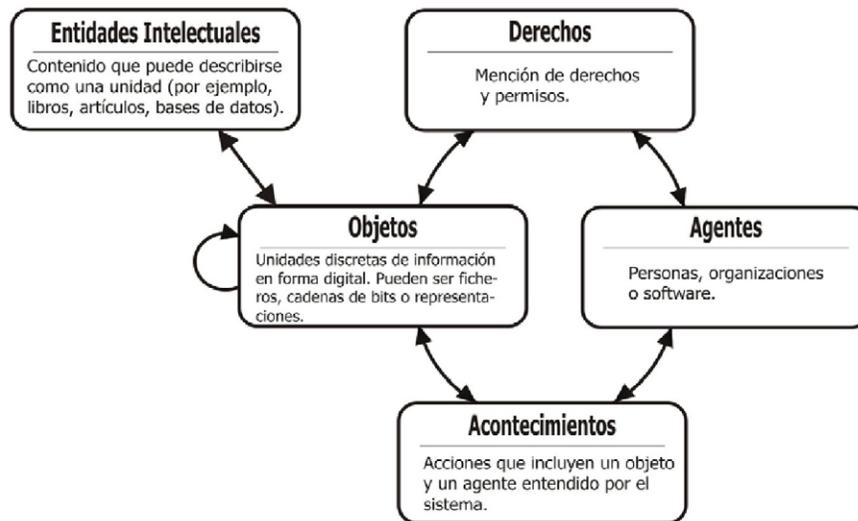
Se trata de metadatos que permitirán definir la política de difusión en función de las limitaciones legales o impuestas por la entidad titular de los derechos. La concreción formal de esta política puede ser una licencia. Hablamos de derechos de uso y derechos de acceso. Ellos definen en función del contexto de uso: ¿Quién? ¿Con qué? (reproductor portátil digital, televisión, ordenador) qué se puede hacer: encontrar, ver, imprimir, copiar, modificar, destruir, y las restricciones de uso: calidad, cantidad, precio.



COMPLEMENTO: DICCIONARIO DE DATOS PREMIS

Elaborado por iniciativa de OCLC (Online Computer Library Center) y RLG (Research Library Group), PREMIS permite identificar los metadatos necesarios para garantizar las principales funciones en un contexto de preservación, definir metadatos cuyo suministro o análisis no requiere intervención humana, elaborar un diccionario de datos totalmente independiente de la implementación y basado en «unidades semánticas» (Semantic units).

PREMIS no es un formato de tipo «cápsula» que se utilizaría para construir los paquetes de información, pero un diccionario que defina elementos de metadatos considerados indispensables para una función de archivo permanente.



Modelo de Datos PREMIS

Fuente: Priscilla Caplan

La principal ventaja de PREMIS, en comparación con otros formatos, es que está diseñado para tener en cuenta las necesidades de preservación digital, y es mantenido por la Biblioteca del Congreso. Se basa en prácticas que tienen en cuenta lo existente y en una fuerte comunidad de usuarios.

Enlace a documento: “Entendiendo PREMIS” de 2021:

https://www.loc.gov/standards/premis/understandingPREMIS_spanish_2021.pdf

6.5. Los metadatos orientados a la gestión de documentos

Los metadatos permiten una calificación cierta para poder así demostrar su autenticidad e integridad.

Por ejemplo, en esta familia de normas se puede consultar la norma ISO 23081-1, denominada “Información y documentación. Procesos de gestión de documentos. Metadatos para la gestión de documentos. Parte 1: 2008”. Define, por ejemplo, además de las principales clases definidas en las partes anteriores, una categoría adicional de metadatos dedicados al comercio electrónico.

Hay diferentes conjuntos de metadatos para el esquema de clasificación, series, carpetas, subcarpetas, volúmenes y documentos, normas de conservación/destrucción, componentes, metadatos de control (después destrucción), los tipos de documentos archivados, extractos, actores, entidades/actores.

Los metadatos incluirán los datos de indexación y otros datos esenciales para un archivo eficaz, especialmente los derechos y restricciones de acceso.

Además, de esta misma norma, parte 2 se destaca que con respecto a la cantidad de metadatos para la gestión de documentos existen límites prácticos a la cantidad de información contextual que puede hacerse explícita y ser incorporada a un sistema determinado en forma de metadatos. El contexto es infinito, mientras un sistema de

información concreto tiene límites finitos. Siempre existirá más información contextual fuera de los límites de cualquier sistema. Una aplicación informática determinada de gestión de documentos solo necesita capturar los metadatos que se consideren necesarios para que el sistema y sus usuarios puedan interpretar y gestionar los documentos durante el tiempo que sean requeridos dentro del sistema y para permitir la migración de los documentos que se tengan que salir del mismo. Los buenos sistemas de metadatos son dinámicos y pueden añadir elementos de metadatos adicionales para la gestión de documentos cuando y como sea necesario a lo largo del tiempo.

Muchos elementos de metadatos para la gestión de documentos pueden obtenerse de otros sistemas de información. Para ser útiles en un sistema de gestión de documentos tienen que estar estructurados y organizados de una forma normalizada. La normalización de los metadatos es un prerrequisito esencial para la interoperabilidad de los sistemas de información tanto dentro de la organización como entre organizaciones. (2011, p. 8).

Capítulo 7. Los metadatos estructurales

Dan los medios para gestionar el árbol de los objetos complejos y devolverlos. En efecto, incluso si un documento simple no plantea ningún problema de estructura en sí mismo, uno puede tener varios niveles de acceso. Permiten así conocer todos los archivos que componen un documento, así como sus relaciones entre ellos.

7.1. Definición y ejemplos de formatos de metadatos de estructura

Los metadatos estructurales sirven para conocer la organización de la información contenida y de los objetos digitales. Hay dos niveles de estructura: un nivel lógico y un nivel físico:

- El nivel lógico define los vínculos entre elementos que tienen sentido para el usuario: número de página, de rangos de audio, título de capítulos, artículos, etc.
- El nivel físico define cómo se guardan los objetos digitales: ¿en qué archivo? ¿En qué repertorio? ¿En qué soporte? Se habla también de mapa de estructura lógica, de estructura física.

En el modelo OAIS, los metadatos de estructura representan la Información de empaquetado OAIS (Ver sección 5 Modelo de Referencia OAIS y Normas Asociadas).



COMPLEMENTO: FORMATOS DE EMPAQUETADO

Los formatos de empaquetado no solo definen metadatos de estructura, sino también paquetes que puedan contener los datos que deban conservarse: los metadatos asociados a dichos datos (descriptivos, técnicos, administrativos) y metadatos (de estructura) que describan la organización lógica y física de este juntos. Los paquetes en cuestión pueden ser reales (todos los datos y metadatos se almacenan juntos), en este caso, el paquete tiene una realidad física. Los paquetes también pueden ser virtuales en el sentido que se define, esencialmente apunta a los datos y metadatos almacenados en diferentes lugares.

Se presenta uno de los formatos de empaquetado: METS



COMPLEMENTO: METS (Metadata Encoding and Transmission Standard)

Mantenido por la Biblioteca del Congreso, METS es un formato diseñado para gestionar cualquier tipo de objeto digital, simple o complejo. Es un formato de tipo «cápsula», por lo que puede integrar cualquier otro formato de metadatos descriptivos o técnicos. Su característica principal es ser modular (una instancia METS se compone de siete secciones) y separar la estructura del objeto y los metadatos (los metadatos se agrupan en secciones específicas, y asociados a los objetos correspondientes a través de enlaces). Permite definir a la vez una tarjeta de estructura física y lógica. METS es así un sobre, un contenedor de metadatos. De metadatos sobre metadatos de alguna manera.

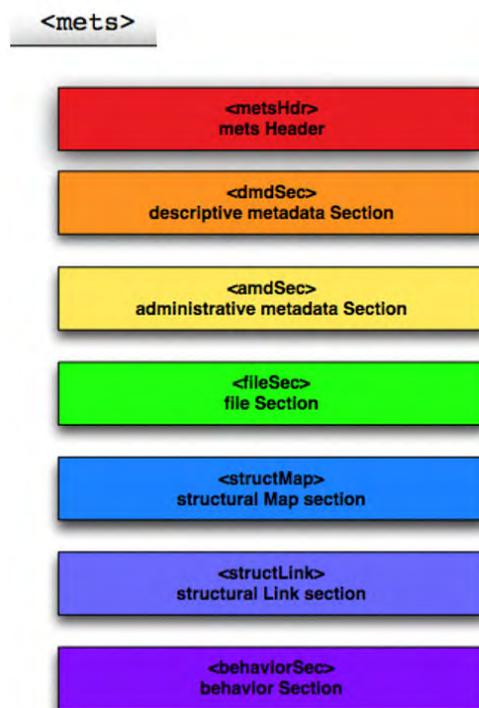
No define qué metadatos deben utilizarse, pero permite categorizar los metadatos y enlazarlos entre sí. Parte del modelo consiste en incluir metadatos o apuntar a metadatos externos. Descripción lógica y física de metadatos se separan en partes distintas del paquete. El estándar define el mecanismo de gestión y organización del sistema de vínculos entre los diferentes elementos. Este principio es extremadamente poderoso, ya que permite adaptarse a todo tipo de organizaciones de datos. La contrapartida reside en la complejidad de la red de vínculos creados.

Se ofrece como un formato basado en XML para codificar los metadatos necesarios para la gestión de objetos digitales y para su intercambio entre repositorios (o entre repositorios y sus usuarios). Dependiendo de cómo se aplique, un documento METS podría usarse como un Paquete de Información de Transferencia (SIP), Paquete de Información de Archivo (AIP), o Paquete de Información de Consulta (DIP) dentro del modelo de referencia OAIS (Sección 5).

Según el sitio web oficial de la Biblioteca del Congreso METS propone organizar la modelización de estos metadatos en siete secciones:

- Cabecera METS - contiene metadatos que describen el propio documento METS, e incluye datos como su creador, editor, fecha de creación etc.
- Metadatos Descriptivos - Esta sección puede: a) apuntar a metadatos descriptivos externos al documento METS (por ejemplo, un registro MARC en un OPAC o un documento EAD disponible en un servidor web); b) contener internamente los metadatos descriptivos, o c) combinar ambas aproximaciones. En la sección Metadatos Descriptivos se pueden incluir múltiples metadatos descriptivos, tanto internos como externos.
- Metadatos Administrativos - ofrece información sobre cómo se crearon y almacenaron los archivos que conforman el objeto digital, derechos de propiedad intelectual, metadatos sobre el objeto original a partir del cual se obtuvo la representación digital, e información sobre la procedencia de los archivos que conforman el objeto digital (es decir, relaciones entre copias maestras y derivadas, migraciones y transformaciones). Al igual que sucede con los metadatos descriptivos, los metadatos administrativos pueden ser externos o codificarse dentro del propio documento METS.

- Sección Archivo - lista todos los archivos con contenidos que forman parte del objeto digital. Los archivos pueden agruparse en elementos <fileGrp>, uno para cada una de las distintas versiones del objeto.
- Mapa Estructural - es la parte principal de un documento METS. Recoge la estructura jerárquica del objeto digital, y enlaza sus secciones con los archivos de contenido y los metadatos correspondientes a cada una de ellas.
- Enlaces Estructurales - permite registrar la existencia de hiperenlaces entre las secciones del mapa estructural. Tiene gran valor cuando se usa METS para archivar sitios web.
- Comportamientos - se puede usar para vincular comportamientos ejecutables con los contenidos del documento METS. Cada comportamiento tiene una definición de interfaz y un "mecanismo" que identifica un módulo de código ejecutable que implementa y ejecuta el comportamiento definido de forma abstracta por la interfaz.



Secciones de METS, imagen en inglés para observar los nombres de las etiquetas normalizadas. Wikipedia

Para ver ejemplos aplicados ingresar al sitio web oficial de METS:

<https://www.loc.gov/standards/mets/mets-examples.html>

Ventajas:

El respeto de los conceptos de OAIS hace del formato METS una buena manera de acoger los metadatos seleccionados para el archivo digital. El estándar prevé incluir, o referenciar, no sólo los metadatos en formato XML, sino que también prevé un mecanismo para incluir otros formatos considerándolos como objetos binarios. Esta capacidad le permite adaptarse a la mayoría de las necesidades en este ámbito.

METS es un formato maduro en el que la comunidad de bibliotecas digitales tiene una buena visibilidad. La mayoría de los proyectos de preservación digital declarados lo utilizan. Desde el punto de vista técnico, el hecho de poder integrar a la vez una estructura física y lógica es una ventaja, ya que permite preservar datos importantes para el acceso y para la «representación» del documento, como las tablas de contenidos. Además, METS permite categorizar los metadatos.

Desventajas:

Para objetos digitales con un gran número de archivos, la instancia METS es difícil de leer y de comprender desde el principio (para un ser humano), debido a una gestión de vínculos compleja. Además, METS no tiene un modelo conceptual, lo que dificulta su migración a otro formato. Por último, algunos elementos son difíciles de interpretar con precisión.

Capítulo 8. Caso de Estudio: España- Esquema de metadatos para la gestión del documento electrónico.

8.1. Objetivos y descripción general

En España, desde 2007 aproximadamente, se han destinado esfuerzos para una legislación Archivística acorde a las necesidades de la administración pública electrónica, lo que dio origen al Esquema Nacional de Interoperabilidad, que comprende el conjunto de criterios y recomendaciones en materia de seguridad, conservación y normalización de la información, de los formatos y de las aplicaciones que deberán ser tenidos en cuenta por las Administraciones Públicas para la toma de decisiones tecnológicas que garanticen la interoperabilidad.

De acuerdo con este trabajo, en el año 2012, se crea la Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos, la cual en su apartado VII menciona la asignación de metadatos y establece (Resolución de 28 de junio de 2012, p. 53779):

1. Las organizaciones garantizarán la disponibilidad e integridad de los metadatos de sus documentos electrónicos, manteniendo de manera permanente las relaciones entre cada documento y sus metadatos.
2. La implementación de los metadatos de gestión de documentos electrónicos para su tratamiento y gestión a nivel interno será diseñada por cada organización en base a sus necesidades, criterios y normativa específica.
3. Los metadatos de gestión de documentos electrónicos se articularán en esquemas de metadatos que responderán a las particularidades y necesidades específicas de gestión de cada organización.

4. El Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE), disponible en el Centro de Interoperabilidad Semántica, que incluye los metadatos mínimos obligatorios, definidos en las Normas Técnicas de Interoperabilidad de Documento electrónico y Expediente electrónico, así como otros metadatos complementarios pertinentes en una política de gestión y conservación de documentos electrónicos, podrá ser utilizado como referencia para la adecuación a los requisitos de interoperabilidad en materia de gestión documental.

8.2. Esquema e-EMGDE

El esquema e-EMGDE es un instrumento que define la incorporación y gestión de los metadatos de contenido, contexto y estructura de los documentos electrónicos a lo largo de su ciclo de vida. La versión 2 establece (2016, p. 9):

- El modelo conceptual en el que se apoya el modelo de metadatos, sus propiedades y su lógica subyacente.
- La descripción de cada uno de los elementos y sub-elementos de metadatos, a través de la descripción de cada una de sus características: definición, propósito, obligatoriedad, etc.
- Los esquemas de valores necesarios para cumplimentar los valores de los elementos de metadatos pertinentes.
- Las referencias a las normas utilizadas como base, así como a otras normas de posible utilidad.

8.2.1. Modelo entidad- relación

El esquema se basa en un modelo entidad-relación que reconoce la existencia de las siguientes entidades:

- Documento: Información estructurada en cualquier formato creada, recibida y mantenida como evidencia por una organización o persona en cumplimiento de obligaciones legales o para actuaciones de gestión.
- Agente: Institución, sistema, persona física o jurídica responsable o involucrada en la creación, producción, custodia o gestión de documentos.
- Actividad: Responsabilidad ejecutada por o asignada a una entidad Agente.
- Regulación: Marco normativo, incluidos los requisitos de gestión de documentos, tales como ordenamiento jurídico, normativa, política, etc. Las posibles categorías para este tipo de entidad son definidas por cada organización en función de sus necesidades y normativas específicas.
- Relación: Asociación entre dos o más entidades que tiene relevancia en un contexto de gestión y/o de gestión de documentos. A efectos de interoperabilidad, en el presente esquema es deseable su aplicación en implantaciones específicas.

6. Este modelo se expresa en la Figura 1:

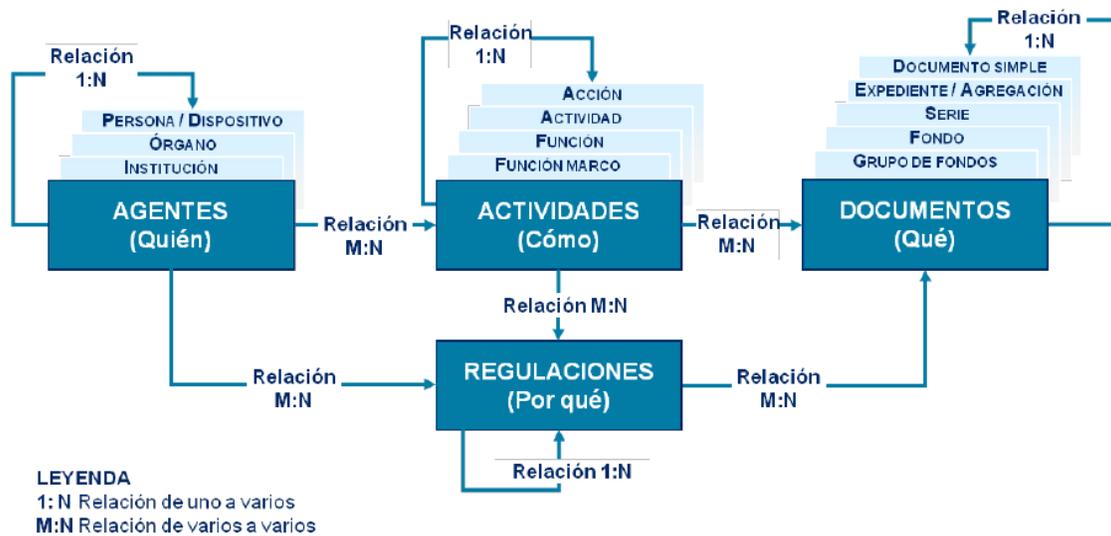


Figura 1. Cobertura de los metadatos para la gestión de documentos electrónicos.

Modelo del Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE) (2016, p. 12)

Se destaca que el esquema describe los elementos de metadatos mínimos necesarios que contribuyen a que los documentos sean auténticos, fiables, íntegros y disponibles, en un momento determinado y a lo largo del tiempo, garantizando su interoperabilidad. También describe algunos de los metadatos necesarios para la conservación a largo plazo. A este respecto, para necesidades de conservación específicas, el esquema es compatible con el PREMIS Data Dictionary. (2016, p. 15), de ahí que en el apéndice 22 del Esquema de metadatos incluya un cuadro de las unidades semánticas que concuerdan entre ambos esquemas de metadatos.

PREMIS			e-EMGDE			
Entidad	Nº	Unidad semántica	Entidad	Elemento	Sub-elemento	Sub-sub-elemento
OBJETO	1.1.1	objectIdentifierType	DOCUMENTO	Identificador	Tipo de identificador	
	1.1.2	objectIdentifierValue		Identificador	Secuencia del identificador	
	1.2	objectCategory				
	1.3.1	preservationLevelValue				
	1.3.2	preservationLevelRole				
	1.3.3	preservationLevelRationale				
	1.3.4	preservationLevelDateAssigned				
1.4.1	significantPropertiesType					

Tabla parcial de concordancias PREMIS y e-EMGDE. Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE) (2016, p. 98)

8.2.2. Características

Algunas de sus características apuntan a que:

Es **independiente de la tecnología**. El esquema prescribe una serie de elementos de metadatos, pero no el procedimiento, o la combinación de procedimientos, para su implantación específica en diferentes entornos de producción. Algunos metadatos se presentarán incrustados en un archivo, otros vinculados al mismo mediante su inclusión en una base de datos o puede que se combinen ambos procedimientos. El esquema no hace previsiones al respecto, siempre que el método o la combinación de métodos elegida respete los principios básicos de relación y de vinculación permanente entre los metadatos de una entidad y la entidad concreta.

Flexibilidad suficiente para permitir su implantación en sistemas de diferente tipo, sin dejar de respetar los criterios mínimos obligatorios del esquema. Por ejemplo, su estructura jerárquica y relacional no impide una implantación plana o lineal.

Interoperabilidad entre distintos sistemas y a lo largo del ciclo de vida de los documentos.

Compatibilidad con otras normas técnicas nacionales, por ejemplo NEDA para la descripción de documentos o internacionales, como ISAD (G), ISAAR (CPF); ISDF y particularmente ISO 23081. En este sentido, el esquema no se estructura en bloques, puesto que los elementos de metadatos pueden tener significados distintos en diversos entornos, pero admite la estructuración mediante los bloques de ISO 23081, si se considera necesaria en implantaciones específicas. La concordancia con ISO 23081, dada la generalidad de esta norma, se realiza sólo en los elementos de nivel superior.

Posibilidad de **reutilización** de los elementos de metadatos dentro de un dominio organizativo y entre dominios.

8.2.3. Estructura del Esquema

El esquema de metadatos consta de 30 elementos. Dependiendo del tipo de implantación se dividen en:

i. **Obligatorios (esenciales):** 16 elementos

[eEMGDE1 - Categoría](#), [eEMGDE2 - Identificador](#), [eEMGDE3 - Nombre](#), [eEMGDE4 - Fechas](#), [eEMGDE6 - Entidad relacionada](#), [eEMGDE13 - Calificación](#), [eEMDGDE14 - Características técnicas](#), [eEMGDE17 - Firma](#), [eEMGDE18 - Tipo documental](#), [eEMGDE20 - Estado de elaboración](#), [eEMGDE22 - Clasificación](#), [eEMGDE23 - Versión NTI](#), [eEMGDE24 - Órgano](#), [eEMGDE25 - Origen del documento](#), [eEMGDE27 - Estado del expediente](#), y [eEMGDE28 - Interesado](#).

Elementos obligatorios. Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE) (2016, p. 17)

ii. **Condicionales** (su uso depende del tipo de entidad que se esté describiendo y el contexto en que ésta funciona): 7 elementos.

[eEMGDE0 - Tipo de entidad](#), [eEMGDE8 - Seguridad](#), [eEMGDE9 - Derechos de acceso, uso y reutilización](#), [eEMGDE10 - Contacto](#), [eEMGDE11 - Idioma](#), [eEMGDE16 - Verificación de integridad](#), y [eEMGDE26 - Identificador del documento origen](#).

Elementos condicionales. Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE) (2016, p. 17)

iii. **Opcionales** (Pueden utilizarse bajo circunstancias en que se requiera una descripción más detallada): 7 elementos.

[eEMGDE5 - Descripción](#), [eEMGDE7 - Ámbito](#), [eEMGDE12 - Puntos de acceso](#), [eEMGDE15 - Ubicación](#), [eEMGDE19 - Prioridad](#), [eEMGDE21 - Trazabilidad](#) y [eEMDGE29 - Asiento registral](#).

Elementos opcionales. Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE) (2016, p. 18)

Parte de los elementos se subdividen en **sub-elementos**, hasta un total de **56**, y éstos a su vez son susceptibles de comprender otros, mediante la aplicación del principio de extensibilidad. Existen **29 sub-sub-elementos**. En un tercer nivel, se encuentran los **sub-sub-sub-elementos**, de los cuales tan solo encontramos 2.

Al igual que los elementos, los sub-elementos y sub-sub-elementos pueden ser obligatorios, condicionales u opcionales, aunque esta posibilidad está condicionada por la aplicación del elemento del nivel superior. Es decir, no pueden utilizarse los sub-elementos por sí solos, sino en dependencia del correspondiente elemento continente. Por ejemplo, para que el sub-elemento 2.1 - Secuencia del identificador, sea utilizado como obligatorio, es preciso utilizar el elemento 2 - Identificador, del que depende.

Así mismo debe tenerse en cuenta que un metadato con subelementos no debe contener un valor propio, sino que deben contenerlo los subelementos inferiores de la jerarquía que incluye.

ELEMENTO	OBLIGATORIEDAD			APLICABILIDAD				
	Obligatorio	Condicional	Opcional	Documento	Agente	Actividad	Regulación	Relación
<u>EEMGDE0 - TIPO DE ENTIDAD</u>		✓		✓	✓	✓	✓	✓
<u>EEMGDE1 - CATEGORIA</u>	✓ _T			✓	✓	✓	✓	✓
<u>EEMGDE2 - IDENTIFICADOR</u>	✓ _E			✓	✓	✓	✓	✓
<u>eEMGDE2.1 - Secuencia de identificador</u>	✓ _E			✓	✓	✓	✓	✓
<u>eEMGDE2.2 - Esquema de identificador</u>		✓		✓	✓	✓	✓	✓
<u>EEMGDE3 - NOMBRE</u>	✓ _T			✓	✓	✓	✓	✓
<u>eEMGDE3.1 - Nombre natural</u>	✓ _T			✓	✓	✓	✓	✓
<u>eEMGDE3.2 - Nombre del fichero</u>			✓	✓	✗	✗	✗	✗
<u>EEMGDE4 - FECHAS</u>	✓ _E			✓	✓	✓	✓	✓
<u>eEMGDE4.1 - Fecha inicio</u>	✓ _E			✓	✓	✓	✓	✓
<u>eEMGDE4.2 - Fecha fin</u>		✓ _T		✓	✓	✓	✓	✓
<u>EEMGDE5 - DESCRIPCION</u>			✓	✓	✓	✓	✓	✓
<u>EEMGDE6 - ENTIDAD RELACIONADA</u>	✓ _T			✗	✗	✗	✗	✓
<u>eEMGDE6.1 - ID de entidad relacionada</u>	✓ _T			✗	✗	✗	✗	✓
<u>eEMGDE6.2 - Esquema de ID de entidad relacionada</u>		✓		✗	✗	✗	✗	✓
<u>eEMGDE6.3 - Rol de la relación</u>	✓ _T			✗	✗	✗	✗	✓

<u>EEMGDE7 - ÁMBITO</u>			✓	✓	✓	✓	✓	✓	×
<u>EEMGDE8 - SEGURIDAD</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.1 - Nivel de seguridad</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.1.1 - Nivel de acceso</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.1.2 - Código de Política de control de acceso</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.2 - Advertencia de seguridad</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.2.1 - Texto de la advertencia</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.2.2 - Categoría de la advertencia</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.3 - Permisos</u>		✓		×	✓	✓	×	×	×
<u>eEMGDE8.4 - Sensibilidad datos de carácter personal</u>		✓		✓	×	✓	✓	✓	×
<u>eEMGDE8.5 - Clasificación ENS</u>		✓		✓	✓	✓	✓	✓	×
<u>eEMGDE8.6 - Nivel de confidencialidad de la información</u>	✓ _T			✓	×	×	×	×	×
<u>EEMGDE9 - DERECHOS DE ACCESO, USO Y REUTILIZACIÓN</u>		✓		✓	×	×	×	×	×
<u>eEMGDE9.1 - Tipo de acceso</u>		✓ _T		✓	×	×	×	×	×
<u>eEMGDE9.2 - Código de la causa de limitación</u>		✓ _T		✓	×	×	×	×	×
<u>eEMGDE9.3 - Causa legal/normativa de limitación</u>		✓ _T		✓	×	×	×	×	×
<u>eEMGDE9.4 - Condiciones de reutilización</u>		✓ _T		✓	×	×	×	×	×
<u>EEMGDE10 - CONTACTO</u>		✓		×	✓	×	×	×	×
<u>eEMGDE10.1 - Tipo de contacto</u>		✓		×	✓	×	×	×	×
<u>eEMGDE10.2 - Dato de contacto</u>		✓		×	✓	×	×	×	×
<u>eEMGDE10.3 - Puesto</u>		✓		×	✓	×	×	×	×
<u>EEMGDE11 - IDIOMA</u>		✓		✓	✓	×	×	×	×
<u>EEMGDE12 - PUNTOS DE ACCESO</u>			✓	✓	×	×	✓	×	×
<u>eEMGDE12.1 - Término punto de acceso</u>			✓	✓	×	×	✓	×	×
<u>eEMGDE12.2 - ID de punto de acceso</u>		✓		✓	×	×	✓	×	×
<u>eEMGDE12.3 - Esquema</u>		✓		✓	×	×	✓	×	×
<u>EEMGDE13 - CALIFICACIÓN</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.1 - Valoración</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.1.1 - Valor primario</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.1.1.1 - Tipo de valor</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.1.1.2 - Plazo</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.1.2 - Valor secundario</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.2 - Dictamen</u>	✓ _T			✓	×	×	×	×	×
<u>eEMGDE13.2.1 - Tipo de dictamen</u>	✓ _T			✓	×	×	×	×	×

eEMGDE13.2.2 - Acción dictaminada		✓ _T		✓	x	x	x	x
eEMGDE13.2.3 - Plazo de ejecución de la acción dictaminada		✓ _T		✓	x	x	x	x
eEMGDE13.3 - Transferencia	✓ _T			✓	x	x	x	x
eEMGDE13.3.1 - Fase de archivo	✓ _T			✓	x	x	x	x
eEMGDE13.3.2 - Plazo de transferencia	✓ _T			✓	x	x	x	x
eEMGDE13.4 - Documento esencial	✓ _T			✓	x	x	x	x
EEMGDE14 - CARACTERÍSTICAS TÉCNICAS	✓ _E			✓	x	x	x	x
eEMGDE14.1 - Formato	✓ _E			✓	x	x	x	x
eEMGDE14.1.1 - Nombre del formato	✓ _E			✓	x	x	x	x
eEMGDE14.1.2 - Extensión del fichero	✓ _E			✓	x	x	x	x
eEMGDE14.2 - Versión de formato			✓	✓	x	x	x	x
eEMGDE14.3 - Resolución			✓	✓	x	x	x	x
eEMGDE14.4 - Tamaño		✓		✓	x	x	x	x
eEMGDE14.4.1 - Dimensiones físicas		✓		✓	x	x	x	x
eEMGDE14.4.2 - Tamaño Lógico			✓	✓	x	x	x	x
eEMGDE14.4.3 - Cantidad			✓	✓	x	x	x	x
eEMGDE14.4.4 - Unidades		✓		✓	x	x	x	x
eEMGDE14.5 - Profundidad de color			✓	✓	x	x	x	x
EEMGDE15 - UBICACIÓN			✓	✓	x	x	x	x
eEMGDE15.1 - Soporte		✓		✓	x	x	x	x
eEMGDE15.2 - Localización			✓	✓	x	x	x	x
EEMGDE16 - VERIFICACIÓN DE INTEGRIDAD		✓		✓	x	x	x	x
eEMGDE16.1 - Algoritmo			✓	✓	x	x	x	x
eEMGDE16.2 - Valor			✓	✓	x	x	x	x
EEMGDE17 - FIRMA	✓ _E			✓	x	x	x	x
eEMGDE17.1 - Tipo de firma	✓ _E			✓	x	x	x	x
eEMGDE17.1.1 - Formato de firma	✓ _E			✓	x	x	x	x
eEMGDE17.1.2 - Perfil de firma		✓		✓	x	x	x	x
eEMGDE17.2 - Rol de firma		✓		✓	x	x	x	x
eEMGDE17.3 - Valor del CSV		✓		✓	x	x	x	x
eEMGDE17.4 - Definición generación CSV		✓		✓	x	x	x	x
eEMGDE17.5 - Firmante			✓	✓	x	x	x	x
eEMGDE17.5.1 - Nombre y apellidos o razón social			✓	✓	x	x	x	x
eEMGDE17.5.2 - Número de identificación de los firmantes			✓	✓	x	x	x	x
eEMGDE17.5.3 - En calidad de			✓	✓	x	x	x	x
eEMGDE17.5.4 - Nivel de firma			✓	✓	x	x	x	x
eEMGDE17.5.5 - Información adicional			✓	✓	x	x	x	x

<u>EEMGDE18 - TIPO DOCUMENTAL</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE19 - PRIORIDAD</u>			✓	✓	x	✓	x	x
<u>EEMGDE20 - ESTADO DE ELABORACIÓN</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE21 - TRAZABILIDAD</u>			✓	✓	✓	✓	✓	✓
<u>eEMGDE21.1 - Acción</u>			✓	✓	✓	✓	✓	✓
<u>eEMGDE21.1.1 - Descripción de la acción</u>			✓	✓	✓	✓	✓	✓
<u>eEMGDE21.1.2 - Fecha de la acción</u>			✓	✓	✓	✓	✓	✓
<u>eEMGDE21.1.3 - Objeto de la acción</u>			✓	✓	✓	✓	✓	✓
<u>eEMGDE21.2 - Motivo reglado</u>			✓	✓	✓	✓	✓	✓
<u>eEMGDE21.3 - Usuario de la Acción</u>		✓		✓	✓	✓	✓	✓
<u>eEMGDE21.4 - Descripción</u>		✓		✓	✓	✓	✓	✓
<u>eEMGDE21.5 - Modificación de los metadatos</u>		✓		✓	✓	✓	✓	✓
<u>eEMGDE21.6 - Historia del cambio</u>		✓		x	x	x	x	✓
<u>eEMGDE21.6.1 - Nombre del elemento</u>		✓		x	x	x	x	✓
<u>eEMGDE21.6.2 - Valor anterior</u>		✓		x	x	x	x	✓
<u>EEMGDE22 - CLASIFICACIÓN</u>	✓ _E			✓	x	x	x	x
<u>eEMGDE22.1 - Código de clasificación</u>	✓ _E			✓	x	x	x	x
<u>eEMGDE22.2 - Denominación de clase</u>	✓ _T			✓	x	x	x	x
<u>eEMGDE22.3 - Tipo de clasificación (SIA/Funcional)</u>	✓ _T			✓	x	x	x	x
<u>EEMGDE23 - VERSIÓN NTI</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE24 - ÓRGANO</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE25 - ORIGEN DEL DOCUMENTO</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE26 - IDENTIFICACIÓN DEL DOCUMENTO ORIGEN</u>		✓		✓	x	x	x	x
<u>EEMGDE27 - ESTADO DEL EXPEDIENTE</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE28 - INTERESADO</u>	✓ _E			✓	x	x	x	x
<u>EEMGDE29 - ASIENTO REGISTRAL</u>			✓	✓	x	x	x	x
<u>eEMGDE29.1 - Tipo de asiento registral</u>			✓	✓	x	x	x	x
<u>eEMGDE29.2 - Código de la oficina de registro</u>			✓	✓	x	x	x	x
<u>eEMGDE29.3 - Fecha del asiento registral</u>			✓	✓	x	x	x	x
<u>eEMGDE29.4 - Número de asiento registral</u>			✓	✓	x	x	x	x

Tabla 6. Listado de elementos de metadato que componen el e-EMGDE.

Listado de Elementos del esquema. Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE) (2016, p. 19-21)

Órgano	Identificador normalizado de la administración generadora del documento o que realiza la captura del mismo.	1:N	Cadena de caracteres	Código alfanumérico único para cada órgano/unidad/oficina extraído del Directorio Común gestionado por el Ministerio de Política Territorial y Administración Pública.
Fecha de captura	Fecha de alta del documento en el sistema de gestión documental.	1	Fecha/hora	Formato: AAAAMMDD T HH:MM:SS <ISO 8601>
Origen	Indica si el contenido del documento fue creado por un ciudadano o por una administración.	1	Lógico	'0' = Ciudadano '1' = Administración
Estado de elaboración	Indica la naturaleza del documento. Si es copia, este metadato indica también si se ha realizado una digitalización o conversión de formato en el proceso de generación.	1	Cadena de caracteres	<ul style="list-style-type: none"> – Original (Ley 11/2007 Ad. 30). – Copia electrónica auténtica con cambio de formato (Ley 11/2007 Art.30.1). – Copia electrónica auténtica de documento papel (Ley 11/2007 Art.30.2 y 30.3). – Copia electrónica parcial auténtica. Otros.
Nombre de formato	Formato lógico del fichero de contenido del documento electrónico.	1	Cadena de caracteres	Valor extraído de la lista de formatos admitidos para ficheros definidos en la Norma Técnica de Interoperabilidad de Catálogo de estándares.
				Documentos de decisión:
				– Resolución.
				– Acuerdo.
				– Contrato.
				– Convenio.
				– Declaración.
				Documentos de transmisión:
				– Comunicación.
				– Notificación.
				– Publicación.
				– Acuse de recibo.
Tipo documental	Descripción del tipo documental del documento.	1	Cadena de caracteres	Documentos de constancia:
				– Acta.
				– Certificado.
				– Diligencia.
				Documentos de juicio:
				– Informe.
				Documentos de ciudadano:
				– Solicitud.
				– Denuncia.
				– Alegación.

METADATO	VALOR	DESCRIPCIÓN
Versión NTI	http://administracionelectronica.gob.es/ENI/XSD/v1.0/documento-e	1.0
Identificador	ES_E04921301_2012_MINHAP0000000000000000000010207	Id del documento
Órgano	E04921301	MINHAP
Fecha captura	2012-12-17T09:30:47Z	17/12/2012
Origen	1	Administración (MINHAP)
Estado de elaboración	EE01	Original
Nombre de formato	PNG	Imagen en formato PNG incluida como contenido codificado en base64.
Tipo documental	TD08	Publicación
Tipo de firma	TF01	CSV
Valor CSV	UjBsR09EbGhjZ0dTQUxNQUFBUUNBRU1tQ1p0dU1GUXhEUzhi	Valor CSV no real
Definición generación CSV	BOE-A-0034-05200	Identificación de la regulación y/o actividad que autoriza la generación del código CSV.

Tabla 1. Metadatos mínimos obligatorios.

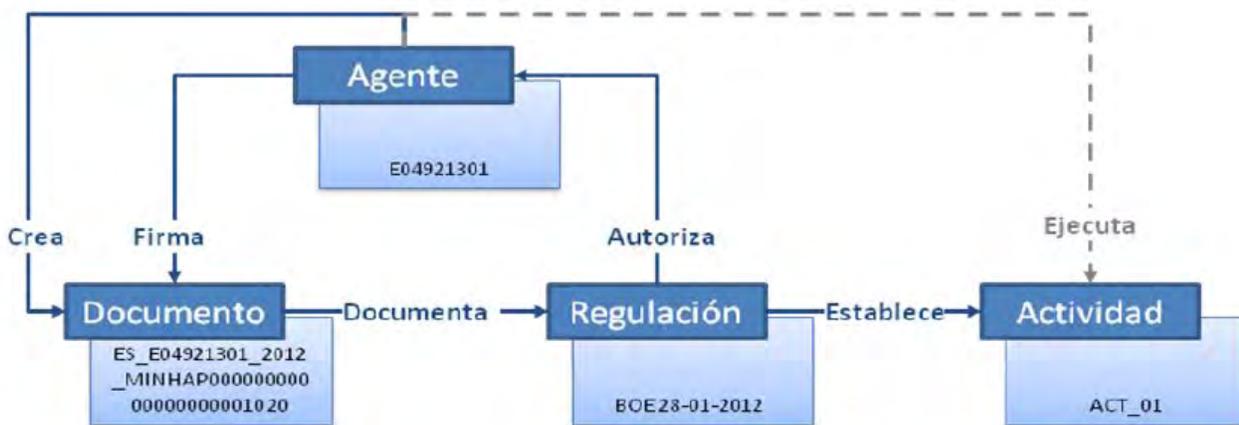


Figura 1. Modelo gráfico del ejemplo de documento electrónico.

Bibliografía

- Archivo Nacional de Costa Rica. Departamento Servicios Archivísticos Externos (2020). Glosario único de términos, definiciones, conceptos y abreviaturas de las normas técnicas nacionales. https://www.archivonacional.go.cr/web/dsae/glosario_%20unico_terminos.pdf
- BANAT-BERGER F., HUC C., DUPLOUY L., *L'Archivage numérique à long terme, les débuts de la maturité?* (Primera obra de síntesis sobre el archivo digital en lengua francesa) Paris, La Documentation française, 2009
- BANAT-BERGER F., HUC C., Module 7 - Gestion et archivage des documents numériques. Portail International Archivistique Francophone. 2011. <https://www.piaf-archives.org/se-former/module-7-gestion-et-archivage-des-documents-numeriques> (Se identifica en el texto como PIAF)
- Barnard, A y Voutssas, J (2014). Glosario de Preservación Archivística Digital Versión 4.0. Universidad Nacional Autónoma de México. https://iibi.unam.mx/archivistica/glosario_preservacion_archivistica_digital_v4.0.pdf
- Digital Library Federation, Metadata Encoding and Transmission Standard: manual de introduction y reference, version 1.6, 2007. https://www.loc.gov/standards/mets/METSOverview_spa.html
- Ministerio de Política Territorial y Administración Pública, Gobierno de España (2011). Norma Técnica de Interoperabilidad de Documento Electrónico. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-13169
- Ministerio de Política Territorial y Administración Pública, Gobierno de España (2012). Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2012-10048
- Ministerio de Política Territorial y Administración Pública, Gobierno de España (2016). Esquema de Metadatos para la Gestión del Documento Electrónico (e-EMGDE). Versión 2.0. Documentación complementaria a la Norma Técnica de Interoperabilidad de Política de gestión de documentos electrónicos.
- Caplan, Priscila (2009). Entender PREMIS. https://www.loc.gov/standards/premis/UnderstandingPREMIS_espanol.pdf



ARCHIVO NACIONAL
COSTA RICA



UNIVERSIDAD DE
COSTA RICA