



BIBLIOTECA
DE GALICIA



AXENCIA PARA A
MODERNIZACIÓN
TECNOLÓXICA DE GALICIA

REQUIRIMENTOS TÉCNICOS PARA PROXECTOS DE DIXITALIZACIÓN DE FONDOS BIBLIOGRÁFICOS EN INSTITUCIÓNS DA MEMORIA DE GALICIA

Outubro 2025



Índice

1. INTRODUCCIÓN.....	3
2. DEFINICIÓNS.....	6
3. DIXITALIZACIÓN.....	11
3.1. Obtención das imaxes.....	11
3.2. Identificación das imaxes.....	14
3.3. Control de calidade.....	16
3.4. Proceso de OCR (<i>Optical Character Recognition</i>).....	17
3.5. Identificación dos ficheiros ALTO-XML.....	18
4. ESTRUCTURA E NOMEADO DOS DIRECTORIOS.....	19
4.1. Arquivos de difusión.....	19
Materiais non seriados.....	19
Materiais seriados.....	22
4.2. Arquivos de preservación.....	26
Materiais seriados e non seriados.....	26
5. METADATOS.....	29
5.1. Ficheiros METS para a inxesta no repositorio de Galiciana-BDG.....	29
5.1.1. METS de carga de materiais non seriados en JPG.....	29
5.1.2. METS de carga de materiais non seriados en PDF.....	33
5.1.3. METS de carga de publicacións seriadas en JPG.....	33
5.1.4. METS de carga de publicacións seriadas en PDF.....	40
5.2. Ficheiros METS para a inxesta no repositorio de preservación da MDG.....	41
ANEXOS.....	43
Anexo A. Cadro resumen requisitos técnicos das imaxes.....	43
Anexo B. Exemplo comentado de METS para a carga de material non seriado en jpg.....	43
Anexo C. Exemplo comentado de METS para a carga de material non seriado en pdf.....	43
Anexo D. Exemplo comentado de METS para a carga de publicacións seriadas en jpg.....	43
Anexo E. Exemplo comentado de METS para a carga de publicacións seriadas en pdf.....	43
Anexo F. Exemplo de METS de preservación para a Memoria Dixital de Galicia.....	43
BIBLIOGRAFÍA.....	44



1. INTRODUCCIÓN

A **Memoria Dixital de Galicia** constitúese como un camiño global e colectivo de transformación, aproveitando a potencialidade das TIC, do modelo de xestión do patrimonio cultural e de dinamización da actividade cultural.

A construción dunha **Memoria Dixital de Galicia** é un proxecto colectivo de conservación, valorización e difusión do patrimonio cultural de Galicia. Este proxecto non se centra unicamente na fase de dixitalización, senón que ten en conta a totalidade do proceso do patrimonio cultural dixital, asumindo unha necesaria coordinación e complementariedade entre a dixitalización, a preservación dixital e difusión dos activos resultantes. A **Memoria Dixital de Galicia** nace coa vontade de ser un proxecto con continuidade no tempo, transformando a relación cos usuarios e favorecendo a implicación da cidadanía.

A **Memoria Dixital de Galicia** é unha estratexia impulsada desde o ámbito público, a través da Consellería de Cultura, Educación, Formación Profesional e Universidades e a Axencia para a Modernización Tecnolóxica de Galicia (Amtega) que busca a colaboración coa sociedade galega e con todos os axentes implicados na xestión do patrimonio cultural galego.

Para construír a **Memoria Dixital de Galicia** como novo modelo de xestión do patrimonio cultural galego baseado nas TIC identificáronse 5 liñas de actuación:

1º - Definir un marco técnico común para todos os axentes e institucións involucrados na conservación e difusión do patrimonio cultural mediante o establecemento de políticas, normas, procedementos e estándares para a recolleita, catalogación, preservación e difusión do patrimonio cultural galego.

2º - Aplicar as novas tecnoloxías para a catalogación, inventariado e xestión do patrimonio, o que implica a dotación de infraestruturas e ferramentas tecnolóxicas nas institucións da memoria que custodian os bens patrimoniais, así como a capacitación no seu uso.

3º - Aproveitar as TIC para facilitar a visibilización e a promoción do patrimonio cultural, fundamentalmente mediante o impulso de políticas de dixitalización do patrimonio e tamén fomentando a aplicación ou desenvolvemento de técnicas novas e interactivas para o acceso, conservación, restauración e posta en valor do patrimonio cultural.

4º - Aplicar a tecnoloxía para mellorar a accesibilidade ao patrimonio desde o punto de vista presencial e virtual. No primeiro caso, para mellorar os



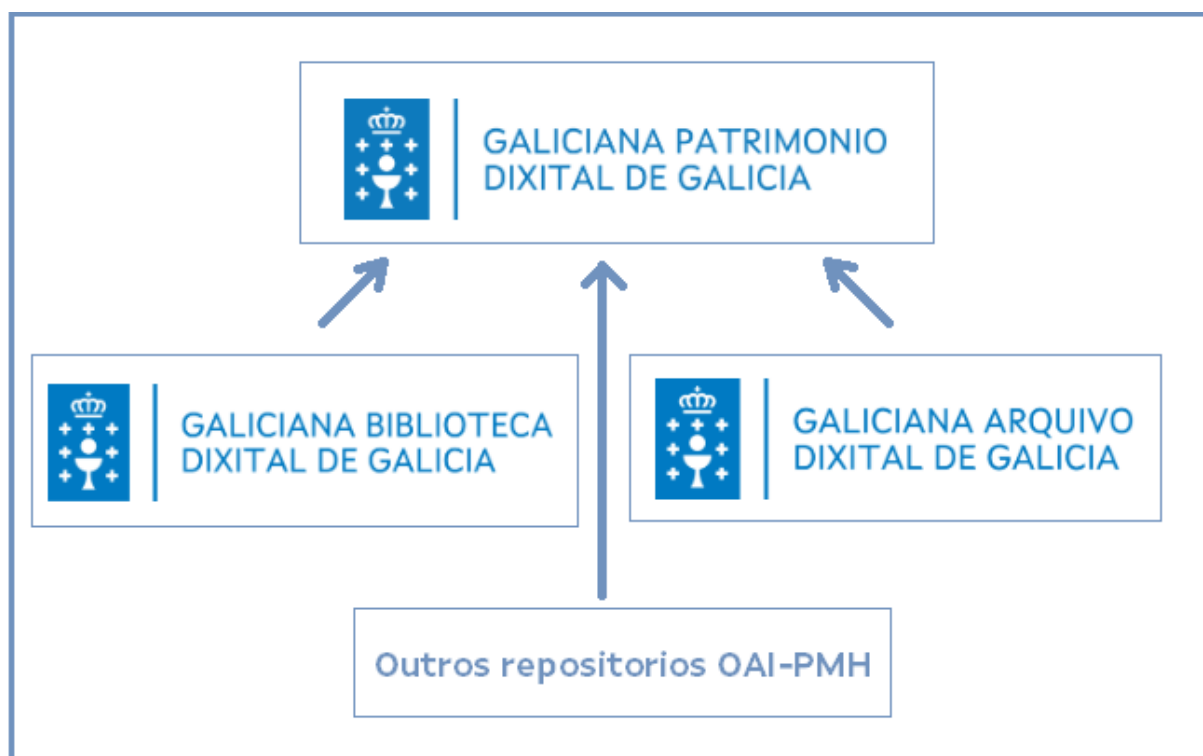
servizos prestados directamente en arquivos, bibliotecas, museos e calquera outra institución que custodie os recursos patrimoniais. No segundo, incrementando e mellorando as solucións dixitais de acceso ao patrimonio, adaptándoas ao perfil e ás necesidades do usuario, desde o público en xeral a profesionais, docentes, investigadores...

5º - Fomentar os uso das TIC como elemento dinamizador da produción cultural, impulsando a creación de contidos dixitais culturais e a súa correspondente promoción, e que contribuirá a xerar novas oportunidades de desenvolvemento económico na sociedade galega.

Tal e como se establece na 1º liña de actuación, un dos obxectivos que pretende acadar a **Memoria Dixital de Galicia** é o *"establecemento de políticas, normas, procedementos e estándares para a recolleita, catalogación, preservación e difusión do patrimonio cultural galego"*. Este documento trata de cumprir con este obxectivo marcándose como finalidade:

- Normalizar as características técnicas da dixitalización e os estándares mínimos de asignación de metadatos para así facilitar o intercambio de recursos e facer posible as buscas a través dunha única plataforma (Galiciana-Patrimonio Dixital de Galicia).
- Asesorar e coordinar os esforzos que as diferentes institucións da memoria de Galicia estean a levar adiante en materia de dixitalización, de forma que se optimicen os recursos persoais, económicos e materiais.

Estrutura de Galiciana-Patrimonio Dixital de Galicia





É importante sinalar que este documento céntrase nas pautas a seguir para unha correcta dixitalización e identificación mediante esquemas de metadatos dos materiais vinculados ao eido bibliotecario (monografías, folletos, mapas, postais, publicacións periódicas, carteis ou fotografías).

Estas recomendacións non pretenden ser un manual de dixitalización senón simplemente un instrumento para o traballo en común e a coordinación dos proxectos que se están a realizar ou se realicen no futuro dentro do ámbito bibliotecario galego. Trátase dun documento adaptado ás características e panorama actual das normativas e directrices nacionais e internacionais vixentes para estes procesos e que requirirá de posteriores actualizacións que permitan adaptar o seu contido a futuros desenvolvementos neste ámbito de aplicación.

A existencia de pautas e directrices xa publicadas e adoptadas tanto a nivel nacional como internacional, aconsella non establecer novas normativas que poidan diverxer dos ámbitos comúns xa establecidos. En consecuencia, este documento segue e respecta esas directrices xa consolidadas e extrae as súas principais ideas das que se citan na bibliografía.

De xeito particular, para a redacción destes requisitos foi fundamental o establecido nos documentos elaborados pola Subdirección General de Coordinación Bibliotecaria, *Requisitos técnicos para proyectos de digitalización de patrimonio bibliográfico y de prensa histórica de la SGCB (v. 2.1 de 26.06.2023)*, e pola Biblioteca Nacional de España, *Proceso de digitalización en la Biblioteca Nacional de España (actualizado a 25.02.2015)*.

En resumo, a función desta publicación está, por un lado, en divulgar no noso ámbito unha serie de conceptos que teñen unha importancia básica para o deseño dunha política común de dixitalización do patrimonio cultural e por outro, concretar ou desenvolver aqueles aspectos desta política que poidan axudar a levar adiante procesos de dixitalización do patrimonio bibliográfico e documental na nosa comunidade.

As especificacións concretas de cada contrato de dixitalización reflectiranse nos pregos de prescricións técnicas. De existir algunha discrepancia entre estes requisitos e o exposto nos pregos de prescricións técnicas, terase en conta en primeiro lugar, o especificado no prego.



2. DEFINICIÓNS

Dixitalización: proceso de converter unha sinal analóxica en dixital.

Imaxe dixital: foto electrónica tomada dunha escena ou escaneada dalgún documento (fotografías, manuscritos, textos impresos, ilustracións).

Captura da imaxe: proceso polo que se obtén unha representación dixital dun orixinal constituída por un conxunto de elementos pictóricos ou píxeles mediante o escaneado ou fotografía dixital.

Píxel (abreviatura de "picture element"): Elemento mínimo dos que, en conxunto, forman unha imaxe dixital.

Ppp (píxeles por polgada)=ppi (pixels per inch): resolución de imaxe/resolución óptica/resolución de captura. Sirve para medir a resolución que é a cantidade de píxeles que entran nunha polgada. A resolución de imaxe é unha das variables que determinan a súa calidade. Unha dixitalización cunha resolución insuficiente faría que perdésemos detalles do orixinal ou xeraría imaxes pouco nítidas. Pola contra, un aumento de resolución aumenta o detalle e a calidade hasta certo punto, a partir do cal non se produce mellora algunha xa que non hai maior detalle no orixinal.

Resolución: indica o número de píxeles utilizados para representar a imaxe, medidos en proporción á unha superficie determinada (píxeles por polgada). A resolución é a capacidade dunha imaxe de representar o detalle. Cantos máis píxeles ou puntos se utilicen para definir unha polgada da imaxe, máis resolución terá esta. En xeral, a máis resolución, máis fiel ao orixinal será a imaxe dixital. A maior resolución a imaxe terá máis peso e tamén máis custo de produción, polo que é importante considerar para cada material a resolución máis adecuada.

Compresión: permite reducir o tamaño do ficheiro da imaxe para o seu almacenamento e transmisión. A compresión pode ser con perda ou sen perda. Sempre se recomenda empregar sistemas de compresión estandarizados.

Compresión con perda: proceso que reduce o espazo de almacenamento necesario para o ficheiro dunha imaxe mediante a eliminación de datos desa imaxe. Ao descomprimir unha imaxe que experimentase unha compresión con perda sempre será distinta da imaxe antes de comprimila, incluso anque a diferenza sexa difícil de detectar para o ollo humano.

Compresión sen perda: proceso que reduce o espazo de almacenamento necesario para o ficheiro dunha imaxe sen perda de datos. Se unha imaxe



experimentou unha compresión sen perda, será idéntica á imaxe antes de comprimirse.

Ficheiros máster: o ficheiro máster é a versión da imaxe dixital de alta calidade sen compresión. A súa función é a de servir como copia de preservación e representar coa máxima fidelidade posible o orixinal. A partir destes ficheiros xeneraranse os outros ficheiros de imaxe de menor calidade.

Ficheiros derivados: obtéñense dos ficheiros máster e son os que se empregan para a consulta e difusión a través das redes. O requisito fundamental destes ficheiros é que son máis lixeiros despois de pasar por un proceso de compresión con perda.

Ficheiros de miniaturas: ficheiros dixitais que representan as imaxes en versións pequenas, de baixa resolución.

Bits: unidade mínima de información dixital. Un bit só pode tomar os valores 0 e 1.

Profundidade de cor: é o número de bits utilizados para representar a cor de cada píxel. A máis profundidade, máis detalles cromáticos.

Bitonal: 1 bit (branco ou negro)

Escala de grises: entre 2 e 8 bits (ata 256 tons de grises)

Cor: 24 bits (16.7 millóns de cores)

O mesmo que coa resolución, cada tipo de material require unha definición óptima que permita explotar ó máximo as súas posibilidades de información é o tempo non desaproveite recursos innecesarios.

Tipo de cor: branco e negro, escala de grises ou color, en relación tamén coa profundidade escollida.

Formato: os datos recollidos polo software do escáner, deben ser almacenados de acordo cun formato. Os formatos de arquivo son unidades lóxicas de almacenamento da información que consisten tanto nos bits que comprende a imaxe como os datos acerca de como ler e interpretar o arquivo. Varían en termos de resolución, profundidade de bits, soporte para compresión e metadatos. A elección do formato dependerá da finalidade da imaxe dixital (preservación, difusión,...). Hai que escoller formatos estándar. Considérase unha boa práctica que na fase de captura o formato de saída sexa de alta calidade e sen compresión, do cal poidan derivarse logo os subprodutos (imaxes para a difusión, miniaturas e PDF).

TIFF (Tagged Image File Format): formato de ficheiros para imaxes con etiquetas. Isto se debe a que os ficheiros TIFF conteñen, ademais dos datos da



imaxe propiamente dita, "etiquetas" nas que se arquiva información sobre as características da imaxe, que serve para o seu tratamento posterior. Este formato é de aplicación xeneralizada para a creación de imaxes de alta calidade, produce ficheiros de gran tamaño, sen perda, útiles como ficheiros máster pero inadecuados para a distribución e acceso público ás coleccións. Pode presentar calquera resolución, branco e negro, escala de grises ou cor.

JPEG (Joint Photographic Experts Group): norma para a compresión da imaxe con calidade fotográfica na World Wide Web. É un formato para o almacenamento e transmisión de imaxes na Web. O seu algoritmo de compresión permite reducir o tamaño dos ficheiros, sen perda ou cunha perda pouco significativa da calidade da imaxe.

PDF (Portable Document Format): é un formato de almacenamento de documentos desenvolvido pola empresa Adobe Systems, especialmente adecuado para a presentación de documentos complexos (múltiples páxinas, combinación de textos e imaxes de diferentes calidades). Este formato ofrece, entre outras vantaxes, opcións de navegación no documento e entre diferentes documentos, fidelidade e seguridade da copia dixital e posibilidades de busca e recuperación a partir dos contidos.

Metadatos: conxunto de informacións relacionadas cos obxectos dixitais. Representan ao obxecto dixital con fins de descrición, xestión e intercambio. Divídense en diferentes grupos segundo a finalidade que sustentan (procura e recuperación da información, xestión de obxectos dixitais, estruturación física e lóxica dos mesmos). O seu obxectivo é garantir a interoperabilidade, visibilidade, autenticidade e preservación dos obxectos dixitais.

Metadatos descritivos: esquemas especialmente destinados a procura e recuperación da información e á recolleita dos datos. Son os metadatos que describen o contido físico e conceptual dun obxecto. Algúns dos esquemas de metadatos descritivos máis empregados son MARCXML, DC, MODS ou ONIX. Poden utilizarse por si mesmos ou formando parte doutros esquemas e protocolos.

Metadatos administrativos: son aqueles que describen a procedencia dun obxecto dixital, os procesos realizados para a súa creación ou xeración, as súas características técnicas, as súas condicións de acceso e dereitos de propiedade intelectual e as accións xa realizadas ou previstas relacionadas coa preservación do obxecto en si mesmo. Comprenden tanto os metadatos técnicos como os de propiedade intelectual e de preservación. Xeralmente empréganse para a xestión interna dos ficheiros dixitais.



Metadatos técnicos: describen os atributos do obxecto dixital. Son os metadatos que proporcionan a información sobre o proceso de captura (data, axentes que dixitalizaron, ...) e os elementos técnicos (hardware e software) empregados para escanear os documentos orixinais. Tamén facilitan datos sobre os formatos e a calidade das imaxes.

Metadatos de preservación: datos que permiten seguir o ciclo de vida dos obxectos dixitais para garantir a súa accesibilidade futura, a súa correcta interpretación e a súa autenticidade e integridade, calquera que sexa o sistema ou metodoloxía que se empregue para elo. Conteñen a información que utiliza un repositorio para soportar o proceso de preservación. Diferéncianse dos metadatos técnicos no sentido de que documentan tamén os procesos que sofren os obxectos dixitais ao longo do tempo (copias, migracións ou calquera outra alteración), incluso despois de importados a un repositorio.

Metadatos estruturais: son os que permiten describir as relacións entre as diferentes partes dun recurso dixital. Achegan información sobre a estrutura intelectual dun obxecto (páxinas, capítulos, ...) e tamén permiten asociar os diferentes formatos que poidan existir dese obxecto.

METS (Metadata Encoding and Transmission Standard): o esquema METS é un estándar para a codificación dos metadatos descritivos, administrativos e estruturais dos obxectos dunha biblioteca dixital. Para elo emprega XML como linguaxe de marcado. Xurde no 2002 como unha iniciativa da *Digital Library Federation* para proporcionar os datos necesarios, nun contorno XML, para a xestión dos materiais dixitais. Permiten tanto o almacenamento dos datos dos obxectos dixitais, como o intercambio de información entre bases de datos ou a difusión aos usuarios finais. É un estándar que mantén a *Network Development and MARC Standards Office* da *Library of Congress*.

MARC 21 XML (MARC XML): esquema que permite reproducir un rexistro MARC completo en XML.

MODS (Metadata Object Description Schema): esquema de metadatos descritivos que se deriva do MARC 21. Os seus elementos teñen unha presentación máis lexible que os do formato MARC XML (<titleinfo> en lugar de <datafiled tag="245" ind1="1" ind2="0"). A súa principal vantaxe é a súa capacidade para establecer relación xerárquicas entre as descrições e proporcionar unha codificación detallada das partes dun recurso (volume, número, capítulo, ...).

PREMIS (Preservation Metadata Implementation Strategies): metadatos de preservación que conteñen a información que utiliza un repositorio para xestionar o proceso de preservación dixital. Están baseados no modelo de referencia OAIS



(Open Archival Information System), o cal establece a necesidade de rexistrar unha serie de datos mínimos que permitan seguir o ciclo de vida dos obxectos dixitais para garantir a súa accesibilidade, interpretación, autenticidade e integridade futuras.

OAI-PMH (Open Archives Initiative–Protocol Metadata Harvesting): é un protocolo de recolección sinxelo para o intercambio de metadatos entre repositorios. Os metadatos a recolectar poder estar en calquera formato establecido por calquera conxunto específico de provedores de datos e provedores de servizos), con independencia de que, como mínimo, deben codificarse en Dublin Core no cualificado para proporcionar un nivel básico de interoperabilidade.

OCR (Optical Character Recognition–Recoñecemento Óptico de Caracteres): tecnoloxía que permite ler os caracteres de texto individuais dunha páxina en formato dixital e converter esa información nun ficheiro de texto que pode ser almacenado de forma electrónica.

ALTO (Analyzed Layout and Text Object) XML: esquema para especificar os metadatos técnicos do deseño e contido dos recursos textuais físicos, tales como as páxinas dun libro ou dun periódico, e que permiten codificar en XML o resultado dun proceso de OCR.



3. DIXITALIZACIÓN

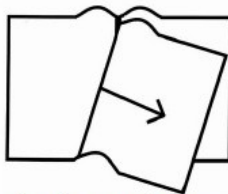
A dixitalización é un proceso no que inciden tanto elementos de carácter informático, referentes especialmente ás características técnicas do proceso e resultados da dixitalización en si, como documentais que afectan aos metadatos que achegamos ás imaxes para facilitar a súa descrición e posterior recuperación polos usuarios das bibliotecas dixitais.

Neste punto trataremos de aclarar aspectos concretos do proceso de obtención das imaxes dixitais mentres que o seguinte centrarase nos elementos documentais da dixitalización.

Contar con perfís técnicos comúns para os proxectos de dixitalización que se desenvolvan en Galicia é imprescindible para que todos poidan estar presentes en Galiciana, ventá de acceso en Internet ao patrimonio dixital de Galicia.

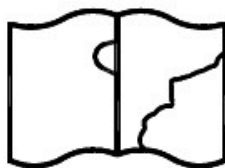
3.1. Obtención das imaxes

- Deberán empregarse escáneres aéreos e de tamaño axustado aos orixinais.
- Asegurarase a correcta manipulación dos exemplares. En ningún caso poderán ser desencadernados ou guillotizados e evitarase o emprego de prensalibros.
- A dixitalización de todos tipo de materiais impresos, sexan de carácter seriado ou monográfico (libros, folletos, manuscritos, fotografías, postais, carteis, planos, gravados, mapas, música impresa, etc.), realizarase en cor cunha profundidade de 24 bits (16,7 millóns de cores).
- A dixitalización das obras realizarase procurando un axuste o máis real posible aos documentos orixinais. Haberá que capturar e xerar imaxes de páxinas enteiras nas que se poidan apreciar os bordos dos documentos. Así, recortaranse as imaxes deixando ao redor da imaxe un bordo que terá como máximo tres milímetros. Para conseguir este resultado, non se poderán utilizar as ferramentas de recorte automático.
- No proceso de dixitalización das obras usarase unha cartolina de cor negro debaixo de cada páxina ou dunhas poucas páxinas para visualizar de forma correcta os bordos das obras. Estas cartolinas, que deben sobresaír do volume, evitan que se vexan na imaxe os bordos do resto de páxinas e das encadernacións. No caso de que o soporte estivese incompleto ou deteriorado interporanse follas de papel xaponés para evitar ver as seguintes páxinas.
- Se no documento que se está a dixitalizar faltase unha páxina, no seu lugar escanearase unha páxina substitutiva cun texto do tipo "Falta páxina". En lugar deste texto tamén se pode usar o símbolo correspondente a esta eventualidade na norma ISO 9878/1990.



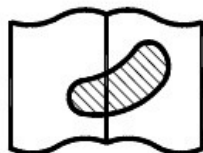
Faltan documentos (páxinas,
cadernos...)
ISO 9878/1990

- Se no documento que se está a dixitalizar houbera unha páxina que debido ao seu estado de conservación non se aconsellase a súa dixitalización, escanearase unha páxina substitutiva cun texto do tipo "Páxina deteriorada". No lugar deste texto tamén se pode usar o símbolo correspondente a esta eventualidade na norma ISO 9878/1990.



Documento deteriorado
(páxinas, cadernos...)
ISO 9878/1990

- Se no documento que se está a dixitalizar houbera unha páxina que resultase ilexible por calquera causa (p. ex. porque está pegada á páxina seguinte), escanearase unha páxina substitutiva cun texto do tipo "Páxina ilexible". No lugar deste texto tamén se pode usar o símbolo correspondente a esta eventualidade na norma ISO 9878/1990.



Documento ilexible
(páxinas, cadernos...)
ISO 9878/1990



- No caso de que o documento que se está a dixitalizar tivese páxinas con textos ou ilustracións cunha orientación distinta á normal (p.ex. unha revista que ten páxinas centrais con carteis ou unha táboa que ocupa dúas páxinas e que se visualiza rotando o documento), deberase rotar as imaxes dixitais correspondentes a esas páxinas para ofrecer estes contidos debidamente orientados de maneira que non sexa necesario rotalas na pantalla para a súa visualización ou lectura.
- Encadernacións facticias. Son aquelas obras individuais que se encadernaron xuntas con posterioridade á súa edición, seguindo criterios diversos que teñen que ver cos intereses dun posuidor concreto, p.ex. obras da mesma temática. Recoméndase dixitalizar todas as obras que constitúen ese volume facticio e asociar cada grupo de obxectos dixitais resultantes da dixitalización á cada rexistro bibliográfico e alta de exemplar correspondente. Especificarase en cada proxecto de dixitalización que facer coas imaxes correspondentes á encadernación facticia: asocialas á primeira e/ou última obra dixitalizada ou non dixitalizalas.
- Para os materiais gráficos que teñan o verso en branco, como carteis ou fotografías, determinarase para cada proxecto de dixitalización, a conveniencia ou non de dixitalizar o verso. Da mesma forma, especificarase a idoneidade ou non, de dixitalizar outras partes dos documentos que non conteñen información, como carpetas ou forros engadidos posteriormente á creación das obras, normalmente con fins de conservación.
- As imaxes das obras non se entregarán a dobre páxina. No caso de que o fondo seleccionado permita a súa dixitalización a dobre páxina (manuscritos, monografías e incunables), do proceso de escaneado obterase un ficheiro máster TIFF en cru e un TIFF recortado en dúas partes, é dicir, un ficheiro por cada páxina que será o ficheiro máster.
- A captura dos ficheiros máster realizarase no formato TIFF 6.0/ISO 12639:2004 sen compresión.
- De ser necesario, realizarase o enderezamento das imaxes para lograr que a copia dixital sexa o máis fidedigna posible ao orixinal.
- Deberán empregarse elementos de control, como cartas de grises ou de cores ou escalas métricas, que se inclúen nos procesos de dixitalización de materiais textuais o gráficos para conseguir unha representación fiel do orixinal.
- As páxinas que se dixitalicen en formato TIFF xunto a estes elementos de control, dixitalizaranse dúas veces, unha cos referidos elementos e outra sen eles.
- A imaxe TIFF na que aparezan as cartas de cores ou escalas métricas irá tamén apropiadamente referenciada no ficheiro METS de preservación asociado ao correspondente obxecto dixital. A imaxe que conteña a carta de cor ou de grises, nomearase co número 0.



Exemplo de nomeado dunha imaxe en formato tiff na que se inclúe unha carta de cor:

ou-bp_ecoo_18830101_0000.tif

- Para a difusión dos obxectos utilizarase o formato JPEG/ISO/IEC 10918:1994 con compresión e optimizado para a súa consulta a través de internet.
- Tamén se proporcionarán miniaturas de cada unha das imaxes dixitalizadas e un arquivo en formato PDF de cada unha das obras.
- Nos formatos derivados JPEG, PDF e MINIATURAS dos ficheiros máster non será necesario incluír a imaxe cos elementos de control empregados para garantir a máxima calidade na dixitalización.
- Os ficheiros cas marcas de auga incluíranse nas respectivas imaxes JPEG e no PDF para difusión en web, nunca nas copias máster. En todo caso, serán de pequeno tamaño e non se superpoñerán ao contido do documento. No caso da dixitalización de materiais gráficos de pequeno tamaño, como postais ou ephemera (etiquetas de produtos, cromos, etc.) especificarase no proxecto de dixitalización concreto o tipo de marca de auga máis conveniente para este tipo de materiais ou incluso se poderá determinar non incluír marca de auga.
- Os requisitos mínimos para a obtención das imaxes segundo o tipo de documento móstranse no **Anexo A**.

3.2. Identificación das imaxes

Todas as imaxes obtidas de proxectos de dixitalización sistemáticos terán que identificarse de forma unívoca para o que se recomenda cumprir coas seguintes pautas:

Para as publicacións non seriadas

- a) O primeiro elemento na identificación das imaxes será o código asignado a institución depositaria dos orixinais, que será facilitado pola Biblioteca Dixital de Galicia.
- b) O seguinte elemento será a sinatura que identifica a obra
- c) A continuación identificarase o número correlativo da imaxe na estrutura do obxecto dixital empregando 4 díxitos.
- d) Por último identificarase a extensión do formato de imaxe correspondente
- e) Os primeiros 3 elementos irán separados entre si por un guión baixo, mentres que a extensión do ficheiro irá precedido por un punto.
- f) O nomeado das imaxes farase empregando letras minúsculas (a-z) e os guións baixos ('_') e medios ('-').



Exemplos

 es-scbg_pb4868_0008.tif	es-scbg → Biblioteca de Galicia pb4868 → O divino sainete: poema en oito cantos / m. Curros Enríquez 0008 → número correlativo da imaxe .tif → extensión do arquivo
 po-v-fp_mp4(12).pdf	po-v-fp → Biblioteca da Fundación Penzol mp4(12) → Carta de la Ría y Puerto de La Coruña número correlativo da imaxe ⁽¹⁾ .pdf → extensión do arquivo
 lu-m-sm_pb4868_0728.jpg	lu-m-sm → Biblioteca do Seminario Santa Catalina. Diócese de Mondoñedo-Ferrol pb4868 → Adagiorum des. Erasmi Roterodami Chiliades quatuor cum sesquicenturia cum diligentia ... 0728 → número correlativo da imaxe .jpg → extensión do arquivo

⁽¹⁾ Os ficheiros PDF que se entreguen nun proxecto de dixitalización sistemática serán habitualmente multipáxina, polo que neste caso non será necesario identificar os números correlativos de cada imaxe. Neste tipo de arquivos podería existir un terceiro elemento que indique o número correlativo da imaxe si polo volume da obra se considerase pertinente a entrega do ficheiro PDF dividido en partes.

Para as publicacións seriadas

- O primeiro elemento na identificación das imaxes será o código asignado a institución depositaria dos orixinais.
- O seguinte elemento será un código, de ata 5 caracteres, que identifique o título da publicación seriada.
- A continuación identificarase o ano, mes e día do número da publicación. No caso de existir erros de imprenta na data do ítem dixitalizado, a data seleccionada para a codificación das imaxes será sempre a da data real de publicación. ⁽¹⁾ ⁽²⁾
- O cuarto elemento será o número correlativo da imaxe na estrutura do obxecto dixital empregando 4 díxitos (para ficheiros tiff, jpg e ALTO-XML). Este número de secuencia non constará no nomeado dos ficheiros mets e pdf.



- e) Por último identificarase a extensión do formato de imaxe correspondente.
- f) Os primeiros 4 elementos irán separados entre si por un guión baixo, mentres que a extensión do ficheiro irá precedido por un punto.
- g) O nomeado das imaxes farase empregando letras minúsculas (a-z) e os guiños baixos ('_') e medios ('-').

⁽¹⁾ Se non existise mes e/ou día concreto no ítem dixitalizado (ex: cabeceiras bimensuais ou semestrais) optarase polo día 01 e polo mes que máis se aproxime a súa data real de publicación

⁽²⁾ Cando a data do ítem dixitalizado non sexa o suficientemente clara para diferenciar os exemplares da publicación entre si poderase engadir o número do ítem correspondente

Exemplos

 <code>es-scbg_eco_18960301_0001.tif</code>	es-scbg → Biblioteca de Galicia eco → El Eco de Santiago: diario independente 18960301 → ítem do día 1 de marzo de 1896 0001 → número correlativo da imaxe .tif → extensión do arquivo
 <code>po-bp_pro_19071020_0003.jpg</code>	po-bp → Biblioteca Pública de Pontevedra pro → El Progreso: semanario independente 19071020 → ítem do día 20 de outubro de 1907 0003 → número correlativo da imaxe .jpg → extensión do arquivo
 <code>po-bmp_tea_19080905.pdf</code>	po-bmp → Biblioteca-Museo de Ponteareas tea → El Tea: semanario independente 19080905 → ítem completo do día 5 de setembro de 1908 .pdf → extensión do arquivo

En todo caso, a Biblioteca de Galicia tratará de facilitar os códigos das institucións da memoria, as sinaturas de cada unha das obras dixitalizadas e os códigos que identifican aos títulos de publicacións seriadas.

3.3. Control de calidade

O control de calidade sobre os resultados do proceso de dixitalización é unha tarefa fundamental. Deberán revisarse todas as imaxes e prestarase atención, sobre todo, aos seguintes aspectos:



- Non existen imaxes borrosas
- Non hai imaxes repetidas
- As imaxes están perfectamente encadradas e verticais
- Os bordes das imaxes están correctamente recortados
- Están dixitalizadas a totalidade das páxinas
- A orde dos obxectos dixitais e das imaxes é o correcto
- Os ficheiros (en tódolos formatos) son perfectamente lexibles

3.4. Proceso de OCR (*Optical Character Recognition*)

A dixitalización de todo tipo de materiais impresos que queiran formar parte do proxecto da **Memoria Dixital de Galicia** terá que levar asociado un proceso de recoñecemento óptico de caracteres (OCR) sobre toda as páxinas dixitalizadas que conteñan texto impreso (normalmente serán imaxes de páxina, pero tamén poderían ser textos en mapas ou materiais gráficos, etc.).

Non será preciso realizar este proceso para materiais gráficos sen texto. No caso dos manuscritos, os técnicos de Galiciana-Biblioteca Dixital de Galicia valorarán para cada proxecto, en función das condicións de conservación de cada manuscrito e da tecnoloxía dispoñible nese momento, a conveniencia ou non de realizar este proceso sobre esta tipoloxía documental.

O resultado do OCR proporcionarase en ficheiros normalizados XML segundo o esquema ALTO (*Analyzed Layout and Text Object*), mantido pola Network Development and MARC Standards Office da Biblioteca do Congreso dos Estados Unidos. Os ficheiros XML deberán incluír unha estrutura xerárquica cos niveis de contido apropiados (tomos, números, páxinas, columnas, seccións, epígrafes e ilustracións) e incluírán por cada división de páxina un vínculo co ficheiro de imaxe correspondente.

Na estrutura dos ficheiros METS incluíranse os elementos e as direccións necesarias para identificar as páxinas co proceso de OCR codificado en ficheiros XML ALTO.

As copias que se entreguen en formato PDF dos materiais dixitalizados deberán conter texto oculto, para que sexa posible a busca por palabras no seu contido textual.



3.5. Identificación dos ficheiros ALTO-XML

Para a identificación de cada un dos arquivos ALTO-XML obtidos do proceso de cada una das páxinas das publicacións seriadas dixitalizadas recoméndase seguir o mesmo esquema que se detalla para as imaxes deste tipo de material no punto 3.2 deste documento. A diferenza virá determinada polo formato do arquivo que será **.xml**.

Exemplo

 po-bmp_agr_18960301_0001.xml	po-bmp → Biblioteca-Museo de Ponteareas agr → El Agro Celta 18960301 → Ítem do día 3 de xaneiro de 1926 0001 → número correlativo do ficheiro ALTO-XML .xml → extensión do arquivo
--	--



4. ESTRUCTURA E NOMEADO DOS DIRECTORIOS

4.1. Arquivos de difusión

Para a súa carga no repositorio de Galiciana-Biblioteca Dixital de Galicia, a Biblioteca de Galicia decidiu que os arquivos destinados á difusión nos proxectos de dixitalización sistemática vinculados á **Memoria Dixital de Galicia** serán os seguintes:

- Imaxes en formato JPG
- Ficheiros PDF
- Miniaturas
- Ficheiros ALTO-XML
- Ficheiros METS de carga de JPGs
- Ficheiros METS de carga de PDFs
- Opcionalmente as imaxes en formato TIFF

Materiais non seriados

Os arquivos destinados á difusión nos proxectos de dixitalización de materiais non seriados deberán entregarse en directorios coa seguinte estrutura:

- 1º nivel – Directorio xeral do proxecto
- 2º nivel – Directorios por formatos
- 3º nivel – Directorios por institución
- 4º nivel – Directorios de cada obra dixitalizada

O **nomeado** deses directorios terá que cumprir as seguintes indicacións:

O directorio de 1º nivel identificará o nome que recibe o proxecto de dixitalización. Esa codificación será proporcionada pola Biblioteca de Galicia. Exemplo:

- **cm_dixi_monografías_2025**

O 2º nivel, que reflectirá cada un dos formatos de ficheiros seleccionados para a difusión, e codificaranse simplemente polos nomes dos formatos respectivos: jpeg, pdf e mets. Exemplos:

- **jpeg**
- **pdf**
- **miniaturas**
- **alto**
- **mets_carga_jpg**
- **mets_carga_pdf**



O 3º nivel, que identificará á institución depositaria dos orixinais dixitalizados, aparecerá nomeada polo código que lle proporcione a Biblioteca de Galicia. Exemplo:

- Para a Biblioteca de Galicia empregárase o código **es-scbg**

O 4º nivel, que identificará cada obra dixitalizada, estará nomeado polo código da institución conservadora seguido da sinatura topográfica que lle teña asignado a institución conservadora e que identifica de maneira unívoca o documento dentro da colección. Exemplo:

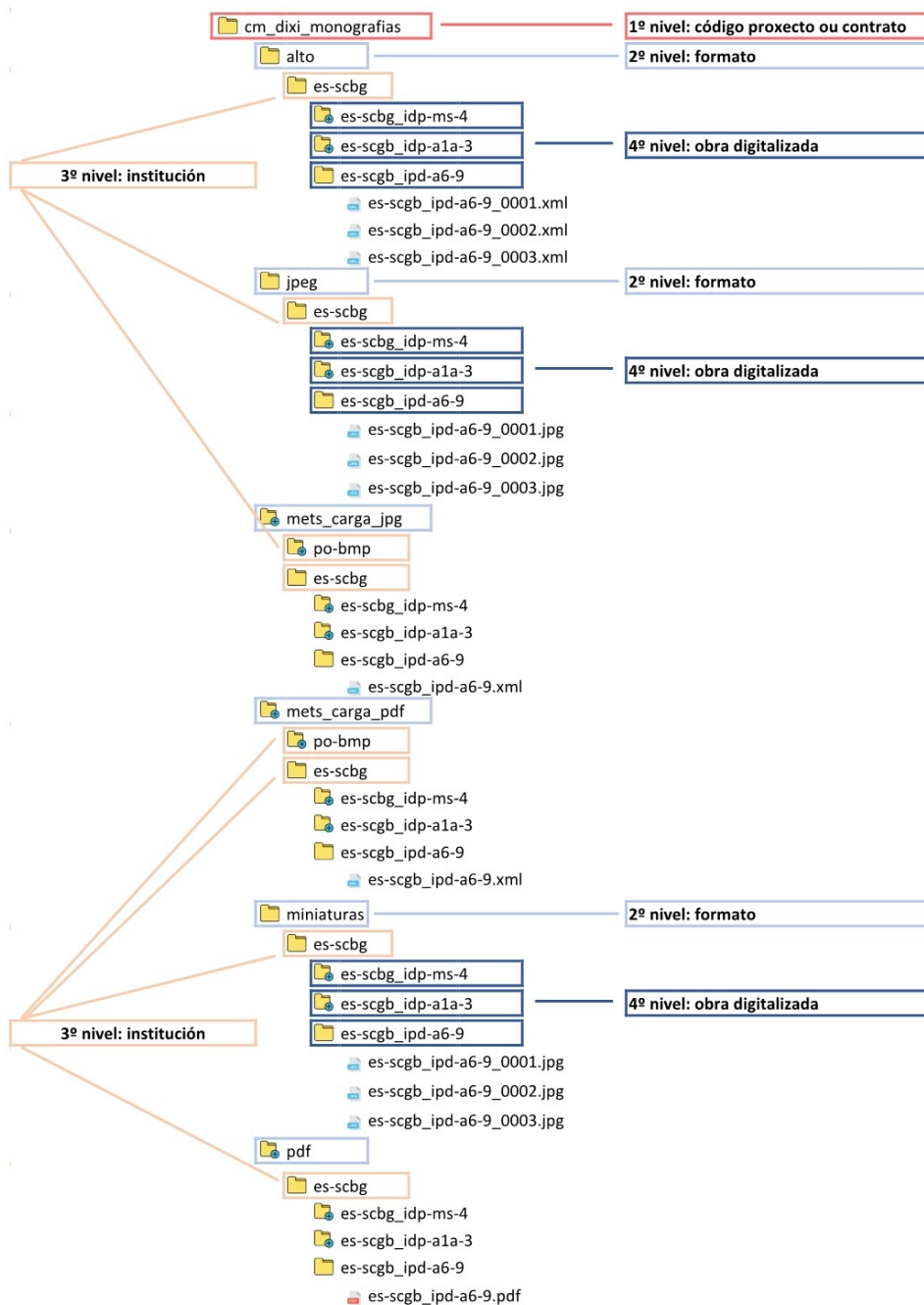
- **es-scbg_idp-ms-4**

Este 4º nivel de codificación na estrutura dos directorios terá que coincidir cos dous primeiros elementos establecidos para a identificación das imaxes. Por exemplo, a ruta para unha imaxe jpg nun ficheiro mets de difusión dunha monografía dixitalizada da Biblioteca de Galicia terá este camiño (estrutura de directorios) e nome do ficheiro da imaxe:

\jpeg\es-scbg\es-scbg_idp-ms-4\es-scbg_idp-ms-4_0001.jpg



Exemplo





Materiais seriados

Os ficheiros de difusión obtidos como resultado de procesos de dixitalización sistemática de prensa histórica almacenaranse seguindo os niveis de directorios establecidos a continuación:

- 1º nivel – Directorio xeral do proxecto
- 2º nivel – Directorios por formatos
- 3º nivel – Directorios por institución
- 4º nivel – Directorios por cada título de publicación seriada dixitalizada
- 5º nivel – Directorios por anos
- 6º nivel – Directorios por cada número de publicación seriada dixitalizado

O **nomeado** deses directorios terá que cumprir as seguintes indicacións:

O directorio de 1º nivel, que identificará o proxecto de dixitalización, será proporcionado pola Biblioteca de Galicia. Exemplo:

- **cm_dix_seriadas_2020**

O 2º nivel, que reflectirá cada un dos formatos de ficheiros seleccionados para a difusión, identificaranse simplemente polos nomes dos formatos respectivos:

- **jpeg**
- **miniaturas**
- **pdf**
- **alto**
- **mets_carga_jpg**
- **mets_carga_pdf**

O 3º nivel, que identificará á institución depositaria dos orixinais dixitalizados, aparecerá nomeada polo código que lle proporcione a Biblioteca de Galicia. Exemplo:

- Para a Biblioteca Municipal de Ponteareas empregárase o código **po-bmp**

O 4º nivel, que identificará cada título de prensa dixitalizado, estará nomeado polo código da institución seguido dun código unívoco de cada título, proporcionado pola Biblioteca de Galicia. Exemplo:

- **po-bmp_tea**

O 5º nivel, que identificará cada ano dixitalizado.



O 6º nivel, que identificará cada número de prensa dixitalizado, estará nomeado polo código da institución seguido do código para cada título e por último coa indicación de ano/mes/día de cada número (aaaammdd). Exemplo:

- Para *El Tea* da Biblioteca Municipal de Ponteareas: **po-bmp_tea_19080905**

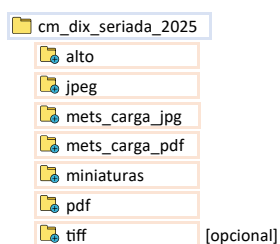
Este 6º nivel de codificación na estrutura dos directorios terá que coincidir cos tres primeiros elementos establecidos para a identificación das imaxes de materiais non seriados. Por exemplo, a ruta para unha imaxe jpg nun ficheiro mets de difusión dunha publicación seriada dixitalizada da BMP terá este camiño (estrutura de directorios) e nome do ficheiro da imaxe:

\jpeg\po-bmp\po-bmp_tea\1908\po-bmp_tea_19080905\po-bmp_tea_19080905_0001.jpg



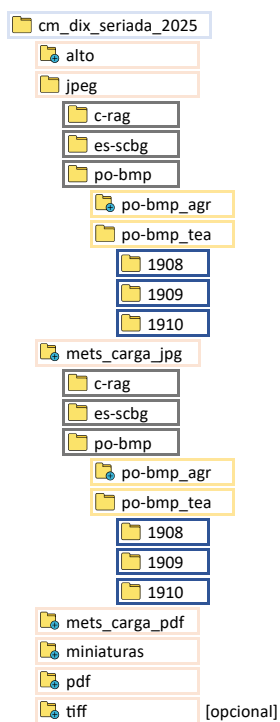
Exemplo

A estrutura da entrega se realizará cun primeiro nivel co código do proxecto ou contrato e un segundo nivel cos formatos



1º nivel: código proxecto ou contrato
2º nivel: formatos de difusión

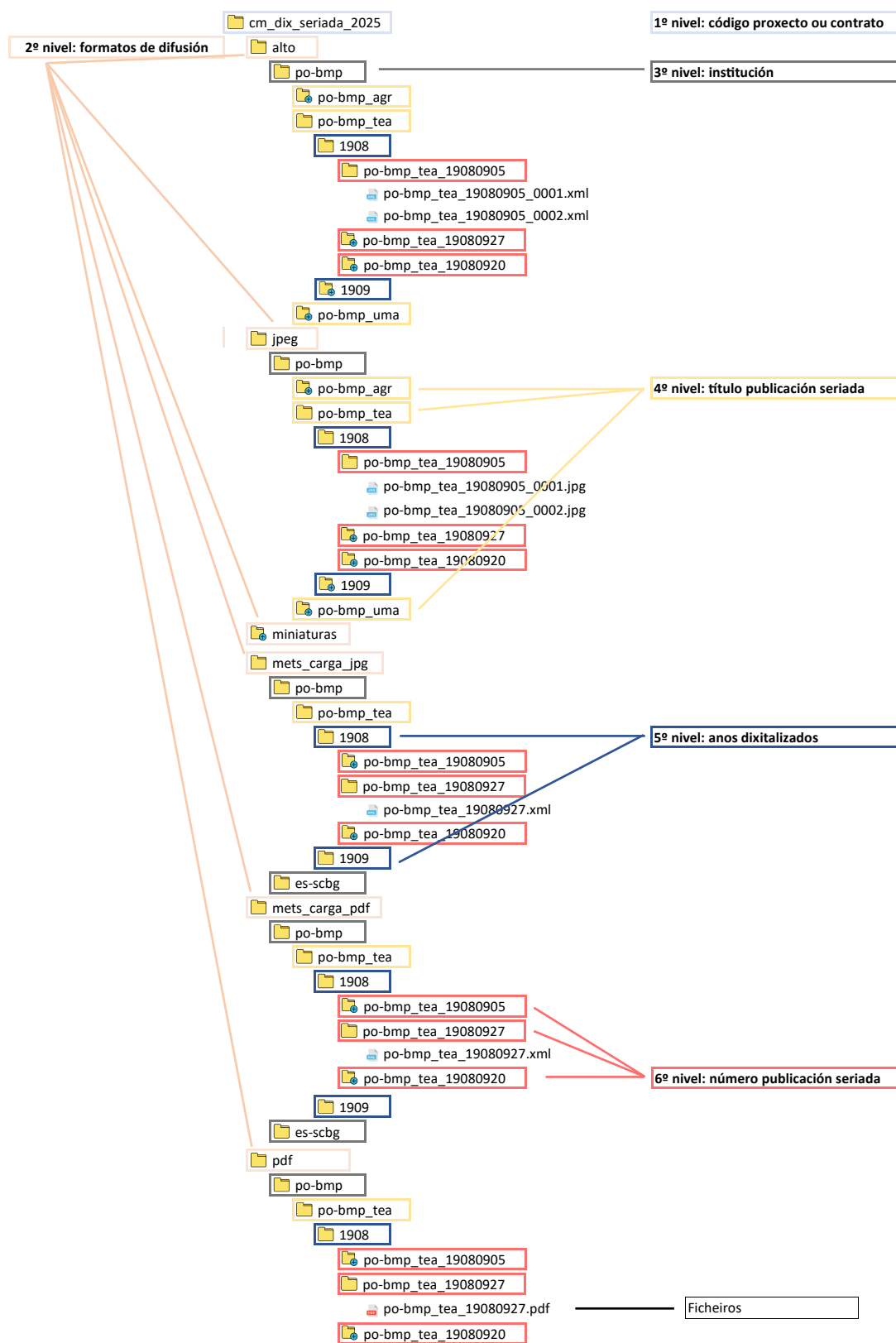
Dentro de cada formato o terceiro nivel será a institución conservadora, o cuarto a cabeceira e o quinto os anos.



1º nivel: código proxecto ou contrato
2º nivel: formatos de difusión

3º nivel: institución
3º nivel: institución
3º nivel: institución
4º nivel: título publicación seriada
4º nivel: título publicación seriada
5º nivel: anos dixitalizados
5º nivel: anos dixitalizados
5º nivel: anos dixitalizados

3º nivel: institución
3º nivel: institución
3º nivel: institución
4º nivel: título publicación seriada
4º nivel: título publicación seriada
5º nivel: anos dixitalizados
5º nivel: anos dixitalizados
5º nivel: anos dixitalizados





4.2. Arquivos de preservación

A Biblioteca de Galicia, seguindo as normas e recomendacións internacionais, decidiu que os arquivos destinados á preservación nos proxectos de dixitalización sistemática vinculados á **Memoria Dixital de Galicia**, sexan obrigatoriamente os seguintes:

- Arquivos máster TIFF editados
- Ficheiros METS de preservación
- Ficheiros ALTO-XML
- De xeito complementario tamén se poderá solicitar a preservación dos arquivos máster TIFF en cru para poder garantir o proceso completo de obtención das imaxes en relación co sinalado no punto 3.1. deste documento. O nome do directorio neste caso será "tiff_master".
- Recoméndase tamén a entrega dos formatos derivados (JPG, MINIATURAS E PDF).

Materiais seriados e non seriados

As imaxes obtidas como resultado do proceso de dixitalización seleccionadas para a preservación en formato TIFF, así como os ficheiros de metadatos asociados, almacenaranse seguindo os niveis de directorios establecidos a continuación:

- 1º nivel – Directorio co nome do proxecto ou contrato
- 2º nivel – Directorio xeral do proxecto por institución da memoria
- 3º nivel – Directorios de cada obra dixitalizada
- 4º nivel – Directorios por formatos dos ficheiros de preservación

O **nomeado** deses directorios terá que cumprir as seguintes indicacións:

O directorio de 1º nivel, que identificará o nome que recibe o proxecto de dixitalización, será proporcionado pola Biblioteca de Galicia. Exemplo:

- **pr606a_2020_2021**

Os directorios de 2º nivel, que identifican a institución depositaria dos orixinais dixitalizados, aparecerá nomeada polo código que lle proporcione a Biblioteca de Galicia. Exemplo:

- Biblioteca de Galicia: **es-scbg**

Os directorios de 3º nivel, que identificarán cada obra dixitalizada, independentemente de que procedan da transformación a formato dixital de



materiais seriados ou non seriados, estarán nomeadas polo código da institución seguido dun código unívoco para cada obxecto dixital. No caso das publicacións seriadas considérase obxecto dixital cada número desas cabeceiras. Exemplo:

- **es-scbg_pug_19420103**

No caso de materiais non seriados, cada obxecto dixital estará nomeado polo código da institución depositaria acompañado da sinatura topográfica que identifique univocamente a obra orixinal. Exemplo:

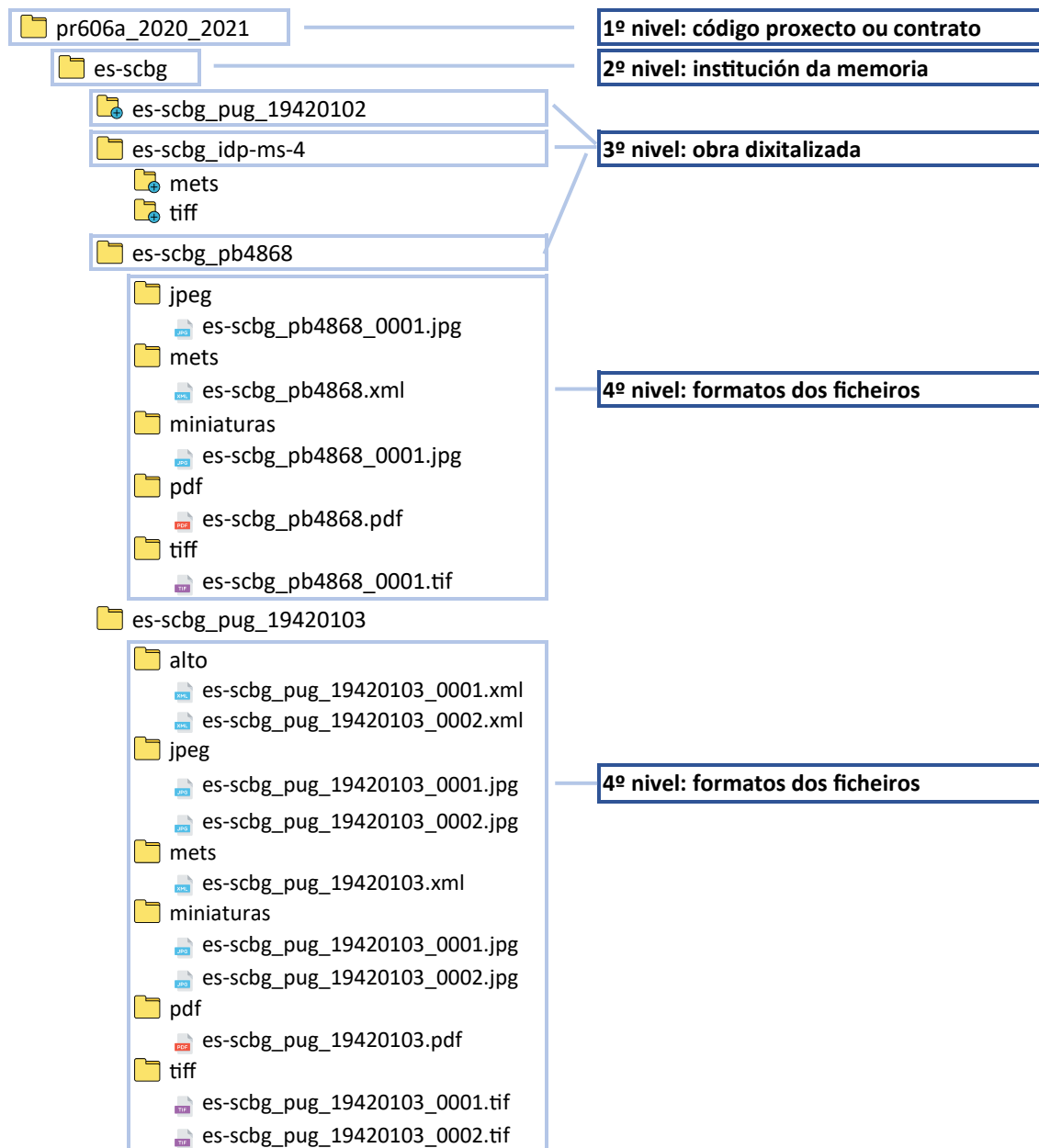
- **es-scbg_pb4868**

O 3º nivel de codificación na estrutura dos directorios de preservación terá que coincidir cos primeiros elementos establecidos para a respectiva identificación das imaxes establecida tanto para materiais seriados como non seriados. Por exemplo, a ruta para unha imaxe tiff nun ficheiro mets de preservación dun número de seriada e dunha monografía dixitalizada da Biblioteca de Galicia terá este camiño (estrutura de directorios) e nome do ficheiro da imaxe:

\es-scbg\es-scbg_pug_19420103\tiff\es-scbg_pug_19420103_0001.tif
\es-scbg\es-scbg_pb4868\tiff\es-scbg_pb4868_0001.tif



Exemplo





5. METADATOS

Os metadatos son o conxunto de informacións que representan ao obxecto dixital. Son as ferramentas que nos permiten especificar a información contextual asociada a cada documento dixital, isto é, o seu contido, os formatos e características de cada ficheiro, o historial tanto da xeración como das transformacións sufridas por cada copia dixital e os diferentes hardwares e softwares vinculados coa súa creación, edición e preservación.

A información dos obxectos dixitais do proxecto da [Memoria Dixital de Galicia](#) ten que estar codificada seguindo o esquema METS. Cada unha das obras dixitalizadas dentro deste proxecto terá que estar asociada a tres arquivos METS, dous deles destinados a DIFUSIÓN (un para a versión en JPG e outro para a versión en PDF) e un terceiro destinado a PRESERVACIÓN.

Os ficheiros METS de carga están configurados para a súa inxesta no repositorio de Galiciana-Biblioteca Dixital de Galicia, mentres que os ficheiros METS de preservación están destinados a súa importación no sistema de preservación da Memoria Dixital de Galicia. Ambos os dous arquivos deberán seguir as especificación que se expoñen a continuación.

5.1. Ficheiros METS para a inxesta no repositorio de Galiciana-BDG

Os perfís técnicos que deberán seguirse para os ficheiros METS de carga en Galiciana-BDG son diferentes en función do tipo de material seleccionado para a súa dixitalización. Nestes requirimentos especifícanse as pautas a seguir para a xeración destes ficheiros de metadatos para material non seriado e para publicacións periódicas.

As especificacións para os ficheiros METS asociados á dixitalización de material audiovisual e de gravacións sonoras proporcionaranse na seguinte versión que se publique destes requirimentos.

5.1.1. METS de carga de materiais non seriados en JPG

As seguintes recomendacións teñen como obxecto orientar na xeración de ficheiros de metadatos asociados aos recursos orixinados a partir da dixitalización de materiais non seriados, isto é:

- monografías
- manuscritos
- material cartográfico
- material gráfico (carteis, fotografías, postais, debuxos, ephemera, láminas e gravados)



- partituras

Os orixinais destes tipos de materiais que queiran formar parte do proxecto da **Memoria Dixital de Galicia** a través de **Galiciana-BDG** terán que estar descritos obrigatoriamente en formato MARC 21 bibliográfico tanto para a súa descrición bibliográfica, como de fondos e localizacións e tamén para as autoridades asociadas a cada rexistro.

A información bibliográfica e de fondos que empregará como base a empresa adxudicataria será a proporcionada polo Catálogo Colectivo de Patrimonio Bibliográfico e/ou o Catálogo Colectivo da Rede de Bibliotecas de Galicia. De non existir esa información nestes catálogos para os documentos seleccionados, a Biblioteca de Galicia proporcionará á empresa ou servizo de dixitalización os rexistros MARC 21 bibliográficos e de fondos cun nivel de catalogación suficiente para que poida realizar as tarefas contratadas. A empresa ou servizo de dixitalización utilizará estes rexistros para embebelos nos rexistros METS, tendo coidado de facer coincidir os números de control, códigos de biblioteca e sinaturas dos rexistros MARC21 e dos METS.

Galiciana-Biblioteca Dixital ten como lingua de catalogación o galego, polo tanto utilizarase obrigatoriamente esta lingua nos campos dos rexistros bibliográficos, de autoridades e de fondos que teñan que ir na lingua do centro catalogador. Así mesmo usarase o galego nos ficheiros METS de carga para o texto de enlace aos obxectos dixitais (atributo "descripcion"=Copia dixital), e de ser necesario, na designación da paxinación do mapa estrutural (p. ex. <div LABEL="[Cuberta anterior]" ORDER="1" TYPE="page">) para que estes datos se visualicen en galego na web de Galiciana-Biblioteca Dixital.

Galiciana-Biblioteca Dixital traballa con MARC21 bibliográfico, de fondos e de autoridades. Tamén se enriquecen os rexistros bibliográficos incluíndo os campos MARC21/RDA 264, 336, 337 e 338.

A empresa adxudicataria estará obrigada a utilización do formato MARC21 bibliográfico para describir as obras. Ademais é recomendable a utilización de rexistros de fondos e de autoridade en MARC21 e o seu enriquecemento segundo o indicado no parágrafo anterior. En todo caso, se non se utilizan rexistros de fondos, é preciso incluír a información de fondos en campos 85X do formato MARC21 bibliográfico. En concreto habería que utilizar o campo 852 co código de institución (no subcampo \$a) e a sinatura (no subcampo \$j). A Biblioteca de Galicia proporcionará o código da institución depositaria dos orixinais para incluír no subcampo \$a do 852, e tamén a sinatura para o \$j do mesmo campo.



A conversión deses rexistros bibliográficos e de fondos e localizacións ao esquema MARC-XML serán os **metadatos descritivos** do obxecto dixital que se embeberán no arquivo METS de carga.

Os obxectos dixitais derivados da dixitalización de calquera tipo de publicación non seriada irán asociados a ficheiros METS de carga que terán como perfil de referencia o *Perfil METS da Biblioteca Virtual de Patrimonio Bibliográfico* accesible en :

- ✓ <http://www.loc.gov/standards/mets/mets-registered-profiles.html>

Os requirimentos mínimos que deben cumprir os ficheiros METS de carga en JPG en **Galiciana-BDG** para este tipo de materiais son os seguintes:

- Incluir as direccións correctas aos namespaces e aos schemas empregados e deben estar ben formados e validados. Para garantir a integridade técnica dos ficheiros se realizará un proceso de caracterización con DROID e a validación con JHOVE, empregando en ambos casos a versión mais recente no momento da entrega.
- Levar unha sección de metadatos descritivos <dmdSec> que embeba o rexistro bibliográfico en formato MARC-XML e, a continuación, o rexistro de fondos en formato MARC-XML de fondos. Poderase substituír o rexistro MARC-XML de fondos por a información do exemplar no campo 852 embebido no rexistro bibliográfico. Esta primeira sección de metadatos descritivos (dmdSec) terá as seguintes etiquetas, atributos e valores:
 - Etiqueta mdWrap con atributos MDTYPE="MARC", que a súa vez conteña unha etiqueta xmlData que embeba o rexistro bibliográfico e o rexistro de fondos e localizacións.
- Incluir unha segunda sección de metadatos descritivos <dmdSec> na que se relacione a imaxe que vai servir como icona representativa da obra (normalmente unha portada ou ilustración). Neste segundo <dmdSec>, con atributo ID="DGBGOM", ten que conter as seguintes etiquetas, atributos e valores:
 - Etiqueta mdWrap con atributos MDTYPE="OTHER" e OTHERMDTYPE="DGBGOM"
 - Etiqueta xmlData
 - Etiqueta grupoObjetoMultimedia con atributos descripcion="Copia dixital", mimeType="image/jpeg", presentacionDef="miniaturas" e tieneOCR="true". Tamén terá un atributo nome cuxo valor será o título da obra seguido da sinatura entre corchetes.



- Os METS deben incluír unha sección de arquivo (fileSec), que nun único fileGrp con atributo USE="reference" agrupe elementos file para cada imaxe JPG das que conforman a obra. Dentro deste fileGrp, para cada elemento file, incluíranse en etiquetas FLocat as direccións ás imaxes correspondentes. Estas direccións apuntarán ás imaxes de difusión que se atopen nos mesmos discos duros que estes METS.
- Poderanse poñer dentro desta sección outros elementos fileGrp para referenciar versións da obra noutros formatos ou con outras características técnicas; para estes casos, utilizaranse atributos USE con valores que describan o tipo de versión (p.ex. USE="thumbnail" para as MINIATURAS). Dentro destes fileGrp incluíranse tantos elementos file como arquivos conformen a versión da obra. Para cada elemento file, incluíranse en etiquetas FLocat as direccións aos arquivos correspondentes. Estas direccións apuntarán aos arquivos (imaxes JPGs para miniaturas, etc.) que se atopen nos mesmos discos duros que estes METS.
- Os ficheiros resultantes do proceso de OCR codificado en ALTO XML, agruparanse nun <fileGRP> independente dentro do <fileSec>, asignándoselle ao seu atributo USE o valor "ocr".
- Cada elemento de agrupación <fileGrp> terá un elemento <file> por cada imaxe dixitalizada que compoña a copia dixital nos diferentes formatos. No elemento <file> indicaranse os atributos correspondentes ao formato (MIMETYPE), o seu tamaño (SIZE) e a data de creación (CREATED). Ademais, no elemento <Flocat> de cada elemento <file> especificarase o atributo LOCTYPE="URL" e o seu valor debe indicar o camiño (estrutura de directorios) e o nome do ficheiro imaxe.
- Incluírase nos METS unha sección de mapa estrutural (strucMap) con atributo TYPE="PHYSICAL", que reflecta a estrutura da copia en JPG. Para materiais textuais este mapa estrutural ha de conter unha etiqueta div con atributo TYPE="book" e da que colguen etiquetas div con atributo TYPE="page", unha para cada páxina. O div de primeiro nivel incluírá tamén un atributo LABEL cuxo valor sexa o título da obra, seguido da sinatura. Os div nivel de páxina incluírán á súa vez atributos ORDER cuxo valor sexa o número de orde da páxina e LABEL con información sobre a paxinación. Consignaranse como mínimo con estes atributos LABEL a portada e todas as páxinas numeradas. Estas etiquetas div nivel de páxinas embeberán etiquetas fptr que enlacen mediante atributos FILEID cos IDs de imaxes de páxina correspondentes da sección de arquivo (fileSec).



No **Anexo B** móstrase un exemplo comentado de arquivo METS de carga para material non seriado en JPG.

5.1.2. METS de carga de materiais non seriados en PDF

Os METS de carga para versións en PDF cinguiranse ás mesmas especificacións que os METS de carga para versións en JPG excepto polo indicado a continuación:

- Na segunda sección de metadatos descritivos (dmdSec), con atributo ID="DGBGOM", terá que reflectir a seguinte información:
-Etiqueta grupoObjetoMultimedia con atributos `descripcion="Versión PDF"`, `mimeType="application/pdf"`, `presentacionDef="abrir"`, `tieneOCR="false"` e `"nombre"`. O atributo `"nombre"` terá como valor o título da obra seguido da signatura entre corchetes, un espazo, guión, espazo e o texto *Versión PDF*.
- Deben incluír unha sección de arquivo (fileSec), que nun fileGrp con atributo USE="reference" e ID="FGX1" inclúa elementos file para cada arquivo PDF dos que conforman a obra (normalmente haberá un só PDF, salvo nos casos de obras en volumes, no que se xeraron un por volume, ou de obras especialmente longas, que se dividiron en partes). Dentro deste fileGrp, para cada elemento file, incluíranse en etiquetas FLocat as direccións aos PDFs correspondentes. Estas direccións apuntarán aos PDFs que se atopen nos mesmos discos duros que estes METS.
- A continuación da anterior incluírase unha sección de mapa estrutural (strucMap) con atributo TYPE="PHYSICAL", que reflicta a estrutura da copia en PDF.

No **Anexo C** móstrase un exemplo comentado de arquivo METS de carga para material non seriado en PDF.

5.1.3. METS de carga de publicacións seriadas en JPG

Os orixinais das publicacións seriadas seleccionados para formar parte do proxecto da **Memoria Dixital de Galicia** terán que estar descritos obrigatoriamente en formato MARC 21 bibliográfico. A información bibliográfica e de fondos que empregará como base a empresa adxudataria será a proporcionada polo Catálogo Colectivo de Patrimonio Bibliográfico e/ou o Catálogo Colectivo da Rede de Bibliotecas de Galicia. De non existir esa información nestes catálogos para as cabeceiras de prensa seleccionadas, a Biblioteca de Galicia proporcionará á empresa ou servizo de dixitalización os rexistros MARC 21 bibliográficos e de fondos cun nivel de catalogación suficiente

1 X será o número de orden que ocupa ese FileGrp no METS.



para que poida realizar as tarefas contratadas. A empresa ou servizo de dixitalización utilizará estes rexistros para embebelos nos rexistros METS, tendo coidado de facer coincidir os números de control, códigos de biblioteca e sinaturas dos rexistros MARC21 e dos METS.

Galicianiana-Biblioteca Dixital ten como lingua de catalogación o galego, polo tanto utilizarase obrigatoriamente esta lingua nos campos dos rexistros bibliográficos, de autoridades e de fondos que teñan que ir na lingua do centro catalogador. Así mesmo usarase o galego nos ficheiros METS de carga para o texto de enlace aos obxectos dixitais (atributo "descripcion"=Copia dixital) para que estes datos se visualicen en galego na web de Galicianiana-Biblioteca Dixital.

Os rexistros de fondos das publicacións periódicas obxecto de dixitalización, deberán incluír os campos 853 e 866 (según ISO 10324:1997/Z39,71:2006) coa información de periodicidade, numeración e fondos existentes. Tamén terán que incluír o campo 852 co código de institución (no subcampo \$a) e a sinatura (no subcampo \$j). A Biblioteca de Galicia proporcionará o código da institución depositaria dos orixinais para incluír no subcampo \$a do 852, e tamén a sinatura para o \$j do mesmo campo. No caso de non utilizar rexistros de fondos, é preciso incluír a información de fondos en campos 85X do formato MARC21 bibliográfico.

A conversión deses rexistros bibliográficos e de fondos e localizacións ao esquema MODS (Metadata Object Description Schema) serán os metadatos descritivos do obxecto dixital que se embeberán no arquivo METS de carga.

Os obxectos dixitais derivados da dixitalización de calquera tipo de publicación seriada teñen que ir asociados a ficheiros METS de carga que terán como perfil de referencia o *METS Profile for Historical Newspapers* accesible en :

- ✓ <http://www.loc.gov/standards/mets/test/ndnp/00000010.xml>

En particular, os METS para a carga de publicacións seriadas en JPG terán que seguir as seguintes especificacións:

- Xerarase un arquivo METS para cada número de prensa ou revista dixitalizado.
- Incluír as direccións correctas aos namespaces e aos schemas empregados e deben estar ben formados e validados. Para garantir a integridade técnica dos ficheiros se realizará un proceso de caracterización con DROID e a validación con JHOVE, empregando en ambos casos a versión mais recente no momento da entrega.



- Deben incluír unha sección de metadatos descritivos (dmdSec) con atributo ID="DMDFICHEROASOCIADO". Nesta sección poderase incluír a ligazón web (URL) a un ficheiro MARC da obra que se describe. Se non se puidese ou quixese incluír esta ligazón, poderase deixar esta sección baleira.

Exemplo de ficheiro de metadatos co primeiro dmdSec cuberto:

```
<dmdSec ID="DMDFICHEROASOCIADO">  
  <mdRef ID="MD1" LABEL="Registro MARC ISO-2709" LOCTYPE="URL"  
  MDTYPE="MARC"  
  xlink:href="http://www.empresa.com/digitalizacionbib_pro/es/consulta/  
  registro.cmd?control=BDG20080003142" xlink:type="simple"/>  
</dmdSec>
```

Exemplo de ficheiro de metadatos co primeiro dmdSec baleiro:

```
<mets:dmdSec ID="DMDEXT1" />
```

- A continuación desta sección de metadatos descritivos incluírase unha segunda sección deste tipo con atributo ID="DMDREGASOCIADO". Nesta sección incluírase o número de control do rexistro MARC como valor do atributo xlink:href da etiqueta mdRef. Se non se puidese ou quixese incluír esta ligazón, poderase deixar esta sección baleira.

Exemplo de ficheiro de metadatos co segundo dmdSec cuberto:

```
<dmdSec ID="DMDREGASOCIADO">  
  <mdRef ID="MD2" LABEL="Nº de control del registro MARC asociado"  
  LOCTYPE="OTHER" MDTYPE="MARC" OTHERLOCTYPE="MARC001BIB"  
  xlink:href="BDG20080003142" xlink:type="simple"/>  
</dmdSec>
```

Exemplo de ficheiro de metadatos co segundo dmdSec baleiro:

```
<mets:dmdSec ID="DMDEXT2" />
```

- A continuación destas dúas seccións de metadatos descritivos levarán unha terceira sección de metadatos descritivos (dmdSec) que embeba o rexistro bibliográfico en formato MODS (Metadata Object Description Schema), con información normalizada de numeración e cronoloxía segundo MARC 21 de fondos e localizacións.
- Aínda que o orixinal non mostre un día concreto de publicación pero si un ano e/ou mes, e non sexa posible deducilo pola periodicidade (p.ex. "se publica los lunes"), haberá que incluír obrigatoriamente unha data real na etiqueta mods:dateIssued para que o programa cree o número no



calendario, do contrario o mostrará como "sen data" e non o asignará no calendario a ningún ano e/ou mes.

Exemplo: unha publicación anual que non mostre o día nin o mes no que se publica, pero si o ano:

```
<mods:originInfo>
```

```
  <mods:dateIssued encoding="iso8601" qualifier="approximate">1927-12-31</mods:dateIssued>
```

```
</mods:originInfo>
```

- Embeberanse os datos da publicación baixo unha etiqueta `mods:relatedItem` con atributo `TYPE="host"`. Esta etiqueta `mods:relatedItem` embeberá, entre outras, unha etiqueta `mods:identifier` cuxo valor sexa o número de control do rexistro de exemplar do número que se está describindo no METS.
- Embeberase así mesmo na etiqueta `mods:relatedItem`, unha etiqueta `recordInfo` que embeba á súa vez unha etiqueta `recordIdentifier` cuxo valor sexa o número de control do rexistro bibliográfico e unha etiqueta `recordOrigin` cuxo valor sexa o código da biblioteca conservadora.
- Embeberase así mesmo na etiqueta `mods:relatedItem`, unha etiqueta `mods:location` que embeba á súa vez unha etiqueta `mods:holdingSimple` que embeba á súa vez unha etiqueta `mods:copyInformation`. Dentro desta última incluírase o patrón de numeración correspondente ao número que se está describindo como valor da etiqueta `mods:enumerationAndChronology`.
- Nas publicacións diarias a orde que se deberá seguir na etiqueta `mods:enumerationAndChronology` é ano/mes/día. É dicir, na etiqueta `mods:enumerationAndChronology` o subcampo `$i` deberá reflectir o ano, o `$j` o mes e `$k` o día.

Exemplo:

```
<mods:enumerationAndChronology unitType="1">marc:
```

```
$aAño$zacrn##$bNúmero$zaaan##$i(Año)$j(Mes)$k(Día)$wd$yodlu</mods:enumerationAndChronology>
```

- Da etiqueta `mods:relatedItem` colgará tamén unha etiqueta `mods:part` que inclúa a numeración e data de publicación de tal forma que respecte o máis posible a designación da numeración que consta na publicación de cara á descrición de cada número. A continuación incluíranse tantas etiquetas `mods:detail` como niveis de numeración (Tomo, Ano, Número...) teña o



número de prensa ou revista que se está describindo. Estas etiquetas conterán atributos "level" que indicarán o nivel de numeración (1, 2, etc.) e terán como valor a numeración correspondente ao nivel indicado. A continuación destas etiquetas incluírase de novo a data de publicación do número de prensa ou revista como valor da etiqueta mods:date. Esta data escribirase en formato aaaa-mm-dd, segundo o indicado na ISO 8601.

Exemplo:

```
<mods:part>
  <mods:detail level="1">
    <mods:number>XXIV</mods:number>
  </mods:detail>
  <mods:detail level="2">
    <mods:number>6542</mods:number>
  </mods:detail>
  <mods:date encoding="iso8601" qualifier="approximate">1940-10-
25</mods:date>
</mods:part>
```

- Utilizarase a etiqueta mods:note, embebida na etiqueta mods:relatedItem, para rexistrar información de edicións especiais, suplementos, numeracións especiais, e demais información de numeración e cronoloxía que se considere relevante e non se puxera noutros campos.

Exemplo:

```
<mods:note>Impresión e ordenación incorrecta das follas no orixinal,
corrixíuse</mods:note>
```

- Para as publicacións periódicas diarias con varias edicións asociadas ao mesmo día, haberá que facer dous ficheiros de metadatos por cada día, que se corresponderán con cada edición. Así o programa poderá xerar dous grupos de obxectos dixitais, que levarán nestes casos o texto de enlace (atributo "descripcion") "Primeira edición", "Segunda edición", "Edición Ferrol", "Edición Santiago de Compostela", ou a que mellor describa estas edicións.

Exemplo:

El Ideal Gallego, durante algún tempo tivo dúas edicións.

Ficheiro de metadatos para a 1ª edición: es-scbg_yde_19410103_1e.xml

```
<grupoObjetoMultimedia                               mimeType="image/jpeg"
presentacionDef="miniaturas" tieneOCR="true" nombre="El Ideal gallego
```



- Año XXV Número 6599 - 1941 xaneiro 3" descripcion="Primeira Edición">

Ficheiro de metadatos para a 2ª edición: es-scbg_yde_19410103_2e.xml

```
<grupoObjetoMultimedia                      mimeType="image/jpeg"
presentacionDef="miniaturas" tieneOCR="true" nombre="El Ideal gallego
- Año XXV Número 6599 - 1941 xaneiro 3"      descripcion="Segunda
Edición">
```

- Incluirán unha cuarta sección de metadatos descriptivos <dmdSec> na que se relacione a imaxe que vai servir como icona representativa do ítem de prensa (normalmente a cuberta ou a primeira páxina do primeiro número). Esta sección deberá incluír as seguintes etiquetas, atributos e valores:
-Etiqueta mdWrap con atributos MDTYPE="OTHER" e OTHERMDTYPE="DGBGOM"
-Etiqueta xmlData
-Etiqueta grupoObjetoMultimedia con atributos descripcion="Copia dixital", mimeType="image/jpeg", presentacionDef="miniaturas" e tieneOCR="true". Tamén terá un atributo "nombre" cuxo valor será o título da cabeceira de prensa seguido da numeración e data do número que se describe.

Exemplo:

```
<mets:dmdSec ID="DGBGOM">
  <mets:mdWrap MDTYPE="OTHER" OTHERMDTYPE="DGBGOM">
    <mets:xmlData>
      <grupoObjetoMultimedia mimeType="image/jpeg"
presentacionDef="miniaturas" tieneOCR="true" nombre="El Miño diario
liberal - Año VIII Número 1832 - 1905 xaneiro 4" descripcion="Copia dixital">
        <imagenFavorita>jpg_ou-sc-bmo_elmi_19050104_0001</imagenFavorita>
      </grupoObjetoMultimedia>
    </mets:xmlData>
  </mets:mdWrap>
</mets:dmdSec>
```

- Os arquivos METS deben incluír unha sección de arquivo (fileSec), que nun fileGrp con atributo USE="reference" agrupe elementos file para cada imaxe JPG das que conforman o número de publicación seriada que se describe. Dentro deste fileGrp, para cada elemento file, incluíranse en etiquetas FLocat as direccións ás imaxes correspondentes. Estas direccións apuntarán ás imaxes de difusión que se atopan nos mesmos discos duros que estes METS.



- Os ficheiros resultantes do proceso de OCR codificado en ALTO XML, agruparanse noutro <fileGRP> independente, asignándoselle ao seu atributo USE o valor "ocr".
- Cada elemento de agrupación <fileGrp> terá un elemento <file> por cada imaxe dixitalizada que compoñe a copia dixital nos diferentes formatos. No elemento <file> indicaranse os atributos correspondentes ao formato (MIMETYPE), o seu tamaño (SIZE) e a data de creación (CREATED). Ademais, no elemento <Flocat> de cada elemento <file> especificarase o atributo LOCTYPE="URL" e o seu valor debe indicar o camiño (estrutura de directorios) e o nome do ficheiro imaxe.
- Incluírse unha sección de mapa estrutural (strucMap) con atributo TYPE="LOGICAL", que reflicta a estrutura da copia en JPG do número que se está a describir e que estableza a relación entre as imaxes JPG e os seus correspondentes ficheiros ALTO XML. O div de primeiro nivel incluírá tamén un atributo LABEL cun valor que coincide co título do material seriado ao que fai referencia o obxecto dixital, seguido da designación da numeración e cronoloxía do número que se está a describir. Os div nivel de páxina incluírán á súa vez atributos ORDER e LABEL. Este atributo LABEL de nivel de páxina reflexará o número correlativo de cada páxina dentro do número que se está describindo.

Exemplo:

```
<mets:structMap ID="SM-1" TYPE="LOGICAL">
  <mets:div TYPE="news:issue" LABEL="Boletín bibliográfico de la librería
gallega - Año II Número 15 Y 16 - 1896 marzo 1 - Copia dixital"
DMDID="DMDISSUE" ORDER="1">
    <mets:div TYPE="news:page" LABEL="1">
        <mets:div TYPE="news:image" LABEL="1" ORDER="1">
            <mets:fptr FILEID="jpg_es-scbg_bblg_18960301_0001" />
        </mets:div>
        <mets:div TYPE="news:alto" LABEL="1">
            <mets:fptr FILEID="alto_es-scbg_bblg_18960301_0001" />
        </mets:div>
    </mets:div>
    <mets:div TYPE="news:page" LABEL="2">
        <mets:div TYPE="news:image" LABEL="2" ORDER="2">
            <mets:fptr FILEID="jpg_es-scbg_bblg_18960301_0002" />
        </mets:div>
        <mets:div TYPE="news:alto" LABEL="2">
            <mets:fptr FILEID="alto_es-scbg_bblg_18960301_0002" />
        </mets:div>
    </mets:div>
</mets:structMap>
```



</mets:div>

No **Anexo D** móstrase un exemplo comentado de arquivo METS de carga para publicacións periódicas en JPG.

5.1.4. METS de carga de publicacións seriadas en PDF

Os METS de carga de números de publicacións seriadas en PDF cinguiranse ás mesmas especificacións que os METS de carga de números de publicacións seriadas en JPG excepto polo indicado a continuación:

- Deben incluír unha sección de metadatos descritivos (dmdSec) con atributo ID="DGBGOM" que conteña as seguintes etiquetas, atributos e valores:
-Etiqueta mdWrap con atributos MDTYPE="OTHER" e OTHERMDTYPE="DGBGOM"
-Etiqueta xmlData
-Etiqueta grupoObjetoMultimedia con atributos descripcion="Versión PDF", mimeType="application/pdf", presentacionDef="abrir", tieneOCR="false" e "nombre". O atributo "nombre" terá como valor o título da cabeceira de prensa seguido da numeración e/ou data do número que se describe, un espazo, guión, espazo e o texto *Versión PDF*.
- Nos METS de carga para PDF, non será preciso utilizar a etiqueta imagenFavorita.

Exemplo:

```
<dmdSec ID="DGBGOM">
  <mdWrap MDTYPE="OTHER" OTHERMDTYPE="DGBGOM">
    <xmlData>
      <grupoObjetoMultimedia      descripcion="Versión      PDF"
      mimeType="application/pdf" nombre="Industria conservera órgano de la
      Unión de Fabricantes de conservas de Galicia - Año XI Número 78 - 1945
      agosto 1 – Versión PDF" presentacionDef="abrir" tieneOCR="false">
    </grupoObjetoMultimedia>
  </xmlData>
</mdWrap>
</dmdSec>
```

- Os arquivos METS deben incluír tamén unha sección de arquivo (fileSec) que nun único fileGrp con atributo USE="reference" e ID="FGX²" inclúa elementos file para cada arquivo PDF dos que conforman a obra (normalmente haberá un só PDF). Dentro deste fileGrp, para cada elemento

² X será o número de orden que ocupa ese FileGrp no METS.



file, incluíranse en etiquetas FLocat as direccións aos PDFs correspondentes. Estas direccións apuntarán aos PDFs que se atopen nos mesmos discos duros que os METS.

- A continuación da anterior incluírase unha sección de mapa estrutural (strucMap) con atributo TYPE="LOGICAL", que reflicta a estrutura da copia en PDF.

No **Anexo E** móstrase un exemplo comentado de arquivo METS de carga para material seriado en PDF.

5.2. Ficheiros METS para a inxesta no repositorio de preservación da MDG

Todas as obras dixitalizadas que formen parte da **Memoria Dixital de Galicia** terán que levar asociado un arquivo METS de preservación para a súa inxesta no sistema de preservación dixital da Xunta de Galicia.

Tal e como se expón no punto 4.2 deste documento, para os procesos de dixitalización sistemática a Biblioteca de Galicia tomou a decisión de preservar:

- Arquivos máster TIFF editados
- Ficheiros METS de preservación
- Ficheiros ALTO-XML
- Formatos derivados (JPG, MINIATURAS E PDF).
- De xeito complementario tamén se poderá solicitar a preservación dos arquivos máster TIFF en cru para poder garantir o proceso completo de obtención das imaxes en relación co sinalado no punto 3.1. deste documento. O nome do directorio neste caso será "tiff_master".

Para garantir a integridade técnica dos ficheiros se realizará un proceso de caracterización con DROID e a validación con JHOVE, empregando en ambos casos a versión máis recente no momento da entrega.

Os arquivos METS de preservación para a inxesta no repositorio de preservación da **Memoria Dixital de Galicia** deberán cumprir, como mínimo, as seguintes características:

- Incluirán os metadatos seleccionados pola Subdirección Xeral de Bibliotecas que describan os termos que se precisan para a preservación da copia dixital. Estes metadatos obtivéronse do Dicionario de Datos PREMIS, na súa versión vixente á firma do contrato, denominado PREMIS Data Dictionary, e elaborado segundo o Modelo de Referencia do Open Archival Information System (OAIS), ISO14721.



- Das cinco entidades ás que pode referirse un arquivo de preservación segundo PREMIS, nos ficheiros de preservación utilizaranse as seguintes: axente, obxecto e evento. Estas unidades semánticas estarán embebidas na sección de metadatos administrativos do ficheiro METS. Obrigatoriamente o METS incluírá os *eventType*³ cap-capture, mig-migration e val-validation
- Farán referencia a información necesaria para a preservación dos arquivos máster TIFF produto da dixitalización. A información deste formato de imaxe irá nun <fileGrp> da Sección Arquivo (fileSec) cos enlaces as imaxes correspondentes. O valor do atributo USE para os ficheiros máster será "archive".
- Os ficheiros resultantes do proceso de OCR codificado en ALTO XML, agruparanse noutro <fileGRP> independente, asignándosele ao seu atributo USE o valor "ocr".
- Incluirán unha sección de metadatos descritivos codificados, como mínimo, nos esquemas MARC 21 XML ou MODS, segundo o tipo de material, para a identificación dos obxectos dixitais.
- Entregaranse no ficheiro XML os metadatos de dereitos de autor e propiedade intelectual dos obxectos dixitais, estruturados segundo o esquema METSRights. Neles detallaranse as características de dereito de uso das publicacións segundo determine a **Memoria Dixital de Galicia**.

No **Anexo F** móstrase un exemplo comentado de arquivo METS para a inxesta no repositorio de preservación da **Memoria Dixital de Galicia**.

3 <https://id.loc.gov/vocabulary/preservation/eventType.html>



ANEXOS

Anexo A. Cadro resumen requisitos técnicos das imaxes

Consultar o documento na seguinte ligazón:

https://amtega.xunta.gal/sites/default/files/pdf/gal_bib_AnexoA_ResumoTecnicoImaxes.pdf

Anexo B. Exemplo comentado de METS para a carga de material non seriado en jpg

Consultar o documento na seguinte ligazón:

https://amtega.xunta.gal/sites/default/files/pdf/gal_bib_AnexoB_CargaMono_jpg.pdf

Anexo C. Exemplo comentado de METS para a carga de material non seriado en pdf

Consultar o documento na seguinte ligazón:

https://amtega.xunta.gal/sites/default/files/pdf/gal_bib_AnexoC_CargaMono_pdf.pdf

Anexo D. Exemplo comentado de METS para a carga de publicacións seriadas en jpg

Consultar o documento na seguinte ligazón:

https://amtega.xunta.gal/sites/default/files/pdf/gal_bib_AnexoD_CargaSerie_jpg.pdf

Anexo E. Exemplo comentado de METS para a carga de publicacións seriadas en pdf

https://amtega.xunta.gal/sites/default/files/pdf/gal_bib_AnexoE_CargaSerie_pdf.pdf

Anexo F. Exemplo de METS de preservación para a Memoria Dixital de Galicia

Consultar o documento na seguinte ligazón:

https://amtega.xunta.gal/sites/default/files/pdf/gal_bib_AnexoF_CargaPreservacion.pdf



BIBLIOGRAFÍA

Recomendaciones para proyectos de digitalización de patrimonio bibliográfico y fotografía histórica – versión de prepublicación

<http://hdl.handle.net/10421/9200>

Directrices para proyectos de digitalización de colecciones y fondos de dominio público, en particular para aquellos custodiados en bibliotecas y archivos

<http://hdl.handle.net/10421/3342>

Requisitos técnicos de los proyectos de digitalización de patrimonio bibliográfico y de prensa histórica de la SGCB

<http://hdl.handle.net/10421/8981>

Proceso de digitalización en la Biblioteca Nacional de España. Biblioteca Digital Hispánica

<http://www.bne.es/webdocs/Catalogos/ProcesoDigitalizacionBNE.pdf>