

iArxiu: Estructura i creació de Paquets d'Informació de Transferència (PIT) utilitzant el model METS

Índex

Índex.....	2
Informació general.....	3
Els participants	3
Glossari.....	4
1. Introducció	5
1.1 El concepte de paquet d'informació	5
1.2 Paquets d'informació de transferència	6
2. Estructura general METS	7
3. Creació d'un METS	9
3.1 Capçalera del fitxer METS	9
3.2 File Section	9
3.2.1 Incrustats en Base64.....	9
3.2.2 Referència a fitxers	10
3.3 Structural Map Section	11
3.4 Descriptive Metadata Section.....	12
3.5 Administrative Metadata Section	14
4. Vocabularis de metadades per iArxiu.....	16
4.1 Vocabulari d'expedient	16
4.2 Vocabulari de document integrat en expedient.....	18
4.3 Vocabulari de document independent	19
4.4 Vocabulari de signatura.....	20
4.5 Altres vocabularis	22
5. Exemple PIT segons el model METS	23
5.1 Paquet amb estructura d'expedient.....	23
5.2 Paquet amb estructura de document independent.....	25
6. Implementació tècnica	27
7. Annexes.....	28
7.1 Taula d'il·lustracions	28

Informació general

Estat formal	Elaborat per:	Aprobat per:
	Equip del projecte iArxiu de CATCert	Grup d'Innovació Tecnològica de la Subdirecció General d'Arxius i Gestió Documental
Data de creació	04/05/09	
Control de versions	Data:	17/12/2008
	Descripció:	Actualització URN i namespace
	Data:	21/04/2009
	Descripció:	Actualització URN, metadada document_essencial, signatures embolcallants i incrustades, eliminació PREMIS
Nivell accés informació	pública	
Títol	iArxiu: Estructura i creació de Paquets d'Informació de Transferència (PIT) utilitzant el model METS	
Fitxer	iArxiu v2 0 - Estructura_PIT_ METS_4.3_20090504.doc	
Drets d'autor	Aquesta obra està subjecta a una llicència Reconeixement-No comercial-Sense obres derivades 2.5 Espanya de Creative Commons. Per veure'n una còpia, visiteu http://creativecommons.org/licenses/by-nc-nd/2.5/es/deed.ca o envieu una carta a Creative Commons, 171 Second Street, Suite 300, San Francisco, California 94105, USA.	

Els participants

El present document ha estat elaborat per l'equip de treball del projecte iArxiu de CATCert, comptant amb l'assessorament i la col·laboració de la Subdirecció General d'Arxius i Gestió Documental del Departament de Cultura i Mitjans de Comunicació de la Generalitat de Catalunya, a través del Grup d'Innovació Tecnològica, que té per missió oferir pautes i criteris per a la gestió de la documentació electrònica en el marc de les polítiques d'e-Governance que impulsa el govern de la Generalitat de Catalunya.

Per part de l'equip de treball del projecte iArxiu les persones que han participat en la redacció del document han estat les següents:

- Marta Cruellas Ibarz
- Joan Mir Rubio
- Raimon Nualart Mercadé
- Ralf Irion (HP)

Els integrants del Grup d'Innovació Tecnològica que han participat en tasques d'assessorament en la redacció del document i en la validació del seus continguts han estat les següents:

- Ferran Agelet, Ajuntament de Barcelona
- Ignacio Alamillo, Secretaria de Telecomunicacions i Societat de la Informació
- Gemma Bayó, Ajuntament de Barcelona
- Lluís-Esteve Casellas, Ajuntament de Girona
- Lluís Cermeno, Subdirecció General d'Arxius i Gestió Documental
- Anna Lorente, Agència Catalana de l'Aigua
- Betlem Martínez, Departament d'Economia i Finances
- Remei Perpinyà, Universitat Autònoma de Barcelona (UAB)
- Julio Quílez, Departament de Justícia, Arxius Judicials de la DGRAJ
- Tomàs Roy, Centre de Telecomunicacions i Tecnologies de la Informació (CTTI)
- Jaume Sardà, Subdirecció General d'Arxius i Gestió Documental
- Jordi Serra, Departament de la Vicepresidència

A tots ells CATCert els vol donar les gràcies per la seva col·laboració i participació en el projecte iArxiu, i en especial en la validació d'aquest document.

Glossari

CATCert: Agència Catalana de Certificació

CMS: Cryptographic Message Syntax. Format estàndard que s'utilitza per a signar electrònicament, autenticar a les parts, xifrar o generar resum de continguts (dades)

METS: Metadata Encoding & Transmission Standard

OAIS: Open archival information systems

PIA: Paquet d'informació d'Arxiu

PIC: Paquet d'informació de Consulta

PIT: Paquet d'informació de Transferència

PREMIS: PREservation Metadata Implementation Strategies)

PRONOM: Registre on-line sobre formats tecnològics desenvolupat per l'Arxiu Nacional del Regne Unit (TNA)

URI: Uniform Resource Identifier. Cadena de caràcters utilitzada per identificar un recurs o nom.

XSD: XML Schema Definition

1. Introducció

1.1 El concepte de paquet d'informació

Per entendre millor el funcionament del servei d'iArxiu i la gestió de la conservació dels documents a llarg termini cal explicar, encara que sigui breument, què entenem per documentació electrònica. En concret, ens estem referint tant a un document electrònic, sigui estructurat o no, com a agrupacions de documents o expedients electrònics, conservats com informació i prova per una organització en el desenvolupament de les seves activitats. Així mateix, un element inseparable i essencial per al tractament i la preservació dels documents electrònics és el de les metadades que descriuen el context, el contingut i l'estructura dels documents, i ens permeten garantir la seva autenticitat, integritat i disponibilitat al llarg del seu cicle de vida. Per tant, quan parlem de documents ens referim tant als documents com objectes d'informació (contingut), com a les metadades que ens ajuden a interpretar i gestionar correctament aquests documents.

Un requisit imprescindible per garantir la preservació de la documentació electrònica és definir, modelar i implementar una forma d'estructurar i agrupar tots els components que són objecte de conservació en un contenidor neutral. Si unim en un mateix contenidor els documents o objectes digitals (fitxers) i les seves metadades, es crea una entitat que pot existir en qualsevol entorn informàtic, ja que inclou tota la informació necessària per a accedir i representar els documents. El format XML és el més adient per a representar i emmagatzemar els documents i les metadades. Seguint la terminologia OAIS¹ existeixen tres tipus d'entitats, anomenades contenidors o paquets d'informació, segons la fase del cicle de vida del document:

1. Paquet d'informació de transferència (PIT).
2. Paquet d'informació d'arxiu (PIA).
3. Paquet d'informació de consulta (PIC).

En primer lloc, ens trobem els documents transferits pels productors de la documentació, que inclouen els continguts a preservar i totes aquelles metadades que faciliten la seva recuperació, el seu tractament i la seva preservació. Els anomenem *paquets d'informació de transferència* (PIT). Quan aquests ingressen a l'arxiu se'ls aplica un conjunt de controls per verificar la seva integritat, autenticitat i idoneïtat amb els requeriments de conservació. Se'ls afegixen altres metadades (per exemple, les relacionades amb la informació de representació) i, si es necessari, i així ho estableix la política de preservació, es pot procedir a la migració a un format més estàndard i estable. A aquesta normalització del PIT se l'anomena *paquet d'informació d'arxiu* (PIA). El PIA s'emmagatzemarà al repositori d'iArxiu el temps que s'estableixi. Quan un usuari vulgui consultar algun document emmagatzemat, el sistema li permetrà obtenir una còpia del PIA més coneguda com a *paquet d'informació de consulta* (PIC).

Davant l'absència d'un estàndard normalitzat d'estructuració i empaquetament de documents electrònics en format XML, ens hem decidit a utilitzar l'esquema METS² que és un model amb una àmplia acceptació, les característiques del qual s'adeqüen molt a les necessitats d'iArxiu per a la

¹ Per a més informació consulteu: ISO 14721:2003 - Space data and information transfer systems. Open archival information systems. Reference model.

Per a més informació sobre el model OAIS desenvolupat pel Consultative Comitè for Space Data Systems de la NASA i que posteriorment es va convertir en la norma ISO 14721:2003 consulteu a: <http://public.ccsds.org/publications/archive/650x0b1.pdf>

² L'esquema METS és un model per a codificar metadades descriptives, administratives i estructurals, d'objectes digitals emmagatzemats en un repositori, en format XML. Consulteu: <http://www.loc.gov/standards/mets/>

fase II. Per aquesta raó, s'ha optat per fer servir aquest com a format d'ingrés i consulta de paquets d'informació. Els motius que ens han dut a escollir METS són els següents:

- Permet estructurar els documents d'una forma flexible que s'adapta a les diferents maneres d'agrupar la documentació (document o expedient).
- Admet diferents diccionaris o vocabularis de metadades.
- Permet el creixement del paquet a mesura que s'apliquen processos de migració als fitxers.
- Permet especificar els fitxers, tant de forma incrustada en base64 com vinculats externament.
- Permet generar un paquet d'informació autònom i documentat en format XML, independent del sistema, i desproveït de les dependències tecnològiques de l'aplicació que el va generar. Inclou:
 - a. Els continguts a preservar.
 - b. Metadades descriptives, administratives i estructurals.

1.2 Paquets d'informació de transferència

Abans de procedir a la transferència de documents entre el sistema productor i iArxiu, el productor ha d'agrupar la documentació (fitxers més metadades) que vol transferir en *paquets d'informació de transferència (PIT)*. L'objectiu del present document és definir el protocol de creació i estructura de PIT seguint el model METS, quan la transferència a iArxiu sigui via webservice. La idea és representar en format XML el contingut dels diferents expedients/documents i que sigui compatible amb el model de PIA d'iArxiu

L'objectiu del present document és desenvolupar una guia que faciliti als productors de la documentació (gestors documentals, gestors d'arxiu, etc.) la creació d'aquests paquets d'informació de transferència.

El document no inclou les diferents operacions de webservice amb iArxiu: operació de transferència, de consulta estat de transferència o de consulta. Tampoc s'inclouen els requeriments o característiques tècniques del client webservice entre l'aplicació i iArxiu. Tots aquests punts es tracten en un altre document.

2. Estructura general METS

L'estructura d'un fitxer METS es compon de diverses seccions, de les quals farem servir les següents:

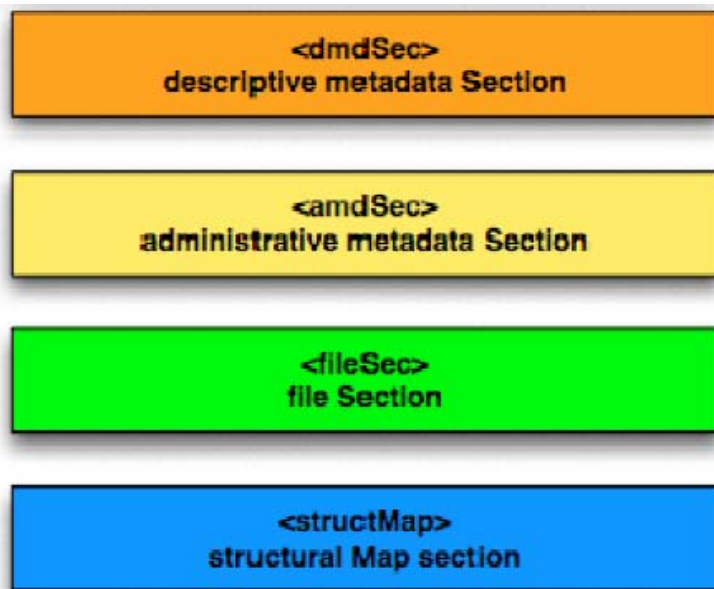


Figura 1: Estructura METS en blocs

Aquesta estructura s'adequa molt bé a les nostres necessitats pels següents motius:

- Permet crear una estructura jeràrquica d'elements (amb carpetes) dins del paquet.
- Permet associar metadades administratives i descriptives, tant a elements estructurals com a fitxers binaris.
- Permet especificar fitxers binaris, tant de forma incrustada com vinculada externament.

A continuació es detallen les seccions:

- Descriptive Metadata Section (<dmdSec>)

Aquesta secció conté tots els vocabularis descriptius que formen part del paquet, independentment del nivell (expedient, document) en què es trobin. Cadascun d'aquests vocabularis porta un identificador únic. Les metadades descriptives les formen uns vocabularis que tenen l'objectiu de descriure i identificar els paquets d'informació i serveixen per a la localització i recuperació de la informació, així com per a entendre el context del paquet.

- Administrative Metadata Section (<amdSec>)

Aquesta secció conté tots els vocabularis administratius que formen part del paquet, independentment del nivell (expedient, document, binari) en què es trobin. Cadascun d'aquests vocabularis porta un identificador únic. Les metadades administratives també les formen uns vocabularis, però no tenen la funció de descriure el paquet, sinó d'administrar-lo per garantir un tractament correcte al llarg del seu cicle de vida (preservació,

representació i ressegellat). Són metadades que faciliten la gestió i el processament dels paquets a llarg termini, la qual cosa cobreix una extensa gama d'informació com ara informació tècnica (formats i especificacions tecnològiques) necessària per mostrar i representar els objectes, metadades de provenença i metadades de cicle de vida i d'accés.

- File Section (<fileSec>)

Aquesta secció conté tots els fitxers binaris que formen part del paquet. Cadascun d'aquests fitxers binaris porta un identificador únic i pot estar referenciat amb una URI o estar incrustat en base64.

- Structural Map Section (<structMap>)

Aquesta secció ens permet estructurar els continguts de forma jeràrquica, tot permetent la creació de carpetes per agrupar els diferents fitxers binaris.

En aquesta mateixa secció també es vinculen els vocabularis de metadades descrits a les seccions de metadades amb les carpetes o fitxers binaris.

3. Creació d'un METS

Per construir un fitxer METS, s'ha de crear un fitxer XML amb les seccions descrites al punt anterior. Tot i que al fitxer METS (XML) creat s'han d'organitzar les diferents seccions en l'ordre indicat, començarem a definir la secció de fitxers (<fileSec>).

Als fragments XML que apareixen als punts següents apareixen certs espais de noms (namespaces) que, en el moment de redactar aquest document, són provisionals. Els espais de noms definitius es definiran durant l'anàlisi tècnic de l'iArxiu.

3.1 Capçalera del fitxer METS

No farem servir el tag *Hdr* que proporciona l'esquema METS, però sí que posarem certa informació al tag arrel (mets):

- L'atribut *TYPE* ha de contenir l'identificador de plantilla iArxiu que es farà servir per ingressar el fitxer (*urn:iarxiu:2.0:templates:catcert:PL_expedient* a l'exemple de sota). La plantilla es fa servir per poder aplicar certes validacions als paquets
- Informar namespace i opcionalment schemaLocation de l'esquema METS, tal com s'indica al següent fragment:

```
<?xml version="1.0" encoding="UTF-8"?>
<METS:mets xmlns:METS="http://www.loc.gov/METS/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd"
TYPE="urn:iarxiu:2.0:templates:catcert:PL_expedient">
...
```

Nota: Els identificadors de les plantilles definitives dependran de la configuració definitiva de la plataforma iArxiu un cop aquesta estigui en producció.

3.2 File Section

3.2.1 Incrustats en Base64

En aquesta secció s'han d'enumerar els diferents binaris que volem incloure al paquet. Suposem, per exemple, que el paquet es compona d'un binari en format Word i d'una signatura CMS. A més, tots dos s'informen en base64 (per simplificar, s'ha reduït molt la llargada de les codificacions b64, tal i com es pot observar a continuació):

```
<METS:fileSec>
<METS:fileGrp ID="BIN_1" >
<METS:file ID="BIN_1.0" MIMETYPE="application/msword"
CHECKSUM="99800b85d3383e3a2fb45eb7d0066a4879a9dad0" CHECKSUMTYPE="SHA-1"
CREATED="2008-05-17T09:30:47.0Z" >
<METS:FContent>
<METS:binData>aG9sYQ==</METS:binData>
</METS:FContent>
```

```

</METS:file>
</METS:fileGrp>
<METS:fileGrp ID="BIN_2" >
  <METS:file ID="BIN_2.0" MIMETYPE="application/pkcs7-signature"
    CHECKSUM="9788b4dffb7f0d350edabbf10d325b47e350def" CHECKSUMTYPE="SHA-1"
    CREATED="2008-05-17T09:30:47.0Z">
    <METS:FContent>
      <METS:binData>aG9sYSBvdHJhIHZleg==</METS:binData>
    </METS:FContent>
  </METS:file>
</METS:fileGrp>
</METS:fileSec>

```

Tal i com es pot observar, s'han d'indicar obligatòriament:

- Identificadors dels fitxers (*bin_1* i *bin_2* a l'exemple): aquests són lliures i es pot seguir qualsevol terminologia, sempre que no es repeteixin dins del mateix METS.
- MimeTypes: MimeTypes segons qualsevol registre de MimeTypes oficial, com per exemple el de PRONOM³ o iana⁴ ("application/msword" i "application/pkcs7-signature" a l'exemple).
- Resum del fitxer (checksum) i algorisme (SHA-1). El resum s'ha de codificar en Hexadecimal, de forma que com és un SHA-1 tindrà una llargària de 40 caràcters.
- Contingut binari en base64.
- Data de creació (atribut *CREATED*). Aquesta data és la data en que s'està incloent el fitxer al METS (no la data de creació del document). Es pot posar la data actual (durant la creació del fitxer METS).

3.2.2 Referència a fitxers

En cas de no voler incrustar els fitxers en Base64 dins del METS, es poden enviar els binaris per separat i indicar una referència des del fitxer METS de la següent forma:

```

</METS:fileGrp>
<METS:fileGrp ID="bin_cms_file" >
  <METS:file ID="bin_cms_file1.0" MIMETYPE="application/pkcs7-signature"
    CHECKSUM=" 9788b4dffb7f0d350edabbf10d325b47e350d" CHECKSUMTYPE="SHA-1"
    CREATED="2008-05-17T09:30:47.0Z">
    <METS:FLocat xlink:href="dir_exp/doc_doc1/sig55.cms" LOCTYPE="URL"/>
  </METS:file>
</METS:fileGrp>

```

L'atribut *xlink:href* ha de contenir la ruta del fitxer relativa a l'arrel del paquet, que és on es troba el fitxer METS.XML.

Per fer l'enviament del paquet a iArxiu, s'ha de crear un fitxer ZIP que contingui el fitxer METS.XML i tots els binaris referenciats.

³ Per a més informació consulteu a: <http://www.nationalarchives.gov.uk/pronom/>

⁴ Per a més informació consulteu a: <http://www.iana.org/assignments/media-types/>

3.3 Structural Map Section

En aquesta secció s'ordenen els binaris dins d'una estructura d'expedient. En l'exemple de sota crearem un expedient que només conté un document que agrupa el fitxer Word i la signatura associada:

```
<METS:structMap>
  <METS:div LABEL="expedient" DMDID="DMD_EXP">
    <METS:div LABEL="document_1" DMDID="DMD_DOC1">
      <METS:div LABEL="document.doc">
        <METS:fptr FILEID="BIN_1"/>
      </METS:div>
      <METS:div LABEL="firma.p7b" ADMID="AMD_SIG">
        <METS:fptr FILEID="BIN_2"/>
      </METS:div>
    </METS:div>
  </METS:div>
</METS:structMap>
```

En aquest fragment tenim 3 nivells de jerarquització:

- expedient (entitat lògica)
 - o document_1 (entitat lògica)
 - fitxers binaris continguts (file)

Per als casos de documents simples que no pertanyen a un expedient, no existeix el nivell d'expedient, sent el primer nivell el de document:

```
<METS:structMap>
  <METS:div LABEL="document_1" DMDID="DMD_DOC1">
    <METS:div LABEL="document.doc">
      <METS:fptr FILEID="BIN_1"/>
    </METS:div>
    <METS:div LABEL="firma.p7b" ADMID="AMD_SIG">
      <METS:fptr FILEID="BIN_2"/>
    </METS:div>
  </METS:div>
</METS:structMap>
```

En aquest fragment tenim 2 nivells de jerarquització:

- document_1 (entitat lògica)
 - o fitxers binaris continguts (file)

S'han de informar obligatòriament:

- A l'últim nivell s'han d'informar les referències als binaris definits a la secció *<fileSec>* amb l'atribut FILEID (*bin_1* i *bin_2*)
- Amb l'atribut DMDID s'han d'identificar els vocabularis de metadades descriptives. En aquesta part només s'indiquen els identificadors del vocabulari. El contingut de les metadades descriptives s'indica a la secció *<dmdSec>*. Anàlogament a les metadades descriptives, l'atribut AMDID conté els identificadors dels vocabularis administratius, informats a la secció *<amdSec>*. Es pot observar que, tant a nivell d'expedient com a nivell

de document, existeixen metadades descriptives (DMD_EXP i DMD_DOC1), mentre que a nivell de binari les metadades són administratives (metadades de la signatura).

- Amb els atributs LABEL es donen noms als diferents elements de l'estructura. Aquests noms apareixen a la consulta de paquet i permeten identificar els diferents documents:
 - o LABEL a nivell d'expedient permet posar un nom a l'expedient
 - o LABEL a nivell de document permet anomenar els diferents documents
 - o LABEL a nivell de fitxer permet anomenar els fitxers. És recomanable posar noms complets incloent l'extensió ("document.doc" a l'exemple de dalt)

Nota: Els identificadors de les plantilles definitives dependran de la configuració definitiva de la plataforma iArxiu un cop aquesta estigui en producció.

Nota: En cas de que no es vulguin informar binaris, es pot eliminar el nivell de fitxer binari deixant la estructura d'aquesta forma:

```
<METS:structMap>
  <METS:div LABEL="expedient" DMDID="DMD_1">
    <METS:div LABEL="document_1" DMDID="DMD_DOC1">
      </METS:div>
    </METS:div>
  </METS:structMap>
```

3.4 Descriptive Metadata Section

Com ja s'ha esmentat anteriorment, en aquesta secció s'informen els valors dels vocabularis descriptius. Al nostre exemple, hi ha dos vocabularis descriptius: DMD_EXP i DMD_DOC1:

```
<METS:dmdSec ID="DMD_EXP">
  <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_expedient" MIMETYPE="text/xml">
    <METS:xmlData>
      <exp:expedient xmlns:exp="http://schemas.user.iarxiu.hp.com/2.0/Voc_expedient">
        <exp:codi_referencia>CAT/AACC/D1161_N-185_05</exp:codi_referencia>
        <exp:numero_expedient>185/08</exp:numero_expedient>
        <exp:codi_classificacio>G0200</exp:codi_classificacio>
        <exp:titol_serie_documental>Expedients de personal</exp:titol_serie_documental>
        <exp:nivell_descripcio>Unitat documental composta</exp:nivell_descripcio>
        <exp:titol>Expedient de contractació per a la compra de quinze ordinadors</exp:titol>
        <exp:data_obertura>2008-05-17T09:30:47.0Z</exp:data_obertura>
        <exp:data_tancament>2008-05-17T09:30:47.0Z</exp:data_tancament>
        <exp:nom_productor>Agència Catalana de Certificació</exp:nom_productor>
        <exp:unitat_productora>Direcció General</exp:unitat_productora>
        <exp:descripcio>String</exp:descripcio>
        <exp:descriptors>String</exp:descriptors>
        <exp:documentacio_relacionada>String</exp:documentacio_relacionada>
        <exp:tipus_relacio>String</exp:tipus_relacio>
        <exp:classificacio_seguretat_acces>Accés públic</exp:classificacio_seguretat_acces>
        <exp:sensibilitat_dades_LOPD>Nivell baix</exp:sensibilitat_dades_LOPD>
      </exp:expedient>
    </METS:xmlData>
  </METS:mdWrap>
</METS:dmdSec>
```

```

<METS:dmdSec ID="DMD_DOC1">
  <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_document_exp"
MIMETYPE="text/xml" >
  <METS:xmlData>
    <doc:document xmlns:doc="http://schemas.user.iarxiu.hp.com/2.0/Voc_document_exp">
      <doc:codi_referencia>CAT/AACC/D1161_N-185_05/01</doc:codi_referencia>
      <doc:numero_document>01</doc:numero_document>
      <doc:titol>Resolució d'adjudicació del contracte de subministrament de 15 ordinadors</doc:titol>
      <doc:data_creacio>2008-05-17T09:30:47.0Z</doc:data_creacio>
      <doc:nivell_descripcio>Unitat documental simple</doc:nivell_descripcio>
      <doc:suport>Electrònic</doc:suport>
      <doc:nom_productor>Agència Catalana de Certificació</doc:nom_productor>
      <doc:unitat_productora>Direcció General</doc:unitat_productora>
      <doc:descripcio>Resolució d'adjudicació signada pel director general adjudicant el contracte de subministrament de 15
ordinadors a l'empresa XTS</doc:descripcio>
      <doc:descriptors>String</doc:descriptors>
      <doc:tipus_document>Acta</doc:tipus_document>
      <doc:classificacio_seguretat_acces>Accés públic</doc:classificacio_seguretat_acces>
      <doc:sensibilitat_dades_LOPD>Nivell baix</doc:sensibilitat_dades_LOPD>
      <doc:nivell_classificacio_evidencial>Sense evidència (identificació al·legada)</doc:nivell_classificacio_evidencial>
      <doc:document_essencial>Sí</doc:document_essencial>
    </doc:document>
  </METS:xmlData>
</METS:mdWrap>
</METS:dmdSec>

```

Es pot observar que hi ha tants grups de metadades administratives com vocabularis diferents apareixen al paquet. En particular, cadascun d'aquests grups conté:

- Identificador de vocabulari (atribut ID). Els identificadors vocabulari poden ser qualsevol, però han de ser únics a tot el METS.
- atribut MDTYPE amb valor "OTHER"
- atribut OTHERMDTYPE amb valor de l'identificador iArxiu del vocabulari. Els valors definitius dependran de la configuració definitiva de la plataforma iArxiu un cop estigui en producció.
- Contingut del vocabulari (format XML).

Aquests vocabularis es poden emplenar tenint en compte determinat esquema XSD, prèviament donat d'alta a la plataforma, o bé poden ser lliures. En aquest segon cas, els vocabularis no s'indexaran i no es podran fer cerques de paquets per aquestes metadades.

Habitualment els vocabularis seguiran un determinat esquema XSD prèviament donat d'alta. en aquest cas, la plataforma farà les validacions estructurals i de contingut corresponents. S'han de tenir en compte els següents punts:

- Per cadascun dels vocabularis s'ha d'informar l'espai de noms (*namespace*) juntament amb un prefix (per exemple *xmlns:exp*) de forma que coincideixi amb *targetNamespace* de l'esquema del vocabulari. Per altre banda, no és recomanable informar els atributs *xmlns:xsi* ni *xsi:schemaLocation*, doncs un cop ingressats, els vocabularis no tindran cap localització definida.
- Els elements als esquemes dels vocabularis estan agrupats amb *xs:sequence*, de forma que és important seguir l'ordre adequats dels elements
- Alguns elements són enumeracions, i s'han de respectar els continguts possibles.
- Si a la plantilla existeixen també limitacions sobre aquest vocabulari, també es tindran en compte.

Al punt 4 es detallen els esquemes de metadades que els clients d'iArxiu han d'emplenar en cas d'enviar expedients i documents seguint el model de vocabulari de metadades proposat per iArxiu.

3.5 Administrative Metadata Section

Finalment, les metadades administratives apareixen en aquesta secció. En aquesta secció només posarem metadades de signatura. En el nostre exemple, en tenim només una: AMD_SIG, que es correspon amb el vocabulari que descriu la signatura.

```
<METS:amdSec ID="AMD_SIG">
  <METS:techMD ID="AMD_SIG_1.0">
    <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_signatura" MIMETYPE="text/xml"
  >
    <METS:xmlData>
      <sig:signatura xmlns:sig="http://schemas.user.iarxiu.hp.com/2.0/Voc_signatura">
        <sig:identificador>0811233232234223</sig:identificador>
        <sig:identificador_document>BIN_1.0</sig:identificador_document>
        <sig:tipus_signatura>Independent (detached)</sig:tipus_signatura>
        <sig:format_signatura>CMS</sig:format_signatura>
      </sig:signatura>
    </METS:xmlData>
  </METS:mdWrap>
</METS:techMD>
</METS:amdSec>
```

Com es pot observar les metadades administratives funcionen de forma anàloga a les descriptives.

És important remarcar que la metadada *identificador_document* de les metadades de signatura conté "BIN_1.0", que és el document al qual pertany aquesta signatura.

Les metadades *tipus_signatura* i *format_signatura* no són obligatòries, però si l'usuari les coneix les pot informar.

A mode de resum, podem descriure un fitxer METS de la següent forma:

4. Vocabularis de metadades per iArxiu

Els vocabularis de metadades que els usuaris clients d'iArxiu han d'emplenar per ingressar un paquet d'informació de transferència (PIT) varien en funció de si es tracta d'un expedient o d'un document sol:

- Per a expedients (unitats documentals compostes), els vocabularis són els següents:
 - Vocabulari d'expedient (*Voc_expedient.xsd*).
 - Vocabulari de documents integrats en expedients (*Voc_document_exp.xsd*).
- Per a documents individuals (unitat de document simple), el vocabulari és el següent:
 - Vocabulari de documents que no formen part d'un expedient (*Voc_document.xsd*).

En funció de si s'envia un expedient o un document és obligatori informar al paquet, com a mínim, el vocabulari d'expedient o el de documentSol, respectivament. És necessari informar als paquets, com a mínim, un d'aquests vocabularis per tal de garantir la preservació i assegurar que es podran cercar els documents.

Adicionalment, si es volen arxivar les signatures dels documents, hi ha un vocabulari per descriure-les (*Voc_signatura.xsd*).

Pel que fa als binaris, aquests tenen com a possible vocabulari a emplenar el de PREMIS com a metadades administratives, però aquest vocabulari és opcional. Les úniques dades a emplenar són les descrites al punt 3.2: Identificador, MIMEType, resum de fitxer i algorisme de resum.

Als següents punts s'inclouen les representacions gràfiques dels esquemes mencionats anteriorment. Als fitxers XSD que s'adjunten amb aquest document es poden observar els tipus de dades dels diferents elements, i si són obligatoris o no.

4.1 Vocabulari d'expedient

URN de vocabulari: urn:iarxiu:2.0:vocabularies:catcert:Voc_expedient

Espai de noms: http://schemas.user.iarxiu.hp.com/2.0/Voc_expedient

L'esquema que descriu un expedient és el següent:

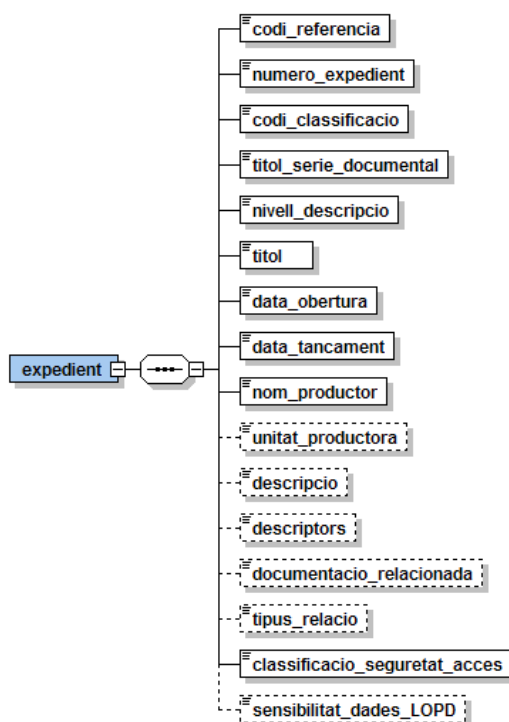


Figura 3: Vocabulari d'expedient

A continuació es mostra un exemple concret de metadades descriptives fent servir aquest vocabulari:

```
<METS:dmdSec ID="DMD_EXP">
  <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_expedient" MIMETYPE="text/xml">
    <METS:xmlData>
      <exp:expedient xmlns:exp="http://schemas.user.iarxiu.hp.com/2.0/Voc_expedient">
        <exp:codi_referencia>CAT/AACC/D1161_N-185_05</exp:codi_referencia>
        <exp:numero_expedient>185/08</exp:numero_expedient>
        <exp:codi_classificacio>G0200</exp:codi_classificacio>
        <exp:titol_serie_documental>Expedients de personal</exp:titol_serie_documental>
        <exp:nivell_descripcio>Unitat documental composta</exp:nivell_descripcio>
        <exp:titol>Expedient de contractació per a la compra de quinze ordinadors</exp:titol>
        <exp:data_obertura>2008-05-17T09:30:47.0Z</exp:data_obertura>
        <exp:data_tancament>2008-05-17T09:30:47.0Z</exp:data_tancament>
        <exp:nom_productor>Agència Catalana de Certificació</exp:nom_productor>
        <exp:unitat_productora>Direcció General</exp:unitat_productora>
        <exp:descripcio>String</exp:descripcio>
        <exp:descriptors>String</exp:descriptors>
        <exp:documentacio_relacionada>String</exp:documentacio_relacionada>
        <exp:tipus_relacio>String</exp:tipus_relacio>
        <exp:classificacio_seguretat_acces>Accés públic</exp:classificacio_seguretat_acces>
        <exp:sensibilitat_dades_LOPD>Nivell baix</exp:sensibilitat_dades_LOPD>
      </exp:expedient>
    </METS:xmlData>
  </METS:mdWrap>
</METS:dmdSec>
```

4.2 Vocabulari de document integrat en expedient

URN de vocabulari: urn:iarxiu:2.0:vocabularies:catcert:Voc_document_exp

Espai de noms: http://schemas.user.iarxiu.hp.com/2.0/Voc_document_exp

L'esquema que descriu un document que forma part d'un expedient és:

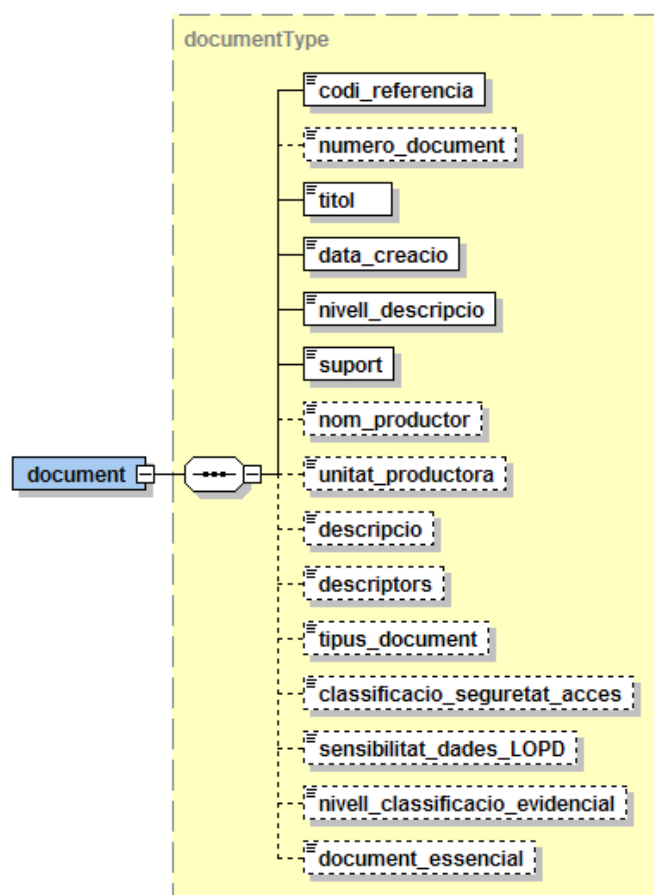


Figura 4: Vocabulari document integrat en expedient

A continuació es mostra un exemple concret de metadades descriptives fent servir aquest vocabulari:

```
<METS:dmdSec ID="DMD_DOC1">
  <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_document_exp"
MIMETYPE="text/xml" >
  <METS:xmlData>
    <doc:document xmlns:doc="http://schemas.user.iarxiu.hp.com/2.0/Voc_document_exp">
      <doc:codi_referencia>CAT/AACC/D1161_N-185_05/01</doc:codi_referencia>
      <doc:numero_document>01</doc:numero_document>
      <doc:titol>Resolució d'adjudicació del contracte de subministrament de 15 ordinadors</doc:titol>
      <doc:data_creacio>2008-05-17T09:30:47.0Z</doc:data_creacio>
      <doc:nivell_descripcio>Unitat documental simple</doc:nivell_descripcio>
      <doc:suport>Electrònic</doc:suport>
      <doc:nom_productor>Agència Catalana de Certificació</doc:nom_productor>
      <doc:unitat_productora>Direcció General</doc:unitat_productora>
      <doc:descripcio>Resolució d'adjudicació signada pel director general adjudicant el contracte de subministrament de 15
ordinadors a l'empresa XTS</doc:descripcio>
      <doc:descriptors>String</doc:descriptors>
```

```
<doc:tipus_document>Acta</doc:tipus_document>
<doc:classificacio_seguretat_acces>Accés públic</doc:classificacio_seguretat_acces>
<doc:sensibilitat_dades_LOPD>Nivell baix</doc:sensibilitat_dades_LOPD>
<doc:nivell_classificacio_evidencial>Sense evidència (identificació al•legada)</doc:nivell_classificacio_evidencial>
<doc:document_essencial>Sí</doc:document_essencial>
</doc:document>
</METS:xmlData>
</METS:mdWrap>
</METS:dmdSec>
```

4.3 Vocabulari de document independent

URN de vocabulari: urn:iarxiu:2.0:vocabularies:catcert:Voc_document

Espai de noms: http://schemas.user.iarxiu.hp.com/2.0/Voc_document

L'esquema que descriu un document independent (que no forma part d'un expedient) és el següent:

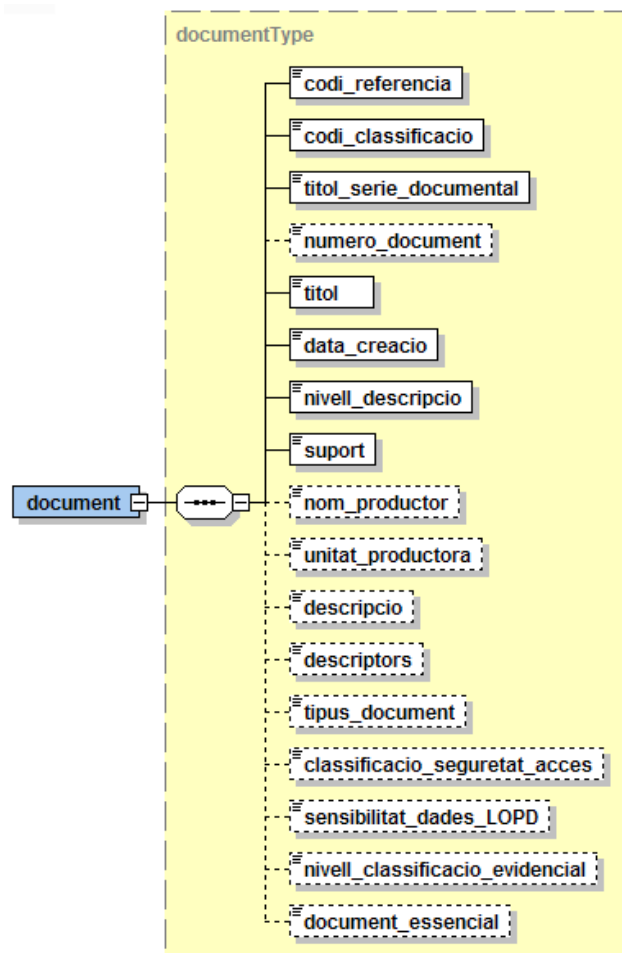


Figura 5: Vocabulari unitat documental simple

A continuació es mostra un exemple concret de metadades descriptives fent servir aquest vocabulari:

```
<METS:dmdSec ID="DMD_1">
<METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_document" MIMETYPE="text/xml">
  <METS:xmlData>
    <doc:document xmlns:doc="http://schemas.user.iarxiu.hp.com/2.0/Voc_document">
      <doc:codi_referencia>AACC= Arxiu de l'Agència Catalana de Certificació</doc:codi_referencia>
      <doc:codi_classificacio>G0200/U0000/E0000</doc:codi_classificacio>
      <doc:titol_serie_documental>Control de l'habilitació</doc:titol_serie_documental>
      <doc:numero_document>01</doc:numero_document>
      <doc:titol>Resolució d'adjudicació del contracte de subministrament de 15 ordinadors</doc:titol>
      <doc:data_creacio>2008-05-17T09:30:47.0Z</doc:data_creacio>
      <doc:nivell_descripcio>Unitat documental simple</doc:nivell_descripcio>
      <doc:suport>Electrònic</doc:suport>
      <doc:nom_productor>Joan Pou Mir</doc:nom_productor>
      <doc:unitat_productora>Àrea tècnica</doc:unitat_productora>
      <doc:descripcio>Resolució d'adjudicació signada pel director general adjudicant el contracte de subministrament de 15
ordinadors a l'empresa XTS</doc:descripcio>
      <doc:descriptors>String</doc:descriptors>
      <doc:tipus_document>Acta</doc:tipus_document>
      <doc:classificacio_seguretat_acces>Accés públic</doc:classificacio_seguretat_acces>
      <doc:sensibilitat_dades_LOPD>Nivell baix</doc:sensibilitat_dades_LOPD>
      <doc:nivell_classificacio_evidencial>Sense evidència (identificació al•legada)</doc:nivell_classificacio_evidencial>
      <doc:document_essencial>Sí</doc:document_essencial>
    </doc:document>
  </METS:xmlData>
</METS:mdWrap>
</METS:dmdSec>
```

4.4 Vocabulari de signatura

URN de vocabulari: urn:iarxiu:2.0:vocabularies:catcert:Voc_signatura

Espai de noms: http://schemas.user.iarxiu.hp.com/2.0/Voc_signatura

L'esquema que descriu una signatura és el següent:

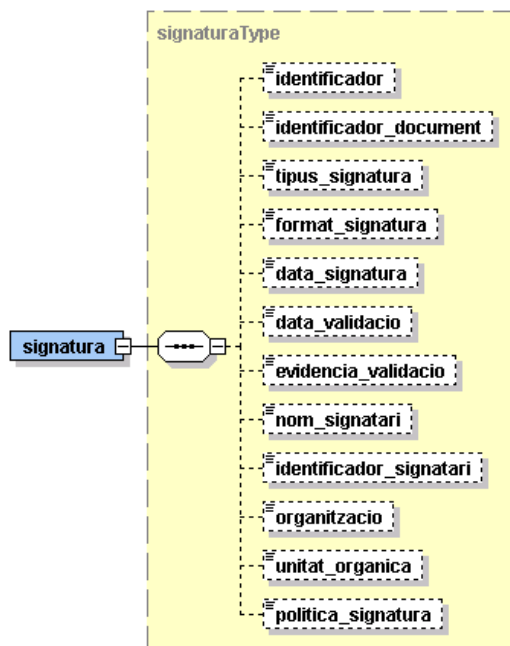


Figura 6: Vocabulari de signatura

A continuació es mostra un exemple concret de metadades administratives fent servir aquest vocabulari:

```
<METS:amdSec ID="AMD_SIG">
  <METS:techMD ID="AMD_SIG_1.0">
    <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_signatura" MIMETYPE="text/xml"
  >
    <METS:xmlData>
      <sig:signatura xmlns:sig="http://schemas.user.iarxiu.hp.com/2.0/Voc_signatura">
        <sig:identificador>08112332322234223</sig:identificador>
        <sig:identificador_document>BIN_1.0</sig:identificador_document>
        <sig:tipus_signatura>Independent (detached)</sig:tipus_signatura>
        <sig:format_signatura>CMS</sig:format_signatura>
      </sig:signatura>
    </METS:xmlData>
  </METS:mdWrap>
</METS:techMD>
</METS:amdSec>
```

Nota: La obligatorietat d'informar d'algunes metadades i dels valors d'algunes enumeracions als fitxers XSD es definirà de forma definitiva durant l'anàlisi tècnic d'iArxiu.

A continuació s'especifiquen alguns casos concrets de com informar aquest vocabulari en funció del tipus de document o signatura:

- **Documents PDF signats** (amb signatura dins del fitxer PDF)

En aquest cas no cal incloure el vocabulari de signatura, doncs la plataforma tracta de forma específica aquests documents i sap tractar la signatura si n'hi hagués.

- **Documents amb signatures independents** (signatures detached)

Per al cas de documents amb signatures independents (CMS, CAdES, XMLDSig o XAdES), s'ha d'informar el vocabulari de signatures al binari que representa la signatura. Respecte a les dades a informar, com a mínim s'ha d'emplenar el tag *identificador_document* amb al binari que està signant la signatura (veure exemple del punt 5). L'usuari pot informar més dades, però la plataforma les obtindrà automàticament si l'usuari no les informa.

- **Tiquets de PSIS**

Per als tiquets de PSIS apliquen els mateixos criteris que per a les signatures independents. Només s'ha de informar el vocabulari de signatura al binari amb el tiquet informant com a mínim el tag *identificador_document*.

- **Signatures incrustades XML** (enveloped)

Aquestes signatures són en format XML i es troben dins d'un document XML. En aquest cas no cal incloure el vocabulari de signatura, doncs la plataforma sap extraure la signatura de dins del document.

- **Signatures embolcallants XML** (enveloping)

Aquestes signatures són en format XML i contenen un document de qualsevol format. En aquest cas no cal incloure el vocabulari de signatura, doncs la plataforma sap detectar la signatura i extraure el document associat.

És important que a la secció de fitxers s'informi el Content Type del document a preservar, i no el de l'embolcall (que en aquest cas és sempre XML). A mode d'exemple, si el document signat és un document Word, s'ha d'informar *MIMETYPE='application/msword'* a l'element *file* que contingui la signatura embolcallant.

- Signatures CMS adjuntes (attached)

Aquestes signatures són en format binari i es poden tractar igual que les embolcallants XML, de forma que no cal incloure el vocabulari de signatura. La plataforma separarà la signatura del document per poder realitzar la seva preservació.

És important que a la secció de fitxers s'informi el Content Type del document a preservar, i no el del contenidor CMS (que en aquest cas és sempre *application/pkcs7-signature*). A mode d'exemple, si el document signat és un document Word, s'ha d'informar *MIMETYPE='application/msword'* a l'element *file* que contingui la signatura CMS attached.

4.5 Altres vocabularis

La plataforma iArxiu permet utilitzar o incloure altres vocabularis de metadades que poden ser específics per a algun tipus de documents o expedients com, per exemple, factures o expedients de subvencions. Aquests vocabularis han de seguir un esquema XSD senzill, tal i com s'ha descrit anteriorment a nivell d'expedient i de document. Ara bé, tenint en compte les següents restriccions:

- Tots els elements han d'estar declarats a l'element principal.
- No es poden definir objectes complexes que contenen altres elements o atributs.
- Només s'accepten elements de text, numèrics, dates (), dates amb hora (), booleans i enumeracions.
- No es permeten atributs als elements
- Els elements dels vocabularis estan agrupats amb *xs:sequence*

Per fer servir aquest tipus de vocabularis, l'administrador de la plataforma iArxiu els haurà d'haver donat d'alta prèviament.

La plataforma també permet fer servir vocabularis lliures que no segueixen cap esquema XSD concret, però aquests no seran indexats i, per tant, no es podran fer cerques per a aquests elements a l'hora de fer consultes.

5. Exemple PIT segons el model METS

A continuació es mostren fitxers METS emplenats amb totes les seccions i vocabularis descrits als punts anteriors.

5.1 Paquet amb estructura d'expedient

Aquest exemple es correspon amb un PIT que inclou un document i una signatura, a més d'incloure metadades a nivell d'expedient, de document i de signatura.

```
<?xml version="1.0" encoding="UTF-8"?>
<METS:mets xmlns:METS="http://www.loc.gov/METS/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd"
xmlns:xlink="http://www.w3.org/1999/xlink" TYPE="urn:iarxiu:2.0:templates:catcert:PL_expedient">

<!--
Metadades descriptives: vocabulari expedient
Aquest vocabulari és propi de dades d'un expedient -->
<METS:dmdSec ID="DMD_EXP">
  <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_expedient" MIMETYPE="text/xml">
    <METS:xmlData>
      <exp:expedient xmlns:exp="http://schemas.user.iarxiu.hp.com/2.0/Voc_expedient">
        <exp:codi_referencia>CAT/AACC/D1161_N-185_05</exp:codi_referencia>
        <exp:numero_expedient>185/08</exp:numero_expedient>
        <exp:codi_classificacio>G0200</exp:codi_classificacio>
        <exp:titol_serie_documental>Expedients de personal</exp:titol_serie_documental>
        <exp:nivell_descripcio>Unitat documental composta</exp:nivell_descripcio>
        <exp:titol>Expedient de contractació per a la compra de quinze ordinadors</exp:titol>
        <exp:data_obertura>2008-05-17T09:30:47.0Z</exp:data_obertura>
        <exp:data_tancament>2008-05-17T09:30:47.0Z</exp:data_tancament>
        <exp:nom_productor>Agència Catalana de Certificació</exp:nom_productor>
        <exp:unitat_productora>Direcció General</exp:unitat_productora>
        <exp:descripcio>String</exp:descripcio>
        <exp:descriptors>String</exp:descriptors>
        <exp:documentacio_relacionada>String</exp:documentacio_relacionada>
        <exp:tipus_relacio>String</exp:tipus_relacio>
        <exp:classificacio_seguretat_acces>Accés públic</exp:classificacio_seguretat_acces>
        <exp:sensibilitat_dades_LOPD>Nivell baix</exp:sensibilitat_dades_LOPD>
      </exp:expedient>
    </METS:xmlData>
  </METS:mdWrap>
</METS:dmdSec>

<!--
Metadades descriptives: vocabulari doc1
Aquestes metadades descriuen el document que hi ha al expedient.
Amb més documents, n'hi haurà més blocs com aquest
-->
<METS:dmdSec ID="DMD_DOC1">
  <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_document_exp"
MIMETYPE="text/xml">
    <METS:xmlData>
      <doc:document xmlns:doc="http://schemas.user.iarxiu.hp.com/2.0/Voc_document_exp">
        <doc:codi_referencia>CAT/AACC/D1161_N-185_05/01</doc:codi_referencia>
        <doc:numero_document>01</doc:numero_document>
        <doc:titol>Resolució d'adjudicació del contracte de subministrament de 15 ordinadors</doc:titol>
        <doc:data_creacio>2008-05-17T09:30:47.0Z</doc:data_creacio>
        <doc:nivell_descripcio>Unitat documental simple</doc:nivell_descripcio>
        <doc:suport>Electrònic</doc:suport>
        <doc:nom_productor>Agència Catalana de Certificació</doc:nom_productor>
        <doc:unitat_productora>Direcció General</doc:unitat_productora>
        <doc:descripcio>Resolució d'adjudicació signada pel director general adjudicant el contracte de subministrament de 15
ordinadors a l'empresa XTS</doc:descripcio>
        <doc:descriptors>String</doc:descriptors>
        <doc:tipus_document>Acta</doc:tipus_document>
        <doc:classificacio_seguretat_acces>Accés públic</doc:classificacio_seguretat_acces>
        <doc:sensibilitat_dades_LOPD>Nivell baix</doc:sensibilitat_dades_LOPD>
      </doc:document>
    </METS:xmlData>
  </METS:mdWrap>
</METS:dmdSec>
```



```

        <doc:nivell_classificacio_evidencial>Sense evidència (identificació al·legada)</doc:nivell_classificacio_evidencial>
        <doc:document_essencial>Sí</doc:document_essencial>
    </doc:document>
</METS:xmlData>
</METS:mdWrap>
</METS:dmdSec>

<!--
Metadades administratives: signatura
-->
<METS:amdSec ID="AMD_SIG">
  <METS:techMD ID="AMD_SIG_1.0">
    <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_signatura" MIMETYPE="text/xml"
  >
    <METS:xmlData>
      <sig:signatura xmlns:sig="http://schemas.user.iarxiu.hp.com/2.0/Voc_signatura">
        <sig:identificador>0811233232234223</sig:identificador>
        <sig:identificador_document>BIN_1.0</sig:identificador_document>
        <sig:tipus_signatura>Independent (detached)</sig:tipus_signatura>
        <sig:format_signatura>CMS</sig:format_signatura>
      </sig:signatura>
    </METS:xmlData>
  </METS:mdWrap>
</METS:techMD>
</METS:amdSec>

<!--
Secció de fitxers: En aquesta secció s'enumeren els diferents fitxers binaris que intervenen al paquet
No han de seguir cap estructura, doncs l'estructura es dona a la secció structMap, que apuntarà
a cadascun d'aquests fitxers. Només és una llista d'identificadors i ubicacions.
-->
<METS:fileSec>
  <METS:fileGrp ID="BIN_1" >
    <METS:file ID="BIN_1.0" MIMETYPE="application/msword" CHECKSUM="99800b85d3383e3a2fb45eb7d0066a4879a9dad0"
CHECKSUMTYPE="SHA-1" CREATED="2008-05-17T09:30:47.0Z">
      <METS:FContent>
        <METS:binData>aG9sYQ==</METS:binData>
      </METS:FContent>
    </METS:file>
  </METS:fileGrp>
  <METS:fileGrp ID="BIN_2" >
    <METS:file ID="BIN_2.0" MIMETYPE="application/pkcs7-signature" CHECKSUM="9788b4dffb7f0d350edabbf10d325b47e350def"
CHECKSUMTYPE="SHA-1" CREATED="2008-05-17T09:30:47.0Z">
      <METS:FContent>
        <METS:binData>aG9sYSBvdHJhIHZleg==</METS:binData>
      </METS:FContent>
    </METS:file>
  </METS:fileGrp>

  <!-- Fitxer referenciat -->
  <!-- <METS:fileGrp ID="BIN_3" >
    <METS:file ID="BIN_3.0" MIMETYPE="application/pkcs7-signature" CHECKSUM="9788b4dffb7f0d350edabbf10d325b47e350def"
CHECKSUMTYPE="SHA-1" CREATED="2008-05-17T09:30:47.0Z">
    <METS:FContent>
      <METS:binData>aG9sYSBvdHJhIHZleg==</METS:binData>
    </METS:FContent>
  </METS:file>
  </METS:fileGrp> -->

  </METS:fileSec>

<!--
Secció de estructura de l'ODP. En aquesta secció indiquen la estructura jeràrquica del paquet.
En aquest cas és senzilla, doncs només reflexa un expedient amb un document, que conté un binari.
En aquesta secció també es relacionen els elements (exp, doc, binari) amb les seves metadades
(enumerades a dalt).
-->
<METS:structMap>
  <METS:div LABEL="expedient" DMDID="DMD_EXP">
    <METS:div LABEL="document_1" DMDID="DMD_DOC1">
      <METS:div LABEL="document.doc">
        <METS:fptr FILEID="BIN_1"/>
      </METS:div>
    </METS:div>
  </METS:div>
</METS:structMap>

```



```
</METS:div>
<METS:div LABEL="firma.p7b" ADMID="AMD_SIG">
  <METS:fptr FILEID="BIN_2"/>
</METS:div>
</METS:div>
</METS:div>
</METS:structMap>

</METS:mets>
```

Per veure el fitxer METS consulteu el document adjunt: *Exemple_generic_PIT_METS.XML*.

5.2 Paquet amb estructura de document independent

Aquest exemple es correspon amb un PIT que inclou un document i una signatura que no formen part de cap expedient:

```
<?xml version="1.0" encoding="UTF-8"?>
<METS:mets xmlns:METS="http://www.loc.gov/METS/" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="http://www.loc.gov/METS/ http://www.loc.gov/standards/mets/mets.xsd"
xmlns:xlink="http://www.w3.org/1999/xlink" TYPE="urn:iarxiu:2.0:templates:catcert:PL_document">

  <METS:dmdSec ID="DMD_1">
    <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_document" MIMETYPE="text/xml">
      <METS:xmlData>
        <doc:document xmlns:doc="http://schemas.user.iarxiu.hp.com/2.0/Voc_document">
          <doc:codi_referencia>AACC= Arxiu de l'Agència Catalana de Certificació</doc:codi_referencia>
          <doc:codi_classificacio>G0200/U0000/E0000</doc:codi_classificacio>
          <doc:titol_serie_documental>Control de l'habilitació</doc:titol_serie_documental>
          <doc:numero_document>01</doc:numero_document>
          <doc:titol>Resolució d'adjudicació del contracte de subministrament de 15 ordinadors</doc:titol>
          <doc:data_creacio>2008-05-17T09:30:47.0Z</doc:data_creacio>
          <doc:nivell_descripcio>Unitat documental simple</doc:nivell_descripcio>
          <doc:suport>suport</doc:suport>
          <doc:nom_productor>Joan Pou Mir</doc:nom_productor>
          <doc:unitat_productora>Àrea tècnica</doc:unitat_productora>
          <doc:descripcio>Resolució d'adjudicació signada pel director general adjudicant el contracte de subministrament de 15
ordinadors a l'empresa XTS</doc:descripcio>
          <doc:descriptors>String</doc:descriptors>
          <doc:tipus_document>Acta</doc:tipus_document>
          <doc:classificacio_seguretat_acces>Accés públic</doc:classificacio_seguretat_acces>
          <doc:sensibilitat_dades_LOPD>Nivell baix</doc:sensibilitat_dades_LOPD>
          <doc:nivell_classificacio_evidencial>Sense evidència (identificació al·legada)</doc:nivell_classificacio_evidencial>
          <doc:document_essencial>Sí</doc:document_essencial>
        </doc:document>
      </METS:xmlData>
    </METS:mdWrap>
  </METS:dmdSec>

  <METS:amdSec ID="AMD_SIG">
    <METS:techMD ID="AMD_SIG_1.0">
      <METS:mdWrap MDTYPE="OTHER" OTHERMDTYPE="urn:iarxiu:2.0:vocabularies:catcert:Voc_signatura" MIMETYPE="text/xml">
        <METS:xmlData>
          <sig:signatura xmlns:sig="http://schemas.user.iarxiu.hp.com/2.0/Voc_signatura">
            <sig:identificador_document>BIN_1.0</sig:identificador_document>
          </sig:signatura>
        </METS:xmlData>
      </METS:mdWrap>
    </METS:techMD>
  </METS:amdSec>

  <METS:fileSec>
    <METS:fileGrp ID="BIN_1">
      <METS:file ID="BIN_1.0" MIMETYPE="text/html" CHECKSUM="ea631551f5569f612dd702b900c596c2a99c0dfd"
CHECKSUMTYPE="SHA-1" CREATED="2008-05-17T09:30:47.0Z">
        <METS:FContent>
          <METS:binData>dGV4dG8=</METS:binData>
        </METS:FContent>
      </METS:file>
    </METS:fileGrp>
  </METS:fileSec>
```

```

</METS:file>
</METS:fileGrp>
<METS:fileGrp ID="BIN_2" >
  <METS:file ID="BIN_2.0" MIMETYPE="application/pkcs7-signature" CHECKSUM="9788b4dffb7f0d350edabbf10d325b47e350def"
CHECKSUMTYPE="SHA-1" CREATED="2008-05-17T09:30:47.0Z">
    <METS:FContent>
      <METS:binData>aG9sYSBvdHJhIHZleg==</METS:binData>
    </METS:FContent>
  </METS:file>
</METS:fileGrp>
</METS:fileSec>

<METS:structMap>
  <METS:div LABEL="document_1" DMDID="DMD_1">
    <METS:div LABEL="document.txt">
      <METS:fptr FILEID="BIN_1"/>
    </METS:div>
    <METS:div LABEL="signatura.p7s" ADMID="AMD_SIG">
      <METS:fptr FILEID="BIN_2"/>
    </METS:div>
  </METS:div>
</METS:structMap>

</METS:mets>

```

6. Implementació tècnica

Per a la creació d'un fitxer METS des de Java plantejem dues opcions:

- Mitjançant tecnologia XMLBeans: XMLBeans és una llibreria genèrica i lliure que ofereix la funcionalitat de generació d'XMLs a partir d'esquemes XSD. Per a més informació sobre el projecte consulteu a: <http://xmlbeans.apache.org/>
- Mitjançant la llibreria específica de generació de fitxers METS: creada per la Universitat de Harvard. La descàrrega es pot fer des de la següent pàgina web: <http://hul.harvard.edu/mets/>

7. Annexes

7.1 Taula d'il·lustracions

Figura 1: Estructura METS en blocs	7
Figura 2: Associacions de metadades i estructura	15
Figura 3: Vocabulari d'expedient	17
Figura 4: Vocabulari document integrat en expedient	18
Figura 5: Vocabulari unitat documental simple	19
Figura 6: Vocabulari de signatura	20