



Diversitat de **BiDes**

# 15es JORNADES CATALANES D'INFORMACIÓ I DOCUMENTACIÓ

Barcelona, 10 i 11 de maig de 2018

## **El nou escenari del dipòsit legal en línia més enllà de l'arxiu web.**

**Eugènia Serra,**  
**Directora de la Biblioteca de Catalunya**  
**Hospital, 56, 08001 Barcelona**  
**93 2702300 Ext. 3122**  
[www.bnc.cat](http://www.bnc.cat)  
[eserra@bnc.ca](mailto:eserra@bnc.ca)

**Karibel Pérez**  
**Coordinadora de l'Àrea de Tecnologia de la Informació de la**  
**Biblioteca de Catalunya**  
**Hospital, 56, 08001 Barcelona**  
**93 2702300 Ext. 3130**  
[www.bnc.cat](http://www.bnc.cat)  
[kperez@bnc.cat](mailto:kperez@bnc.cat)

Resum:

A partir del nou marc legal del 2015 per al dipòsit legal dels documents nascuts digitals en línia, s'ha creat un nou escenari en la gestió del dipòsit legal que implica a les biblioteques que actuen com a centres de conservació del dipòsit legal, als editors i també a la forma d'accés per als usuaris. Més enllà de la captura de les pàgines web que la Biblioteca de Catalunya du a terme des del 2005 amb PADICAT, ara l'àmbit d'actuació legal afecta a una producció de llibres, revistes i altres tipus de materials digitals intangibles que necessàriament requereixen d'un model de gestió diferent a l'utilitzat fins fa poc temps. La comunicació descriu aquest nou escenari i les gestions en curs que realitza la Biblioteca de Catalunya per donar-hi resposta.

Paraules clau: Biblioteca de Catalunya, Dipòsit legal en línia, Preservació digital, Sistemes d'informació per a la gestió, Recursos electrònics

Resumen:

A partir del nuevo marco legal de 2015 para el depósito legal de los documentos nacidos digitales en línea, se ha creado un nuevo escenario en la gestión del depósito legal que implica a las bibliotecas que actúan como centros de conservación del depósito legal, a los editores y también a la forma de dar acceso a los usuarios. Más allá de la captura de las



pàgines web que la Biblioteca de Catalunya lleva a cabo desde el 2005 con PADICAT, ahora, el ámbito de actuación legal afecta a una producción de libros, revistas y otros tipos de materiales digitales intangibles que necesariamente requieren de un modelo de gestión diferente al utilizado hasta hace poco tiempo. La comunicación describe este nuevo escenario y las gestiones en curso que lleva a cabo la Biblioteca de Catalunya para darle respuesta.

Palabras clave: Biblioteca de Catalunya, Depósito legal en línea, Preservación digital, Sistemas de información para la gestión, Recursos electrónicos

Abstract:

Since the new legal framework passed in 2015 for the legal deposit of online born-digital documents, a new scenario has been created in the management of the legal deposit that involves the libraries which act as preservation centers of the legal deposit, the editors and also the way of giving access to users. Beyond the harvesting of web pages that the Biblioteca de Catalunya carries out since 2005 through PADICAT, nowadays, the scope of legal action affects to an output of books, magazines and other types of intangible digital materials that necessarily require a model of management different from the one used until recently. This communication describes this new scenario and the ongoing efforts of the Biblioteca de Catalunya to satisfy it.

Key words: Biblioteca de Catalunya, Online legal deposit, Digital preservation, Management information systems, Electronic resources

## 1. ANTECEDENTS

El dipòsit legal és el mecanisme que tenen les biblioteques nacionals de bona part del món per recollir la producció bibliogràfica del seu país. Fins a finals del segle passat, disposaven d'una legislació més o menys antiga i més o menys obsoleta dissenyada per a un món analògic, en front d'un escenari que ja es basava en sistemes de producció digitals.

Des de llavors i especialment, en l'actual mil·lenni, el creixement de la producció digital nascuda a Internet i, per tant sense suport tangible, ha estat continu. Ha estat doncs indispensable un canvi de normativa, que en el cas espanyol havia d'afectar no només

al fet de contemplar el nou model de documents nascuts digitals no tangibles, sinó que també al subjecte responsable de dipositar que, fins a aquell moment -com a herència d'una tradició d'impressors/editors exercint un mateix rol- era encara principalment, l'impressor, una figura que avui en dia no té cap responsabilitat intel·lectual en la producció bibliogràfica.

Així doncs, progressivament, els països europeus van aprovar noves legislacions que responien al nou escenari digital.

Paral·lelament, un grup creixent de països iniciava l'arxiu web, la web contenia informació extremadament útil, però amb una mitjana de pervivència molt baixa (de promig, un mes) i per tant, amb un risc de desaparició molt elevat.

Suècia<sup>1</sup> i Dinamarca<sup>2</sup> a Europa, o Austràlia<sup>3</sup>, van ser dels primers països que van posar en marxa sistemes d'informació que recollien i conservaven versions de les webs produïdes als seus territoris, en uns casos de forma exhaustiva i en altres, de forma selectiva. Cas a banda, però també obligatori de citar és l'Internet Archive<sup>4</sup>, que comença al mateix temps que Suècia i Austràlia, però amb una voluntat focalitzada més en l'accés que no pas en la funció de preservació característica de les biblioteques nacionals.

La Biblioteca de Catalunya va ser la primera institució a Espanya a iniciar aquesta recollida de l'arxiu web mitjançant el sistema PADICAT (Patrimoni Digital de Catalunya), amb el suport com a soci tecnològic del CESCO (Centre de Supercomputació de Catalunya), actualment transformat en CSUC (Consorci de Serveis Universitaris). PADICAT, ha presentat un tret diferencial respecte a altres projectes d'arxiu web, que és el fet de ser un portal obert i accessible, des dels seus inicis, davant la situació d'alegalitat existent.

## 2. LA LEGISLACIÓ

<sup>1</sup> Kulturarw3. <<http://www.kb.se/english/find/internet/websites/>> iniciat el 1996.

<sup>2</sup> Netarkivet.dk. <<http://netarkivet.dk/in-english/>> iniciat el 1998.

<sup>3</sup> Pandora.< <http://pandora.nla.gov.au/>> iniciat el 1996.

<sup>4</sup> Internet Archive. <<https://archive.org/>>

A l'estat espanyol, la llei de dipòsit legal vigent és la llei 23/2011, de 29 de juliol, en la qual, a més del canvi de responsabilitat del dipòsit de l'impressor a l'editor, apareixia per primera vegada l'obligatorietat de dipositar publicacions nascudes digitals sense suport tangible (és a dir, a la xarxa), tot i que no especificava encara el detall.

Van passar alguns anys i molts debats fins a aprovar el Reial decret 635/2015, de 10 de juliol, que regula el dipòsit legal de les publicacions en línia. Aquest decret dona cobertura en bona mesura al que ja feia la Biblioteca de Catalunya amb PADICAT<sup>5</sup> des del 2005, i als projectes posteriors iniciats a l'estat de característiques similars, Ondarenet<sup>6</sup> al País Basc, liderat pel Departament de Cultura del Govern Basco el 2007, amb la col·laboració de la EJIE (Societat Informàtica del Govern Basco); i la recollida a l'Estat espanyol, feta per la Biblioteca Nacional de Espanya<sup>7</sup> des del 2009.

Els trets característics de la nova legislació de dipòsit legal de les publicacions en línia són:

- Continua l'obligació de dipositar si bé no s'assigna número de DL a les publicacions, donada la dificultat de delimitar correctament les unitats documentals a les webs, bases de dades, repositoris, etc.
- Els centres de preservació de DL, com la Biblioteca de Catalunya, tenen la possibilitat de capturar els recursos en línia, tant els que són d'accés obert, com restringit.
- Existeix l'obligació de dipositar per part del responsable de la publicació en línia, si el centre de conservació així ho demana.
- S'obre la possibilitat a realitzar una preservació distribuïda mitjançant convenis de col·laboració amb entitats posseïdores de repositoris que compleixin una sèrie de requisits de reservació fixats per la Biblioteca de Catalunya.

La nova regulació ha portat a la Biblioteca de Catalunya a dues accions principals:

- Dissenyar un sistema de dipòsit de publicacions nascudes digitals en línia, via web, amb destinació al seu repositori de preservació COFRE (COntservem per al Futur Recursos Electrònics), que inclou la gestió des de la recepció fins a la catalogació i consulta; sistema que exposarem més endavant.
- Definir un protocol de requeriments mínims per a subscriure convenis de col·laboració amb repositoris d'entitats que vulguin contribuir a la preservació digital; del qual en farem un resum en un altre apartat.

<sup>5</sup> PADICAT (Patrimoni Digital de Catalunya). <<http://www.padicat.cat>>

<sup>6</sup> OndareNet. <<http://www.ondarenet.kultura.ejgv.euskadi.eus:8085/ondarenet/>>. El País Basc no disposa de biblioteca nacional, de manera que el mateix govern va impulsar l'arxiu web.

<sup>7</sup> Archivo de la Web Española. <<http://www.bne.es/es/Colecciones/ArchivoWeb/>>

### 3. PADICAT (Patrimoni Digital de Catalunya)

No ens estendrem a explicar el sistema PADICAT donat que existeix força bibliografia en obert<sup>8</sup>, el mateix sistema es troba accessible en línia a <http://www.padicat.cat>, i no és l'objecte directe de la comunicació, però si farem un breu resum.

Iniciat el 2005, es va dissenyar com un sistema híbrid de recol·lecció:

- Captures automàtiques periòdiques del .cat, gràcies a un conveni amb la Fundació PuntCat pel qual s'obtenen periòdicament les url del domini que són actives.
- Captures periòdiques de webs recomanades per usuaris
- Captures periòdiques d'institucions, administracions i entitats públiques i privades amb qui s'ha signat una acord de col·laboració.
- Captures selectives monogràfiques, per exemple, de les eleccions municipals, al Parlament i al Congrés.

És d'accés obert, donat que tot el que es captura es pot consultar. Conté 338.580 captures de 86.241 webs diferents, amb un total 493 milions de fitxers, i un espai d'ocupació de 19,9 TB.

La crisi econòmica dels darrers anys ha afectat significativament PADICAT, que només s'ha pogut mantenir a mínims; el repte de futur immediat és incrementar-ne el pressupost i recuperar un equip bibliotecari permanent que s'hi dediqui, especialment pel que fa a validacions de continguts i disseny de monogràfics.

### 4. DIPÒSIT DE LES PUBLICACIONS EN LÍNIA

Els primers anys de la publicació electrònica (llibre electrònic però també publicacions en altres formats textuais, imatges o objectes complexos com CD's interactius) i en bona part dels casos com a mètode alternatiu i convivint amb la versió en paper, la BC rebia per dipòsit legal aquestes publicacions en suports com CD's, DVD's i inclús alguna memòria USB.

En la producció sonora i audiovisual que arribava també per dipòsit legal ja era habitual gestionar una gran diversitat de contenidors físics per albergar continguts tant analògics com digitals. Gràcies a aquest fet, els procediments interns per a gestionar i conservar aquests tipus de continguts estaven i continuen fortament imbricats amb

---

<sup>8</sup> A l'apartat Àrea Professional de la web de PADICAT s'hi troben referències a una selecció dels projectes més significatius d'arxiu web, així com un recull d'articles i ponències sobre el mateix PADICAT.  
<<https://www.padicat.cat/ca/coneix-nos/area-professional/ponencies-i-publicacions>>.

l'emmagatzematge i conservació dels seus contenidors. Estem parlant de, per exemple, cintes magnètiques, discs òptics, DVDs, etc.

En canvi en l'entorn de la producció en paper convertit a versió electrònica aquests suports arribaven a uns circuits "clàssics" de materials en paper no preparats per a gestionar, emmagatzemar i conservar aquests suports.

Diversos estudis<sup>9</sup> adverteixen de la finitud i degradació d'aquests contenidors (sobretot en condicions ambientals adverses) i per tant és ben conegut el fet que per assegurar la pervivència del seu contingut, s'haurà –entre altres tècniques- de transferir a un entorn de preservació segur, o almenys "refrescar el suport" com es denomina en preservació digital a la tècnica de transferir les dades d'un suport a un altre.

Ja en els darrers anys, aquests suports físics van ser substituïts per mètodes d'enviament purament electrònic com annexes de correu o transferències de fitxers via FTP o emmagatzematge al núvol com Dropbox, Drive, etc. Aquests enviaments suposaven per a la biblioteca fer front a l'emmagatzematge i conservació de tot un seguit de publicacions digitals de dipòsit legal que mentre no existien uns procediments establerts i un repositori de preservació a la institució, es desaven en volums de dades de la xarxa interna de l'organització i per tant no eren reconduïts a cap procediment de validació, d'emmagatzematge apropiat ni tampoc eren catalogats o posats a l'abast del públic (respectant els drets vigents).

Aquests fets van fer palesa la necessitat d'unificar i normalitzar el mètode de transferència de la producció digital des dels editors. A partir del moment que el repositori de preservació digital de la Biblioteca de Catalunya, COFRE (COservem per al Futur Recursos Electrònics)<sup>10</sup> ja era un sistema en funcionament, la BC va decidir implementar un mòdul que facilités als editors el dipòsit en línia dels documents

---

<sup>9</sup> Consultar com a exemple: Brown, Adrian. *Digital Preservation Guidance Note 2: Selecting storage media for long-term preservation*. The National Archives, 2008

<<https://www.nationalarchives.gov.uk/documents/selecting-storage-media.pdf>>

<sup>10</sup> La BC va posar en marxa el repositori COFRE (COservem per al Futur Recursos Electrònics) el 2011 amb diversos objectius: conservar els màsters resultat del seu programa de digitalització, guardar una còpia periòdica de PADICAT, i preservar el DL de publicacions nascudes digitals. A l'informe *Repositori de preservació digital de la Biblioteca de Catalunya*. <<http://www.recercat.cat/handle/2072/97251>> és descriu el plantejament i disseny inicial, fins i tot quan encara no era anomenat encara COFRE,

i proporcionés a més una metodologia per a la seva validació pel personal de la biblioteca: el mòdul DL-UP.

Aquest desenvolupament està format per un frontal web que guia als editors en el procés de càrrega dels documents digitals i els permet visualitzar l'estat de les seves càrregues, i per una base de dades on es desen les metadades relatives a les càrregues (data, estat, dades bibliogràfiques, enllaç amb COFRE). El frontal interactua per una banda amb el programari de gestió electrònica del Dipòsit Legal<sup>11</sup> per a validar les credencials de l'editor i per una altra amb el COFRE per a crear el document i transferir els fitxers que el componen. La figura 1 mostra el diagrama de flux del procediment des de la petició de càrrega en el mòdul DL-UP per part de l'editor, fins que els documents són validats a COFRE i catalogats al catàleg de la BC. O pel contrari, rebutjats per errors en els fitxers carregats. A continuació veurem el procés de càrrega en detall.

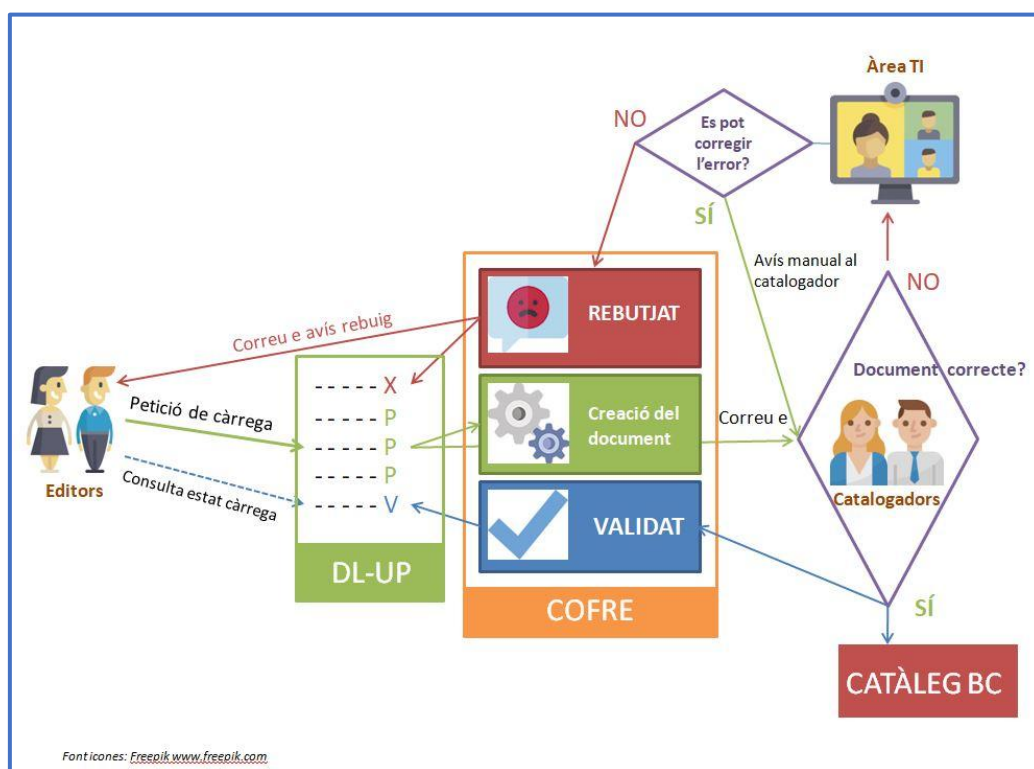


FIG. 1 CÀRREGA DE DOCUMENTS DIGITALS DE DIPÒSIT LEGAL AL REPOSITORI COFRE

<sup>11</sup> Més informació a <http://www.bnc.cat/Professionals/Diposit-legal>

#### 4.1. Creació d'una petició de càrrega

La petició de càrrega d'un document digital al mòdul DL-UP és un procés de 4 passos que permet a l'editor incorporar la informació requerida i carregar els fitxers que componen el document (figura 2).

Primer pas: Demana a l'editor si la publicació és una monografia o una publicació periòdica.

Segon pas: Aquest pas varia depenent de la resposta al pas1.

En el cas de publicacions periòdiques el mòdul demana a l'editor indicar si és la primera vegada que carrega un número (sigui el primer o no) de la publicació en el repositori. Si no ho és, li permet seleccionar un títol de la llista de totes les publicacions carregades pel mateix editor amb anterioritat. Aquesta informació és necessària per a localitzar lògicament el document al repositori, ja que tots els exemplars d'una mateixa publicació en sèrie es localitzen sota un mateix contenidor lògic.

En el cas de monografies demana a l'editor indicar el tipus de material: llibre, audiovisual, sonor, gràfic, cartogràfic o partitura. Aquesta informació permet per una banda identificar el grup de catalogadors, experts en el tipus de material designat, als quals se'ls enviarà una notificació de càrrega i per una altra banda permet al repositori COFRE dipositar el document sota un contenidor lògic específic.

Tercer pas: Permet a l'editor donar informació relativa al document. Només un conjunt mínim de metadades són obligatòries: títol, any de publicació, domini públic o accés restringit, any o número de l'exemplar per a la publicació en sèrie. La resta d'informació bibliogràfica del document serà proporcionada posteriorment pels catalogadors de la biblioteca, com ja es fa per a les publicacions en suport tangible.

Quart pas: L'últim pas permet a l'editor carregar els fitxers que componen els documents. Actualment no existeixen limitacions quant al format del fitxer, però es requereix als editors que dipositin la versió final del document en la major qualitat possible.



FIG. 1 PROCÉS DE PETICIÓ DE CÀRREGA AL MÒDUL DL-UP

## 4.2. Procés de càrrega a COFRE i validació

El mòdul DL-UP empaqueta la informació i els fitxers, i ho transfereix a una localització temporal, mentre realitza un seguit de crides seqüencials al repositori COFRE amb l'objectiu de cercar o crear el contenidor lògic, crear l'estructura del nou objecte (document, versió, esdeveniment de creació) i transferir els fitxers a la seva localització física definitiva.

Una vegada la transferència s'ha realitzat correctament, s'envia un missatge al grup de catalogadors expert en el tipus de material especificat, els quals són responsables de comprovar i validar els fitxers carregats i la informació proporcionada per l'editor, catalogar el document al catàleg de la BC i registrar el número de registre bibliogràfic al document a COFRE.

COFRE mentrestant realitzarà els processos de comprovació habituals a qualsevol càrrega en el repositori: detecció de virus, càlcul de la signatura digital dels fitxers carregats, identificació de format. Una vegada la validació manual es completa, l'objecte al repositori serà tractat seguint els mateixos procediments que la resta d'objectes al repositori de preservació: comprovació cíclica de virus i de signatura digital, còpia remota, creació de fitxer sidecar de metadades, logs d'accés i modificació.

Si la validació presenta problemes: impossibilitat d'obrir els fitxers, qualitat insuficient, falta de metadades per a identificar els fitxers, etc., el catalogador rebutja el document. Immediatament es genera un avís que arriba a l'Àrea de Tecnologia de la Informació de la BC, des d'on es realitza una segona comprovació amb l'objectiu de solucionar el problema: instal·lant visors apropiats, demanant contrasenyes, etc. Si el problema es resol, s'avisava manualment al catalogador, el qual pot continuar amb el procés normal de validació.

Si el problema no es pot resoldre, es rebutja la càrrega definitivament. Aquest pas implica l'eliminació de l'objecte a COFRE i la generació d'un avís per correu-e a l'editor perquè procedeixi a resoldre l'error i faci de nou la càrrega.

### 4.3. Llistat de càrregues

La segona opció del mòdul DL-UP mostra el llistat de les peticions de càrrega realitzades per l'editor (figura 3) i li permet per tant, gestionar les seves peticions de càrrega, realitzar-ne de noves a partir de les dades d'una càrrega anterior o resoldre els errors de càrregues rebutjades.

| Data                | Títol              | Típus material       | Any/Número | Estat de la càrrega                       |                              |
|---------------------|--------------------|----------------------|------------|---|------------------------------|
| 2018-03-13 13:40:10 | Time Out Barcelona | Publicació periòdica | 2018/496   | En revisió <a href="#">+Afegir número</a> | <a href="#">Canvia títol</a> |
| 2018-03-06 12:49:24 | Time Out Barcelona | Publicació periòdica | 2018/493   | Validada                                  | <a href="#">Canvia títol</a> |
| 2018-03-06 12:48:10 | Time Out Barcelona | Publicació periòdica | 2018/495   | Validada                                  | <a href="#">Canvia títol</a> |
| 2018-03-06 12:37:02 | Time Out Barcelona | Publicació periòdica | 2018/496   | Validada                                  | <a href="#">Canvia títol</a> |

FIG. 2 LLISTAT DE PETICIONS DE CÀRREGA AL MÒDUL DL-UP

El llistat mostra la data i hora de càrrega, el títol del document, el tipus de material, any/número de la publicació en sèrie i l'estat de la càrrega. Les peticions de càrrega poden ser-hi en un dels següents estats:

- En procés: la càrrega s'ha realitzat al repositori i es troba pendent de la comprovació manual i aprovació per part dels catalogadors.
- Validada: la comprovació manual i la conseqüent validació s'han realitzat.
- En revisió: La revisió manual s'ha realitzat però han sorgit problemes que requereixen comprovacions addicionals.
- Rebutjada: la càrrega ha estat rebutjada per la biblioteca. S'ha enviat un missatge a l'editor. Un enllaç a la dreta de l'estat permet realitzar de nou la càrrega fent servir la informació proporcionada en la càrrega inicial.

#### **4.4. Enllaç amb COFRE i amb el catàleg**

Com ja hem vist en el procés de càrrega el mateix mòdul DL-UP fa una crida al repositori de preservació COFRE per a poder crear l'entorn per a la càrrega dels fitxers i les metadades. I una vegada carregats, aquest passen a formar part del repositori de preservació i s'inclouen en les seves rutines de control.

A la base de dades del COFRE es guarden a més dades relatives a la càrrega que permeten, des de la perspectiva del repositori de preservació, millorar els processos de gestió de la validació d'aquestes peticions. Per exemple els catalogadors poden, a través de llistats, comprovar l'estat de validació dels documents dins d'un contenidor específic (que equival a un tipus de material específic o als números d'un mateix títol de publicació periòdica, veure figura 4), extreure únicament llistats de peticions pendents de validació o consultar estadístiques relatives als documents de preservació provinents de Dipòsit Legal.




































| <b>Time Out Barcelona (DL)</b>  |                   |   |                  |
|---|-------------------|---|------------------|
| <b>Títol</b>  | <b>Topogràfic</b> | <b>Bibliogràfic</b>   | <b>Validació</b> |
|  Time Out Barcelona ( 2018 / 493 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2018 / 495 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2018 / 496 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2018 / 494 )   |                   |   | Pendent          |
|  Time Out Barcelona ( 2018 / 492 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2018 / 490 )   |                   | b26104908    | Pendent          |
|  Time Out Barcelona ( 2018 / 491 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2018 / 489 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2018 / 488 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2017 / 475 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2017 / 477 )   |                   | b26104908    | Validat          |
|  Time Out Barcelona ( 2017 / 479 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 481 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 482 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 485 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2018 / 487 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 486 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 484 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 483 ) |                   | b26104908  | Validat          |
|  Time Out Barcelona ( 2017 / 480 ) |                   | b26104908  | Validat          |

FIG. 3 LLISTAT DE CÀRREGUES DE LA REVISTA TIME OUT BARCELONA AL REPOSITORI COFRE

Per una altra banda, els catalogadors, una vegada validat un document de DL a COFRE, s'encarreguen de catalogar-lo al catàleg de la BC i informar el número de registre bibliogràfic al document de COFRE. Aquest enllaç permetrà, una vegada desenvolupat el mòdul de visualització de documents de Dipòsit Legal, generar a través d'un procés automàtic els enllaços des del catàleg al mòdul de visualització.

#### 4.5. Alguns números

Des de la posada en marxa del mòdul de Dipòsit Legal, durant l'últim trimestre de l'any 2015, s'han realitzat 1.174 peticions de càrrega de documents de 32 editorials.

La figura 5 mostra les càrregues realitzades els anys 2016 i 2017 per tipus de documents, on es pot apreciar que la major part de les càrregues corresponen a publicacions periòdiques i llibres. Actualment s'han realitzat càrregues de 67 publicacions periòdiques amb un total de 735 exemplars i 382 monografies llibre. Cal destacar també l'entrada a partir del 2017 de documents audiovisuals i l'existència de 102 partitures provinents de l'editorial Clivis.

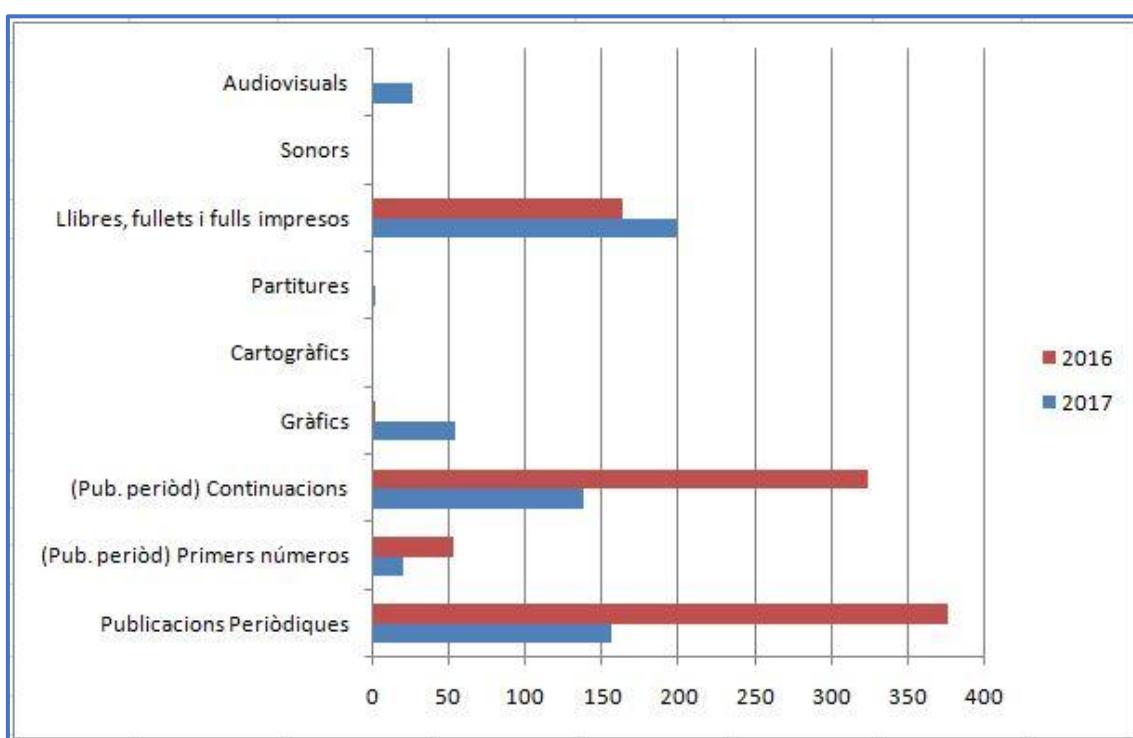


FIG. 4 SOL·LICITUDS DE CÀRREGA DE DL. DADES 2016-2017

La figura 6 mostra les càrregues des de la perspectiva del repositori de preservació COFRE, que aporta informació addicional, com el total de documents en accés restringit –el 30%- o domini públic. També ens dóna informació sobre el total de fitxers inclosos i l'ocupació en disc segons el format dels fitxers, on veiem que més del 60% de documents carregats contenen fitxers en format PDF. Els audiovisuals, en format mpeg4 i mov, ocupen el 20% del volum total. També podem trobar fitxers en format epub o mp3. Aquestes dades ajuden a la BC a fer una estimació de les necessitats d'emmagatzemament necessàries per a donar cabuda a les noves incorporacions de

documents. Dels resultats es pot despendre que el número de càrregues al repositori provinents de Dipòsit Legal encara pot créixer molt. Hi ha editors que han mostrat una predisposició gran i una actitud molt col·laborativa per a portar a terme les càrregues i millorar la interfície. D'altres tenen una actitud més reticent, per raons relacionades amb la seguretat i l'accés restringit als documents i cal per tant fer pedagogia i explicar-los molt bé el model de gestió del mòdul el qual es respectuós amb els drets vigents dels documents.

Per altra banda, la majoria d'editors amb una producció digital elevada, quant al seu volum, distribueixen les seves edicions digitals a través de plataformes, amb les quals encara cal trobar l'encaix per a que facin el dipòsit legal a la BC.

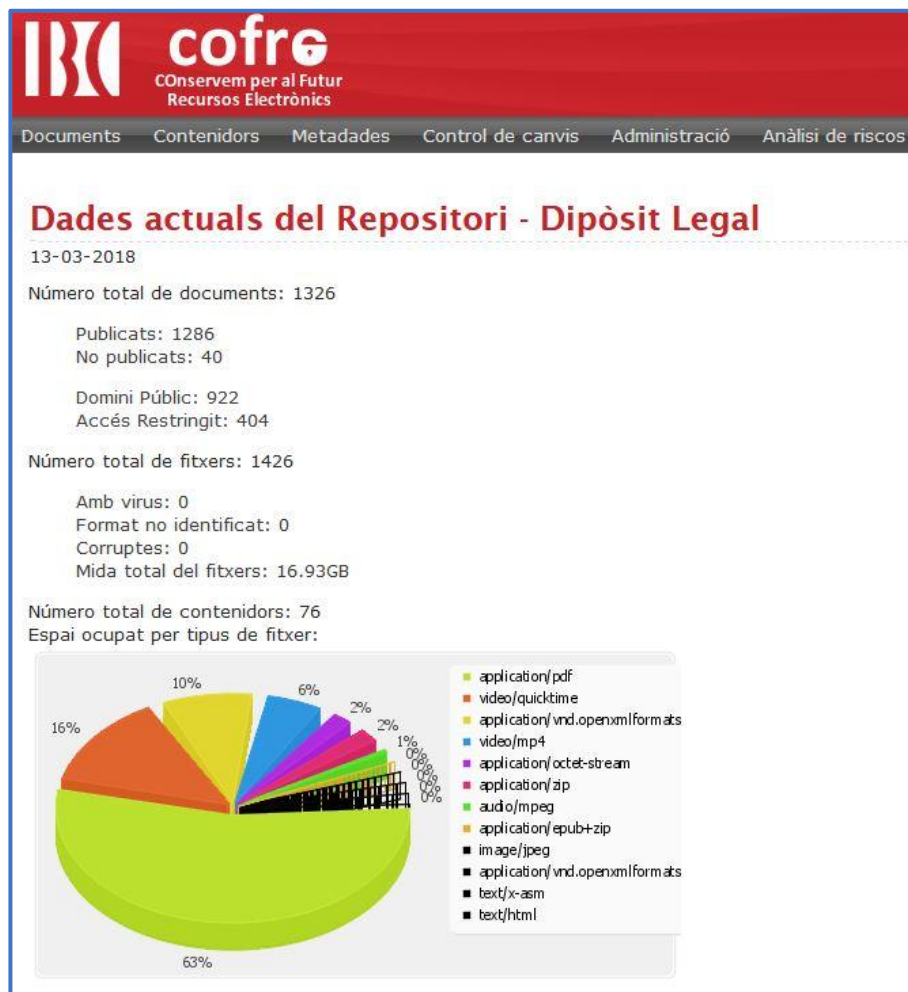


FIG. 5 DOCUMENTS CARRREGATS VIA DL AL REPOSITORI COFRE. MARÇ 2018

#### 4.6. Problemes detectats

Del total de càrregues del 2017, un 6% van ser rebutjades definitivament. El motiu de rebuig principal són càrregues que apareixen sense fitxer. Aquesta situació s'explica en la major part dels casos per un tancament prematur del navegador durant el procés de càrrega, quan una part dels processos ja han estat realitzats. Altres situacions que han comportat un rebuig de la càrrega han estat imatges de baixa qualitat o càrregues duplicades. Els rebuigs parcials se solen donar al carregar publicacions periòdiques com monografies i viceversa, que es reassignen manualment per a ser enviades als catalogadors que corresponen.

Una part d'aquestes situacions es podran prevenir fent canvis en el mòdul de càrrega que ajudin als editors a fer una selecció més acurada dels paràmetres i un ús més eficient de l'eina.

Cal destacar el fet que ara per ara no s'han trobat problemes relacionats amb l'accés als documents deguts a formats desconeguts o característiques de seguretat dels documents. Aquest fet s'explica per la tendència general a fer servir formats molt estesos com pdf, mpeg4, mp3, etc. Si bé esperem que en créixer el nombre de càrregues les casuístiques també s'incrementin.

#### 4.7. Millores i futurs desenvolupaments

El fet de ser, el mòdul de càrrega, un desenvolupament del personal de l'àrea de tecnologia de la Biblioteca de Catalunya, permet aplicar molt ràpidament correccions a errors o petites modificacions a partir dels comentaris dels editors: com la creació de nous camps de metadades o algun canvi visual per a facilitar la càrrega o que generava confusions.

Actualment el mòdul DL-UP només permet carregar documents d'un a un (amb un o més fitxers). Una de les peticions més demandada pels editors és la possibilitat de fer càrregues de paquets de diversos documents, adjuntant un document amb les metadades. Aquesta petició està en estudi actualment, així com la creació d'un servei web que permeti als editors interconnectar el mòdul DL-UP amb els seus programaris de gestió específics.

A més està planificada la creació per l'any 2018 d'un mòdul que permeti als usuaris que es trobin a les sales de lectura de la BC, la cerca i visualització de tots els documents digitals de Dipòsit Legal carregats a COFRE, ja siguin d'accés restringit o públics. Aquests punts de consulta, inicialment dos, s'aïllaran i adequaran per a impedir la còpia, enviament o impressió dels documents consultats, respectant així els drets d'autors i editors vigents, però equiparan la seva consulta als dels exemplars físics de la biblioteca que també són actualment d'accés restringit.

## 5. COL-LABORACIÓ EN LA CONSERVACIÓ DE LES PUBLICACIONS EN LÍNIA

El Reial decret 635/2015<sup>12</sup>, de 10 de juliol, com ja hem apuntat prèviament, és la regulació vigent a l'estat espanyol per al dipòsit legal de les publicacions en línia. Al seu article 10. explicita que:

- Els centres de conservació poden establir acords de col·laboració amb les entitats públiques o privades que disposin de plataformes de distribució en línia de les publicacions i recursos digitals, que siguin considerades «repositoris segurs», als efectes de conservació,
- Els acords que es portin a terme entre els centres de conservació i les entitats públiques o privades que mantinguin les plataformes esmentades han d'incloure entre altres aspectes els següents:
  - Les característiques del repositori de l'entitat col·laboradora.
  - Els procediments per a la conservació de les publicacions o els recursos de l'entitat col·laboradora.
  - El període de vigència de la col·laboració.
- En aquests casos, el centre de conservació no ha de capturar els recursos de l'entitat, i aquesta queda obligada a conservar les publicacions i recursos, així com a mantenir la plataforma activa i facilitar l'accés als seus recursos en les mateixes condicions que les establertes per als centres de conservació.
- En cas de desaparició de les entitats públiques o privades, els recursos en línia conservats en aquestes entitats s'han de lliurar al centre de conservació estipulat a l'acord.

La Biblioteca de Catalunya té la voluntat d'aplicar aquesta opció que proporciona la legislació, per a fomentar un sistema de preservació del dipòsit legal més sostenible, motiu pel qual ha establert recentment els requeriments mínims que les entitats

<sup>12</sup> Reial decret 635/2015, de 10 de juliol, pel qual es regula el dipòsit legal de les publicacions en línia. Boletín Oficial del Estado. Suplement en llengua catalana al núm. 177.  
[https://www.boe.es/boe\\_catalan/dias/2015/07/25/pdfs/BOE-A-2015-8338-C.pdf](https://www.boe.es/boe_catalan/dias/2015/07/25/pdfs/BOE-A-2015-8338-C.pdf).



públiques o privades que disposen de plataformes de distribució en línia i vulguin subscriure acords de col·laboració com a repositoris segurs, han de complir.

L'objectiu de la BC amb l'exigència dels requeriments fixats –des del punt de vista de la perdurabilitat dels recursos- és trobar un equilibri entre la garantia de preservació i la viabilitat de que les institucions i entitats s'hi puguin sumar. Per aquest motiu, abans d'establir de forma definitiva els requisits, s'ha testejat el document de mínims amb el repositori TDX del CSUC<sup>13</sup>, amb resultats satisfactoris.

Com a criteri general, les entitats han de fer el possible per a complir amb els estàndards internacionals existents sobre repositoris segurs (Data Seal of Approval<sup>14</sup>, DIN 31644<sup>15</sup> o ISO 16363<sup>16</sup>) més enllà dels requeriments mínims fixats.

Quant al requeriments mínims concrets són:

#### Infraestructura de dades

- Sistema de recuperació de dades davant possibles pèrdues degudes a avaries en el propi sistema (errors de disc, etc.) o per causes externes (inundacions, incendis,...). Cal disposar de còpies íntegres de la informació en diferents localitzacions geogràfiques, ja sigui amb baixa o alta disponibilitat. Es tracta de garantir la recuperació de tots els documents i metadades associades.
- Comptar amb processos de monitoratge de l'obsolescència dels suports físics i sistemes d'emmagatzemament.

#### Integritat de les dades

- Garantir la integritat de les dades durant la ingesta i a través de la comprovació periòdica de la signatura digital dels fitxers o altres mètodes alternatius que es donin en un futur.
- Comprovació de virus de tot el contingut.

<sup>13</sup> TDX (Tesis Doctorals en Xarxa). [http://tdx.cat/?locale-attribute=ca\\_ES](http://tdx.cat/?locale-attribute=ca_ES). És un repositori cooperatiu acadèmic de tesis doctorals llegides a les universitats de Catalunya i d'altres comunitats autònomes. Donat que està coordinat pel CSUC, col·laborador habitual de la Biblioteca de Catalunya, va semblar un cas adequat per a realitzar el testeig de requeriments.

<sup>14</sup> Data Seal of Approval. <http://www.datasealofapproval.org/en/>

<sup>15</sup> DIN Standards Committee Information and Documentation. *DIN 31644. Information and documentation - Criteria for trustworthy digital archives*. <http://www.din.de/en/getting-involved/standards-committees/nid/standards/wdc-beuth:din21:147058907>

<sup>16</sup> International Organization for Standardization. *ISO 16363:2012. Space data and information transfer systems -- Audit and certification of trustworthy digital repositories*. [http://www.iso.org/iso/catalogue\\_detail.htm?csnumber=56510](http://www.iso.org/iso/catalogue_detail.htm?csnumber=56510)



# 15es JORNADES CATALANES D'INFORMACIÓ I DOCUMENTACIÓ

Barcelona, 10 i 11 de maig de 2018

- Capacitat per a recuperar o reparar les dades danyades.

## Seguretat de la informació

- Sistema de control d'accessos amb diferents nivells d'autorització.
- Manteniment d'un registre de qui ha realitzat quines accions sobre els fitxers, incloent eliminacions i accions de preservació.
- Eliminació segura: control i validació en l'eliminació de continguts – al menys dues persones diferents confirmen que l'eliminació és procedent – .

## Metadades

- Guardar metadades de preservació (tècniques, descriptives, administratives, transformatives i d'esdeveniments sobre les dades) en formats estàndards de metadades.

## Format dels fitxers

- Mantener un inventari dels formats dels fitxers en ús, mantenir una actitud de vigilància respecte a l'obsolescència dels formats dels fitxers i realitzar migracions de format, emulacions o accions similars segons sigui necessari per a mantenir l'accés íntegre a les dades.

## **6. A MODE DE CONCLUSIÓ**

El dipòsit legal de publicacions digitals nascudes en línia, a la Biblioteca de Catalunya i la resta del món té molt camí a recórrer i, molt probablement, la manera de gestionar-lo i preservar-lo requerirà una permanent capacitat evolutiva per part dels centres conservadors. El canvi de formats, productes, models de producció, gestió, i mitjans de difusió que s'observa al món digital condicionarà aquesta evolució.

El creixement de la producció digital, la tendència a la cultura col·laborativa i la necessitat de solucions de sostenibilitat, recomanaran posar l'accent en la creació d'una bona xarxa d'entitats col·laboradores tal i com preveu la legislació. Al final, l'encert raurà en l'adaptabilitat dels sistemes propis que la BC utilitzi i en la seva capacitat de promoure una comunitat sòlida, tecnològicament competitiva i adaptativa que contribueixi en la tasca de conservació de la producció digital.