

## MDC. Com carregar llibres antics o incunables

(Ciència Oberta\3 Repositoris\Repositori MDC\Docs. tecnicos\ProcedimentLlibresAntics\_1805.docx, 17.05.18)

### Introducció

En aquest vídeo (<http://www.youtube.com/watch?v=9T8m7ONYrjc>) es veurà com publiquem a la Memòria Digital de Catalunya (MDC) una col·lecció digitalitzada de llibres antics o incunables. La publicació es realitza a través del Project Client que és el programari que usa ContentDM per introduir les metadades i les imatges dels documents.

### Procediment

Els llibres cal carregar-los a l'MDC com objectes compostos. És a dir, una portada (la del llibre) amb una sèrie de pàgines que el formen.

Es desplega l'opció "Add" del menú superior del Project Client i a continuació es selecciona l'opció "Compound Objects".

Apareix una pantalla amb un format de graella on, un cop carregats els objectes, apareixeran llistats. Cal assegurar-se que en el desplegable superior estigui seleccionada l'opció "Compound Object Wizard" i es fa clic a "Add".

La següent pantalla és per escollir el tipus d'objecte compost, en aquest cas s'escolleix la segona opció: Monograph. Document with hierarchy, such as chapters in a book (Monografia. Document amb jerarquia, com els capítols d'un llibre).

En la mateixa pantalla, hi ha una altra opció que serveix per informar les metadades de l'objecte a través de fitxers de text amb tabulacions al mateix temps que aquest es carrega. En aquest cas seleccionem l'opció NO, ja que no disposem d'aquests fitxers amb les metadades.

Després de fer clic a "Next" apareix la pantalla des d'on carregarem les imatges. Fent clic a "Browse" es desplega una altra pantalla on seleccionarem el directori que conté les imatges digitalitzades que componen l'objecte.

Com es pot veure aquest directori conté tant les imatges com els fitxers de text amb la transcripció corresponent a les imatges digitalitzades (OCR). Abans de carregar aquest directori, cal separar les imatges i els fitxers de text en dues carpetes diferents. Així, amb l'explorador de Windows, seleccionem el directori que conté la col·lecció; creem dues carpetes, una corresponent a les imatges i l'altra als fitxers de text i movem el contingut corresponent a cada carpeta.

Repetim aquest procés per tots els objectes compostos que vulguem publicar. Per exemple: anem a la carpeta corresponent a "Francesc Ferrer Nogués", creem les carpetes 'img' i 'txt', ordenem el contingut segons el tipus (per així poder moure amb més facilitat els fitxers a la carpeta corresponent) i movem el contingut. Recordeu que per a que el programa pugui relacionar la imatge amb el text corresponent cal que els fitxers tinguin el mateix nom però cadascun amb la seva extensió (per exemple: imatge1.jpg –fitxer que correspon a la imatge- i imatge1.txt –fitxer que correspon al text-).

Un cop separat el contingut en imatges i text, ja podem tornar al Project Client i continuar el procés de publicació de continguts. Seleccionem la carpeta de les imatges, acceptem i passem a la següent pantalla, en la qual escollim l'opció que ContentDM NO creï una miniatura automàticament. D'aquesta manera la miniatura que es generarà serà la de la imatge que nosaltres carreguem. Fem clic a "Next".

La següent pantalla està dividida en 3 parts: La primera, on especifiquem que s'utilitzi el nom del fitxer com a nom de la imatge publicada. La segona, on seleccionem el directori corresponent als fitxers de text i la tercera, opció que crea automàticament un PDF de les imatges. Recomanem activar aquesta opció per aquest tipus de col·lecció. Cal tenir en compte però, que aquesta generació automàtica de PDF funciona només òptimament per a formats DIN-4.

La última pantalla del procés és un resum que se'ns mostra abans d'aprovar tot el treball fet fins al moment. Si hi estem d'acord procedim amb la creació de l'objecte compost per a ser publicat clicant a "Finish". Com a resum del procés de creació, apareix una pantalla amb errors o, si no n'hi ha, amb el nombre de pàgines que contindrà l'objecte.

Tancant aquesta pantalla, ja podem veure que s'ha carregat l'objecte compost en la graella del principi del procés. Ara podem continuar carregant altres objectes compostos seguint el mateix procediment que fins ara. Continuem, per exemple, amb la càrrega del llibre antic d'en Francesc Ferrer Nogues..... No es detallarà el procés ja que és el mateix que l'explicat anteriorment (a mode de resum):

- Afegir objecte compost seleccionant monografia
- Seleccionar la carpeta d'imatges d'aquesta col·lecció
- Seleccionar l'opció que ContentDM NO creï la miniatura automàticament
- Seleccionar el directori de fitxers de text corresponents a l'OCR

En aquest cas, un cop finalitzat el procés, apareixen una sèrie de missatges d'alerta, avisant que hi ha fitxers de text que no tenen correspondència amb cap imatge. S'hauria de revisar ja que hi hauria d'haver el mateix nombre d'imatges que de fitxers de text.

Suposem que ja no volem pujar cap més objecte, finalitzem el procés fent clic al botó "Finish" de la pantalla resum d'objectes compostos. Un cop processats aquests objectes apareix una altra pantalla amb un resum de l'operació o, si fos el cas, amb els missatges d'error que hi hagués hagut.

En aquest punt ja es poden veure els 2 objectes compostos carregats: 2 llibres antics digitalitzats.