

Digitizarea documentelor de patrimoniu

Drd. Nicoleta-Roxana Dinu

Școala Doctorală de Litere, Universitatea din București,

Str. Edgar Quinet, nr. 5-7, sector 1, București. E-mail: dnicole_2005@yahoo.com

Patrimoniul cultural constituie identitatea unei națiuni, de-a lungul istoriei. În „Recomandarea privind protecția bunurilor culturale mobile”, adoptată pe 28 noiembrie 1978 de către UNESCO, noțiunea de „bunuri culturale mobile” cuprinde „toate bunurile mobile care sunt expresia sau mărturia creației umane sau a evoluției naturii și care au o valoare sau un interes arheologic, istoric, artistic, științific sau tehnic [1]”.

Conform legislației românești [2], patrimoniul cultural național este definit în sens larg și anume: „ansamblul bunurilor identificate ca atare, indiferent de regimul de proprietate asupra acestora, care reprezintă o mărturie și o expresie a valorilor, credințelor, cunoștințelor și tradițiilor aflate în continuă evoluție.” De asemenea, „cuprinde toate elementele rezultate din interacțiunea, de-a lungul timpului, dintre factorii umani și cei naturali”.

În vederea păstrării și valorificării memoriei colective, instituțiile culturale au adoptat metoda digitizării, aceasta devenind una dintre activitățile-cheie. Tot mai multe biblioteci includ, printre obiectivele fundamentale, și digitizarea documentelor (cu precădere, a celor patrimoniale), cu scopul de a furniza acces la conținutul digital creat (stimulându-se, astfel, procesul de învățare/ cercetare) și de a prezerva documentele extrem de valoroase, aflate într-o stare degradabilă. Totodată, prin punerea la dispoziția utilizatorilor a conținutului patrimonial în spațiul virtual și prin încorporarea aplicațiilor web 2.0, biblioteca transcende limitele instituționale, devenind vizibilă comunităților academice internaționale.

O definiție complexă a digitizării este oferită de Ghidul Calimera: digitizarea presupune conversia documentului pe suportul inițial, analogic, în format digital, descrierea și reprezentarea obiectelor și a documentației referitoare la acestea, procesarea, asigurarea accesului la conținutul digitizat și prezervarea pe termen lung [3].

Cu ajutorul tehnologiilor informației și comunicării, digitizarea poate fi percepută „mai mult decât un substitut al unei pagini consultate, ea poate genera, de asemenea, noi oportunități de cercetare [4]”. Imaginile unui text pot face trimiteri la alte texte, în relație cu aceeași temă, la alte puncte de vedere ale diferiților autori.

Conform specificațiilor Consiliului Canadian al Arhivelor (Canadian Council of Archives) [5], digitizarea:

- încurajează prezervarea prin limitarea manevrării documentelor originale;
- este parte a strategiei de acces;
- nu trebuie să provoace daune documentelor originale;
- nu trebuie să încalce drepturile intelectuale ale unei persoane;
- trebuie să ajute la prezervarea documentelor, prin crearea de copii destinate consultării și a unei alte copii (copia master), la care să nu aibă acces decât un număr restrâns de persoane (personalul bibliotecii).

Digitizarea este văzută ca un întreg set de avantaje aduse bibliotecilor, printre care: acces non-stop și simultan la documente, promovarea colecțiilor și atragerea de noi utilizatori, acces la materiale inaccesibile până atunci, prezervarea de materiale fragile și rare, stimularea inovării și colaborării în mediul academic.

Deși orice instituție culturală își propune să deruleze proiecte de digitizare a celor mai valoroase colecții ale ei, adesea, întâmpină obstacole diverse, precum:

- costuri foarte ridicate;
- obsolescența rapidă a tehnologiilor digitale;
- riscul încălcării dreptului de autor;
- potențial mare de acces neautorizat;
- posibilitatea pierderii sau introducerii de date incorecte;
- insuficiența metadatelor (fapt ce reduce valoarea materialului);
- dificultăți în citirea pe ecran;
- greutatea cu care se pot realiza adnotări, direct pe ecran.

Digitizarea nu trebuie realizată în izolare, iar din componența echipei proiectului trebuie să facă parte și experți în prezervare/ conservare, întrucât aceste aspecte ar trebui să fie parte integrantă a procesului.

Conform Ghidului Calimera, în selecția documentelor se au în vedere: cadrul legislativ; tipul de public căruia se adresează documentele digitizate; importanța materialelor în patrimoniul cultural al comunității, regiunii, țării, spațiului european, universal; unicitatea materialelor și prezervarea originalelor [6].

Modalitățile de punere la dispoziție a utilizatorilor a documentelor tipărite pe hârtie, prin digitizare, pot fi fotografierea sau scanarea (în varianta imagine sau varianta text). Fiecare dintre cele două variante prezintă anumite particularități.

În privința Imaginii, aceasta garantează reproducerea identică a documentului. Reprezentarea în modulul imagine, în fișiere în format PDF, permite indexarea conținutului,

dând posibilitatea motoarelor de căutare să caute în interiorul documentului. Ca formate ale imaginilor fixe, putem enumera:

- *TIFF* (Tagged Image File Format), [extensie .tif] este formatul destinat prezervării, de aceea pentru fiecare operațiune de digitizare se recomandă stocarea rezultatului sub formă de fișiere tif. Acest format păstrează integritatea informațiilor generate prin digitizare (prin acest format se salvează imagini de calitate excepțională), fișierele conservate putând servi, astfel, ca bază pentru utilizări ulterioare. Spațiul de stocare nu reprezintă o problemă;

- *JPEG* (Joint Photographic Experts Group), [extensie .jpg]: deoarece permite comprimarea anumitor informații conținute în imagine, formatul acesta este destinat, mai degrabă, consultării decât prezervării;

- *GIF* (GraPhical Interchange Format), [extensie .gif]: este un format comprimat pentru imagini, realizat de către CompuServe Interactive Services, în anul 1987. Fotografiile salvate în format gif. pierd foarte multe informații despre culoare, nu se poate face deosebirea subtilă dintre nuanțele aceleiași culori. Este utilizat pentru logo-uri, desene, grafică;

- *PNG* (Portable Network GraPhics), [extensie .png]: este un alt format de compresie fără pierdere de informații. Este recomandabil pentru imagini din mediul web;

- *BMP* (Bitmap): este unul din formatele dezvoltate de către Microsoft Windows. De aceea, multe aplicații tip imagine, grafice care funcționează în sistemul de operare Windows ajută la crearea fișierelor în format BMP;

- *PICT*: dezvoltat în 1984 cu scopul de a fi folosit pentru calculatoarele Apple Macintosh, acest format este folosit, în principal, pentru schimbul de grafice între aplicațiile diverse ale Macintosh.

În această tipologie, poate fi inclus și formatul *PDF* (Portable Document Format) [extensie .pdf]. Deși nu este un format de fișier imagine, este foarte utilizat deoarece permite conservarea afișării formei documentului original, indiferent care este platforma, softul sau tipul de caractere folosite pentru diseminarea lui. Conținutul unui fișier PDF poate fi indexat pentru a se realiza căutări. De obicei, există diferite nivele de protecție a unui document PDF, de exemplu, se poate împiedica sau se poate permite extragerea unei părți din document. Formatul PDF/ A (Archival) este destinat prezervării.

În privința modului Text, după scanarea/ fotografierea documentelor, textul poate fi reutilizat, prin folosirea unui software de Recunoaștere Optică a Caracterelor - Optical Character Recognition (OCR). Alegerea modului text prezintă avantaje precum: identificarea rapidă și exactă a ediției utilizate (prin adăugarea unei note bibliografice atașate la document).

Apărut în anul 1950, în Statele Unite ale Americii, OCR poate fi folosit: în mod automat (după fiecare scanare) sau manual (pentru documentele scanate cu ajutorul altui program).

Printre sistemele OCR din prezent se numără ABBYY Fine Reader [7]. Versiunea 8.0 oferă ca facilități: procesarea ușoară a documentelor în format pdf, recunoașterea fotografiilor realizate cu camera digitală și recunoașterea optică a caracterelor, printr-un singur click. Versiunea 11 (la care s-a ajuns în prezent) simplifică citirea și publicarea cărților electronice. De asemenea, poate recunoaște documente în 1710 limbi, cu verificarea ortografiei în 36 de limbi.

În privința documentelor multimedia, care trebuie reformatate, se poate alege unul dintre standardele pentru fișierele audio/ video digitale:

- *AVI* (Audio Video Interleave), [extensie .avi]: creat de către compania Microsoft, în anul 1992, constituie primul format de fișiere video de dimensiune mare;

- *FLV* (Flash Video Format), [extensie .flv]: produs în 2002, poate fi vizualizat fie prin folosirea unui program de redare separat, fie printr-un program inserat în alt program principal (așa-numitele „plug-in”-uri);

- *MOV* [extensie .mov]: este un format pentru fișiere video, realitate virtuală, animații, conținut în 3D;

- *MP4* [extensie .mp4] și *MPEG 4* (Moving Pictures Expert Group): permit o comprimare avansată a fișierelor mai vechi MPEG2.

O etapă foarte importantă în procesul de digitizare o reprezintă crearea metadatelor, o activitate destul de costisitoare, întrucât antrenează timp îndelungat în elaborarea lor și surse financiare (pentru a avea personal calificat în acest sens).

Metadatele reprezintă date despre date care au rolul de a identifica și descrie orice resursă digitală. Metadatele se împart în trei categorii [8]:

1. Metadate descriptive - utilizate în descoperirea și identificarea unui obiect. Pot conține elemente precum: titlu, abstract, autor, cuvinte-cheie;

2. Metadate structurale - utilizate pentru a afișa un obiect și a naviga către un anumit obiect. Ele includ, de asemenea, informații despre partea organizatorică a obiectului (de exemplu, componența cărții: prefață, cuprins, capitole, index);

3. Metadate administrative - cuprinde informații care ajută la gestionarea resursei. De exemplu: data la care a fost creat obiectul, formatul conținutului (de exemplu: TIFF, JPEG, PDF), rezoluția scanării, informații privind drepturile de utilizare. Aici există câteva subseturi de metadate administrative, dintre care două sunt prezentate adesea ca metadate distincte: metadatele gestionării proprietății intelectuale și metadatele de preservare care conțin informații necesare arhivării și preservării resursei.

Procesul de digitizare nu se încheie odată cu generarea documentului digital, ci se continuă cu asigurarea accesului, în timp, la informațiile deținute de acesta. Preservarea digitală presupune realizarea unui set complet de măsuri care să protejeze și să furnizeze acces atât la

documentele digitizate, cât și la cele create numai în mediul digital. Se pot observa câteva particularități ale prezervării documentelor digitale:

- dependența automatizării – materialele digitale cer utilizarea anumitor hardware-uri și software-uri pentru a putea fi consultate;

viteza de schimbare a tehnologiilor constituie o amenințare la procedura de acces la un document;

- fragilitatea noilor medii de stocare a informațiilor – numeroase astfel de suporturi digitale, deși fizic nu prezintă nicio neregulă, pot deveni instabile în accesarea informațiilor pe care le dețin;

- criteriile de prezervare a documentelor pe suport digital sunt mult mai severe decât pentru cele pe suport tradițional, deoarece o resursă digitală care nu a fost selectată în programul de prezervare va fi mai mult ca sigur pierdută sau inaccesibilă în viitorul apropiat.

Pentru ca întreg procesul de digitizare să fie unul de succes, trebuie stabilită o strategie la nivel instituțional, prin care să fie stabilite toate etapele necesare derulării acestuia și prin care să se țină cont de cadrul legal privind drepturile de proprietate intelectuală, cerințele tehnice, personalul ce alcătuiește echipa și resursele financiare.

Deși este considerat un proces destul de costisitor, digitizarea a creat noi oportunități de regândire a procesului de prezervare, prin reducerea manevrării documentelor fragile, și a condus la crearea de noi metode de acces la documentele de patrimoniu, fapt ce a revoluționat cercetarea științifică, la nivel mondial.

Note

¹ UNESCO. *Recommandation pour la protection des biens culturels mobiliers*, 28 novembre 1978, art. 1 (a) [online]. [Accesat la data de 13.02.2012]. Disponibil pe internet la adresa: http://portal.unesco.org/fr/ev.php-URL_ID=13137&URL_DO=DO_TOPIC&URL_SECTION=201.html.

² Legea nr. 182/2000 privind protejarea patrimoniului cultural național mobil, republicată în 2008. În: *MO*, Partea I, nr. 828, 09/12/2008, art. 1 (2) Legea nr. 182/2000 privind protejarea patrimoniului cultural național mobil, republicată în 2008. În: *MO*, Partea I, nr. 828, 09/12/2008, art. 1 (2).

³ PROIECT Calimera. *Ghiduri de bune practici* [online]. Traducere din engleză a Ghidului *Calimera Guidelines: cultural applications: Local Institutions Mediating Electronic Resources*, p. 198. [Accesat la data de 17.03.2012]. Disponibil pe internet la adresa: <http://www.bjc.ro/new/files/calimera/ghidurile%20calimera.pdf>.

⁴ HIRTLE, Peter B. The impact of digitization on special collections in libraries. În: *Libraries and Culture*, vol. 37, nr. 1, 2002, p. 44.

⁵ CANADIAN Council of Archives. *Digitization and Archives*, p. 1, 2002. [online]. [Accesat la data de 22.03.2012]. Disponibil pe internet la adresa: http://www.cdncouncilarchives.ca/digitization_en.pdf.

⁶ PROIECT Calimera. *Op. cit.*, p. 201.

⁷ Pagina-gazdă: <http://www.abbyy.com/>. [Accesat la data de 23.03.2012].

⁸ UNIVERSITY of Kansas. *Digital Initiatives – Recommended Standards and Best Practices for Digital Projects* 2003, p. 3. [online]. [Accesat la data de 21.12.2011]. Disponibil pe internet la adresa: http://kudiglib.ku.edu/planning/dli_standards.pdf.

Bibliografie

CANADIAN Council of Archives. *Digitization and Archives*, 2002. 12 p. [online]. Disponibil pe internet la adresa: http://www.cdncouncilarchives.ca/digitization_en.pdf.

HIRTLE, Peter B. The impact of digitization on special collections in libraries. În: *Libraries and Culture*, vol. 37, nr. 1, 2002, p. 42-52.

Legea nr. 182/2000 privind protejarea patrimoniului cultural național mobil, republicată în 2008. În: MO, Partea I, nr. 828, 09/12/2008.

PROIECT Calimera. *Ghiduri de bune practici* [online]. Traducere din engleză a Ghidului *Calimera Guidelines: cultural applications: Local Institutions Mediating Electronic Resources*, 384 p. Disponibil pe internet la adresele: <http://www.bjc.ro/new/files/calimera/ghidurile%20calimera.pdf> și <http://www.abbyy.com/>.

UNESCO. *Recommandation pour la protection des biens culturels mobiliers* [online], 28 novembre 1978. Disponibil pe internet la adresa: http://portal.unesco.org/fr/ev.php-URL_ID=13137&URL_DO=DO_TOPIC&URL_SECTION=201.html.

UNIVERSITY of Kansas. *Digital Initiatives – Recommended Standards and Best Practices for Digital Projects* [online]. 2003, 8 p. Disponibil pe internet la adresa: http://kudiglib.ku.edu/planning/dli_standards.pdf.