

STRUCTURA ȘI CONȚINUTUL UNEI BIBLIOTECI DIGITALE

Valentina Nastas

Biblioteca U.T.M.

Prezervarea digitală este o problemă vitală pentru societatea informațională, unde cantitatea informației crește rapid și unde conținutul devine din ce în ce mai dinamic. Sarcina de bază în digitizarea colecțiilor revine bibliotecilor, în special celor academice.

Cuvinte cheie: *digitizare, bibliotecă digitală, sistemul de metadate Dublin Core, proiectul SINRED, structuri GRID.*

Digitizarea reprezintă un proces de conversie a unui element – fie că acesta este un text tipărit, un manuscris, o imagine, un sunet, un film sau o înregistrare video – de la un format analog la un format digital. Rezultatul digitizării este numit **reprezentare digitală**, imagine digitală pentru un obiect sau formă digitală pentru un sunet.

Digitizarea într-o bibliotecă devine parte integrantă a proceselor de conservare, prelucrare și distribuție a informațiilor.

O bibliotecă digitală este o bibliotecă în care colecțiile sunt stocate în formate digitale (text, audio, video, documente compuse, obiecte digitale sau colecții). Bibliotecile digitale sunt sisteme informatice complexe care acoperă aspectele legate de crearea, stocarea, procesarea, distribuția și accesul la date.

Termenul de **bibliotecă digitală** a fost utilizat prima dată în anul 1988 într-un raport al Corporației pentru Cercetarea Inițiativelor Naționale (*Corporation for National Research Initiatives*), dar a fost popularizat în anul 1994 de către Inițiativa Bibliotecilor Digitale (*Digital Library Initiative*) NSF/DARPA/NASA. Denumirile de **bibliotecă electronică** sau **bibliotecă virtuală** sunt folosite ocazional și sunt denumiri care preced în denumirea de bibliotecă digitală. Site-urile agențiilor guvernamentale deseori fac referire la **bibliotecă electronică** (ex. *Florida Electronic Library*).

Multe din cele mai cunoscute biblioteci digitale sunt mai vechi decât Web-ul (de ex. *Project Perseus*, *Project Gutenberg*, și *ibiblio*). Pe măsura dezvoltării tehnologiilor de lucru în rețea, a creșterii potențialului motoarelor de căutare bibliotecile digitale se dezvoltă în mediul bazat pe Web (ex. *European Library*, *Library of Congress*).

Decizia de a digitiza publicațiile dintr-o bibliotecă are la bază următoarele motive:

- sporirea accesului: acesta este motivul principal și cel mai evident, acolo unde există un număr ridicat de cereri din partea utilizatorilor, iar bibliotecă sau arhivă dorește să îmbunătățească accesul la o anumită colecție.
- reducerea utilizării materialelor fragile sau foarte uzate, și pentru crearea unei copii „de rezervă” pentru materialele aflate în pericol, precum cărți sau documente fragile.

- crearea colecțiilor virtuale care pot fi partajate în scopul colaborării cu alte instituții, pentru a lărgi parteneriatul.

Bibliotecile digitale organizează colecții cu conținut digital și le fac disponibile pentru public. Colecțiile pot conține atât materiale care au fost digitizate, cum ar fi copiile digitale ale publicațiilor “fizice”, cât și materiale bazate pe informații produse original în format digital (digital-născute). Acesta este în mare parte cazul informației din domeniul științific, unde publicațiile digitale sunt stocate în depozitarii digitale. Bibliotecile care dețin colecții fizice și colecții digitale sunt numite **biblioteci hibride** (de ex. *American Memory* este o bibliotecă digitală în cadrul *Library of Congress*). Unele biblioteci digitale dețin arhive stocate pe termen lung cum ar fi *EPrint*, *arXiv*, *Internet Archive*.

Multe biblioteci academice se implica activ în crearea Repozitelor Academice (*Academic Repositories*). Acestea conțin publicații (cărți, manuale, lucrări, teze ș.a) care au fost convertite din mediul fizic în cel digital sau sunt într-un format digital – născut.

Pentru crearea de repozite pot fi utilizate soft-uri speciale cum ar fi: *DSpace*, *EPrints*, *Fedora*, *dLibra* și *Greenstone Digital Library Soft*.

Arhivele digitale diferă de bibliotecile digitale. Tradițional arhivele digitale conțin colecții care cuprind surse primare de informații (ex. scrisori, rapoarte ș.a. lucrări) produse direct de autor (persoană fizică sau juridică). Entitățile într-o arhivă sunt grupate conform originii dar nu sunt catalogate ca entități individuale cum ar fi în cazul unei biblioteci digitale. Cea mai veche arhivă digitală de surse primare academice este *Oxford Text Archive*. *Project Gutenberg*, *Google Book Search*, *Windows Live Search Books*, *Internet Archive*, *CMU's Universal library* sunt considerate ca lideri în domeniul creării și gestionării arhivelor digitale.

Sunt inițiate proiecte de digitizare pe scară largă de către *Google - Million Book Project*, *MSN*, și *Yahoo*.

Digitizarea necesită o muncă intensă și costisitoare, cere investiții considerabile care în cele mai multe cazuri depășesc mijloacele instituțiilor deținătoare de informații.

O problemă tehnică cheie este necesitatea îmbunătățirii tehnologiilor de digitizare din punct de vedere al eficienței și accesibilității.

Digitizarea presupune efectuarea unei copii, care poate fi problematică vizavi de dreptul de proprietate intelectuală. Așa cum prezervarea depinde de copyright, problema trebuie examinată în lumina legislației respectării dreptului de autor. Nu este posibilă crearea unei biblioteci online care să ofere materiale ce nu sunt în domeniul public, fără schimbări esențiale în legislația copyright, sau acorduri cu proprietarii de drepturi.

Publicațiile electronice disponibile în rețele se supun exigențelor impuse de metadata, adică de acel sistem de reprezentare a elementelor de identificare, de reprezentare a conținutului și de acces a documentelor în mediul electronic.

Sunt trei categorii de metadata ce sunt frecvent utilizate pentru a descrie obiecte în biblioteca digitală:

1. descriptive – informația care descrie conținutul intelectual al obiectului, ca de exemplu înregistrări de catalogare de tip MARC;
2. structurale - informația care indică legătura fiecărui obiect cu alte obiecte în scopul alcătuirii unității logice (ca ex. informația care leagă imaginile din pagina unei cărți cu alte imagini din carte);
3. administrative - informația se folosește la administrarea obiectului sau controlul accesului la acesta; acesta poate cuprinde informația despre cum a fost scanat obiectul, formatul de stocare, informația privind dreptul de autor sau licență precum și informația de preservare pe lungă durată a obiectului digital.

Sistemul de metadata **Dublin Core** are cea mai largă răspândire și este acceptat și dezvoltat de specialiștii în științele informării și comunicării. Acest sistem integrează atât formate MARC cât și formate nonMARC; permite căutarea în text integral dar și navigarea hypertextuală; este foarte bine adoptat pentru date bibliografice text ce sunt structurate. Este independent de platformă și de programe; permite reutilizarea de date. Principalul dezavantaj: este dificil de aplicat pentru texte juridice și tehnice ce au o structură logică cu multiple niveluri și trimiteri.

În vederea stimulării dezvoltării societății informaționale la nivel național trebuie de încurajat crearea de conținut digital atât la nivelul bibliotecilor, arhivelor cât și a muzeelor. Pentru aceasta este necesară elaborarea unei politici de digitizare a resurselor culturale care să propună un model de sistem informatizat de management a resurselor digitale.

Biblioteca Digitală a României va reprezenta unul dintre cele 27 de noduri naționale, componente ale Bibliotecii Digitale Europene, ce urmează să fie create.

În România, există deja o serie de inițiative în domeniul digitizării resurselor culturale. Dintre aceste inițiative se remarcă o serie de proiecte derulate de către Ministerul Culturii și Cultelor și unele Biblioteci. Consorțiul creat în cadrul proiectului CEEEX, SINRED și-a propus să definească și să realizeze un sistem național unitar de management al resurselor digitale în știință și tehnologie bazat pe structuri GRID.

Beneficiile scontate prin derularea acestui proiect sunt: acces printr-o interfață unitară la informațiile publice, tehnice și științifice; dezvoltarea de proceduri de creare, prelucrare, utilizare, diseminare a resurselor digitale full-text; facilitarea unor parteneriate naționale și internaționale bazate pe schimbul de informații; susținerea activității didactice și cercetării; dezvoltarea domeniului Științelor Informării și Comunicării la nivel național; diversificarea tipologiilor de produse și servicii de informare și documentare.

Abordarea pe bază de ontologii a bibliotecilor digitale pentru domeniul științific și tehnic din prisma reprezentării prin structuri GRID cuprinde următoarele concepte de bază:

- **Obiecte digitale:**
 - asociere de conținut, metadata și proceduri de prelucrare și de acces a procedurilor;
- **Colecții digitale:**
 - asocierea pe baza unui anumit criteriu a mai multor obiecte digitale;
- **Evenimente:**
 - conținut asociat unei anumite manifestări de scurtă durată (exemple: conferințe, workshop-uri, seminarii);

- **Procese:**
 - continut asociat unei activitati de durata (exemple: proiecte, cursuri);
 - **Organizații virtuale:**
 - Roluri;
 - Utilizatori;
- Digitizarea tuturor materialelor relevante este imposibilă, deci trebuie de selectat ce este necesar de digitizat.

Politica publică privind digitizarea resurselor culturale naționale și crearea Bibliotecii Digitale a României propunere următoarele criterii de selecție a documentelor/colecțiilor ce urmează a fi digitizate:

1. Valoarea documentară (valoarea/importanța conținutului intelectual, valoarea bibliofilă);
2. Reprezentativitatea pentru un anumit domeniu, o anumită perioadă de timp sau o anumită regiune geografică, atât la nivel național, cât și la nivel internațional;
3. Gradul de interes: grupul de utilizatori (reali sau potențiali) trebuie să fie relativ mare pentru a justifica selecția unui document în lotul celor propuse pentru scanare;
4. Prezervarea documentelor originale;
5. Regimul juridic (dreptul de autor și dreptul de difuzare);
6. Criteriul financiar;
7. Soluția de scanare și tehnologia implicată.

Sunt necesare servicii adecvate care să permită utilizatorului să descopere conținutul și să lucreze cu el. Aceasta implică descrierea și indexarea calitativă a conținutului.

Biblioteca digitală trebuie să permită utilizatorului să poată localiza documentele digitizate oricare ar fi tipul de document și să le consulte în mod liber, fie integral în cazul operelor neprotejate prin dreptul de autor, fie sub formă de scurte fragmente în cazul operelor protejate, cu acordul titularilor drepturilor.

Internetul ne incită la o abordare colaborativă, cooperativă inclusiv pentru activitățile de învățământ și de cercetare.

Web-Bibliografie:

<http://en.wikipedia.org/wiki/Digitisation> - Digitizing, from Wikipedia, the free encyclopedia.

http://en.wikipedia.org/wiki/Digital_library - Digital library, from Wikipedia, the free encyclopedia.

http://en.wikipedia.org/wiki/Institutional_repository - Institutional repository, from Wikipedia, the free encyclopedia.

<http://www.bnrm.md/memoria/b10.7.php> - COMUNICATUL COMISIEI EUROPENE "i 2010: Biblioteci Digitale"

<http://www.cultura.ro/Files/GenericFiles/PPP-Digitizare.pdf> - Propunere de politică publică privind digitizarea resurselor culturale naționale și crearea Bibliotecii Digitale a României.

http://www.ici.ro/RRIA/ria2006_4/index.html - Sistem național de management al resurselor digitale în știință și tehnologie, bazat pe structuri GRID – SINRED.