

# GHID DE DIGITIZARE – PILONUL TEMATIC ”BIBLIOTECI” - versiunea 01 / 30.10.2009 -

## I. Aspecte teoretice

1. Introducere.....	2
2. Etapele desfășurării procesului de digitizare.....	3
3. Resurse informaționale și documentare în bibliotecile din România.....	3
4. Criteriile de selecție.....	3
5. Modelul Europeană.....	4

## II. Formate. Fișiere. Metadate.

1. Terminologie.....	6
2. Formate.....	7
3. Denumirea fișierelor .....	10
3.1 Excepție: Documente din arhiva istorică.....	11
4. Tipologie documentară.....	12
5. Descrierea obiectelor digitizate. Metadate.....	13

## III. Aspecte tehnice

1. Caracteristici tehnice recomandate.....	18
2. Norme de conservare pentru scanarea colecțiilor de patrimoniu.....	21

Concluzii.....	23
----------------	----

*”Ghid de digitizare – pilonul tematic ”Biblioteci”, versiunea 01/30.10.2009 a fost realizată în cadrul Comisiei de specialitate pentru digitizare - pilonul tematic ”Biblioteci”, prin contribuția bibliotecilor din Sistemul Național de Biblioteci, deținătoare de colecții digitale și/sau patrimoniu cultural scris: B.C.U ”Eugen Todoran” Timișoara, B.C.U. ”Lucian Blaga” Cluj-Napoca, BCU ”Mihai Eminescu” Iași, BCU ”Carol I” București, Asociația Națională a Bibliotecilor Publice din România, Biblioteca Academiei Române, Biblioteca Națională a României.*

*Documentul postat pe site-ul Bibliotecii Naționale a României urmează fi supus dezbaterii, actualizării și completării.*

## **II. Aspecte teoretice**

### **1. Introducere**

Constituirea bibliotecii digitale naționale, parte componentă a bibliotecii digitale europene, are ca principal scop conservarea și protecția patrimoniului cultural național existent în bibliotecile din cadrul sistemul național de biblioteci, promovarea colecțiilor și lărgirea accesului la informație.

Prin digitizarea colecțiilor de documente și constituirea bibliotecii digitale naționale se asigură o bună promovare a valorilor naționale, o mai bună diseminare a informației și o valorificare superioară, la nivel național și internațional, a colecțiilor speciale, a documentelor rare.

Obiectivele care stau la baza acestui demers se referă la:

- transpunerea în format electronic a patrimoniului cultural scris;
- promovarea patrimoniului cultural scris la nivel european;
- protejarea valorilor de carte bibliofilă și manuscrise;
- protejarea documentelor aflate într-o stare avansată de deteriorare;
- îmbunătățirea posibilităților de acces la documente, local sau la distanță, cu impact asupra creșterii numărului de utilizatori și a categoriilor acestora;
- posibilitatea consultării simultane de către mai mulți utilizatori a aceluiași document;
- oferirea unui mod de consultare a documentelor modern, în acord cu noile tehnologii, independent de spațiul și programul de funcționare al bibliotecii (cu respectarea restricțiilor de copyright);
- îmbunătățirea calității procesului de consultare a documentelor.

## 2. Etapele desfășurării procesului de digitizare

- a) identificarea proiectelor de digitizare în curs;
- b) inventarierea documentelor digitizate în bibliotecile din țară;
- c) identificarea priorităților de digitizare ale bibliotecilor și a documentelor propuse spre digitizare;
- d) selecția documentelor/colecțiilor ce urmează a fi digitizate;
- e) metodologia de lucru;
- f) stabilirea soluției de digitizare;
- g) identificarea echipamentelor și a software-ului;
- h) crearea de conținut digital;
- i) realizarea și administrarea portalului Biblioteca Digitală a României;
- j) integrarea Bibliotecii Digitale a României în Biblioteca Digitală Europeană.

## 3. Resurse informaționale și documentare în bibliotecile din România

Bibliotecile din sistemul național de biblioteci sunt componente definitorii ale sistemului național informațional. Prin structuri, funcționalități, servicii și patrimoniu info-documentar ele trebuie să răspundă cerințelor complexe și variate de informare.

Sistemul național de biblioteci, în raport cu funcțiile și atribuțiile acestora, este structurat astfel:

- a) Biblioteca Națională a României;
- b) Biblioteca Academiei Române;
- c) biblioteci universitare;
- d) biblioteci specializate;
- e) biblioteci publice;
- f) biblioteci școlare.

Prin participarea tuturor bibliotecilor din cadrul sistemului național de biblioteci, deținătoare de documente și colecții, se va putea crea o bibliotecă digitală cu un conținut eterogen din punct de vedere al tipului de documente (cărți, manuscrise, publicații seriale), al perioadei de publicare, al adresabilității.

Biblioteca Digitală Națională se poate realiza respectând principiul partajării resurselor.

## 4. Criteriile de selecție

Criteriile de selecție pot fi și de cele mai multe ori trebuie să fie combinate pentru obținerea unui soluții optime.

Se va avea în vedere respectarea regimului juridic referitor la documentele propuse pentru scanare (Legea nr. 8/1996 privind dreptul de autor și drepturile conexe; Legea nr. 334/2002 privind bibliotecile; Legea nr. 182/2000 privind protejarea patrimoniului cultural; Legea nr. 135/2007 privind arhivarea documentelor în formă electronică; Legea nr. 186/2003 privind susținerea și promovarea culturii scrise).

Criteriile de selecție sunt:

- valoarea documentară;
- reprezentativitatea pentru un anumit domeniu, o anumită perioadă de timp sau o anumită regiune geografică, atât la nivel național, cât și la nivel internațional;
- gradul de interes și adresabilitate;
- prezervarea documentelor originale;
- regimul juridic (dreptul de autor și dreptul de difuzare);
- criteriul financiar;
- soluția de scanare și tehnologia implicată.

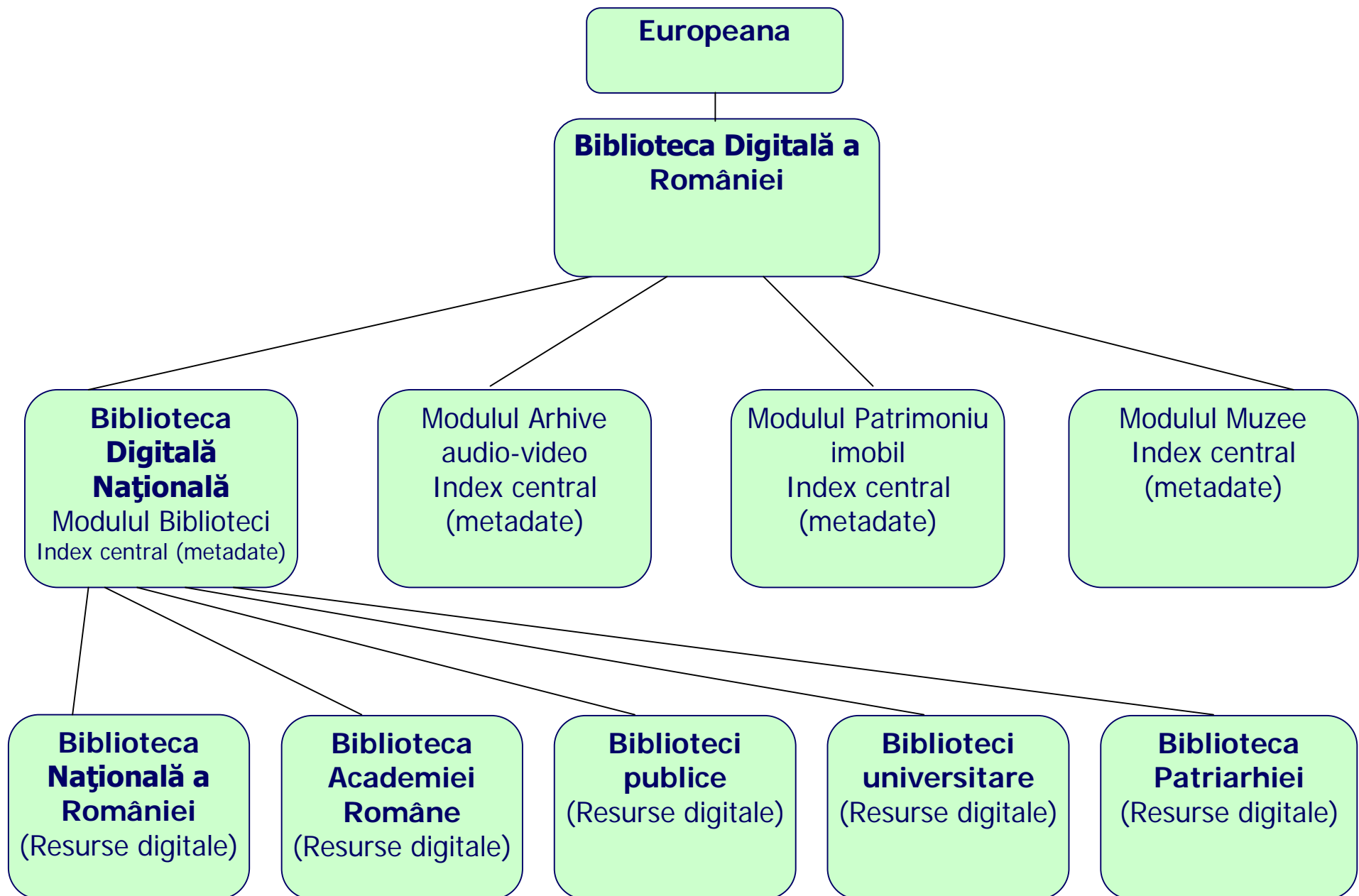
La nivelul bibliotecilor județene se manifestă un interes deosebit în digitizarea publicațiilor cu caracter local, care reflectă istoria, cultura și civilizația zonei respective.

La nivelul bibliotecilor universitare accentul se pune pe publicațiile academice (se are în vedere curricula universitară), teze de doctorat.

## **5. Modelul Europeana**

Modelul Europeana este modelul descentralizat promovat de toate bazele de date europene. Interrelaționarea sistemului asigură posibilitatea portalului de a culege metadatele documentelor digitale din bazele de date ale deținătorilor și de a le introduce în indexul general, punându-le astfel la dispoziția utilizatorilor. Acesta este modelul utilizat la nivel european, în cadrul TEL, Europeana, Manuscriptorium.

Metadatele se vor găsi în cadrul portalului, iar documentele digitale vor rămâne la deținător.



## II. Formate. Fișiere. Metadate.

### 1. Terminologie

**Digitizare** reprezintă procedeul prin care informația este capturată în format digital (imagine, document text, fișier audio, etc.) cu ajutorul unui echipament tehnic digital (cameră digitală, scanner, etc.). Când vorbim despre **digitizarea documentelor**, de cele mai multe ori ne referim la imaginea paginii capturată de un astfel de echipament - pur și simplu o poză a documentului – sau o versiune full-text, în care documentul este stocat folosind caractere text/scrise. Forma neprocesată a documentului (“plain-text”), reprezintă varianta integrală a documentului, folosind caractere ASCII sau Unicode, pentru acestea existând posibilitatea efectuării unei căutări în text (cuvinte sau fraze), însă se pierde structura și aspectul original al documentului.

O versiune „codificată” (encoded) a documentului va include informații suplimentare sau „markup” de diferite feluri, pentru a exprima structura documentului, formatarea sau alte informații pe care creatorul a dorit să le evidențieze și să-i ofere acestuia funcții speciale. Acest tip de codificare „markup” este folosit frecvent în asociere cu limbajele SGML sau XML și acest gen de informație este aplicată documentelor cu text integral.

**Imagine digitală** reprezintă setul organizat de elemente fixe, rectangulare, bidimensionale denumite pixeli, care definesc o imagine pe ecranul unui computer sau care sunt transformată într-un alt format, precum pagina tipărită.

**Documentul / Fișierul imagine digitală master** este rezultatul materialului analog transformat, respectiv convertit printr-un proces de reformatare digitală (prin scanare, fotografiere cu aparat digital, etc.). Acesta ar trebui să reprezinte cât mai fidel posibil informația vizuală din documentul original. Un document digital master, poate servi ca înlocuitor pentru cel original sau poate fi folosit ca siguranță în cazul pierderii sau deteriorării acestuia. Pentru „documentele născute digital”, adică cele care nu au fost supuse procesului de reformatare digitală, documentul imagine digitală master reprezintă documentul original în sine. Deoarece procesul capturării unei imagini presupune o muncă intensă, durata de viață a unui document imagine master ar trebui să fie de cel puțin 50 de ani. Prin urmare, cei care organizează colecția digitală trebuie să aibă în vedere o varietate largă de utilizări ulterioare și să stabilească de la început o calitate înaltă/foarte bună a imaginii.

**Fișiere / Documente derivate** sunt create din documentul imagine digitală master cu scopul de a fi editate ulterior, de a fi mărite, transformate în diferite formate, dar și pentru a fi reprezentate și transmise prin intermediul rețelelor (ca de exemplu imagini create pentru accesare și imagini timbru)

**OCR** (Recunoaștere Optică a Caracterelor), este procesul de recunoaștere cu acuratețe a literelor prin care un program citește imaginea scanată a unui text și o transpune într-un fișier text.

**Metadate:** orice date referitoare la datele specifice creării unui obiect digital și care se referă la descrierea informațiilor caracteristice (informații legate de publicarea acestuia, formatul, drepturi, etc.).

**SGML** (Standard Generalized Markup Language) standard internațional (ISO 8879/1986) care este folosit ca metalimbaj pentru definirea limbajelor de codificare (markup).

**XML** (Extensible Markup Language) extensie a SGML , apărut ca recomandare a W3C în 1998.

**TEI** (Text Encoding Initiative) un consorțiu internațional care publică indicații referitoare la schimbul și codificarea textelor electronice, conform specificațiilor limbajelor SGML și XML pentru prezervarea textelor literare și lingvistice folosite în mediile culturale și academice.

**EAD** (Encoded Archival Description), un limbaj adaptat cerințelor standardelor SGML și XML folosit pentru prezervarea informațiilor unei arhive, dar și de către biblioteci.

**METS** (Metadata Encoding and Transmission Standard) un standard de codificare care îndeplinește cerințele XML pentru metadatele obiectelor deținute de o bibliotecă digitală.

## 2. Formate

Procesul de digitizare trebuie să țină cont de o serie de factori, precum natura materialului, scopul exemplarului digital, relevanța în cadrul colecției, suportul tehnic necesar prezervării dar și bugetul estimat și finanțările posibile. Este important de știut care sunt limitele formatelor disponibile, pentru a vedea dacă acestea corespund scopului colecției, dar și pentru a-i asigura accesul în viitor. Formatele digitale pentru documentul text diferă foarte mult, deoarece acestea pot fi convertite cu ușurință în alte formate dar și datorită capacităților multiple de expunere.

Sistemele care dețin drepturi de proprietate pot fi potrivite pentru tipărirea unui document, prin instrumentele de procesare text sau formatele descrierilor de pagină (de exemplu, Microsoft Word, PDF) și permit publicarea acestora pe web, prin intermediul funcției "Salvare ca HTML". În cazul în care este necesară mutarea/migrarea datelor pe o altă platformă software, există riscul de a pierde informațiile de formatare. Pentru că astfel de sisteme depind de existența unor programe cu drept de proprietate - ai căror termeni de licențiere și existență nu pot fi definiți pe termen lung - acestea nu sunt recomandate pentru scopuri de arhivare sau pentru definirea unei existențe durabile în timp. De asemenea este indicat să fie identificate formatele acceptate pentru transfer, stocare și distribuție către utilizator, deoarece acestea pot diferi, să nu fie utilizate metode de criptare sau comprimare pentru fișierele care vor fi păstrate în arhivă și trebuie micșorat numărul tipurilor de formate, pentru ca documentele digitizate să fie gestionate cât mai ușor.

Principalele formate de fișier utilizate sunt:

- **Imagine:** TIFF, GIF, JPG, JPG-2000, PNG, PDF
- **Text:** HTML, XML, PDF, UTF-8, ASCII
- **Audio, Video:** WAVE, MPEG, AVI, WMA, OGG
- **Arhivare:** RAR, ZIP

<b>Extensie</b>	<b>Denumire</b>	<b>Descriere</b>	<b>Puncte tari/Puncte slabe</b>
<b>.tiff, .tif</b>	TIFF (Tagged Image File Format)	Document la rezoluție maxim recomandată; formatul a fost creat la început pentru publicarea electronică 1 – 64 bit adâncime. Folosit în general pentru imagini de înaltă calitate și pentru stocare arhivată.	Calitate înaltă; Documente foarte mari; Majoritatea programelor care recunosc acest format ating un maxim de 24-bit culoare. Distribuția pe web este limitată de dimensiunea fișierului/documentului.
<b>.gif</b>	GIF (Graphics Interchange Format)	Acest format de 8-bit suportă compresie fără pierdere de date, interrelaționare și transparență.	Compresie fără pierdere de date. Format de livrare pentru web folosit cu regularitate.
<b>.jpg, .jpeg</b>	JPEG (Joint Photographic Experts Group)	Imagine cu o rezoluție comprimată/restrânsă, 8-24 bit. Acesta variază în funcție de calitatea și dimensiunea fișierului/documentului.	Compresie fără pierdere de date. Cel mai folosit format pentru expunere pe monitor. Flexibil.
<b>.png</b>	PNG (Portable Network Graphics)	Compresie fără pierdere de date 24 bit. Poate înlocui liber formatul GIF și poate fi folosit în locul formatului .tiff Suportă interrelaționare, transparență, gamma.	Standard liber , poate fi recunoscut doar de unele navigatoare de internet; combină avantajele oferite de formatul TIFF și de formatul GIF furnizând aceeași calitate pentru un document dar la dimensiuni mai mici.
<b>.pdf</b>	PDF (Portable Document Format)	4-64 bit adâncime. Dimensiune maxim recomandată. Folosit în principal pentru livrarea/ expunerea documentelor imagine.	Este necesară aplicația Adobe pentru vizualizare sau un program adițional care extinde funcționalitatea aplicației.
<b>.wav</b>	Wave . Windows media.	Un format necomprimat care poate fi folosit de obicei fără a fi nevoie de un suport adițional de 8, 16, & 32 bit. Înaltă calitate. Fișiere de dimensiuni mari, pot fi folosite și de calculatoare Mac precum și de cele obișnuite PC.	



<b>HTML</b>	Hypertext Markup Language	HTML este o aplicație SGML (SGML: Standard Generalized Markup Language, standard internațional creat în 1986 (ISO 8879) este un meta-limbaj care definește sistemul de adnotare al unui text folosit în mediul electronic) a fost dezvoltat pentru a ușura modul în care o pagină web este redată de navigatorul folosit.	HTML nu oferă informații despre structura textului; câteva caracteristici depind de navigatorul web folosit
<b>XML</b>	Extensible Markup Language	Standard liber creat de W3C care conține un set de reguli pentru codificarea documentelor electronice. Asigură longevitate datelor și flexibilitate la transferul pe o altă platformă software.	
<b>.mpg</b>	Moving Picture Experts Group	Format digital standardizat de către Moving Picture Experts Group (MPEG); încorporează formatele audio și video MPEG-1 sau MPEG-2	Folosit adesea pentru crearea înregistrărilor care sunt distribuite/care pot fi accesate pe internet.
<b>.avi</b>	Audio/Video Interleave	Audio-video standard creat de Microsoft și folosit cu precădere de către Windows' Media Player. Poate fi înlocuit de formatul MPEG	Calitate bună, dimensiune scăzută a fișierului.
<b>.ogg</b>	Format audio video de tip container	Formatul de fișiere Ogg creat de Fundația Xiph.org, a fost creat drept un cadru de lucru pentru o inițiativă mai largă cu ținta dezvoltării unui set de componente pentru codificarea și decodificarea conținutului multimedia.	Atât codificarea cât și decodificarea sunt disponibile gratuit.

### 3. Denumirea fișierelor

Documentele rezultate în urma procesului de digitizare trebuie denumite astfel încât alcătuirea colecțiilor, stabilirea relațiilor logice să nu fie îngreunată. Există două posibilități:

1. folosirea unei scheme alfa-numerice care să reflecte numerele folosite deja în catalogarea documentului
2. utilizarea unor termeni reprezentativi pentru documentul respectiv.

Se pot stabili convenții pentru nume, tabele cu semne sau reguli (de ex. reguli referitoare la punctuație și la majuscule) uniformizarea acestora fiind un aspect foarte important.

**Biblioteca Națională a României** propune următoarea structură a numelui de fișier:

***Cota/nr. inv\_TD\_secvența.ext***

unde:

**- cota sau nr.inventar**

- reprezintă cota obiectului digitizat (sau numărul de inventar);
- câmp alfanumeric, număr variabil de caractere;

**- TD**

- tipologia documentară;
- identifică tipul documentului (vezi mai jos codurile);
- câmp alfanumeric din două poziții:
  - o o majusculă pentru clasa de documente;
  - o un număr pentru tipul documentului în cadrul clasei (este 0 dacă nu există subclasificare).

**- secvența**

- poziția obiectului digitizat în cadrul unei structuri (de exemplu identifică numărul paginii într-o carte, sau al unei melodii pe un CD);
- câmp numeric din minim 2 cifre;
- este 00 dacă obiectul digitizat nu este structurat;
- se completează obligatoriu zerourile (de ex. 07, 095, 0087 etc.).

**- .ext**

- reprezintă extensia (tipul fișierului);
- ex.: .tiff, .jpg etc.

### 3.1 Excepție: Documente din arhiva istorică

#### **Cota(Fondul)\_Dosar\_docnnn\_secvența.ext**

Unde:

**- cota sau fondul**

- reprezintă cota obiectului digitizat (fondul din care provine);
- câmp alfanumeric, număr variabil de caractere;

**- dosar**

- reprezintă numărul dosarului în cadrul fondului;
- câmp alfanumeric, număr variabil de caractere;
- se vor completa zerourile

**- docnnn**

- reprezintă reprezintă numărul documentului din cadrul dosarului;
- câmp alfanumeric;
- nnn este un număr
- exemplu: doc005, doc157 etc.;

**- secvența**

- reprezintă numărul paginii din cadrul documentului (nu al filei);
- câmp numeric, număr variabil de caractere;

**- .ext**

- reprezintă extensia (tipul fișierului);
- ex.: .tiff, .jpg etc.

**NOTĂ:** grupul **cota(fond)\_dosar\_docnnn** constituie **cheie de regăsire** a obiectelor ce aparțin aceluiași dosar.

Exemple:

Biblioteca Națională a României

BNR\_XI-92\_M0\_001 (reprezintă coperta de la un manuscris)

BNR\_CR-XVII-I-63\_C1\_101.TIF (reprezintă a 101 înregistrare dintr-o carte veche digitizată – nu neapărat pagina 101)

BNR\_AI-27-1\_M3\_00.jpg (reprezintă o scrisoare din arhiva istorică)

Biblioteca Academiei Române

- periodice **RAL\_Multimedia\_Fond\_Titlu\_AN de vârsta sau VOLUM\_\_AN calendaristic \_Nr.\_ P.tif** respectiv extensia **.jpg**
- cărți/manuscrite in volum **RAL\_Multimedia\_Fond\_COTA\_P** sau **F.tif**, respectiv **.jpeg**
- fotografii, documente grafice, hărți **RAL\_Multimedia\_Fond\_Cota.tif** respectiv **.jpeg**

## 4. Tipologie documentară

Tipologia documentară propusă, este cea elaborată de Biblioteca Națională și aprobată de Comisia de specialitate pentru Digitizare - pilonul biblioteci <http://www.bibnat.ro/dyn-doc/Studiu%20Fezabilitate/Tipologie-documentara.pdf> .

Literele identifică grupa de documente, iar cifrele indică tipul documentului.

### **C. Carte tipărită**

1. *carte veche*
2. *incunabul*
3. *carte rară/bibliofilă*
4. *carte curentă*

### **P. Publicații periodice**

1. *periodice vechi*
2. *periodice rare/bibliofile*
3. *periodice curente*

### **M. Manuscrise**

1. *carte manuscris a unei singure opere*
2. *carte manuscrisă-miscelaneu*
3. *manuscrise de autor*
4. *corespondență*

### **F. Microformate**

1. *microfișă*
2. *microfilm*

### **H. Documente cartografice**

1. *hărți (plane, în relief sau globulare);*
2. *atlase*
3. *planuri tipărite*

### **A. Documente audio-video**

1. *muzicale tipărite*
2. *audio muzică*
3. *audio voce*
4. *video*

### **G. Documente grafice**

1. *studiu*
2. *desen tehnic*
3. *program*
4. *anunț*
5. *proclamație*
6. *planșă*

### **V. Arte vizuale:**

1. *gravură*
2. *desen*
3. *ex libris*
4. *reproducere de artă*
5. *fotografie*
6. *afiș*
7. *pictură*
8. *sculptură*

#### T. Alte tipuri de documente:

1. teze de doctorat și rezumate ale acestora;
2. broșuri;
3. cursuri universitare tipărite sau multigrafiate;
4. manuale școlare;
5. documente numismatice și filatelice;
6. standarde și norme tehnice și de funcționare;
7. orice alte documente tipărite sau multiplicare prin proceduri grafice sau fizico-chimice de tipul: litografieră, fotografieră, fono și videografieră.

### 5. Descrierea obiectelor digitizate. Metadate

Imaginile digitale sunt descrise cu ajutorul metadatelor, acestea reprezentând date ce permit gestionarea și utilizarea altor date. Folosirea acestora este strâns legată de reguli de descriere și catalogare specifice documentelor tradiționale, înregistrărilor de arhivă și artefactelor. Diferența constă în prezența metadatelor adiționale necesare pentru a facilita navigarea și managementul fișierelor de date.

#### Tipuri de metadate:

- **Metadate DESCRIPTIVE** descriu conținutul resursei digitale astfel încât aceasta să poată fi regăsită/localizată când este căutată și totodată să poată fi diferențiată de alte resurse similare.
- **Metadate de PREZERVARE**, anume
  - **Metadate TEHNICE:** formatul, data digitizării, caracteristici tehnice ale obiectului digital, etc.
  - **Metadate ADMINISTRATIVE:** aspecte ce țin de gestionarea documentului digital, livrare, distribuție, informații legate de drepturi asupra obiectului digital.

Dacă aceste metadate nu se regăsesc, accesul la obiectul digital poate fi îngreunat, iar prezervarea acestuia nu se poate realiza. Metadatele care oferă informații despre contextul tehnic al resursei (formatul fișierului, posibilitățile de acces, etc.) și a altor informații (legate de drepturile de autor, drepturile de reproducere) sunt foarte importante pentru gestionarea continuă a obiectului digital. Ele oferă informații utile în procesul de prezervare, asemenea măsurilor de securitate ce trebuie luate pentru un document fizic. Fără aceste metadate nu se poate ști ce reprezintă resursa, cum poate fi folosită, care a fost procesul de realizare.

Metadatele pot fi stabilite pentru o întreagă colecție sau pentru fiecare exemplar în parte. Metadatele de prezervare sunt importante datorită faptului că prin intermediul lor un obiect digital poate să ofere informații despre el însuși o perioadă îndelungată de timp, facilitând astfel prezervarea și accesul, deținătorului, custodelui, tehnologiei, și chiar comunităților utilizatorilor care se află într-o permanentă schimbare.

**Metadate DESCRIPTIVE**

<b>ELEMENT</b>	<b>STATUT</b>
<b>Titlu (Denumirea înscrisă pe document)</b>	<i>Obligatoriu</i>
<b>Autor/ Alta mențiune de responsabilitate</b>	<i>Obligatoriu</i>
<b>Editor</b>	<i>Obligatoriu</i> Editura, Loc – complet SAU parțial, acolo unde există
<b>Anul apariției</b>	<i>Obligatoriu</i>
<b>Număr de pagini</b>	<i>Obligatoriu (formatul)</i>
<b>Limba textului</b>	<i>Recomandat</i>
<b>Descrierea documentului</b>	<i>Recomandat</i> Pentru cărțile vechi acest câmp va fi completat cu descrierea fizică a documentului original, de ex: dacă este complet sau nu, dacă documentul conține ilustrații, informații despre calitatea hârtiei, dacă documentul a fost restaurat, dacă poate fi consultat fizic.
<b>Tipul documentului</b>	<i>Obligatoriu</i>
<b>Formatul documentului original</b>	<i>Obligatoriu</i>
<b>Deținător</b>	<i>Obligatoriu</i> Țara/Instituția/Filiala deținătoare
<b>Proveniența</b>	<i>Recomandat</i> Colecția din care face parte/Fondul etc.
<b>Subiectul/Tematica</b>	<i>Recomandat</i>
<b>Alte legături/Relația</b>	<i>Recomandat</i>
<b>Data creării</b>	<i>Recomandat</i>

**Metadata de PREZERVARE (TEHNICE și ADMINISTRATIVE)**

<b>ELEMENT</b>	<b>STATUT</b>
<b>Identificator</b>	<i>Obligatoriu</i> Folosit pentru stabilirea unui cod de identificare unic documentelor digitale care au fost create cu scopul de a fi expuse pe web - prin intermediul unui program de gestionare a obiectelor digitale, stocate pe un server sau orice altă modalitate
<b>Formatul documentului digital</b>	<i>Recomandat</i> (.jpg, .pdf, etc.)
<b>Mărimea /Rezoluția fișierului</b>	<i>Recomandat</i> (dacă e comprimat, etc.)
<b>Aplicația folosită</b>	<i>Recomandat</i> Numele, Versiunea, Creatorul acesteia
<b>Software folosit</b>	<i>Recomandat</i> (Denumire, caracteristici)
<b>Hardware folosit</b>	<i>Recomandat</i> (Denumire, caracteristici)
<b>Tehnica de reproducere</b>	<i>Recomandat</i> (Scanner, Cameră digitală - caracteristici,etc.)
<b>Data efectuării reproducerii/ creării obiectului digital</b>	<i>Recomandat</i>
<b>Instituția care a efectuat reproducerea</b>	<i>Recomandat</i> In-house, externalizare; evidența deținătorilor obiectului digital, pentru eventualele modificări
<b>Metoda de siguranță</b>	<i>Recomandat</i> Procedeul de codificare: watermark, semnătură digitală, etc.
<b>Autenticitatea documentului digital</b>	<i>Recomandat</i> Metadatale de preservare trebuie să includă informație suficientă pentru validarea unui obiect digital, pentru confirmarea scopului digitizării acestuia dar și pentru confirmarea faptului că acesta nu a fost alterat/modificat fie intenționat sau neintenționat
<b>Relație / Legătură cu</b>	<i>Recomandat</i> Structurală – pentru stabilirea relațiilor dintre obiecte; Derivare – replică sau transformare a unui alt obiect digital
<b>Locația / Directorul</b>	<i>Obligatoriu</i> (server, hard extern, etc.)
<b>Drepturile de management</b>	<i>Recomandat</i> În context depozit digital, se referă la acțiunile efectuate în scopul prezervării (numele persoanei autorizate, poziția ocupată)
<b>Drept de acces</b>	<i>Obligatoriu</i> Restricții referitoare la drepturile de utilizare, reproducere, distribuire; modalități folosite

	pentru restricționare, de ex. : vizualizare în interiorul instituției, accesibil doar utilizatorilor înregistrați, etc. sau document cu acces liber (pentru documente care aparțin domeniului public), durata dreptului de acces, etc.
<b>Drept de autor/de proprietate intelectuală</b>	<i>Recomandat</i> Numele autorului, Date de contact; Numele instituției, Date de contact
<b>Declarația de copyright</b>	<i>Recomandat</i>

**Pentru a satisface nevoile prototipului Europeana\*, s-au adăugat următoarele elemente:**

<b>dc:relation</b> (relația cu altă resursă)	<i>Obligatoriu</i> <b>europena:isShownBy</b> – referința clară URL către pagina web a instituției care deține obiectul digital disponibil la cea mai bună rezoluție FIE <b>europena: isShownAt</b> - referința URL către descrierea completă a obiectului digital pe pagina web a instituției deținătoare
<b>europena:userTag</b>	Etichete create de către utilizatori înregistrați prin interfața Europeana. Pentru această informație furnizorul de conținut <b>nu are nicio responsabilitate.</b>
<b>europena:unstored</b>	<i>Recomandat</i> Locul unde se află stocate informațiile referitoare la obiectul digital, fără de care acestuia nu i se pot atribui elemente specifice Europeana. Include informații necesare indexării. Vor fi incluse toate elementele care nu se regăsesc în câmpurile deja stabilite. Nu va fi afișat ca rezultat la o căutare completă.
<b>europena:object</b>	<i>Adăugat în mod obligatoriu de către Europeana</i> Referința URL la obiectul digital cu cea mai bună rezoluție disponibilă pe pagina web a furnizorului de conținut pentru a genera o imagine timbru. În cazul în care nu există deja o imagine timbru Europeana va crea câte una pentru fiecare obiect în parte iar dacă sunt mai multe imagini incluse, va fi aleasă prima. În cazul unui document în format PDF, se va alege drept reprezentativă pagina de titlu.
<b>europena:language</b>	<i>Recomandat</i> Limba atribuită resursei cu referință la furnizorul de conținut. Este obținut automat din fișierul de autoritate al furnizorilor din Europeana. Pentru această informație furnizorul de conținut <b>nu are nicio responsabilitate.</b>

\* Specification for the Europeana Semantic Elements V 3.1, 25/02/2009+



<b>europena:provider</b>	<i>Adăugat în mod obligatoriu de către Europeana pentru uz intern</i> Numele organizației care deține obiectul digital (digitizat sau născut digital). Este controlat de fișierul de autoritate al furnizorilor din Europeana.
<b>europena:type</b>	<i>Obligatoriu</i> Tipul resursei, ținând cont de cele patru tipuri suportate de Europeana <b>TEXT, IMAGINE, AUDIO și VIDEO.</b> Această informație este folosită pentru descrierea obiectului în interfața Europeana și în căutări. Furnizorii de conținut trebuie să creeze o tipologie locală a documentelor deținute.
<b>europena:uri</b>	<i>Adăugat în mod obligatoriu de către Europeana</i> Referința URI (nu URL) către resursă în contextul Europeana. Reprezintă identificatorul înregistrării și este creat pe baza identificatorilor unici furnizați în metadatele însoțitoare. Pentru situațiile în care sunt trimise două sau mai multe obiecte care au același identificator, doar primul va fi păstrat de Europeana, celelalte vor fi eliminate.
<b>europena:year</b>	<i>Recomandat</i> Anul sau perioada asociată cu documentul original în format tradițional sau cu obiectul născut digital. Va fi scris în forma completă cu patru cifre (ex. 1523) și va reieși din valorile trimise în metadatele originale. Va fi folosit ca informație în navigatorul cronologic. Portalul Europeana nu suportă informații cronologice de dinaintea anului 0 , spre ex. 250 î. Hr., 600 î. e.n sau 10,000 î.e.n. însă se dorește a fi introduse în viitor.
<b>europena:hasObject</b>	<i>Adăugat în mod obligatoriu de către Europeana pentru uz intern</i> Indică disponibilitatea imaginilor timbru aferente obiectelor digitale.
<b>europena:country</b>	<i>Adăugat în mod obligatoriu de către Europeana</i> Țara căreia aparține furnizorul de conținut sau va fi aleasă varianta “Europa” pentru proiecte derulate la nivelul Europei. Numele va fi extras automat din informațiile deținute de fișierul de autoritate al furnizorilor de conținut din Europeana; această informație nu trebuie trimisă separat de conținutul digital.

### III. Aspecte tehnice

#### 1. Caracteristici tehnice recomandate

##### Terminologie

**Calitatea imaginii** capturate poate fi definită prin suma dintre rezoluția scanării, profunzimea biților imaginii scanate, procesul de compresie aplicat, echipamentul de scanare și tehnica folosită, și îndemânarea operatorului.

**Rezoluția** este determinată de numărul de pixeli care alcătuiesc o imagine și este exprimată în dpi „dots pe inch” **dpi** sau „points pe inch” **ppi** (număr de puncte pe inch). Diferența între ele este dată de faptul că în **dpi** se măsoară **rezoluția documentului la imprimare** iar **ppi**, **rezoluția documentului pe ecranul** calculatorului. Creșterea numărului de pixeli pentru capturarea unei imagini nu va duce doar la o rezoluție mai mare și o mai mare capacitate de a delimita detalii fine, însă pe măsură ce rezoluția crește, aceasta nu va determina o calitate mai bună ci doar un fișier mai mare. Important este a determina punctul în care rezoluția este considerată suficientă pentru a captura toate detaliile semnificative în documentul original/sursă.

**Profunzimea (Bit depth).** Un bit (pl. biți) reprezintă cea mai mică unitate de date (**0** sau **1**, alb sau negru). 8 biți reprezintă un **byte**. Profunzimea biților reprezintă numărul de biți folosiți pentru a defini un pixel. Cu cât aceasta este mai mare cu atât este mai variat numărul de tonuri și nuanțe.

**Comprimarea** este folosită pentru a reduce mărimea fișierului/documentului digital pentru procesare, stocare sau transmitere. Calitatea unei imagini poate fi afectată atât de către tehnica folosită cât și nivelul de compresie aplicat acesteia. Tehnica de comprimare poate fi „loss less” - ceea ce înseamnă că o imagine decomprimată va rămâne identică cu cea inițială, deoarece nu există informații pierdute când fișierul este redus – sau „lossy”, atunci când are loc pierdere de informație în procesul de comprimare.

##### Tehnici de scanare:

- **scanare bitonal** - *bitonal scanning* – se utilizează un bit per pixel pentru a reprezenta culorile negru sau alb
- **scanare cu nuanțe de gri** - *greyscale scanning* - mai mulți biți per pixel pentru reprezentarea mai multor nuanțe de gri; nivelul preferat de scară gri este de 8 biți per pixel, iar la acest nivel imaginea afișată poate selecta 256 de la diferite niveluri de gri.
- **scanare color** - *colour scanning* - se utilizează mai mulți biți per pixel pentru a reprezenta culoare; 24 biți per pixel reprezintă adevăratul nivel de culoare și face posibilă o selecție din 16,7 milioane de culori. (24-bit RGB)

## **Tipuri de fișiere**

În funcție de scopul creării fișierelor digitale, s-au conturat trei tipuri de fișiere:

- **fișier master:** copie identică cu obiectul digitizat, de înaltă rezoluție, necomprimat, utilizat pentru prezervare.
- **fișier pentru acces:** fișier derivat din mașter, care se aproprie de original, utilizat pentru publicare și care asigură o viteză rezonabilă la livrare.
- **fișier pictogramă (thumbnail):** fișier derivat din master, de dimensiuni și rezoluție mică, utilizat pentru previzualizare.

## Caracteristici tehnice recomandate

Tipuri de fișier	Caracteristici tehnice	Tipuri de documente								
		Manuscrise Carte rară	Text tipărit		Hărți	Fotografii	Artă grafică	Audio		Video
			imagine	OCR				muzică	voce	
<b>Master</b>	profundzime	Gri 8bit/pixel Color 24bit/pixel	Gri 8bit/pixel Color 24bit/p		Gri 8bit/pixel Color 24bit/p	Gri 8bit/pixel Color 24bit/p	Gri 8bit/pixel Color 24bit/p	Mono/Stereo 16-24 bit	Mono/Stereo 8-16 bit	Norma: PAL
	format fișier	Tiff	Tiff		Tiff	Tiff	Tiff	Wave/Ogg	Wave/Ogg	Mpeg2 /DVD
	compresie	nu	nu		nu	nu	nu	nu	nu	nu
	rezoluție	400/600 dpi sau 4000-6000 pixel pe dimensiunea mai mare	300/400 dpi sau 3000-5000 pixel pe dimensiune. mai mare		400 dpi sau 4000-6000 pixel pe dimensiune. mai mare	300 dpi sau 3000-6000 pixel pe dimensiune. mai mare	400 dpi sau 3000-6000 pixel pe dimensiune. mai mare	128-200 kbps /44,1-48 kHz	192 kbps/ 22 kHz	Size: 720x480 sau mai mare la 48 kHz și 24 biți calitate audio
<b>Acces</b>	profundzime	Gri 8bit/pixel Color 24bit/pixel	Gri 8bit/pixel Color 24bit/p	Gri 8bit/pixel	Gri 8bit/pixel Color 24bit/p	Color 24bit/p	Color 24bit/p	Mono/Stereo	Mono/Stereo	Norma: PAL
	format fișier	Jpeg	Jpeg	Pdf	Jpeg	Jpeg	Jpeg	MP3	MP3	Mpeg1/avi
	compresie	da	da	da	da	da	da	da	da	da
	rezoluție	120-300 dpi sau de la 500x400pixel la 1024x768pixel	96-120 dpi sau de la 500x400pixel la 1024x768 pixel	400 dpi	200-300 dpi	250 dpi	250 dpi	128 kbps /22-44,1 kHz	28-128kbps/ 11 kHz	Size: de la 160x120 la 640x480 la 22 kHz
<b>Pictogramă</b>	profundzime	8bit/pixel	8bit/pixel		8bit/pixel	8bit/pixel	8bit/pixel			
	format fișier	Gif	Gif		Gif	Gif	Gif			
	compresie	da	da		da	da	da			
	rezoluție	72 dpi sau de la 150x100 la 200x200	72 dpi sau de la 150x100 la 200x200		72 dpi sau de la 150x100 la 200x200	72 dpi sau de la 150x100 la 200x200	72 dpi sau de la 150x100 la 200x200			

## **2. Norme de conservare pentru scanarea colecțiilor de patrimoniu**

### **Aparatura de scanare**

Scanarea colecțiilor de patrimoniu (manuscrise, cărți vechi, lucrări de grafică, fotografii, materiale de arhivă, hărți etc.) este permisă numai cu aparatură profesională, de înaltă performanță tehnică, special proiectată pentru asigurarea protecției maxime a documentelor față de factorii fizico – chimici și mecanici, pe toată durata fluxului de lucru .

Trebuie obligatoriu îndeplinite următoarele cerințe esențiale:

- camera digitală să aibă o rezoluție foarte bună, astfel încât calitatea imaginilor obținute să corespundă tuturor exigențelor ulterioare, deoarece normele de conservare a colecțiilor de patrimoniu nu permit scanări repetate pe același document original;
- captarea imaginii digitale de pe document să se facă pe verticală, de sus în jos, de la o distanță de cel puțin 70 cm, fără atingerea suprafeței documentului;
- sistemul de iluminare să asigure protejarea totală a documentelor față de radiațiile ultraviolete și infraroșii;
- suportul pe care se așează documentele în timpul scanării trebuie să prevină orice fel de deteriorare fizico-mecanică a cotoarelor și a suprafețelor legăturilor. De exemplu - pentru scanarea cărților masa – suport va fi compusă din două module reglabile, tip balanță, care se vor adapta ușor la grosimile variabile ale cotoarelor și la modificările de greutate ale corpului de carte pe parcursul foietării pentru digitalizare; se vor preveni , astfel, tensionările, plesniturile sau ruperile cotoarelor interioare și exterioare ale volumelor în timpul mânăuirii pe scanner.

### **Pregătirea documentelor pentru scanare**

În această etapă se vor efectua următoarele proceduri:

- evaluarea stării de conservare a documentelor selectate pentru scanare, de către specialistul în conservarea colecțiilor de patrimoniu;
- conservatorul va decide care dintre volumele sau documentele selectate pot fi scanate, având în vedere evitarea riscurilor privind integritatea lor și deteriorarea stării lor de conservare inițiale, în timpul procesului de scanare;
- scanarea documentelor din categoria patrimoniu național, a celor propuse pentru a fi clasate, a rarităților, a unicitelor și a celor foarte fragile, se va face numai cu asistența conservatorului specializat;
- la scoaterea documentelor din depozite și trecerea lor în spațiile destinate procesului de scanare se vor avea în atenție, obligatoriu, diferențele de temperatură și de umiditate relativă între aceste incinte. Dacă există diferențe de temperatură mai mari de 2 – 3 grade Celsius și diferențe de umiditate relativă mai mari de 4 – 5 % , trebuie luate măsuri de acclimatizare a pieselor, pentru a se evita apariția condensării vaporilor de apă pe materialele componente ale acestora;
- se vor măsura zilnic, cu ajutorul unui termohigrometru digital, valorile temperaturilor și ale umidităților relative în depozite și spațiile de lucru;

- dacă se constată diferențe riscante, fiecare volum va fi învelit în material neutru, cu rol de tampon la umiditate ( de ex. finet alb de bumbac) și numai astfel va fi dus în sala de lucru, iar aici va fi lăsat, pentru aclimatizare, 45 – 60 minute;
- transportarea documentelor din depozite în spațiile de lucru și retur se va face în coșuri de dimensiuni medii, din material plastic bine finisat, care să poată fi mânuite și de o singură persoană;
- înainte de începerea scanării, toate volumele trebuie desprăfuite pe toate suprafețele exterioare (coperte, cotor, tranșe), iar în interior filă cu filă, în zona cotorului, fără a se atinge oglinda textului sau ornamentațiile; această operațiune se va face după recomandările conservatorului, în funcție de particularitățile fiecărui obiect; se vor folosi pensule de diverse dimensiuni, din păr animal, cârpe de bumbac și va fi amenajat un spațiu optim pentru această operațiune;
- personalul care desprăfuieste trebuie să fie echipat cu halat cu mâneci lungi, cu mască de reținere a pulberilor ( așezată peste nas și gură), cu ochelari de protecție la praf și cu mănuși chirurgicale din latex natural, de unică folosință, care se vor trage peste mânele halatului.

### **Mănuirea documentelor pe scanner**

- mănuirea și foilețarea documentelor, pe toată durata procesului de scanare, se va face cu foarte multă atenție și îndemânare, pentru a nu produce nicio deteriorare;
- mănuirea și foilețarea documentelor pe scanner se va face cu mască de protecție a căilor respiratorii (obligatorie la manuscrisele pe pergament!) și cu mănuși foarte fine, care să se muleze strâns pe mâini, pentru a asigura atât protecția obiectului foilețat cât și a mănuitorului, dar și o dexteritate foarte bună din partea acestuia; mănușile pot fi din latex natural (chirurgicale) sau din bumbac;
- persoanele care sunt alergice nu trebuie să facă foilețare pe scanner;
- în timpul foilețării nu trebuie atinse: oglinda textului, miniaturile, adnotările marginale etc.; pot fi atinse doar marginile albe ale filelor;
- scanarea unui volum se va face cu continuitate, până la terminarea lui;
- după finalizarea scanării volumul/documentul va fi dus în depozit, aplicând, atunci când este cazul, procedura de aclimatizare.

## **Concluzii referitoare la punerea în practică a procesului de digitizare pentru pilonul tematic ”Biblioteci”**

- procesul de digitizare, prin obiectivele și scopurile lui, constituie un element strategic în dezvoltarea societății informaționale la nivel național, interconectată la nivel internațional.
- programul național de digitizare va fi realizat prin colaborarea tuturor instituțiilor deținătoare de documente ce pot fi incluse în Biblioteca Digitală a României.
- procesul de digitizare la nivel național trebuie să aibă la bază principiul partajării.
- Biblioteca Națională a României prin atribuțiile conferite de cadrul legal, prin structură și organizare poate coordona, la nivelul sistemului național de biblioteci, activitatea de digitizare și de constituire a Bibliotecii Digitale a României.
- se va aplica principiul evitării duplicării eforturilor
- se impune recuperarea rezultatelor proiectelor anterioare și integrarea lor într-un sistem unitar, în cadrul Bibliotecii Digitale Naționale și inventarierea tuturor documentelor scanate la nivel național.
- soluția de digitizare poate fi: digitizarea în cadrul bibliotecii, digitizarea prin externalizare și o soluție mixtă.
- corpusul documentar propus pentru digitizare va fi constituit atât din colecții de tip patrimonial, cât și din colecții contemporane, cu un grad mare de adresabilitate și interes.
- instituția deținătoare a documentului original care a fost inclus în Biblioteca Digitală a României acordă utilizatorilor dreptul de acces și dă drept de difuzare a documentului scanat.
- modelul descentralizat, promovat și de The European Library – Europeana, este modelul tuturor bazelor de date europene și reprezintă varianta optimă pentru Biblioteca Digitală a României.