

SECȚIUNEA 1

RAPORTUL ȘTIINȚIFIC ȘI TEHNIC (RST)

ETAPA DE EXECUȚIE NR. 3

CU TITLUL: Dezvoltarea componentelor depozitelor de conținut digital tehnic

- RST- Raport științific și tehnic în extenso
- Proces verbal de avizare internă
- Procese verbale de recepție a lucrărilor de la parteneri

CUPRINS

| | |
|--|----------|
| Obiectivele generale ale proiectului..... | 4 |
| Obiectivele etapei de execuție..... | 7 |
| Rezumatul etapei..... | 8 |
| 1. Selectarea tipurilor de documente tehnice pe suport tradițional care vor constitui conținutul depozitului digital..... | 10 |
| 1.1. Carte cu profil tehnic..... | 10 |
| 1.2. Catalog..... | 12 |
| 1.3. Buletin informativ..... | 16 |
| 1.4. Manual de utilizare..... | 17 |
| 1.5. raport de cercetare..... | 20 |
| 1.6. Brevet de invenție..... | 23 |
| 1.7. Standard, normă..... | 24 |
| 1.8. Teză de doctorat..... | 30 |
| 2. Identificarea tipurilor de sisteme de informare și documentare care gestionează documente tehnice..... | 33 |
| 2.1. Exemple de cataloage „Noua Generație”..... | 34 |
| 2.1.1. AquaBrowser..... | 34 |
| 2.1.2. VuFind..... | 36 |
| 2.1.3. Primo..... | 40 |
| 2.1.4. Encore..... | 43 |
| 2.2. Alte exemple de posibile cataloage "Next Generation"..... | 45 |
| 3. Prelucrarea resurselor tehnice în format tradițional și în alte formate decât cel hârtie..... | 46 |
| 3.1. Aspecte tehnice privind digitizarea documentelor tradiționale..... | 46 |
| 3.1.1. Programe folosite în procesul de digitizare din formatul tradițional..... | 49 |
| 3.2. Aspecte privind digitizarea documentelor audio și video..... | 49 |
| 3.2.1. Tehnici de digitizare ale documentelor audio..... | 49 |

| | |
|---|-----|
| 3.2.2. Arhitectura proceselor de digitizare..... | 54 |
| 3.2.3. Formate de intrare..... | 56 |
| 3.2.4. Dispozitive hardware..... | 60 |
| 3.2.5. Software de prelucrare video..... | 65 |
| 3.2.6. Formate de stocare..... | 67 |
| 4. Realizarea modelului experimental al depozitelor cu conținut digital tehnic..... | 76 |
| 4.1. Aplicație practică-conversia unei cărți vechi din format static jpg în format pdf..... | 76 |
| 4.1.1.Sistem de regăsire a informației în documentele digitizate în format ststic jpeg cu transformarea în pdf cu posibilitate de căutare după cuvânt în textul documentului..... | 76 |
| 4.2. Desfășurarea aplicației practice. Conversia cărții din format jpeg în Format pdf folosind AbbyFineReader 9.0..... | 83 |
| 5. Propunere pentru realizarea unui model conceptual al sistemelor de informare și documentare cu conținut tehnic..... | 99 |
| 5.1. Descrierea unui model conceptual pentru sistemele de informare și documentare cu conținut tehnic..... | 99 |
| 5.2. Specificații de proiectare a modulelor. Sistemul integrat al bibliotecii digitale universitare..... | 104 |
| 5.2.1. Modul de digitizare al documentelor..... | 104 |
| 5.2.2. Modul de autentificare al utilizatorilor..... | 105 |
| 5.2.3. Modul de arhivare a documentelor digitizate..... | 107 |
| 5.2.4. Modul de întreținere a Catalogului bibliotecii digitala..... | 109 |
| 5.2.5. Modul de căutare a documentelor în sistemul UDL..... | 110 |
| 5.2.5. Modul de afișare a documentelor în sistem UDL..... | 112 |
| 5.3. Infrastructura tehnică de implementare a sistemelor de arhivare electronică..... | 114 |

| | |
|---------------------------------|-----|
| 6. Concluzii..... | 116 |
| 7. Bibliografie, webgrafie..... | 118 |
| 8. Anexă..... | 122 |

OBIECTIVELE GENERALE ALE PROIECTULUI

Prin cercetările ce se vor întreprinde pe parcursul derulării proiectului TEMACOD, acesta își propune realizarea următoarelor obiective generale:

- a) Îmbunătățirea, coordonarea și eficientizarea proceselor de digitizare a resurselor informaționale din domeniul tehnic din universități și alte tipuri de organizații;
- b) Sporirea numărului de resurse informaționale tehnice reprezentative digitizate, diversificarea și conservarea acestora;
- c) Creșterea gradului de accesibilitate a publicului utilizator de Internet la resursele informaționale tehnice.

Pentru atingerea acestor obiective proiectul TEMACOD are în vedere și realizarea următoarelor obiective specifice în cadrul fiecărui obiectiv general:

Obiectivul general (a)

1. Identificarea sistemelor de informare și documentare care vehiculează documente tehnice și coordonarea modalităților de digitizare a resurselor informaționale în organizații;
2. Soluționarea și coordonarea aspectelor logistice și documentare pe care le presupune digitizarea pentru înlăturarea paralelismelor și suprapunerilor, pe fiecare arie tematică în parte;
3. Crearea unor mecanisme de coordonare, gestionare și întreținere a procesului de digitizare, care să permită o arhitectură coerentă a întregului sistem, la nivel de proiect.

Obiectivul general (b)

1. Identificarea unui segment documentar reprezentativ de resurse informaționale, pe fiecare arie tematică;

2. Dezvoltarea unui sistem performant și eficient în domeniul tehnologiei informației, care să permită crearea unei platforme unice și unitare de acces la resursele digitizate, precum și conservarea materialului digitizat.

Obiectiv general (c)

1. Aplicarea reglementărilor legale referitoare la drepturile de autor și la proprietatea intelectuală a resurselor informaționale;
2. Instituirea unui sistem unic și unitar de catalogare și/sau clasificare și descriere a resurselor informaționale, pe fiecare arie tematică în parte.

Programul național de informatizare a societății românești, cu un sens mai larg de integrare a culturii și științei naționale în circuitele mondiale, generează în mod natural și logic un nou tip de acces la valorile tehnice și științifice din țară și din lume, un nou tip de relații între marea categorie a celor dornici de a ști și grupul restrâns ca dimensiuni al celor care oferă informații: instituțiile educaționale, de cercetare, de cultură și artă.

În cadrul acestui program, un rol important revine procesului de informatizare a bibliotecilor, care a trecut de la faza de înregistrare a datelor în cataloage, la cea de dezvoltare de sisteme integrate de bibliotecă. În prezent ne aflăm în etapa de constituire și dezvoltare a colecțiilor de date digitale în biblioteci, care va permite ulterior interconectarea bibliotecilor digitale în rețele, bazate pe infrastructuri de comunicații.

Proiectul propus în cadrul Programului 4, „Parteneriate în domeniile prioritare”: *Tehnici pentru managementul conținutului digital-TEMACOD* are ca obiective principale: dezvoltarea de biblioteci digitale, bazate-Web, la nivelul instituțiilor de învățământ superior implicate în contract, interconectarea lor și accesarea rapidă via Internet, printr-o interfață prietenoasă și eficientă.

Tema abordată este de mare actualitate în întreaga lume, se pune problema construirii unei Biblioteci digitale la nivel mondial, accesibilă prin Internet, gratuită și în format multilingv, care să conțină materiale semnificative.

În constituirea acestei biblioteci digitale mondiale, un rol important va reveni mediului academic, bogat în resurse educaționale, de cercetare și care are deja o mare parte din fondul infodocumentar digitizat.

Spre deosebire de bibliotecile tradiționale, o bibliotecă digitală are colecțiile stocate în format digital și accesibile prin calculatoare conectate local sau la distanță, prin rețele de calculatoare. De asemenea, ea oferă un sistem performant de regăsire a informației.

Ca rezultat al dezvoltării Internetului și a potențialului său de căutare, bibliotecile digitale cum ar fi Biblioteca Europeană și Biblioteca Congresului sunt acum dezvoltate într-un mediu bazat-Web. Bibliotecile publice, cele școlare și academice sunt de asemenea capabile să dezvolte website-uri care să permită preluarea de informație științifică digitizată, din categoria cărților scrise, a celor audio sau video, schimbând fundamental noțiunea de resursă educațională.

Multe biblioteci academice sunt activ implicate în constituirea depozitelor de cărți, buletine științifice, teze de doctorat și alte lucrări care pot fi digitizate sau au fost de la început preluate în format digital. Aceste depozite sunt accesibile în rețeaua universității, dar și în afara ei, cu anumite restricții, ce se referă în special la protejarea rezultatelor cercetării și care sunt în concordanță cu politica adoptată de instituție privind accesul liber.

OBIECTIVELE ETAPEI DE EXECUȚIE

ACTIVITATE: Selectarea tipurilor de documente tehnice pe suport tradițional care vor constitui conținutul depozitului digital.

**Obiectiv: Identificarea ariilor tematice din domeniul tehnic. Lista ariilor tematice.
Exemplificarea tipurilor de documente specifice.**

ACTIVITATE: Identificarea tipurilor de sisteme de informare și documentare care gestionează documente tehnice.

**Obiectiv: Prezentarea programelor specifice realizării depozitului digital.
Constituirea înregistrărilor electronice.**

ACTIVITATE: Prelucrarea resurselor tehnice în alte formate decât cel hârtie.

Obiectiv: Constituirea înregistrărilor pentru resursele audiovizuale.

ACTIVITATE: Realizarea modelului experimental al depozitelor cu conținut digital tehnic.

Obiectiv. Crearea modelului experimental și al specificațiilor.

REZUMATUL ETAPEI

Proiectul de cercetare **TEHNICI PENTRU MANAGEMENTUL CONTINUTULUI DIGITAL** și-a propus în etapa 3, conform Planului de activități, atingerea obiectivelor prin finalizarea a patru activități de bază și anume *Selectarea tipurilor de documente tehnice pe suport tradițional care vor constitui conținutul depozitului digital* ca primă activitate, *Identificarea tipurilor de sisteme de informare și documentare care gestionează documente tehnice*, *Prelucrarea resurselor tehnice în alte formate decât cel hârtie* și *Realizarea modelului experimental al depozitelor cu conținut digital tehnic*. ca activitate care va constitui punctul de plecare al sistemului informatic proiectat.

Activitățile de cercetare efectuate au avut ca scop identificarea ariilor tematice din domeniul tehnic și stabilirea celor mai frecvent întâlnite tipuri de documente tehnice din sistemele de informare și documentare. Accentul s-a pus pe identificarea conținutului tehnic. De asemenea, cercetarea s-a îndreptat către evidențierea programelor specifice realizării depozitelor digitale și construirea unui model experimental de creare și gestionare a unui deposit digital de profil tehnic.

Metodologia de cercetare a avut la bază un studiu teoretic și o cercetare practică care s-a desfășurat în diferite structuri de informare și documentare, în principal a partenerilor implicați în proiect. Au avut loc diverse activități de schimb de păreri, dezbateri cu privire la identificarea posibilelor resurse supuse digitizării, organizare de mese rotunde, workshopuri, etc.

Raportul de cercetare pentru această etapă poate fi considerat ca fiind o sinteză din punct de vedere teoretic și practic a activităților întreprinse pe parcursul desfășurării etapei. Raportul științific conține:

Selectarea tipurilor de documente tehnice pe suport tradițional care vor constitui conținutul depozitului digital

Luând în considerare faptul că partenerii implicați în desfășurarea proiectului apar în universităților unde funcționează facultăți cu profil tehnic, majoritatea tipurilor de documente au fost selecționate din bibliotecile care le deservește. S-a pornit de la documentele primare și d-s trcut către documentele special de tipul brevetelor, standardelor, buletinelor informative, manualelor de utilizare, etc. a conținutului digital.

Identificarea tipurilor de sisteme de informare și documentare care gestionează documente tehnice. A pus în evidență sisteme de informare și documentare din "Noua Generație". Acestea oferă acces la conținut integrat și mai mult decât atât funcții sociale prin comunicare directă.

Prelucrarea resurselor tehnice în alte formate decât cel hârtie S-a pornit de la digitizarea documentelor în format tradițional exemplificând principalele programe utilizate. Accentul s-a pus pe prelucrarea documentelor în format audio și video, care prin procedee specifice de digitizare vor completa depozitul digital al documentelor tehnice.

Ultima activitate a avut ca scop *Realizarea modelului experimental al depozitelor cu conținut digital tehnic.* Astfel s-a realizat o aplicație practică pentru conversia unei cărți vechi din format static jpeg în format pdf . S-a evidențiat și un sistem de regăsire a informației în documentele digitizate în format static jpeg cu transformare în format pdf cu posibilitate de căutare după cuvânt în textul documentului. Deasemenea s-a realizat și o propunere pentru realizarea unui model conceptual al sistemelor de informare și documentare cu conținut tehnic care să implice toți partenerii.

Aparatul critic al raportului de cercetare este constituit din Bibliografie Selectivă, Webgrafie Selectivă, Anexă – exemplificare asupra metodelor de digitizare a documentelor tehnice (cărți vechi și rare).

1. Selectarea tipurilor de documente tehnice pe suport tradițional care vor constitui conținutul depozitului digital

1.1. Carte cu profil tehnic

Cartea face parte dintre documentele primare utilizate în informare și documentare. În cercetarea noastră, cartile cu profil tehnic vor constitui elementele de interes. Pentru o primă etapă în prelucrarea acestora, se va întocmi o descriere bibliografică într-un sistem dedicat de bibliotecă. Apoi, acestuia îi se poate atașa o bază de date full-text, obținută prin tehnologiile prezentate mai jos.

Exemplu:

- Nedelcu, Anișor - **Tehnologii de prelucrare mecanică și neconvenționale** / Anișor Nedelcu . - Brașov : Editura Universității "Transilvania", 2003- . - Vol. ; 24 cm 621.9(075.8)

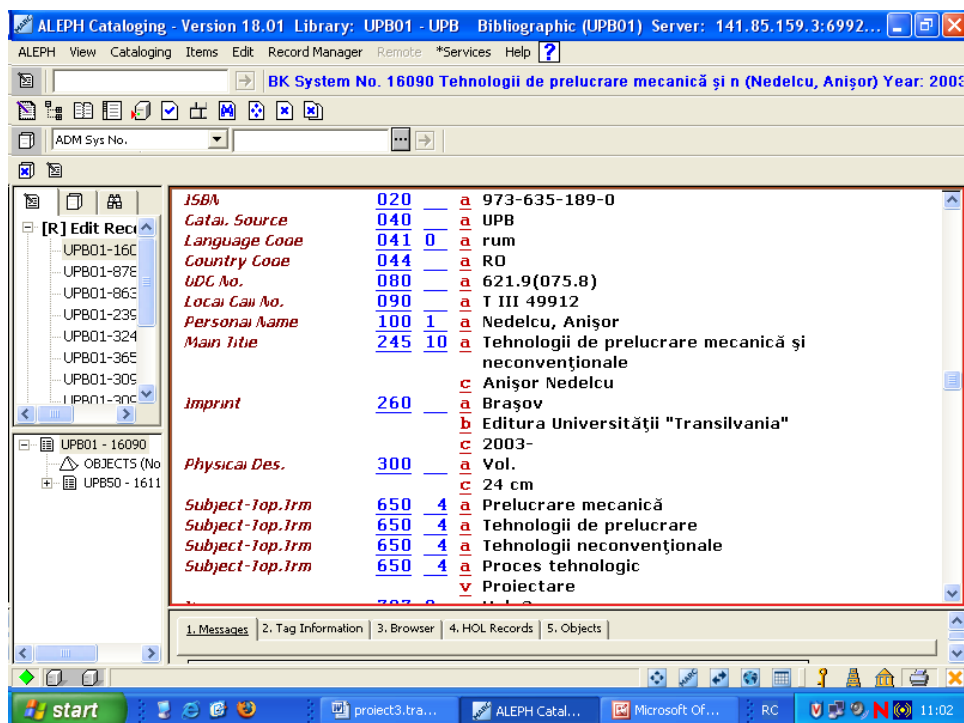


Fig. 1.1. Descrierea bibliografică a cărții „Tehnologii de prelucrare mecanică și econvenționale”

- Popescu, Gheorghe ; Ghimiși, Ștefan Sorinel - **Automatizarea proceselor tehnologice** [Text tipărit] / Gheorghe Popescu, Ștefan Sorinel Ghimiși . - Târgu-Jiu : Academica Brâncuși, 2009 . - 267 p. : fig., tab. ; 26 cm 681.5:621 (T III 51265, Sala Centrală de Lectură)

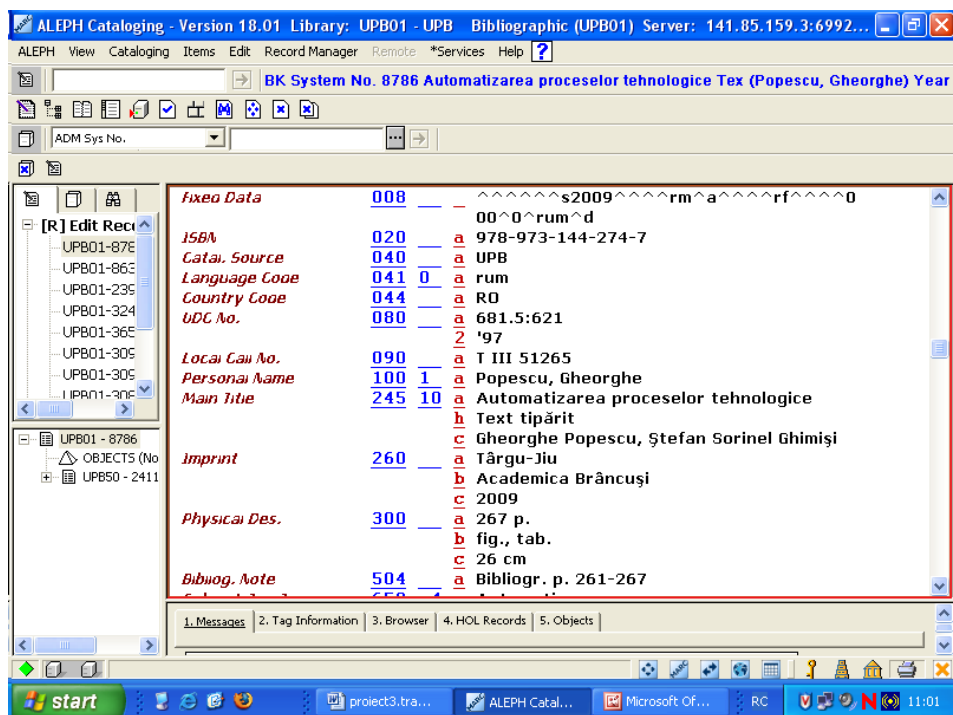


Fig. 1.2. Descrierea bibliografică a cărții „Automatizarea proceselor tehnologice”

Amza, Gheorghe - **Procese de operare** / Gheorghe Amza, Cătălin Gheorghe Amza, Constantin Radu, Alexandru Dumitrache-Rujinski - București : Bren, 2001-2004 . - 3 vol. ; 24 cm 621.9 (T III 49352, *Biblioteca de Mecanică*)

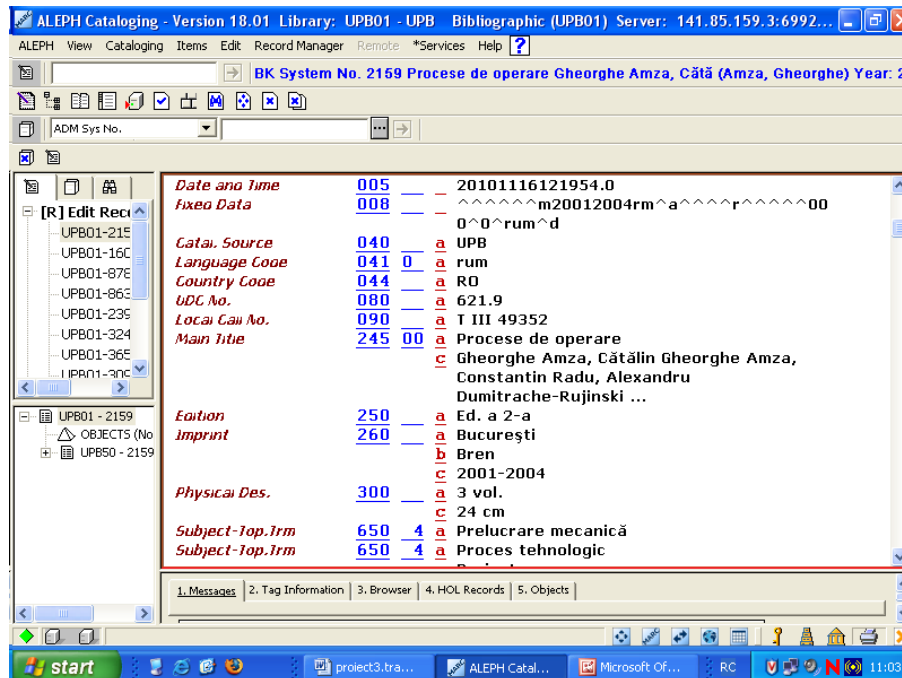


Fig. 1.3. Descrierea bibliografică a cărții „Procese de operare”

1.2. Catalog

Alt tip de documente tehnice avute în vedere îl reprezintă cataloagele. Conform Dictionarului explicativ al Limbii Române, catalogul reprezintă o „Listă, caiet, registru, sistem de fișiere etc. care conține o însirare metodică, după anumite criterii și cu anumite scopuri, de nume de ființe sau de obiecte, titluri de cărți etc. – din fr. **catalogue**, lat. **catalogus**.”

Și ele beneficiază de descrierea conform normelor biblioteconomice și realizarea înregistrărilor în sistemul informatic dedicat.

Punți redresoare monofazate 1 PM : [catalog] . - București : IPRS Băneasa, 1982-1983 . - pag. spec. 621.3.03 (CS 1398, Biblioteca de Electronică și Telecomunicații)

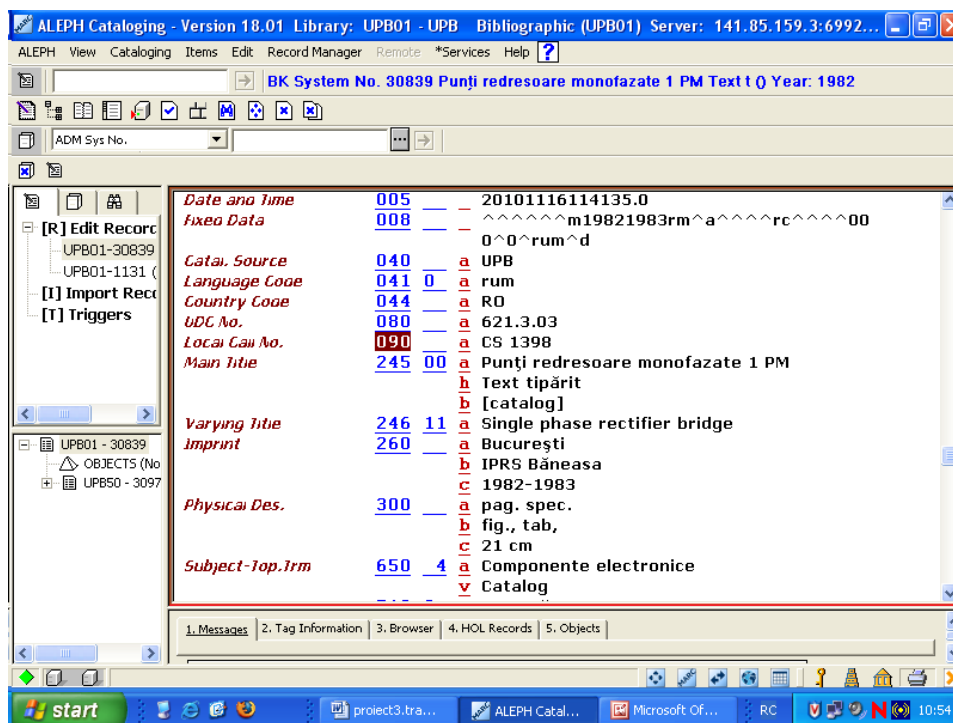


Fig. 1.4. Descrierea bibliografică pentru catalogul „Punți redresoare monofazate 1 PM”

- **Tranzistoare NPN : [catalog] . - București : IPRS Băneasa, 1983 . - pag. spec. :**
fig. 621.382.3 (CS 1397, *Biblioteca de Electronică și Telecomunicații*)

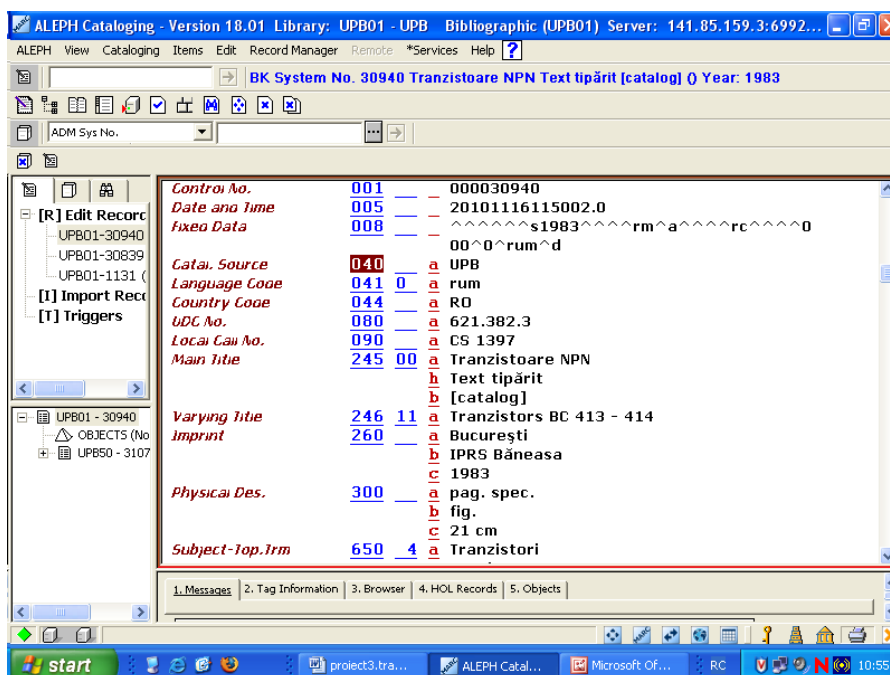


Fig. 1.5. Descrierea bibliografică pentru catalogul „Tranzistoare PNP”

- **Catalogul reglementărilor și prescripțiilor tehnice valabile în sectorul energetic în anul 2002** / Oficiul de Documentare în Energetică (ed.) . - București : ICEMENERG, 2002 . - 132 p. ; 20 cm

62(083.1) (CS 1861, *Biblioteca de Energetică*)

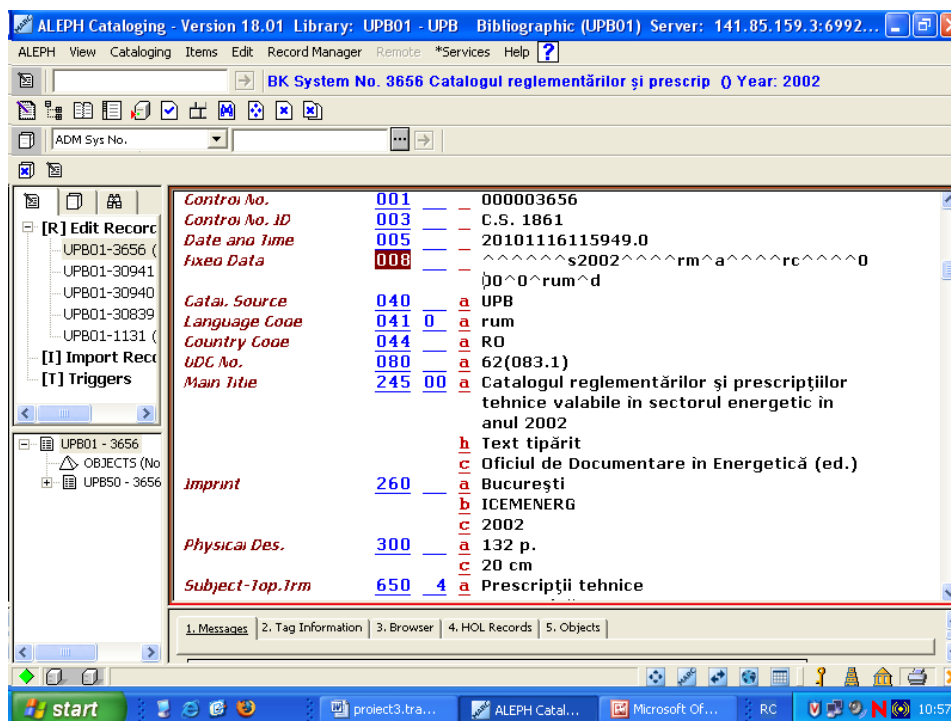


Fig. 1.6. Descrierea bibliografică pentru "Catalogul reglementărilor din sectorul energetic"

- **Catalogul standardelor române** [Text tipărit] : 2008 / Asociația de Standardizare din România ; red.: dr. ing. Speranța Stomff . - București : Standardizarea, 2009 . - LX, 1684 p. ; 30 cm

006"2008"(498)(083.8)

(CS 1912, *Sala Centrală de Lectură, Biblioteca de Chimie Industrială, Biblioteca de Mecanică*)

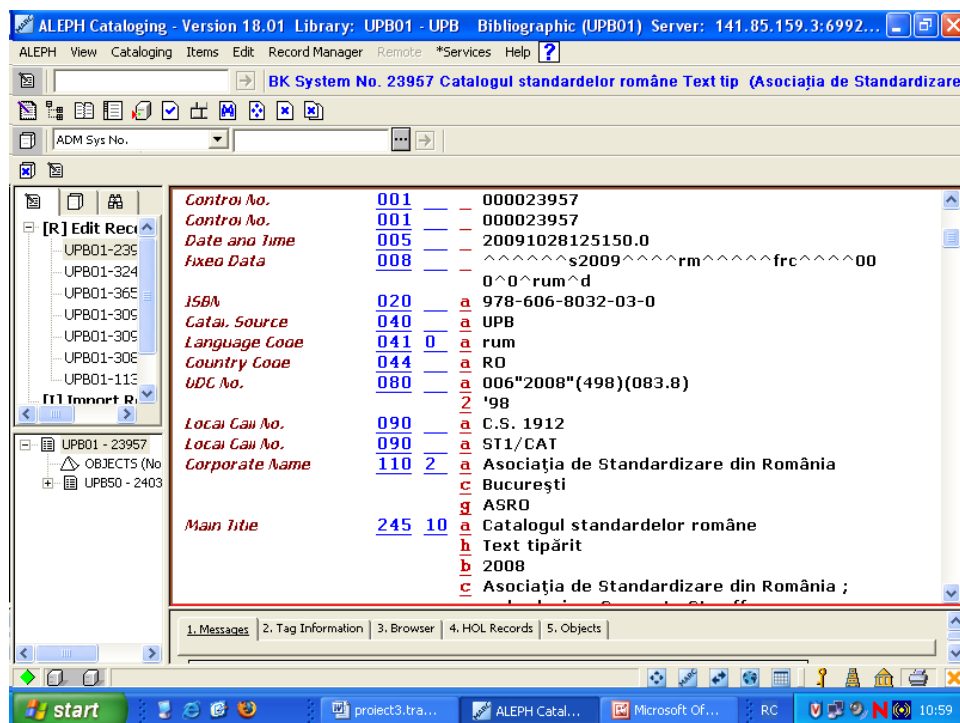


Fig. 1.7. Descrierea bibliografică pentru „catalogul standardelor române”

- **Scheme electronice pentru televizoare : Catalog . - București : General Partner, 1997 . - 160 p. : sch.**

621.397.132(083.8): 003.63

(CS 1765, Sala Centrală de Lectură, Biblioteca de Electronică și Telecomunicații)

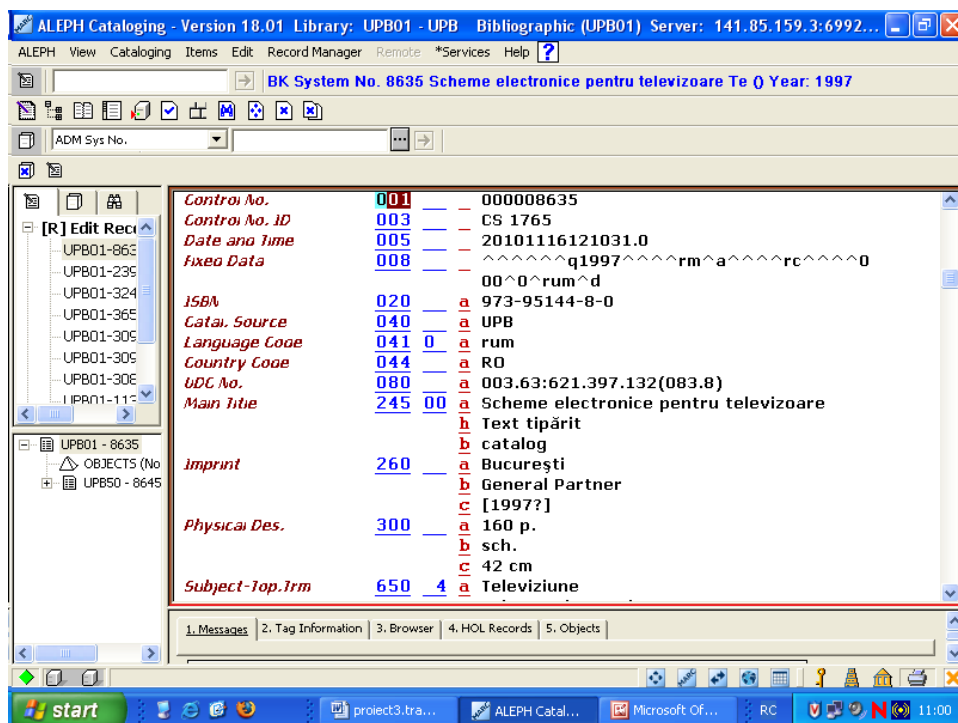


Fig. 1.8. Descrierea bibliografică pentru "Scheme electronice pentru televizoare..."

1.3. Buletin informativ

Conform Dictionarului explicativ al Limbii Romane

Buletin Informativ . – București : Administrația prezidențială, 2001

002(05)(498)București (P 6743, *Sala de Referințe*)

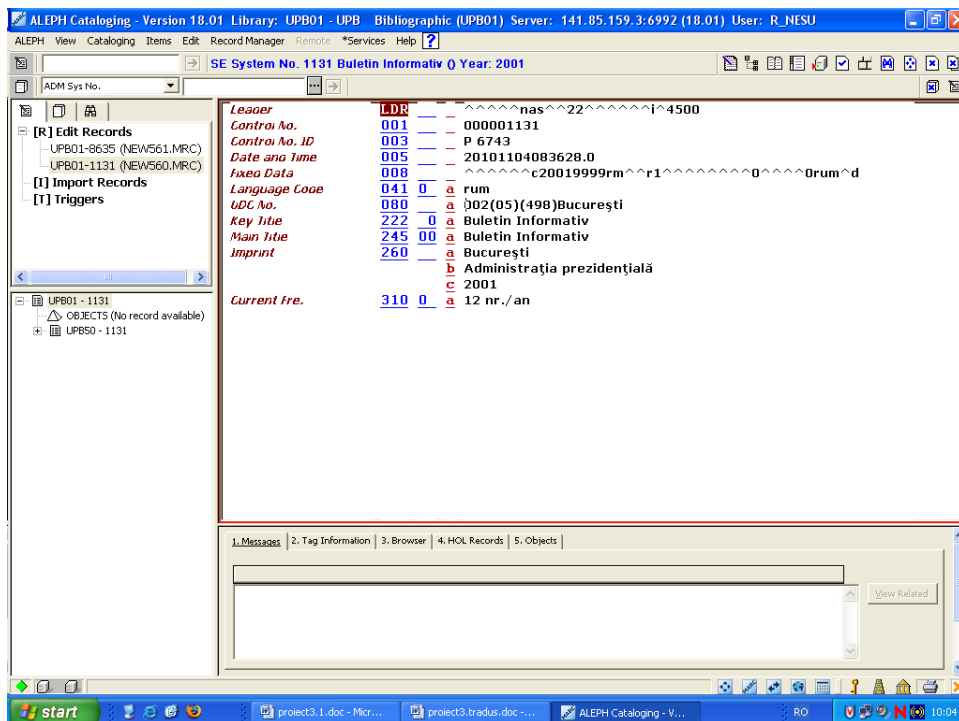


Fig. 1.9. Descrierea bibliografică pentru "Buletinul informativ"

1.4. Manual de utilizare

- Munteanu, Florin – **Alimentarea cu energie electrică a consumatorilor : manual de utilizare a programelor de instruire (CAT) și proiectare (CAD) asistată** / dr. ing. Florin Munteanu . – Iași : Universitatea Tehnică „Gh. Asachi”, 1998 . – 92 p. Multigr. : fig. 621.311:681.3.006.2(075.8) (T II 35641, *Biblioteca de Electrotehnică*)

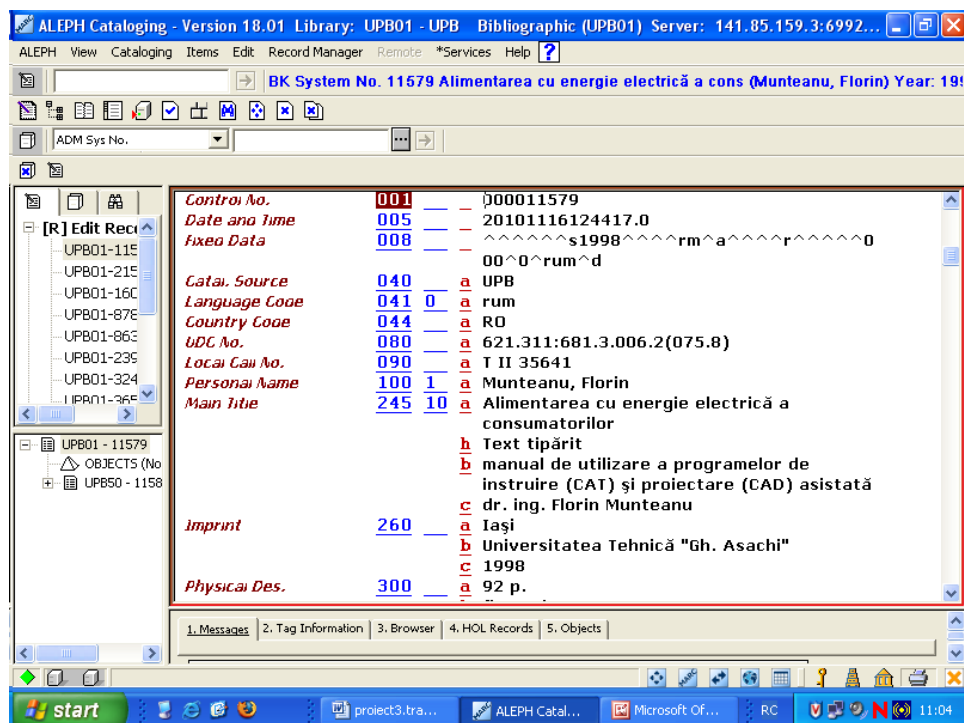


Fig. 1. 10. Descrierea bibliografică pentru “Alimentarea cu energie electrică a consumatorilor”

- Edelhauser, Eduard ; Ionică, Andreea - **Sisteme de gestiune a bazelor de date Access, FoxPro : manual de utilizare** / Eduard Edelhauser, Andreea Ionică . - Petroșani : Universitas, 2002 . - 173 p. : fig. ; 24 cm 004.65

(T III 49066, Sala Centrală de Lectură)

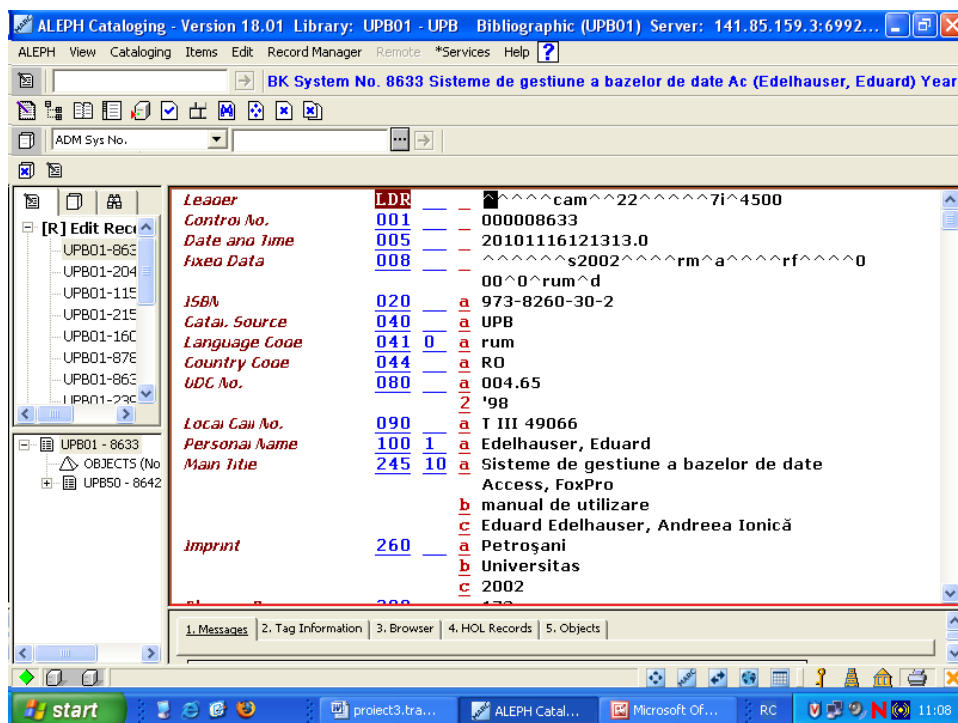


Fig. 1.11. Descrierea bibliografică pentru „Sisteme de gestiune a bazelor de date...”

- Plahteanu, Boris ; Horodincă, Mihăiță - **Parametrii încărcării energetice - Resurse de monitorizare, diagnoză și conducere automată în sistemele de fabricație** [Text tipărit] : manual de utilizare / prof. univ. dr. ing. Boris Plahteanu, dr. ing. Mihăiță Horodincă . - Iași : Performantica, 2007 . - 69 p. : fig., graf. ; 24 cm + 1 CD-ROM (12 cm)

681.518.5(075.8)

(T III 50945, Sala Centrală de Lectură)

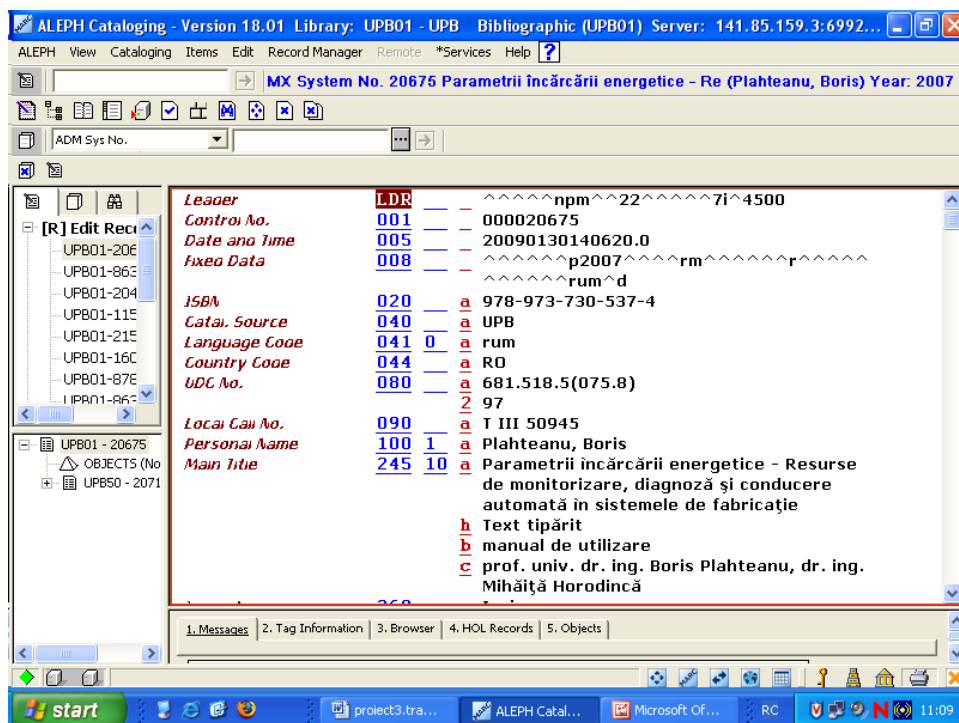


Fig. 1.13. Descrierea bibliografică pentru “Parametrii încărcării energetice...”

1.5. Raport de cercetare

Proiectarea filtrelor de semnal utilizând programarea semidefinită : raport de cercetare / director de proiect: conf. dr. ing. B. Dumitrescu . - [s.l.] : [s.n.], 2002 . - [20] p. : fig. ; 30 cm 621.372.542 (CS 1863, *Biblioteca de Automatică și Calculatoare*)

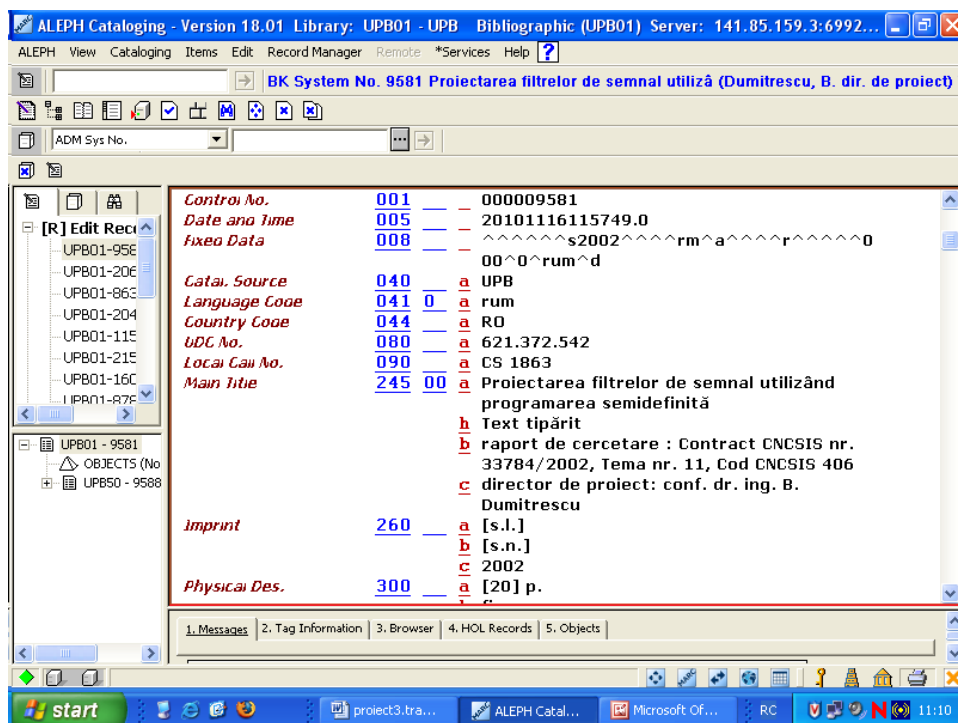


Fig. 1.13. Descrierea bibliografică pentru „Proiectarea filtrelor de semnal...”

- **Cuantificarea obiectivă a calității și similarității imaginilor color prin modelarea percepției vizuale umane :** Grant CNCISIS 13 / director de grand : conf. dr. ing. Constantin Vertan . - [s.l.] : [s.n.], 2003 . - Vol. ; 30 cm

004.81:159.937.51 * 004.932 (CS 1891, *Biblioteca de Electronică și Telecomunicații*)

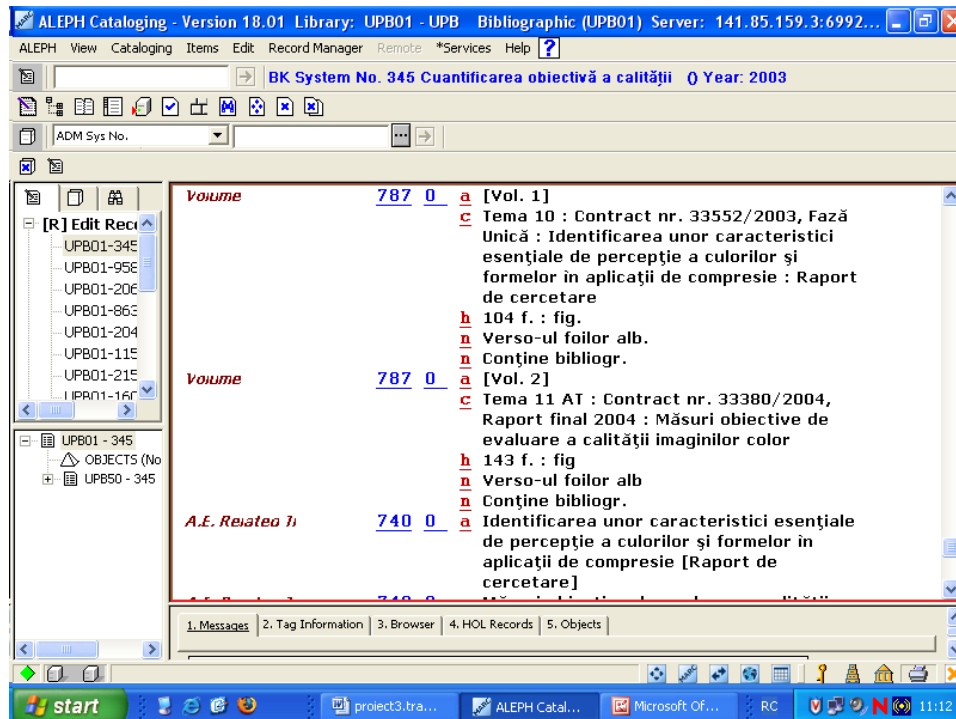
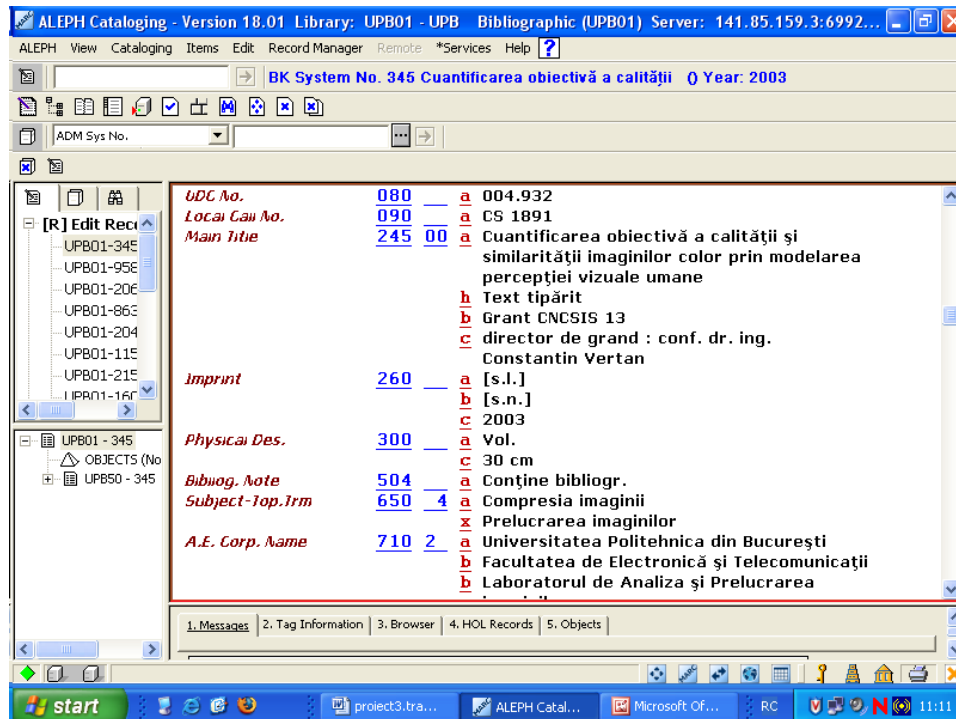


Fig. 1.14. Descrierea bibliografică pentru „Cuantificarea obiectivă a calității...”

1.6. Brevet de invenție

Oficiul de Stat pentru Invenții și Mărci (O.S.I.M.) elaborează și supune spre aprobare Guvernului strategia dezvoltării protecției proprietății industriale în România și aplică politica în domeniu. ♦ <http://www.osim.ro>

O.S.I.M. administrează următoarele legi și acte normative care definesc legislația română în domeniul protecției proprietății industriale: Legea nr. 64/1991 privind brevetele de invenție;

- H.G. nr. 152/1992 cu privire la Regulamentul de aplicare a Legii nr.64/1991 privind brevetele de invenție;

- Legea nr. 129/1992 privind protecția desenelor și modelelor republicată în temeiul art. IV din Legea nr. 280/2007, publicată în M.Of.nr.876/20.12.2007

- Legea nr. 16/1995 privind protecția topografiilor produselor semiconductoare;

- Legea nr. 84/1998 privind mărcile și indicațiile geografice;

- O.G. nr. 41/1998 privind taxele în domeniul protecției proprietății industriale și regimul de utilizare a acestora;

- Legea nr. 255/1998 privind protecția noilor soiuri de plante;

- Legea nr. 75/1999 privind recunoașterea internațională a depozitului de microorganisme în scopul procedurii de brevetare, conform Tratatului de la Budapesta la care România a aderat;

- Legea nr. 93/1998 privind protecția tranzitorie a brevetelor de invenție

- Normele nr.242/1999 privind sprijinirea brevetării în străinătate a invențiilor românești.



Fig. 1.15. Criterii de căutare în baza de date RoPatent

1.7. Standarde

Conform Ghidului ISO/CEI2:1996, standardul este un document, stabilit prin consens și aprobat de către un organism recunoscut, care asigură, pentru uz comun și repetat, reguli, linii directoare sau caracteristici pentru activități sau rezultatelor lor, cu scopul de a se obține gradul optim de ordine într-un anumit context. Reprezintă o normă sau un ansamblu de norme care reglementează calitatea, caracteristicile, forma etc. unui produs, un document în care sunt consemnate aceste norme.

Un standard înseamnă printre altele, un nivel de experiență și tehnologie care face ca prezența industriei în elaborarea sa să fie indispensabilă. Acesta este un document de referință. Standardul trebuie să se bazeze pe rezultatele consolidate ale științei, tehnologiei și experienței și să fie orientat spre atingerea optimului de beneficii pentru comunitate.

Standardul este o operă colectivă. Includerea unui standard național în programul național de standardizare și elaborarea acestuia se face sub autoritatea organismului național de standardizare, care îl și publică. Prin urmare, standardul național este o operă

protejată încă din faza de proiect, dreptul de autor aparținând organismului național de standardizare.

Un standard reprezintă un nivel de experiență și tehnologie care face ca prezența industriei în elaborarea sa să fie indispensabilă. Acesta este un document de referință.

I.S.O (International Organization for Standardization) este cel mai mare creator de standarde. Deși principala activitate a I.S.O este dezvoltarea standardelor tehnice, standardele ISO au de asemenea importante repercursiuni economice și sociale. Standardele ISO creează o diferență pozitivă și contează nu numai pentru inginerii și producătorii pentru care rezolvă probleme de bază în producție și distribuție, ci și pentru societate ca întreg. Standardele ISO sunt adoptate, traduse și difuzate în România de ASRO - Asociația de Standardizare din România care participă prin specialiști în cadrul comitetelor tehnice internaționale ale ISO.

Principalele caracteristici ale standardelor:

- Standardele sunt documente primare, fixează direct conținutul și rezultatele activității științifice și tehnice;
- Stabilesc pentru utilizări comune și repetate reguli, prescripții sau caracteristici pentru activități sau pentru rezultatele lor în scopul obținerii unui grad optim de ordine într-un context dat;
- Asigură calitatea tehnică și permit accesul prompt la informația științifică și tehnică în domeniu;
- Sunt documente neperiodice, nu au o apariție regulată, nu apar în serie; ele apar o singură dată;
- Sunt documente speciale, destinate unor domenii de interes restrâns;
- Includ prescripții obligatorii sau cu caracter de recomandare ;
- Sunt valabile pe teritoriul unei țări sau pe plan internațional;
- Sunt actualizate : standardele sunt revizuite periodic sau după cum dictează circumstanțele pentru a le asigura actualitatea și, de aceea, evoluează împreună cu progresul social și tehnologic.

Standardele apar sub formă de foi, broșuri sau volum conținând standarde referitoare la aceeași temă.

Principalele elemente ale unui standard sunt :

- indicativul standardului (poate fi sub formă de siglă, număr de ordine și an de apariție) ex. :

a) STAS - standarde românești aprobate înainte de 26 august 1991 (fig.1)

The image shows the cover page of a Romanian Standard (STAS) document. At the top left, it says "IRS STANDARD ROMÂN" with handwritten initials "Joc" and "P. Popescu". To the right, it says "STAS 13101" and "Iunie 1992". Below this, there is a box labeled "INVENTARIAT" with "Nr. 191/952" and "ms". The title of the standard is "PANOURI PREFABRICATE DIN BETON ARMAT PENTRU VATRA LUCRĂRILOR MINIERE SUBTERANE". Below the title, there are three lines of text in Romanian, English, and French. The English text reads: "Precast reinforced concrete panels for the floor of the underground mining works". There are sections for "APROBARE" (Approved by the General Director of IRS on 26 November 1991) and "CORESPONDENȚĂ" (On the date of approval, there is no international standard). At the bottom, there is a "DESCRIPTORI TIT" section with the text "Construcție; panou prefabricat; beton armat; lucrare minieră; vatră".

Fig. 1.16. Standard de tip STAS

b) SR - standarde românești aprobate după 26 august 1992 (fig.2)

The image shows the cover page of a Romanian Standard (SR) document. At the top left, it says "IRS STANDARD ROMÂN" with handwritten initials "3ex" and "FB". To the right, it says "SR ISO 254" and "Ianuarie 1996". Below this, there is a box labeled "INVENTARIAT" with "Nr. 191/952" and "ms". The title of the standard is "Transmisii prin curele RCTI Calitate, starea suprafeței și echilibrare". Below the title, there are three lines of text in Romanian, English, and French. The English text reads: "Belt drives - Pulleys - Quality, finish and balance". There are sections for "APROBARE" (Approved by the General Director of IRS on 11 July 1995) and "CORESPONDENȚĂ" (This standard is identical with the international standard ISO 254:1990). At the bottom, there is a "DESCRIPTORI TIT" section with the text "Transmisie prin curele, roată, condiție, starea suprafeței, echilibrare".

Fig. 1.17. Standard de tip SR

c) SR EN - standarde europene adoptate ca standarde românești



Fig. 1.18. Standard de tip SR EN

d) SR ISO (STAS ISO) – standarde internaționale adoptate ca standarde românești



Fig. 1. 19. Standard de tip SR ISO

- e) DIN – standarde germane
- f) AFNOR – standarde franceze
- g) GOST – standarde ruse

- titlul standardului; subtitlul
- aprobare
- corespondență
- descriptori
- adresa instituției care a elaborat standardul

- dreptul de autor
- ediția
- metoda de adoptare
- consimțământul propriu-zis
- bibliografie
- comitetul tehnic
- modificari după publicare

Clasificarea standardelor:

- a) ISO (clasificare internațională) este o clasificare pe subiecte ; standardele ISO sunt numerotate, și au un format de tipul: "ISO 99999:yyyy: Titlul", unde "99999" este numărul standardului, "yyyy" este anul publicării, și "Titlul" descrie obiectul
- b) CAS (clasificare alfanumerică) cuprinde grupe, subgrupe și sectoare; sectorul este notat cu o literă , grupa cu o cifră de la 0 la 9, iar sugrupa este notată cu o a doua cifră de la 0 la 9, de ex.: B02 (elemente de control metalografic)

Descrierea bibliografică a standardelor

Colecția de standarde a unei biblioteci universitare cu profil tehnic acoperă o arie tematică vastă, reprezentată de domeniile de interes specifice: chimie, electrotehnică, energetică, electronică, mecanică, metalurgie, transporturi, aeronautică, ș.a.

Catalogarea standardelor, descrierea bibliografică corectă și precisă, identificarea metadatelor care ajuta la identificarea și localizarea standardelor în baze de date, in cataloage online, in depozite digitale sunt activități importante care contribuie la creșterea valorii conținutului sistemelor de stocare și regăsire a informației.

Elementele bibliografice se iau de pe prima pagina a standardului și sunt reprezentate de (fig.5) :

- vedetă: indicativul complet al standardelor (sigla, numărul și anul de apariție); indeplinește rol de cotă
- titlul standardului: se scrie în formă completă

- subtitlul: definește mai precis conținutul standardului și specifică înlocuirea unui standard cu altul
- ediția: se trece obligatoriu (cu excepția primei ediții)
- datele de publicare: localitate, editură, an
- descriere fizică: nr.pagini, figuri, tabele, etc.
- data aprobării, data intrării în vigoare (când apar)
- observații: ref.anulare/suspendare/ retipărire standard

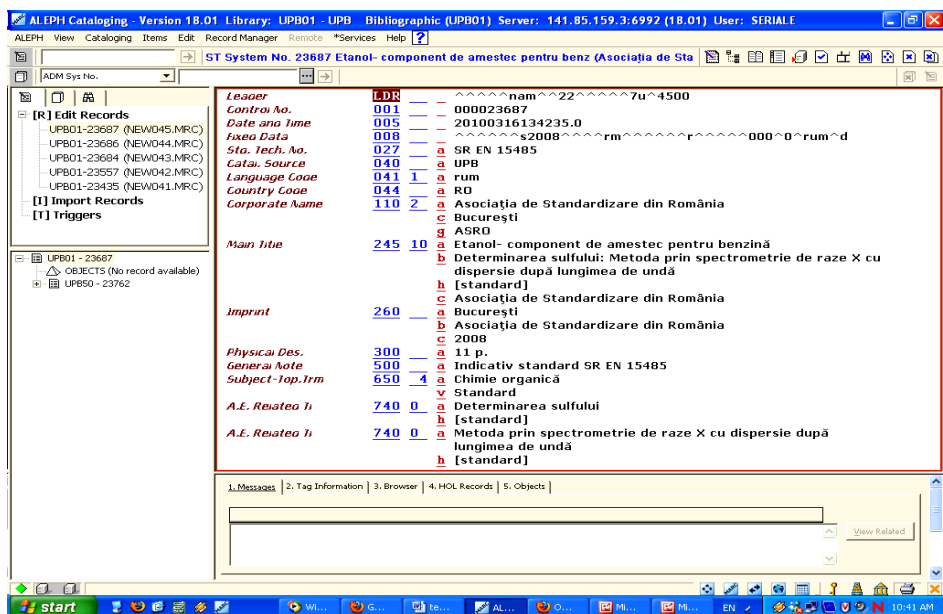


Fig. 1.20. Descrierea bibliografică pentru standard de tip SR

Regăsirea standardelor într-o bază de date se face după o multitudine de criterii :

- format document (format ST)
- indicativ standard
- titlu
- cuvinte cheie
- an apariție
- editură



Fig. 1.21. Căutare standarde după criteriu “cuvant cheie”

1.8. Teză de doctorat

Tezele de doctorat sunt documente primare de informare și documentare, alături de manuale, monografii, tratate, rapoarte științifice, rapoarte tehnice, brevete, invenții, standarde, comunicări științifice.

Sunt documente cu caracter științific destinate unor domenii de interes științific bine precizate, au caracter unicat și au o circulație relativ restrânsă.

De
 asemene

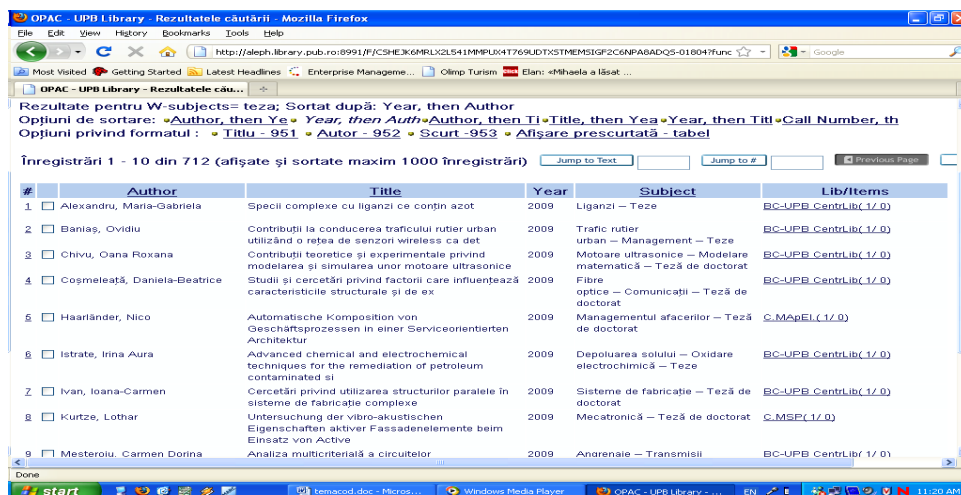


Fig. 1.23. – Căutare titluri de teze după subiect

2. Identificarea tipurilor de sisteme de informare și documentare care gestionează documente tehnice

Cele mai cunoscute sisteme de informare și documentare care pot gestiona documente tehnice sunt bibliotecile și centrele de informare și documentare. Sistemele informatice utilizate în acest scop sunt identificate cu programele dedicate pentru biblioteci, care permit o descriere bibliografică și din ce în ce mai mult sisteme din „noua generație”.

Sistemele de informații digitale dispun de diferite instrumente care facilitează accesarea, căutarea, navigarea, indexarea, stocarea, organizarea și difuzarea de informații digitizate.

Sistemele tradiționale de regăsire a informațiilor (OPACs, bazele de date bibliografice) au dezvoltat instrumente mai sofisticate pentru regăsirea de informații de la arhivele digitale (bibliotecă digitală, depozite instituționale, „Noua Generație” de cataloage - catalogul "Next Generation").

Online Public Access Catalog (OPAC) se referă la conceptul de bibliotecă electronică, fiind o bază de date online a materialelor deținute de o bibliotecă sau un grup de biblioteci și care oferă o accesul de la distanță la resursele bibliotecii. Într-un catalog de bibliotecă, utilizatorii caută de obicei cărți, periodice, materiale audio / video sau alte articole deținute de aceasta.

Catalogul OPAC (Open Public Acces Catalog) cuprinde înregistrari curente și retrospective, permite vizualizarea fișelor bibliografice, oferă acces la informații legate de titlurile publicațiilor, cota acestora, locația/repartiția, numărul și disponibilitatea exemplarelor. Regăsirea informațiilor din baza de date se face după criteriile de căutare multiple: titlu, cuvânt din titlu, subiect, autor, editură, ISBN, ISSN, număr standard.

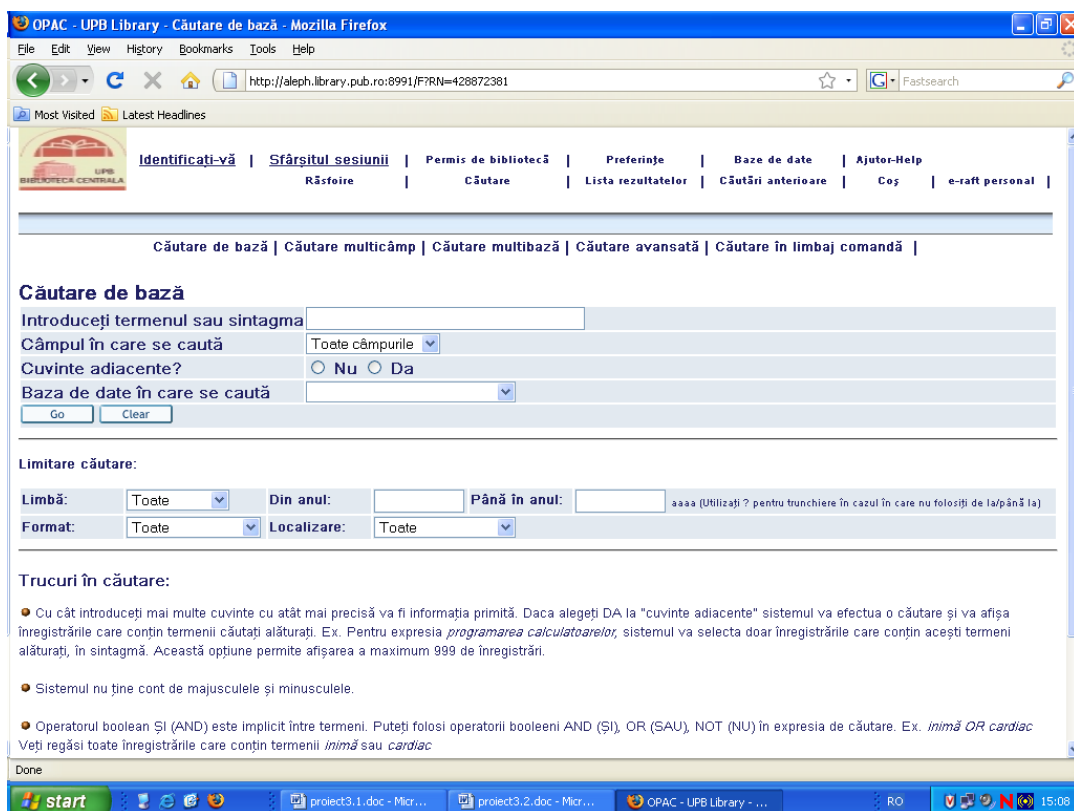


Fig.2.1. Exemplificare a sistemelor dedicate pentru biblioteci. Sistemul ALEPH

Cataloagele "Next Generation" reprezintă o soluție pe care bibliotecile o pot folosi pentru a face materialele lor mai ușor accesibile. Acestea oferă acces la bazele de date ale unei biblioteci sau consorțiu de biblioteci și permite utilizatorilor să se informeze singuri în legătură cu colecțiile bibliotecii. Acest tip de catalog prezintă avantaje față de catalogul clasic de bibliotecă:

- conținut integrat - se pot încorpora date din surse multiple, în mai multe formate;
- funcții sociale – utilizează caracteristici media, cum ar fi etichetarea sau crearea de liste și comunicarea cu alți utilizatori;
- vizualizarea datelor – cataloagele au o interfață prietenoasă și prezintă un mod de căutare simplu, utilizează iconițe pentru a indica rezultatele căutării

2.1. Exemple de cataloage „Noua generație”

2.1.1. *AquaBrowser® Library*

Bibliotecile membre RCLS (Ramapo Catskill Library System) au lansat un nou tip de catalog, dezvoltat de MediaLab Solutions în 2005 ca un instrument local pentru Queens Library.

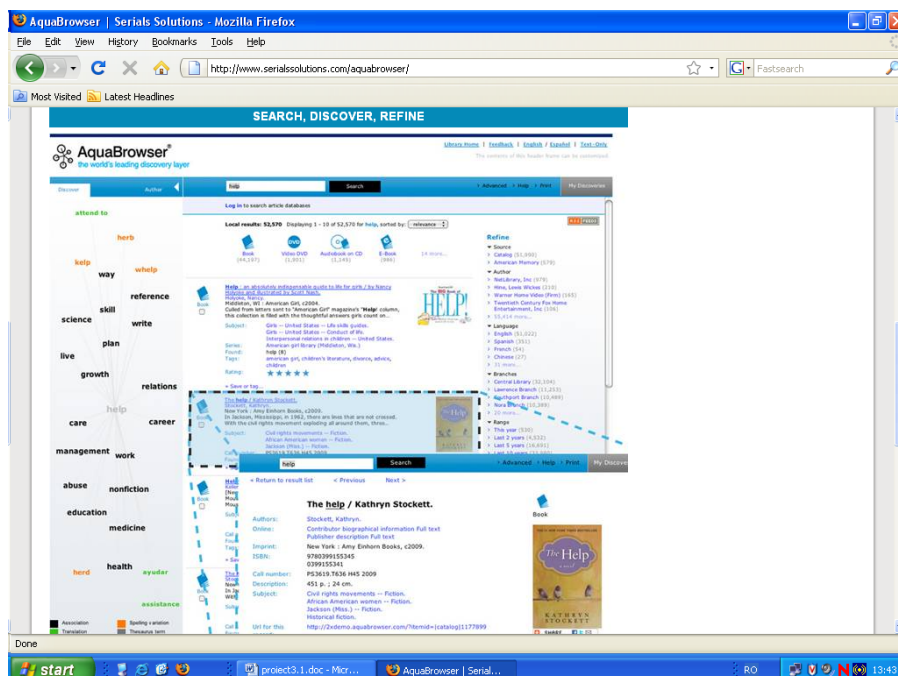


Fig. 2.2. Programul AquaBrowser®

AquaBrowser® oferă o interfață simplă și modernă pentru utilizatori, cu o singură casetă de căutare, astfel încât aceștia pot regăsi rapid informația de care au nevoie utilizând o singură interogare, rezultatele fiind ordonate în funcție de relevanță, în mai multe limbi de circulație internațională: engleză, spaniolă, germană, franceză, rusă, suedeză, norvegiană, olandeză.

De asemenea, utilizatorii pot printa sau trimite prin email articolele sau le pot exporta către instrumente de gestionare bibliografice.

AquaBrowser include:

- rețeaua Web 2.0;
- căutarea subiectului de autoritate;
- două seturi suplimentare de date locale/colecții locale.

Tot mai multe biblioteci utilizează AquaBrowser (120 de sisteme de bibliotecă din Statele Unite, biblioteci din Marea Britanie, Noua Zeelandă, Germania, Belgia, Spania și Italia și 80% dintre bibliotecile din Olanda, aproximativ 2.000 de locații bibliotecă deservind 60 milioane cetățeni din întreaga lume).

AquaBrowser este un instrument de căutare util și inteligent, putându-se conecta la orice număr de surse de date, toate bibliotecile fiind socotite ca surse superioare de

informații și cunoștințe, oferind utilizatorilor bibliotecii posibilități îmbunătățite de căutare și informații suplimentare despre titlurile care le găsesc. El oferă experiența primei căutări interactive, utilizatorii putându-se conecta la multe sisteme de bibliotecă prin crearea de rating, comentarii, și tag-uri pentru titlurile în catalog și integrează cu ușurință multe alte surse într-o interfață rapidă și plăcută. În plus, AquaBrowser permite utilizatorilor să efectueze și să partajeze notele lor și impresiile despre titluri prin caracteristicile „Descoperirile mele”.

Punerea în aplicare este relativ ușoară, nefiind necesară nici o schimbare în sistemele și procesele existente. Aceasta va îmbunătăți în mod automat rezultatele de căutare utilizând metadate suplimentare deja prezente. AquaBrowser poate funcționa împreună cu catalogul automatizat existent pentru a oferi utilizatorilor de bibliotecă posibilități îmbunătățite de căutare, informații suplimentare despre titlurile care le găsesc, putându-se utiliza în continuare interfața tradițională Aleph OPAC (denumit catalog clasic).

Odată ce utilizatorii selectează un anumit titlu pentru a vizualiza, ei pot vedea recenzii de carte, fragmente din primul capitol, sinteze, tabele de cuprins, imagini din carte, multe din aceste înregistrări neregăsindu-se în catalogul tradițional

2.1.2. VuFind®

VuFind® dezvoltat de Universitatea Villanova este un portal de resurse de bibliotecă. ♦

<http://vufind.org>,

♦ <http://www.library.illinois.edu/catalog/vufind/>

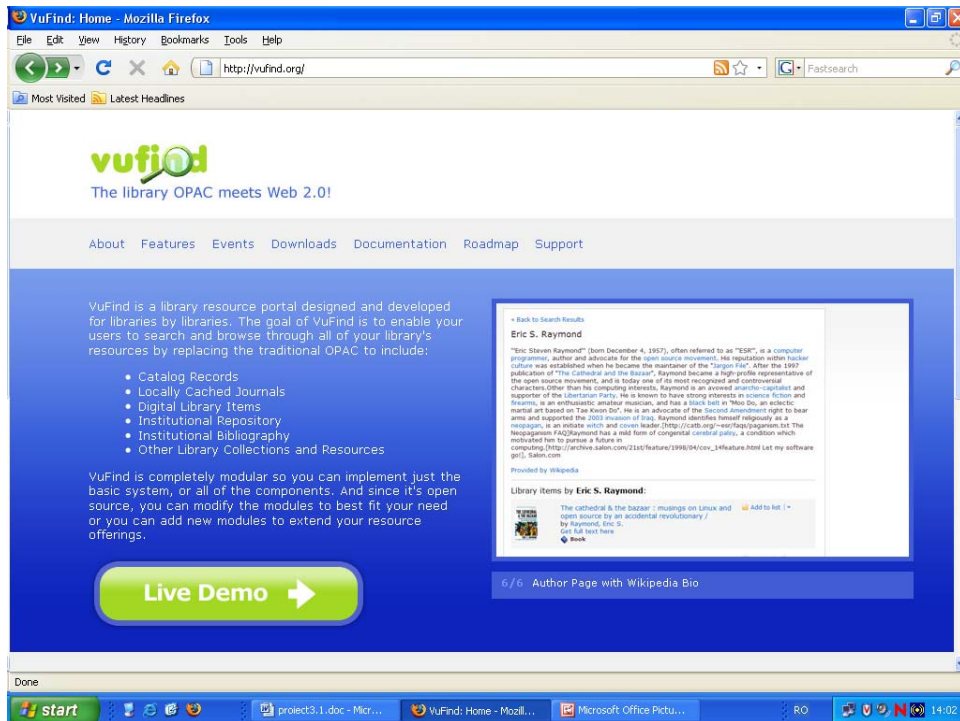


Fig.2.3. Portalul VuFind®-prezentare generală

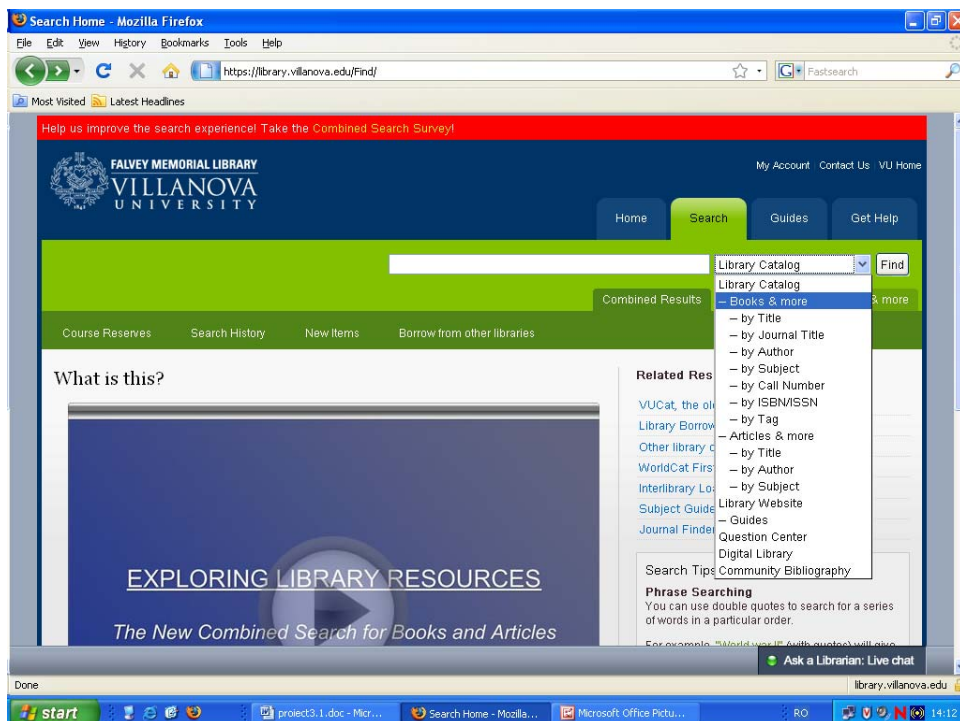


Fig2.4. Portalul VuFind®-facilități de căutare

Scopul VuFind este de a permite utilizatorilor să caute și să navigheze prin toate resursele bibliotecii prin înlocuirea tradiționalului OPAC, incluzând:

- înregistrări în catalog
- produsele bibliotecii digitale
- depozite instituționale
- bibliografii instituționale
- alte categorii de resurse de bibliotecă.

VuFind este complet modular, astfel încât se poate pune în aplicare doar sistemul de bază sau toate componentele. Deoarece este „open source”, se pot modifica sau adăuga module pentru a se potrivi cel mai bine nevoilor bibliotecii. O gamă largă de opțiuni configurabile permite personalizări extinse fără a schimba vreun cod. VuFind rulează pe Solr Energy.

Apache Solr, motorul de căutare open source, oferă o performanță e și scalabilitate pentru a permite VuFind să răspundă la interogări de căutare într-o perioadă de timp de milisecunde și are capacitatea de a fi distribuit pentru a răspândi înregistrarea din catalog pe mai multe servere.

Caracteristici:

- căutare cu rezultate fațetate. Sistemul de căutare permite utilizatorului să caute de la o casetă de căutare de bază și apoi să fie capabil de a restrânge rezultatele, făcând clic pe diversele fațete ale rezultatelor.

- Live Record Status și interogare Ajax .Pagina de rezultate ale căutării este capabilă să afișeze statusul înregistrării prin interogarea catalogului la momentul exact. Și, deoarece acest lucru se face prin intermediul AJAX, după ce rezultatele s-au încărcat, pagina nu va încetini pentru orice motiv.

- mai multe sugestii de resurse. Când se vizualizează o înregistrare, utilizatorului îi sunt oferite sugestii de resurse care sunt similare cu resursele curente.

- salvează resurse pentru liste organizate. Utilizatorul are posibilitatea de a salva resursele atât de la pagina de căutare și rezultatele de la pagina de înregistrare în vederea creării propriilor liste personalizate. Listele pot fi recuperate în orice moment și vor fi mereu disponibile pentru utilizator. Acest lucru ajută la eliminarea nevoii de gestionarea citatelor pe desktop care tinde să fie prea dificil pentru utilizatorii de bază.

- căutarea de resurse. Utilizatorul are posibilitatea de a naviga prin catalogul bibliotecii ceea ce le permite să exploreze colecțiile bibliotecii, decât doar să vadă un spectru foarte îngust de rezultate.

- biografii de autor. Utilizatorul poate afla mai multe despre un autor, cu informații contextuale și a poate vedea toate cărțile pe care acesta le-a scris și care se află în bibliotecă.

- URL-uri persistente. Permite utilizatorului să marcheze interogarea sau înregistrarea pentru a avea acces permanent la o pagină pe care au fost o dată

- compatibil Zotero (*Zotero = instrument de cercetare „next generation”*)

Utilizatorii pot salva și eticheta orice înregistrări cu Zotero astfel încât înregistrările să poată fie stocate într-un singur loc.

- internațional

Interfața are traduceri disponibile în portugheza braziliană, chineză, olandeză, engleză, franceză, germană, japoneză, spaniolă etc.

- acces de date: Open Search, OAI, Solr

VuFind are multe API-uri pentru a interacționa cu căutările, datele și multe alte caracteristici. Se pot reuni datele înregistrărilor cu alte instituții printr-un server OAI. Se pot face căutări folosind algoritmi VuFind prin OpenSearch și dacă se dorește acces complet la datele indexate, există posibilitatea interacționării cu Solr.

Interfața catalogului de bibliotecă VuFind este una foarte prietenoasă (fig. 2.5):

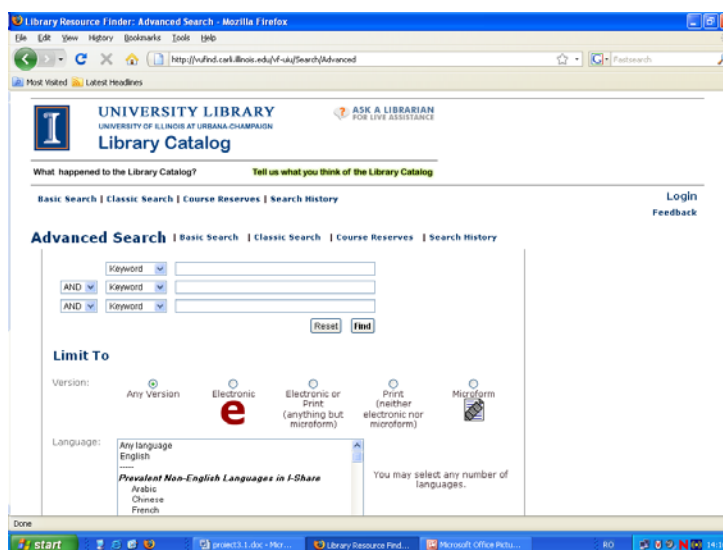


Fig. 2.5. Interfața catalogului VuFind

Câteva dintre caracteristicile VuFind:

- căsuță de căutare simplă
- căutare după titlu, subiect, limbă, format, etc.
- căutare în catalogul bibliotecii sau cataloagele comune odată, incluzând starea documentului dorit și informații de localizare
- înregistrarea cu propriul username și parolă
- conectarea cu propria pagină „My Account”
- salvarea și etichetarea înregistrări cu Zotero
- preyintă instrumente de rețele sociale, cum ar fi lăsarea de comentarii și marcare
- previzualizări în Google Book Search

Caracteristici ale VuFind indisponibile deocamdată:

- căutarea rezervării de cursuri
- căutarea după numărul de apel
- posibilitatea de a limita elementele de bibliotecă
- nu poate exporta direct către RefWorks

2.1.3. PRIMO

Primo este un produs nou și modern al firmei Ex Libris (furnizoare de soluții software și servicii complementare pentru biblioteci și centre de informare). <http://www.exlibrisgroup.com/category/PrimoOverviewferințe>. Se bazează pe o arhitectură deschisă și pe interoperabilitate, fiind proiectat să funcționeze cu sisteme standard integrate de bibliotecă (ILS) dar și cu alte aplicații de bibliotecă. Este construit să extragă datele din cataloage de bibliotecă (ILS), din arhive electronice și depozite digitale, din cataloage virtuale bazate pe SFX și METALIB (knowledge bases) (Fig. 2.6) www.exlibrisgroup.com/.../PrimoForYourLibrary

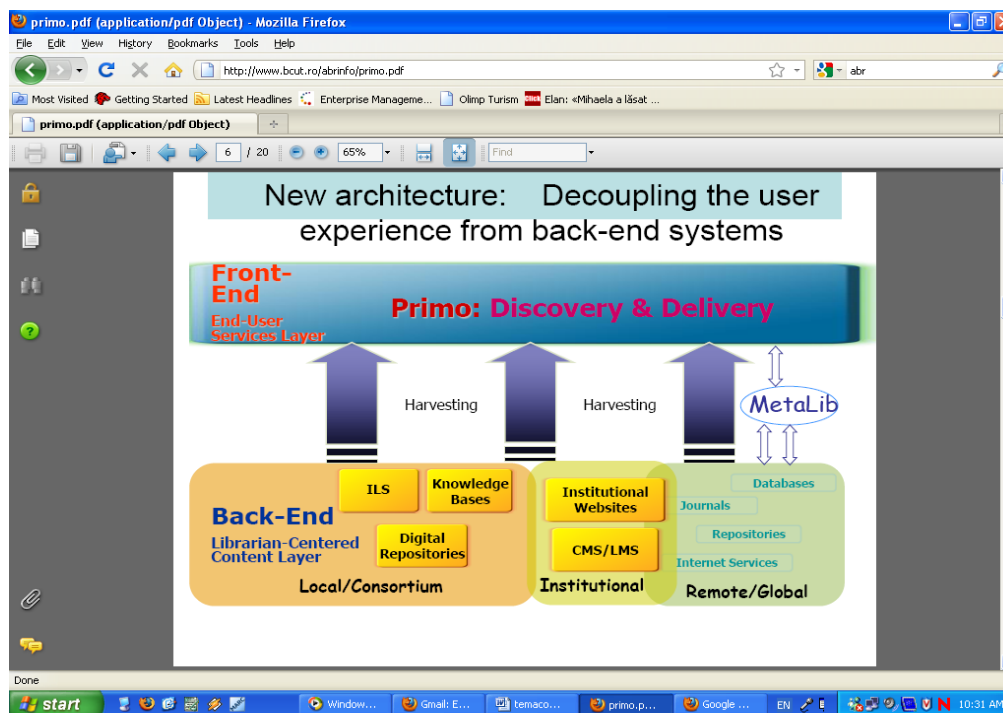


Fig. 2.6. – Arhitectura PRIMO

Primo este o soluție unitară (one-stop solution) pentru descoperirea și furnizarea informației din resurse locale sau din rețea (de la distanță), și care vine în întâmpinarea nevoilor cititorilor, a așteptărilor lor în regăsirea informației, făcând o punte între interfețele sistemelor de bibliotecă, care sunt considerate destul de greoaie, complicate și modul de căutare în Internet.

Primo permite bibliotecilor să-și expună colecțiile bogate (print, e-resurse, resurse digitale) atrăgând cititorii cu tehnicile noi, adaptate la cerințele de ultimă oră, în descoperirea și livrarea informațiilor.

Astăzi, biblioteca nu mai este apreciată atât pentru colecțiile sale, cât pentru tipul de acces pe care îl oferă. Utilizatorul are nevoie de servicii de bibliotecă independente de spațiu și timp, ceea ce face ca centrul de greutate să se mute dinspre bibliotecă spre utilizator. Cu *Primo*, utilizatorii pot accesa colecțiile bibliotecii de la distanță cu aceeași ușurință și viteză ca și resursele locale, instituționale.

Primo oferă utilizatorului o soluție rapidă, care simplifică procesul de căutare și care localizează rapid informația exactă și de calitate, sortând-o după criteriile de relevanță (fig.2.7).

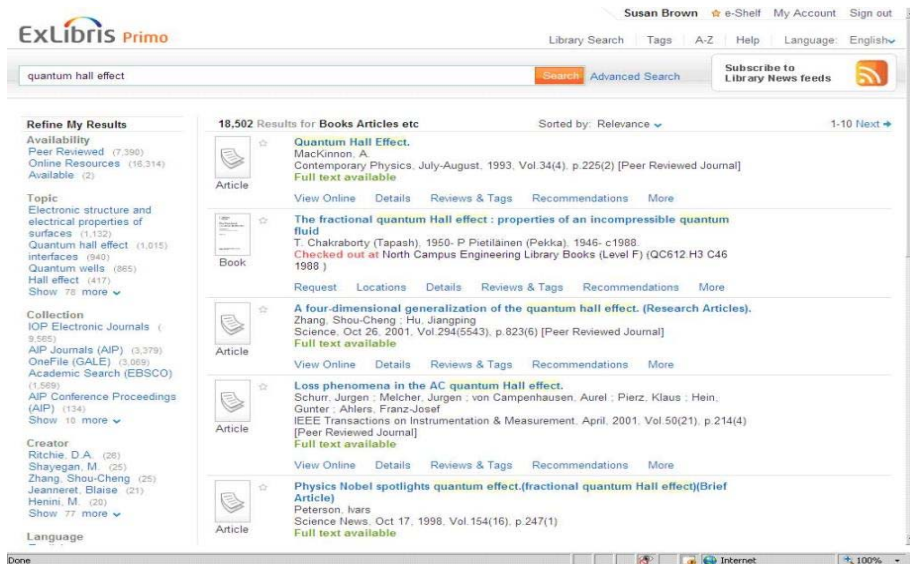


Fig.2.7 –Afișarea rezultatelor căutării în Primo

Tendința actuală este de a face cataloagele de bibliotecă să semene cât mai mult cu Google și Amazon. Primo, este un exemplu în acest sens.

Dintre caracteristicile lui Primo pot fi enunțate:

- renunțarea la duplicarea înregistrărilor bibliografice prin utilizarea unuia dintre principiile FRBR, anume principiul colocării
- interfețele de bibliotecă sunt restructurate și dezvoltate în sensul dorințelor utilizatorului
- regulile de catalogare, atât descriptivă cât și tematică sunt rafinate ; terminologia privind indexarea subiectelor este și ea rafinată
- utilizează mult mai mult full text-ul și se bazează pe furnizare de servicii
- serviciile *Primo* pot fi înglobate în locații utilizate în mod obișnuit: site-uri Web, portaluri instituționale, bloguri, rețele de socializare, dispozitive mobile de tip telefon mobil (mobil web interface)
- cu *Primo*, utilizatorii pot accesa colecțiile bibliotecii de la distanță

- cu *Primo*, utilizatorii pot căpăta informații referitoare la disponibilitatea documentelor, pot plasa o cerere de ILL, pot salva interogările și transmite prin E-mail sau le pot salva într-o listă e-Shelf
- nu în ultimul rând, PRIMO permite bibliotecilor să-și expună întreaga gamă de resurse de informare în același timp (de la cărți și reviste tipărite la e-books, e-journals, obiecte digitale) ceea ce conduce la creșterea gradului de utilizare a acestora.
- *Primo* reduce decalajul dintre sistemele de bibliotecă actuale și așteptările utilizatorilor

Primo permite bibliotecilor să-și expună întreaga sferă a colecțiilor ajutând utilizatorii să descopere resurse ale bibliotecii neexploatate încă. Rezultatul este o mai mare utilizare și dezvoltarea continuă a colecțiilor bibliotecii bazate pe nevoile utilizatorilor.

2.1.4. ENCORE

ENCORE este o interfață nouă, mai prietenoasă, de căutare în catalogul unei biblioteci. El utilizează ca și criteriu de căutare unul sau mai multe cuvinte cheie (din acest punct de vedere este similar cu Google) (fig.2.8.).

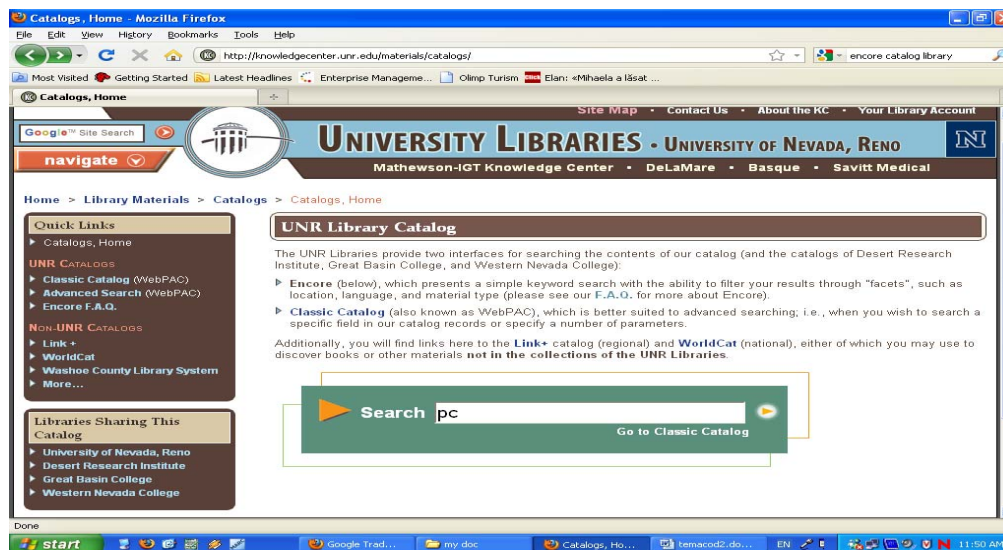


Fig.2.8. – Buton căutare în ENCORE

Ceea ce particularizează **ENCORE** în raport cu alte interfețe este aceea că toate operațiile de căutare încep cu o căutare de bază, după cuvinte cheie (fig.2.8) **ENCORE** afișează un singur buton de căutare de tip „search” (fig.2.8), funcționând ca un motor de căutare în catalogul bibliotecii.

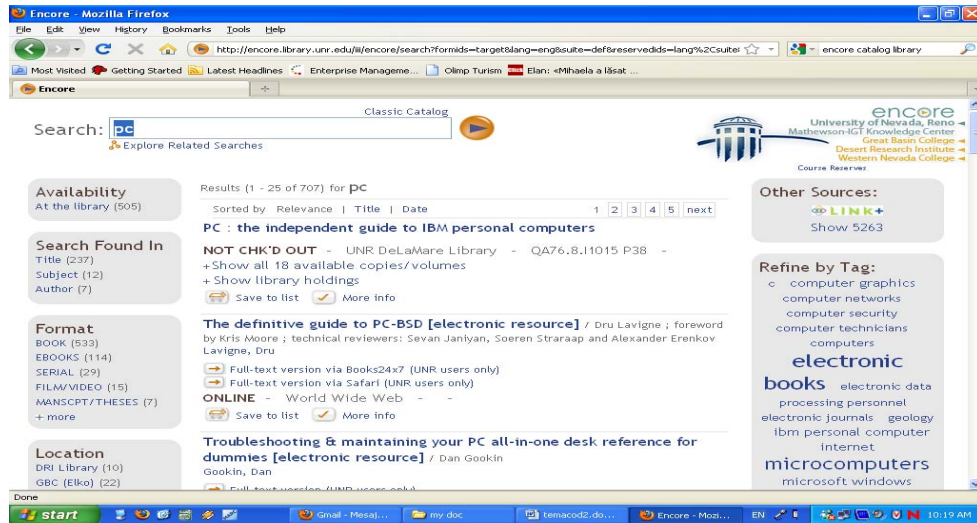


Fig.2.9. Afișarea rezultatelor căutării după cuvânt cheie

ENCORE afișează o listă de lucrări aflate în relație cu termenul după care s-a efectuat căutarea, oferind apoi posibilitatea de sortare/rafinare a căutării prin utilizarea **fațetelor** și **tagurilor** mai sus amintite. **ENCORE** răsfoiește într-o varietate de colecții incluzând cărți, reviste, material audio-video, e-books, arhive digitale, ș.a.. Aria de căutare poate fi restrânsă prin accesarea fațetelor referitoare la format, limbă, loc de publicare, dată de publicare, locație, sau prin adăugarea mai multor cuvinte cheie în câmpul de căutare „search” (fig.2.10)

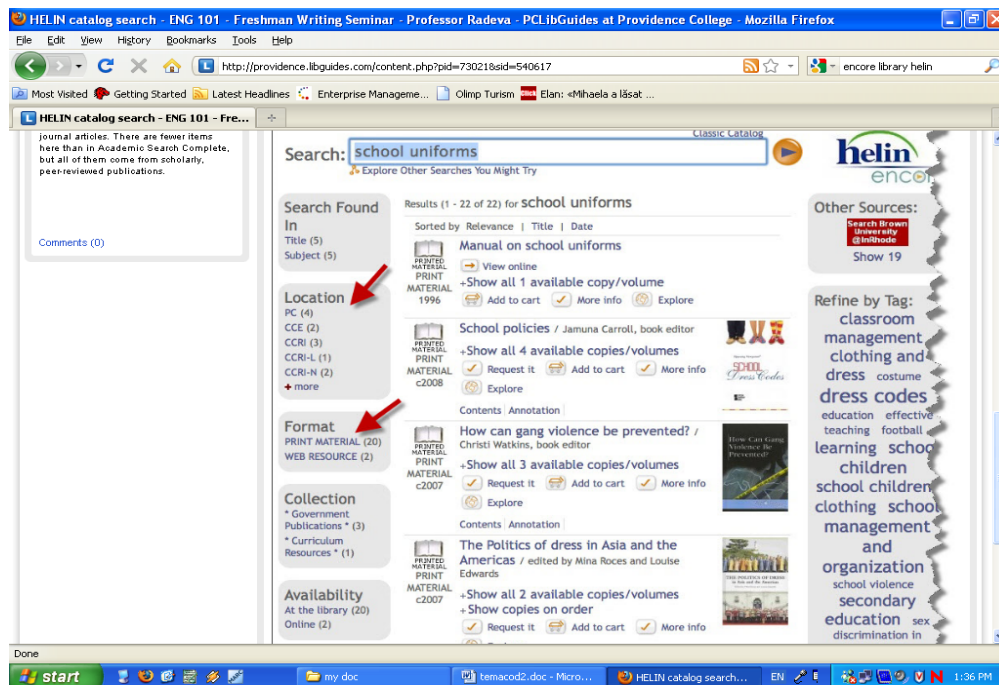


Fig.2.10. Restrângerea ariei de căutare prin utilizarea fațetelor

Rolul acestor fațete și taguri este tocmai limitarea, restrângerea, numărului de înregistrări găsite, permițându-i astfel utilizatorului să se concentreze asupra celor mai relevante.

ENCORE nu înlocuiește vechiul catalog, el se constituie într-o interfață nouă de căutare; vechiul catalog apare cu denumirea „Classic Catalog”, „Classic Search”. Utilizatorul poate opta pentru utilizarea ambelor opțiuni: ENCORE sau „classic search”.

- 1) providence.libguides.com/content.php?pid=73021&sid=540617
- 2) knowledgecenter.unr.edu/encore/Default.htm
- 3) www.library.nashville.org/.../gen_highlight_encore.asp - Statele Unite ale Americii -

2.2. Alte exemple de posibile cataloage „Next generation”

Cataloage de bibliotecă unice:

- [Santa Clara County](#) (folosește interfața Bibliocommons cu rețele sociale)
- [North Carolina State University Library Catalog](#)
- [E41ST](#)
- [Aquabrowser at Queens Library](#)

- [Lamson Library](#)
- [Lexington Public Library Catalogue](#) Aquabrowser
- [UCLA Film & Television Archive catalog](#)
- [Hennepin County Library's Bookspace](#)
- [Danbury Public Library](#) integrează “alte ediții și traduceri”, tag-uri, și “cărți similare” de la Biblioteca Thing

- [Blacklight @ University of Virginia](#) (experimental)
- [McMaster University](#) (utilizează a doua casetă de căutare)
- [Phoenix Public Library](#)
- [Nelsonville Public Library](#) Athens County (utilizează Koha ILS open-source)
- [Georgia Pines](#) (Un site al consorțiului bibliotecilor publice care utilizează ILS Evergreen open-source. Utilizează fațete create din rezultate pentru prezentarea de noi căutări, în loc de rafinarea de căutării actuale)
- Portalul [MyResearch al Universității Villanova](#)
- [Galileo Digital Library of Georgia](#), University of Georgia

Cataloage comune:

- [WorldCat.org](#)
- [National Library of Australia Lucene Demonstrator](#)
- [Florida Center for Library Automation \(FCLA\)](#)
- [LIBRIS: catalogul comun suedez al bibliotecilor academice și de cercetare](#)

Alte cataloage / site-uri comerciale fațetate:

- [Google Books](#)
- [Amazon - Books](#)
- [Open Library](#), un proiect al Internet Archive
- [WalMart Books](#)
- [Chapters](#) Canadian bookstore.

3. Prelucrarea resurselor tehnice în format tradițional cât și în alte formate decât cel hârtie

3.1. Aspecte tehnice privind digitizarea documentelor tradiționale

Procesul de digitizare este un proces complex. Pentru ca el să se desfășoare cât mai eficient și să dea rezultate concrete, trebuie analizate cât mai amănunțit toate aspectele tehnice legate de acest proces.

Procesul de digitizare presupune copierea documentului tradițional original prin intermediul scannerului sau al aparatului foto digital, prelucrarea și distribuirea documentului obținut prin intermediul calculatorului utilizând programe specifice (programe pentru redactare texte, prelucrare imagini, recunoaștere optică a caracterelor, generatoare de pagini web etc.).

Digitizarea trebuie să aibă loc în timp util și trebuie să asigure conservarea și diseminarea rapidă și eficientă a documentelor digitizate, de aceea este absolut necesar să fie ales un calculator cât mai performant. Pentru a aprecia performanțele unui sistem de calcul trebuie să se țină cont de viteza de execuție, fiabilitate, tipul arhitecturii, capacitatea memoriei și timpul de acces la memorie.

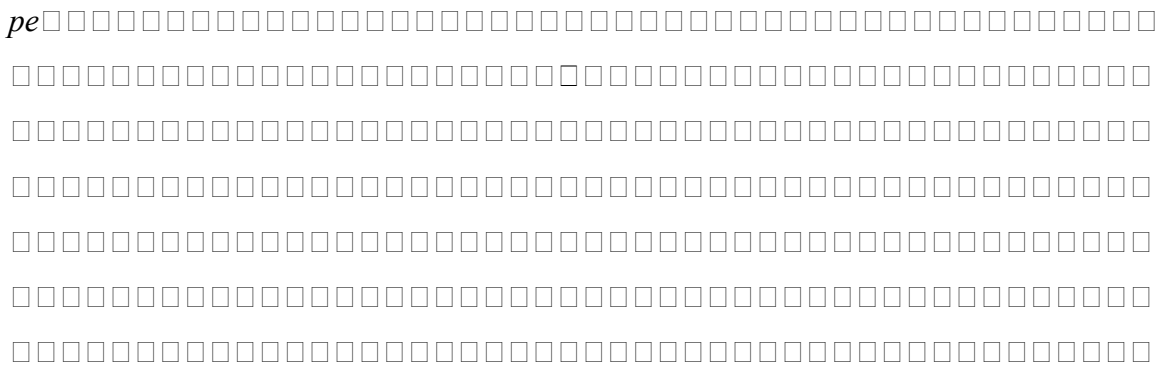
Recent au fost dezvoltate noi tehnici de digitizare, care combină în mod optim robotica, electronica digitală și tehnica de calcul. Folosind aceste tehnici s-a creat o linie de digitizare care oferă în premieră o soluție complet integrată. *Digitizing Line* (linia de digitizare) este un sistem de scanare de carte automat conceput de către firma elvețiană 4Digitalbooks. *Digitizing line* are încorporat un leagăn automat al cărții, astfel asigurând o poziționare optimă a cărții pentru a putea fi scanate foile. De asemenea, *Digitizing Line* deține și un detector automat al parametrilor cărții. *Digitizing line* este dotat cu un scanner puternic, oferind o calitate foarte bună a scanării prin mișcarea traversei luminoase. *Digitizing Line* dispune de un sistem automat de întoarcere a paginilor, acesta acționând prin vacuum, pentru a păstra integritatea cărților. *Digitizing Line* folosește lumină rece pentru scanare. Pentru o mai bună conservare a documentului scanat, viteza brațului care se ocupă cu întoarcerea paginilor poate fi ajustată în funcție de preferințe. Tot pentru a asigura integritatea documentului, *Digitizing Line* este dotat cu *bookcradle* automat, astfel cotorul cărților păstrându-se intact. Acest sistem de scanare folosește

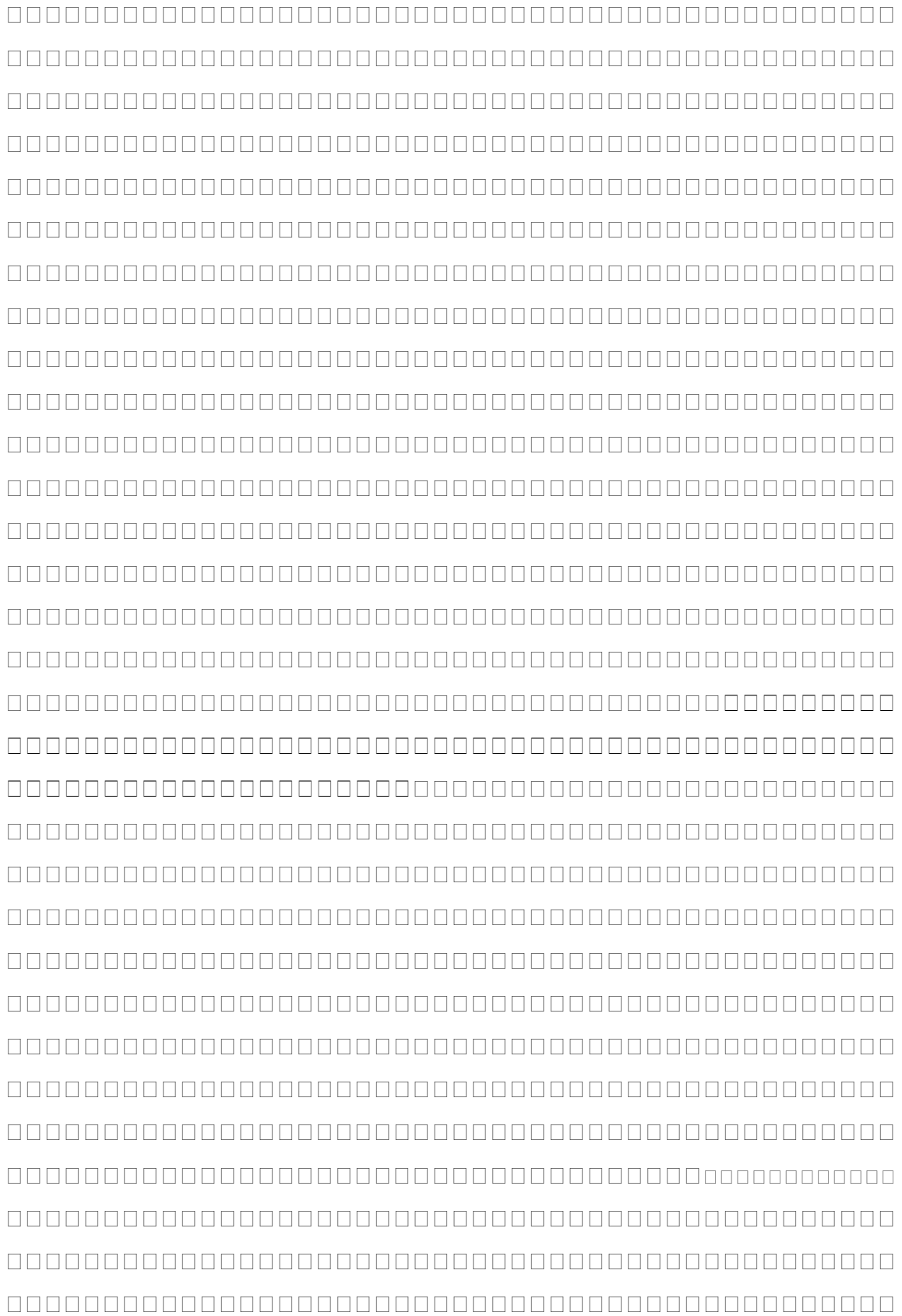
lumină LED, fiind recomandat pentru scanarea de documente vechi. Scannerul poate recunoaște optic caracterele. Digitizing Line folosește camere digitale pentru capturarea imaginilor. Acest model are integrat un sistem care detectează eventuale defecte ale paginii (*page bending detector*), cum ar fi paginile îndoite. Un avantaj al liniei de digitizare constă în gradul foarte ridicat al productivității. Alt avantaj important este menținerea integrității fizice a cărții supuse procesului de digitizare. Documentele sunt manipulate, transferate și digitizate automat, astfel asigurându-se prezervarea informației, precum și a formei originale a documentului. Vorbind despre sistemul *Digitizing Line*, Michael Kellner, director al Bibliotecii Universității din Stanford, explică: „Un avantaj este acela că documentele tradiționale nu părăsesc niciodată mâinile noastre. Nu trebuie să le retragem din circuit șase sau chiar 10 săptămâni. Dacă pentru procesarea unei cărți de 300 pagini sunt necesare 20 de minute, atunci un raft pe zi e o nimica toată. Produsul final e sub controlul nostru.”(DigiBook, nr. 8/2004, pag.4)

BOPCRIS este un proiect al Bibliotecii Universității din Southampton care are ca activitate principală scanarea documentelor parlamentare, rare și istorice, folosind *Digitizing Line*. Simon Brackenbury, director de proiecte de digitizare la Southampton, afirmă:

„Într-un anume fel, acest proiect este un studiu de caz în automatizare. Linia de digitizare are o productivitate ridicată (scanează 600 de pagini / oră, propunându-ne să scanăm 1 000 000 într-un an), permițându-ne scanarea unui număr mare de documente, ceea ce va face posibil accesul la documentele istorice” (inaccesibile utilizatorilor până în acel moment)” (DigiBook, nr. 9/2004, pag.1).

„Facilitățile de care dispune această linie de digitizare permite ca milioanele de pagini ale documentelor istorice tipărite să devină accesibile ca informații full-text





- programe speciale pentru cărți digitale (Calibre-E-book management, DeskTop Author, Flipalbum, Glogster, Kerpoof, KooBits, Book Restorer, etc.).

. Însă, în componența software-ului intră, pe lângă programele propriu-zise, și materialul auxiliar, cum ar fi datele grafice.

3.2.Aspecte privind digitizarea documentelor audio și video

3.2.1. Tehnici de digitizare ale documentelor în format audio și video

Audio

Ceea ce noi numim „sunet” reprezintă, de fapt, o perturbație (energie mecanică) propagată prin mediu (aer) sub forma unor unde. Când undele sonore ajung la timpan, nervii auditivi sunt stimulați și trimit impulsuri electrice către creier, acesta realizând o interpretare a informației primite. În mod asemănător funcționează, spre exemplu, un microfon: mica membrană a microfonului, de care este fixată o bobină, este pusă în vibrație de undele sonore care se ciocnesc de ea și, datorită unui magnet permanent aflat în apropiere, se induce curent electric alternativ în bobină, „interpretat” mai apoi de un dispozitiv fie de amplificare, fie de înregistrare.

Înregistrarea digitală constă în următoarele etape:

- Semnalul analog (curentul sinusoidal) este transmis de către dispozitivul de captare al sunetului către un convertor analog→digital (en. „**ADC**” – *analog to digital converter*);
- ADC-ul transformă semnalul primit într-un șir de biți (1 și 0), prin procedeul de **eșantionare** – transformarea semnalului continuu în eșantioane. Numărul de eșantioane „ produse” pe secundă poartă numele de **SAMPLE RATE** (rata de eșantionare).
- Informația este apoi stocată, spre exemplu, pe hard-disk sau CD-ROM.

Pentru a reda sunetul, se parcurg pașii în sens invers:

- Secvența binară este trimisă unei unități „**DAC**” (en. *digital to analog converter*), ce transformă biții înapoi, în semnal analog.
- Semnalul este eventual amplificat pentru a fi trimis apoi spre difuzoare.

Salvarea sunetelor într-un calculator se bazează pe procedeul de eșantionare a semnalului audio, de reținere a valorilor amplitudinii (voltajului), la anumite intervale de timp. Numărul de biți folosit în formatul audio pentru reprezentarea binară a valorii unei măsurări de amplitudine se numește **rezoluție**, sau **adâncime de biți** (en. resolution, bits per sample, bit depth). Intervalele de timp la care se măsoară această amplitudine poartă

numele de perioadă, ajungându-se astfel la noțiunea de **frecvență** (en. sample rate) – numărul de măsurări într-un interval de o secundă.

Datele pot fi stocate necomprimate sau comprimate, pentru a se reduce dimensiunea fișierului.

Bit Depth & Sample Rate

Calitatea unei secvențe audio este determinată de doi factori:

- De câte ori pe secundă este măsurată amplitudinea semnalului analog (**SAMPLE RATE**);
- Intervalul numeric în care se află datele fiecărei măsurători (**BIT DEPTH**).

Primul factor se măsoară în kHz, kilohertzi. Spre exemplu, CD-urile audio sunt înregistrate la o rată de 44.1 kHz, adică fiecare secundă de audio s-a constituit prin măsurarea de 44,100 de ori a amplitudinii semnalului.

Într-o secvență audio nu poate fi captat un sunet ce are o înălțime mai mare decât jumătate din valoarea „sample rate” a secvenței respective. Prin urmare, pentru o valoare de 44.1 kHz, nu pot fi captate sunete ce depășesc frecvența de 22.05 kHz.

Experții spun că majoritatea oamenilor nu pot auzi oricum peste pragul de 20 kHz și de aceea s-a stabilit valoarea de 44.1 kHz pentru CD-urile audio.

Cel de-al doilea factor, „bit depth”, determină intervalul valorilor numerice ce sunt asociate fiecărei citiri de amplitudine. Bineînțeles, o creștere a „bit depth”-ului implică o reprezentare digitală a sunetului de o mai mare acuratețe și finețe. Și, fiindcă urechea umană este sensibilă la mici variații ale intensității sunetului, este de dorit un interval cât mai mare de reprezentare. Pentru CD-urile audio, intervalul este cuprins între 0 și 65,535, „bit depth” având valoarea 16 (16 biți = 2 octeți).

Iată o comparație în funcție de valorile factorilor „Sample rate” și „bit depth” (tabel 3.1.):

Tabel 3.1. comparație în funcție de valorile factorilor „Sample rate” și „bit depth

| | Sample rate | Bit depth | Avantaje | Dezavantaje |
|------------|-------------|-----------|--|---|
| minimum | 44.1 kHz | 16 biți | <p>Nu este necesară conversia pentru „arderea” secvenței pe un CD audio.</p> <p>Ocupă mai puțin spațiu per secundă înregistrată.</p> <p>Necesită puțin timp de procesare (ex. la redare).</p> <p>Reprezintă standardul internațional pentru compact-discuri (așa-zisul Red Book Standard).</p> | <p>Calitatea oferită ar putea fi insuficientă pentru formate audio viitoare.</p> |
| recomandat | 44.1 kHz | 24 biți | <p>O reprezentare de o mai mare acuratețe a sursei sonore originale.</p> <p>„Dinamicitate” sonoră sporită.</p> <p>Reprezintă standardul profesioniștilor în domeniul audio.</p> | |
| optimum | 96 kHz | 24 biți | <p>Standardul pentru DVD/HD Audio.</p> <p>Domeniu al frecvențelor reprezentabile mult mai mare.</p> <p>O calitate superioară, alegerea ideală pentru înregistrări muzicale.</p> <p>Mai mult potențial pentru aplicații viitoare.</p> <p>Reflectă noile tendințe pentru profesioniștii în domeniul audio.</p> | <p>Ocupă mult spațiu per secundă înregistrată.</p> <p>Timp de procesare mare.</p> <p>Pentru unele secvențe audio nu prezintă îmbunătățiri perceptibile.</p> <p>Necesită conversie pentru folosirea pe un CD Audio standard.</p> |

Următorul tabel prezintă dimensiuni ale unor fișiere ce conțin **O ORĂ** de sunet în format digital, pentru valori diferite ale ratei de eșantionare (sample rate) și rezoluției de biți (bit depth) (Tabel 3.2.)

Tabel 3.2. Valori ale ratei de eșantionare

| Sample Rate | Bit depth | Număr de canale | Mărime, în MB |
|-------------|-----------|-----------------|---------------|
| 44.1 kHz | 16 | 2 (stereo) | 591 MB |
| 44.1 kHz | 16 | 1 (mono) | 296 MB |
| 44.1 kHz | 24 | 2 (stereo) | 887 MB |
| 44.1 kHz | 24 | 1 (mono) | 444 MB |
| 96 kHz | 24 | 2 (stereo) | 1931 MB |
| 96 kHz | 24 | 1 (mono) | 966 MB |

Codec-uri

Un format de fișier audio este un format dedicat stocării datelor audio pe un sistem de calcul. Un astfel de format se poate compune dintr-un *flux brut de biți*, dar, în general, constă într-un **format-container** cu o structură fixă. Acesta conține, în afara datelor audio propriu-zise și o serie de informații necesare (ex. **un antet, metadata**), dispuse sub aspectul unui strat ce „învelește” datele.

Este important să știm să facem diferența între un format de fișiere și un codec. Un **codec** este un mic program ce se ocupă cu compresia și decompresia fluxului brut de biți, în timp ce **formatul de fișiere** audio este containerul în care se află „depozitat” acest flux de biți. Cuvântul codec s-a format prin legarea celor două cuvinte-cheie: **compress-decompress**. Iată definiția codecului, preluată de pe Wikipedia:

Un codec este un program sau o bibliotecă de software, eventual chiar și un aparat hardware corespunzător, care asigură codarea și decodarea unei informații. Termenul este folosit în general referitor la mediile audio sau video analoage, care trebuie transformate într-un format digital. Cu această ocazie datele rezultate sunt de obicei și

comprimate, de exemplu pentru a accelera transmiterea lor prin rețeaua de telefonie mobilă sau prin Internet. Prin extensie se numesc tot "codec" și unele programe care transformă informația deja digitalizată dintr-un format în altul ("transcodare").

Majoritatea formatelor audio suportă multiple codecuri. Să zicem că un fișier cu formatul audio XYZ are extensia „.xyz”. Vom crea două astfel de fișiere ce conțin exact același sunet. Dacă nu ar fi aplicat niciun codec, atunci datele din ambele fișiere ar fi perfect identice. Însă un fișier îl salvăm folosind codecul A, iar pe celălalt folosind codecul B. Fiecare fișier va conține la o anumită adresă numele codecului folosit, iar datele nu vor mai fi aceleași, ele fiind salvate prin algoritmi diferiți. Mărimea celor două fișiere nu va mai fi nici ea aceeași, depinzând de gradul de comprimare al fiecărui codec. Cu toate că datele sunt diferite, informația (sunetul) rămâne, de obicei, neschimbată. Când vom rula un player audio pentru redarea secvenței înregistrate, acesta „observă” codecul specificat și va importa proceduri din librăriile dinamice ale fiecărui codec (un exemplu ar fi fișierele „DLL”, pentru platforma Windows) pentru interpretarea datelor din fișierele noastre. În cazul în care codecul specificat nu este instalat pe calculator, atunci datele nu vor putea fi citite, deci sunetul nu va fi redat.

3.2.2 Arhitectura procesului de digitizare

O privire de ansamblu asupra unui proces pentru translatarea datelor in formatele digitale dorite este dat in figura urmatoare (fig. 3.1):

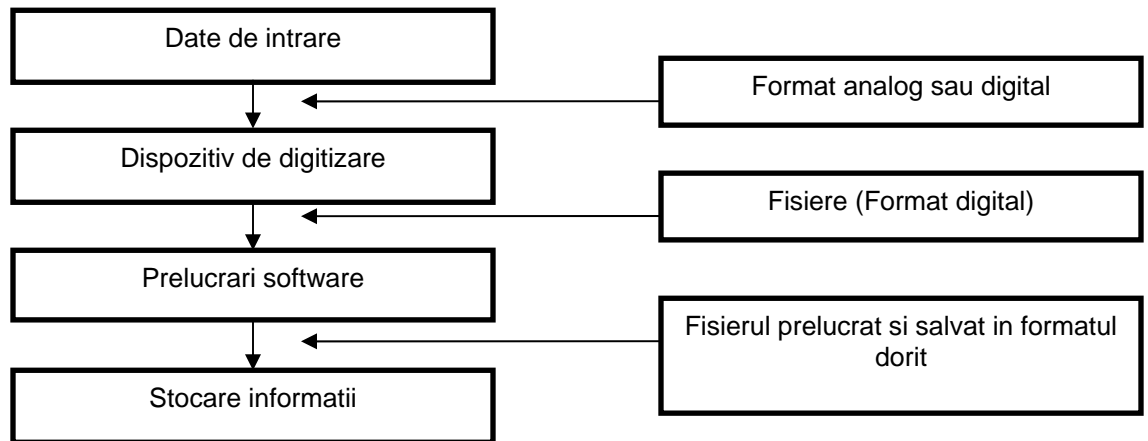


Fig3.1. Procesul de translatare a datelor în formate digitale

In figură avem:

- Date de intrare - Casete (MiniDV, HDV) conținând material audio-video înregistrat folosind o camera audio-video;
- Dispozitivele de digitizare (DV-CAM, HDV) folosite pentru citirea informațiilor de pe dispozitivele de intrare;
- Prelucrarea software – captura materialului într-un format digital, editarea acestuia și salvarea produsului final;
- Stocarea de date, salvarea de meta-informații în baza de date, transmiterea fișierelor către servere (pentru mirroring), etc.

3.2.3 Formate de intrare

DV

DV este un format pentru înregistrarea și redarea de materiale audio-video digitale . Acesta a fost lansat în 1995, cu eforturi comune ale principalilor producători de camere video. DV este un sistem de compresie un intraframe, care utilizează transformata cosinus (DCT) pentru a comprima video frame cu frame. Sunetul este stocat necomprimat.

Urmand indeaproape standardul ITU-R Rec.601, DV foloseste scanare întretesuta (interlaced) cu frecvența de eșantionare a luminanței de 13,5 MHz. Acest lucru duce la 480 scanlines pe un cadru complet pentru sistemul de 60 Hz, și 576 scanlines per cadru complet pentru sistemul de 50 Hz. În ambele sisteme zona activă conține 720 pixeli per scanline, cu 704 pixeli utilizati pentru conținut și 16 pixeli pe partea stângă pentru decupari digitale. Aceeasi dimensiune este folosita raporturile 4:3 si 16:9, rezultând în diferite raporturi de aspect ale pixelilor pentru fullscreen și widescreen video. Sunetul poate fi stocat in oricare dintre cele două forme: 16-bit Linear PCM stereo la 48 kHz rata de eșantionare (768 kbit / s pe canal, 1,5 Mbit / s stereo), sau patru canale PCM neliniare pe 12 biti la rata de eșantionare 32 kHz (384 kbit / s pe canal, 1,5 Mbit / s pentru patru canale). În plus, specificatiile DV suporta de asemenea audio pe 16 biti la 44,1 kHz (706 kbit / s pe canal, 1.4 Mbit / s stereo), aceeași rata de eșantionare folosita pentru CD-ul audio. În practică, de 48 kHz stereo modul este folosit aproape exclusiv.

Digital Interface Format

Audio, video, și metadatele sunt ambalate în blocuri de 80 bytes DIF (Digital Interface Format), care sunt multiplexate într-o secvență de 150 blocuri. Blocurile DIF sunt unitățile de bază ale fluxurilor DV și pot fi stocate ca fișiere în formă brută sau în formate de fișiere, cum ar fi AVI , QuickTime și MXF . Un cadru video este format fie din 10 sau 12 astfel de secvențe, în funcție de rata de scanare, ceea ce duce la o rată de date de aproximativ 25 Mbit / s pentru video, și o suplimentare de 1.5 Mbit / s pentru audio.

Sony și Panasonic au creat versiunile lor proprietare de DV, care utilizează aceeași schema de compresie, dar îmbunătățită la robustețe, capabilități de editare liniară, redarea culorilor și mărimea raster. Toate variantele DV, cu excepția DVCPRO Progressive sunt înregistrate pe bandă în flux video întrețesut. Sunetul poate fi desincronizat cu videoul cu +/- o treime de frame.

DVCPRO

DVCPRO, de asemenea, cunoscut sub numele de DVCPRO25, este o variație de DV dezvoltat de Panasonic și introdus în 1995 pentru utilizarea în colectarea de știri electronice (ENG). Spre deosebire de DV, DVCPRO folosește audio blocat și subeșantionare de croma 4:1:1 pentru variantele 50 Hz și 60 Hz în scopul reducerii pierderilor de generare. Sunetul este disponibil doar în varianta 16-bit/48 kHz. Când este înregistrat pe caseta, DVCPRO folosește track-pitch mai larg - 18 μm vs 10 μm al DV de bază, ceea ce reduce șansele de erori atunci când materialul este înregistrat.

DVCAM

În 1996, Sony a răspuns cu propria sa versiune profesională de DV numită DVCAM. Ca și DVCPRO, DVCAM folosește audio blocat audio, care împiedică desincronizarea audio care se poate întâmpla pe DV în cazul în care mai multe generații de copii sunt făcute. Când înregistrate pe bandă, DVCAM folosește track-pitch de 15 μm, care este cu 50% mai mare față de valoarea originală (DV). În consecință, banda este transportată cu 50% mai rapid, ceea ce reduce timpul de înregistrare cu o treime în comparație cu DV.

DVCPRO50

DVCPRO50 a fost introdus în 1997 de Panasonic pentru ENG și cinema digital, și este adesea descris ca fiind două DV-codec-uri care lucrează în paralel. DVCPRO50 dublează rata de date video codificate la 50 Mbit / s, înjumătățind timpul de înregistrare față de DVCPRO. Rezoluția cromă este îmbunătățită prin utilizarea eșantionării cromă de 4:2:2. Formatele similare Sony includ Digital Betacam lansat în 1993 și MPEG IMX din 2001.

DVCPRO Progressive

DVCPRO progresivă a fost introdus de Panasonic pentru colectarea de știri, sport, jurnalism și cinema digital. Aceasta a oferit 480 sau 576 de linii de înregistrare cu scanare progresivă, având subeșantionare chroma de 4:2:0 și patru canale audio PCM de 16-bit / 48 kHz. Ca HDV-SD , a fost menit ca un format intermediar în timpul perioadei de tranziție de la definiție standard la video de înaltă definiție. Formatul oferea șase moduri de înregistrare și redare: 16:9 progresivă (50 Mbit / s), 4:3 progresivă (50 Mbit / s), 16:9 interlaced (50 Mbit / s), 4:3 interlaced (50 Mbit / s), 16:9 interlaced (25 Mbit / s), 4:3 interlaced (25 Mbit / s). Formatul a fost înlocuit cu DVCPRO HD.

DVCPRO HD

DVCPRO HD, de asemenea, cunoscut ca DVCPRO100 este un format de înaltă definiție care poate fi gândit ca patru DV codec-uri care funcționează în paralel. Rata de date video depinde de rata de cadre și poate fi de 40 Mbit / s pentru 24 frame / s și de 100 Mbit / s pentru 50/60 cadru / s. La fel ca DVCPRO50, DVCPRO HD are rata de esantionare a culorii de 4:2:2. DVCPRO HD folosește raster de dimensiune mai mică decât televiziunea de înaltă definiție: 960x720 pixeli pentru 720p, 1280x1080 pentru 1080/59.94i și 1440x1080 pentru 1080/50i. Sub-eșantionare orizontală similară este utilizată în multe alte formate HD. Pentru a menține compatibilitatea cu HDSDI, echipamentul DVCPRO100 supra-esantionează video-ul în timpul redării.

HDV

HDV este un format pentru înregistrarea video high-definition pe caseta DV. Formatul a fost inițial dezvoltat de JVC și susținut de către Sony , Canon și Sharp . Cele patru companii au format consorțiul HDV în septembrie 2003. Văzut ca un format de înaltă definiție pentru camere video digitale accesibil ca pret, HDV a devenit repede adoptat de cameramani amatori și profesioniști din cauza costului scăzut, portabilității și calității imaginii acceptabile pentru multe producții profesionale.

Codare video si audio

HDV video și audio sunt codate în format digital, folosind compresie lossy. Video-ul este codat cu codecul MPEG-2, folosind chroma de 8 biti și probe în cazul componentei luma cu subeșantionare cromă de 4:2:0. Sunetul stereo este codat cu codecul MPEG-1 Layer 2. Audio și video compresate sunt multiplexate într-un flux de transport MPEG, care este de obicei înregistrat pe bandă magnetică, dar poate fi de asemenea stocat într-un fișier de calculator.

Rata de date atât pentru audio și video este constantă și este aproximativ aceeași ca rata de date DV. Rata de date relativ scăzută poate cauza bit rate starvation în scene care au o mulțime de detalii fine, mișcare rapidă sau alte activități complexe, cum ar fi lumini intermitente, și poate avea ca rezultat artefacte vizibile, cum ar fi blockiness și blurring. În contrast cu video, bitrate-ul sunetului HDV este relativ generos. La un bitrate codat pe 384kbit / s, sunetul MPEG-1 Layer 2 este considerată ca fiind fără pierderi perceptuale;

Două versiuni majore ale HDV sunt HDV 720p și HDV 1080i. Prima este utilizată de JVC și este cunoscută și ca HDV1. A doua este preferată de Sony și Canon și este uneori menționată ca HDV2. HDV 1080i definește opțional moduri de înregistrare progresivă.

HDV 720p

Formatul HDV 720p permite înregistrare video de înaltă definiție (HDV-HD), precum și scanare progresivă pentru video standard de înaltă definiție (HDV-SD). HDV-HD urmărește scanarea progresivă 720p în ceea ce privește tipul de scanare, dimensiunea cadrului, raportul de aspect și rata de date. Camerele video HDV 720p anterioare puteau filma doar la 24, 25 și 30 de cadre pe secundă. Modelele următoare oferă atât rate de cadru (framerate) film-like (24p, 25p, 30p) cât și reality-like (50p, 60p). HDV-SD este un modul pentru înregistrarea progresivă-scanare de video standard - definiție. Un astfel de video este uneori numit video enhanced definition video (EDTV), dar este considerat video de înaltă definiție în Australia. În funcție de regiune, HDV-SD video sunt înregistrate fie ca 576p50 sau 480p60. La fel ca DVCPRO Progressive, HDV-SD a fost menit ca un format intermediar în timpul perioadei de tranziție de la definiție standard la

video de înaltă definiție. Modelele de camere video recente HDV 720p nu înregistrează în acest mod.

HDV 1080i

Sony a adaptat HDV, inițial conceput ca format cu scanare progresiva de JVC, pentru video întretesut. Acesta a fost un compromis util pentru zeci de ani, datorită capacității sale de a afișa mișcare fără probleme reducând în același timp de latimea de banda a transmisiei și înregistrarea. Video-ul întretesut este încă folosit în achiziționare și radiodifuziune, dar dispozitivele de afișare întretesut sunt eliminate. Ecranele plate de televiziune care utilizează plasma și tehnologie LCD, sunt în mod inerent progresive. Toate monitoarele moderne folosesc de asemenea scanare progresivă. Înainte de afișarea unui video întretesut pe un dispozitiv de scanare progresivă, acesta trebuie să fie convertit la progresiv prin intermediul procesului cunoscut sub numele de deinterlacing .

Unele camere video HDV 1080i sunt capabile să înregistreze video progresiv într-un flux întretesut, cu condiția ca rata de cadru nu depășește jumătate din rata de teren (field rate) . Prima camera video HDV 1080i care a implementat astfel de scanare progresivă a fost Sony HVR-V1.

Scanarea progresivă video trebuie să fie corect de-întretesută pentru a obține rezoluție maximă verticală și pentru a evita artefacte interlace. Video 25P și 30P trebuie să fie de-întretesut cu weave sau algoritmul no deinterlacing, ceea ce înseamnă asamblarea a două câmpuri ale fiecărei cadru împreună într-un cadru progresiv. Această operație se poate face în cele mai multe instrumente de editare, pur și simplu prin schimbarea proprietatilor de proiect din întretesut în progresiv. Video 24P trebuie să treacă prin film-mode deinterlacing, de asemenea, cunoscut sub numele de inverse telecine.

3.2.4. Dispozitive hardware

Captura video este doar un proces de convertire a semnalului video analog de la un dispozitiv cum ar fi o camera video într-un format digital și apoi stocarea datelor video digitale pe un computer. Captura video de la dispozitive analogice necesită o placă specială de captură video care convertește semnalele analogice în format digital și

comprima datele. **Plăcile de captura video** permit înregistrarea de semnale video de la camere video sau VCR pe hard-diskul computerului. Aceste plăci folosesc compresie (codec) hardware și/sau software pentru a digitaliza fișierele video pe hard-disk. În timp ce programul de editare video va permite să creați și să editați în format video, placa de captura video este cea care determină calitatea înregistrării și implicit a produsului final. Există de asemenea dispozitive digitale video care pot captura imagini și apoi să le transfere într-un computer cu ajutorul unei interfețe standard serial/paralel, USB sau chiar FIREWIRE.

Înainte de achiziționarea unui sistem hardware dedicat digitizării informației audio/video, câteva aspecte ar trebui luate în considerare:

- tipul sistemului de calcul (PC,Mac)
- tipul intrării (DV-Firewire, analog sau amândouă)
- suport de DVD Authoring
- software de editare inclus
- posibilitate a prelucrărilor în timp real
- alte facilități de performanță specifice
- cost

O lista a celor mai populare dispozitive hardware de captură video este prezentată în tabelul următor(Tabel3)

Tabelul 3. Dispozitive hardware de captură video

| Dispozitiv | PC | Mac | FireWire | Analog | DVD | Software de Editare | captură de timp real | Playback | Plug-in-uri | Pret aprox. US\$ |
|------------------------------|----|-----|----------|--------|-------|------------------------------|----------------------|----------|-------------|------------------|
| Pinnacle Studio AV version 8 | Y | | Y | | | Pinnacle Studio 8 | | | Y | 70-90 |
| Pinnacle Studio DV version 8 | Y | | Y | | | Pinnacle Studio 8 | | | Y | 100-125 |
| Pinnacle Studio Deluxe | Y | | Y | Y | Y | Pinnacle Studio 8 | | | Y | 270-340 |
| Pinnacle DV-200 | Y | | Y | | | Adobe Premiere 6 | | Y | Y, a lot | 240-300 |
| Pinnacle DV-500 Plus | Y | | Y | Y | Y | Adobe Premiere 6 | Y | Y | Y, a lot | 450-550 |
| Pinnacle DV-500 DVD | Y | | Y | Y | Y | Adobe Premiere 6 | Y | Y | Y, a lot | 580-660 |
| Pinnacle Pro-One | Y | | Y | Y | Y | Adobe Premiere 6 | Y | Y | Y, a lot | 750-900 |
| Matrox RT2500 | Y | | Y | Y | Y | Adobe Premiere 6 | Y | Y | Y, a lot | 700-820 |
| Matrox RTMac | | Y | Y | Y | | | Y | Y | | 540-600 |
| ADS Pyro Basic DV | Y | Y | Y | | Trial | Ulead VideoStudio 6 SE Basic | | | | 45-65 |
| ADS Pyro DV | Y | | Y | | Y | Ulead Video | | | | 60-80 |

| | | | | | | | | | | |
|-------------------------------------|---|---|---|---|---|---|--|--|-------------|---------|
| | | | | | | Studio 6 | | | | |
| ADS Pyro Platinum DV | Y | Y | Y | | Y | Adobe Premiere 6 | | | Y, a lot | 250-320 |
| ADS Pyro Platinum SE | Y | | Y | | Y | Adobe Premiere 6 LE | | | Y | 80-110 |
| ADS Pyro ProDV | Y | | Y | | | Ulead MediaStudio Pro 6 | | | Y, a lot | 180-220 |
| ADS USB Instant Video | Y | Y | | Y | | Ulead VideoStudio 5 for PC, MovieWorks LE for Mac | | | | 60-80 |
| ADS USB Instant DVD | Y | | | Y | Y | Ulead VideoStudio 6 | | | | 160-200 |
| Digital Origin MotoDV | Y | Y | Y | | | Premiere 6 LE, EditDV Unplugged | | | | 300-360 |
| Digital Origin EditDV | Y | | Y | | | EditDV 2.0 | | | Y, a lot | 400-480 |
| Dazzle DV Editor SE | Y | | Y | | Y | Dazzle MovieStar 5 | | | | 40-60 |
| Dazzle DV Editor for Notebook | Y | | Y | | | Ulead VideoStudio | | | | 70-90 |
| Dazzle Digital Video Creator | Y | | Y | | | MGI VideoWave4 | | | | 50-70 |

| | | | | | | | | | | |
|--------------------------------------|---|---|---|---|---------|--------------------------|---|---|----------|-----------|
| DVC80 | | | | | | | | | | |
| Dazzle Digital Video Creator 2 DVCII | Y | | | Y | Y | Dazzle MovieStar | | | Y | 240-290 |
| Dazzle DVnow AV | Y | | Y | Y | Y | Adobe Premiere 6 | | Y | Y | 430-500 |
| Dazzle Hollywood DV Bridge | Y | Y | Y | Y | PC only | MovieStar 5 PC only | | | | 240-280 |
| Canopus DV Converter | Y | Y | Y | Y | | | | | | 300-360 |
| Canopus DVStorm SE | Y | | Y | Y | Y | StormEdit | Y | Y | Y, a lot | 1280-1400 |
| Aist MovieDV Suite | Y | | Y | | Y | Aist MovieDV 4 | | | Y, a lot | 95-120 |
| Aist MoviePack VE | Y | | Y | | Y | Aist MoviePack VE 4 | Y | | Y, a lot | 380-450 |
| Aist MoviePack Pro | Y | | Y | | Y | Aist MoviePack Pro4 | Y | | Y, a lot | 680-780 |
| Aist MoviePack eXtreme | Y | | Y | | Y | Aist MoviePack eXtreme 4 | Y | | Y, a lot | 1200-1350 |

3.2.5. Software de prelucrare video

Tabel 4. Caracteristici principale ale programelor de editare

| Software | HD | Editare nedistructivă | Playback fullscreen | Mod Storyboard | Piste Video | Piste Audio / nr maxim canale | Afișarea marcajelor de timp |
|-------------------------|-----|-----------------------|--------------------------|----------------------------|--------------------------------|-------------------------------|-----------------------------|
| Adobe Premiere Pro | Yes | Yes | Yes, on separate monitor | Yes | 99 | 99 | Yes |
| Avid Xpress Pro | Yes | Yes | Yes | Yes | 24 tracks, unlimited layers | 24 tracks, unlimited layers | Yes |
| Cyberlink PowerDirector | Yes | Yes | Yes | Yes | 10 | 4 | Yes |
| FinalCut Studio | Yes | Yes | Yes | Yes (in the Media Browser) | 99 per sequence | 99/2 per sequence | No |
| Sony Vegas | Yes | Yes | Yes | No | Unlimited | Unlimited / 16 | Yes |
| Pinnacle Studio | Yes | No | Yes | Yes | 2 (3 via glitching the system) | 4 | Yes |
| Ulead MediaStudio Pro | Yes | Yes | Yes, on separate monitor | N/A | 99 | ? | Yes |
| VirtualDub | N/A | Yes | No | No | 1 | 1 | Yes |
| Windows Movie Maker | Yes | No | Yes | Yes | 1 | 2 | Yes |

Tabelul următor indică posibilitatea fiecărui program de editare video de a importa video HD pentru editare:

Tabel 5: Posibilitatea de import a video

| Software | 720p HDV | 1080i HDV | DVCPRO HD | HD necomprimat | IMX | AVCHD | XDCAM HD | HDV | AVC-1 |
|-------------------------|----------|-----------|-----------------|----------------|-----------------|----------------|----------------|-----|-------|
| Adobe Premiere Pro | Yes | Yes | Yes | Yes | Yes (3rd party) | Yes (from CS4) | Yes (from CS3) | Yes | N/A |
| Avid Xpress Pro | Yes | Yes | Yes | No | No | Yes | Yes | Yes | Yes |
| Cyberlink PowerDirector | Yes | Yes | ? | Yes | ? | Yes | ? | Yes | Yes |
| FinalCut Studio | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Sony Vegas | Yes | Yes | Yes (3rd party) | Yes | Yes | Yes | Yes | Yes | ? |
| Pinnacle Studio | Yes | Yes | No | No | No | Yes | ? | ? | ? |
| Ulead MediaStudio Pro | Yes | Yes | ? | Yes | No | ? | ? | ? | ? |
| VirtualDub | ? | ? | ? | ? | ? | ? | ? | ? | ? |
| Windows Movie Maker | Yes | Yes | No | No | No | Yes | ? | Yes | ? |

3.2.6. Formate de stocare

Formate de stocare audio:

Există trei mari grupe de formate-container audio:

- Formate ce conțin (în general) **date audio necomprimate** (en. uncompressed audio formats) – cum ar fi WAV, AIFF, AU;
- Formate ce conțin **date audio comprimate**, însă **fără pierdere de informație** (en. **lossless** compression) – cum ar fi FLAC, Monkey's audio (extensia .ape), WavPack (extensia .wv), Windows Media Audio Lossless (WMA Lossless).
- Formate ce conțin **date audio comprimate**, cu **pierdere de date și informație** (en. **lossy**)– MP3, Vorbis (extensia .ogg), AAC, Windows Media Audio (WMA).

Formate fără compresie

Există un format principal de reprezentare necomprimată a sunetului, și anume PCM (Pulse-code modulation – „împărțirea” curentului sinusoidal pe eșantioane dreptunghiulare, un mod de aproximare fidelă a sunetului), folosit în fișierele WAV (platforma Windows) și AIFF (pentru Mac OS). Formatele WAV și AIFF sunt deosebit de flexibile și sunt concepute astfel încât să poată stoca flux de date la diverse combinații de rezoluție și rate de eșantionare (en. bit rate, sample rate). Sunt folosite pentru stocarea înregistrărilor originale.

Formatul AIFF (Audio Interchange File Format) se bazează pe formatul IFF (Interchange File Format) dezvoltat de Apple. Formatul WAV se bazează pe RIFF (Resource Interchange File Format), replica Microsoft la IFF-ul celor de la Apple. BWF (Broadcast Wave Format) este un format audio standard creat de European Broadcasting Union, ca un succesori al lui WAV. BWF permite metadatelor (metadata = date despre date) să fie stocate în fișier. Este foarte utilizat de către profesioniștii în domeniul audio, în televiziune și industria cinematografică. Fișierele conțin și o linie temporală ce permite sincronizarea perfectă a sunetului cu anumite imagini, etc.

Formate cu compresie, fără pierdere de informație

O secvență de date dintr-un fișier cu un astfel de format necesită mai mult timp de procesare înainte de a fi încărcat în memoria principală, dar ocupă mai puțin spațiu pe hard-disk.

Formatele fără compresie rețin același număr de biți pentru reprezentarea fiecărei sample (eșantion) astfel încât, spre exemplu, un minut de liniște deplină rezultă într-un fișier de aceeași mărime cu un fișier ce reține un minut de muzică orchestrală. În cazul comprimării fără pierdere de informații, lucrurile stau altfel. Minutul de muzică va ocupa ceva mai puțin spațiu decât inițial (necomprimat), iar minutul de liniște aproape că nu va ocupa deloc spațiu. Formatele comprimate fără pierdere de informații asigură o rată de compresie de aproximativ 50%. Dezvoltarea algoritmilor de comprimare are ca scop reducerea timpului de procesare, menținând totuși o rată bună de compresie.

Formate cu compresie, cu pierdere de informație

Codarea cu pierdere de informație se folosește pentru a reduce cantitatea de date, eventual pe seama calității. Intenția este de a elimina numai informațiile pe care omul obișnuit nu le percepe (de exemplu detaliile ultrafine ale unei interpretări muzicale, sunetele foarte înalte sau cele cu o amplitudine foarte mică). Pe măsură ce se dorește obținerea unei cantități din ce în ce mai mici a materialului codat, pierderea de informație poate deveni totuși vizibilă / audibilă și chiar deranjantă (de exemplu, în cazul codării unei imagini, cerul cu nori poate fi înlocuit pur și simplu printr-o pată albastră uniformă).

Fișierele cu o astfel de compresie pot duce la o dimensiune a datelor de 5-20% din fluxul de date audio inițial. Exemple: MP3, OGG, WMA.

Formate audio și codec-urile asociate lor

- **WAV** – formatul standard pentru platforma Windows. În general, folosește stocarea necomprimată, PCM, ceea ce înseamnă că fișierele WAV pot avea dimensiuni mari, de aproape 10 MB pe minut. Pot fi totuși folosite o varietate de codec-uri pentru reducerea dimensiunii. Formatul WAV este un „descendent” al formatului RIFF.
- **OGG** – un format gratis, cu sursa deschisă editării publice (en. open-source), ce suportă o mulțime de codec-uri, cel mai folosit fiind codec-ul Vorbis, care oferă o

compresie similară celei oferite de MP3, dar nu este chiar atât de răspândit ca acesta.

- **MPC** – Musepack (denumit inițial MPEG+; extensii: .mpp, .mpc) este un format cu sursa deschisă editării publice, ce oferă o compresie cu pierdere de informație. Musepack și Vorbis sunt considerate de mulți cele mai bune codec-uri existente, oferind o bună compresie și o calitate foarte mare a sunetului. Cu toate acestea, Musepack e foarte puțin popular, fiind folosit în principal de audiofili.
- **FLAC** – (Free Lossless Audio Codec) - Specificațiile acestui format (codec-ul implicit) pot fi găsite la adresa <http://flac.sourceforge.net/format.html>.
- **AIFF** – formatul standard pentru platforma Mac OS.
- **AU** – formatul standard folosit de Sun, Unix și Java. Fluxul audio poate fi PCM (necomprimat) sau comprimat cu codecurile „μ-law”, „a-law” sau G729.
- **MIDI** (Musical Instrument Digital Interface) – protocolul standard de comunicare dintre instrumentele muzicale electronice și computer. Un fișier MIDI conține mai multe piste (en. tracks) paralele de date ce rețin, fiecare, înălțimea, intensitatea, durata fiecărei note muzicale transmise de un instrument electronic muzical. Spre exemplu, un fișier MIDI ce „înregistrează” un duet claviatură-tobe electronice va avea două piste. De reținut este faptul că MIDI nu reține semnale audio, ci doar caracteristici ale notelor.
- **GSM** (codec) – creat pentru a fi folosit în telefonie (comunicații) în Europa, și este foarte practic pentru că reprezintă un bun compromis între dimensiunea datelor și calitatea sunetului. Fluxurile audio din fișierele WAV pot fi comprimate cu codecul GSM.
- **VOX** – acest format folosește în general codecul ADPCM (Adaptive Differential Pulse Code Modulation) pentru stocarea fluxului de date. Fișierele VOX se aseamănă cu cele WAV, însă nu conțin un antet cu informații despre rata de eșantionare sau compresia folosită, ele trebuind specificate înainte de redarea fișierului. Dacă nu sunt specificate, se consideră că fișierul este „codat” cu ADPCM, iar rata lui de eșantionare este de 8000 Hz.

- **AAC** (Advanced Audio Coding) – acest format se bazează pe standardele MPEG2 și MPEG4. Se regăsește în fișiere cu extensiile .acc, .m4a, .mp4, .3gp, etc.
- **MP3** (MPEG-1 Layer 3) – cel mai popular format de stocare și răspândire a muzicii, dezvoltat de Institutul Fraunhofer. Eliminând date redundante din fluxul de date (sunetele greu de perceput de urechea umană), poate determina o compresie de aproape 90% a unui fișier WAV de exemplu (în format PCM), menținând în același timp o calitate foarte bună.
- **WMA** (Windows Media Audio) – format cu compresie, deținut de Microsoft.
- **RA** (Real Audio) – format conceput pentru transmiterea fluxurilor audio prin Internet (poate fi redat în timpul descărcării – en. streaming).

Formate de stocare Video

Standardele MPEG

MPEG este acronimul pentru **Moving Picture Experts Group** (grupul experților în filme) creat de organizația internațională pentru standardizare (ISO). Grupul a fost format pentru a stabili standarde de compresie și transmisie audio și video. Prima întâlnire a grupului a avut loc în orașul Ottawa din Canada în mai 1988.

Metodologia de compresie MPEG este considerată asimetrică deoarece codorul este mai complex decât decodorul. Codorul trebuie să fie algoritmic sau adaptabil pe când decodorul execută acțiuni fixe. Acest fapt este considerat un avantaj în aplicațiile de difuzare în masă unde numărul de codoare scumpe este mic dar numărul decodoarelor ieftine este mare.

Standardele MPEG sunt compuse din mai multe părți. Fiecare parte acoperă un anumit aspect al întregii specificații. Standardele specifică de asemenea și profilele și nivelele. Profilele sunt destinate să definească un set de instrumente care sunt valabile, iar nivelele definesc gama de valori potrivite pentru proprietățile acestora. MPEG a standardizat următoarele formate de compresie și standarde auxiliare:

- **MPEG-1**: Acest standard a apărut în anul 1993. Utilizat la codarea filmului și a sunetului asociat pentru medii digitale de stocare cu viteze până la 1,5 Mbit/s. Este primul standard pentru compresie audio și video. Principiul de proiectare a fost

permiterea codării filmului și a sunetului asociat la rata de biți a unui disc compact (CD). Pentru aceasta MPEG-1 reduce rata semnalului de transmisie pentru imagini, dar utilizează și o frecvență a pozelor de doar 24-30 Hz, rezultând o calitate moderată. Include popularul format de compresie audio MP3.

- **MPEG-2:** Standardul a fost lansat în anul 1995. Era utilizat pentru codarea generică a filmelor și a informațiilor audio asociate. MPEG-2 suportă întrepătrundere și o înaltă definiție. Standardul este considerat important deoarece a fost ales ca și schemă de compresie pentru televiziunile digitale ATSC, DVB și ISDB, servicii digitale prin satelit precum Dish Network, SVCD, DVD și Blu-ray.
- **MPEG-3:** Se ocupă de standardizarea compresării scalabile și de rezoluții multiple și a fost creat pentru compresie HDTV însă a fost descoperit că este redundant și a fost contopit cu MPEG-2, astfel că nu există standardul MPEG-3. A nu se confunda MPEG-3 cu MP3 care este stratul audio 3 al standardului MPEG-1.
- **MPEG-4:** A apărut în 1998. A fost creat pentru codarea obiectelor audio-vizuale. MPEG-4 folosește unelte de codare cu complexitate sporită pentru a atinge factori de compresie mai mari decât MPEG-2. Pe lângă codarea video mai eficientă, MPEG-4 se apropie mai mult de aplicațiile grafice. În profile mai complexe, decodorul MPEG-4 devine un eficient procesor de interpretare grafică, iar fluxul de date compresate descrie forme tri-dimensionale și texturi ale suprafețelor.

Formatul AVI

AVI, acronimul de la „**Audio Video Interleave**” (Interclasare audio video), este un format de stocare multimedia digital introdus de Microsoft în 1992 ca parte a tehnologiei „Video pentru Windows”. Fișierele AVI pot conține atât date audio cât și date video într-un fișier care permite redare sincronizată audio-video, însă nu e necesar ca acestea să apară împreună (exemplul unui clip fără sunet).

AVI este o derivare a formatului intitulat RIFF care împarte datele unui fișier în bucăți. Fiecare bucată este identificată de o etichetă FourCC(cod de patru caractere) . Un fișier AVI ia forma unui fișier de tip RIFF format dintr-o singură bucată care este divizată în două părți mai mici obligatorii și o parte opțională.

Prima parte obligatorie este identificată de eticheta „hdrl” și reprezintă antetul fișierului care conține informații legate de video precum înălțimea, lățimea și rata cadrelor. A doua parte obligatorie este identificată de eticheta „movi” și conține toate datele audio-video care alcătuiesc filmul. Partea opțională este identificată de eticheta „idx1” și indexează adresele relative ale bucăților de date din fișier.

Asemănător formatului RIFF, datele audio-video conținute în partea „movi” pot fi codate sau decodate de un software numit **codec**. La crearea fișierului, codecul translatează între datele neprelucrate și tipul de date compresate folosit în fișier. Un fișier AVI poate stoca datele audio-video în aproape toate formele de compresie, inclusiv : Full Frame (necompresat), Intel Real Time (Indeo), Cinepak, Motion JPEG, Editable MPEG, VDOWave, ClearVideo/RealVideo, QPEG și MPEG-4 video.

Formatul MOV

Formatul **Quicktime (.mov)** funcționează ca un recipient care conține una sau mai multe piste, fiecare conținând un singur tip de date: audio, video, efecte sau text. Fiecare pistă conține un flux de date codate digital (folosind un codec specific) sau o referință către fluxul de date aflat într-un alt fișier. Pistele sunt menținute într-o structură de date ierarhizată care este formată din obiecte numite atomi. Un atom poate fi părintele altor atomi sau poate conține date media, dar nu poate fi ambele.

Abilitatea de a conține referințe abstracte de date pentru datele media și separarea datelor media de listele de editare a pistelor duc la faptul că formatul Quicktime este printre cele mai potrivite pentru editare, fiind capabil să importe și să modifice datele pe loc fără necesitatea rescrierii întregului fișier.

Formatul OGG

OGG este un format standard gratuit, deschis dezvoltării publice, întreținut de fundația Xiph.Org. Creatorii formatului susțin că acesta nu este restricționat de brevete de software și este proiectat să producă un eficient flux de date și o manipulare eficientă a datelor multimedia digitale de calitate înaltă.

Numele „Ogg” se referă la formatul care poate multiplexa mai multe codecuri gratuite, deschise dezvoltării publice, pentru date audio, video, text și meta.

În cadrul arhitecturii multimedia Ogg, Theora furnizează stratul video, în timp ce codecul orientat spre muzică Vorbis acționează ca strat audio. Ca și strat audio mai pot fi utilizate și Speex, FLAC sau OggPCM.

Termenul „Ogg” este deseori folosit ca referință la formatul audio Ogg Vorbis care este de fapt audio codat folosind Vorbis în formatul de stocare Ogg. Anterior, extensia .ogg era folosită pentru orice conținut distribuit prin Ogg, însă începând cu anul 2007, fundația Xiph.Org a cerut ca extensia .ogg să fie utilizată doar pentru codecul Vorbis pe baza îngrijorărilor legate de compatibilitate inversă. Fundația a decis să creeze un nou set de extensii ale fișierelor și tipurilor media pentru a descrie diferitele tipuri de conținut precum .oga pentru fișiere cu conținut exclusiv audio, .ogv pentru video cu sau fără audio (inclusiv Theora), și .ogx pentru aplicații.

Formatul FLAC

FLAC este un format folosit pentru compresia datelor audio fără pierderea calității datelor originale necompresate. În timpul compresării, FLAC nu pierde din calitatea fluxului de date audio precum fac MP3, AAC și Vorbis. Principalul autor al acestui format este Josh Coalson.

FLAC reduce lățimea de bandă și cerințele de stocare fără sacrificarea integrității sursei audio. O înregistrare audio digitală (exemplu o melodie de pe un CD audio) codată în FLAC poate fi decodată (decompresată) într-o copie identică a sursei audio. Sursele audio codate în FLAC sunt, în majoritatea cazurilor, reduse la 50%-60% din mărimea lor originală.

FLAC este potrivit pentru redarea audio uzuală și pentru păstrare în timp îndelungat, cu suport pentru etichetare, poză de fațadă (de obicei pentru albume se folosește) și salt rapid în intervalul audio. Suportul de redare FLAC în sisteme audio portabile sau sisteme audio dedicate este limitat comparabil cu formate care pierd din calitate precum MP3, însă FLAC este suportat fără cerințe software de mai multe componente hardware decât alte formate asemănătoare lui precum Wavpack. Este posibilă redarea fișierelor FLAC pe un iPod, spre exemplu, folosind Rockbox.

Tabel 6. Comparație între formate video

| Numele și extensia, dacă diferă | AVI | Fișier video MPEG (.mpg .mpeg) | Quicktime (.mov, .qt) | OGG |
|---------------------------------|--|--------------------------------|--|--|
| Proprietarul | Microsoft | MPEG | Apple | Xiph.org, nepatentată |
| Suport pentru cadre B | Da | Da | Da | Da |
| Audio cu rată de biți variabilă | Da | Da | Da | Da |
| Rata cadrelor variabilă | Da | Da | Da | Da |
| Capitole | Da, prin modificări externe | Nu | Da | Da |
| Subtitrări | Da, prin modificări externe | Nu | Da | Ogg Writ, Ogg Kate |
| Formate video suportate | Aproape orice prin VFW (Video pentru Windows); H.264/AVC este problematic datorită suportului limitat de cadre B | MPEG-1, MPEG-2 | Limitat la ce este valabil pentru managerul de codec Quicktime | Theora, Dirac, OggUVS, MNG și altele, aproape orice prin VFW |
| Formate audio suportate | Aproape orice prin ACM (Manager de compresie audio); Vorbis este problematic | MPEG-1 straturile 1,2,3(mp3) | Limitat la ce este valabil pentru Sound Manager sau CoreAudio | Vorbis, FLAC, Speex, CELT, OggPCM, aproape orice prin ACM |
| Suport pentru etichete | Nu oficial | Nu | Da | CMML, Ogg Skeleton, comentariu Vorbis |

| | | | | |
|-----------------------------------|----|----|----|----|
| Suportă meniuri (ca DVD-urile) | Nu | Nu | Nu | Nu |
|-----------------------------------|----|----|----|----|

4. Realizarea modelului experimental al depozitelor cu conținut digital tehnic

4.1. Aplicație practică-conversia unei cărți vechi din format static jpeg în format pdf

4.1.1. Sistem de regăsire a informației în documentele digitizate în format static jpeg cu transformare în format pdf cu posibilitate de căutare după cuvânt în textul documentului

Aplicația practică constă în conversia unei cărți tehnice vechi din formatul JPEG în format PDF cu căutare după cuvânt cheie, folosind recunoașterea optică a caracterelor. Motivul principal pentru digitizarea documentelor tradiționale este acela de a-l face disponibil utilizatorilor într-un mediu online. În unele cazuri, digitizarea nu doar face mai accesibil conținutul documentului, ci chiar asigură supraviețuirea acestuia. Acesta este cazul documentelor vechi și rare în pericol de degradare. Digitizarea cărților tehnice vechi și rare este un proces foarte complicat și destul de costisitor. Vechimea documentului este un factor foarte important în digitizare. Pentru a digitiza o carte veche este recomandată folosirea unui scanner special destinat digitizării documentelor vechi și rare. Documentele cu conținut tehnic vechi și rare trebuie transportate și depozitate în cutii speciale pentru a nu se deteriora la transport. Trebuie evitată expunerea documentelor la lumină caldă, lumina de scanare recomandată este cea rece fără radiații UV și IR.. Dacă cartea veche și rară este scanată cu un scanner care nu dispune de opțiunea de recunoaștere optică a caracterelor, aceasta se va transforma în format digital format dintr-o serie de imagini (JPG, JPEG, GIF). Căutarea după cuvânt cheie nu este posibilă în documentele format JPG, JPEG sau GIF. De aceea, documentul digitizat trebuie convertit într-un format ce permite căutarea după cuvânt cheie (PDF cu căutare după cuvânt cheie, WORD etc.). Scopul aplicației de față constă în evidențierea pașilor care trebuie urmați pentru a obține un document digital care să păstreze atât aspectul fizic, cât și conținutul informațional al documentului original. Rezultatul final al aplicației practice este carte tehnică veche și rară în format digital (PDF) care permite căutarea după cuvânt cheie.

Catalog de carte veche și rară la Universitatea Transilvania din Brasov (fig.4.1.), disponibil la adresa: http://www.unitbv.ro/faculties/biblio/carte_veche/index.htm.

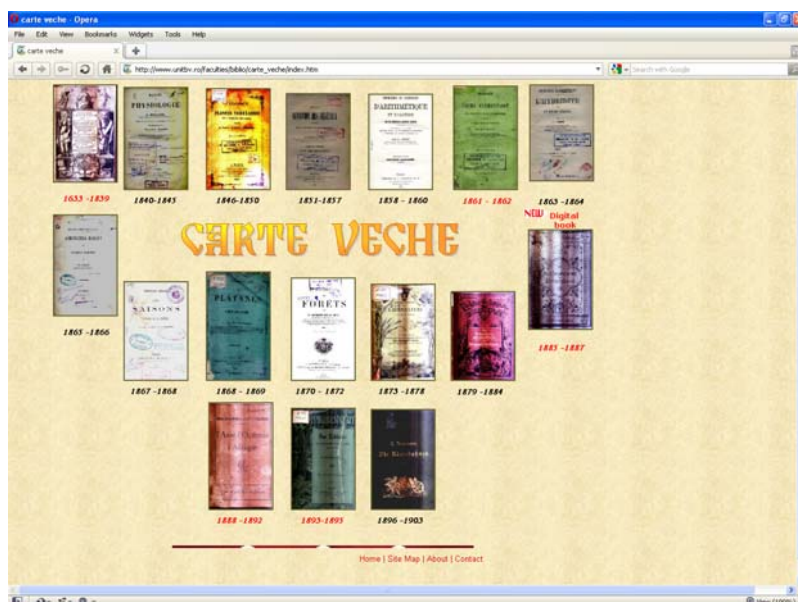


Fig. 4.1. – Catalog de carte veche și rară

Implementarea proiectului are la bază dorința de mărire a accesului la documentele bibliotecii, în special la cele vechi și rare. S-a urmărit reducerea considerabilă a manipulării și folosirii materialelor originale fragile sau uzate și crearea unor copii cât mai fidele pentru materialele aflate în pericol. De asemenea, s-a dorit dezvoltarea infrastructurii tehnice a bibliotecii și îmbunătățirea abilităților personalului acesteia. Un alt scop al proiectului este acela de a crea colecții virtuale și acces lărgit din dorința de a dezvolta resurse comune, în parteneriat cu alte instituții. Această formulă evită limitările de spațiu pe care le presupune un catalog tipărit, astfel încât utilizatorului i se furnizează o mulțime de informații despre fiecare carte.

Instrumentele folosite în cadrul proiectului sunt: un scanner cu alimentare automată a documentelor, un calculator pentru prelucrarea datelor, programele de digitizare și, bineînțeles, operatorul uman. Pentru realizarea proiectului au fost selectate în primul rând publicațiile care urmau să fie digitizate, apoi ca metodă de digitizare a fost aleasă scanarea pagină cu pagină a documentelor tradiționale, proces de conversie a imaginilor și textului din format grafic în format electronic. Linia de digitizare propusă în cadrul acestui proiect este formată dintr-un scanner cu alimentare automată a documentelor, un calculator pentru prelucrarea datelor, softul de digitizare și, bineînțeles, operatorul uman (fig.4.2). S-a utilizat scannerul COPYMATE 18 CONTEX. Scannerul

reprezintă un element foarte important în cadrul acestui proiect. Prin scanarea documentului original se obține o imagine raster. Imaginile scanate au fost salvate în fișiere diferite în format GIF sau JPEG și au fost prelucrate în programul de procesare grafică Adobe Photoshop. Pentru a obține imagini de calitate au fost folosite următoarele funcții ale programului Adobe Photoshop: funcția de corecție geometrică pentru a restaura imaginile la forma lor originală, funcția de corecție a luminozității pentru a uniformiza lumina pe toată suprafața documentului, funcția crop pentru a obține imagini tăiate la suprafețele dorite și funcția de conversie a documentelor color sau greyscale la alb-negru. Tot cu ajutorul programului Photoshop s-a rotit sau redimensionat imaginea.

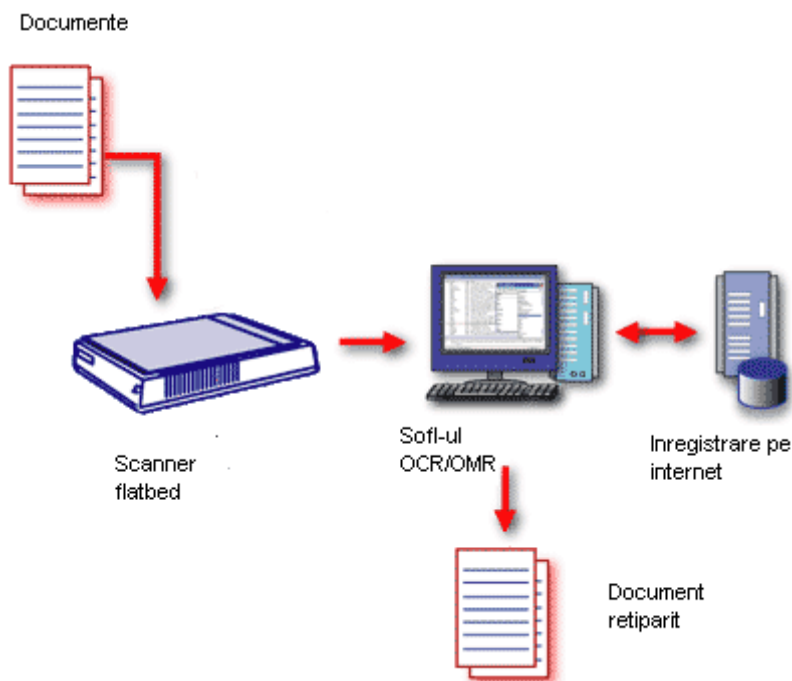


Fig.4.2 – Linia de digitizare folosită în cadrul proiectului Conservarea documentelor vechi și rare prin metode digitale. Linie de digitizare

Pentru formatul de stocare a imaginilor a fost ales DjVu datorită spațiului mic pe care îl ocupă, cărțile în acest format fiind de dimensiuni mult mai mici (de 3 până la 8 ori) decât echivalentul lor PDF. Formatul DjVu este recomandat pentru manipularea imaginilor scanate cu maximum 300 - 400 dpi. Prin DjVu imaginile sunt separate pe nivele (multicolor, rezoluție mică, rezoluție mare, background), fiecare nivel fiind comprimat la rândul lui, astfel fiind minimizate dimensiunile fișierelor. După salvarea cărților în format DjVu, s-a creat un site ce conține catalogul de carte veche, utilizându-se

limbajul HTML. Cărțile electronice din catalogul de carte veche sunt de fapt fișiere executabile care conțin paginile formatate în limbaj HTML și, de aceea, realizarea ei se aseamănă foarte mult cu crearea unui site web. Deci, pentru început trebuie realizate paginile respective în format HTML. Catalogul de carte veche a fost construit prin programul Microsoft Front Page, program de tip WYSWYG (What You See Is What You Get). Pentru fiecare pagină a cărții scanate s-a construit câte o pagină HTML. Apoi s-au creat butoanele de comandă (fig.4.3) ce permit navigarea de la o pagină la alta a cărții electronice cu ajutorul limbajului javascript. La deschiderea acestui fișier, se lansează fereastra cu o interfață de tip browser.

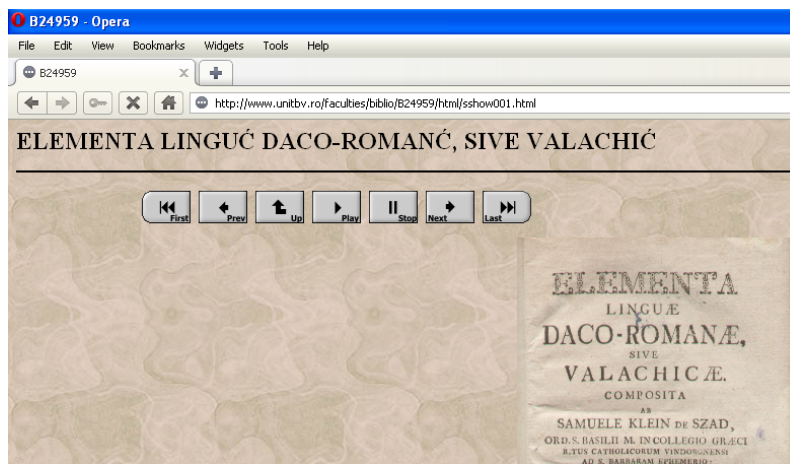


Fig.4.3. – Butoanele de comandă

Odată ce paginile web au fost create, ele au fost transformate ușor într-o carte electronică. Procesul de creare a unei cărți electronice din pagini web se numește compilare. Programul folosit în digitizarea cărților se numește DeskTop Author. Etapele procesului de digitizare sunt următoarele:

- crearea paginilor html folosind editorul disponibil;
- lansarea programului DeskTop Author și selectarea fișierelor pe care dorim să le includem în e-book;
- selectarea opțiunilor vizuale și de securitate;
- apăsarea butonului „compile” și încheierea procesului de compilare care de obicei durează mai mult de 10 minute.

Pentru a descrie procesul de digitizare, se folosește o schemă de digitizare ce reiese din figura 4.4.

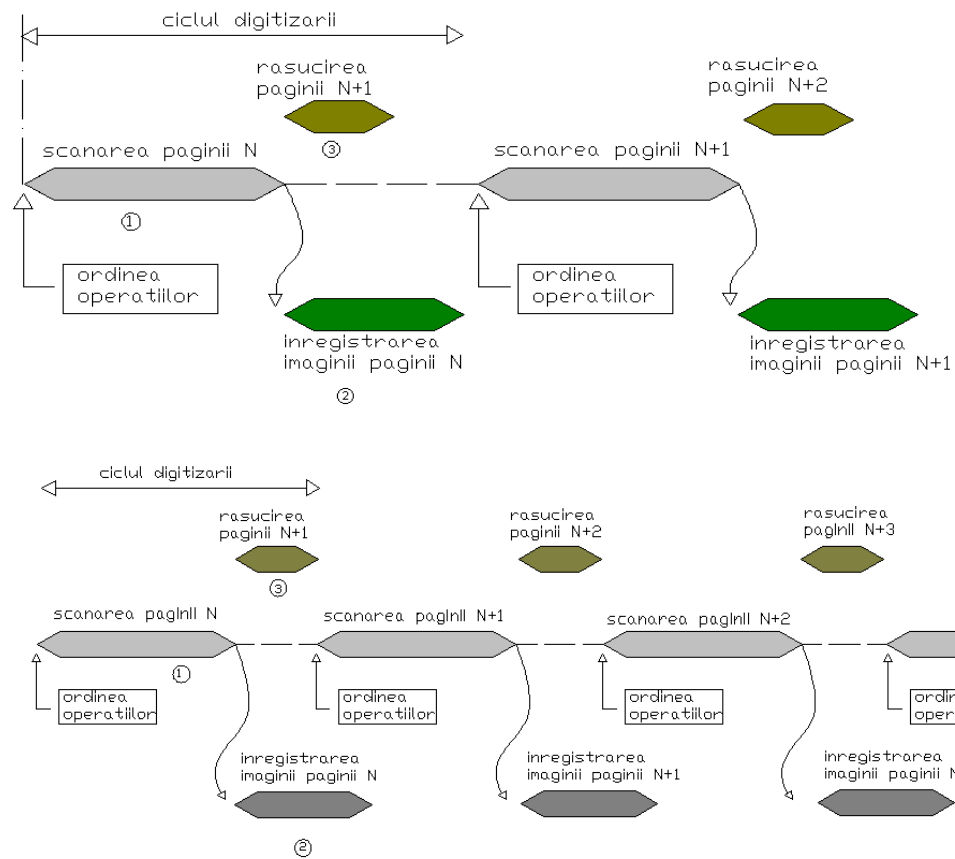


Fig.4.4 – Schema procesului de digitizare

Copiile digitale ale documentelor pot fi catalogate cu ajutorul metadatelor. Metadatele sunt date ce permit gestionarea și utilizarea altor date. Imaginile digitale sunt descrise cu ajutorul acestor metadate. Metadatele folosite în proiect au fost metadate adiționale necesare pentru a facilita navigarea și managementul fișierelor de date.

Prin proiectul de digitizare inițiat de Biblioteca Universității Transilvania se asigură conservarea cărților vechi și rare, precum și accesul rapid la informațiile conținute de acestea. Catalogul online de carte veche și rară oferă cititorilor posibilitatea de a consulta documente la care în mod normal nu ar avea acces. Un singur inconvenient al proiectului constă în faptul că nu se poate căuta după cuvânt cheie în documentele digitizate, formatul JPEG neavând această opțiune. Pentru a remedia acest inconvenient s-a folosit programul ABBY FineReader 9.0.

Un factor important care influențează calitatea cărții digitizate, constă în alegerea programelor cele mai eficiente. Pentru conversia documentului original, din format JPEG în format PDF, a fost nevoie de un program de recunoaștere optică a caracterelor de ultimă generație. ABBYY FineReader 9.0 este un sistem inteligent de recunoaștere optică a caracterelor. Cu ABBYY FineReader 9.0 se pot transforma documentele din format fizic, fotografiile digitale ale documentelor și fișierele PDF în formate editabile în care se pot identifica elemente, păstrându-le formatarea și structura. Acest program a fost ales, dintr-o serie de alte programe OCR, pentru digitizarea unui document tehnic vechi și rar. Comparat cu celelalte programe, ABBYY FineReader 9.0 s-a dovedit a fi cel mai eficient datorită acurateței și preciziei cu care recunoaște textul dintr-un document, redând cu succes și imaginile de calitate scăzută. Celelalte aplicații de recunoaștere optică a caracterelor analizate s-au limitat la procesarea paginilor documentului, neputând menține unui stil comun în toate paginile documentului. ABBYY FineReader 9.0 pune la dispoziția utilizatorului o serie de instrumente foarte utile pentru corectarea textului identificat prin care se poate asigura o corespondență de 100 % între original și textul recunoscut. Programul se bazează pe tehnologia ADRT, astfel el analizează documentul în ansamblu și îi poate reda structura logică. ABBYY FineReader 9.0 corectează automat neclaritățile și distorsiunile de imagine. Programul recunoaște 186 de limbi, 39 de limbi au suport de dicționar. De asemenea, programul deține funcții de procesare automată: îndreptarea rândurilor de text, detectarea orientării, despărțirea paginilor duble, îndreptarea și curățirea imaginii. ABBYY FineReader 9.0 poate prelua formatele: PDF, PDF/A, DJVU, TIFF, JPEG 2000, JPEG, JBIG2, GIF, BMP, DCX, PCX, PNG, XPS, DIB, RLE, WDP și poate salva în formatele: DOC, DOCX, XLS, XLSX, PPTX, PDF, PDF/A, TXT, RTF, HTML, CSV, TIFF, JPEG 2000, JPEG, JBIG2, BMP, DCX, PCX, PNG, DIB, RLE.

Documentul introdus în ABBYY FineReader 9.0 este analizat ca o entitate singulară, păstrând fidel layout-ul documentului original, structura logică și elementele de formatare. ABBYY FineReader 9.0 oferă mai multe tipuri de formate PDF de ieșire: PDF-uri cu căutare după cuvânt cheie, PDF-uri securizate, PDF-uri cu șampilă, PDF-uri cu tagg uri sau PDF/A. Obținerea documentului electronic este mult mai ușor de realizat datorită interfaței intuitive. Imaginile pot fi convertite ușor cu un singur click de mouse,

în format PDF cu căutare după cuvânt cheie, fără lansarea programului. Pentru ca programul ABBYY FineReader 9.0 să ruleze în condiții optime, este nevoie de un calculator cu următoarele cerințe minime hardware:

- PC cu procesor 1GHz sau superior ;
- Sistem de operare Windows Vista, Windows 7, Microsoft Windows Server 2003, Windows XP. RAM 512 MB ;
- Spațiu pe disc de 650 MB la instalare și 612 MB pentru operare ;
- Scanner compatibil TWAIN- sau WIA-, cameră digitală ;
- Rezoluție recomandată XGA (1024x768) sau mai mare ;
- Tastatură, mouse.

Majoritatea cărților în format digital se află stocate în formatul PDF (portable document format), deținut de către compania Adobe. Pentru citirea cărților în format PDF, este nevoie de instalarea programului Adobe Reader 9. Adobe Reader 9 este un vizualizator de PDF-uri care poate deschide și poate interacționa cu toate documentele PDF. Adobe Reader 9 permite vizualizare, căutare, semnare digitală, verificare, tipărire și colaborare pentru documentele electronice. Folosind Adobe LiveCycle cartea digitală poate fi protejată prin parolă, pot fi restricționate anumite funcții, cum ar fi copierea sau editarea textului. Documentele PDF pot primi semnături digitale, astfel asigurându-se autenticitate și integritate lor. Adobe Reader 9 deține instrumente care ajută la descărcarea și organizarea cărților digitale. Portofoliile PDF sunt instrumente folosite pentru a asigura o navigare mai ușoară atunci când se lucrează cu mai multe documente PDF și alte tipuri de documente. Portofoliile oferă posibilitatea de a lucra cu o colecție de materiale (desene, mesaje e-mail, foi de calcul și clipuri video). Toate materialele apar sub forma unui singur fișier, astfel asigurându-se distribuirea, stocarea, recuperarea și colaborarea pentru utilizatorii finali. Adobe Reader 9 este soluția cea mai bună pentru a gestiona drepturile asupra documentelor pentru confidențialitate, autorizare și monitorizare. Interfața programului este simplificată. Adobe Reader 9 permite selectarea modului de citire. Modul paginii alăturate este folosit pentru a vizualiza două pagini alăturate pe ecran. Pagina poate fi mărită, micșorată sau se poate utiliza funcția lupă

pentru o analizare în detaliu a imaginii. În Adobe Reader 9 se pot căuta cuvinte în interiorul paginilor. Adobe Reader 9 necesită următoarele cerințe de sistem:

-Procesor Intel de 1.3 GHz sau echivalent;

-Microsoft Windows 2000 cu Service Pack 4, Windows Server 2003, 2008 si 2008 R2; Windows XP Professional, Home Edition sau Tablet PC Edition cu Service Pack 2 sau 3 (32 de biti si 64 de biti); Windows Vista Home Basic, Home Premium, Business, Ultimate sau Enterprise Service Pack 1 sau 2 (32 de biti si 64 de biti), Microsoft Windows 7 Starter, Home Premium, Professional, Ultimate sau Enterprise (32 de biti si 64 de biti);

-128 MO de RAM (256 MO recomandat);

-335 MO de spatiu disponibil pe hard disk (spatiu suplimentar necesar pentru instalare);

-Microsoft Internet Explorer 6.0, 6.0 cu Service Pack 1, 7.0 sau 8.0; Firefox 2.0, 3.0 sau 3.5.

4.2. Desfășurarea aplicației practice. Conversia cărții din format jpeg în format pdf folosind Abby FineReader 9.0

Pentru început s-a ales o carte electronică deja existentă în catalogul online de carte veche și rară al Bibliotecii Universității Transilvania. Este vorba despre *Elementa linguae daco-romanae, sive Valachicae* (vezi fig.4.5), o lucrare foarte valoroasă pentru patrimoniul cultural, tipărită la 1780 la Viena. Cartea reprezintă prima gramatică tipărită a limbii române și a fost elaborată de Gheorghe Șincai alături de Samuel Micu Klein. Cartea este scrisă în limba latină, dar conține și porțiuni scrise în limba română.

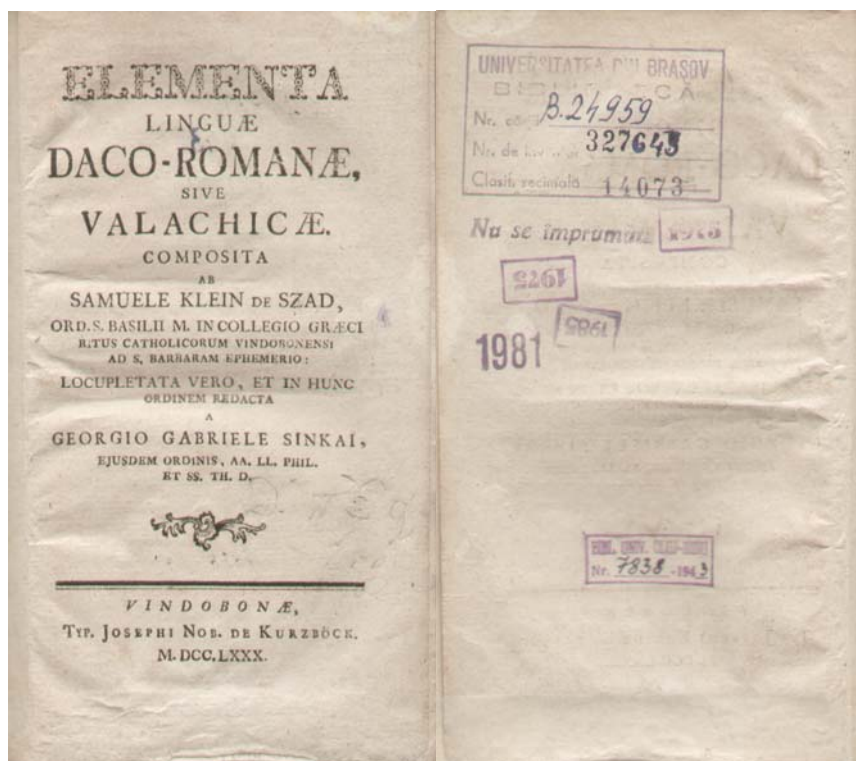


Fig.4.5 – Coperta și prima pagină a cărții “Elementa linguae daco-romanae, sive Valachicae”

Cartea a fost digitizată prin scanarea documentului original și salvată în format JPEG. Cartea digitizată conține 107 pagini reprezentate de 170 de imagini JPEG. Documentul electronic a fost extras din catalogul de carte veche și rară al Bibliotecii Universității Transilvania din Brașov.

A fost necesar un calculator pentru a procesa textul, respectiv imaginile și de două programe: ABBYY FineReader 9.0 pentru a prelucra cartea electronică și Adobe Reader 9 pentru a vizualiza cartea electronică în format PDF. Configurația hardware a calculatorului folosit pentru realizarea cărții digitale este următoarea: Intel Celeron, procesor de 2.3GHz, memorie RAM de 512 MB, sistem de operare Windows XP Professional Service pack 2. Programul ABBYY FineReader 9.0 și vizualizatorul Adobe Reader 9 au fost descărcate prin intermediul Internetului, ambele fiind variante gratuite (trial).

După stabilirea instrumentelor necesare s-a trecut la instalarea pe calculator a programelor și copierea într-un singur folder a paginilor cărții electronice. După instalarea programelor și crearea folderului care conține paginile cărții digitizate, s-a

lansat aplicația ABBYY FineReader 9.0 printr-un simplu dublu click pe iconul aflat pe desktop sau în bara de start. După deschiderea aplicației, a apărut o fereastră principală (fig.4.6).

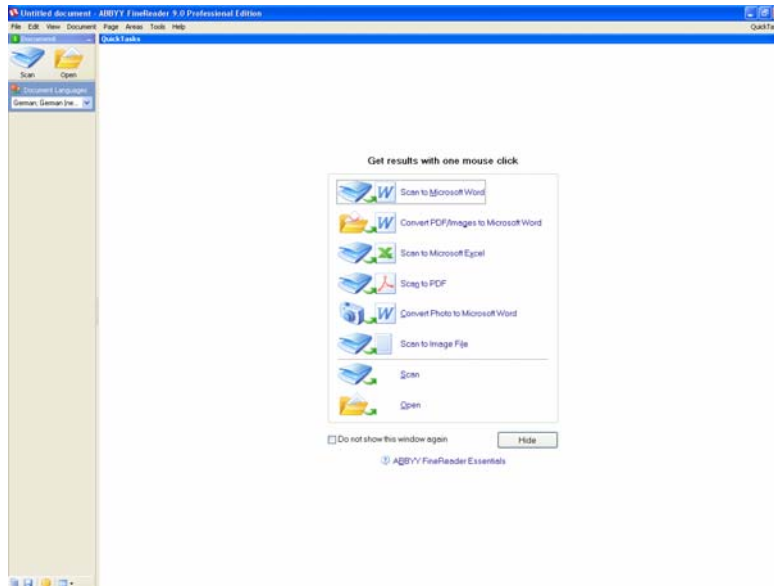


Fig.4.6. – Fereastra de întâmpinare

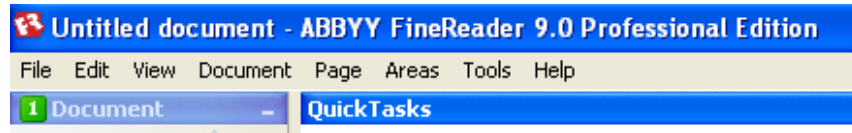


Fig.4.7 – Bară de meniu – Menu Bar

Fereastra de întâmpinare conține o bară de meniu (fig.4.7.) în partea de sus. Această bară conține o serie de meniuri: File, Edit, View, Document, Page, Areas, Tools, Help. Prin intermediul meniului File se poate lansa scanarea paginilor, se poate deschide un PDF, o imagine sau un fișier, un document FineReader deja existent sau un nou document FineReader. Tot în meniul File se găsește opțiunea de salvare a documentului și o altă opțiune prin care se poate exporta paginile în Microsoft Word, Excel, Power Point, Acrobat Reader, Open office etc. Din meniul File se selectează funcția de printare a documentului, se salvează și se trimite prin mail imaginile conținute de document. În meniul Edit se află funcții pentru editarea de text (Undo, Redo, Cut, Copy, Paste, Clear, Select all) și pentru căutarea în text. Din meniul View se setează proprietățile ferestrelor și cele ale barelor cu unelte (Toolbars). Accesând meniul Document, se poate selecta

funcția citire a documentului folosind tehnica OCR și se poate analiza formatul documentului. Tot din meniul Document se controlează navigare în cadrul documentului, putându-se schimba maniera de numerotare a paginilor. Meniul Page conține funcții ca: citirea paginii folosind tehnica OCR, analiza formatului acesteia, editarea imaginii pagini și rotire pagina. De asemenea din meniul Page se lansează unealta de ștergere a anumitor arii de text și a anumitor pagini din document. Din meniul Page se poate alege dacă se vrea sau nu afișarea proprietăților paginii respective prin bifare sau debifarea opțiunii Page properties. Accesând meniul Area, se pot prelucra anumite arii din document. Meniul Tools este meniul ce conține uneltele pe care le pune la dispoziția utilizatorului, programul ABBYY FineReader 9.0. Din meniul Tools se lansează Task-urile automate prin care se poate scana și converti automat în anumite formate. Accesând meniul Tools putem lansa editoarele de stil, pattern și limbă. Prin intermediul meniului Tools se pot vizualiza dicționarele de limbă de care dispune programul. Tot prin intermediul acestui meniu se poate lansa unealta de verificare a ortografiei (Check Spelling). Din meniul tools se poate lansa fereastra Options, care conține o serie de opțiuni legate de scanarea, salvarea, deschiderea și limba documentului. Ultimul meniu din bara de meniu este meniul Help care conține o serie de instrucțiuni de folosire care vin în ajutorul utilizatorului. Din bara de meniu se intră în meniul Tools și se lansează fereastra Options (fig.4.8), unde s-au realizat următoarele setări:

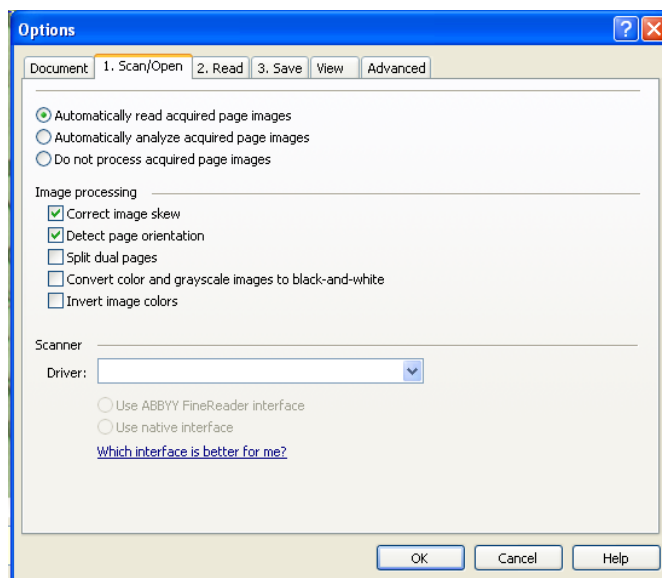


Fig.4.8. – Fereastră de opțiuni

-în fereastra Document s-a introdus titlul (*Elementa linguae daco-romanae, sive Valachicae*), subiectul (Prima gramatică tipărită) și numele autorilor (Gheorghe Șincai alături de Samuel Micu Klein);

-în fereastra Scan/Open se selectează citirea automată a imaginilor și îndreptarea imaginii;

-în fereastra Read se selectează opțiunea de citire *Thorough* pentru documente complexe;

-în fereastra Save se selectează PDF și setarea automată a mărimii paginii. Tot în această fereastră s-a ales ca textul recunoscut să fie suprapus imaginii pentru ca documentul final să păstreze ca fundal imaginea paginii originale. Programul deține o funcție de recunoașterea și procesarea automată a fundalului. În figura 4.9. se poate observa că documentul convertit prin recunoașterea optică a caracterelor păstrează atât aspectul fizic al documentului original, cât și întreg conținutul informațional al acestuia. Așa cum se poate observa în figura 4.10., ABBYY FineReader 9.0 păstrează notele de subsol din finalul documentului.

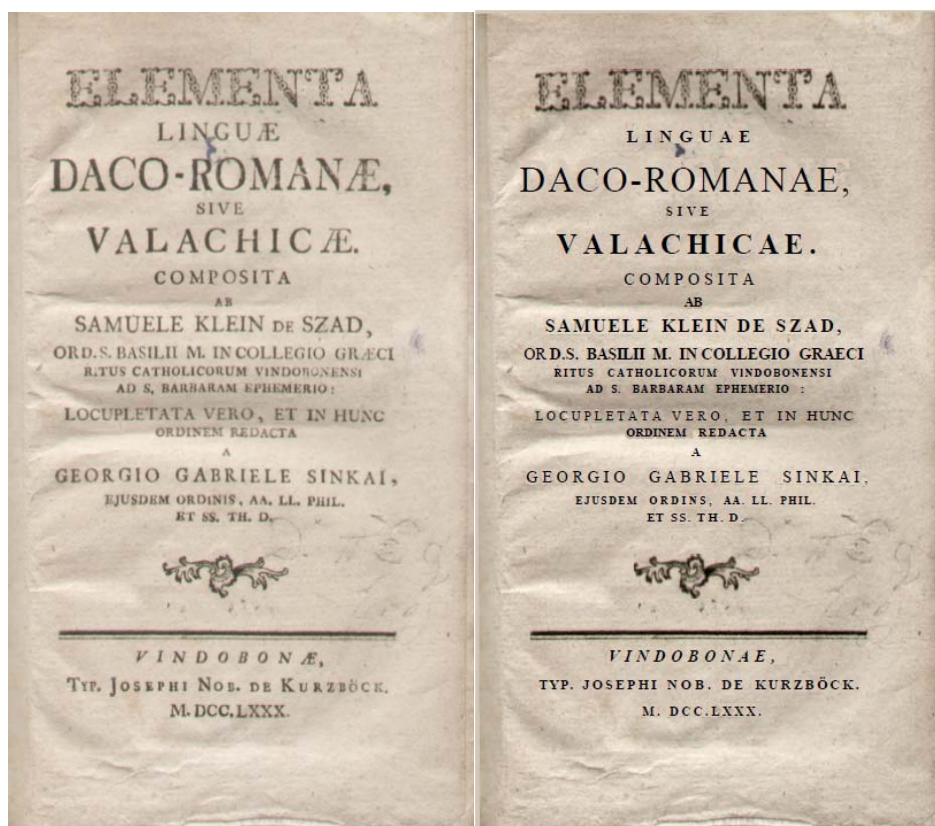


Fig.4.9 – Pagină din documentul original și pagină din documentul rezultat în urma prelucrării celui original folosind Recunoașterea Optică a Caracterelor

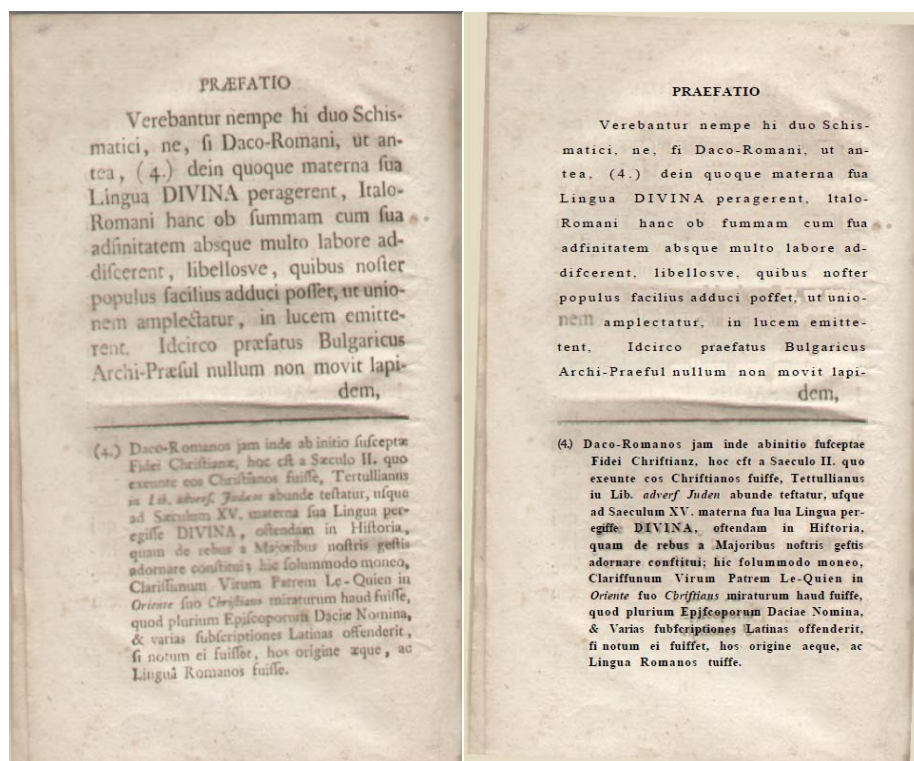


Fig.4.10 – Exemplu de recunoaștere optică a notelor de subsol

Pe lângă bara de meniu, fereastra de întâmpinare conține alte două ferestre: o fereastră intitulată Document și o fereastră dedicată comenzilor rapide accesibile printr-un singur click de mouse (quick tasks). După cum se poate observa în figura 4.11, în prima fereastră se disting două butoane, unul pentru a deschide procesul de scanare iar cel de-al doilea pentru a selecta documentul pe care dorim să îl prelucrăm în ABBYY FineReader 9.0. Tot în prima fereastră putem selecta limba documentului. Cea de-a doua fereastră conține o serie de comenzi rapide: scanarea direct în format Microsoft Word, conversia imaginilor/pdf-urilor în format Microsoft Word, scanarea direct în format Microsoft Excel, scanarea direct în format PDF, conversia unei fotografii digitale în format Microsoft Word, scanare imagine, scanare și deschidere fișier.

După ce s-au setat opțiunile documentului, se selectează limba în care textul original a fost redactat. Textul fiind un text bilingv, limba documentului a fost limba română și limba latină (Fig.4.6). După apăsarea butonul Open din prima fereastră, se selectează paginile scanate ale cărții electronice *Elementa linguae daco-romanae, sive Valachicae*. (fig.4.12). Pentru a selecta o altă limbă se selectează opțiunea More languages care lansează o altă fereastră ce conține toate limbile recunoscute de ABBYY FineReader 9.0.

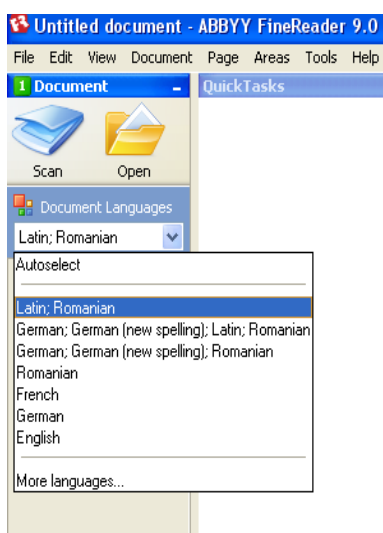


Fig.4. 11 – Selectarea limbii documentului

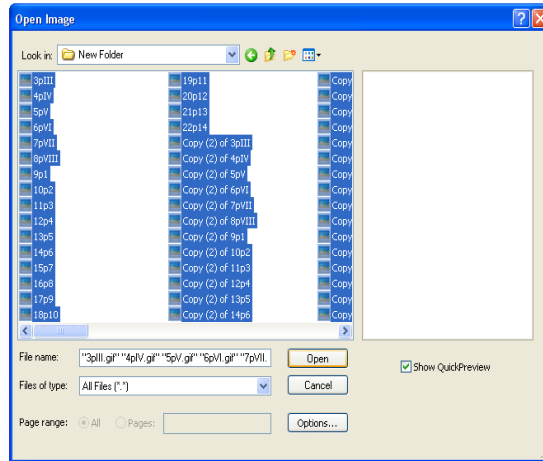


Fig.4.12. Selectarea și deschiderea imaginilor

La apăsarea butonului Open apare o fereastră de încărcare, un loading screen care arată progresul citirii imaginilor introduce (fig.4.13). Atunci când citește o imagine, ABBY FineReader 9.0 descoperă o serie de erori și oferă sfaturi pentru a le putea rezolva. În figura 4.13 se poate observa că programul a întâmpinat o serie de probleme în recunoașterea caracterelor datorită rezoluției imaginii, a limbii de recunoaștere și a paternului selectat.

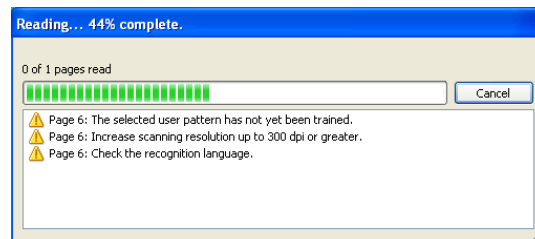


Fig. 4.13. Loading screen

Imaginile în format JPEG ce se află în componența cărții electronice au fost importate în programul ABBY Fine Reader 9.0. Programul a efectuat automat recunoașterea optică a caracterelor. După deschiderea imaginilor, fereastra principală s-a împărțit în trei ferestre (vezi fig.4.4): prima fereastră numită *Document* ce conține o listă a paginilor cărții electronice, a doua fereastră numită *Image* ce conține prima pagină a cărții electronice și o a treia fereastră numită *Text* ce conține textul care a rezultat în urma procesului de recunoaștere optică a caracterelor. În fereastra *Text* întâlnim un instrument de corecție pentru caracterele cu un nivel scăzut de recunoaștere, care permite confruntarea cu imaginea originală aflată în fereastra *Image*. În partea de jos a ferestrei

unu și ferestrei doi există o a patra fereastră opțională care este de fapt un instrument de corectare a greșelilor ortografice folosind tehnica de mărire a imaginii originale.

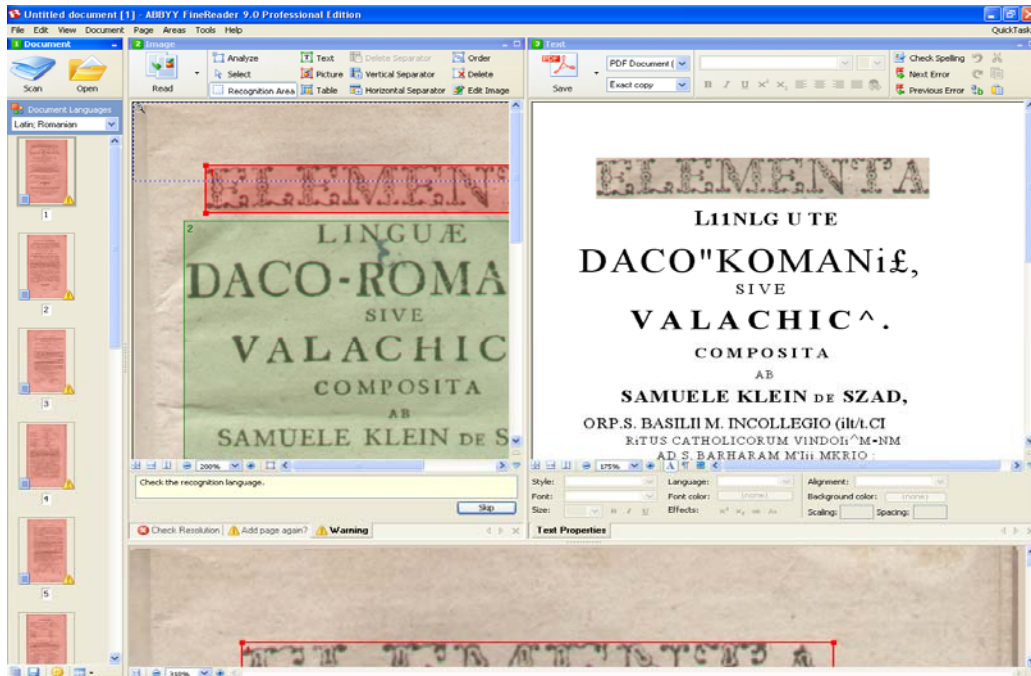


Fig.4.14 – Pagina de lucru

Fereastră *Document* permite selectarea paginii dorite. Pagina poate fi selectată printr-un singur click, apoi va fi deschisă în fereastra *Image*. Pentru a citi imaginea se acționează butonul *Read*, fiind posibilă atât citirea unei singure pagini, cât și citirea întregului document.

După ce imaginile au fost încărcate, s-a trecut la editarea lor folosind opțiunea *Edit Image* din fereastra *Image*. Pentru a deschide această opțiune am apăsat butonul *Edit Image*. Astfel, s-a deschis o altă fereastră în care am putut edita imaginile pentru o mai bună recunoaștere optică a caracterelor (vezi fig.4.15). În fereastra *Edit Image* întâlnim o serie unele folosite pentru editarea de imagini:

- *Deskew & Straighten* este o unealtă folosită pentru a elimina aspectul distorsionat al paginii;
- *Rotate* este o unealtă folosită pentru a roti pagina;
- *Split* este o unealtă de separare a două pagini, în cazul în care ambele pagini au fost scanate într-o singură imagine;

- Crop este o unealtă care poate fi folosită pentru a decupa anumite porțiuni din imagine;
- Invert este o unealtă care transformă imaginea în negativul ei prin inversia culorilor;
- *Image resolution* este o unealtă prin care se selectează rezoluția documentului;
- Eraser este o unealtă cu ajutorul căreia se pot șterge anumite porțiuni din text.

Imaginile cărții *Elementa linguae daco-romanae, sive Valachicae* au fost editate aplicându-se funcția de îndreptare a linilor de text (straighten text lines) și funcția Deskew pentru a înlătura aspectul distorsionat al paginii scanate. Apoi s-a modificat rezoluția imaginilor de la 220 dpi la 300 dpi.

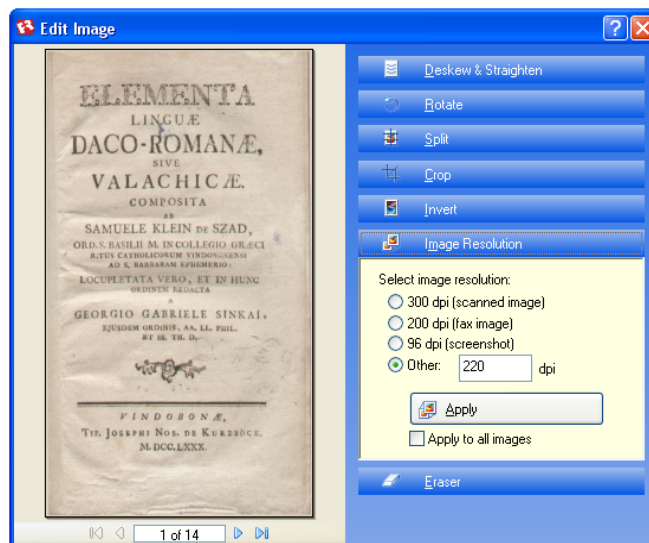


Fig.4.15 – Meniu editare imagine

După editarea imaginilor am apăsut butonul read din fereastra Image pentru o recitare a imaginilor editate.

Al cincilea pas a constat în corectarea manuală folosind unealta *Check Spelling* a posibilelor greșeli care au rezultat în urma prelucrării imaginii, erori pe care ABBY Fine Reader nu le-a putut remedia datorită nerecunoașterii caracterelor respective. Asemenea tuturor programelor OCR, ABBYY FineReader 9.0 nu are o acuratețe de 100% în recunoașterea caracterelor. Cartea electronică a prezentat unele pagini greu lizibile sau de calitate inferioară care nu au putut fi recunoscute cu exactitate. Pentru a corecta posibilele greșeli de conținut am folosit unealta *Check Spelling*, unealtă de verificare a ortografiei

(vezi fig.4.16). *Check Spelling* se poate selecta din meniul *Tool*. Unealta prelucrează fiecare cuvânt recunoscut din imaginea originală, oferind o serie de variante considerate corecte. *Check Spelling* are la bază un dicționar de limba Română încorporat în program. Atunci când cuvântul din text a corespuns celui din imagine, s-a apăsat tasta *Ignore*, în caz contrar s-a corectat cuvântul fie manual, fie folosind o sugestie deja existentă. Pentru a introduce cuvântul corectat în text s-a apăsat tasta *Replace*. Unealta *Check Spelling* este ușor de folosit și foarte utilă pentru asigurarea unei copii cât mai fidele a documentului original. După ce am corectat cuvintele greșite, am închis unealta *Check Spelling* prin apăsarea tastei *Close*.

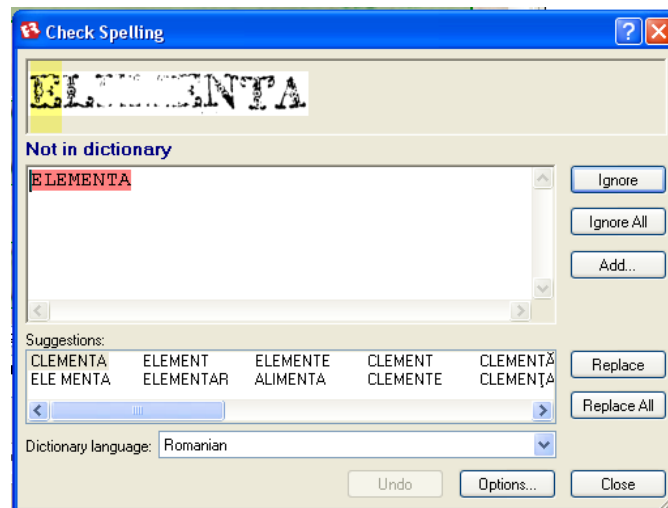


Fig.4. 16 – Check Spelling, unealtă de verificare a ortografiei

După verificarea ortografiei textului (*Check Spelling*), se efectuează o a doua verificare pentru a exclude în totalitate erorile. Așa cum se poate observa în figura 4.17, a doua verificare a constat în compararea textului din fereastra *Text* cu imaginea din fereastra *Image*, folosindu-se unealta de mărire a imaginii (*Zoom*).

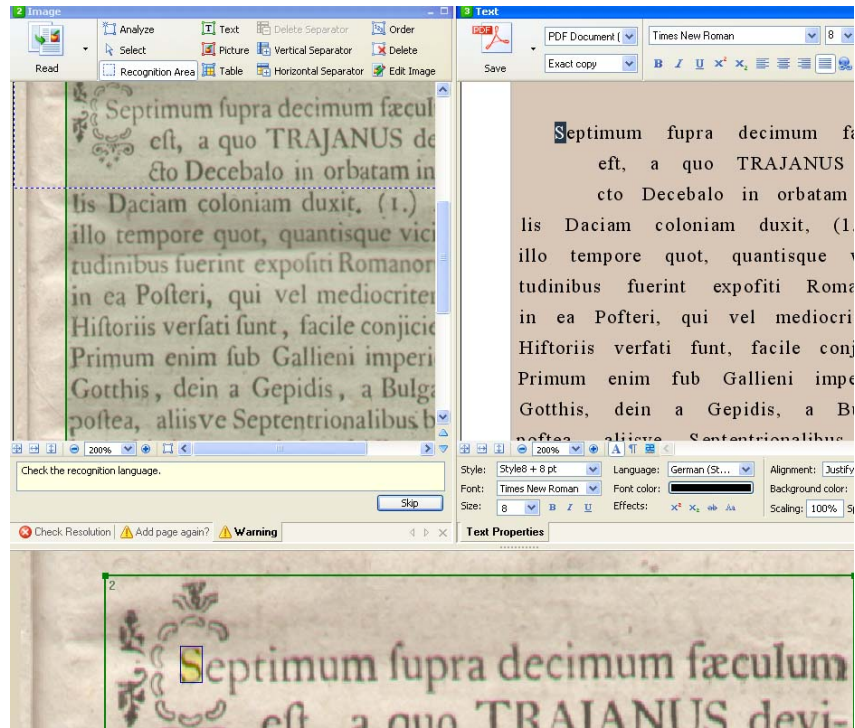


Fig.4. 17 – Verificare ortografică folosind unealta Lupă (Zoom)

Pentru o verificare precisă a ortografiei textului s-a selectat opțiunea de subliniere a literelor pe care ABBY FineReader 9.0 le consideră a fi recunoscute greșit. Pentru a selecta această opțiune se utilizează butonul *Highlight Uncertain Characters and Show Spelling Errors* situat în partea stângă jos a ferestrei Text, după cum se poate observa în figura 4.18 . După cum se poate sesiza în figura 4.19 atunci când s-a apăsat butonul *Highlight Uncertain Characters and Show Spelling Errors*, s-au selectat toate literele ce ar putea prezenta eventuale erori în urma recunoașterii.

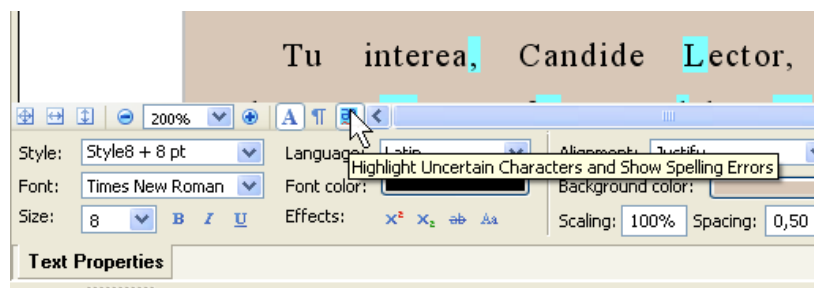


Fig.4.18. – Locația butonului Highlight Uncertain Characters and Show Spelling Errors

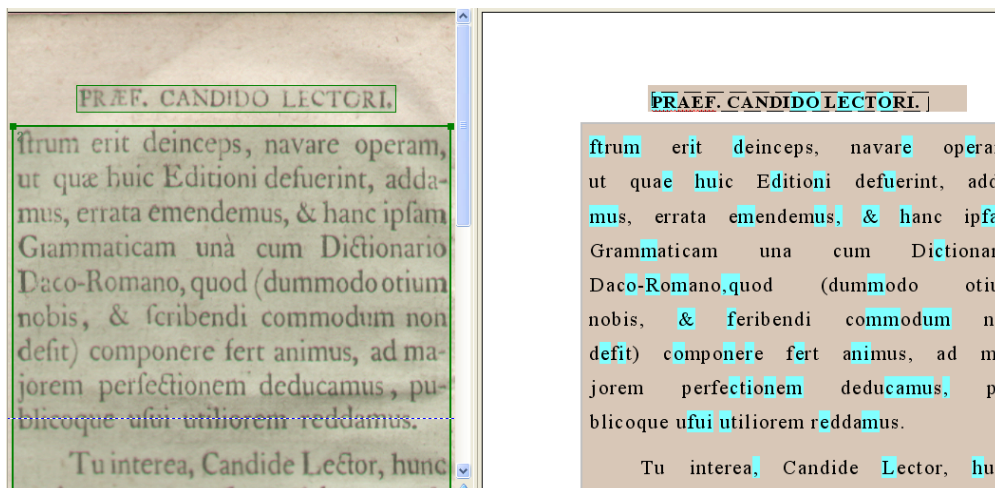


Fig.4. 19 - Funcția Highlight Uncertain Characters and Show Spelling Errors

Tot în fereastra *Text* s-a realizat așezarea în pagină a textului - justify, s-au selectat fontul, culoarea și mărimea fontului- times new roman, mărime caracter 8.

După ce s-a asigurat că textul este identic cu cel din imaginea originală, a urmat al șaselea pas care a constat în salvarea în format PDF a cărții electronice rezultate în urma prelucrării prin recunoașterea optică a caracterelor. Pentru a salva documentul s-a folosit comanda *Save*. Apoi se denumește fișierul *Elementa linguae daco-romanae, sive Valachicae*, se alege locația salvării și se execută operațiunea propriu-zisă.

Pentru a asigura accesul online la cartea digitizată *Elementa linguae daco-romanae, sive Valachicae*, paginile documentului pot fi transformate în limbaj HTML (fig.4.20), fiind posibilă crearea unui site care să conțină cartea. Pentru a lansa această aplicație, trebuie accesat meniul *File* din bara de meniuri, apoi *Sent Selected Pages To* și trebuie să se selecteze *Web Browser*. O dată cu transportarea paginilor documentului în *Web Browser*, acestea își schimbă aspectul. Paginile își pierd fundalul și astfel asemănarea cu documentul original este diminuată. Documentul postat online contribuie foarte mult la îmbunătățirea accesului la informații.

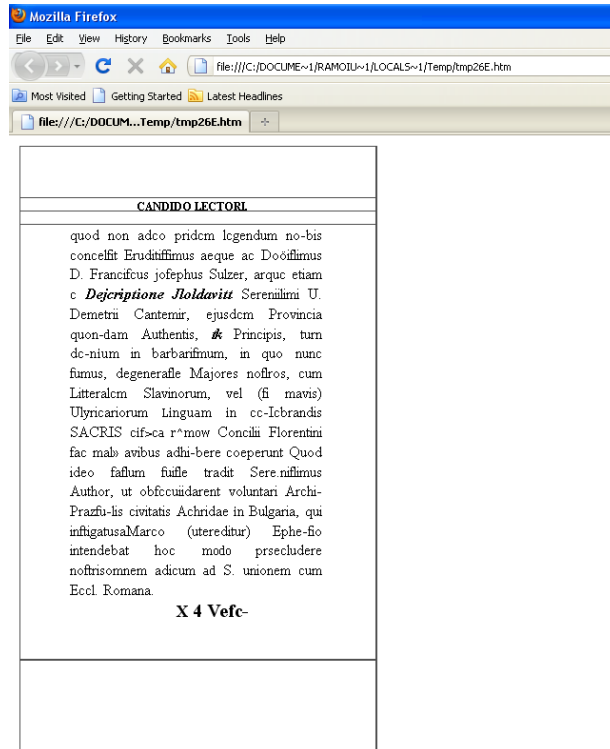


Fig.4.20.- Paginile cărții în limbaj HTML

Prin salvarea în format PDF, s-a putut crea o nouă carte digitală, cursivitatea paginilor documentului fiind asigurată automat de program. Datorită vechimii, textul din documentul original este greu de citit, unele din caractere fiind ilizibile. Cartea digitală în format PDF oferă o lizibilitate mai bună a textului, precum și o rezoluție mai înaltă. Cel mai important avantaj al cărții digitale în format PDF, este posibilitatea de a căuta după cuvânt cheie, Adobe Reader 9 având un motor de căutare încorporat. Pentru a căuta după cuvânt cheie în cartea digitală *Elementa linguae daco-romanae, sive Valachicae*, trebuie în primul rând introdus cuvântul sau chiar fraza căutată. După introducerea cuvântului în căsuța de căutare (vezi fig.4.21), se apasă tasta Enter și cuvântul sau fraza căutată va fi evidențiată prin marcarea acestuia printr-un chenar de culoare albastră, așa cum se poate observa în figurile 4.22 și 4.23.

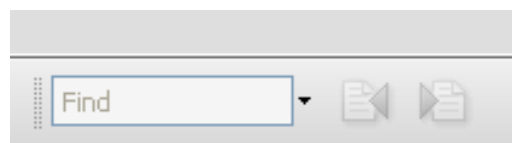


Fig.4.21 – Căsuța de căutare

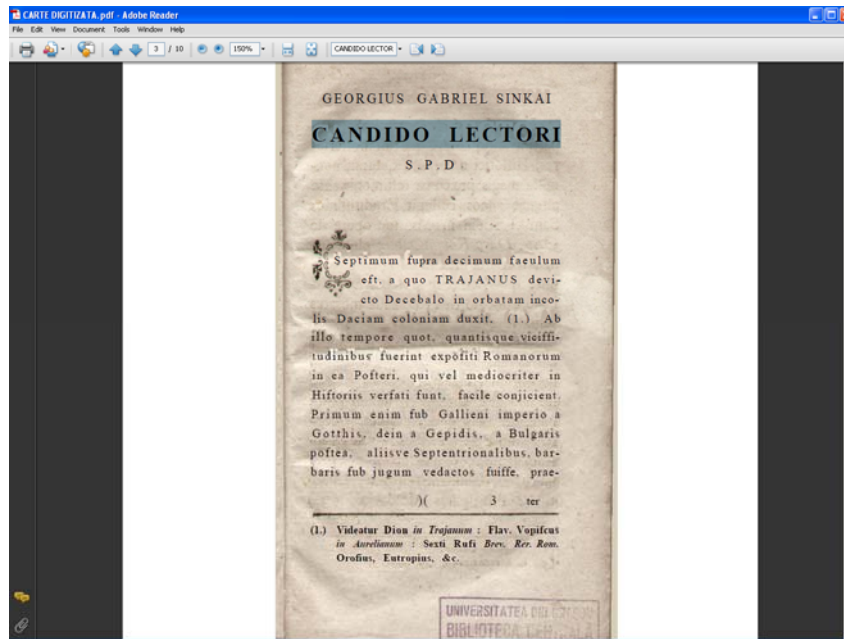


Fig.4.22 – Căutare după cuvânt

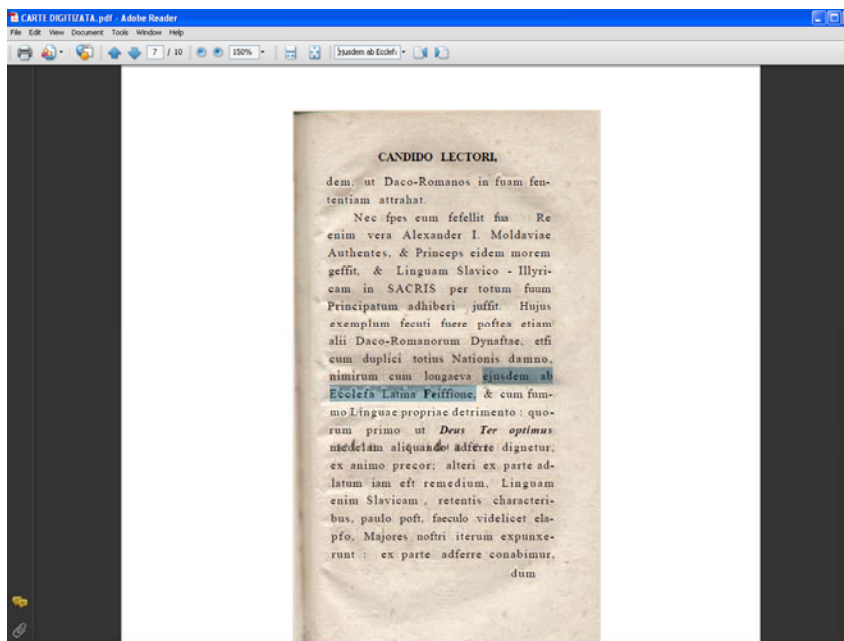


Fig.4.23 – Căutare după frază

ABBYY FineReader 9.0 oferă posibilitatea de a salva documentul în format PDF cu anumite reguli de securitate. Pentru a accesa acesta opțiune trebuie lansată fereastra de opțiuni pentru documentul PDF (fig.4.24). După apăsarea butonului se va deschide fereastra de opțiuni (fig.4.25) Se poate pune parolă pentru deschiderea documentului. În fereastra de opțiuni se va apăsa butonul PDF *Security Settings* care va deschide o nouă

fereastră (fig.4.26) în care se poate alege prin bifare dacă documentul PDF va cere parolă la deschidere și dacă va avea anumite restricții pentru copiere și editare.

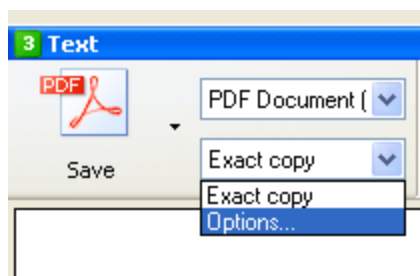


Fig. 4.24 – Butonul de lansare a ferestrei de opțiuni

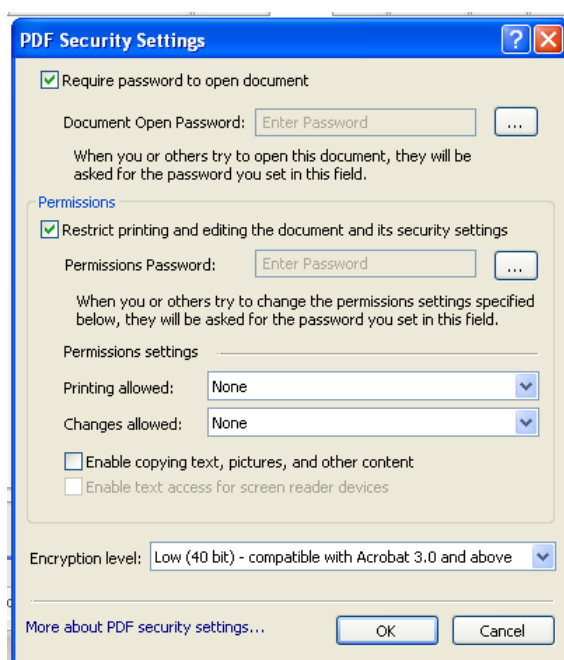


Fig.4.25. – Fereastra de opțiuni pentru securitatea documentului PDF

Cartea electronică în format PDF, rezultată în urma procesului de recunoaștere optică a caracterelor, permite utilizatorului căutarea după cuvânt cheie, astfel accesul la informație realizându-se într-un timp mult mai rapid. Cartea digitizată va fi arhivată cu ajutorul unui program de arhivare, urmând să fie introdusă în bazele de date ale bibliotecii pentru a putea fi consultată sau descărcată. Se vor folosi metadatele pentru descrierea documentului.

5. Propunere pentru realizarea unui model conceptual al sistemelor de informare și documentare cu conținut tehnic

5.1. Descrierea unui model conceptual pentru sisteme de informare și documentare cu conținut tehnic

Pornind de la ideea fundamentală că Biblioteca digitală universitară este un produs al tehnologiei informației pe care instituția îl oferă membrilor comunității sale pentru gestionarea și difuzarea materialelor digitale, s-a elaborat un model conceptual complex destinat bibliotecilor universitare din România.

Acest produs informatic este responsabil de gestionarea documentelor cu caracter științific realizate în cadrul universității, de organizarea, accesul și distribuția lor, dar și conservarea pe termen lung. Prin acest software se vor colecta și distribui toate documentele științifice create de membrii universității, dar se va asigura accesul și la documentele create în cadrul altor universități, pe baza de acorduri de colaborare.

Ne propunem ca software-ul rezultat să se integreze în același Web-Portal cu platformele de e-learning, implementate în universități (ex. IBM Workplace Collaboration Learning prin proiectului POSCCE Axa 3: „Implementarea unei platforme de e-learning la Universitatea din Pitești”, cod SMIS – CSNR 4415) (<http://elearning.upit.ro>). În felul acesta, se introduce un suport de tip SSO (Single Sign On) pentru autentificarea și autorizarea accesului în rețeaua universității a cadrelor didactice și a studenților, în scopul utilizării atât a instrumentelor de instruire oferite de platforma e-learning pentru un număr maxim de utilizatori, cât și a depozitului digital de documentație pus la dispoziție de Biblioteca universității, pentru extinderea și completarea câmpului informațional.

În figura 5.1. este prezentată schema modelului conceptual pentru sistemul format din bibliotecile universitare participante la proiectul de cercetare “tehnici pentru managementul conținutului digital”-TEMACOD.

Pornind de la aceste considerente, noi propunem un model nou de accesare a bibliotecilor universitare, împreună cu o nouă abordare care să asigure studenților facilitatea de a citi orice carte virtuală existentă în oricare dintre bibliotecile conectate în sistem.

UDL System (University Digital Libraries) este o soluție bazată pe website pentru accesul larg al tinerilor la publicațiile virtuale din bibliotecile mediului academic și în același timp, o propunere pentru un sistem deschis oricăror biblioteci care dețin resurse virtuale.

Sistemul central este conceput pentru a lucra pe platforma LAMP (Linux + Apache + MySQL + PHP), dar poate fi adaptat la platforma WAMP (W - Windows). De asemenea, sistemul poate fi implementat și pe o platforma comercială furnizată de IBM (serverul Internet HTTP Server - IHS, baza de date DB2) pentru a beneficia de facilitățile oferite de infrastructura hardware ce găzduiește întreaga platformă. Astfel, baza de date va cuprinde doar metadatele aferente fiecărui document disponibil, împreună cu nivelul de acces necesar și o referință la poziția sa reală în sistemul de fișiere de pe dispozitivul de stocare externă (serverele IBM utilizează un sistem de stocare externă conectat prin fibră optică NAS - Network Attached Storage, scrierea fiind executată în paralel pe mai multe discuri SCSI conectate în matrice RAID - Redundant Array of Independent Disks).

Prima abordare este descrisă în figura 5.1 – arhitectura sistemului UDL și cuprinde 2 niveluri:

- *Sistemul central* de verificare a autentificării corecte a utilizatorilor și de dirijare a accesului către bibliotecile locale;
- *Sistemul local* de creare, gestionare și accesare a depozitului digital în cadrul fiecărei biblioteci universitare.

Sistemul se bazează pe o configurație de tip portal de acces la o bază de date centralizată, cuprinzând informațiile de redirectare către serverele bibliotecilor implicate în proiect (IP adresa server, login name, password, numele bazei de date). Utilizatorul cititor ia contact cu UDL-ul printr-o interfață grafică sugestivă și nu accesează direct bazele de date ale universităților.

Pe primul nivel al arhitecturii, proiectul focusează pe următoarele funcții:

- identificarea utilizatorilor folosind Serverele de LDAP existente la nivelul fiecărei universități;
- stabilirea conexiunii la baza de date solicitată (UPB, UPIT, ULBS, UTBV) ;
- selectarea catalogului formatat standard și căutarea publicației digitale.

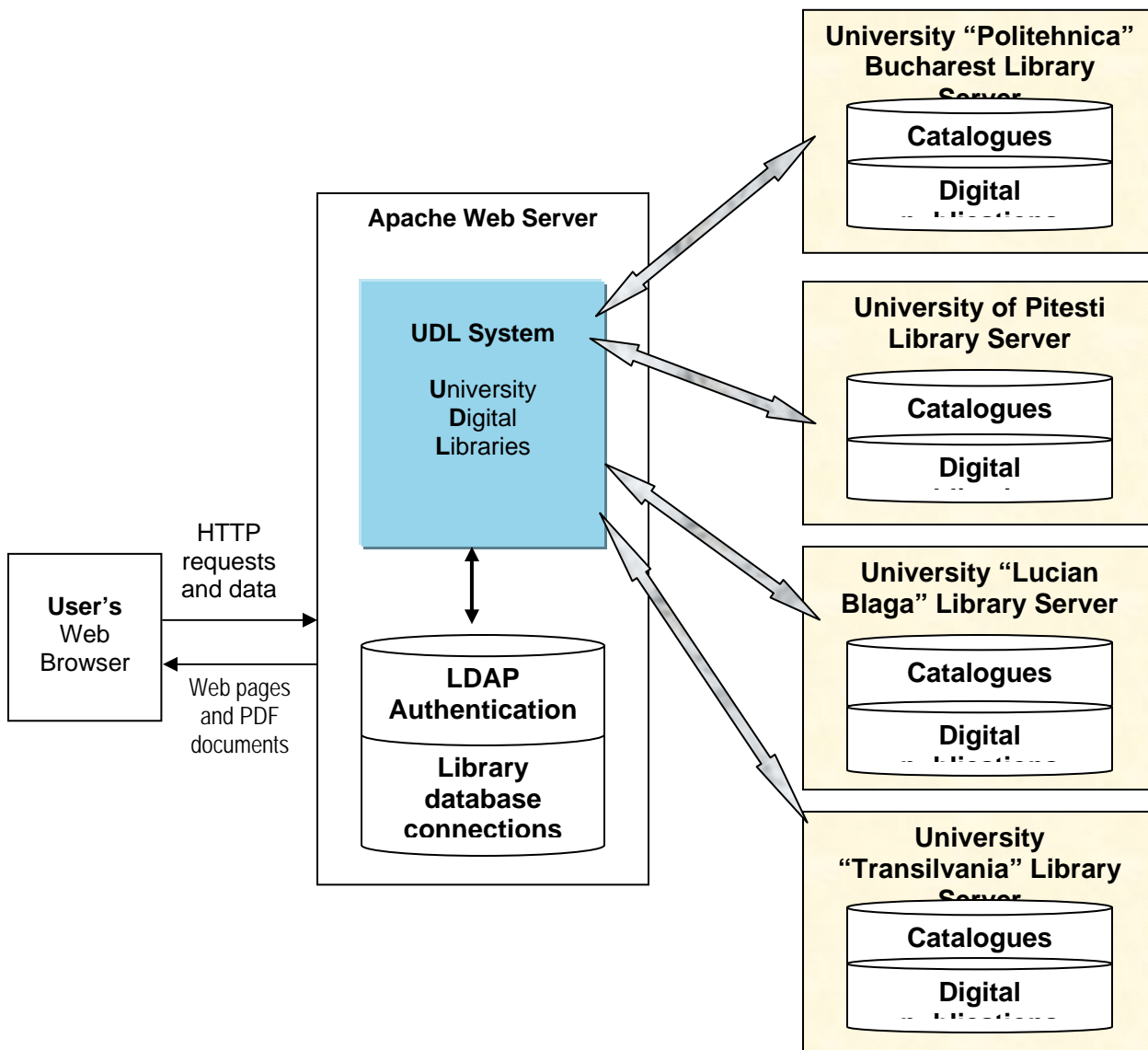


Figura 5.1. Arhitectura sistemului UDL

Pe nivelul al 2-lea, sistemul cuprinde modulele prezentate în figura 5.2:

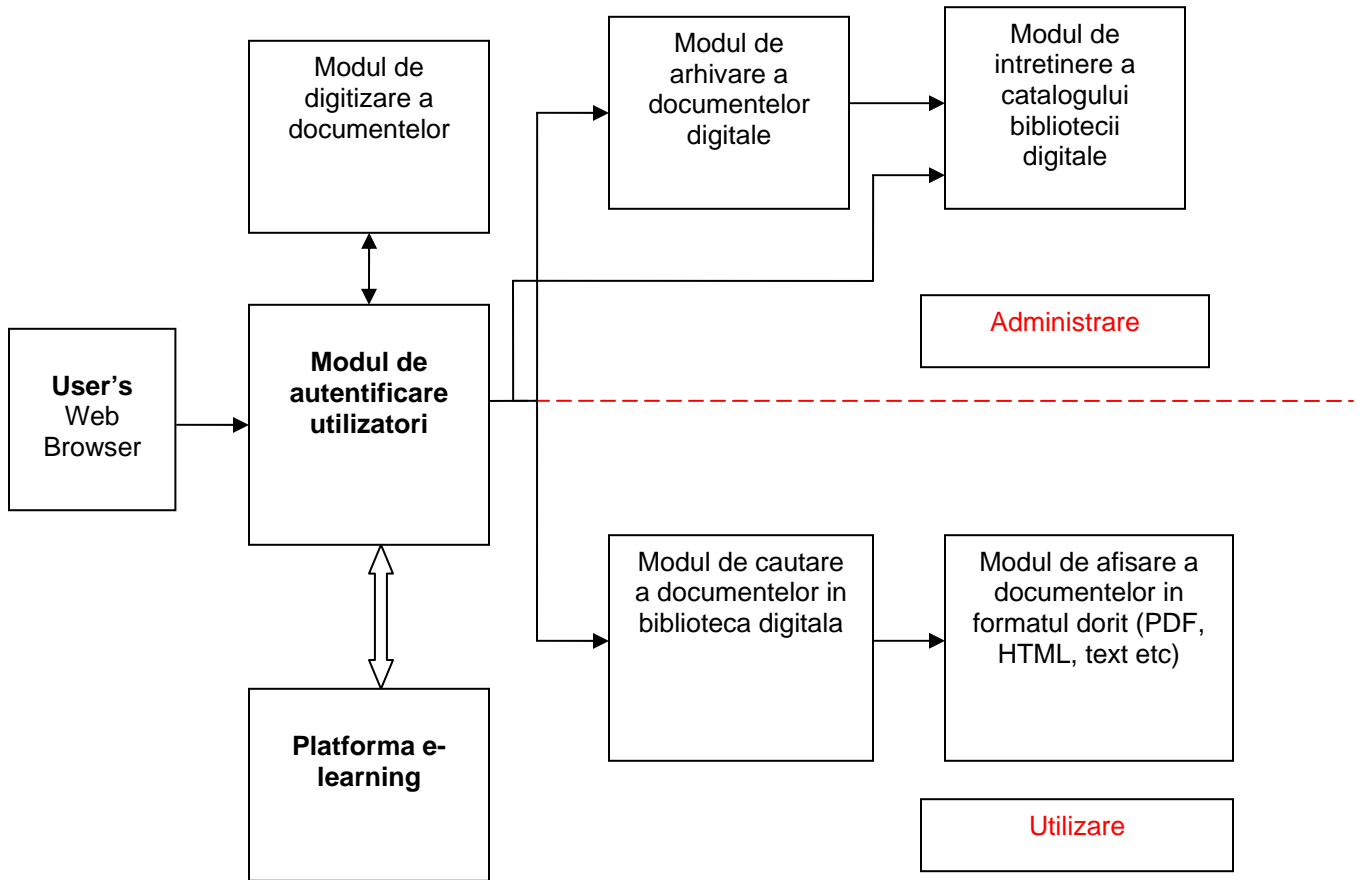


Fig. 5.2 Sistem integrat la nivelul Universității

Sistemul dezvoltat în cadrul universităților va cuprinde două subsisteme (Administrare și Utilizare), care la rândul lor sunt formate din mai multe module, având următoarele funcții:

- digitizarea documentelor cu adaugarea unor informații (metadate) standard pentru fiecare tip de document;
- arhivarea (stocarea) documentelor digitizate;
- adaugarea unei intrări caracteristice documentului arhivat în catalogul bibliotecii digitale;
- căutarea în bibliotecă digitală a oricărui document depozitat;
- afisarea documentului în format .pdf sau html.

Subsistemul de Administrare este compus din mai multe secțiuni care vor putea fi accesate fie de utilizatorii care vor fi desemnați ca administratori de documente în Modulul de Arhivare. Administratorii pot vizualiza grupurile care au drepturi de acces pentru fiecare tip de document și pot vizualiza componența fiecărui grup de acces.

Subsistemul Utilizare este accesibil oricărui utilizator care se autentifică prin user și parolă, o singură dată, la lansarea interfeței principale a sistemului UDL.

Prin integrarea cu *Platforma de e-learning* existentă la nivelul fiecărei universități, operațiunile administrative de gestionare a utilizatorilor și a drepturilor acestora se realizează centralizat, prin intermediul serverului de LDAP. De asemenea și partea de arhivare se poate realiza prin platforma de e-learning, dacă aceasta are prevăzut un modul de arhivare.

Corespunzator modulelor prezentate in schema din fig. 5.2, din punct de vedere functional, sistemul va pune la dispozitia utilizatorilor urmatoarele servicii:

- La nivelul de prezentare avem :
 - Serviciul interfeței cu utilizatorul;
 - Serviciul de publicare.
- Managementul spațiului informațional:
 - Serviciul de căutare ;
 - Serviciul de indexare ;
 - Serviciul colecțiilor ;
 - Serviciul de navigare ;
 - Serviciul de stocare metadate.
- Managementul sistemului funcționează cu:
 - un serviciu de autorizare și autentificare,
 - serviciul de informare și
 - serviciul de management:
- La nivelul colectiv avem serviciul de culegere și agregare. Serviciul de bază este acela de depozitare și stocare a documentelor

Organizarea conținutului depozitului digital se poate face pe departamente de cercetare sau pe facultăți, la nivelul cărora se va lua și decizia asupra tipului de materiale

ce vor fi depozitate. Pe lângă fondul propriu de informație științifică, bibliotecile universitare vor achiziționa și alte lucrări, a caror arhivare va fi stabilită în urma unor verificări de calitate, oportunitate și proprietate intelectuală. Astfel, bibliotecile devin veritabile *noduri informationale de referinta* în documentarea publicului interesat de cercetare, deschizând noi orizonturi celor ce doresc să realizeze lucrări bine documentate și cu un standard calitativ înalt, corespunzător cerințelor societății digitale actuale.

5.2. Specificații de proiectare a modulelor Sistemului integrat al bibliotecii digitale universitare

5.2.1. Modul de digitizare al documentelor

Funcțiuni: trecerea documentelor în format electronic, folosind echipamente speciale de digitizare: scanner, aparat foto digital, scanarea filmelor folosite în fotografia clasică;

Documentele create trebuie să aibă următoarele caracteristici:

- trebuie să conțină atât imagini cât și text;
- să existe posibilitatea de căutare;
- să poată fi citite de către clienți web uzuali;
- să aibă conceptul de pagină;
- să fie de dimensiuni reduse ;
- să fie suportate de către sistemele OCR existente (ca documente de ieșire).

Metode și tehnici utilizate : Modulul conține o componentă de achiziție de imagini (scanner sau cameră foto digitală), un sistem software de prelucrare a imaginilor, un sistem software de recunoaștere a caracterelor și un sistem de înmagazinare a informațiilor acumulate în diversele faze ale procesului.

Structuri de date: baza de date local sau stocare temporară într-o bază de date ARHIVARE_TEMP ;

Intercorelație cu alte subsisteme : este un modul independent, care are ca scop final o prezentare a documentelor în format standard și se utilizează înaintea *modulului de Arhivare*.

Intercorelație cu *Modulul de autentificare utilizatori* pentru acces autorizat în modulul Arhivare.

5.2.2. Modul de autentificare al utilizatorilor

Funcțiuni: identificarea utilizatorilor folosind Serverele de LDAP existente la nivelul fiecărei universități; pentru a beneficia de o securitate îmbunătățită, se poate utiliza protocolul HTTPS la autentificare; permite crearea de noi intrări de către rolul administrator.

Metode și tehnici utilizate : Pentru configurarea conturilor de utilizator și a rolurilor pe serverul LDAP s-a utilizat clientul Apache Directory Studio.

Accesul unui utilizator în orice zonă logică a arhivei se face numai pe bază de autentificare și numai dacă:

- utilizatorul este definit la nivelul LDAP și s-a autentificat prin user și parola corectă;
- utilizatorul aparține unui grup de acces cu rol aferent adăugării/stergerii de documente în arhivă;
- utilizatorul nu este blocat de administrator sau de aplicație (prin setarea timpului, blocarea adresei IP sau blocarea în urma încercării de a se conecta de mai mult de „n” ori succesiv, cu credențiale invalide).

Lightweight Directory Access Protocol (LDAP) - este un protocol de aplicație folosit pentru interogarea și modificarea serviciilor de directoare prin intermediul TCP/IP. Un director este un set de obiecte cu atribute organizate într-o structură ierarhică. Un arbore director LDAP de cele mai multe ori reflectă limite politice, geografice sau alte organizații, depinzând de modelul ales. Utilizatorii de LDAP din momentul de față preferă folosirea DNS pentru structurarea nivelelor unei ierarhii. În cadrul directorului pot apărea mai multe intrări reprezentând persoane, organizații, imprimante, documente, grupuri sau orice altceva care reprezintă o intrare arborescentă.

În cadrul proiectului, încărcarea utilizatorilor se va face numai de către administratori, fără a se permite auto-inscrierea de noi utilizatori ai sistemului.

Utilizatorii bibliotecii digitale, ca și cei ai platformei de e-learning, vor fi mai întâi creați în LDAP, după care se activează de către administratorii sistemului. Se va avea în vedere ca administrarea accesului la sistemul de arhivare a documentelor să se facă în principal prin propriile facilități (drepturi de acces la nivelul DB2CM), fără

modificări frecvente în grupurile LDAP. Este posibil să se creeze noi intrări în ierarhia LDAP utilizată în cadrul sistemului (de exemplu: crearea de noi grupuri și roluri), cu condiția ca acestea să fie limitate ca număr și să nu presupună actualizări frecvente.

Urmărind structura organizatorică a unei facultăți și rolurile existente, s-au stabilit următoarele atribute pentru încărcarea unui utilizator:

- tipUtilizator(0 pt student/1 pt profesor)
- login
- nume
- prenume
- email
- facultate
- specializare
- an
- serie
- grupa
- CNP
- catedra
- gradStiintific
- asociat(true/false)
- functie

Atributele sunt elementele responsabile pentru stocarea informațiilor într-un director și schema definește regulile pentru care atributele pot fi folosite într-o intrare, tipul valorilor pe care le pot lua acele atribute și felul în care clienții pot interacționa cu acele valori.

Structuri de date:

Intrari:

a) Se pot folosi fișiere în format CSV pentru încărcarea unui număr mare de utilizatori sau se poate face o încărcare individuală.

Iesiri: a) catalogul de utilizatori localizat pe Serverul de LDAP existent la nivelul universităților.

Intercorelație cu alte module : Modulul de digitizare, Modulul de întreținere a catalogului bibliotecii digitale, Modul de arhivare, Modul de regasire a documentelor.

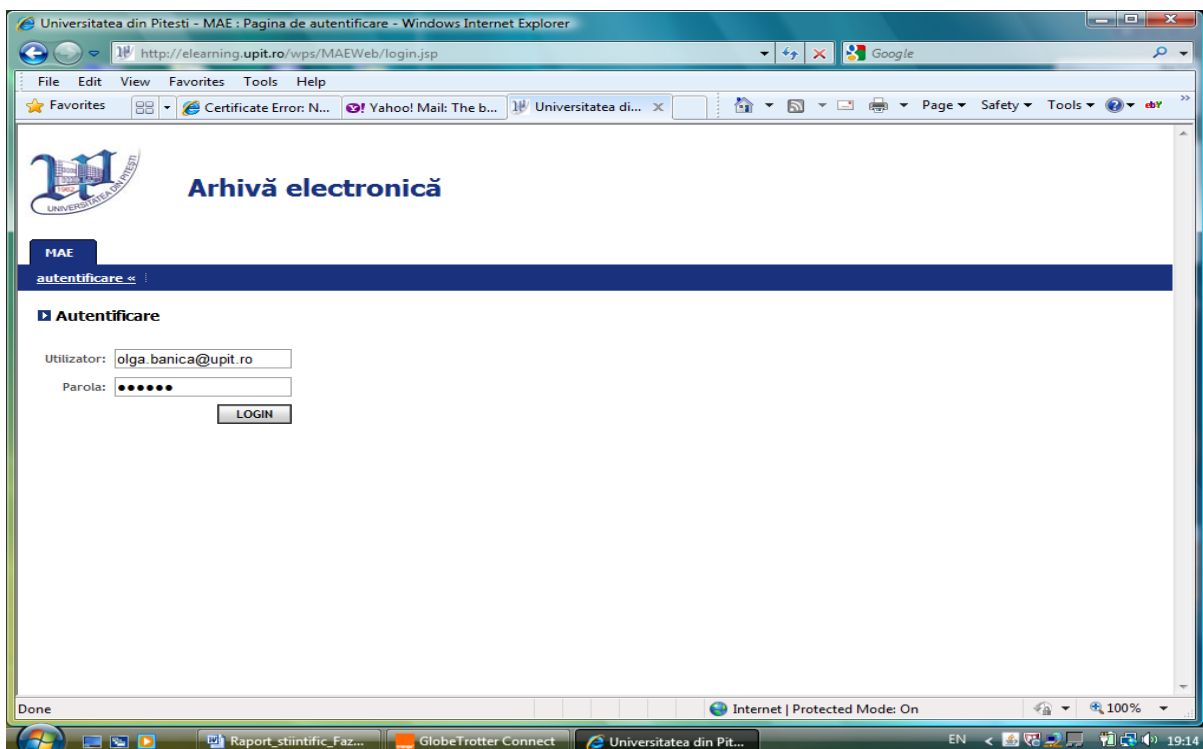


Fig. 5.4. Autentificarea utilizatorilor

5.2.3. Modul de arhivare a documentelor digitizate

Funcțiuni. Configurarea tipurilor de documente și a condițiilor de accesare a lor, stocarea documentelor într-o arhivă unică, la care accesul se face prin interfața web; arhivarea este o operație permisă unui utilizator autentificat, care are drepturi de arhivare. Arhivarea se aplică documentelor în format electronic, prin completarea metadatelor aferente tipului de document ales (teza de doctorat, articol științific, lucrare de licență sau de disertație etc).

Metode și tehnici utilizate. Modulul folosește un catalog de machete de documente, care au seturi de metadata predefinite; un utilizator autentificat completează aceste date, adaugă chei de identificare a documentului și face arhivarea propriu-zisă.

Metadatale pot fi de următoarele tipuri:

- șir de caractere;
- numeric;
- boolean;
- timestamp.

Se vor completa automat data arhivării și numele operatorului care a realizat arhivarea documentului și se vor atașa ca metadata implicite documentului respectiv.

De asemenea, permite arhivarea documentelor formate dintr-un singur fișier sau din mai multe fișiere (multi-part).

Metadata

| NUME METADATA | TIP METADATA |
|---------------|--------------|
| data | timestamp |

STERGE **ADAUGA**

Tip metadata

sir de caractere
 numeric
 boolean
 timestamp

Nume metadata

ADAUGA METADATA

PASUL URMATOR >>

<< Inapoi

Fig. 5.,5, Definirea metadataelor pentru un tip de document

Pentru un tip de document nou creat se stabilesc câmpurile de metadata obligatoriu de completat, se definesc grupurile de acces, extensiile posibile pentru atașamentele documentelor care fac parte din acest tip, dimensiunea maximă a fișierelor atasate, permisiunea de adăugare a mai multor atașamente în cadrul unui singur document.

Implementarea tehnică a accesului la datele din Arhiva electronică se face prin intermediul componentei Resource Manager (RM) a platformei IBM DB2 CM.

Depozitarea datelor arhivate se va face cu ajutorul componentei Library Server (LS) a IBM DB2 CM.

Modulul de Arhivare electronica este realizat pe platformele IBM DB2 Content Manager (DB2CM) și IBM WebSphere Application Server.

Toate documentele arhivate trebuie să aibă un identificator unic la nivelul sistemului, acordat în mod automat de către sistem, care trebuie să fie înscris în Catalogul_Documentelor.

Structuri de date:

Intrări:

- a) baza de date locala sau baza de date ARHIVARE_TEMP ;
- b) catalogul de machete de documente CATALOG_TIP_DOCUMENTE

Ieșiri:

- a) documentul arhivat în baza de date ARHIVARE

Intercorelație cu alte module : Modulul de digitizare, Modulul de intretinere a catalogului bibliotecii digitale.

5.2.4.Modul de intretinere a Catalogului bibliotecii digitale

Funcțiuni: adaugarea/ștergerea intrărilor în Catalog_documente, odată cu adăugarea/ștergerea unui document în Arhiva electronică. Fiecare intrare va conține tipul documentului arhivat, informațiile specificate în metadatele caracteristice și cuvinte cheie de identificare.

Accesul unui utilizator cu drepturi de administrator pentru adaugarea manuală sau ștergerea unei înregistrări din Catalog.

Metode și tehnici utilizate : Se adaugă o nouă înregistrare în tabela Catalog și se preiau informațiile din metadatele asociate tipului de document arhivat.

Structuri de date:

Intrări:

- a) documentul arhivat în baza de date ARHIVARE;
- b) catalogul de machete de documente CATALOG_TIP_DOCUMENTE

Ieșiri: a) catalogul de documente arhivate CATALOG_DOCUMENTE

Intercorelație cu alte module : Modulul de Arhivare, Modulul de autentificare utilizatori.

5.2.5 Modul de căutare a documentelor în sistemul UDL (University Digital Libraries)

Funcțiuni: în cadrul sistemului integrat la nivel de consorțiu de universități - UDL (University Digital Libraries) modulul de căutare implică următoarele funcțiuni:

- autentificarea utilizatorilor în sistemul central;
- conectarea la una dintre bibliotecile universitare participante,
- afișarea Catalogului de documente digitale,
- filtrarea după tip document, titlu, autor, ISBN, editură, data depozitării;
- accesul la documentul digital identificat.

Formatul prevăzut pentru încărcare poate fi .pdf, .html, .txt, dar este preferabil formatul .pdf deoarece el oferă mai multe avantaje, printre care cele mai importante ar fi cele legate de volumul mic al memoriei ocupate și de facilități foarte accesibile de securizare la copiere și imprimare, acces care ar putea rămâne la nivelul de decizie al autorilor.

Modulul de arhivare integrat în platforma de e-learning WCL (Workspace Collaborative Learning) are prevăzute funcțiile specificate anterior la nivelul sistemului local (al Universității).

Metode și tehnici utilizate: accesarea serverului LDAP centralizat permite autentificarea utilizatorului și stabilirea drepturilor pe care le are. Pe baza opțiunilor utilizatorului se face conexiunea cu baza de date a bibliotecii digitale selectate și se transferă controlul sistemului local.

Regăsirea documentelor se va face prin intermediul interfeței web de interogare a arhivei. Modulul de căutare accesează inițial CATALOG_DOCUMENTE pentru identificarea documentului.

Tipul documentului se alege dintr-o listă de selecție, urmat de afișarea unui filtru de căutare, constând în setul de metadate definite pe tipul de document pe care se realizează căutarea, pe baza căruia s-a realizat indexarea documentelor.

În urma căutării, va fi obținută o listă de documente (vor fi afișate metadatele) ce îndeplinesc criteriile de selecție, pentru care utilizatorul respectiv are drept de acces. Accesul la documentele arhivate se face prin intermediul unei aplicații J2EE personalizate, având o interfață utilizator web, accesibilă printr-un browser.

Această interfață va avea mai multe module care vor permite:

- adăugarea documentelor;
- căutarea documentelor arhivate;
- consultarea Catalogului_documentelor.

Structuri de date:

Intrări:

a) Pentru configurarea și accesarea LDAP s-a utilizat clientul Apache Directory Studio;

b) catalogul de documente arhivate CATALOG_DOCUMENTE

Ieșiri: a) lista de documente ce îndeplinesc criteriile de selecție (extrasă din CATALOG_DOCUMENTE)

Intercorelație cu alte module : Modulul de Arhivare, Modulul de autentificare utilizatori.

Exemplu de cautare in depozitul digital, dupa tipul de document si criterii de filtrare: data arhivarii documentului, autorul, cuvinte cheie etc.



Figura 5.6. Selectarea tipului de document

În funcție de tipul documentului selectat va fi afișat un filtru de căutare, constând în setul de metadate definite pe tipul de document pe care se realizează căutarea, pe baza căruia s-

a realizat indexarea documentelor. Este obligatorie completarea valorii cel puțin a unei metadate obligatorii din filtru.

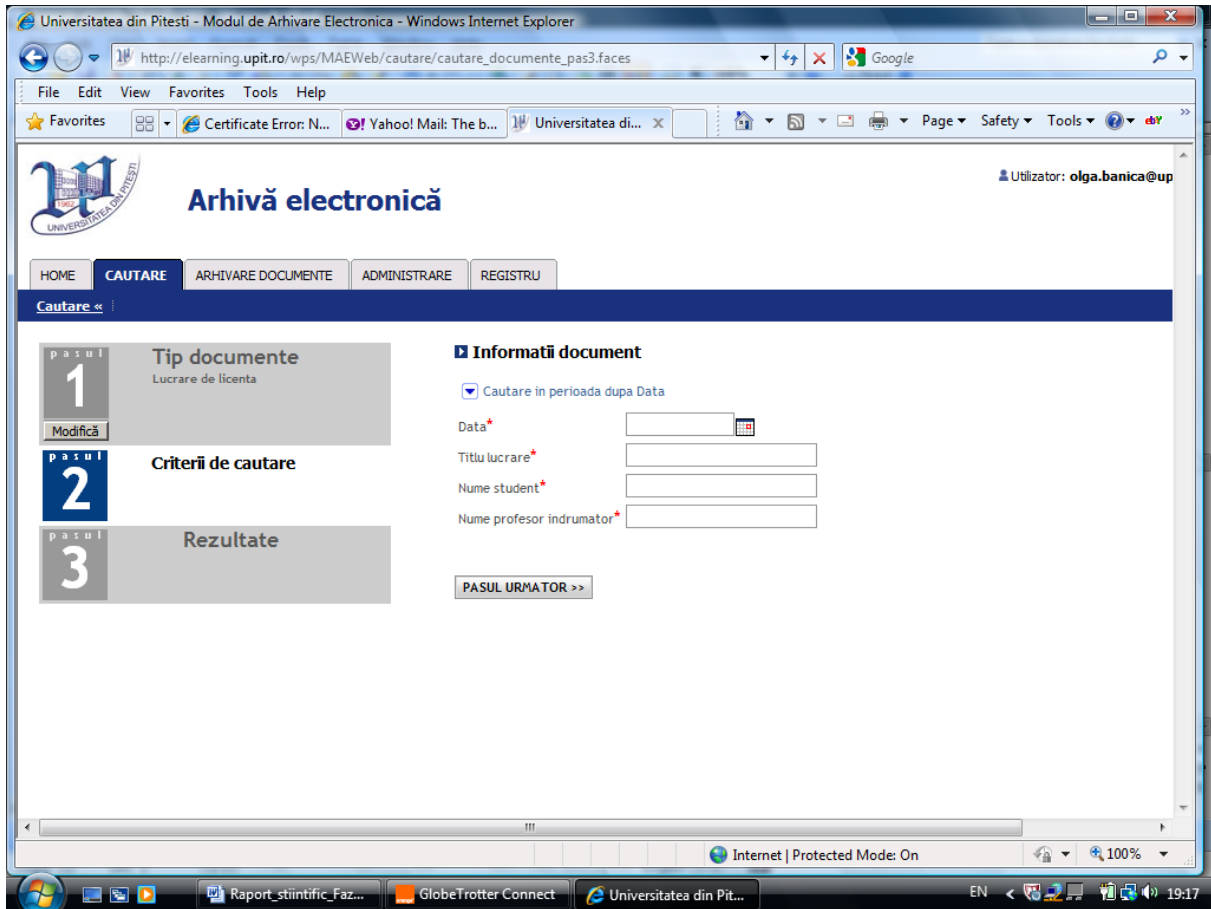


Fig. 5. 7. Completare criteriilor de cautare

5.2.6. Modul de afișare a documentelor în sistemul UDL

Funcțiuni: Din lista de documente selectată pe baza criteriilor de selecție, se selectează o înregistrare din lista și se descarcă din baza de date ARHIVA documentul respectiv în format .pdf, .html, .txt etc.

Metode și tehnici utilizate: Toate documentele arhivate au un identificator unic la nivelul bazei de date ARHIVA, acordat în mod automat de către sistem, care este înscris și în CATALOG_DOCUMENTE. Lista afișată în urma lansării Modulului de căutare conține un link către documentul propriu-zis.

Modulul de Arhivare atașat software-ului de e-learning WCL este realizat pe platformele IBM DB2 Content Manager (DB2CM) și IBM WebSphere Application Server, fiind în sine o aplicație enterprise separată, realizată în tehnologie J2EE și gestionată de platforma.

Structuri de date:

Intrări: a) Lista selectată pe baza de criterii din CATALOG_DOCUMENTE

Ieșiri: a) afișarea documentului prin activarea link-ului asociat

Intercorelație cu alte module : Modulul de cautare a documentelor în biblioteca digitală..

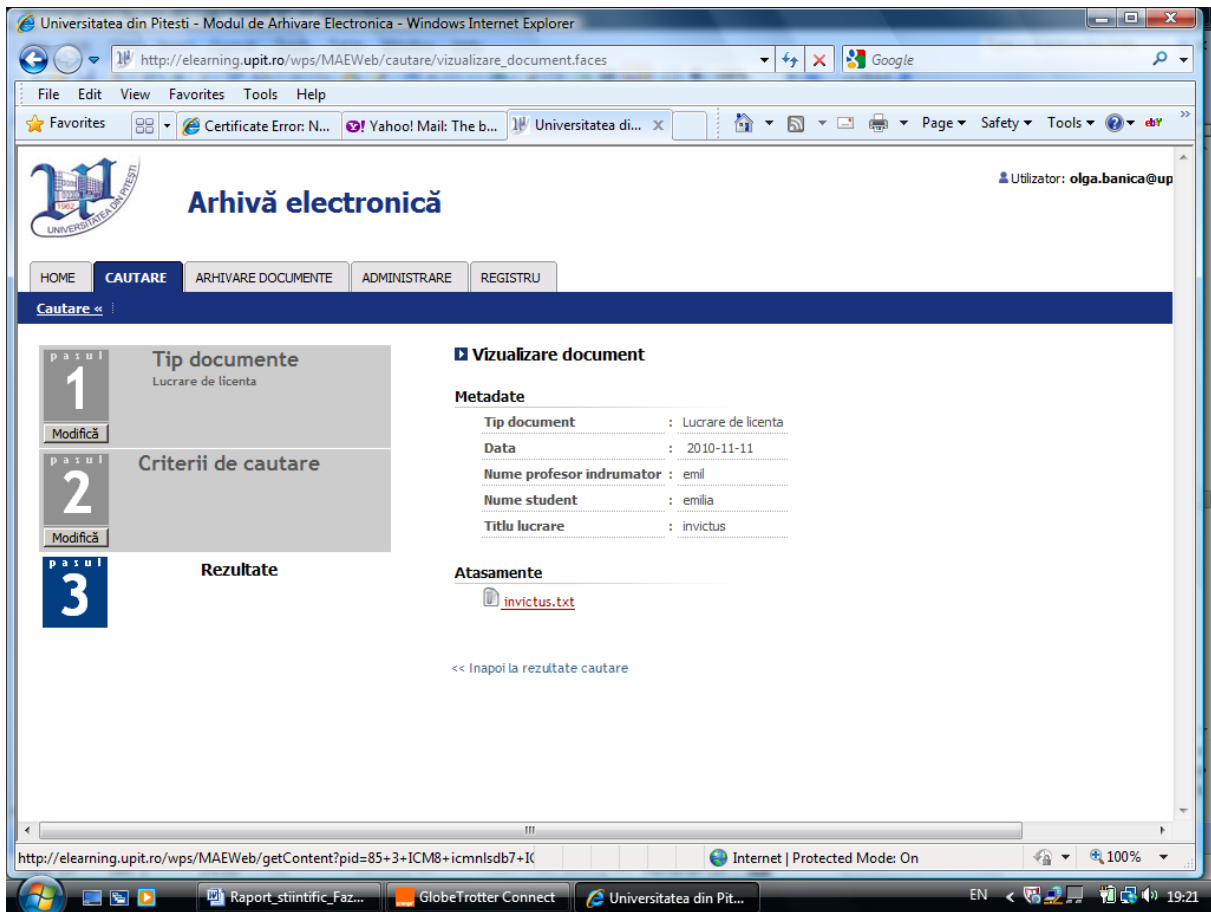


Figura 5. 8. Regăsirea documentului și vizualizarea lui

În urma căutării rezultă o listă de documente ce îndeplinesc criteriile de selecție, pentru care utilizatorul respectiv are drept de acces. Selectându-se o înregistrare din lista, se vor putea vedea informațiile conținute în metadate și se va putea descărca din arhivă

documentul respectiv. Pentru documentele care au mai multe părți (atașamente) vor fi afișate linkuri pentru fiecare parte (atașament).

5.. 3 Infrastructura tehnica de implementare a sistemului de arhivare electronica

Infrastructura tehnica se bazeaza pe o platforma IBM DB2CM, care ofera securitate, integrarea proceselor, servicii pentru ciclul de viata al documentelor, precum si un framework deschis si consistent pentru managementul, partajarea si arhivarea tuturor tipurilor de continut digital. Continutul poate varia de la fisiere HTML sau XML, imagini, fisiere scanate, documente de tip office, pana la documente de tip media, cum ar fi fisiere audio sau video.

CM are mai următoarele componente:

- Library Server
- Mid-tier Server
- Resource Manager

Library Server este sursa centrala pentru indexarea, descrierea, localizarea, organizarea si managementul continutului (arhivei). El suporta definirea unui model de date obiectual, controleaza accesul la continut si ofera un sistem tranzactional pentru pastrarea integritatii datelor. Este integrat cu motorul de cautare al bazei de date relationale pe care se bazeaza (IBM DB2) oferind suport pentru cautari in continut (arhiva). Cautarile se fac in metainformatia asociata documentelor (ex: campul subject pentru un email). De asemenea el ofera servicii de acces securizat la continut, suportand definirea de privilegii pentru useri, precum si access control list (ACL).

Mid-tier Server furnizeaza un API pentru conectarea la CM. De asemenea el este raspunzator pentru managementul conexiunilor la Library Server si Resource Manager.

Resource Manager este depozitarul specializat optimizat pentru stocarea, arhivarea, regasirea si livrarea continutului. El suporta managementul documentelor pe mai multe niveluri:

- se pot asocia obiecte la un document (ex. adnotari)
- documetele pot fi grupate in dosare
- dosarele pot contine si alte dosare

- utilizatorii pot indexa documentele pentru a facilita stocarea și regăsirea lor

Din clientul de administrare CM se poate defini suportul unde este stocată informația, precum și politici de migrare a documentelor arhivate.

CM are în componență și un motor de fluxuri. Rutarea documentelor este inclusă și construită în interiorul CM. Sunt furnizate capacități de rutare a documentelor cum ar fi procesarea paralelă sau serială, puncte de decizie și colectare și ramificarea automată. Definirea unui proces de rutare se face foarte ușor cu ajutorul lui DB2 Content Manager Workflow Builder, ce oferă o interfață simplă, grafică, de tipul drag-and-drop.

În figura de mai sus este prezentat un flux de documente creat cu ajutorul interfeței puse la dispoziție de CM.

CM folosește pentru Library Server și Resource Manager o bază de date relatională, IBM DB2.

CM furnizează un API toolkit pentru programare. Se poate folosi pentru programare:

- Java
- Java Beans
- C++
- Interfața Web Services

Stocarea și accesul la documente

Implementarea tehnică a accesului la datele din Arhiva electronică se face prin intermediul componentei Resource Manager (RM), componentă a platformei IBM DB2 CM.

Depozitarea datelor arhivate se va face cu ajutorul componentei Library Server (LS) a IBM DB2 CM.

Modulul de Arhivare electronică este realizat pe platformele IBM DB2 Content Manager (DB2CM) și IBM WebSphere Application Server.

Accesul la documentele arhivate se face prin intermediul unei aplicații J2EE (soft personalizat), dezvoltată de ***Software Development & Consulting (SDC)***.

Interfața utilizator pentru această aplicație J2EE este web, accesibilă printr-un browser. Această interfață conține mai multe module care vor permite:

- adăugarea manuală a documentelor;
- căutarea în MAE a oricărui document arhivat (fie automat, fie manual);

- executarea de operațiuni administrative (gestiunea utilizatorilor și a drepturilor acestora etc.);

6. Concluzii

Proiectul propus în cadrul Programului 4, „Parteneriate în domeniile prioritare”: *Tehnici pentru managementul conținutului digital-TEMACOD* are ca obiective principale: dezvoltarea de biblioteci digitale, bazate-Web, la nivelul instituțiilor de învățământ superior implicate în contract, interconectarea lor și accesarea rapidă via Internet, printr-o interfață prietenoasă și eficientă.

Bibliotecile publice, cele școlare și academice trebuie să fie capabile să dezvolte website-uri care să permită preluarea de informație științifică digitizată, din categoria cărților scrise, a celor audio sau video, schimbând fundamental noțiunea de resursă educațională.

Tema abordată este de mare actualitate în întreaga lume, se pune problema construirii unei Biblioteci digitale la nivel mondial, accesibilă prin Internet, gratuită și în format multilingv, care să conțină materiale semnificative despre culturile existente, inclusiv manuscrise, hărți, cărți rare, note muzicale, înregistrări, filme, documente, fotografii etc.

În constituirea acestei biblioteci digitale mondiale, un rol important va reveni mediului academic, bogat în resurse educaționale, de cercetare și care are deja o mare parte din fondul info-documentar digitizat.

Fiind o tendință nouă în procesul de furnizare a rezultatelor cercetărilor academice, depozitele digitale ar trebui să fie implementate în cadrul câtor mai multe instituții, pentru ca astfel toți cei implicați în munca de cercetare să beneficieze de mijloace moderne de diseminare a cunoașterii.

Modelul propus în cadrul acestei faze porneste de la premisa unei stranse colaborari între patru universitati, care au aceeasi viziune referitoare la constituirea unui sistem de biblioteci digitale universitare, dar care au sisteme informatice diferite și de aceea trebuie să definească și să realizeze o platformă de integrare adaptabilă la nivelul fiecăreia dintre universitățile implicate.

7. Bibliografie

1. **BANCIU**, Doina. Proiecte de cercetare pentru dezvoltarea bibliotecilor digitale din Romania. În: *ABR-Revista Română de Biblioteconomie și Știința informării*, nr.2/2008, pag.123
2. **BANCIU**, Doina. Regăsirea unitară a informațiilor digitale în bibliotecă. În: *Conferința internațională de Biblioteconomie și știința informării*, 2007
3. **BAKKER**, Trix. Biblioteca virtuală. Consecințe asupra dezvoltării colecțiilor în Olanda. În: *Biblioteconomie: Culegere de traduceri prelucrate*, nr. 1/1998
4. **BIBLIOTECA NAȚIONALĂ A ROMÂNIEI**. *Studiu de fezabilitate privind digitizarea, prezervarea digitală și accesibilitatea on-line a resurselor bibliotecilor*, <http://www.bibnat.ro/dyn-doc/Studiu%20Fezabilitate/Studiu-de-fezabilitate-digitizare.pdf> [accesat în 27.03.2010]
5. **BIBLIOTECA UNIVERSITĂȚII TRANSILVANIA**. *Catalog de carte veche și rară*, http://www.unitbv.ro/faculties/biblio/carte_veche/index.htm [accesat în 25.03.2010]
6. **BRACKENBURY**, Simon. At the University Library of Southampton, the Digitizing Line digitises rare and historical parliamentary documents. În: *DigiBook Mag*, nr. 9/2004, pag. 1, accesibilă online la http://www.i2s-digibook.com/upload/digibook_mag_no9_en.pdf
7. **Calimera Guidelines**. *Cultural Applications: Local Institutions Mediating Electronic Resources*, <http://www.calimera.org/Lists/Guidelines/Digitisation.htm> [accesat 11.03.2008]
8. **CHIVU**, Cristina. *Noile tehnologii informaționale: provocare și necesitate în managementul dezvoltării colecțiilor*, http://www.bcub.ro/continut/unibib/cristina_chivu.php#7 [accesat în 14.04.2010]
9. **CRISTEA**, Luciana; **REPANOVICI**, Angela. Tehnologia informației-motorul trecerii la biblioteca digitală. În: *Conferința internațională de Biblioteconomie și știința informării*, 2007

- 10. CIOCOIU, Adina.** Noua tendință- biblioteca digitală. Prezentare și valorificare a patrimoniului cultural. În: *Biblioteca*, nr 11/12 2009, pag. 284 – 288
- 11. Comisia de specialitate pentru digitizare.** *Ghid de digitizare – pilonul tematic "BIBLIOTECI"*, http://www.bibnat.ro/dyn/doc/Ghid%20de%20digitizare_Pilonul%20tematic_Biblioteci_versiunea01_25_11_2009.pdf [accesat în 27.03.2010]
- 12. ERICH, Agnes.** Biblioteca digitală: cale de transfer dinspre biblioteca tradițională spre biblioteca virtuală. În: *ABR-Revista Română de Biblioteconomie și Știința informării*, nr.2/2007, Pag. 11 - 14
- 13. FRIPTU, Mihaela.** Copyright-ul documentelor electronice. În: *Buletinul ABIR*, Vol. 14, nr. 1/2003
- 14. GARNES, Kari; LANDOY, Ane; REPANOVICI, Angela.** *Aspects of the Digital Library : Alvheim & Eide Akademisk.* Norway Forlag: Norway, 2006
- 15. HARVAT, Săluc.** *Introducere în biblioteconomie.* București: Editura Grafoart, 1996
- 16. KELLER, Michael.** Route of a world unique automatic digitizing device: the DIGITIZING LINE. În: *DigiBook Mag*, nr. 8/2004, pag.4, accesibilă online la http://www.i2s-digibook.com/upload/digibook_mag_no8_en.pdf
- 19. OSTAFE, Doina.** Biblioteca viitorului- colecțiile digitale. În: *ABR-Revista Română de Biblioteconomie și Știința informării*, nr.2/2007, pag15 – 18
- 20. PERIȘINARU, Aurelia.** Digitizarea colecțiilor- o prioritate pentru bibliotecile romanesti. În: *ABR-Revista Română de Biblioteconomie și Știința informării*, nr.2/2008, pag. 57-58
- 22. REPANOVICI, Angela.** Argument. În: *ABR-Revista Română de Biblioteconomie și Știința informării*, nr.3 2007
- 23. REPANOVICI, Angela, MAILAT, Gabriela.** Conservarea, protejarea și diseminarea documentelor vechi pe suport electronic. În: *Buletinul AGIR*, nr. 4/2003.
- 24. REPANOVICI, Angela.** *Informația științifică. Tehnici de documentare.* Brașov: Editura Infomarket,2000

- 25. REPANOVICI, Angela.** Proiect de marketing al Bibliotecii. Cartile vechi și rare: Păstrarea vie a memoriei. În: *Bulletin of the Transilvania University of Brasov*. Vol. 11(46) – New Series, Series B. 2004
- 26. REPANOVICI, Angela.** Quality of electronic publishing. În: *Bulletin of the Transilvania University of Braşov*. vol.10(50)-New Series. Series B.2004
- 27. REGNEALĂ, Mircea.** Copyright-ul în preocupările IFLA. În: *Studii de Bibliologie și Știința Informării*, nr. 2/1996. București: Editura Universității, 1997, pag. 45
- 28. REGNEALĂ, Mircea.** *Dicționarul explicativ de biblioteconomie și știința informării*. București: FABR, 2001. vol.I; vol.II
- 29. REGNEALĂ, Mircea.** *Studii de biblioteconomie*. Constanța: Ex Ponto, 2000
- 30. SOROIU, Claudiu.** *Istoria tehnologiei informației*,
http://www.ginfo.ro/revista/15_1/profil.pdf [accesat în 09.05.2010]
- 31. STAN, Laurențiu.** *Proiect privind integrarea documentelor pe suport electronic în oferta documentara a Bibliotecii Facultatii de Drept*,
http://www.bcub.ro/continut/unibib/integrarea_doc_electronice.php [accesat în 24.04.2010]
- 32. TÎRZIMAN, Elena.** Digitizarea documentelor de bibliotecă-Perspective europene și naționale. În: *Biblioteca*, nr. 5/2008, pag.143-144
- 34. ZĂNESCU, Gabriel.** Accesul la informație, conservarea documentelor și hotărârea de a digitiza. În: *Conferința internațională de Biblioteconomie și știința informării*, 2007
- 35. BARUCH, Zoltan Francisc.** *Arhitectura calculatoarelor*. Cluj-Napoca: Editura **TODESCO, 2000**

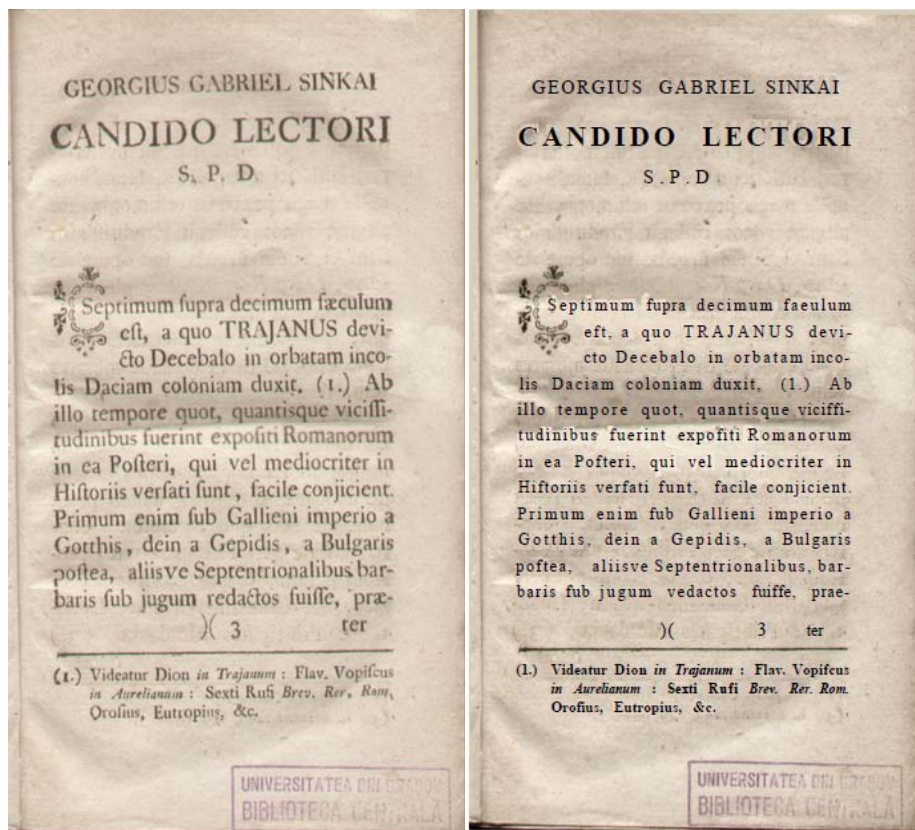
Webgrafie

1. http://blogs.usd.edu/library/entry/aquabrowser_library_catalog Aqoa Browser
2. <http://en.wikipedia.org/wiki/Ogg>
3. http://en.wikipedia.org/wiki/Free_Lossless_Audio_Codec
4. http://en.wikipedia.org/wiki/.mov#QuickTime_file_format
5. http://en.wikipedia.org/wiki/Audio_Video_Interleave
6. <http://en.wikipedia.org/wiki/FourCC>

7. http://en.wikipedia.org/wiki/RIFF_%28File_format%29
8. http://en.wikipedia.org/wiki/Moving_Picture_Experts_Group
9. <http://futurelib.pbworks.com/w/page/13686653/Examples>
10. <http://msdn.microsoft.com/>
11. <http://piermontlibrary.org/whats-new/aquabrowser-library-new-online-catalog/>
Aqoa Browser
12. <http://pumanet.codeplex.com/>
13. http://www.adobe.com/devnet/flv/pdf/video_file_format_spec_v10.pdf
14. http://websrv.bcut.ro/dyn_img/ExLibrisABR.pdf PRIMO
15. http://www.asro.ro/romana/standard/scopstd/scopstd_1.html
16. <http://www.digitalsquad.ro/>
17. <http://www.cvisiontech.com/>
18. <http://www.htmlgoodies.com>
19. <http://www.musitek.com/X/default.html>
20. <http://www.serialssolutions.com/aquabrowser/>
21. <http://www.thorpe.com.au/products/AquaBrowser.htm>
22. [http://www.unitbv.ro/LinkClick.aspx?fileticket=qJ2wBS9akhk%3D&tabid=4516
&language=en-US](http://www.unitbv.ro/LinkClick.aspx?fileticket=qJ2wBS9akhk%3D&tabid=4516&language=en-US)
23. <http://www.w3.org>
24. <http://www.xml.com>

ANEXE

ALTE IMAGINI ALE CARTILOR DIGITIZATE



PRAEFATIO

ter Jornandem (2.) sexcenti alii, illo partim antiquiores, partim posteriores, Historici testantur: quorum nonnulla magis praecipua testimonia ante aliquot annos collegit Eruditissimus Samuel Klein in Ms. suo opere de *Orig. Daco-Romanorum*; plura ego in adversaria mea toto illo quinquennio, quo Romae in Celeberrimo Urbano Collegio de Prop. Fide moratus sum, retuli.

Hujuscemodi servitute plurimum corruptam fuisse Latinam Majorum nostrorum Linguam, ecquis inficias ire audeat? corruptam tamen prout nunc est falso quis adfereret. Constat enim e Scriptoribus rerum Daco-Romanarum, & praesertim e *Chronico Ms. Constantis Miron, Logothetae Principatus Moldaviae*, (3.) quod

(2.) De rebus Geticis.
(3.) In Alexandrum I. Cap. 18. §. 1.

PRAEFATIO

ter Jornandem (2.) sexcenti alii, illo partim antiquiores, partim posteriores, Historici testantur: quorum nonnulla magis praecipua testimonia ante aliquot annos collegit Eruditissimus Samuel Klein in Ms. suo opere de *Orig. Daco-Romanorum*; plura ego in adversaria mea toto illo quinquennio, quo Romae in Celeberrimo Urbano Collegio de Prop. Fide moratus sum, retuli.

Hujuscemodi servitute plurimum corruptam fuisse Latinam Majorum nostrorum Linguam, ecquis inficias ire audeat? corruptam tamen prout nunc est falso quis adfereret. Constat enim e Scriptoribus rerum Daco-Romanarum, & praesertim e *Chronico Ms. Constantis Miron, Logothetae Principatus Moldaviae*, (3.) quod

(2.) De rebus Geticis.
(3.) In Alexandrum I. Cap. 18

CANDIDO LECTORI.

quod non adeo pridem legendum nobis concessit Eruditissimus aequae ac Doctissimus D. Franciscus Josephus Sulzer, atque etiam e *Descriptione Moldaviae* Serenissimi D. Demetrii Cantemir, ejusdem Provinciae quondam Authentis, & Principis, tum demum in barbarissimum, in quo nunc fumus, degenerasse Majores nostros, cum Litteralem Slavinorum, vel (si mavis) Illyricanorum Linguam in celebrandis SACRIS circa tempus Concilii Florentini fat malis avibus adhibere coeperunt. Quod ideo factum fuisse tradit Serenissimus Author, ut obsecundarent voluntati Archi-Praefulis civitatis Achridae in Bulgaria, qui instigatus a Marco (ut creditur) Ephesio intendebat hoc modo praeccludere nostris omnem aditum ad S. unionem cum Eccl. Romana.

) 4 Vere-

CANDIDO LECTORI.

quod non adeo pridem legendum nobis concessit Eruditissimus aequae ac Doctissimus D. Franciscus Josephus Sulzer, atque etiam e *Descriptione Moldaviae* Serenissimi D. Demetrii Cantemir, ejusdem Provinciae quondam Authentis, & Principis, tum demum in barbarissimum, in quo nunc fumus, degenerasse Majores nostros, cum Litteralem Slavinorum, vel (si mavis) Illyricanorum Linguam in celebrandis SACRIS circa tempus Concilii Florentini fat malis avibus adhibere coeperunt. Quod ideo factum fuisse tradit Serenissimus Author, ut obsecundarent voluntati Archi-Praefulis civitatis Achridae in Bulgaria, qui instigatus a Marco (ut creditur) Ephesio intendebat hoc modo praeccludere nostris omnem aditum ad S. unionem cum Eccl. Romana.

) 4 Vere-

CANDIDO LECTORI.

dem, ut Daco-Romanos in suam sententiam attrahat.

Nec spes eum fefellit sua. Re enim vera Alexander I. Moldaviae Authentis, & Princeps eidem morem gessit, & Linguam Slavico-Illyricam in SACRIS per totum suum Principatum adhiberi jussit. Hujus exemplum secuti fuere postea etiam alii Daco-Romanorum Dynastae, etsi cum duplici totius Nationis damno, nimirum cum longaeva ejusdem ab Ecclesia Latina scissione, & cum summo Linguae propriae detrimento: quorum primo ut *Deus Ter optimus* medelam aliquando adferre dignetur, ex animo precor; alteri ex parte adlatum iam est remedium, Linguam enim Slavicam, retentis caracteribus, paulo post, saeculo videlicet elapso, Majores nostri iterum expunxerunt: ex parte adferre conabimur, dum

CANDIDO LECTORI.

dem, ut Daco-Romanos in suam sententiam attrahat.

Nec spes eum fefellit sua. Re enim vera Alexander I. Moldaviae Authentis, & Princeps eidem morem gessit, & Linguam Slavico-Illyricam in SACRIS per totum suum Principatum adhiberi jussit. Hujus exemplum secuti fuere postea etiam alii Daco-Romanorum Dynastae, etsi cum duplici totius Nationis damno, nimirum cum longaeva ejusdem ab Ecclesia Latina scissione, & cum summo Linguae propriae detrimento: quorum primo ut *Deus Ter optimus* medelam aliquando adferre dignetur, ex animo precor; alteri ex parte adlatum iam est remedium, Linguam enim Slavicam, retentis caracteribus, paulo post, saeculo videlicet elapso, Majores nostri iterum expunxerunt: ex parte adferre conabimur, dum

PRÆF. CANDIDO LECTORI.

strum erit deinceps, navare operam, ut quæ huic Editioni defuerint, addamus, errata emendemus, & hanc ipsam Grammaticam unâ cum Dictionario Daco-Romano, quod (dummodo otium nobis, & scribendi commodum non deficit) componere fert animus, ad majorem perfectionem deducamus, publicoque usui utiliorem reddamus.

Tu interea, Candide Lector, hunc qualemcunque nostrum laborem, & fatigium æqui bonique consulito, & valeto.



ELE.

PRAEF. CANDIDO LECTORI.

frum erit deinceps, navare operam, ut quae huic Editioni defuerint, addamus, errata emendemus, & hanc ipsam Grammaticam una cum Dictionario Daco-Romano, quod (dummodo otium nobis, & feribendi commodum non deficit) componere fert animus, ad majorem perfectionem deducamus, publicoque usui utiliorem reddamus.

Tu interea, Candide Lector, hunc qualemcunque nostrum laborem, & farigium aequi bonique consulito, & valeto.



ELE



ELEMENTORUM
LINGUÆ DACO-ROMANÆ,
SIVE
VALACHICÆ
PARS PRIMA
DE ORTHOGRAPHIA.

CAPUT I.

De Litteris Daco-Romanis, earumque Pronuntiatione.

§. I.

Veteres Daco-Romanorum Litteræ eadem sunt, quæ & Latinorum, seu antiquorum Romanorum, a quibus originem ducunt suam. Videlicet:

Majuscula.

A. B. C. D. E. F. G. H. I. K. L. M. N.
O. P. Q. R. S. T. U. X. Y. Z. J. V.

A Mi-



ELEMENTORUM
LINGUÆ DACO-ROMANÆ,
SIVE
VALACHICÆ
PARS PRIMA
DE ORTHOGRAPHIA.

CAPUT I.

De Litteris Daco-Romanis, earumque Pronuntiatione.

§. I.

Veteres Daco-Romanorum Litteræ eadem sunt, quæ & Latinorum, seu antiquorum Romanorum, a quibus originem ducunt suam. Videlicet:

Majuscula.

A. B. C. D. E. F. G. H. I. K. L. M. N.
O. P. Q. R. S. T. U. X. Y. Z. J. V.

A Mi-

Minusculæ.

a. b. c. d. e. f. g. h. i. k. l. m. n. o. p.
q. r. s. t. u. x. y. z. j. v.

§. 2. Harum Litterarum potestatem, seu potius Pronuntiationem, a modernâ Latinâ pronuntiatione in quibusdam vocibus non nihil diversam, sequentes paginae, quoad fieri poterit absque viva Magistri voce, indicabunt. Itaque

A. a. Accentû acutô ' notatum, aut duplicatum sonat, sicut in Lingua Latinâ. e. g. *Câp*, caput. *Nâs*, nasus. *Câpra*, capra. *Lege* Cap, Nas, Capra.

Absque ullô Accentû positum pronuntiat, sicut Littera Ruthenica *а*, a Daco-Romanis pronuntiaci solet in ultimâ syllabâ hujus vocis *Аstrumen*, lumen.

Ab his regulis excipe *A.* ante MB, MP, & N. simplex positum, aut Accentû circumflexô notatum. In his enim circumstantiis *A* non pronuntiat, neque ut Accentû acutô notatum, neque ut sine Accentû positum, sed mollius, ac quasi per nasum. e. g. *Trambitia*, tuba. *Camp*, Campus. *Pane*, panis. *Targ*, nundinae. De quarum vocum pronuntiatione, nec non de Littera *а* in priori Regula citatâ præstabit consulere vivam Magistri vocem.

Hic tamen notes velim, priorem Exceptionem in vocibus peregrinis, præsertim Græcis propriis, atque in voce *Lancea*, lancea,

Minusculæ.

a. b. c. d. e. f. g. h. i. k. l. m. n. o. p.
q. r. s. t. u. x. y. z. j. v.

§. 2. Harum Litterarum potestatem, seu potius Pronuntiationem, a modernâ Latinâ pronuntiatione in quibusdam vocibus non nihil diversam, sequentes paginae, quoad fieri poterit absque viva Magistri voce, indicabunt. Itaque

A. a. Accentû acutô ' notatum, aut duplicatum sonat, sicut in Lingua Latinâ. e. g. *Cap*, caput. *Nas*, nasus. *Capra*, capra. *Lege* Cap, Nas, Capra.

Absque ullô Accentû positum pronuntiat, sicut Littera Ruthenica *Т*, a Daco-Romanis pronuntiaci solet in ultimâ syllabâ hujus vocis *Трумен*, lumen.

Ab his regulis excipe *A.* ante MB, MP, & N. simplex positum, aut Accentû circumflexô notatum. In his enim circumstantiis *A* non pronuntiat, neque ut Accentû acutô notatum, neque ut sine Accentû positum, sed mollius, ac quasi per nasum. e. g. *Trambitia*, tuba. *Camp*, Campus. *Pane*, panis. *Targ*, nundinae. De quarum vocum pronuntiatione, nec non de Littera *т* in priori Regula citatâ præstabit consulere vivam Magistri vocem.

Hic tamen notes velim, priorem Exceptionem in vocibus peregrinis, præsertim Græcis propriis, atque in voce *Lancea*, lancea,

