

ARHIVELE MUZEULUI - DOCUMENTELE ISTORICE DEPARTE DE A FI DOAR SURSE DE INFORMAȚIE

**Jon HOLMEN,
Christian-Emil ORE,
Ryvind EIDE**

PREAMBUL

Norvegia este divizată, pe criterii istorice, în cinci districte arheologice, fiecare dintre ele fiind reprezentată de un muzeu. Patru dintre aceste cinci muzee aparțin unei universități norvegiene. În 1991, facultățile de artă a patru universități norvegiene au inițiat un mare proiect numerizat, de colaborare (1992-1997), numit Proiectul Național de Documentare, cu scopul de a realiza baze de date și de a administra retro-conversia și numerizarea arhivelor analoge - cărți, imagini și alte tipuri de informații media - din departamentele de colecții ale facultăților, inclusiv din muzeele de arheologie. Proiectul de Documentare a fost continuat de un nou proiect, Proiectul Muzeu (1998-2005), centrat pe muzeele universitare (istoria culturală și naturală).

Proiectul Muzeu este motivat de ambiția de a dezvolta sisteme de tehnologie IC&T, care vor oferi utilizatorilor acces centralizat și eficient la informațiile legate de patrimoniul cultural și natural al Norvegiei. Cu ajutorul interfețelor și al hiperlegăturilor de informații către diferite domenii de studiu va fi posibil generarea de combinații informaționale și abordări noi, în discipline variate. Proiectul Muzeu se ocupă, de asemenea, de dezvoltarea sistemelor comune de baze de date privind managementul colecțiilor din toate muzeele universitare norvegiene. În mod ideal, aceste sisteme de baze de date ar trebui să fie capabile să

folosească toate informațiile de referință legate de colecțiile de artefacte și specimene, din muzee și din afara muzeelor, referitoare atât la istoria naturală, cât și la cea culturală.

Întrucât Norvegia nu a avut niciodată un muzeu național, multe colecții pe care, în mod normal, le găsim în muzee naționale în alte țări, se află aici în muzee universitare. În ceea ce privește arheologia, muzeele implicate în Proiectul Muzeu se ocupă de patru districte culturale. Astfel, dezvoltând un sistem comun de arhivare a informațiilor din aceste patru muzee și incluzând informații importate din al cincilea district, sistemul de bază de date devine un registru central arheologic al Norvegiei.

Unitatea de Documentare Numerică (DOK) este o unitate permanentă a Facultății de Arte a Universității din Oslo, o continuare a grupului central al Proiectului de Documentare. DOK se ocupă de dezvoltarea continuă a bazelor de date din domeniile lexicografiei, corpusurilor de texte, toponimie, medievală și istoria artei, dar și de partea de IC&T a Proiectului Muzeu.

Prin cooperare internațională, ca, de exemplu, proiectul Arena (<http://ads.ahds.ac.uk/arena/>), dezvoltăm, de asemenea, sisteme transnaționale de schimb de informație și de interconexiune între colecțiile naționale. Scopul pe termen lung este un sistem arheologic informațional global. Pe de altă parte, scopul trans-tematic al

proiectului nostru deschide, de asemenea, inte-grarea trans-tematică a acestor colecții naționale.

DOK are o experiență unică în dezvoltarea și aplicarea metodelor baze pe IC&T, într-o gamă largă de discipline ce studiază texte ale științelor umaniste. Infrastructura textului ne-a condus la aplicarea și la dezvoltarea mai departe a metodelor luate din domeniul editării textelor erudite, chiar atunci când lucrăm cu arhive și cataloage de muzeu. În timpul celor zece ani de lucru, am dezvoltat un sistem bine testat pentru crearea colecțiilor numerice bazate pe piese analoge (în principal, cele din hârtie). Eforturile noastre de a armoniza practica cu aceste linii de ghidaj sunt subiectul principal al acestui studiu.

O NOTĂ METODOLOGICĂ: EDIȚIILE TEXTELOR SCOLASTICE

În filologia tradițională, scopul este de a crea „cel mai bun text” dintr-un set de texte – martori. O ediție de texte scolare ar fi, în mod obișnuit, o versiune tipărită a „celui mai bun text”, însoțit de o variație de aparate critice, documentate în texte – martori și, poate, de un set de facsimile de (o parte din) manuscrise.

În ultimii 10-20 de ani, computerele au fost introduse în filologia textuală. Astăzi, o culegere electronică modernă de texte filologice poate să nu conțină în mod necesar „cele mai bune texte”. O „ediție de texte” poate fi o bază de date bibliografice, facsimile electronice și transcripse electronice ale textelor – martori, combinate cu câteva unelte în plus de căutare și cu texte explicative. Transcriptele sunt, de obicei, marcate în XML/SGML, ceea ce împiedică hiperconexiunile și vizionarea diferită a textelor.

Faptul că filologii sunt interesați de textele deținute de muzee, acest fapt, în mod normal, nu se trece

dincolo de texte, care sunt obiecte muzeale de primă clasă. Rareori ei consideră că textele ce fac, în mod normal, obiectul activității muzeelor, cum ar fi arhivele, rapoartele de săpături arheologice ș.a.m.d., îi pot interesa. În munca noastră, noi am mers în direcția opusă. Arhivele interne, rapoartele, achizițiile și alte cataloage sunt principala sursă în dezvoltarea bazelor de date științifice / scolare, pentru studenți, cercetători și muzeografi dintr-un muzeu. Drept rezultat, în ultimii zece ani am experimentat o procedură de lucru, în patru etape, pentru textele aflate în muzee, dezvoltată din și inspirată de contactul nostru apropiat cu alte discipline umaniste.

Primii trei pași corespund pașilor aplicați, în mod normal, edițiilor obișnuite de texte. Ultimul pas, ultima etapă, în care am creat o poziționare de la conținutul semantic al textului la o ontologie, nu este unul obișnuit folosit de editorii de text:

Primul pas este înregistrarea meta-informațiilor documentelor, cum ar fi Titlul, Autorul, Anul publicării, Numărul ediției etc. Cu alte cuvinte, noi creăm o înregistrare bibliografică a documentului original.

Următorul pas sau procesul de înregistrare este realizarea unei imagini pentru fiecare pagină a publicației. Rezultatul procesului este un facsimil electronic care ar trebui păstrat neschimbat și privit ca o importantă sursă în sine.

Apoi creăm un transcript, aplicând facsimilului procedeul OCR sau, dacă este vorba de un manuscris, transcrierea manuală a textului. Transcriptul este apoi marcat, indicând pagina, așezarea specială în pagină, fișele tip, ilustrația, inserările etc. Dacă nu este vorba de însemnări de mână în textul original, ar trebui, în principiu, să putem face o copie tipărită, care este mai mult sau mai puțin identică cu documentul original înregistrat. De

asemenea, vom executa un tip de lematizare, adică adăugăm termeni normalizați acolo unde aceștia variază în text, cum ar fi pronunții diferite ale numelor de locuri. Rezultatul va fi un text XML.

În fine, vom face un pas care va extinde modelul descriptiv minimal, introducând codări semantice ale conținutului. Aceasta înseamnă că vom căuta elementele informaționale din text și vom încerca să le codăm în relație cu o concepție dată sau cu ontologia.

În cazul nostru, ontologia este un set de entități și relații pe care le considerăm că pot descrie cel mai bine descoperirile arheologice și obiectele. Rezultatul este un text XML pe de-a întregul marcat, care poate fi accesat nu numai ca text liber, dar și ca o bază de date fundamentată pe un document XML. Mai departe, marcarea poate fi folosită pentru a extrage informația într-o bază de date relaționată mai tradițional.

CATALOAGELE DE ACHIZIȚII

De aproape 170 de ani, toate muzeele arheologice din Norvegia au publicat, anual, informații despre artefactele pe care le-au obținut, în

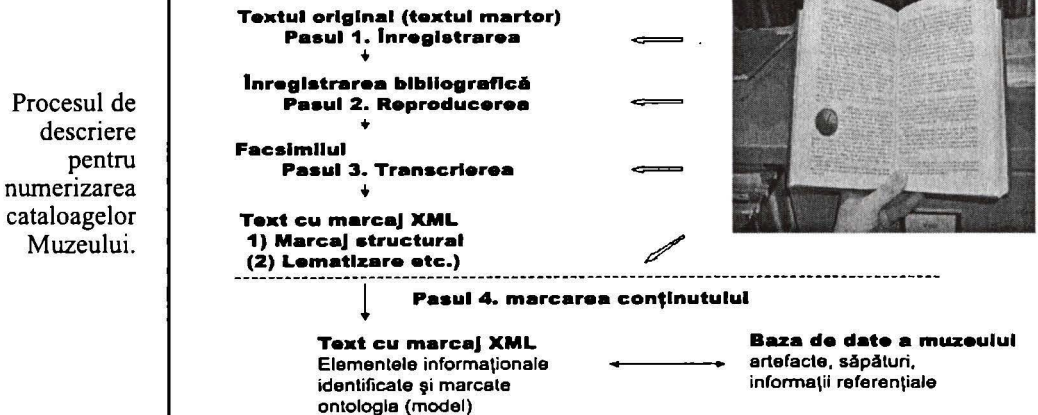
cataloge de achiziții special pregătite. Descrierea descoperirilor în aceste cataloge este aproape prolixă, incluzând informații extinse ale domeniului, contextul descoperirii, locul și timpul, descoperitorul sau arheologul, cât și descrieri detaliate și clasificări. Astfel, acest set de cataloge servește ca principal inventar pentru fiecare muzeu.

Când Proiectul de Documentare a pornit în 1992, s-a decis ca acesta să creeze o bază de date de artefacte, bazată pe catalogele tipărite și pe doar câteva adăugiri mai vechi, nepublicate.

În ce măsură s-a făcut procesarea acestor documente conform indicațiilor subliniate mai sus? Noi chiar am creat o înregistrare bibliografică pentru fiecare catalog, așa cum este descrisă în prima etapă. Din nefericire, facsimilele paginilor nu au fost păstrate datorită costului de depozitare la începutul anilor 1990. Catalogele au fost citite prin OCR și corectate față de original, iar toponimele și termenii de obiecte au fost lematizate.

Mai târziu, variantele electronice au fost marcate cu etichete pentru conținutul lor semantic, conform ontologiei noastre. Peste toate, noi ne-am apropiat de metodologia ideală.

Documentele muzeului - Ideal vs. Catalogele muzeului - realitate



<CAT AN>
<INTRO> <M NUME>
UNIVERSITETETS
OLDSAKSAMLING TILVEKST

<AN>1989</INTRO>

[...]

<NRPAR> <CATNR
nrId="37267">C.37267.

<INFORMAȚIE ARTEFACT>
<ARTEFACT> Topor </ARTEFACT>
de <MAT> fier </MAT>

din <EPOCA>EVUL MEDIU
TÂRZIU </EPOCA>.

<INFORMAȚIE
ARTEFACT><MĂSURĂ> L: 141mm
</MĂSURĂ>, <MĂSURĂ> capăt
W:109mm </MĂSURĂ>.

TOPOR DE DULGHER cu
<FORMĂ> cu formă specială
</FORMĂ> lamă special construită
pentru degete, fiind ținută chiar „în
spatele” capătului.

În special necesară pentru lucruri
fine, când este folosită într-un unghi
spre capăt, sau ca găuritor.

<DISTRIBUȚIE> descoperită
<LOCUL DESCOPERIRII> pe deal,

la aproximativ 300 m deasupra

<LOC> LROS KAPELL </LOC>.

<GOSPODĂRIE> SJRGLRRT

</GOSPODĂRIE>.

<DISTRICT D> <DISTRICT>

LROS </DISTRICT> </DISTRICT

D>, <MUNICIPALITATE> RRYKEN

</MUNICIPALITATE>.

<COMITAT> BUSKERUD

</COMITAT> </LOCUL

DESCOPERIRII>, în 1959 de

<GĂSITOR> Berge Narvik

</GĂSITOR>.

Tjernsrudveien 24, Jar, Oslo.

</DISTRIBUȚIE> </NRPAR>

Cataloagele codate rezultate au
fost indexate prin cuprinsul etichetelor

și pot fi căutate ca texte independente.

În plus, cuprinsul etichetelor a fost

importat într-o bază de date, baza de

date de regărire, ca versiune originară

a clasificării obiectelor. Un total de

Documentele muzeului - ideal vs. Arhivele topografice - realitate

Textul original (textul martor)
Pasul 1. Înregistrare

Înregistrarea bibliografică
Pasul 2. Reproducerea

Facsimil
Pasul 3. transcrierea

Text cu marcaj XML
1) Marcaj structural
(2) Lematizare etc.)

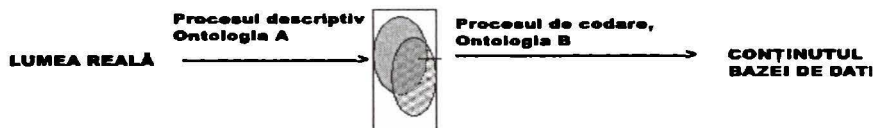
Pasul 4. marcarea conținutului

Text cu marcaj XML
Elemente informaționale
identificate și marcate
ontologie (model)



Exemplu de
paragraf de
catalog
complet
codat.

PROBLEMA CODĂRII CONȚINUTURILOR:



SOLUȚII PE MATERIALE VECHE:

Documente istorice;
Folosește gama largă de informații derivate
datorită metodologiei folosite

MATERIALE NOI:

Păstrează, de asemenea, standardul folosit

lor direct într-o bază de date. Aceasta, desigur, ne-a furnizat o referință mai puțin concretă a documentelor.

Fiecare document a fost grupat în cadrul unei singure activități istorice și a fost organizat într-o secvență cronologică în cadrul cazului. De asemenea, am făcut referințe la toate gospodăriile relevante, la descoperiri și vechi monumente din documente. În baza de date, am creat un index de căutare, atât pentru informațiile de arhivă, cât și pentru informațiile de conținut, cât și pentru facsimile. Un total de aproximativ 200.000 de pagini de documente au fost procesate în acest mod.

SCHIMBARE ONTOLOGICĂ

Ontologia a fost și este un termen filozofic referitor la studiul ființei. Comunitatea inteligenței artificiale tinde să o interpreteze ca o specificare a conceptualizării sau, mai simplu, un model al anumitor părți ale realității. În știința informaticii, termenul de „ontologie” este deseori sinonim cu „model de date”.

Materialul arhivistic acumulat în ultimii 180 de ani ilustrează schimbări ontologice în înțelesul pur filozofic, datorită viziunii proprii a cercetătorilor asupra lumii și datorită viziunii

implicite asupra trecutului. Viziunea cercetătorilor asupra scrierilor colegilor lor din trecut se schimbă de-a lungul timpului și este o schimbare corespunzătoare în ceea ce este considerat a fi important într-un raport vechi.

Etapa a patra a metodologiei, codarea conținutului semantic, va reflecta în mod necesar viziunea noastră asupra a ceea ce constituie o importantă informație arheologică. Aceasta nu va coincide totalmente cu concepția autorului de la sfârșitul secolului 19, autor al materialului original. Autorul includea informații conform concepției sale, din care noi codăm sub-setul care corespunde concepției sau ontologiei noastre.

Pentru că materialul vechi a fost creat în ultimii 180 de ani, el variază semnificativ atât în ceea ce privește tipul de informație care este pus în evidență, cât și în ceea ce privește stilul.

Metodologia folosită în convertirea acestui material ne furnizează informații adiționale asupra acestui aspect. Au fost generate registre de persoane care au jucat un rol important în istoria muzeelor. Avem, de asemenea, secvențele activităților din care este alcătuită acesta istorie,

Biți de informații identificate din Documentul de Arhivă, exemplu

dându-ne ocazia să cunoaștem în ce măsură fiecare „stil” individual influențează conținutul. Prin relaționarea tuturor acestora vom fi capabili să relatăm istoria instituțiilor.

Odată cu materialele și activitățile noi ale muzeelor, am creat un standard mult mai unitar. Dar pentru că acest tip de standard trebuie să exprime o formă a unei școli arheologice și care, probabil, va fi schimbat în viitor, am dezvoltat un sistem în care standardele, reprezentate ca definiții ale formei/DTD și taxemele, sunt conservate și salvate.

Noile materiale incluse în ultima fază vor face astfel referire la un standard specific care este prezent în baza de date, împreună cu informații, cum ar fi autorul și data. O consecință a acesteia este că atunci când actualizăm standardul, nu este nevoie să actualizăm materialul mai vechi, care a fost introdus conform unui standard vechi. Va fi nevoie de o aranjare între standardele noi și cele vechi.

Aceasta va asigura conservarea informațiilor în forma lor originală și ne va asigura că nici o informație nu va fi pierdută în timpul actualizării. Ne va da, astfel, nouă și viitorilor utilizatori ai sistemelor, o viziune generală asupra

dezvoltării arheologie ca știință și va arăta cum sunt scrise diferite rapoarte de către fiecare școală în parte.

CONCLUZII

Prin procesul convertirii cataloagelor și arhivelor de pe suport de hârtie, din muzeele arheologice, în formă numerică și, pe cât posibil, ținând cont de principiile noastre, definite pentru acest tip de conversie, noi am creat baze de date relaționate, de încredere, ale arhivelor arheologice și ale informațiilor despre artefacte. Am dezvoltat, de asemenea, ocazia de a lucra direct cu materialele – surse, prin realizarea unei arhive electronice de text, cu indexuri clare, atât pentru cataloage, cât și pentru documente. Aceasta face posibilă studierea lor ca obiecte istorice individuale.

Prin codificarea materialului-sursă într-o ontologie comprehensivă noi, în plus față de funcționalitatea „normală” a bazei de date, am creat terenul pentru oportunitatea de a relata istoria muzeelor și istoria arheologiei norvegiene. De asemenea, am stabilit un sistem în care principiile, pe care arheologii de astăzi le folosesc în munca lor de documentare, pot fi continuu înregistrate, creându-se, astfel, ceea ce va fi, în viitor, istorie.

Traducere: Mihaela Murgoci

ABSTRACT

This paper describes some methodological aspects of creating electronic versions of 200 years of paper-based documentation in the archaeological university museums in Norway. We focus on a few selected methods and positive and negative

experiences during this work as well as some aspects of the relationship between these documents as important parts of the non-material cultural heritage in the museums and the material heritage described in them.