INFORMATIZAREA DICŢIONARULUI LIMBII ROMÂNE – UN VIS PE CALE DE REALIZARE

THE CYBERNATION OF THE DICTIONARY OF ROMANIAN LANGUAGE – A DREAM TURNING TO REALITY

(Abstract)

The article presents the general framework of the project of cybernation of the *Dictionary of Romanian Language*. The decision of the Board of the Romanian Academy to resume work on the first part of the Dictionary (DA) according to the norms of the new series (DLR) constituted a fitting opportunity to start work on writing the academic *Dictionary* in an XML format. The digitalization of the dictionary is accomplished with the help of the Oxygen program put forward by a computer specialist who achieved the work interface. This will offer the possibility of the dictionary's being posted on the internet.

Cuvinte-cheie: dicţionar, cibernetică, programul oxygen, interfaţă, câmpuri de redactare.

Key-words: dictionary, cybernation, oxygen program, interface, section of the writing formula.

Obiective

La Institutul de Lingvistică al Academiei Române „Iorgu Iordan–Al. Rosetti" urmează să se redacteze, în urma hotărârii Prezidiului Academiei Române, prima parte a *Dicţionarului limbii române* după normele celei de a doua părţi.

Dicţionarul limbii române a constituit, alături de gramatică şi ortografie, unul dintre cele trei deziderate istorice ale înfiinţării Societăţii Academice Române, devenită ulterior Academia Română. El este cunoscut sub numele de *Dicţionarul Academiei,* întrucât este elaborat sub egida acestei prestigioase instituţii.

BDD-V2853 © 2015 Editura Univers Enciclopedic Gold
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 22:59:08 UTC)

Dicţionarul limbii române este un dicţionar istoric, explicativ, etimologic şi normativ şi este denumit şi dicţionarul *tezaur,* graţie marii sale bogăţii lexicale. Redactarea primei ediţii a început în anul 1906 şi a fost finalizată în anul 2010. Prima parte, elaborată între 1906 şi 1949, cuprinde literele *A–L*, iar a doua parte, numită Serie nouă, literele *M–Z*, precum şi literele neredactate sau nefinalizate (*E, D* şi *L*). *Dicţionarul limbii române* a fost publicat, între 1913 şi 2010, într-un număr de 37 de volume, însumând peste 175.000 de cuvinte şi variante în cca 18.000 de pagini.

Cu sprijinul Băncii Naţionale a României, în anul 2010, a fost retipărită o ediţie anastatică a întregului dicţionar, în 19 tomuri, în condiţii grafice excelente.

O tehnologie nouă. O structură nouă

În noul context al societăţii informaţionale, se impune cu necesitate ca noua ediţie a *Dicţionarului limbii române* să fie redactată în format XML, pentru a putea oferi posibilitatea consultării lui pe internet şi a intra astfel în circuitul valorilor mondiale.

Digitalizarea noii ediţii a *Dicţionarului* se va face în cadrul unui proiect de anvergură cu ajutorul programului Oxygen, produs în România.

Institutul a luat legătura cu compania producătoare a programului Oxygen, un program de editare XML, bine cunoscut şi foarte apreciat, cu scopul de a îndeplini obiectivul convertirii formatului nestructurat al dicţionarului într-un format structurat.

În cursul acestei operaţiuni, institutul a hotărât adoptarea unor standarde tehnice larg răspândite şi acceptate, bazate pe XML, respectiv TEI (Text Encoding Initiative = Iniţiativa de codare a textelor), adecvate pentru redactarea dicţionarului.

Instrumente noi

Programul Oxygen permite redactorilor să utilizeze abordări accesibile pentru utilizator în ceea ce priveşte editarea documentelor în format XML, respectiv folosirea de interfeţe similare formularelor web.

Această modalitate de editare este extrem de utilă şi de facilă, întrucât redactorii pot astfel să adauge conţinut folosind un formular şi nu un arbore XML, a cărui structură ar trebui să o cunoască, aşa cum este cazul pentru majoritatea editoarelor XML existente.

Perioada de învăţare a modului în care se realizează redactarea dicţionarului este foarte scurtă, interfaţa este uşor de utilizat şi folosirea sa a întrunit consensul tuturor redactorilor care îl vor utiliza.

Această abordare ne-a permis să satisfacem chiar şi cele mai sofisticate doleanţe, prin crearea unui cadru specializat pentru conţinutul dorit al *Dicţionarului,* un cadru care conţine module de redactare, fişiere de adaptare a interfeţei utilizatorului şi clase Java pentru funcţii specializate. Dintre aceste

BDD-V2853 © 2015 Editura Univers Enciclopedic Gold
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 22:59:08 UTC)

doleanţe enumerăm: necesitatea de căutare şi consultare a serverului dicţionarului pentru interogări complexe privind conţinutul în curs de redactare, adăugarea de fragmente de conţinut şi de şabloane de conţinut în funcţie de valorile selectate de utilizator din liste derulante, componente noi pentru formularul de redactare, altele decât cele oferite în mod obişnuit de programul Oxygen, redarea datelor în formatele dorite etc.

Redactare versus redare

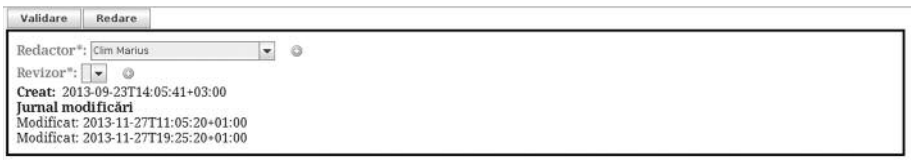
Elaborarea fomularului de redactare s-a făcut, pe cât a fost posibil, cu respectarea ordinii secţiunilor unei intrări în dicţionar.

Cu toate că redactarea folosind acest formular nu este identică cu redactarea anterioară, care foloseşte editoare de text obişnuite, noi considerăm că formularul este un progres şi un avantaj, din mai multe considerente, dintre care enumerăm următoarele:

1. conţinutul rezultat este structurat din punct de vedere semantic, nu doar stilistic, aşa cum este cazul când conţinutul este redactat cu editoare de text;
2. conţinutul structurat poate fi redat în formate multiple,: HTML (pagină web), PDF, EPUB (format de carte electronică) etc.;
3. structura conţinutului poate fi oricând modificată, după necesităţi;
4. se pot face căutări complexe în conţinut, lucru care nu este posibil în cazul conţinutului nestructurat obţinut cu editoare de text;
5. se pot înfiinţa servicii care să furnizeze conţinut în orice format dorit, pentru terţi;
6. redactarea este uşurată prin folosirea de liste derulante care conţin valori predefinite, standardizate;
7. în cursul introducerii de conţinut sau în cursul etapei de validare a conţinutului, există posibilitatea verificării existenţei unor corelaţii complexe, atât între diferitele elemente ale intrării curente, cât şi între elemente ale intrării curente şi elemente ale altor intrări din dicţionar.

Prezentarea secţiunilor formularului de redactare

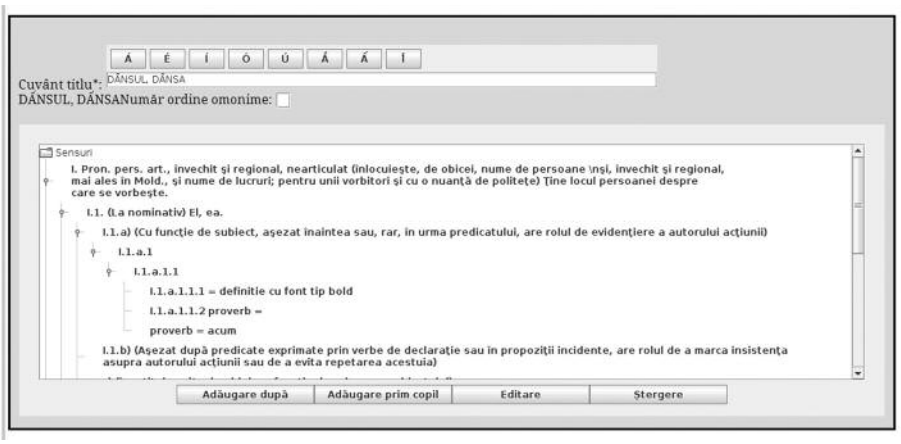
Secţiunea cu metadate ale intrării



Această secţiune permite introducerea numelor redactorilor şi a revizorilor care au contribuit la redactarea intrării şi afişează data creării intrării, precum şi un jurnal cu data şi ora modificărilor intrării, validate prin apăsarea butonului Validare.

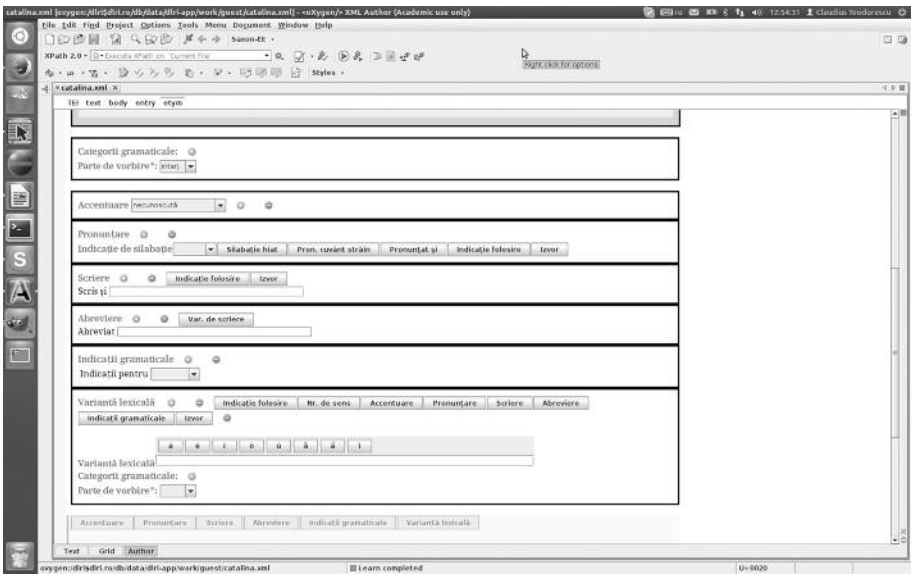
BDD-V2853 © 2015 Editura Univers Enciclopedic Gold
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 22:59:08 UTC)

Secţiunea cu cuvântul-titlu şi cu arborele de sensuri



Această secţiune permite introducerea cuvântului-titlu şi a numărului de omonim, precum şi redarea în formă sintetică a arborelui de sensuri ale intrării, redare care permite o viziune de ansamblu asupra sensurilor, precum şi diferite operaţii asupra lor, cum ar fi adăugarea, ştergerea sau editarea de sensuri.

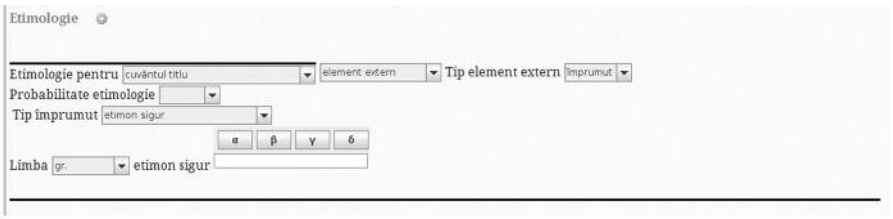
Secţiunea cu informaţii despre cuvântul-titlu



BDD-V2853 © 2015 Editura Univers Enciclopedic Gold
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 22:59:08 UTC)

Această secţiune permite selectarea categoriei gramaticale a cuvântului-titlu şi, după caz, introducerea de informaţii privind formele de accentuare, de pronunţare, de scriere, de abreviere, indicaţiile gramaticale şi variantele lexicale ale cuvântului-titlu.

Secţiunea etimologie



Această secţiune permite clasificare etimologiei cuvântului-titlu, precum şi introducerea de alte informaţii etimologice.

Structura de clasificare etimologică este de departe cea mai complexă dintre toate secţiunile unei intrări, dar credem că s-a reuşit obţinerea unei interfeţe „prietenoase”, prin folosirea de liste derulante cu tipuri de clasificări şi prin introducerea de conţinut specializat o dată cu selectarea unei intrări dintr-o asemenea listă.

Concluzii

Considerăm ca fiind excelentă iniţiativa colectivului redacţional de a continua redactarea acestui dicţionar într-un format structurat, folosind standarde consacrate în domeniu, şi suntem siguri că în viitor multiple reuşite vor fi posibile, printr-o cercetare inteligentă a datelor astfel structurate şi prin corelarea acestor date cu datele structurate conţinute în alte dicţionare, fie româneşti, fie străine.

Monica BUSUIOC
Institutul de Lingvistică al Academiei Române
„Iorgu Iordan – Al. Rosetti”, Bucureşti
Claudius TEODORESCU

BDD-V2853 © 2015 Editura Univers Enciclopedic Gold
Provided by Diacronia.ro for IP 216.73.217.142 (2026-11-14 22:59:08 UTC)