

## REZUMATUL PROIECTULUI

In contextul actual al dezvoltarii economiei bazate pe cunoastere, problema elaborarii de continut electronic (eContent) si a gestionarii acestuia devine esentiala. Continutul poate fi procesat automat, de catre sisteme IT inteligente pentru a furniza cunostinte, doar daca datele inglobate sunt imbogatite semantic. Tema propusa are un caracter interdisciplinar si se incadreaza in directia de cercetare **1.2.6 Sisteme informatice pentru gestiunea si dezvoltarea de continut digital interactiv (e-content)**.

Proiectul isi propune studiul si dezvoltarea unui sistem integrat de generare, gestiune si procesare de continut digital provenit din surse arhivistice folosind ontologii de domeniu si adnotari semantice. Pe baza continutului generat si adnotat se vor putea regasi informatii relevante semantic. Proiectul isi propune si realizarea unui subsistem de rationare asupra continutului semantic stocat pentru identificarea de corelatii noi in vederea imbogatirii cunostintelor de domeniu.

Sistemul propus include atat achizitia de date din surse arhivistice primare sau din biblioteci digitale externe existente, cat si structurarea acestor folosind un model propriu de date care sa permita adnotarea semantica. In vederea adnotarii semantice se va studia si dezvolta o ontologie de domeniu. Informatiile vor fi transformate in cunostinte prin procese complexe de rationare folosind ontologia de domeniu si adnotarea semantica. Pentru exploatarea cunostintelor, ne propunem sa dezvoltam algoritmi performanti de cautare ghidata de ontologie in spatiul cunostintelor in vederea regasirii de informatie relevanta semantic pe de o parte, si mecanisme de rationare in vederea cautarii de noi informatii (information mining) pe de alta parte. In cadrul proiectului se vor dezvolta si oferi servicii unei game largi de utilizatori incluzind cercetatori si administratori de arhive, studenti, agenti software si publicul larg. Astfel:

- Cercetatorii (istorici) vor avea la dispozitie unele software de regasire de date, informatii si cunostinte in care interogarea se va face in limbaj natural.
- Administratorii (arhivisti) vor avea la dispozitie unele de rationare bazate pe algoritmi de information mining pentru identificarea si analiza de corelatii / informatii / cunostinte noi, imbogatind astfel intr-o maniera interactiva cunostintele specifice de domeniu precum si ontologiile de domeniu.
- Pentru cautarile programatice prin intermediul agentilor software se vor oferi servicii Web adnotate

semantic precum si ontologia de adnotare.

- Publicul va putea accesa sistemul prin intermediul unei interfete Web, folosind interogari bazate pe limbaj natural.

Pentru realizarea sistemului ArhiNET se vor studia si elabora modele de reprezentare datelor complexe in format multimedia (text, voce, film) si a cunostintelor, modele de reprezentare a ontologiilor, algoritmi de regasire a informatiilor adnotate semantic, modele de rationare asupra cunostintelor.

Sistemul propus pentru realizare în cadrul proiectului ArhiNET va avea valente multiple fiind util atat institutiilor detinatoare de continut de interes general (arhive, muzee, biblioteci, institutii de administratie publica, etc.), cercetatorilor si publicului interesat de accesul la acest continut. Mai mult, acest sistem permite integrarea continutului cultural, istoric de interes in patrimoniul european.