

Sinteză activități sub-proiect 5 : Dezvoltarea experimentală a unei ontologii specifice limbii române și testarea efectelor căutărilor bazate pe web-ul semantic la nivelul rețelelor sociale online” (Semantic Media)

-Sinteză 2019-

Activitatea 5-2-1: Formularea cerințelor funcționale și non funcționale ale ontologiei specifice limbii române

O definiție clară a cerințelor funcționale și non-funcționale reprezintă o condiție esențială pentru dezvoltarea cu succes a oricărui produs software. În cazul ontologiei FutureWeb au fost definite șase cerințe funcționale și douăsprezece cerințe non-funcționale. Cerințele funcționale (eng. functional requirements) definesc funcționalitățile unui sistem sau ale componentelor acestuia, fiecare funcționalitate fiind descrisă prin specificarea comportamentului sistemului între ieșiri și intrări. Astfel, în procesul de dezvoltare al ontologiei propuse s-au avut în vedere următoarele cerințe funcționale: **CF1**. Ontologia va permite identificarea conceptelor generale în raport cu un anumit concept, **CF2**. Ontologia va permite identificarea conceptelor specifice pentru un anumit concept, **CF3**. Ontologia va permite identificarea conceptelor aflate în relație cu un anumit concept, **CF4**. Ontologia va permite identificarea conceptelor utilizând termeni diferiți, **CF5**. Ontologia va facilita identificarea conceptelor utilizând abrevierile uzuale, **CF6**. Ontologia va permite identificarea conceptelor în cazul unor erori uzuale de scriere. Cerințele non-funcționale (eng. nonfunctional requirements) au în vedere diverse tipuri de constrângeri impuse asupra sistemului, ortogonale față de cerințele funcționale. În cazul ontologiilor, cerințele non-funcționale vizează caracteristicile, calitățile sau aspectele generale ale ontologiei, precum ușurința în utilizarea, scalabilitate și mentenanța ulterioară (Noy and McGuinness, n.d.; Allemang and Hendler, 2011; Horridge, n.d.; Bergman, 2010b), care nu au legătură cu conținutul efectiv al acesteia (Suárez-Figueroa, Gómez Pérez and Villazón-Terrazas, 2009). În cazul ontologiei propuse, s-a avut în vedere satisfacerea următoarelor cerințe non-funcționale: **CNF1**. Utilizarea de concepte și proprietăți din ontologii recunoscute sau standardizate de W3C, **CNF2**. Ontologia trebuie să poată fi extinsă pentru alte limbi, **CNF3**. Utilizarea unei structuri modulare, **CNF4**. Respectarea unor convenții pentru denumirea claselor și a proprietăților, **CNF5**. Versionarea ontologiei, **CNF6**. Descrierea ontologiei utilizând sintaxa Turtle, **CNF7**. Utilizarea unor prefixe pentru ontologie și modulele acesteia, **CNF8**. Separarea între TBox și ABox, **CNF9**. Declararea claselor ca fiind disjuncte, **CNF10**. Declararea de la proprietăți inverse, **CNF11**. Includerea de etichete și descrieri pentru clasele și proprietățile definite și CNF12. Evitarea inconsistențelor în ontologie.

Activitatea 5-2-2: Dezvoltarea unui model (prototip) experimental al ontologiei specifice limbii romane la nivelul unei rețele sociale online

Dezvoltarea prototipului ontologiei a avut în vedere îndeplinirea cerințelor funcționale și non funcționale, ținând cont de bunele practici pentru crearea de ontologii și de metodologiile din literatura de specialitate. Ontologia dezvoltată își propune să modeleze pe de o parte conceptele relevante pentru rețeaua socială FutureWeb, precum conturile de utilizatori sau mesajele publicate, iar pe de altă parte noțiunile din domeniul dezvoltării software, de interes pentru studenții de la facultățile de profil. Ontologia propusă pentru este denumită în restul documentului *Future Web* și utilizează prefixul fw, precum și URI-ul (Uniform Resource Identifier) <http://www.futureweb.ro/fw>

Ținând cont de bunele practici din domeniu, s-a optat pentru dezvoltarea etapizată a ontologiei: Etapa 1. Stabilirea domeniului și a obiectivelor ontologiei, Etapa 2. Analiza ontologiilor existente, Etapa 3. Alcătuirea unei liste cu termenii care vor apărea în ontologie, Etapa 4. Definirea claselor și a ierarhiei de clase, Etapa 5. Definirea proprietăților claselor, Etapa 6. Definirea tipurilor proprietăților, Etapa 7. Crearea instanțelor. Pe parcursul dezvoltării modelului experimental de ontologie s-a verificat respectarea tuturor cerințelor funcționale și non-funcționale.

Activitatea 5-2-3: Realizarea metodologiei de testare și evaluare a prototipului dezvoltat

În cadrul metodologiei de testare și evaluare a prototipului de ontologie dezvoltat au fost identificați următorii pași necesari: **Pasul 1.** Realizarea unei fișe de testare în care sunt incluse întrebări cu privire la corectitudinea rezultatelor returnate de platformă prin compararea acestora cu datele din ontologie. În acest sens, în fișa de testare vor fi incluse explicarea scopului platformei, elemente precum conceptele cheie care vor fi testate, pașii necesari a fi urmați, rezultatele așteptate pentru fiecare concept cheie, pentru fiecare funcționalitate în parte, modalitatea de înregistrare a îndeplinirii fiecărui element testat și de descriere a situațiilor în care platforma nu se comportă conform specificațiilor. **Pasul 2.** Selectarea unei echipe de patru persoane cu competențe în testarea de aplicații web în vederea testării platformei și completării fișei de evaluare. **Pasul 3.** Parcurgerea împreună cu membrii echipei de testare a fișei de evaluare și clarificarea eventualelor nelămuriri cu privire la conținutul acesteia. **Pasul 4.** Testarea efectivă a platformei de către echipa de testare și înregistrarea răspunsurilor pe fișele de testare individuală. **Pasul 5.** Colectarea și analiza răspunsurilor înregistrate pentru fiecare element din fișa de testare. **Pasul 6.** Realizarea unui raport cu privire la eventualele deficiente descoperite în aplicație și a modalităților de corectare a acestora.

În situația în care sunt identificate neconcordanțe între rezultatele așteptate și cele obținute pentru cel puțin unul dintre elementele avute în vedere în fișa de testare, se va specifica acest lucru în raportul de testare. **Pasul 7.** Remedierea eventualelor deficiente constatate și reluarea testării pentru validarea corectării acestora.

În vederea testării și validării prototipului de ontologie, a fost implementată o aplicație web dezvoltată în limbajul C#, cu ajutorul platformei Microsoft ASP.NET Core, ce implementează pattern ul Model-View-Controller. Pentru rularea interogărilor SPARQL a fost utilizat serverul Apache Jena Fuseki. Testarea eventualelor inconsistențe în declararea claselor și proprietăților a fost realizată încă din etapa de dezvoltare prin utilizarea unui motor de inferență.

Activitatea 5-2-4: Testarea experimentală a căutărilor bazate pe web-ul semantic la nivelul rețelei sociale online după dezvoltarea ontologiei specifice limbii române

În vederea testării experimentale s-a realizat o fișă de evaluare organizată în două secțiuni principale, o secțiune vizând testarea și evaluarea în cazul căutării unui singur concept, respectiv o secțiune dedicată testării și evaluării în cazul căutării mai multor concepte. Pentru fiecare test au fost precizați pașii de parcurs și s-a asigurat înregistrarea rezultatelor obținute. Rezultatele obținute în urma testării efectuate cu ajutorul a patru utilizatori au fost incluse într-un tabel centralizator. În cazul căutării unui singur concept au fost verificate aspecte precum: Testarea identificării conceptelor cu ajutorul denumirii, Testarea identificării conceptelor cu ajutorul acronimelor, Testarea identificării conceptelor utilizând termeni diferiți, Testarea identificării conceptelor în cazul unor erori uzuale, Testarea identificării conceptelor utilizând termeni la singular / plural, Testarea identificării conceptelor utilizând termeni articulați, Testarea identificării conceptelor utilizând termeni la plural articulați, Testarea afișării conceptelor cu un grad mai ridicat de generalitate, Testarea afișării conceptelor cu un grad mai scăzut de generalitate, Testarea afișării conceptelor similare. În cazul căutării mai multor concepte au fost verificate aspecte precum: Testarea identificării conceptelor cu ajutorul denumirii, Testarea identificării conceptelor în cazul utilizării cuvintelor de legătură, Testarea identificării conceptelor cu ajutorul acronimelor, Testarea afișării conceptelor cu un grad mai ridicat de generalitate în cazul căutării unor concepte similare, Testarea afișării conceptelor cu un grad mai ridicat de generalitate în cazul în care unul dintre concepte are un grad mai ridicat de generalitate în raport cu celelalte concepte căutate, Testarea afișării conceptelor cu un grad mai ridicat de generalitate în cazul căutării mai multor concepte diferite.

Activitatea 5-2-5: Realizarea de cercetări calitative pentru determinarea problemelor de funcționalitate și de utilitate ale ontologiei la nivelul rețelei sociale online

Pentru determinarea problemelor de funcționalitate și de utilitate ale ontologiei

au fost realizate șase studii. Primul studiu a vizat a vizat înțelegerea opiniei utilizatorilor cu privire la deficiențele prezente în motoarele de căutare clasice, cu privire la avantajele oferite de un motor de căutare semantic, precum și evidențierea nivelului de importanță al funcționalităților adiționale pe care un motor de căutare semantic le poate oferi. Cel de al doilea studiu a vizat înțelegerea opiniei utilizatorilor cu privire la utilitatea și îmbunătățirile necesare pentru funcționalitatea de determinare a conceptelor cu un grad mai ridicat de generalitate, implementată în prototipul dezvoltat. Cel de al treilea studiu a vizat înțelegerea opiniei utilizatorilor cu privire la utilitatea și îmbunătățirile necesare pentru funcționalitatea de determinare a conceptelor cu un grad mai redus de generalitate. Cel de al patrulea studiu a vizat înțelegerea opiniei utilizatorilor cu privire la funcționalitatea de identificare a conceptelor similare cu conceptul căutat de utilizator, implementată în prototipul dezvoltat. Cel de al cincilea studiu a vizat înțelegerea opiniei utilizatorilor cu privire la funcționalitatea de identificare a conceptelor pornind de la utilizarea de abrevieri. Cel de al șaselea studiu a vizat înțelegerea opiniei utilizatorilor cu privire la capacitatea prototipului implementat de a oferi funcționalități precum identificarea conceptelor căutate, identificarea conceptelor cu un grad mai redus sau mai ridicat de generalitate, identificarea conceptelor similare, căutarea după acronime. S-a urmărit obținerea de sugestii de îmbunătățire și de observații din partea participanților.

Activitatea 5-2-6: Realizarea unor specificații de îmbunătățire la nivelul ontologiei pe baza rezultatelor cercetării

Pe baza studiilor efectuate s-a desprins necesitatea extinderii atât a listei de concepte, cât și a listelor de termeni și abrevieri aferente acestora. Pentru remedierea deficiențelor observate se va avea în vedere parcurgerea a două etape. În cadrul primei etape se va avea în vedere includerea în ontologie FutureWeb a conceptelor și abrevierilor identificate prin intermediul studiilor efectuate: Pasul 1 - Lărgirea listei de concepte prin includerea conceptelor menționate de respondenți în cadrul studiilor; Pasul 2 - Extinderea listei de abrevieri prin includerea abrevierilor menționate de respondenți în cadrul studiilor; Pasul 3 - Validarea versiunii actualizate a ontologie. Cea de a doua etapa presupune identificarea de noi concepte din domenii conexe informaticii și includerea acestora în ontologia FutureWeb: Pasul 1 - Prelucrarea automată de documente din respectivele domenii, precum materiale didactice (suporturi de curs / seminar); Pasul 2 - Se vor elimina apoi semnele de punctuație; Pasul 3 - Se va trece la împărțirea textului în cuvinte individuale; Pasul 4 - Cuvintele individuale vor fi reduse apoi la forma de bază utilizând un algoritm de stemming; Pasul 5 - Se va utiliza apoi tehnica numită TF-IDF (Term Frequency-Inverse DocumentFrequency) pentru a diferenția cuvintele specifice domeniului de cuvinte de legătură și termeni generali; Pasul 6 - Organizarea termenilor într-o ierarhie cu ajutorul unor experți din domeniu; Pasul 7 - Includerea noilor concepte în ontologia dezvoltată prin crearea instanțelor aferente și a relațiilor necesare; Pasul 8 - Validarea versiunii actualizate a ontologiei.

Activitatea 5-3-1: Construirea unei rețele sociale online test

În vederea construirii unei rețele sociale online test au fost parcurse următoarele etape: Etapa 1 - Compararea rețelelor sociale online pentru care codul sursă este disponibil gratuit; Etapa 2 - Analiza avantajelor și dezavantajelor dezvoltării unei rețele sociale noi, în raport cu adaptare uneia existente; Etapa 3 - Implementarea rețelei sociale test. Astfel, într-o primă etapă au fost analizate rețelele sociale online pentru care codul sursă este disponibil gratuit. Dintre acestea s-au evidențiat prin comunitatea activă, prin documentația detaliată, prin posibilitatea de implementare a unor noi funcționalități sub forma unor module rețelele HumHub și Elgg. În cadrul celei de a doua etape s-a realizat o analiză a avantajelor și dezavantajelor dezvoltării unei rețele sociale noi, în raport cu adaptare uneia dintre rețele sociale existente, precum HumHub. În timp ce dezvoltarea unei rețele sociale noi oferă avantajul posibilității de a ajusta complet rețeaua socială la cerințele proiectului, adaptarea unei rețele sociale existente prin construirea de noi module oferă beneficii importante din prisma diseminării rezultatelor proiectului. Astfel, dacă în cazul construirii unei noi rețele sociale șansele adoptării acesteia de către comunitate sunt foarte reduse, în cazul dezvoltării de noi module pentru rețele sociale populare, acestea pot fi ușor preluate de către comunitate. Pe baza comparației realizate în cadrul etapei 1 și a analizei efectuate în etapa 2 s-a optat pentru adaptarea rețelei sociale HumHub la cerințele proiectului. În cadrul celei de a treia etape s-a trecut la dezvoltarea modulelor necesare utilizând limbajele PHP și JavaScript, alături de librăriile jQuery și Bootstrap.

Activitatea 5-4-3: Diseminarea pe scară largă a rezultatelor proiectului

Pentru diseminarea rezultatelor obținute în cadrul proiectului, au fost folosite toate mediile de comunicare avute la dispoziție: Website, Social Media, Comunicări în cadrul fiecărei facultăți partenere, precum și workshop-uri și conferințe.

A fost publicat un articol științific în cadrul unui jurnal cu factor de impact 0.3:

- C. Delcea, L.-A. Cotfas, C. L. Trică, L. Crăciun, and A. G. Molanescu, "Modeling the Consumers Opinion Influence in Online Social Media in the Case of Eco-friendly Products," *Sustainability*, vol. 11, no. 6, p. 1796, Jan. 2019.

În urma participărilor la conferințe au fost publicate și prezentate următoarele articole:

- Liviu-Adrian COTFAS, Mihai ORZAN, Camelia DELCEA, Chuanmin MI - Corporate Social Responsibility Evaluation on Social Media using Machine Learning and Semantic Web, 29th EBES Conference, Lisbon, 2019.
- Camelia DELCEA, Liviu-Adrian COTFAS, Rafal Mierzwiak, Mihai ORZAN - Consumers Contagion in Online Social Networks Regarding Recycling Habits, 29th EBES Conference, Lisbon, 2019.
- Liviu Adrian COTFAS, Ionut Costinel NICA - Uncovering Social Media Users Emotions towards Companies using Semantic Web Technologies, 28th EBES Conference, Coventry, 2019.

- Camelia DELCEA, Liviu-Adrian COTFAS, Ionut Costinel NICA - Analyzing Customers' Opinions towards Product Characteristics using Social Media, 28th EBES Conference, Coventry, 2019.
- Liviu-Adrian COTFAS, Ioan ROXIN, Camelia DELCEA - SEMANTIC SEARCH IN SOCIAL MEDIA ANALYSIS. In: Proceedings of the 18th International Conference on Conference on Informatics in Economy (IE 2019), Bucharest University of Economic Studies Press, Bucharest, 2019
- Alexandra Cristina Dinu, Raluca Giorgiana Chivu, Alexandru Valentin Teodorov, Otilia Platon, Gheorghe ORZAN - The impact of semantic web in user-machine interaction, SCM 4 ECR Technology and Innovation in Supply Chain Management for Creating New Value for Consumers, Valahia University of Târgoviște, 2019
- Alexandra Cristina Dinu, Violeta Rădulescu, Anca - Francisca Cruceru & Mihai Orzan, Necessity for Semantic Web development in user interaction. Emerging Trends in Marketing and Management International Conference (ETIMM), 4, Bucuresti, 2019
- Alexandra Cristina Dinu, Violeta Rădulescu, Anca – Francisca Cruceru & Mihai Orzan, How Semantic Web Can be used in Better Machine Decision