

IKTA-126/02.

## Logikai alapú ontológiakezelés

Magyar nyelvű összefoglaló

A projektjavaslat egy olyan logikai alapokra épülő ontológiakezelő informatikai eszköz elkészítését tűzi ki célul, amely a résztvevők évtizedes hazai kutatási tevékenységére és sikeres európai kutatási projektjeinek eredményeire támaszkodik. A megcélzott ún. LOBO (LOGic-Based management of Ontologies) munkapad egy alkalmazás-független eszköz ontológiák kezelésére, amelynek orvosbiológiai területre való kísérleti felhasználása a projekt kiemelten fontos része. Az orvosbiológia ma egyik legnagyobb „termelője” és felhasználója az ontológiáknak: az ezen a területen szerezhető tapasztalatok és a kialakítandó megoldások túlnyomó része várhatóan más területekre is jól átvihető.

A projektjavaslat épít az egyik konzorciumi tag által koordinált IST-1999-11135 jelű SILK (System Integration via Logic and Knowledge) nemzetközi EU-projekt eredményeire. A SILK projekt keretében az információintegráció alkalmazási szintű támogatására egy **alaptechnológia** kialakítása történt meg. Ezt az alaptechnológiát a jelen pályázatban a konzorcium a konceptuális – ontológia szintű – integráció irányába kívánja továbbfejleszteni. A projektjavaslat – követve a kapcsolódó EU-projekt technikai irányvonalát – a technológia kidolgozásánál alapvető követelménynek tekinti a gyakorlatban is bevált informatikai szabványok felhasználását és a megfelelést ezeknek (pl. Unified Modelling Language, Object Constraint Language, Prolog, Java, eXtended Markup Language, XML Model Interchange). Mivel a tervezett fejlesztés az említett EU-projekt keretében létrehozott technológiát veszi kiindulási alapul, a pályázat megvalósíthatósági kockázata jelentősen lecsökken.

Az alaptechnológia továbbfejlesztése azonban csak valós alkalmazási környezetben képzelhető el. Ezt az alább részletezett okok miatt az orvosbiológiai területre való alkalmazással tervezi megvalósítani a konzorcium.

A 90-es években bekövetkezett gyors fejlődés következtében az orvosbiológia (pl. a kapcsolódó molekuláris biológia, génszabvány stb.) ma már **exponenciálisan növekedő biológiai adatmennyiséget** „termel ki”. A pusztán adatmennyiségen túl a mögöttes **biológiai komplexitás**, ha lehet, még nagyobb kihívást jelent az ilyen típusú információk kezelésénél, amely párosul az **elosztottság** problémájával is: pl. több száz géninformációt tartalmazó adatbázist kezelnek a szakértők világszerte, és hasonló a helyzet más területeken is. Ezek egységes központi adatbázisba való szervezése (azaz a „szokásos” megoldás követése) az ismeretek/információk rendkívüli elosztottsága és gyors növekedése miatt teljes mértékben lehetetlen. Más oldalról az orvosok és biológusok sürgető igénye, hogy az **adatok elemzésénél** támogatást, útmutatást kapjanak, másokkal szakmailag szorosan együttműködhessenek. A különböző biológiai és orvosi tudományágak jobb integrációjához, a különböző szakterületek információinak és ismereteinek összekapcsolásához és egységes keretek között történő alkalmazásához az információtechnológia innovatív felhasználása szükséges valamilyen praktikus, **közös platform** biztosításával.

A problémakör konkrét eszköztárában kiemelt szerepet kapnak az **orvosbiológiai ontológiák**, amelyek mintegy kodifikálják egy-egy szakterület explicit ismeretanyagát, és strukturált formában bocsátják rendelkezésre azt a szakértők számára. Az ontológiák felhasználása ma már egyre nélkülözhetetlenebb segédeszköz az orvosbiológiai kutatások, kísérletek során és a napi gyakorlatban.

Az ontológiák elterjedése, sokoldalú felhasználása előtt azonban jelenleg még számos akadály áll: 1) heterogén, nem szabványos tárolási módok, 2) a hiányos ill. nem kellően specifikus támogatás a szakértői szinten, 3) a további felhasználási módokat megnehezítő különböző szempontú és szintű leírások. További problémát jelent az orvosbiológiai ontológiák szokásosnál nagyobb méretének kezelése, a konzisztenciájuk, újrahasonosíthatóságuk biztosítása és nyelvfüggetlenségük megőrzése.

A tervezett ontológiakezelési technológia és eszköz elsősorban az ontológiák **konzisztencia-ellenőrzésének**, a szakterületi ontológiák és a szakértő- ill. szakértőcsoport-specifikus ontológiák **összekapcsolásának**, integrálásának és a szakterületi információforrások ontológiaalapú **(fogalmi szintű) lekérdezésének** területein fog újat adni. Emellett figyelemre méltó az ontológiák automatikus (de emberi felügyelet mellett történő) **átstrukturálásához** adandó támogatás, az ontológiák szerkezetének és tartalmának, valamint a lekérdezett információk grafikus, a szakértők számára **intuitív megjelenítése**.

A projekt eredménye nemcsak egy jól használható általános (szakterület-független) eszközként jelenik meg, hanem fontos hozzájárulást lehet az orvosbiológia hazai művelőinek informatikai eszköztárában is. Az eszköz használatával helyi szakértő csoportok saját fogalmi rendszerüket könnyen meg tudják szervezni ontológia formájában és össze tudják vetni a „hivatalos” ontológiákkal hiányok, hibák, ellentmondások kiszűrése érdekében. Felgyorsulhat ezáltal a szakmai együttműködés a különböző szakértő ill. kutató csoportok között az újonnan megjelenő fogalmak elnevezése és szerkezete területén. Ez előrelépést jelenthet az egyre gyorsuló ütemben keletkező biológiai adatok értelmezésében és felhasználásában is.

A LOBO munkapad technológiai alapjai (**Prolog** alapú ismerettárolás és kapcsolódó programozási felület, illetve **Java** alapú adatelérési és felhasználói felületek) lehetővé teszik, hogy az ontológiákat intelligens, korlát-logikai elven működő, szakterület-specifikus következtetési, ismeretkezelési modulokkal lehessen viszonylag könnyen, konkrét igényekhez illeszkedően kibővíteni, és ezzel továbbtágítani a szakértőcsoportok kutatási tevékenységének informatikai támogatását.

A projekt eredményeit két kísérleti alkalmazás megvalósításával verifikálja. Ezek keretében létrehozunk egy sebészeti beavatkozásokat nyilvántartó adat/tudásbázist és mammográfiás felvételek kiértékelését támogató adat/tudásbázist az orvosi partner intézményében.

Az 1. konzorciumi tag a hazai szoftverpiac és különösen az alkalmazásfejlesztési és –integrálási piac egyik vezető cége. Ugyanakkor a cég komoly kutató-fejlesztő részleggel is rendelkezik, amely az elmúlt évtizedben mintegy tucatnyi sikeres hazai és külföldi projektben bizonyított. A javasolt projektben ez a pályázó látja el a **technológiai központ** szerepét.

A 2. konzorciumi tag sokéves tapasztalattal és alkalmazási gyakorlattal rendelkezik az ontológiák felhasználására vonatkozóan különböző szakterületeken, különösen az orvosbiológia területén. A javasolt projektben ez a pályázó tölti be a **tudásközpont** szerepét be.

A 3. konzorciumi tag az orvosi szakterületet képviseli, és így ez a pályázó tölti be a szakterületi **tudásközpont** illetve az **alkalmazási központ** szerepét. Ez a konzorciumi tag központi szerepet kap a kísérleti alkalmazások létrehozásában és kiértékelésében.