

Szemantikailag szervezett lexikai hálózat és internetes tartalomkeresés

Szakmai összefoglaló

A projekt célja egy szemantikai szerveződésű, értelmezett, lexikai hálózat kifejlesztésére épülő, internetes tartalomkeresésre alkalmazható, újfajta technológia létrehozása. Az újfajta keresési technológia és alkalmazás a hagyományos elektronikus keresésbe mintegy életet lehel, azt meggyorsítja. A projekt ezen célját a lexikai hálózat alapegységének tekintett, már kifejlesztett, ún. jelentésközpontok¹ lehetséges kapcsolódásainak kutatásával és a kapcsolóelemek kiépítésével kívánja elérni. A jelentésközpont az azonos jelentés köré szerveződő, értelmezett természetes nyelvi kijelölők (szavak, szó szerkezetek, mondatértékű kifejezések) egy struktúrában összefogott és kezelt egysége. A jelentésközpontok egymással ily módon való összekötésével, ún. linkek létrehozásával létrejövő, szemantikailag szervezett, kommunikatív lexikai hálózat a projekt olyan kutatási szempontok alapján fejleszti ki, hogy az képes legyen nyelvtechnológiai alkalmazásokban (természetes nyelvi szövegfeldolgozó-rendszerek, értelmezett információkeresés elektronikus szövegekben és strukturált szövegtestekben, tartalomfigyelés, gépi fordítás, kontextus- és stílusérzékeny helyesírás-ellenőrző) értelmezetten és hatékonyan működni.

A kutatócsoport által előzőleg szerkesztett és fejlesztett, jelentésalapú szótári adatbázisok (mint a jelentésközpontok felszínre jutásai, kommunikatív nyelvi tükröződései), valamint a projekt során kifejlesztendő alkalmazás segítségével, az internetes tartalomkeresés tesztelése során, a felhasználó kiválaszthatja és azonosíthatja az általa keresett szót annak általa akkor egyedülként keresett jelentésében. Ezáltal a keresés valós találati aránya várhatóan megnő, illetve a hamis találatok száma bizonyosan lecsökken, mivel a létező, karaktersorra működő keresőszolgáltatások ez esetben (a kifejlesztendő alkalmazásnak és adatbázis-struktúrájának köszönhetően) csak a felhasználó által beazonosított és kijelölt egyetlen jelentésközpontnak kapcsolódó lexikai elemekre végzik el a keresést.

A projekt várható eredményessége a valódi, téma vagy fogalomkör szerinti, illetve az értelmes gépi tartalomkeresés közeles ígéretét is nyújtja az informatika tudománya és a számítógépes felhasználó megváltozott igényei számára egyaránt.

¹ Az elnevezés magyarázatát lásd a 2. kötet 1.2.2. részében.