**Názov výsledku:** Spracovanie a vyhľadávanie informácií z Internetu
**Anglický názov výsledku**: Information Processing and Search from the Web
**Číslo projektu a agentúra**: Spolupráca s podnikateľskou sférou
**Hlavní odberatelia**: SCS GROUP, Švajčiarsko
**Autori**: Ladislav Hluchý, Martin Šeleng, Michal Laclavík, Štefan Dlugolinský, Marek Ciglan, Emil Gatial, Zoltán Balogh

Dielo predstavuje súbor nástrojov na sťahovanie webového obsahu (textu) za účelom ďalšieho spracovania ako je: rozpoznanie názvoslovných entít, dátumov, geografických lokalít a iných objektov a ich vyhľadávanie. V rámci takto navrhnutého systému boli dodané 2 aplikácie: vyhľadávanie nad pracovnými ponukami a vyhľadávanie nad novinovými správami. Overenie aplikácie na vyhľadávanie pracovných ponúk prebehlo nad pracovnými ponukami stiahnutými so stránky LinkedIn.com, kde sme spracovali viac ako 100 tisíc dokumentov, identifikovali sme cca 70 tisíc pracovných ponúk a 200 tisíc rôznych objektov. Cieľom tejto aplikácie bolo nájsť pre určitú pracovnú ponuku vhodného kandidáta na základe porovnania životopisu a pracovnej ponuky. Druhá aplikácia určená na vyhľadávanie nad novinovými správami bola overená nad dátami získanými z portálu BBC.com. Obsahuje okolo 19 tisíc novinových článkov, kde sme identifikovali okolo 100 tisíc objektov (ľudí, organizácií, krajín, miest a iných entít) nad ktorými je možné vyhľadávať jednotlivé dokumenty, vzťahy medzi objektmi ako aj vyhľadávať dokumenty s obmedzením na geografickú oblasť definovanú používateľom. Projekt bol rozdelený na tri etapy. Celková cena zákazky bola 70 000€.

**Hlavné scientometrické[1] výstupy**:
**ADEB05, AEC27**