

Laboratorium Web Mining PJWSTK

<http://www.pjwstk.edu.pl/~msyd/webmining.html>

zaprasza na wykład seminarium „Web Mining”

Poniedziałek 7 stycznia, 15:10, sala senatu

EXPRESS

**Extraction Pattern Engine
and Specification Suite**

-

**Nowy język specyfikacji reguł ekstrakcji
z tekstów ciągłych i jego interpreter**

dr Jakub Piskorski

Joint Research Centre of European Commission, Ispra Italy

Podczas wykładu przedstawiony zostanie nowy formalizm EXPRESS, służący do specyfikacji reguł ekstrakcji z tekstów. Został on wykorzystany między innymi w systemie Europe Media Monitor (<http://press.jrc.it/NewsBrief>) na potrzeby ekstrakcji informacji o różnego typu zdarzeniach z sieci WWW. Prezentacja o nieco technicznym charakterze będzie zawierała omówienie składni formalizmu, podstawowych zasobów przetwarzających w EXPRESS oraz implementacji interpretera reguł.

Reguły w formalizmie EXPRESS są wyrażeniami regularnymi nad 'płaskimi' strukturami atrybutów. Formalizm EXPRESS stanowi 'połączenie' dwóch języków specyfikacji reguł ekstrakcji: JAPE (platforma GATE) i XTDL (system SProUT). Główną motywacją powstania EXPRESS były problemy z efektywnością przetwarzania reguł w dwóch wyżej wymienionych językach.(slides will be in English)

Serdecznie zapraszamy,
Prof. dr hab. Witold Kosiński i dr Marcin Sydow
Katedra Systemów Inteligentnych