

**POZVÁNKA na třetí setkání PRAŽSKÉHO INFORMATICKÉHO SEMINÁŘE**

# **KAREL OLIVA:** **Korektor české gramatiky**

*po přednášce bude následovat diskuse*

## **ANOTACE PŘEDNÁŠKY**

Zatímco český i anglický spelling-checker ("korektor pravopisu") mohou být založeny na stejných principech, konkrétně na kontrole toho, zda se určité slovo vyskytuje v (rozsáhlém, ale konečném) seznamu slovních tvarů, korektor českého pravopisu nemůže vyhledávat chyby pouhým srovnáváním textu s předem danými chybovými konfiguracemi: důvodem je především typologická odlišnost češtiny, v níž dominantní roli ve skladbě věty hraje tvarosloví, zatímco pořádek slov je gramatikou ovlivněn jen spíše výjimečně. Tvorba korektoru českého pravopisu proto vyžaduje alternativní, zásadně odlišný přístup.

V přednášce se proto budeme nejprve věnovat obecně teoretickým úvahám o tom, co je „zaručeně gramaticky chybná konstrukce (jakéhokoliv) jazyka, a dále pak technikám, jak takové negramatické konstrukce a věty (i) definovat, (ii) explicitně popsat a (iii) detekovat v případě češtiny. Na závěr si ukážeme, že návrh oprav chyb je sice s hledáním chyb tematicky svázaný, ale fakticky velmi odlišný úkol, a uvedeme i jisté teoretické (a pro rodilé mluvčí češtiny snad dokonce zábavné) dopady celé práce.

## **O PRAŽSKÉM INFORMATICKÉM SEMINÁŘI**

Seminář se schází vždy 4. čtvrtek v měsíci v 16 hod. (s výjimkou letních měsíců a prosince), a to buď v budově FEL ČVUT na Karlově náměstí, nebo v budově MFF UK na Malostranském náměstí.

Jeho program je tvořen hodinovou přednáškou, po níž následuje časově neomezená diskuse. Základem přednášky by mělo být něco (v mezinárodním měřítku) mimořádného nebo aspoň pozoruhodného, na co přednášející přišel a co vysvětlí způsobem srozumitelným a zajímavým i pro širší informatickou obec. Přednášky jsou standardně v angličtině.

Idea Pražského informatického semináře vznikla z rozhovorů představitelů několika vědeckých institucí na téma, jak odstranit zbytečnou fragmentaci informatické komunity v ČR. Formát semináře připravil přípravný výbor ve složení Michal Chytil (ÚI AVČR), Pavel Kordík (FIT ČVUT), Jan Kybic (FEL ČVUT), Michal Pěchouček (FEL ČVUT), Jiří Sgall (MFF UK), Vojtěch Svátek (FIS VŠE), Michal Šorel (ÚTIA AV ČR), Filip Železný (FEL ČVUT).



**Doc. RNDr. Karel Oliva, Dr.** je mezinárodně uznávanou vědeckou osobností v oblasti formální syntaxe přirozených jazyků, v níž uplatňuje i inovativní přístupy vedoucí k praktickým (počítačovým) aplikacím. Absolvoval MFF UK, byl dlouhodobě vědecky činný v Bulharské akademii věd v Sofii, v Ústavu počítačové lingvistiky Sárské univerzity v Saarbrücken (SRN) a v oddělení matematické lingvistiky Rakouského ústavu pro výzkum umělé inteligence (OeFAI) ve Vídni. Publikoval přes 85 původních vědeckých prací v časopisech a sbornících odborných konferencí, byl vedoucím výzkumných týmů a spoluautorem implementací řady softwarových celků (mj. korektoru české gramatiky pro MS Office). Je ředitelem Ústavu pro jazyk český AV ČR. Pedagogicky působí na Přírodovědecké fakultě UJEP.

**Kontakt:** [info@praguecomputerscience.cz](mailto:info@praguecomputerscience.cz)

**Informace:** [www.praguecomputerscience.cz](http://www.praguecomputerscience.cz)

