

Софийски университет Св. Климент Охридски, Факултет по славянски филологии

ИНТЕНЗИВЕН КУРС – СВОБОДНО ИЗБИРАЕМА ДИСЦИПЛИНА
(за студентите от бакалавърските и магистърските програми)

Компютърна компаративистика

Езикови технологии за славистични изследвания



ст.н.с. II ст. д-р **Елена Паскалева**
Институт по паралелна обработка на информацията, БАН

PD Dr. phil. habil. **Таня Августинова**
DFKI Language Technology Lab / Saarland University

Курсът ще се проведе през първата половина на октомври 2007

URL: <http://www.coli.uni-saarland.de/~tania/block-kurs/>

e-mail: SlavicNLP@online.de (регистрация до 2 октомври 2007)

АНОТАЦИЯ (ОБЕМ: 30 лекционни часа)

Предлаганият курс се изгражда на базата на (i) общи положения на компютърната лингвистика; (ii) съдържание и обхват на курсовете по компютърна лингвистика в Саарландския университет; (iii) конкретни лингвистични технологии, разработени в ИПОИ БАН. Постепенното въвеждане в теорията на лингвистичните технологии и компютърните модели се осъществява чрез обработка на конкретни равнища от езиковото описание на славянските езици (текстово, морфологично, синтактично) и цели не само да запознае студентите с един модерен клон от езикознанието, но и да ги мотивира да усвоят основни съпоставителни параметри в описанието на тези езици. Студентите се запознават и с налични компютърни ресурси – граматически речници, текстови корпуси, статистически и лингвистични процедури, които ще им бъдат предоставени за лабораторна работа. Така се постига опознаване на механизма на контрастивността чрез компютърното му моделиране.

ПЪРВА СБИРКА: 1 октомври 2007 (13:00) Деканата на ФСФ

МЯСТО НА ПРОВЕЖДАНЕ: зала 124А на ФСФ от 17:00 до 19:30

- Основи на компютърната лингвистика и формални подходи към славянските езици (02.10.2007)
- Граматически речници (03.10.2007)
- Анотационни маркери (04.10.2007)
- Обработка на паралелни текстове (05.10.2007)
- Измерване на лексикална близост между славянските езици (08.10.2007)
- Приложения, изискващи (пълен или частичен) синтактичен анализ (09.10.2007)
- Модулно разработване на компютърни граматика за славянските езици (10.10.2007)
- Използване на синтактично анотирани корпуси в славистиката (11.10.2007)
- Машинен превод (12.10.2007)