

Linguateca: Polo de Coimbra - Plantando Florestas e Criando Papel

Paulo Gomes, CISUC / DEI-FCTUC

Enquadramento

■ Início:

- 16 de Fevereiro de 2007

■ Pessoas:

- Cláudia Freitas
- Paulo Rocha
- Hugo Oliveira
- Nuno Seco
- Paulo Gomes

Enquadramento

■ Objectivos:

- Ferramentas de extracção de conhecimento.
- Criação de uma ontologia lexical.
- Metodologia de avaliação.

Recursos

■ Floresta

- “Treebank” para a língua Portuguesa.
- Manutenção e evolução do recurso.

■ Papel

- Extracção de uma ontologia lexical para o Português a partir do dicionário da Porto Editora.
- Criação do recurso.

Recursos

■ PEN Parser

- Chart Parser que implementa o algoritmo de Earley [Java].

■ Etiquet(H)AREM

- Aplicação utilizada para anotar textos de acordo com a sintaxe do Segundo HAREM.

■ SílabasPT

- API Java que implementa métodos para a separação de sílabas e identificação da sílaba tónica de palavras portuguesas.

Outras Tarefas

- Participação no Harem
 - Avaliação das Participações
 - Anotação da Colecção Dourada
- Workshop da Floresta
 - Introdução à Floresta Sintáctica
- Workshop no Talc
 - Introdução e utilização da Floresta

Lições e Experiências

- Três comunidades de utilizadores diferentes:
 - Linguística
 - Uso, criação e estudo de recursos.
 - Computacional
 - Criação de recursos.
 - Engenharia
 - Uso de recursos.