

AvalON'2003: Abertura

Diana Santos, www.linguateca.pt

AvalON'2003

- Encontro dedicado à avaliação conjunta do português
- Organizado pela Linguateca
- Em preparação desde o encontro preparatório em Faro, Junho de 2002

O que é a Linguateca

- Uma iniciativa para melhorar o processamento do português
 - Disseminação
 - Criação de recursos
 - Avaliação
- Uma organização virtual com quatro pólos
 - Oslo, Braga, Lisboa, Porto, ...
 - Colaboração com outras localizações: Odense, Lisboa, São Carlos, Porto Alegre, ...
- A continuação do projecto *Processamento Computacional do Português*, lançado em 1998, pelo então Ministério da Ciência e da Tecnologia (projecto POSI até 2006)

Razões de existir da Linguateca

- Começar pelo princípio
 - Quais são os maiores entraves ao processamento do português?
- Entidades internacionais não podem resolver os nossos problemas
 - De qualquer maneira, não melhor do que nós
 - A construção de recursos é cara, e o português nunca tem prioridade...
- Língua (e não região, or nação) devia ser a unidade no processamento da linguagem natural
 - Ou seja, Brasil e Portugal deviam ter uma cooperação estreita
- Recursos públicos são obrigatórios para o progresso científico
 - Já há suficientes barreiras...
- É preciso um enorme esforço organizacional para efectuar avaliações conjuntas

Pressupostos do AvalON'2003

- É em conjunto que uma comunidade científica progride
- Não é trivial chegar a uma especificação consensual
 - conhecimento implícito
 - decisões não documentadas
 - opções nunca postas em causa
- Comunicação é cara, morosa e difícil
 - combater/criticar é incomparavelmente mais simples
 - cooperar é complexo
 - é preciso encontros em que as pessoas se conheçam
- Avaliação conjunta é o único caminho a seguir para garantir que a investigação foca problemas verdadeiros e não reinventa a roda

Objectivos

- Critérios objectivos Forma de medir o progresso
- Reconhecimento É bom sair da toca, obter reconhecimento, ter outras pessoas a olhar para o resultado
- Auto-reflexão Reflectir sobre os pressupostos e eventualmente justificá-los
- Cartografia Medir a contribuição dos vários problemas
- Momento histórico Orgulhar-se de participar na primeira vez; estar desde a primeira hora envolvido

Programa

- Morfolimpiadas
A primeira avaliação conjunta para o português
- Avaliação de recolha de informação em português
Vários actores diferentes discutem a compilação de uma colecção genérica
- Avaliação de tradução automática de inglês para português
Descrição da obtenção conjunta de dados para analisar sistemas de TA
- Avaliação de alinhamento
Proposta de avaliar alinhadores de frases, etc.

Programa (cont.)

- Avaliação de reconhecimento de entidades citadas
Ensaio de uma avaliação conjunta de identificação de entidades citadas
- Discussão sobre a Floresta
Discussão de um recurso para avaliação de análise sintáctica em português. Como prosseguir?
- Encerramento
Planos para o futuro