



UiO : Det humanistiske fakultet

Diana Santos

Balanço do Págico

algumas notas





O que aprendemos

- A dificuldade de formalizar uma resposta, a sua utilidade, e a sua comparabilidade, são problemas semânticos interessantes
- Lidar com coleções grandes, mesmo de recursos “multilingues” está sempre demasiado relacionada com o inglês. Para as outras línguas, há remendos e adaptações
- Não é fácil obter impacto de uma iniciativa de PLN ou RI



Ideias para o futuro

- A questão da classificação e das categorias da wikipédia é algo que pode dar informação interessante sobre o processo de categorização em linguagem natural
- Diferentes categorias em diferentes línguas?
- Que tipo de marcas os wikipedistas costumam usar?
- Como obter semi-automaticamente tópicos “semelhantes” para futuras edições?



Trabalho de casa

- Em relação aos tópicos do Págico
 - Além de verificar o progresso na versão corrente
 - Adicionar o que (sabemos que) falta
 - Corrigir as incorreções encontradas
- Melhoria da coleção
- Estudo dos erros/fontes de problemas



Agradecimentos

- Aos participantes no Págico
- À comissão científica da edição especial da Linguamática
- Ao Fernando Perdigão pelo convite
- Aos membros da mesa redonda e a toda a audiência do encontro
- Aos financiadores



FCCN Fundação para a Computação Científica Nacional
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

FCT Fundação para a Ciéncia e a Tecnologia
MINISTÉRIO DA CIÉNCIA, TECNOLOGIA E ENSINO SUPERIOR

MCTES MINISTÉRIO DA CIÉNCIA, TECNOLOGIA E ENSINO SUPERIOR

