Português computacional

- O processamento da língua portuguesa por computador foi identificado como precisando de atenção especial
- Há três intervenientes
  - A comunidade científica
  - O Estado através da sua política de gestão da investigação
  - A indústria

1

O MCT tomou a iniciativa

- Encomendando um documento para discussão a uma pessoa fora do "establishment"
- Iniciando o processo de criar um perfil desta área

Agora é a altura das instituições, das empresas e da comunidade científica se pronunciarem

2

Objectivo último

- os portugueses deveriam ter sistemas informáticos (com que cada vez mais se vêem forçados – ou tentados – a interagir) que os tratassem na sua própria língua

3

Causas múltiplas

- é uma área que não é reconhecida
- em que os seus (poucos) intervenientes na sua maioria guardam para si os frutos do seu trabalho e a sua competência (recursos partilháveis são mínimos; as publicações são quase inexistentes)
- em que os mecanismos de avaliação até agora não funcionaram de todo
- em que existe demasiada ênfase na teoria (desleixo da prática)
- em que não há meios, ou vontade, de ultrapassar a investigação fundamental (ou seja, não se desenvolvem aplicações relacionadas com a população em geral)
- em que a formação é francamente insuficiente
- em que as relações entre as pessoas não primam pela transparência

4

Tentativa de resolução

O meu texto

- sugere algumas formas de contrariar cada um dos pontos mencionados
- adianta algumas medidas concretas
- refere que é preciso haver um consenso sobre as medidas a tomar e insiste na necessidade de colaboração

5

Processamento computacional da língua portuguesa: documento de trabalho

1. Definição da área como área prioritária
2. Disponibilização de recursos (nas suas múltiplas vertentes)
3. Avaliação e controlo de qualidade em relação ao português
4. Serviços de desenvolvimento de recursos e ferramentas partilháveis
   - serviço de tradução
   - serviço de terminologia
   - rede de fala
   - rede de processamento da língua escrita
5. Dinamização dos métodos empíricos
6. Ligação da investigação fundamental com as tecnologias

6

Processamento computacional da língua portuguesa: documento de trabalho (cont.)

7. Desenvolvimento de aplicações relacionadas com o trabalho de todos os dias no sector da informação
8. Política de formação na área
9. Áreas relacionadas
   - o direito e o processamento computacional de uma língua
   - a internet e o processamento computacional de uma língua
   - bases de dados, arquivos e o processamento computacional de uma língua
   - a informatização da educação e o proc. computacional de uma língua
10. Transparência, participação e colaboração de todos
11. Contributos para um perfil

7

Medidas concretas

- a criação de serviços dedicados a facilitar o trabalho na área
- a aposta no interrelacionamento com outras áreas
- a criação de um curso de processamento computacional do português na rede (Internet)
- algumas formas de prevenir a falta de transparência
- algumas maneiras de dinamizar os métodos empíricos, estimular o desenvolvimento das aplicações e incentivar a avaliação

8

Contributos para um perfil

Através dos dados oficiais

- Uma lista de **projectos**, divididos em projectos na área e projectos relacionados 53
- Uma lista de **bolsas**, também divididas entre bolsas na área ou com ela relacionadas 52 / 280
- Uma lista de **doutorados** potencialmente associados ou associáveis à área 93 / 541

Através da informação publicada na Internet

- um catálogo de **recursos** e de **actores**

9