

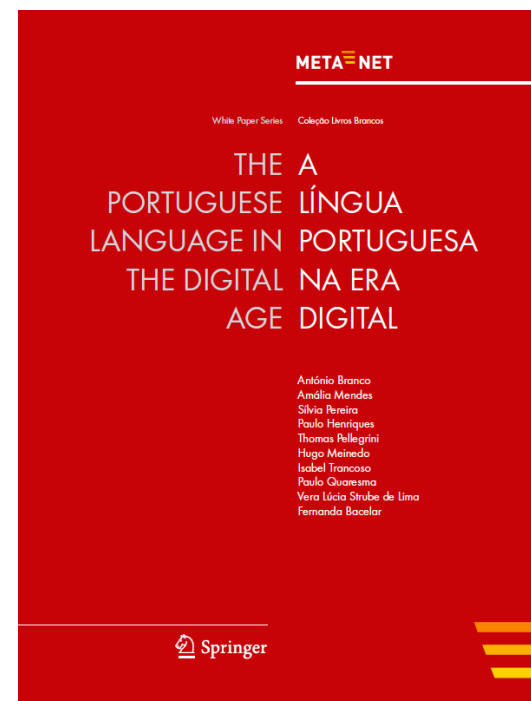
Preparação da língua portuguesa para a sociedade de informação

António Branco



Livro Branco

- **Língua portuguesa**
 - 10 coautores; Portugal e Brasil
- **Coleção**
 - 30 línguas; 200+ peritos
 - divulgação, avaliação, recomendação
- **Rede europeia de excelência em I&D**
 - META-NET; 60 centros; 34 países



Preparação tecnológica dos falantes?

- **5ª** língua em nº de utilizadores na internet
 - 5ª língua em nº de falantes no mundo

- **3,8%** utilizadores de internet no mundo
 - 3,7% população mundial



- **34,1%** penetração da internet nos países da CPLP
 - 32,7% penetração no mundo

- 31/12/2012, internet world stats (inglês, chinês, castelhano, japonês)

Preparação tecnológica da língua?

- **Processamento computacional da língua portuguesa**
 - processamento das expressões linguísticas e do seu significado
- **Tecnologia crucial**
 - Interagir com serviços e dispositivos artificiais em linguagem natural
 - desde robôs sofisticados a electrodomésticos banais
 - Comunicar com falantes de outras línguas
 - sem haver língua veicular comum
 - Aceder a acervos de informação
 - registados em outras línguas

Avaliação qualitativa

| Apoio excelente | Apoio bom | Apoio médio | Apoio fragmentário | Pouco/nenhum apoio |
|-----------------|-----------|---|--|---|
| | Inglês | Alemão Espanhol Francês Italiano Neerlandês | Basco Búlgaro Catalão Checo Dinamarquês Eslovaco Esloveno Finlandês Galego Grego Húngaro Norueguês Polaco Português Romeno Sueco | Croata Estónio Irlandês Islandês Letão Lituano Maltês Sérvio |

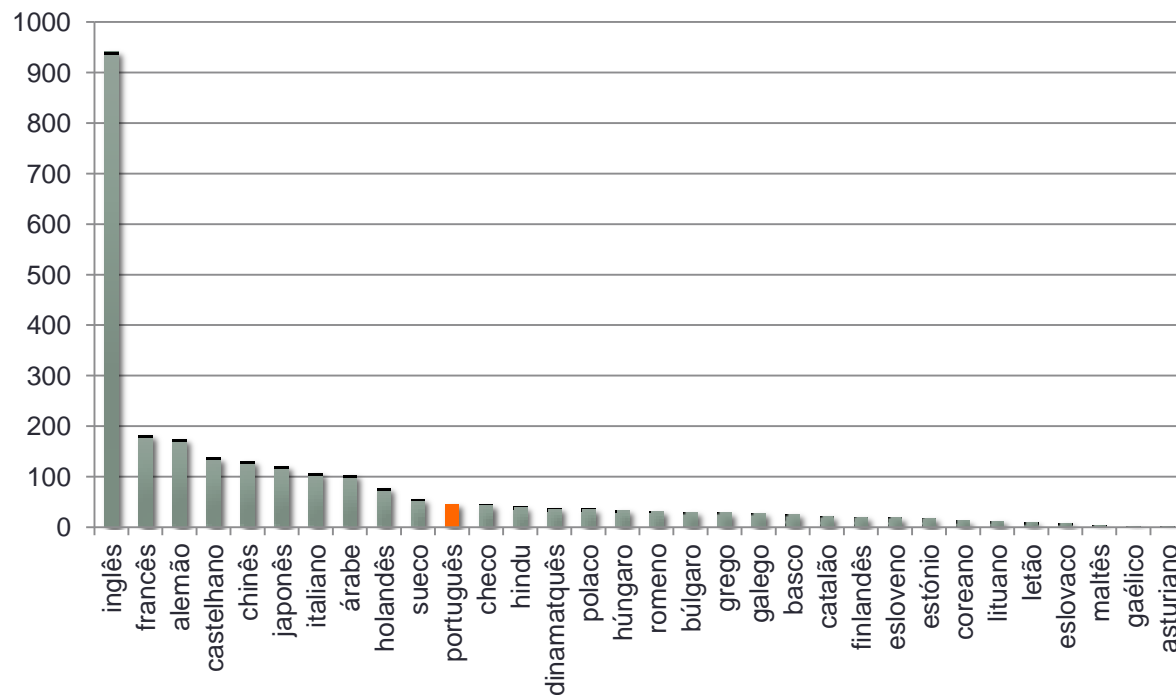
10: Análise do Texto: estado da tecnologia da linguagem para 30 línguas europeias

Avaliação quantitativa

- Esforço de investigação

- Referências a línguas em artigos científicos da área, 10 conferências de topo, 2010-2012

- ≈ Basco, Romeno, ...
- Holandês 2 x mais, Italiano 2,5 x, Castelhana 3 x, Alemão e Francês 4 x mais
- **Inglês 22 x mais**



Desafio

- **Processamento Computacional da Língua Portuguesa**
 - **Impacto social**
 - cidadania plena na sociedade da informação
 - **Económico**
 - inovação e internacionalização na economia digital
 - **Cultural**
 - soberania, património e mundividência
 - **Geoestratégico**
 - língua internacional de comunicação global
 - **Científico**
 - plataforma de cooperação com Brasil e restantes países da CPLP

Desafio estratégico prioritário

- **Impactos**

- **Social**

- cidadania plena na sociedade da informação

- **Económico**

- inovação e internacionalização na economia digital

- **Cultural**

- soberania, património e mundividência

- **Geoestratégico**

- língua internacional de comunicação global

- **Científico**

- plataforma de cooperação com Brasil e restantes países da CPLP

- **Processamento Computacional da Língua Portuguesa**

- Plano nacional de I&D para 2014-2020

Oportunidade

- Fundos estruturais / Europa 2020
- Programa quadro Horizonte 2020
- Países/regiões: Especialização inteligente
 - 9 temas em Tecnologias de Informação e Comunicação :
- Processamento Computacional das Línguas Nacionais



S3 Platform – language resources

s3platform.jrc.ec.europa.eu/language-resources/

Legal notice | Search | Contact | Sitemap | English | Sign In

European Commission

SMART SPECIALISATION PLATFORM

European Commission > S3 Platform > Digital Agenda > language resources

Home | About S³P | News and events | RIS³ Guide | Guidance & Documents | Eye@RIS³ | Digital Agenda | Peer review | Thematic activities

Language resources

The META Language white papers series, published in autumn 2012 indicated that as many as 21 European languages (most of which are official EU languages) are facing "digital extinction" – meaning that relevant and up-to-date online content and online services are not available in those languages (or that the level of service in such "endangered" languages is significantly lower than English or other major languages). However, language technologies such as machine translation (MT) offer opportunities to gap language barriers in the online domain. The problem is that for the smaller languages (especially for the 21 languages referred to above), the availability of MT solutions is limited, due to limited availability of language resources. This is because language resources are the indispensable "raw material" for producing high-quality MT systems. While EU-funded actions have been undertaken (see META: box below) to create a single access

Useful resources

Thematic Fiche

[Language resources](#)

Further reading

[Language White Papers](#)

[META members \(national centres\)](#)

Desafio estratégico prioritário: linhas de ação 1/2

- **Estímulo interdisciplinar específico**
 - Programas coordenados de estímulo a esta área interdisciplinar
 - Para cobrir ciência da e tecnologia para a língua portuguesa :
 - investigação básica
 - desenvolvimento tecnológico
 - empreendedorismo inovador
 - Envolvendo um leque amplo de instrumentos, incluindo entre outros:
 - chamadas de propostas de projetos dirigidas à área
 - programas de pós-graduação específicos
 - bolsas de pós-graduação agrupadas para esta temática
 - incentivos à inovação nesta área
 - ... etc

Desafio estratégico prioritário: linhas de ação 2/2

- **Reforçar cooperação internacional**
 - Aderir à recém-criada **infraestrutura internacional de investigação científica** da área:
 - CLARIN – Common Language Resources Infrastructure
 - Programas de intercâmbio e cooperação:
 - entre **Brasil** e Portugal
 - outros países da **CPLP** e Portugal

Inscrever

o Processamento Computacional da Língua Portuguesa

como Desafio Estratégico Prioritário

no Plano Nacional I&D para 2014-2020

tema específico recomendado pela CE

para aplicação dos fundos estruturais / especialização Inteligente

Muito obrigado.

António Branco

