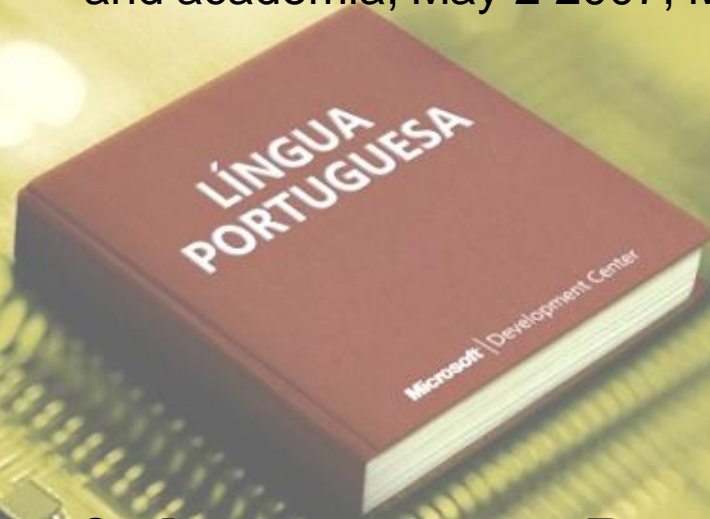


I Microsoft Workshop on Speech Technology - Building bridges between industry and academia, May 2 2007, MLDC



Microsoft Language Development Center in Portugal: an overview

Miguel Dias, Director
Miguel.Dias@microsoft.com

- Outline:
 1. Microsoft Language Development Center in Portugal: launch
 2. Mission and action lines
 3. Results:
 - Speech Lab
 - Technology
 - Speech ecosystem
 4. Challenges for 2007 and 2008

MLDC - Microsoft Language Development Center, Portugal

Microsoft | Development Center
Portugal

□ A Microsoft | Development Center in Portugal

- <http://www.microsoft.com/portugal/mldc>
- Miguel Dias (Director, FTE)
 - Staff: PM, Dev, Test and Linguist
- Founded in Nov 2005
- One of the four Microsoft Development Centers established in Europe and the first, outside the US, dedicated to local language development
- Citizenship initiative
- Long-term plan to bring key language component product development to Europe
- Expansion branch of the Redmond-based product development group (Speech Component Group), under Fil Alleva and Richard Sprague, responsible for speech in Microsoft



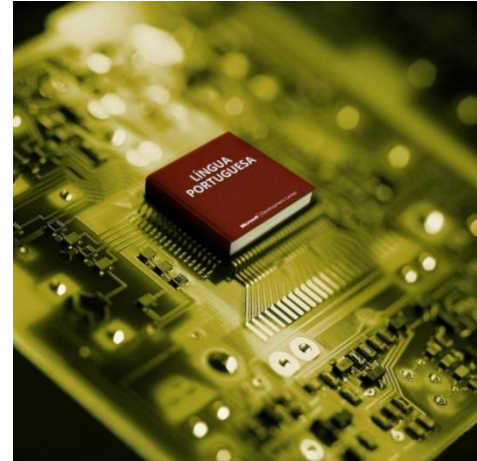




Long Term Mission of MLDC

Microsoft | Development Center
Portugal

- First Microsoft Development Center outside of Redmond dedicated to key Speech developments
- Clear demonstration of Microsoft citizenship efforts of stimulating a strong software industry in Europe
- Promotes language expansion in Europe for a range of Microsoft Products (Vista, Exchange UM, Communicator, Office, Mobility, Media Center, Xbox)
 - Starting with European Portuguese

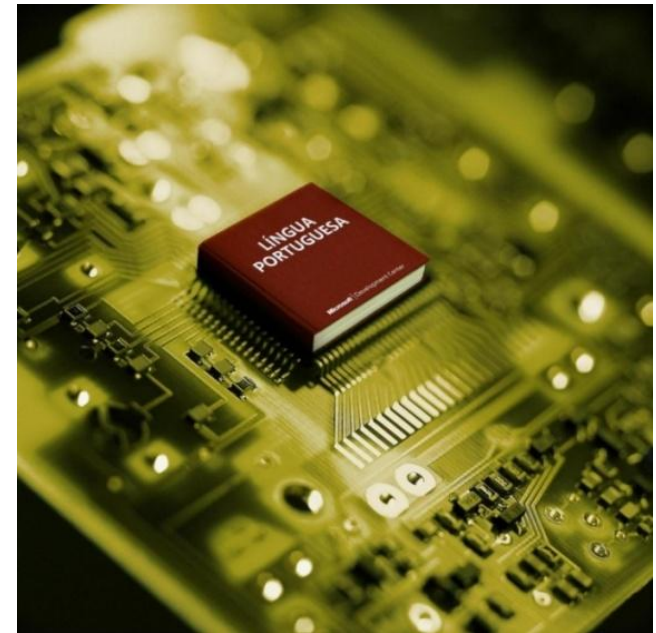


Microsoft **R&D facilities** in Europe

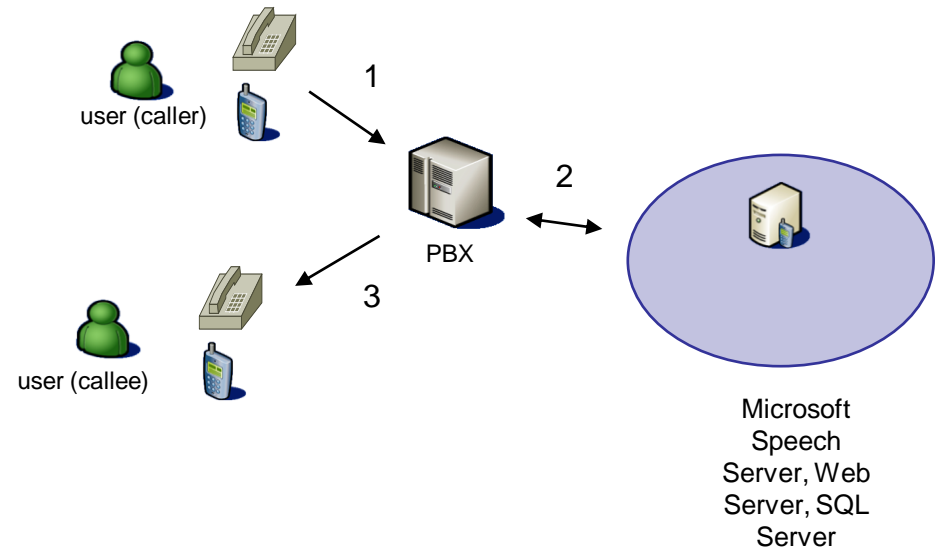
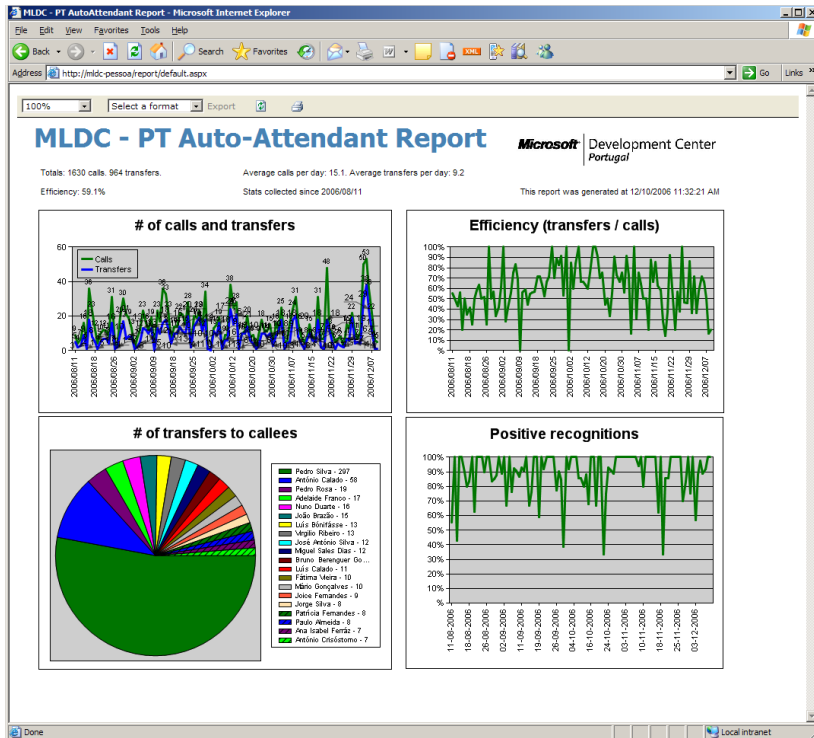


Action Lines of MLDC

1. **Close relationships** with academia, R&D laboratories, companies, government and European institutions:
2. **Basic speech development:** Speech recognition and Synthesis in Portugal, Western Europe and Latin America
3. **Citizenship:** Collaborative R&D:
 - National: FCT (1 submitted), ADI (**2 approved – 200 KEuro grant**)
 - European: FP7 (2 proposals under submission to FP7 Call 1)

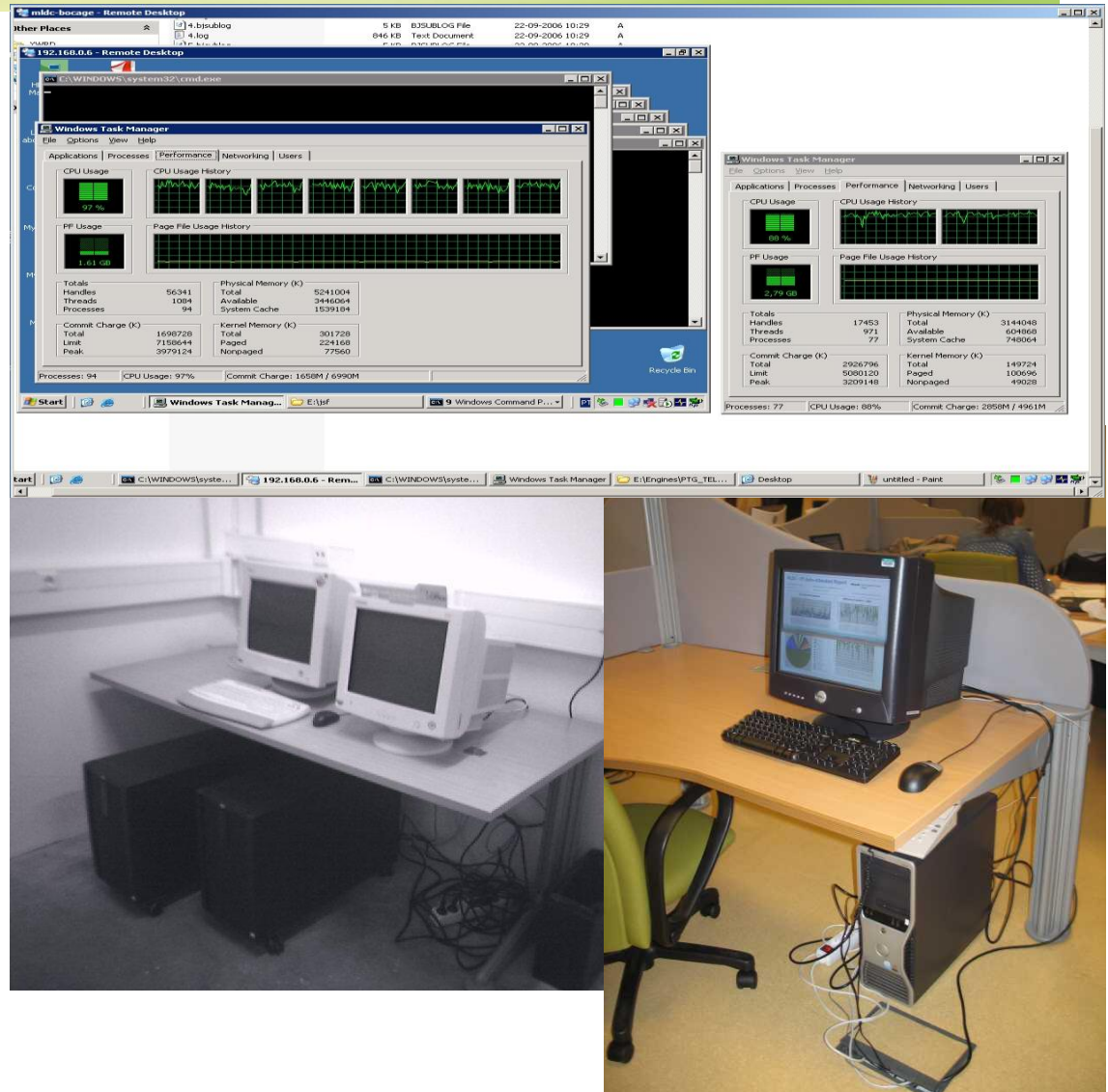


Results - Technology



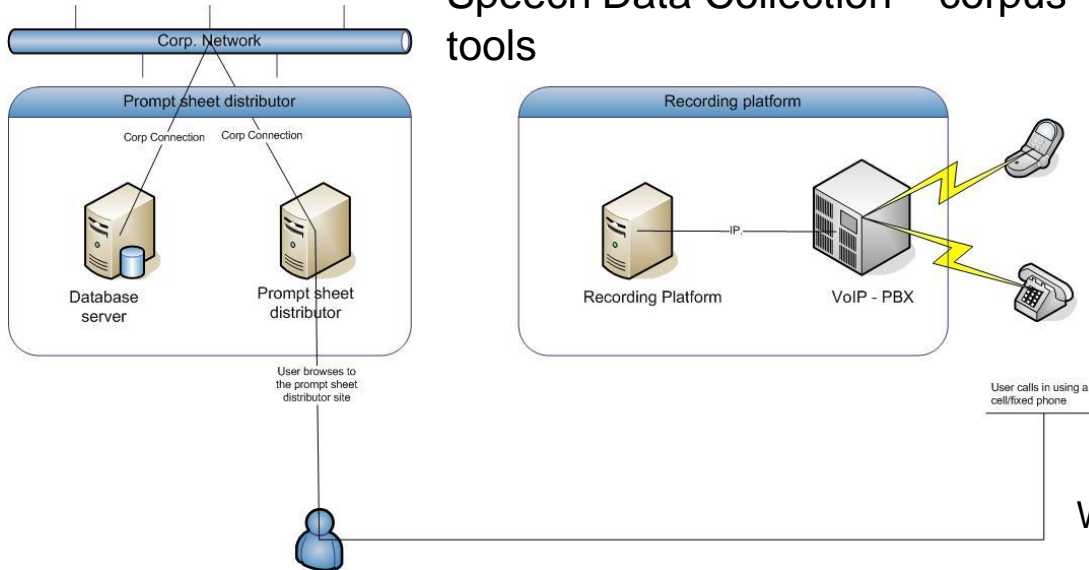
Results - Speech Processing Lab

- Acoustic Modeling
- TTS systems
- Data and code repository
- 12 CPUS
- ISDN and PTN telephony access
- Telephony speech applications (based in Speech Server)



Results - Technology

Speech Data Collection – corpus tools



Windows Media Center controlled by Speech



PTG News Reader



PTG Voice Command



Exchange Unified
Messaging



Results– Speech Ecosystem in Portugal

- Collaboration with virtually all **MSFT** divisions
- Contacts 1on1 with **11 Academic Institutes + 11 Companies**
- **11 I&D collaboration protocols** (some being finalised): INESC-ID (IMMI and L2F), GPCG, FCUL, FLUL, UBI , INESC INOV, ISEL, ADETTI-ISCTE, FEUP, ILTEC
- **NITEC** Contract with **ADI** – Agência de Inovação – Development of a Portuguese Speech Platform
- **IDEIA – UTILVOX** Contract with **ADI** – Portuguese Text-to Speech (TTS) Synthesis System – with FEUP + ILTEC
- Participation in **collaborative PF7 proposals** regarding Speech for accessibility and health care

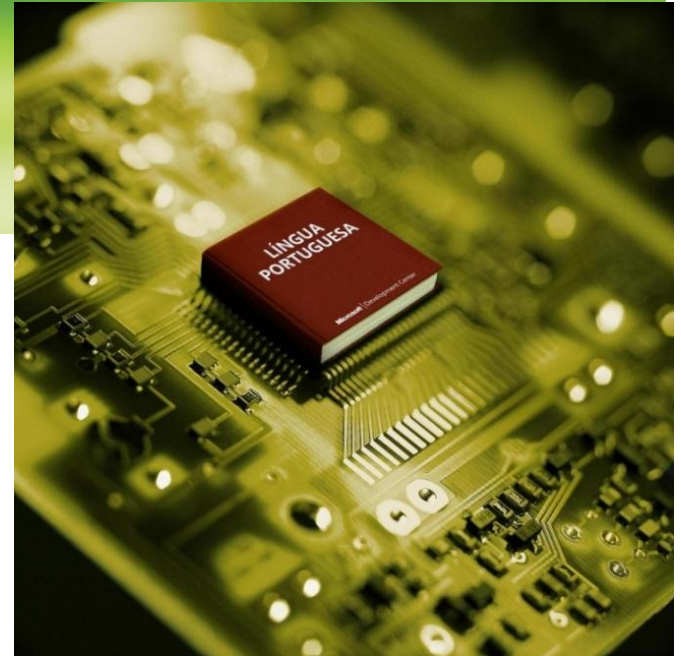
Challenges for 2007 and 2008

Microsoft | Development Center
Portugal

- **Language Resources and Tools:**
 - **Tools for data collection, validation and cataloging** (internal)
 - **Speech International Program (SIP):** in-house (Microsoft) Telephony Speech Data Collection Campaign in Western Europe for **Exchange 14 Unified Messaging** for 12 languages and 13 countries
 - **Language Resources** for the **Mirandese Language** - “Lhéngua” Mirandesa (with ALM, CELGA – UC, CLUL, INESC-ID)
 - **Language Resources** for **PALOP**
- **Modeling:**
 - **PTG Dictation model** (with LabEL-FLUL and PRIBERAM)
 - **PTG Desktop model** for Windows Vista wave 7
- **TTS:**
 - **PTB TTS**
 - **PTG TTS** (under ADI UTILVOX with FEUP and ILTEC)
 - Participation in **ECESS** - International Consortium for the exchange of TTS language resources, modules and systems
- **Products and Applications:**
 - **Exchange UM (PTB, PTG)**
 - **Speech Server 2007 (PTB, PTG)**
 - **Voice Command**
 - **Windows Media Center** controlled by Speech (Mic Array for stereo and surround)
 - **Speech Demos**
 - **PF7 consortia**



Microsoft Language Development Center in Portugal: an overview



I Microsoft Workshop on Speech Technology -
Building bridges between industry and academia, May
2 2007, MLDC

Miguel Sales Dias, Director MLDC

miguel.dias@microsoft.com

Muito obrigado pela vossa atenção!

PERGUNTAS ?

www.microsoft.com/portugal/mldc

Microsoft | Development Center
Portugal

