

### MLDC - Microsoft Language Development Center, Portugal

**Microsoft | Development Center Portugal**

- Criado um Microsoft | Development Center em Portugal
  - Criação em Novembro de 2005 e início da operação em Março de 2006
  - <http://www.microsoft.com/portugal/mldc>
  - Miguel Dias (Director, FTE) + 8 colaboradores (Engenheiros e Linguistas)
  - Um dos 4 Centros de Desenvolvimento Microsoft na Europa e o 1º fora de Redmond (EUA) dedicado ao desenvolvimento local da linguagem.
  - Expansão do grupo de componentes de processamento de fala da Microsoft, baseado em Redmond, EUA
  - Co-suportado pelo PRIME –NITEC





### Missão de longo termo e linhas de acção do MLDC

- Missão de Longo Termo
  - Expandir o desenvolvimento local das tecnologias de linguagem na região da EMEA, para um conjunto de produtos e plataformas Microsoft (Vista, Exchange, Office, Mobilidade, Media Center, Xbox)
  - Início na língua Portuguesa
- Linhas de Acção :
  1. Linhas de cooperação com as universidades e institutos de I&D mais inovadores em Portugal e na região da EMEA, nos domínios da fala e da língua natural
  2. Desenvolvimento de recursos e tecnologias da linguagem em Portugal e na EMEA
  3. Participação em projectos de I&D em consórcio nos programas Nacionais (FCT, PRIME-IDEA, PRIME-NITEC) e Europeus (FP7)




### Resultados - Tecnologia

- Pacote de Linguagem em Português Europeu e Brasileiro para aplicações telefónicas e Windows
- Euro PT Connect : Sistema de atendimento telefónico automático na subsidiária Portuguesa baseado em Microsoft Speech Server 2007 (300 nomes)
- Versão Brasileira pronta e em instalação na respectiva subsidiária (2007)




## Desafios para 2007 e 2008

### Projectos a decorrer:

1. **VoxCorp:**
  - Contrato com INESC-ID (L2F) para transferência de metodologias de aquisição de corpus de fala telefónica, visando a localização do Exchange UM – Comunicações Unificadas
  - Obtenção de outros corpus de Fala em Português Europeu, Brasileiro (com ISEL)
2. **Voice Command:** localização em Português Europeu e Brasileiro (com ISEL)
3. **TTS** ("Text-to-Speech"): Sistemas de Síntese de Fala em Português Europeu e do Brasil (com a Microsoft ATC - China)
4. **Media Center** controlado por fala (Português + Inglês) com cancelamento de eco e de áudio (com FCUL e Microsoft Research)
5. **Acoustics:** Reconhecimento de Fala de Inglês pronunciado por Portugueses (com FCUL)
6. Desenvolvimento de **ferramentas** internas para Anotação (TransSpeech) e Modelação Acústica (Lingo 1.0), com o grupo de fala de Redmond – EUA
7. **SpeechApps:** Demonstradores de Fala em Português Europeu para Windows (com Univ Coimbra, INESC-ID – IMMI, ISCTE – ADETTI e SBH – Solutions by Heart)



# DEMO

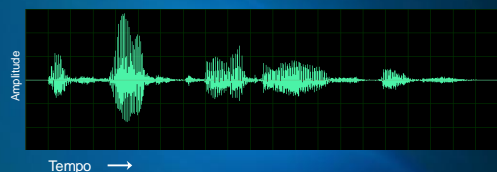
TalkToMe

## Agenda

- O Microsoft Language Development Center (MLDC)
- Tecnologias de fala:
  - Reconhecimento de fala.
  - Síntese de texto-para-fala.
- Desenvolvimento de aplicações Speech (client-side)
  - A managed API SpeechFX.
  - Desenvolvimento e demos.
- Desenvolvimento de aplicações Speech (server-side).
  - Office Communications Server 2007: Speech Server.
  - Desenvolvimento e demos.
- Downloads públicos e recursos:
  - Cliente + servidor
  - Bits da fala para Português!

## Speech Recognition

### Transformar



### Em:

“ É fácil nós reconhecemos fala”

## Speech Recognition - SR

- Ou Reconhecimento Automático de Fala.
- Características de um sistema de SR:
  - Modos de operação:
    - Comando e controlo,
    - Ditado (ou fala espontânea)
  - Dependência ao falante.
  - Adaptação ao falante.
  - Principais métricas de avaliação: **precisão e velocidade**

## Speech Recognition

- Como funciona
  - Hidden Markov Models: modelos estatísticos baseados em probabilidades.
  - A Fala é uma série de palavras.
  - Cada palavra consiste numa série de sons (fonemas).
  - Confidence scoring.

## Text-to-speech synthesis

- Transformar  
“É fácil sintetizar fala”
- Em



## Text-to-speech synthesis - TTS

- Síntese de texto para fala.
- Produção artificial de **fala humana**.
- Tipicamente, convertendo uma representação textual para fala num formato de audio.
- Como funciona? Técnicas:
  - Concatenative synthesis
  - Formant
  - Articulatory
  - HMMs
- A **voice font**: a fala do talento de voz armazenado como um conjunto de segmentos de sons individuais.

## Engines and Language Packs

- Microsoft **Speech** Technology.
- Dois principais core engines:
  - O engine de **SR**.
  - O engine de **TTS**.
  - Independentes da língua.
- **Speech Language Packs**: ficheiros específicos por língua.
- Tipicamente, LPs contêm:
  - language-dependent **recognizer** data.
  - language-dependent **synthesizer** data.

Tech  
2007 DAYS

Esteja presente no seu futuro!  
Inspire-se!

## Desenvolvimento de aplicações Speech (client-side)

## A nova API de Speech



- A API managed **SpeechFX**.
- O que está no namespace **System.Speech**:
  - *System.Speech.Recognition*
  - *System.Speech.Synthesis*
- Disponível publicamente no **.NET Framework 3.0**

## A nova API de Speech

- O que já vem no **Vista**:
  - Runtime .NET Framework 3.0, incluindo **SpeechFX**.
  - O **reconhecedor** de Inglês (Francês, Alemão, Espanhol, Japonês e Chinês também disponíveis).
  - O **sintetizador** de Inglês – a voz “Anna”.
  - O “Windows Speech Recognition User Experience”
- Para XP: download do .NF3.0



## System.Speech.Synthesizer

### • Como usar?

```
SpeechSynthesizer synthesizer = new SpeechSynthesizer();
synthesizer.Speak("Olá mundo!");
```

- Inclui suporte para custom spoken pronunciations, standard XML SSML W3C, gravar output para wave file, alterar velocidade de síntese e volume.

## System.Speech.Recognition

### • Como usar?

- Construir uma gramática.
- Carregar a gramática no reconhecedor.
- Registrar eventos (*SpeechRecognized*, *SpeechHypothesized*, *SpeechDetected*, ...)
- Começar o reconhecimento...!
- Inclui suporte para gramáticas complexas, semantic values, standard XML SRGS W3C, input de wave file, recognition confidence value, recognition alternate choices.

## System.Speech.Recognition

### • Como usar?

```
SpeechRecognitionEngine recognizer = new SpeechRecognitionEngine(new System.Globalization.CultureInfo("en-us"));

void init()
{
    Choices pizzaChoices = new Choices();
    pizzaChoices.AddPhrase("I'd like a cheese pizza");
    pizzaChoices.AddPhrase("I'd like a pepperoni pizza");
    pizzaChoices.AddPhrase("I'd like a large pepperoni pizza");
    Grammar pizzaGrammar = new Grammar(new GrammarBuilder(pizzaChoices));
    recognizer.LoadGrammar(pizzaGrammar);

    pizzaGrammar.SpeechRecognized += new EventHandler<RecognitionEventArgs>(PizzaGrammar_SpeechRecognized);
    recognizer.Recognize(.,);
}

void PizzaGrammar_SpeechRecognized(object sender, RecognitionEventArgs e)
{
    MessageBox.Show(e.Result.Text);
}
```

# DEV

Um cliente simples com SR + TTS.

# DEV

Integrar Speech numa aplicação existente: "Speech Wiki"

# DEMO

Viatecla "Brisa"



Esteja presente no seu futuro!  
*Inspire-se!*

## Desenvolvimento de aplicações Speech (server-side).

## Speech Server 2007 no Office Communications Server 2007

- Speech Server (2007) está incluído no Microsoft® Office Communications Server 2007
- Principais componentes:
  - Authoring and debugging
  - Reporting, Analysis and Tuning
  - Telephony
  - Operations, Administration, Maintenance

Microsoft®  
Office Communications Server 2007  
Speech Server

## Supported Languages

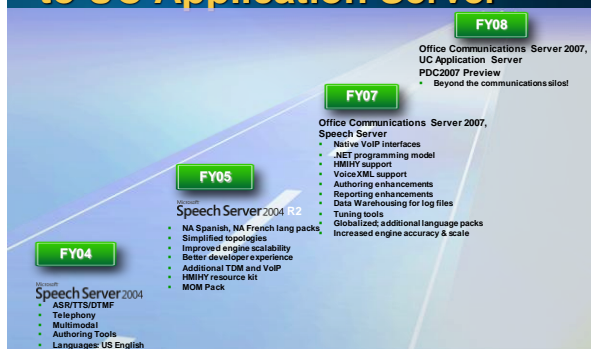
• Speech Server (2007) suporta **Speech Recognition** em:

- North American English
- American Spanish
- Canadian French
- UK English
- German

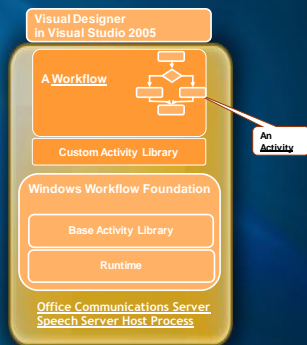
• Suporte planeado de **Speech Synthesis** no OCS2007 RTM:

Language	Country/Region	Ex UM TTS	Ex UM SR	OCS SR	OCS TTS
Chinese (Mandarin)	China	Available			Planned
Chinese (Traditional)	Taiwan	Available			Planned
Dutch	Netherlands	Available			
English	Australia	Available	Available	Available	Available
English	United Kingdom	Available	Available	Available	Available
English	North America	Available	Available	Available	Available
French	Canada	Available	Available	Available	Available
French	France	Available			Planned
German	Germany	Available		Available	Available
Italian	Italy	Available			Planned
Japanese	Japan	Available			Planned
Korean	Korea	Available			Planned
Portuguese	Brazil	Available			Planned
Spanish	Spain	Available			Planned
Spanish	Americas	Available		Available	Available
Swedish	Sweden	Available			

## From Speech Server to UC Application Server



## Speech Server (2007) uses the Windows Workflow Foundation



## Key Workflow Concepts

- Workflows are a set of activities
  - The workflow itself is an Activity
- Activities are the **building blocks**
  - Unit of Reuse
  - Unit of Execution
- An Activity has associated **properties, conditions and events**
- Developers can build their own **Custom Activity Libraries**
- A Workflow runs within a **Host Process**: a server or an application



## Dialogue Flow == Workflow

Like a sequential flowchart....



Or a state diagram....



**Speech Server only supports  
Sequential Workflow Development!**

## Speech Application Development Steps

1. Define the **Dialogue Flow**
  - Statements, Questions and Answers...
  - Other Activities
2. Specify possible Answers (**Grammars**)
3. Record Questions (**Prompts**)
4. Integrate into the back-end (**Web Services**)
5. Deploy, Test and **Tune** Application

# DEV

A "Agenda Telefónica do  
TechDays 2007"

## Get the bits!

- Para **cliente desktop**:
  - API **SpeechFX** no .NET Framework3 (incluído no Vista; download necessário para XP).
  - Language Packs estão incluídos no Vista.
- Para **servidor: Speech Server!**
  - Beta pública OCS 2007 Beta: 27 Março!
  - Language Packs estão incluídos.

## Language Packs de Português

- MLDC está a abrir **Programas de Avaliação Beta** da sua tecnologia!
- LPs Beta para **Português Europeu** em:
  - Client desktop:
    - Reconhecedor de Fala de Português.
  - Speech Server:
    - Reconhecedor + Sintetizador de Fala de Português.
- Invitation code: MLDC-BKBY-DTBD
- <http://connect.microsoft.com>
- i-pedros@microsoft.com

## E para terminar...

- Eu gostava...

## Questionário de Avaliação Passatempo!

- Complete o questionário de avaliação e devolva-o no balcão da recepção.
- Habilite-se a ganhar uma Xbox 360 por dia!



**DEV017**  
 Novas APIs para Reconhecimento e Síntese de Fala em Português para Aplicações Cliente e Servidor

**DEV017**  
 Novas APIs para Reconhecimento e Síntese de Fala em Português para Aplicações Cliente e Servidor

Muito obrigado pela vossa atenção

Miguel Dias  
 Miguel.Dias@microsoft.com, Director

Pedro Silva Santos  
 i-pedros@microsoft.com, Lead Software Development Engineer



**Microsoft** | Development Center  
 Portugal

**Tech**  
 2007 **DAYS**

Esteja presente no seu futuro!  
 Inspire-se!

© 2007 Microsoft Corporation. All rights reserved. This presentation is for informational purposes only. Microsoft makes no warranties, express or implied, in this summary.

