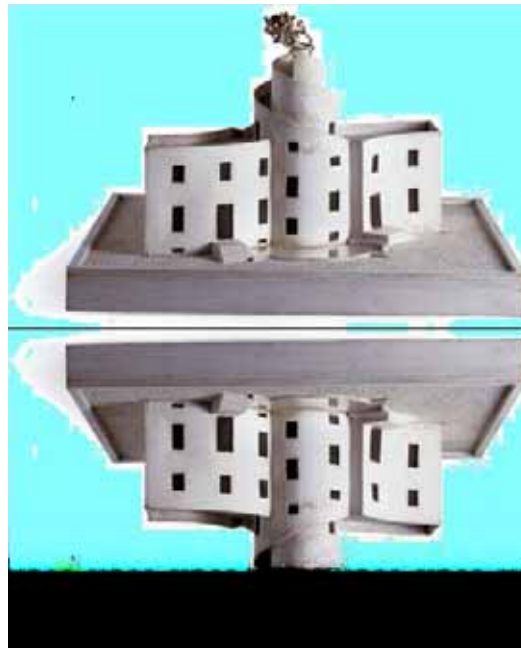


# IMMI イミ

Institute for Multilingual and Multimedia Information  
Institut des technologies Multilingues et Multimédias de l'Information  
Institut für multilinguale und multimediale Informationsverarbeitung  
[immi-labs.org](http://immi-labs.org)

A joint Research Unit between  
CNRS, RWTH and KIT



20 may 2010

<ul style="list-style-type: none"><li>• <b>Introduction</b></li><li>• Research Topics</li><li>• Structure</li><li>• Status</li><li>• Means</li><li>• Goals</li><li>• Actions</li></ul>
--

# Introduction

- Background (2005)
  - Political will at the highest level (President / Chancellor) to reinforce scientific cooperation between France and Germany
- Mission
  - Conduct research on
    - Multimedia (text, speech, sound, (image)...) document processing
    - Multilingual document processing
  - Major European/International challenge : access to information, whatever the media and the language in which it has been encoded
- 3 founding partners
  - LIMSI-CNRS, HLTPR-RWTH, Interact-Karlsruhe Inst. of Technology
  - Important participation in Quaero program (LIMSI responsible for Core Technology Cluster, RWTH for Corpus project)
  - Experience in research and innovation
  - Similar approach / resources and evaluation
- Location
  - Choice to settle the Institute in France

# Research topics

- Introduction
- **Research Topics**
- Structure
- Status
- Means
- Goals
- Actions

- « Natural » language processing
  - Text analysis, filtering, indexing, automatic summarization,...
- Machine Translation based on statistical approaches
  - text to text, speech to text, speech to speech,...
- Acoustic speech processing
  - Speech recognition, language identification, speaker diarization (who speaks when ?), speech understanding, audio classification, acoustic scene analysis, audiovisual speech processing,...
- Multilingual documents processing
  - Written and spoken language identification, crosslingual information retrieval, document translation,...
- Multimedia documents indexing
  - Information content extraction, content-based indexing, media surveys (alert...),...
- In perspective: Sign Language, image and video processing

# Structure

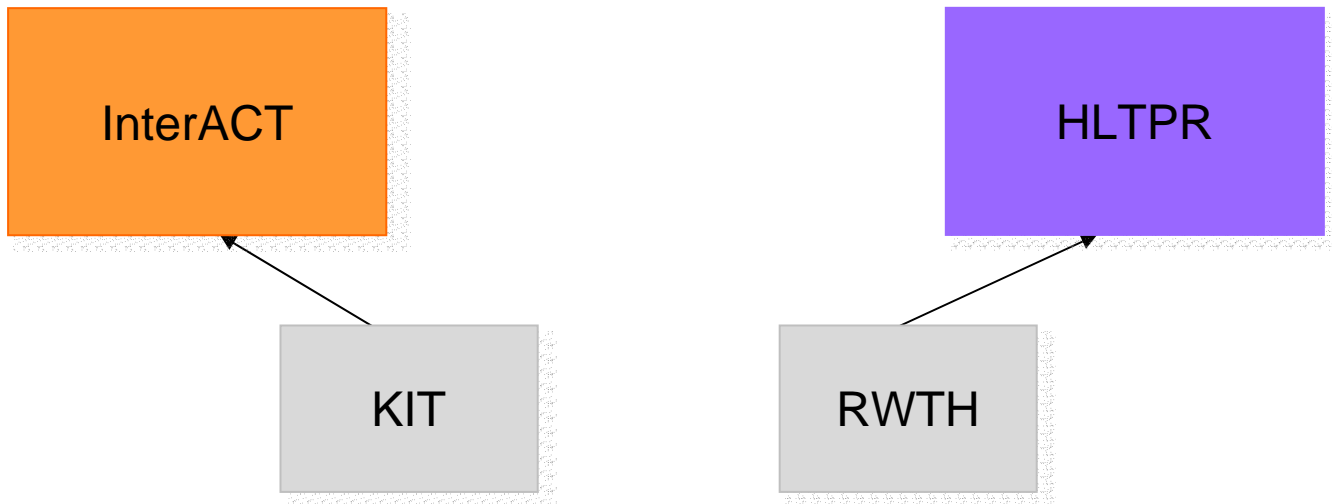
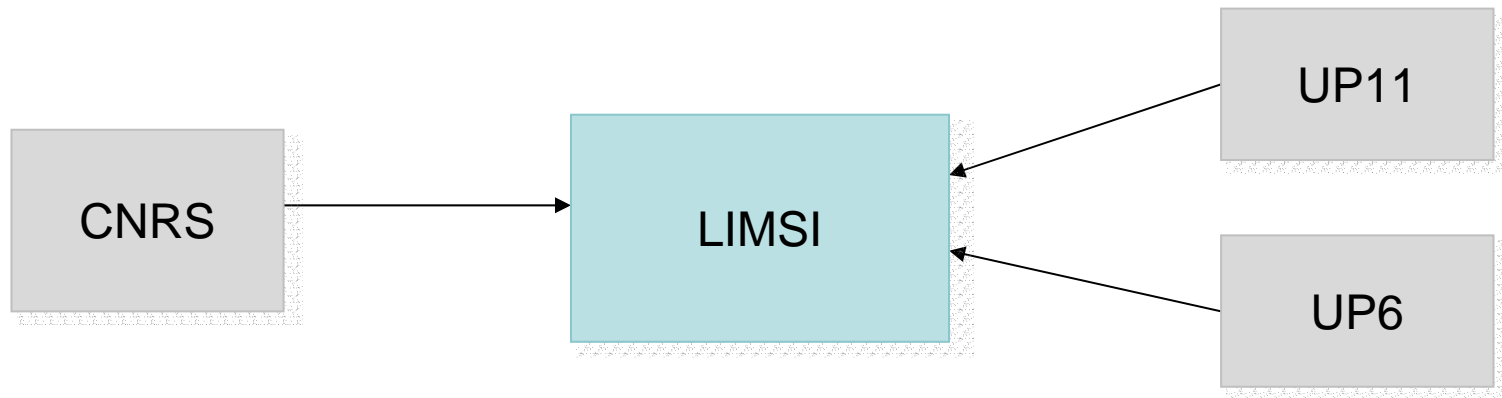
- Introduction
- Research Topics
- **Structure**
- Status
- Means
- Goals
- Actions

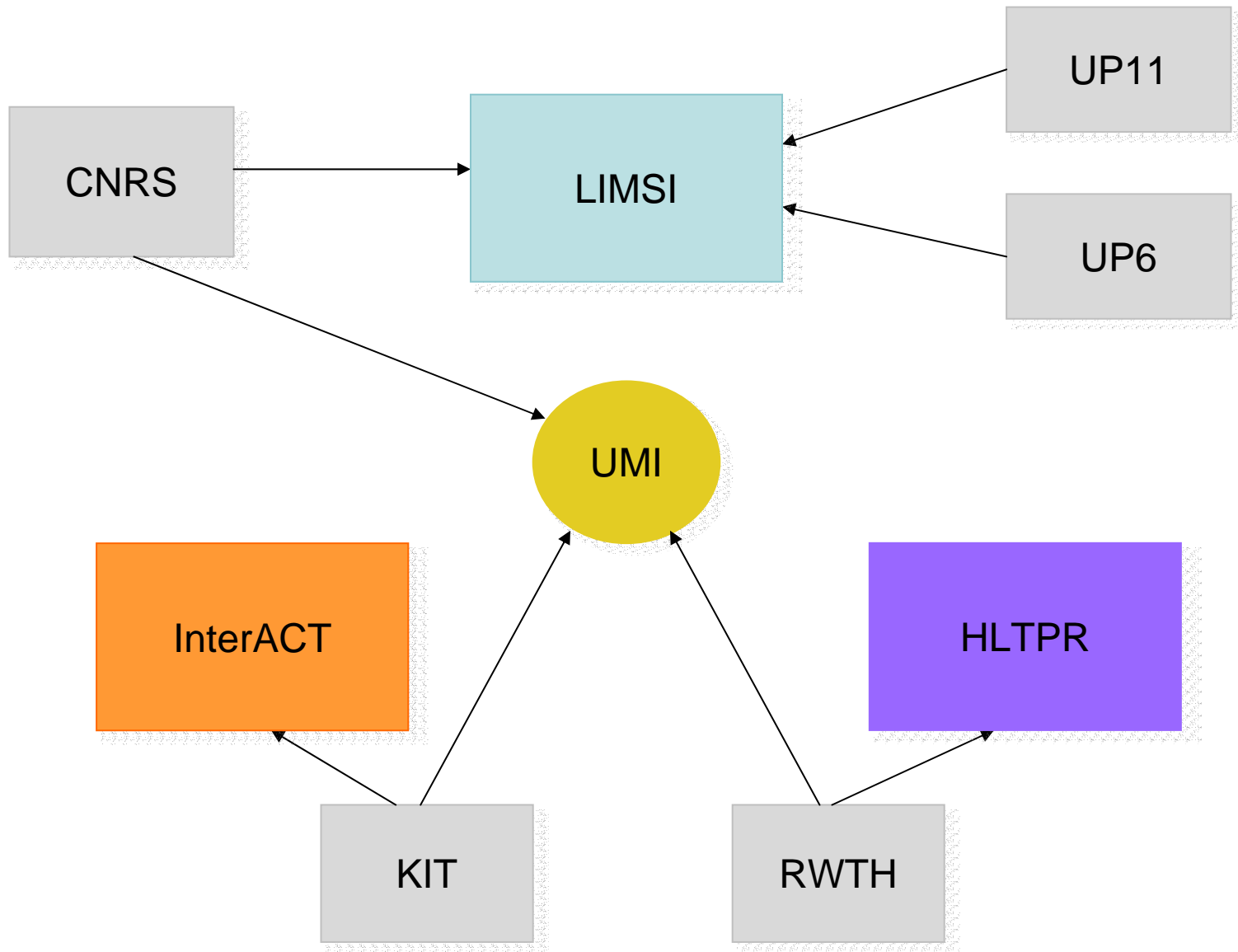
## ■ UMI

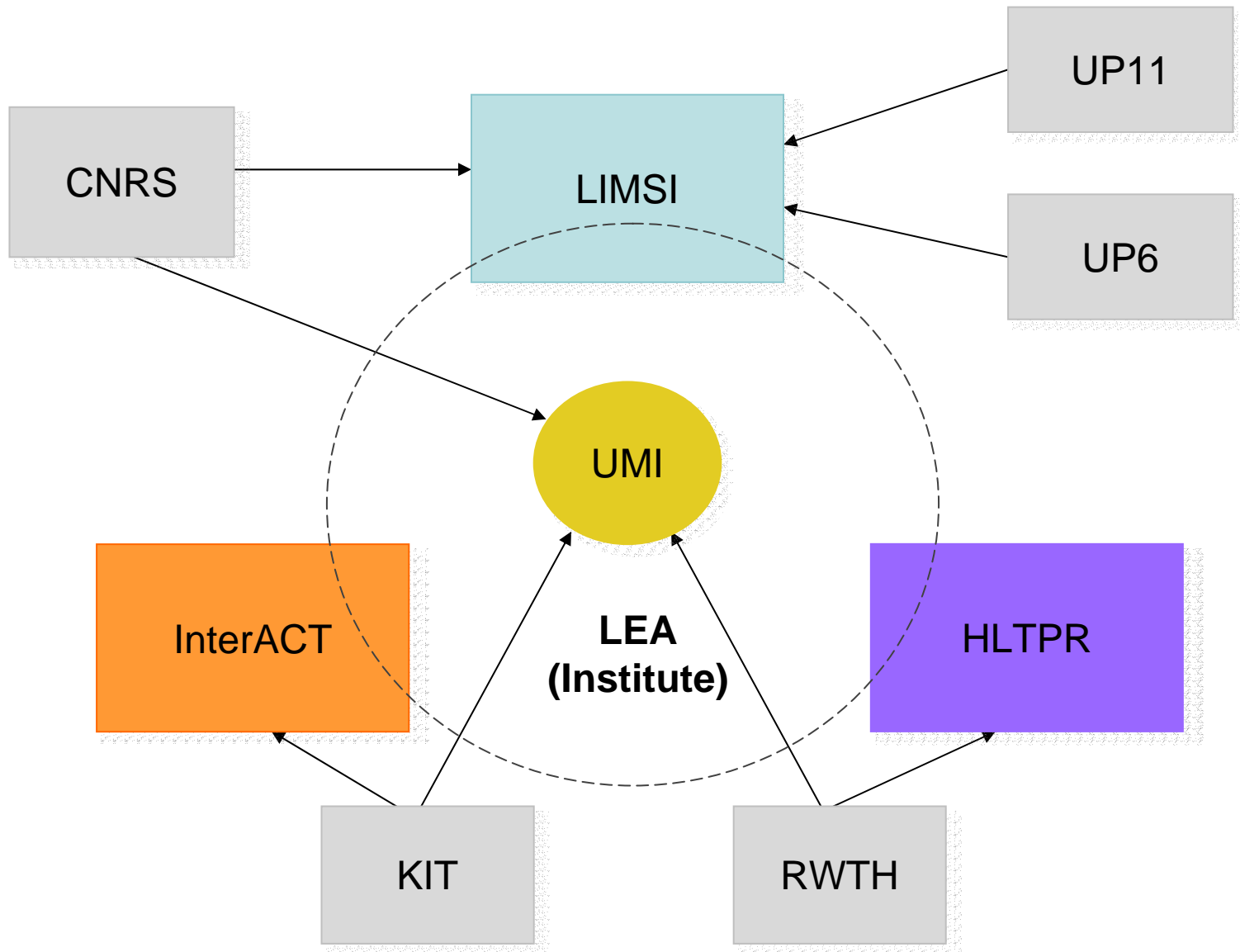
- Unité Mixte Internationale (*International Joint Unit*)
- Managerial, administrative and technical support to the Institute : Human Resources Management, projects management, industrial partnership, international cooperations..., computer facilities, data storage and management...
- 3 partners : CNRS, RWTH, KIT

## ■ LEA

- Laboratoire(s) Européen(s) Associé(s) (*European Associated Labs*)
- Coordinate the 3 laboratories research activities in the related scientific domain, with the support of the UMI
- 4 partners: CNRS, RWTH, KIT, University Paris Sud 11







# Status

- Introduction
- Research Topics
- Structure
- **Status**
- Means
- Goals
- Actions

## ■ UMI

- Agreement signed on December 19,2007
- Staff : 8
- J. Mariani (Director), A. Waibel (Associate Director), H. Ney (Scientific Dir.), J.L.Gauvain (Scientific Dir.)
- M. Garnier-Rizet (Scientific Information Coordinator), L. Barreteau (Management/Communication), A. Azhar (Computers), G. Francopoulo (Expert Researcher)

## ■ LEA

- Agreement signed on December 16, 2008



# Means

- Introduction
- Research Topics
- Structure
- Status
- **Means**
- Goals
- Actions

## ■ Staff

- 124 people in the 3 sites, as of Dec. 2009
- LIMSI: 1 Prof. 2007, 3 Assistant Prof. 2008-2009, 2 CNRS researchers (2008 + 2010) + Quaero recruitments (about 40 people)

## ■ Support to scientific cooperation

- 100 K€ per year from CNRS (INSIS + INS2I + DAE), KIT and RWTH

# Means

- Introduction
- Research Topics
- Structure
- Status
- **Means**
- Goals
- Actions

- Benefits from the dynamics of Paris region in ICT
  - Digiteo Cluster of Excellence (RTRA) in ICT
  - Ile de France Competitiveness Clusters (System@tic, Cap Digital)
  - National Priority Program (OIN), Campus Program, Universud & ParisTech Academic Clusters (PRES)...
- State-Region Project Contract (CPER) (2007-2013)
  - Building (3000 m<sup>2</sup>) : 9.8 M€
  - Computer equipment (computers, storage...) : 2.8 M€
  - Signed in march 2007
  - CNRS 7.6 M€, CG Essonne : 2.9 M€, RTRA Digiteo : 2.1 M€
- Quaero Program
  - LIMSI + RWTH + UKA = 18.5 M€ funding over 5 years
  - UMI IMMI = 1.7 M€ funding over 5 years
  - 20.2 M€ total over 5 years

• Introduction
• Research Topics
• Structure
• Status
• Means
• <b>Goals</b>
• Actions

# Goals

- Address the Multilingualism challenge
  - Preserve the culture for each country/region
  - Allow for communication among humans
- Develop corresponding technologies
  - Multilingual recognition / synthesis systems for many languages
  - (Spoken) translation systems for many language pairs
  - Crosslingual Information retrieval
  - Processing of low / non resourced languages, oral languages, regional languages...
- Fully participate in the European effort
  - Cf. FP7 FlaReNet Thematic Network
  - Cf. FP7 T4ME Network of Excellence
  - Cf. Multilingualism Summit (Etats Généraux) (26 Sept. 2008)
  - Cf. Euro-mediterranean program
- Play a major role on the International scene
- Ensure relationship research-society
  - International Conference on «Translators and MT» (Paris, 2011)

# Goals

- Address the Multimedia challenge
  - Video processing (subtitling, dubbing...)
  - Radio/TV broadcast processing...
  - Live performances (Surtitling...)
  - Sign Language Processing  
(analysis/synthesis/translation) ?

- Introduction
- Research Topics
- Structure
- Status
- Means
- **Goals**
- Actions

# Actions

- Introduction
- Research Topics
- Structure
- Status
- Means
- Goals
- **Actions**

- Ensure the availability of a state-of-the-art research environment
- Support the organization of joint workshops
  - International Workshop on Spoken Language Translation 2010 (December 2-3, Paris)



- Organize exchanges of researchers
  - Digiteo Senior Chair of Excellence for H. Ney (part time, 2010)
- Welcome postdoc researchers in an international environment
- Extend the partnership ?

# Actions

- Introduction
- Research Topics
- Structure
- Status
- Means
- Goals
- **Actions**

- First results
  - 90 publications (2008-2009)
  - Best results in Evaluation campaigns
    - - 1st in audio tracks at QA@CLEF (2008-2009)
    - - 1st in ASR at ESTER'08 (TechnoLangue)
    - - 1st in ASR at N-Best'08 (Stevin)
  - Innovation : Voice search
    - - Voxalead News (Exalead, Vecsys Research, LIMSI)
    - - Most Practical Demo Award (ACM MM Challenge'09, Beijing)
    - - Transcription in France24 HD Player
    - - Transcription Sarkozy's talks on Presidency Web site

# Thank you.

