



Asia-Pacific  
Economic Cooperation

---

**2006/SOM1/ECSG/SYM/005**

Agenda Item: 3

## **Human-Human Communication in Globalizing & Computerized World**

Purpose: Information  
Submitted by: United States



**APEC Symposium on Information  
Privacy Protection in E-Government  
and E-Commerce  
Ha Noi, Viet Nam  
20-22 February 2006**

# Human-Human Communication in a Globalizing & Computerized World

February. 20, 2006

Alex Waibel

International Center for  
Advanced Communication Technologies

Carnegie Mellon University

University of Karlsruhe

<http://www.interact.cs.cmu.edu>



....will know more about us  
....should know more about us  
  
... to serve us better.



- InterACT
  - International Center for Advanced Communication Technologies
  - Joint Center between Carnegie Mellon and University of Karlsruhe
  - Emerged from 15 year Collaboration
  - Launched January, 2004
- Mission of Center
  - To Develop Advanced Communication Technologies
  - To Facilitate Student Exchange and Training
- Major Ongoing Projects
  - CHIL – Computer Supported Human-Human Interaction
  - TC-STAR & STR-DUST & TRANSTAC & GALE –  
Speech Translation
  - TIDES & ASSIST &... - Text, Image Translation



# Phone Calls During Meetings





JEFF'S CONTEXT INFO		
Context	environment	UNKNOWN
	environment model	
	in smartroom? situation	YES MEETING
Current State		MEETING
Availability	Contact	Talk    Message
	personal	<input type="checkbox"/> <input type="checkbox"/>
	business	<input type="checkbox"/> <input checked="" type="checkbox"/>
	VIP	<input checked="" type="checkbox"/> <input checked="" type="checkbox"/>
Phone Alert	personal	MUTE
	business	MUTE
	VIP	EXCLUSIVE



# Memory Jog

....What was his name? ...Where did I meet him? ...What did we discuss last time?



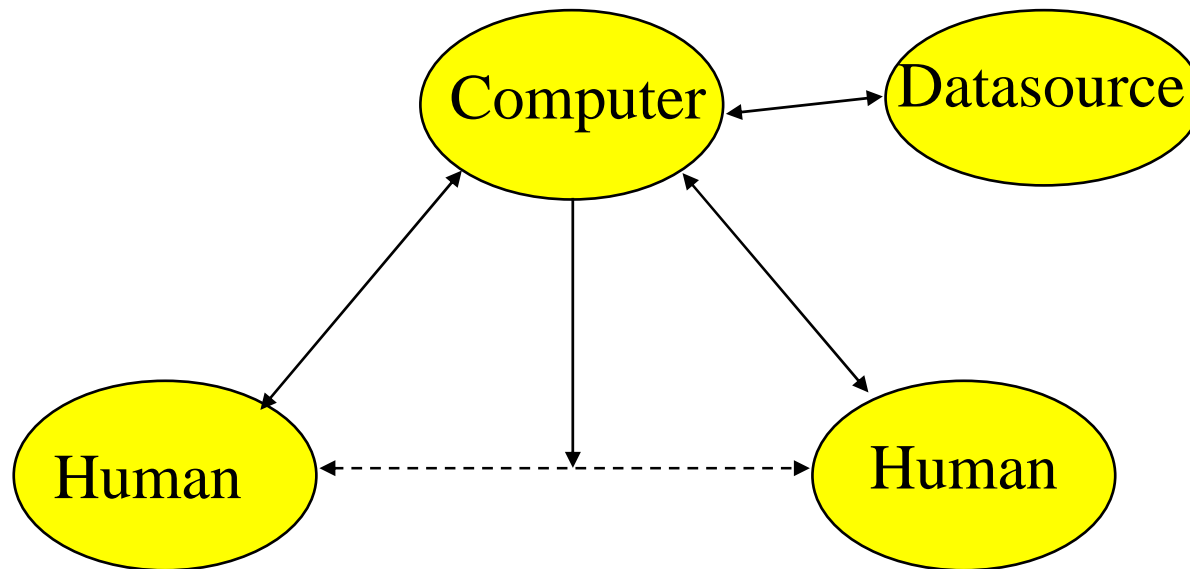




....and what is he saying?



你们的评估准则是什么



Goal: Technology to Improve Human Communication

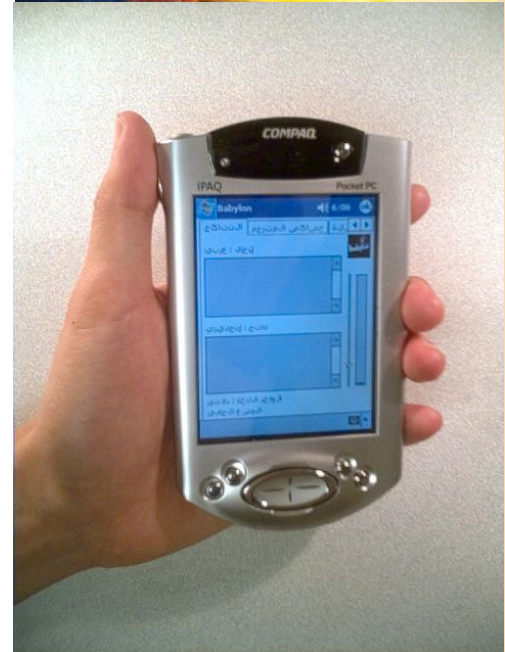
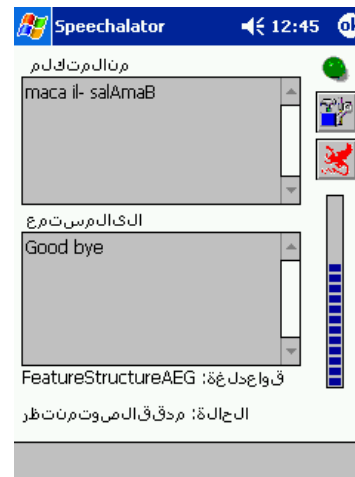
- **Making Computer Services Implicit/Invisible**
  - Computers in the Human Interaction Loop
  - Implicit Computing Services Supporting Human Communication
  - Observe, Use, Understand Human Communicative Context
  - Projects: CHIL
- **Speech Translation to Bridge the Language Divide**
  - From Domain Limited Fieldable Systems to Domain Unlimited Speech Translation
  - Usable by Anyone to Communicate Anywhere with Anyone
  - Projects: STR-DUST, TC-STAR, GALE
- **Support**
  - European Commission FP-6, Integrated Projects
  - US: NSF, DARPA



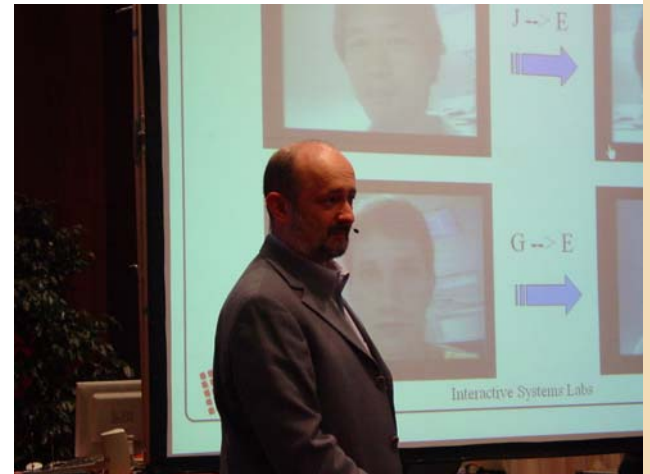
## Fieldable Systems:

### PDA Speech Translators

- Tourism
  - Conferences
  - Business
  - Olympics
- Humanitarian
  - Refugee Registration
  - First Responder
  - Healthcare
    - USA, Latino Population
    - Europe, Expansion
    - Third World
- Government
  - Peace Keeping, Police



- Applications:
  - TV, Radio, Lectures, Speeches, Meetings,...
- Technical Difficulty:
  - Open Domain, Open Vocabulary, Open Speaking Style
  - Spontaneous Speech, Disfluencies, Ill-Formed Sentences
  - Too Complex to Program Rules
- How it is Done:
  - Develop Statistical Learning Algorithms
  - Learn Speech and Translation Mappings from Large Example Corpora
  - With Increasing Data on the Internet, Improving Performance & Generalization
- Performance:
  - Depends on Language, but already Generally Understandable





MR PRESIDENT

señor presidente

*“Why did Joe get angry at Bob about the budget ?”*

Need Recognition and Understanding of Multimodal Cues

- Verbal:

- Speech
  - Words
  - Speakers
  - Emotion
  - Genre
- Language
- Summaries
- Topic
- Handwriting

- Visual

- Identity
- Gestures
- Body-language
- Track Face, Gaze, Pose
- Facial Expressions
- Focus of Attention



We need to understand the: **Who, What, Where, Why and How !**



- **Who & Where ?**

- Audio-Visual Person Tracking
- Tracking Hands and Faces
- AV Person Identification
- Head Pose / Focus of Attention
- Pointing Gestures
- Audio Activity Detection

- **What ? (Input)**

- Far-field Speech Recognition
- Far-field Audio-Visual Speech Recognition
- Acoustic Event Classification

- **What ? (Output)**

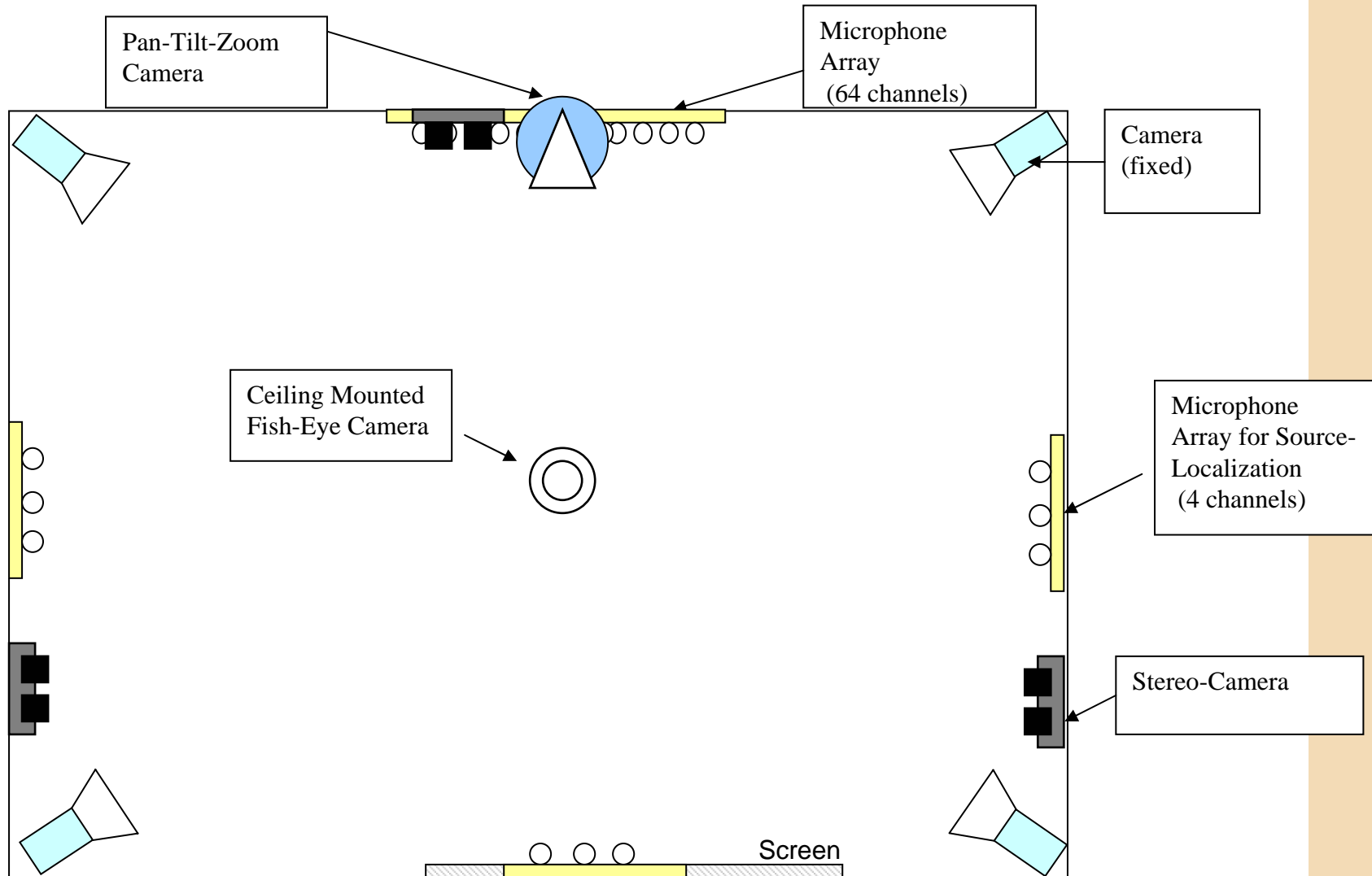
- Animated Social Agents
- Steerable targeted Sound
- Q&A Systems
- Summarization

- **Why & How ?**

- Classification of Activities
- Emotion Recognition
- Interaction & Context Modelling
- Vision-based posture recognition
- Topical Segmentation



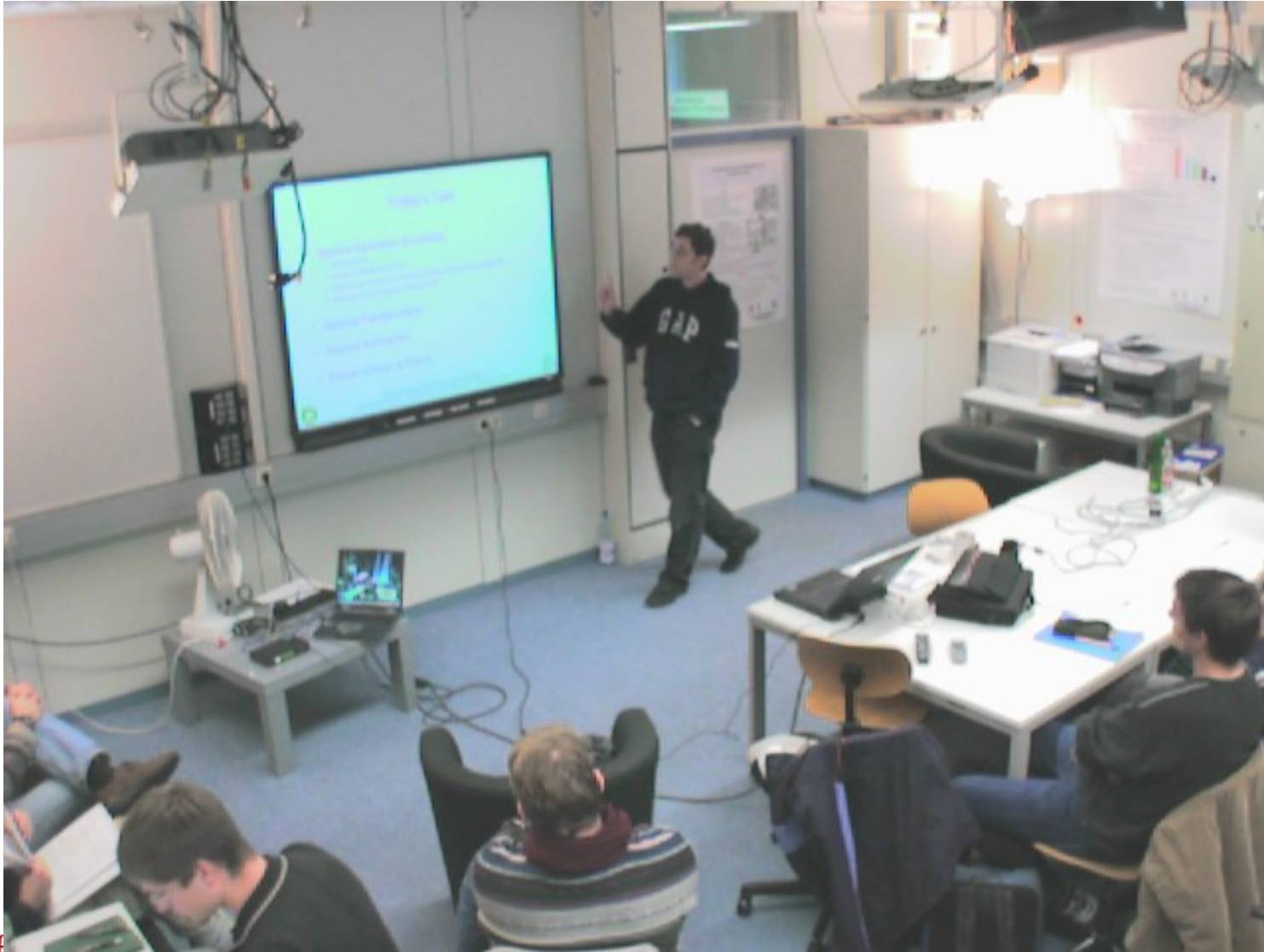
# Sensors in the CHIL Room

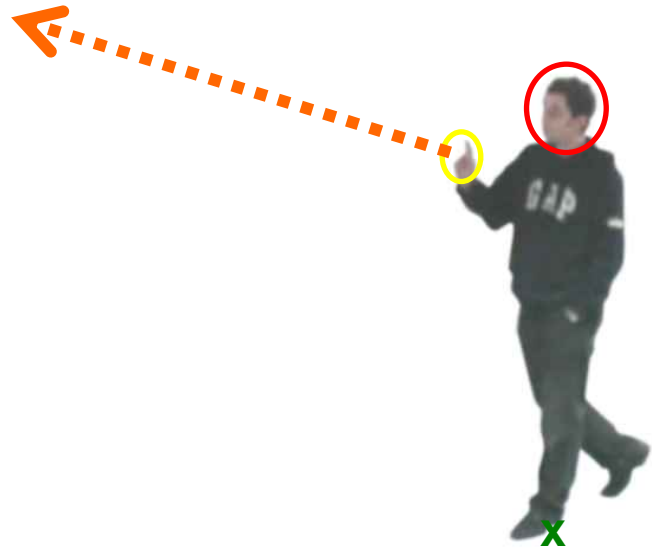


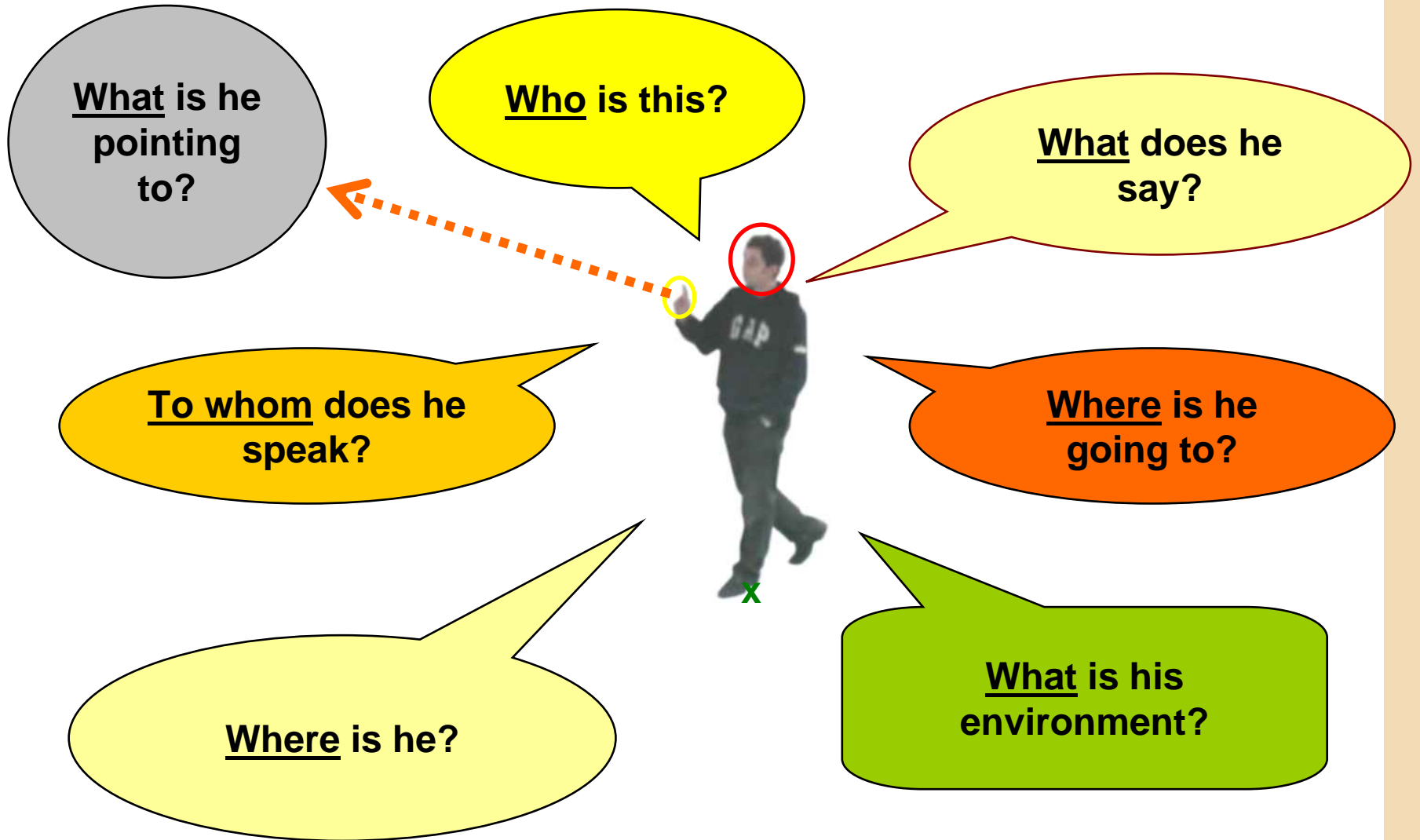


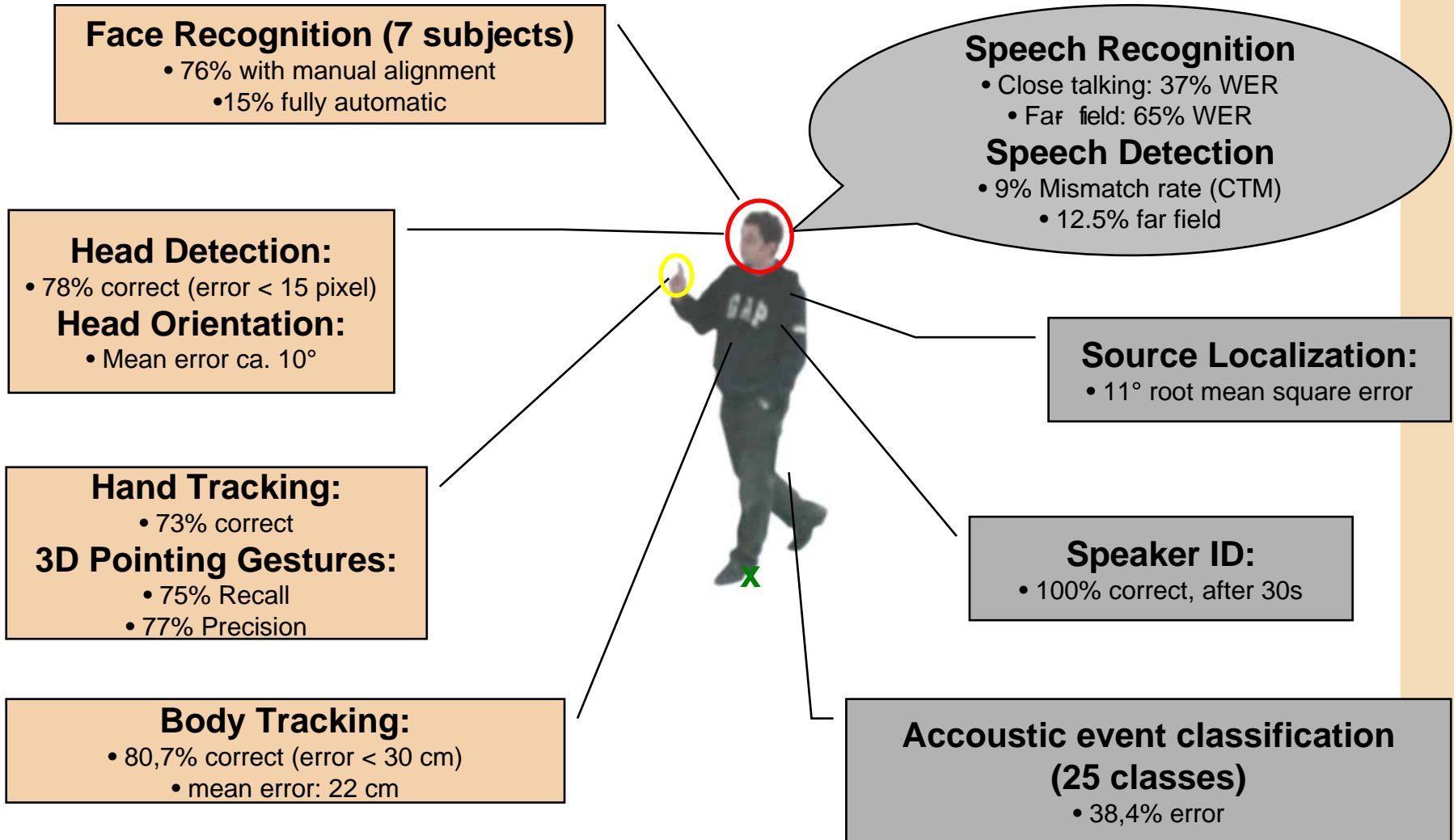


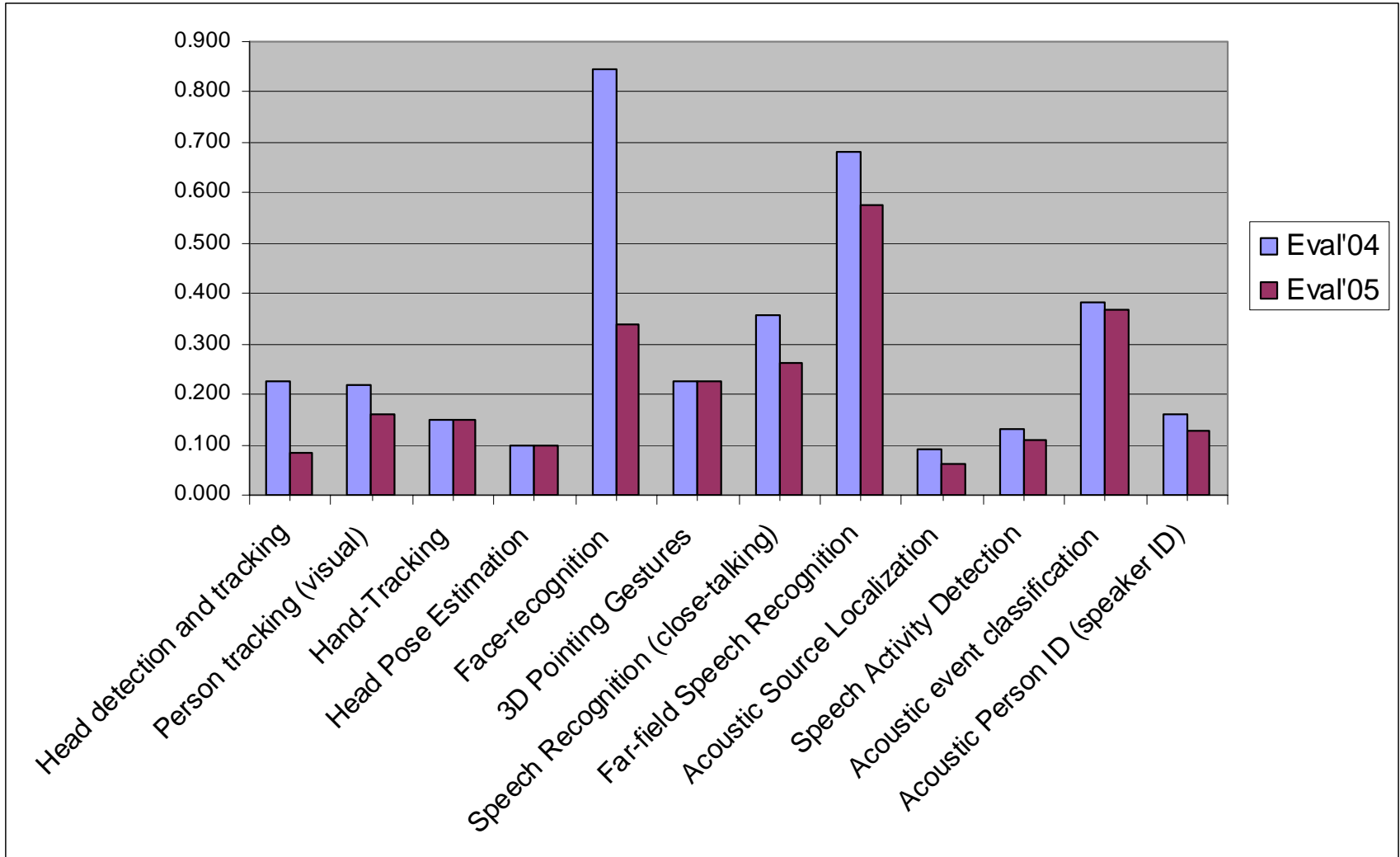
# Describing Human Activities













## Joint US (NIST) and EC Programs

- RT-Meeting'06 – Rich Transcription
  - Emerges from established DARPA activity
  - MLMI Workshops, AMI/CHIL
  - Evaluated Verbal Content Extraction
  - Chair: Garofolo (NIST)

- CLEAR'06 –

### Classification of Locations, Events, Activities, Relationships

- Emerging from European program efforts (CHIL, etc.) and US-Programs (VACE,..)
- First Joint Workshop to be Held in Europe after Face & Gesture Reco WS, April 6 & 7, Southampton
- Chair: Stiefelhagen (UKA)





## – Connector

- Connects people through the right device at the right moment

## – Memory Jog

- Unobtrusive service. Helps meeting attendees with information
- Provides pertinent information at the right time (proactive/reactive)
- Lecture Tracking and Memory

## – Attention cockpit

- Informs the current speaker about interest/boredom of audience

## – Socially Supportive Workspaces

- Physically shared infrastructure aimed at fostering collaboration

## – Simultaneous Translation Services

- Detect Language Need and Deliver Services Inobtrusively

## Private and Public Information Delivery

- CHIL phone
- Steerable Camera Projector
- Targeted Audio
- Retinal and Heads-Up Displays



- Trust
  - Is Translation/Information Believable?
    - Must Measure and Communicate Confidence
- Control
  - Is Implicit Service Correct/Desired/Wanted?
    - Must Occasionally Control, Communicate Intent
    - Balance Between Autonomy and Control
- Privacy
  - Is Information Secure and Private?
    - Personalized
    - Access Control

Interactive Machine Learning a Key Technology!

## Human-Human Communication Technologies

- Paradigm Shift in Human-Computer Interaction
- Potential for Significant Improvements for
  - Productivity Improvements
  - Comfort, Convenience, Safety
  - Humanitarian, Social, Multi-Cultural Integration
- Must Study Social Issues and Trade-Offs
  - Privacy, Trust, Control



