

# Newsletter

www.amiproject.org

AMI c/o IDIAP Research Institute, Simplon 4, P.O. Box 592, CH-1920 Martigny  
info@amiproject.org - www.amiproject.org

## Contents

### COVER STORY

AMI Stand at CeBIT 2006 1

### FOCUS

• Overview of AMI demonstrators 2  
• Overview of AMI demonstrators 3

### INSIDE AMI

• News and Upcoming Events 4

## News

### MLMI'06

The 3rd Joint Workshop on Multimodal Interaction and Related Machine Learning Algorithms (MLMI'06), jointly sponsored by AMI, Swiss NSF, and several other related EU projects will be held in Washington DC, USA, 1-3 May 2006.

**Additional information:**  
<http://groups.inf.ed.ac.uk/mlmi06>

© AMI Integrated Project, all rights reserved

## AMI Stand at CeBIT 2006

CeBIT, held each year in Hanover, Germany, is the world's largest technology fair. The AMI Project was offered an opportunity, by way of DFKI's Rienhard Karger, to participate in the BMBF pavilion on Human Computer Interaction at CeBIT 2006 March 9-15 in Hall 9 (Futures). AMI partners worked independently and collaborated for several months prior to the fair to develop an excellent set of materials and demonstrations representing the ongoing research on which this project focuses. At the event, 16 volunteers from DFKI, Philips, Brno Technical University, IDIAP, University of Twente, Fastcom Technologies and TNO staffed the booth and had many engaging conversations and demonstrations with members of the general public.

### CeBIT Quantified

Over the course of the week-long event more than 400,000 visitors came to CeBIT.

Many visitors were eager to see the research projects in the Futures Hall. For example, in the Human Computer Interaction pavilion there was the "smart automobile" and various types of robots. In the middle of this exciting environment, the AMI Project had 25 square meters in which to demonstrate to the CeBIT attendees how technologies under development in the partner institutions will impact their meetings in the future. Large plush-filled animals seated at a meeting table, high quality video on large Philips displays and large clear signs inviting people to learn about the AMI project in the booth attracted the attention of those walking in the aisles.

Approximately 480 flyers describing the AMI project in general and the demonstrations in the booth, were distributed to those who stopped by the booth. Half of the visitors to the booth were people in business who are managing technology for

their companies. Approximately one quarter of those with whom we spoke were the general public, including students, families and retired persons. Another quarter were from European governments, consultants, media/analysts and companies who could recommend AMI technologies or potentially transfer the AMI technology into their future products or services. One hundred and ten qualified leads were collected.

### Enthusiastic responses

All visitors to the booth were intrigued by the future we portrayed in the demonstrations of real time media processing and meeting archive browsing, and in the scenario depicted in the AMI vision video. Many were enthusiastic about the breakthroughs on which the AMI Project is working



From left to right:  
Tim Siglin (consultant), Marc Ruisjendaal (TNO), Pierre Wellner (IDIAP), Reinhard Karger (DFKI) and Tilman Becker (DFKI).

and wanted to know when there would be products available or where they could purchase or use such products today. We received numerous requests to evaluate the technologies in commercial environments.

We also conducted interviews with members of the press who can communicate our project's message to wider audiences and venture capitalists who are interested in funding companies which improve employee productivity.

Cover Story

## Overview of AMI demonstrators

### Overview

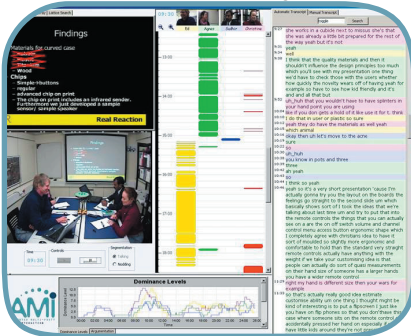
Some of the research underway in the AMI project lends itself to visual or aural representation. To more fully understand the way a user might benefit from a particular technology, an interactive demonstration or even a short movie is very helpful. Interactive demonstrations can be arranged with the most appropriate scientist or groups at AMI research facilities on an appointment basis. Many research initiatives are cross institutional. This article provides short summaries of the demonstrations currently available in the AMI project and which can be shown. To arrange a demonstration, contact AMI Technology Transfer. We will assist you to identify the AMI scientist or group best placed to answer your questions and to schedule an appointment.

### Key Word Spotting with Automatic Speech Recognition

One of the limitations of most commercial speech-to-text systems is that they need to be trained. They also are limited in that they usually work with only the person who trained (one participant). This is suitable for applications like broadcast news transcripts, however, in meetings there are many people and there is no training possible. In AMI, several partners are working on Large Vocabulary Conversational Speech Recognition (LVCSR) systems. In this demonstration the participant can type in a word that should be recognized and then begin recording a multiparty meeting. The LVCSR will recognize the word in real time and alert the system user. This is particularly compelling when in a meeting because approximately 40% of the verbal communication between people is overlapping. This means that the LVCSR has to detect a word and attribute it correctly to the person who spoke it.

### Meeting Archive Browsing

Finding elements of interest within a recorded meeting is time-consuming. Using the AMI meeting archive browser, JFerret, we support this process by displaying many types of data. Data types include synchronized media, slides and transcripts. We can also display processing results, such as a "synthesized" agenda, speaker segmentations, diarization and argumentation. Users interact with these visualizations to find elements and control synchronized playback of the recorded meeting. The extension of this is to be able to browse a corpus of meetings and to search among many meetings to find the most relevant pieces to answer a particular question.



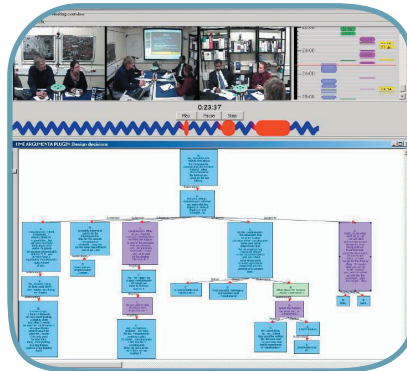
### Face Tracking

The face is one of the most expressive and important elements in human-to-human communication. Finding where a face is, counting the number of faces in a particular field of vision, and eventually identifying a person by their facial features are all important elements to rich meeting augmentation systems. The face tracking technology developed by AMI uses a web camera or any video capture device. We can demonstrate real time implementations of face tracking as well as several movies, showing the face tracker in action.

### Focus of Attention Tracking

One of the most important elements of a human-to-human communication analysis system is knowing what has the attention of the people (meeting participants). Using machine learning models based on the orientation of participant heads, the AMI project is developing technology that identifies who has the focus of attention (and who or what does not). We can show that the model is very sensitive to the head orientation and highly reliable at predicting the participant's focus of attention.

### Meeting Dialog and Argumentation Tracking



One of the most important and difficult-to-understand parts of a meeting is the discussion. People want to remember, indeed they frequently need to understand agreement, disagreement and the arguments given by different constituents, but this is not available today using existing technologies or the notes of individual meeting participants. At University of Twente, a method to capture and represent the discussion content in an intuitive and crisp manner has been developed. From the AMI meeting corpus, we can see how the argumentation representation is available in a typical recording of a future meeting. The lines in the tree depict relations between various points (positive, negative, request for more information) and permit the user to follow a discussion thread in a new and meaningful manner. Our demonstration also shows how the technology has been adapted for use in the JFerret meeting browser.

### Distributed Virtual Meeting Room

In the future, people may choose to meet in virtual meeting rooms using three dimensional rendering and systems developed by AMI to augment the verbal and non-verbal communications. Imagine, for example, that people wanted to meet in a "neutral" space, or that they did not wish to disclose information about their current surroundings. In the Distributed Virtual Meeting Room, all participants can benefit from verbal and non-verbal communications; there can be pro-active and reactive meeting support based on real time processing of the media streams, detecting events and activities. Virtual reality will permit higher and more comprehensive levels of visualization than is presently possible using current commercial technologies.



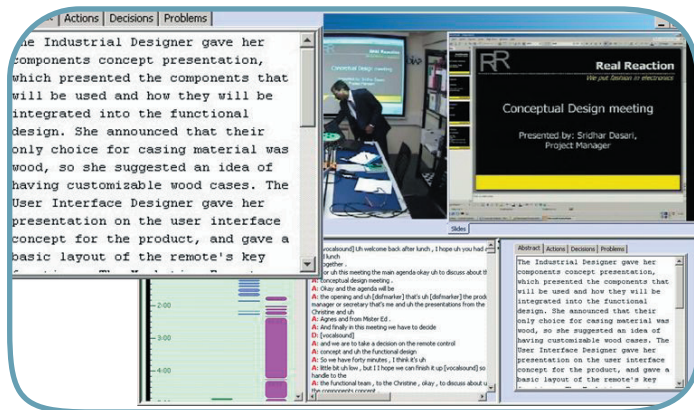


## Overview of AMI demonstrators

CONTINUED FROM PAGE 2

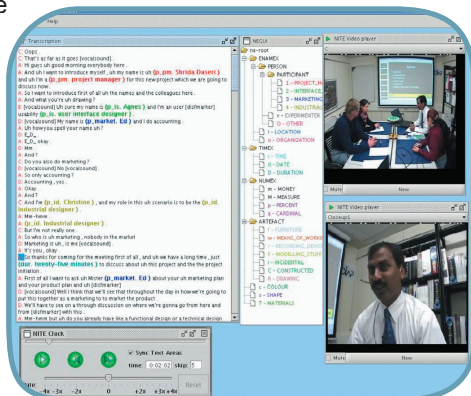
### Compressed Meeting Player

If we are to “catch up” or to synthesize meetings, we need systems which reduce the amount of time required to listen to a recording whilst retaining the important content contained in the meeting. The AMI project is working on several approaches to this problem. In the case of an audio-only archive or a situation where the user does not have a multimedia machine but only a low capability (audio only) player, the gist of a meeting can still be extracted via a novel compressed meeting recording. In this demonstration we show a research prototype in which a variety of compression techniques are available and compared. Furthermore, the system will allow for simple navigational controls and the ability to switch between the compressed and the uncompressed recording.



### NITE XML Toolkit for annotations

The NITE XML Toolkit (NXT) is new open source software for working with multimodal, spoken, or text language corpora. It is specifically designed to support the tasks of human annotators and analysts of heavily cross-annotated data sets. Written in Java, it includes data handling and search; libraries upon which to base end user interfaces for hand-annotation and data analysis that can highlight displayed data corresponding to the current signal time or that matches the result of a search; configurable GUIs for some common hand-annotation tasks such as creating simple time-stamped labellings and markup of named entities, co-reference, and dialogue acts; and utilities for data transforms that make it easier to combine use of NXT with NLP applications and other corpus toolkits.

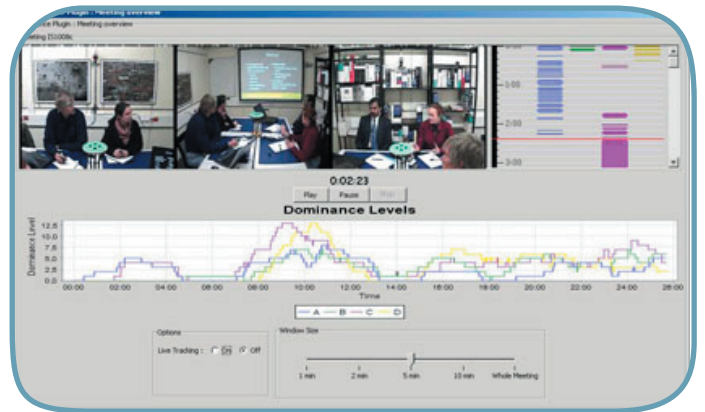


### Mobile Meeting Capture System

In some environments, for example when all people are remote, the meeting participants are already using a dedicated camera and a dedicated microphone for each person. Elsewhere, there may be several people sharing a room. In order to capture the audio and video of all the participants, a 360 degree capture system was developed using off-the-shelf technologies. In addition, the AMI project demonstration uses special algorithms that remove distortion caused by the parabolic mirror. It also is tightly integrated with real-time key word and event detection systems so that when using this system the speaker is visually identified and their words automatically captured and synchronized with the audio and video.

### Participant Influence

In many cases it is beneficial for the effectiveness of a meeting if people assume a cooperative stance. Cooperative conversations can be measured objectively using media processing. When cooperation is threatened due to one or more participants seeking control over the conversation, the facilitator or chair should facilitate the participants to have their say, to cut off people who make their contribution too long or to intervene when contributions are not relevant to the discussion at hand. Discussions should be properly organized to have arguments develop, so that all positions are put to the fore, and relevant pros and cons are raised. People that are too dominant or influential in meetings may violate one or more of the cooperative maxims and thereby frustrate the process of collective decision making for which many meetings are intended.



The HMI group of Twente University has developed a tool that is able to measure the influence levels of the participants. Aspects such as the interruption frequency and the number of floorgrabs are used to calculate the relative performance for each of the participants. The resulting measures can be used as input for live meeting assistance, as well as become part of accessible organizational memory. For the JFerret framework, a plugin has been developed that displays the conveyed influence levels over the whole meeting.

## News and Upcoming Events

### CESNET grant will increase CPU power at Brno University of Technology

CESNET is a non-profit organization unifying Czech Universities, responsible for high-speed University backbone and for research in network- and media-related technologies. The Development Fund of CESNET has granted Brno University of Technology (AMI partner) 770,000 CZK (26,000 EUR) to support its computing resources related to multimodal data processing. The financial means will be used to purchase 4 IBM blade servers each with two Opteron 270 processors (16 CPU-cores total). The machines will complete 28 servers with 56 CPUs (total value of 3.2M CZK, around 100K EUR) already on-line at BUT - these were purchased from AMI, other CESNET projects, and financial contribution of the Faculty of Information Technology of BUT.

The computing infrastructure at BUT (including more than 8 TeraBytes on disk servers) is serving for large-scale computations in speech and video processing. Discriminative training of acoustic models for large vocabulary continuous speech recognition, training of neural networks for phoneme recognition and different tasks in video processing are among the biggest CPU time «consumers». The new infrastructure will help BUT in two NIST evaluations organized this year:

- Rich Transcription 2006 Spring Meeting Recognition Evaluation where BUT participates on development of AMI-LVCSR system coordinated by the University of Sheffield,
- 2006 NIST Speaker Recognition Evaluation.

For more information, contact Honza Cernocky, [cernocky@fit.vutbr.cz](mailto:cernocky@fit.vutbr.cz)



*Tomas Kasperek, BUT's system administrator and Linux guru in BUT server room soon to be completed by new blade servers.*

### AMI-Partner TUM one of most successful candidates after pre-selection for German excellence initiative

On 23 June 2005, the German federal and state governments agreed on an initiative to promote top-level research in Germany. The Excellence Initiative aims to strengthen science and research in Germany in the long term, improve its international competitiveness and raise the profile of the top performers in academia and research. The total budget of the initiative will be €1.9 billion for the period 2006 through 2011, which will be split between three lines of funding:

- Graduate schools to promote young researchers
- Clusters of Excellence to promote world-class research
- Institutional strategies to promote top-level university research

Just recently, the results of the first pre-selection phase (based on the evaluation of short proposals) have been published and Munich has confirmed impressively its reputation of Germany's number 1 location for science and research: Out of 157 draft proposals for clusters of excellence, only 41 have been positively evaluated, and among those, seven come from Munich with three of them originating from TUM. Each university that has successfully passed the pre-selection stage is requested to submit full proposals by 20 April 2006. These proposals will be reviewed over the summer and final funding decisions will be made by the Excellence Initiative Grants Committee on 13 October. Eventually, 15 research clusters will be funded in this first phase. The most interesting TUM cluster application for AMI will be the research cluster "Cognitive Technical Systems". Although not directly focused on meeting analysis, this cluster will be still very closely related to the methods promoted in AMI, such as e.g. machine learning, perception and multimodal human machine communication, and thus could open interesting collaboration perspectives with the AMI project. Each successful cluster will receive an average funding of EUR 6.5 million annually.

More information can be found in the following press release:

[www.dfg.de/en/news/press\\_releases/2006/press\\_release\\_2006\\_03.html](http://www.dfg.de/en/news/press_releases/2006/press_release_2006_03.html)

This programme will make a significant contribution towards improving the quality of research at the selected universities, thus boosting their international profile significantly.

### ACM Multimedia Workshop on Human-Centered Multimedia Santa Barbara, CA, USA October 28, 2006

#### General information

Computing plays an essential role in supporting many important human activities, and it is at the core of the explosion of multimedia information (text, image, audio, video, and other sensor data) and devices (PDAs, cell phones, sensor networks) that are available or have even become part of human daily life, ranging from personal consumption and health care to professional activities and to social exchange. However, the methodologies and systems created by computer scientists often ignore issues related to the human use of the technology, whether it relates to the human-computer-interaction loop (e.g. in the development of methods to search, organize, and integrate multimedia data for human use, to make computers more accessible during interaction) or to the computer-in-the-human-interaction-loop (e.g. in the development of systems to support human communication that incorporate the social and cultural context in which such systems are deployed).

Human-centered computing (HCC) lies at the crossroads of multiple disciplines and research areas that are concerned both with understanding human beings and with the design of computational methods. Researchers and designers of human-centered computing methods and systems include engineers in multimedia information processing and systems, scholars in psychology, cognitive science, sociology, and graphic designers. The research in this area deals with understanding humans, both as individuals and in social groups, by focusing on the ways that human beings adopt, adapt, and organize their lives around computational technologies. The focus of the workshop is the multimedia aspects of the HCC paradigm.

#### Workshop Chairs

- Daniel Gatica-Perez, IDIAP Research Institute, Switzerland
- Alejandro Jaimes, FujiXerox, Japan
- Nicu Sebe, University of Amsterdam, The Netherlands

#### Workshop webpage:

<http://staff.science.uva.nl/~nicu/HCM2006/>