



- AMI @ a glance  
- Who's who: ICSI & USFD

- Tech News: AMI Meeting Rooms  
- Events & Conferences

## AMI @ a glance

The cover story in this issue of the AMI newsletter reports on the successful MLMI'04 Workshop, which was held in Martigny in June. The continuing series of partner portraits sees the spotlight on the teams at ICSI and Sheffield University, and the Tech News outlines progress on the AMI meeting rooms. Finally, recent and upcoming events are highlighted on the Events and Conferences page.

## MLMI 04 Multimodal Interaction and Related Machine Learning Algorithms Centre du Parc, Martigny, Switzerland, 21-23 June 2004

Building upon bringing together researchers from several related projects, (IM)2, the Integrated Project AMI, and the Network of Excellence PASCAL, IDIAP took the initiative to organize a large joint international workshop and summer institute on machine learning and multimodal interaction, held in Martigny (Hotel du Parc) on June 21-23, 2004. Fully sponsored by (IM)2, AMI, and PASCAL, this workshop was a major success, involving more than 200 participants, mainly from the sponsoring projects.



While including peer reviewed oral presentations (23), the workshop also included numerous (around 89) poster presentations, mainly by students. The workshop also featured 9 invited talks from prominent speakers such as: Dr. Bill Buxton (Buxton Design, USA), Dr. Jordy Cohen (VoiceSignal, USA), Prof. Roddy Cowie (Queen's University, Belfast), Prof. Stephen Cox (University of East Anglia, UK), Dr. Jonathan Foote (FX Palo Alto Laboratory, USA), Dr. Rafah Hosn (IBM Research, USA), Dr. Mats Ljungqvist (EU representative), Dr. Nuria Oliver (Microsoft Research, USA), and Prof. Yorick Wilks (Sheffield University, UK).

The workshop was considered by all attendees as a major success in terms of scientific quality, networking, and opportunities for the students. The EC representative also heartily applauded the initiative, while encouraging other related EC projects to join the future MLMI events.

The Proceedings of this first MLMI workshop will be published as a book by Springer, as part of the Lecture Notes on Artificial Intelligence (LNAI).



Based on the exceptional success of this first workshop, several candidates volunteered to organize the future MLMI workshops. NIST also expressed their interest to join MLMI and their "Rich transcription and meeting recognition workshop". Consequently, it has already been decided that the next MLMI'05 will be held in Edinburgh, jointly with the NIST workshop, while MLMI'06 will be organized by NIST and held in the US.



*The whole workshop was recorded, to provide further raw material for meeting processing.*



# AMI NEWS

Discover who's behind AMI's sucess!

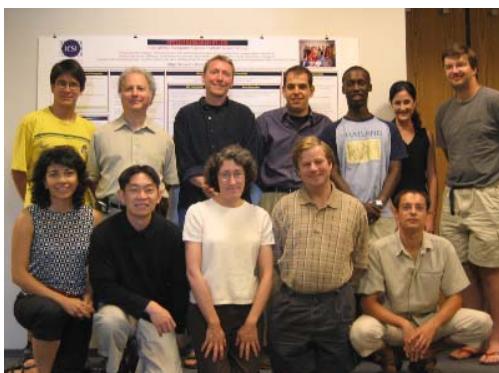
Each quarter, we will present two of the 15 partners that bring together their world class, skills and contribution to the AMI consortium. In the spotlight: ICSI and USFD's teams.



## International Computer Science Institute (ICSI), Berkeley, California - USA

The International Computer Science Institute is a non-profit research institute focused on fundamental, open research in computer science. Its speech laboratory is a center for innovation in speech technologies, with expertise in a broad array of speech processing activities, including robust signal processing, automatic speech recognition, speaker recognition and modelling, and higher-level analyses such as dialogue and discourse modelling.

ICSI has been a partner in numerous collaborations with universities, research labs, and industry, both at the national and international level, including several projects on multi-modal human-human and human-computer interaction. In addition, ICSI has hosted an international exchange program for over a decade.



From left to right, back row: Arlo Faria, Nelson Morgan, Chuck Woorters, Xavier Anguera, Kofi Boakye, Liz Shriberg, and Andreas Stolcke  
Front row: Nikki Mirghafori, Jeremy Ang, Barbara Peskin, Adam Janin, and Marc Ferras.

Of particular relevance to the AMI project, ICSI is one of the leading laboratories in the US exploring Meeting data, and its Meeting Corpus is the first substantial publicly-available corpus of Meeting Room recordings. ICSI's current research efforts in the Meetings domain include:

- automatic transcription of spontaneous multi-party speech
- exploration of novel signal processing techniques and associated acoustic modelling to provide greater robustness, especially in the presence of reverberation and noise
- automatic extraction of "metadata", such as speaker and sentence segmentation, disfluency detection and clean-up, and speaker modelling
- automatic annotation of dialogue acts and high-intensity "hot spots"
- structuring information from meetings to create "meeting maps" on a number of different levels, such as topic/content, speaker interactions, etc.

ICSI serves three main roles in the AMI project: as a data and resource provider, as a research collaborator, and as the host of an international visitor program. As a data provider, ICSI draws on years of experience in collecting and annotating meeting data, and has contributed its Meeting Corpus to the consortium. Though audio-only, this 75-meeting, 72-hour corpus is fully transcribed and we (and other labs now using the data) continue to enrich the corpus with new layers of annotation. As a research partner, we are actively engaged in a number of core technologies, developing constituent speech, speaker, and dialogue modelling capabilities; are regular participants in NIST's series of Meetings Recognition Evaluations; and serve as a connector to other US labs engaged in meetings research. Finally, ICSI is hosting an international exchange program, building on its existing visitor programs and providing educational and collaborative opportunities for visiting AMI students and scientists.

<http://www.icsi.berkeley.edu>



## Sheffield University (USFD), UK AMI Training Coordinator

Sheffield University is located in the heart of the UK's 4th largest city and is home to more than 23,000 students, from 116 countries, and 5,500 staff. The university's research partners and clients include Boeing, Rolls Royce, Unilever, Boots, AstraZeneca, GSK, ICI, Slazenger, and many more household names, as well as UK and overseas government agencies and charitable foundations. Our academic partners include leading universities around the world. International partnerships include Worldwide Universities Network (USA and China) and our partnership with Leeds and York Universities (the White Rose Consortium) has combined research power greater than that of either Oxford or Cambridge.

Work on the AMI project from Sheffield comes from two research groups in two different departments in the university. The speech and hearing group in the department of computer science focuses on work package four and its main areas of expertise are in large-vocabulary speech recognition, robustness in speech recognition and computational modelling of auditory and speech perceptions in humans and machines. The IR group in the department of information studies is focussed on work package six. The primary research areas of the group are cognitive and user-centric methods for design and evaluation, new statistical IR techniques, multimedia browsing and retrieval and personal information management and retrieval.

Contributions from the speech and hearing group to AMI are in the fields of emotion recognition and location tracking using audio-visual fusion.



# TECH NEWS

**Following much work over the first months of the project, the AMI meeting rooms are now operational. Below is an overview of the meeting rooms installed at the University of Edinburgh, TNO Human Factors, and the IDIAP Research Institute.**

The group is also a major contributor to the automatic speech recognition (ASR) track in work package four, and is currently involved in running a workshop to discuss issues surrounding the eventual ASR system for the AMI project. The group was a contributor to the Multimodal Meeting Manager project (M4) and will be building on the work carried out on the speech recognition system used in M4 for use in the AMI project. Contributions from the information retrieval group to AMI are in the fields of meeting browser evaluation and in the construction of browser components. The group is also interested in audio-only browsers and techniques for presenting meetings reduced amounts of time whilst retaining the significant content.

Sheffield also manages work package 10, the AMI training programme, whose objective is to provide opportunities for undergraduates, masters students, Ph.D. students and post-doctoral researchers to take part in AMI. The training programme allows AMI members and other researchers a bursary to visit AMI partner sites and assist the group at the site with their research.

SPandH and the IR group combined consists of 5 professors, 7 lecturers, 8 research associates and 14 students. Contributing to the AMI project are Phil Green (Member of the AMI Project Board, Sheffield Coordinator, Head of AMI training programme), Steve Whittaker (Browser Design), Guy Brown (Location Tracking), Thomas Hain (ASR Design), Yoshi Gotoh (Summarisation), Jon Barker (Tracking), Vincent Wan (Emotion Recognition), Stu Wrigley (Tracking) and Simon Tucker (Browser Design).

Spandh: <http://www.dcs.shef.ac.uk/spandh/>  
IR: <http://ir.shef.ac.uk/>

## AMI Meeting Rooms

Following much work over the first months of the project, the AMI meeting rooms are now operational. The meeting rooms are located at three sites: the University of Edinburgh, TNO Human Factors, and the IDIAP Research Institute.

To ensure consistency between data across sites, a basic common hardware setup was agreed upon. This was carefully designed to provide data for all the research needs within the project.



University of Edinburgh Meeting Room

## Augmented Multi-party Interaction

As a result, each meeting room contains a minimum of: 4 head-set microphones, 2 wide-angle camera views, a table-top unit consisting of an 8-element circular microphone array and 4 close-up cameras, a whiteboard capture device, screen capture device for a data projector, and 4 pen devices.

Additional devices at different sites include lapel microphones, extra microphone arrays, other wide-angle camera views, rear-projection screens, and a binaural manikin. All data sources are acquired synchronously and at high resolutions to facilitate research of emerging multimodal technologies. To ensure diversity of recorded data, the rooms differ in physical layout and views provided by wide-angle cameras. Future extensions to the rooms include the installation of VNC software on laptops to record desktop sessions.



IDIAP Meeting Room

With completion of the meeting room installation, data collection is currently commencing across the sites. Pilot recordings are being made throughout September and October, with the aim to start collecting meetings suitable for the AMI Hub Corpus by late October. This corpus will contain 100 hours of meetings, with a third of the data coming from each room. A first series of pilot meetings following the AMI Remote Control Design scenario have been made in the IDIAP room, and are available to partners via the Media File Server.

Several recordings of AMI Spoke Data have also commenced, including a set of meetings made using flock-of-birds sensors to facilitate research into gestures and focus of attention. Beyond the meeting rooms discussed here, a low-cost portable system developed at the Brno University of Technology is also ready to collect AMI Spoke Data.



TNO Meeting Room





# EVENTS & CONFERENCES

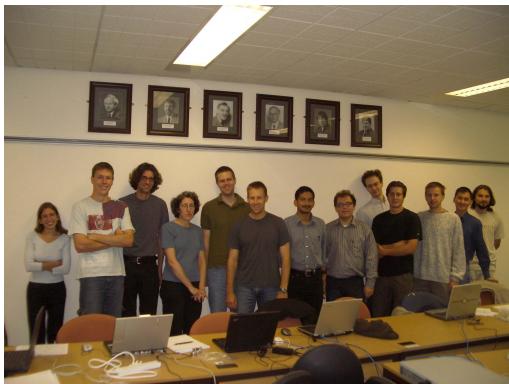
## Summer School

This year AMI has sponsored the Summer School organised by the Euromasters in Language and Speech, in Bonn.

Next year the plan is to attach a tutorial day to MLMI in Edinburgh, and again sponsor the Euromasters SS, which is being held in Edinburgh as well.

## AMI - ASR Workshop in Sheffield

A workgroup was held by the AMI ASR team at the University of Sheffield on the 13-14 September to continue developing a baseline recognition system for the project.



From left to right: Giulia Garau, John Dines, Roeland Odelmann, Barbara Peskin, Darren Moore, Iain McCowan, Jithendra Vepa, Thomas Hain, Jean-Christophe Lacroix, Martin Karafiat, Petr Schwarz, Vincent Wan, Lukas Burget

## W3C Workshop on Multimodal Interaction Sophia Antipolis, France, July 19-20, 2004

On 19 and 20 July, W3C organised a workshop on Multimodal Interaction, under the auspices of the IST Programme's MWeb project.

As the W3C Multimodal Interaction Activity is entering its third year, much progress has been made defining the W3C MMI framework, and the workshop was organised as an opportunity for discussing W3C's current plans and to provide feedback and suggestions for future work. Another goal of the workshop was to look for feedback from multimodal user communities that are currently less well represented in the W3C Multimodal work (e.g. automotive industry/telematics, home entertainment industry, healthcare, aerospace and manufacturing) as well as feedback and input from the Multimodal research community.

41 people attended the workshop and 19 papers were presented, both from academia and industry, contributing to a better understanding of each other's experience and requirements. The workshop concluded with separate "break-out" sessions, where participants gathered to discuss problems of more specific interest to them, such as multimodal interaction on mobile devices or advanced academic research.

The organisers are confident that the results of these discussions as well as the quality of the papers and presentations made this workshop very successful. All papers and slides are available online, but the organisers are still working on editing the minutes of the discussions, for publication in the next few weeks.

Workshop Page:

<http://www.w3.org/2004/02/mmi-workshop-cfp.html>

## UPCOMING EVENTS

### AMI Internal Review

On November 22nd AMI will hold its first Internal Review. The aim of this meeting is to review the project program with external boards, and update the program of work for the 2nd 18 month period.

The outline agenda is as follows

- Introduction
- Overview of Y1 scientific achievements
- for each work package manager:
  - Y1 achievements
  - Y2 plans and update of work package manager contribution to technical annex
- Feedback from advisory board, corrective actions, etc.
- General discussion and summary of possible major updates to the technical annex.

### NIPS Workshop on Multimodal Signal Processing. Vancouver, December 17-18, 2004

During this year's NIPS post-conference workshops, we are organizing a workshop on Multimodal Signal Processing. The workshop concerns the challenge of handling several sources of information at the same time (during training and decoding), mainly for sound and image integration, as well as wearable computing.

The aim of the workshop is to:

- Introduce problems related to multimodal integration to machine learning researchers.
- Identify and compare different integration strategies.
- Determine the major challenges in using more than one modality.
- Propose novel methods to improve the state of the art.

Submission deadline October 15.  
More information on workshop is available at :  
<http://www.imm.dtu.dk/multimodal>