Issue N° 17 / January 2009
Editor : Valérie Devanthéry, valerie.devanthery@idiap.ch

Newsletter

www.amiproject.org
AMI c/o Idiap Research Institute, Centre du Parc,
Rue Marconi 19, P.O. Box 592, CH-1920 Martigny
info@amiproject.org - www.amiproject.org

Contents

News

Community of Interest Workshop

5-6 February 2009 Edinburgh, Scotland

AMI Consortium partners are preparing for the next important Technology Transfer event—the 2009 Community of Interest (COI) Workshop.

Cover Story

Labeling and Analysing Speaker Turns in Conversational Speech

MINI-PROJECT: IDIAP RESEARCH INSTITUTE AND DEV/AUDIO

dev/audio produces software and peripheral devices that apply innovative computer speech technologies to enable more natural interaction between people and computers. A core aim of the company is to improve the usability of computers as tools for collaboration and communication, and therefore a focus is on technologies that support groups of users, rather than individuals.

As the basis of their initial product range, dev/audio is developing an intelligent microphone for group conversations. From a single table-top unit, the product automatically produces a separated high-quality audio track for each conversation participant, together with timing and location information that indicates where and when each person is speaking.

Dr. Iain McCowan, Director of dev/audio

Through its participation in the AMIDA Community of Interest, dev/audio is engaged in a project with the Idiap Research Institute to evaluate their voice identification software for inclusion in this product. This will allow the speaking person's name to be automatically associated with each audio track.

During the project, Idiap will first customise their existing software library to dev/audio's requirements. dev/audio will then integrate the software with their product and both partners will work together to confirm the accuracy of the voice labeling.

In addition, the project will investigate the inclusion of recent research performed at Idiap to automatically detect the dominant people in a conversation. dev/audio will first implement research methods that have been published by Idiap, and then both parties will conduct a trial to validate the automatic output from the product.

For more information about this work, contact Dr. Johnny Mariethoz at Idiap Research Institute, *johnny.mariethoz@idiap.ch* or Dr. Daniel Gatica-Perez at Idiap Research Institute, *gatica@idiap.ch*.

For further information about the AMI Consortium, visit: *http://www.amiproject.org/business-portal/*

If you have any questions about how to work with the AMI Consortium, please contact Christine Perey at *cperey@perey.com*

Dr. Iain McCowan

Augmented Multiparty Interaction Distance Access (AMIDA) is an Integrated Project funded by the EC's 6th Framework Program, jointly managed by Idiap (CH) and the University of Edinburgh (UK).

Issue N° 17 / January 2009
Editor : Valérie Devanthéry, valerie.devanthery@idiap.ch

Newsletter

Automatic storyboard generation for moderated large group chats

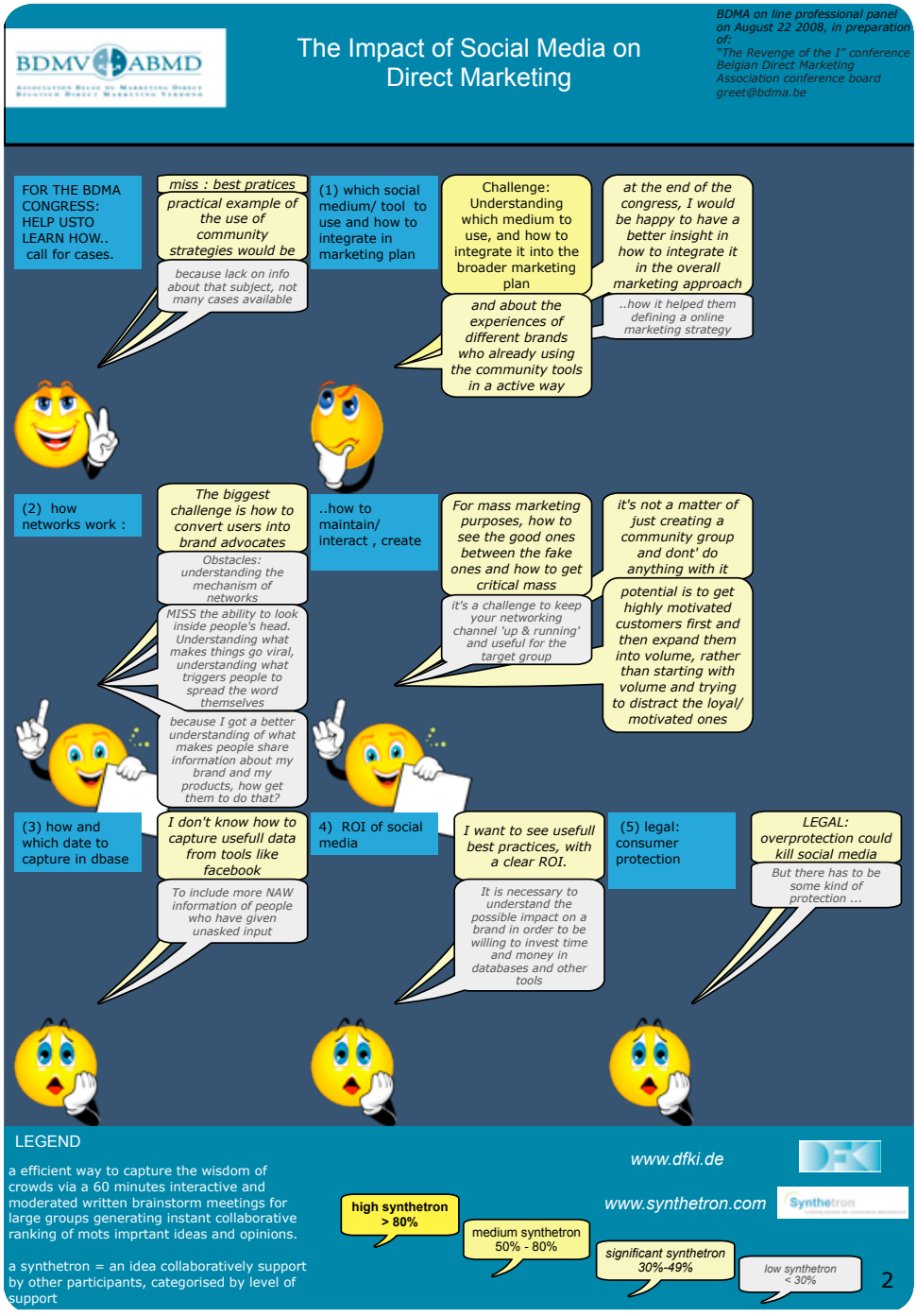
MINI-PROJECT: GERMAN RESEARCH CENTER FOR ARTIFICIAL INTELLIGENCE (DFKI) AND SYNTHETRON

Synthetron helps clients (corporations, consulting firms and communications agencies) to listen to, and get insights from, large groups of people in an efficient and timely way. Synthetron discussions are anonymous, interactive and conclusive. Participants join a discussion via their own PC from anywhere they like. The results of the discussion are collected in an automated way and then manually distilled by Synthetron consultants.

Synthetron has been looking for new ways of presenting the results of the group chats to their clients and was attracted by the storyboard summaries developed in the AMIDA project. These summaries have been presentend in issue 14 (April 2008, p.3) of this newsletter. Within the scope of this AMIDA mini-project, DFKI and Synthetron have designed and implemented a system that generates comic-like summaries from the final results of Synthetron discussions. To this end we have used the Storyboard Summary visualizer (SuVi) SuVi and have extensively modified it during the project in order to fit the needs of Synthetron. An example of the output of our system is shown here.

The system reads the structured discussion results from an Excel sheet and generates a comic-like layout in a fully automated process that can be controlled by a number of parameter settings. The layout process follows the structure of the input data and uses design knowledge to find an optimal distribution of the available data, typically on a two or four page leaflet.

The design knowledge which includes esthetical as well as usability rules is encoded in mathematical constraints and SuVi uses constraint-solving techniques to find an optimal solution. The results as shown here are generated fully automatically, however finishing touches can be applied manually in an external, visual editor.



Information is communicated by various visual clues: The style of the speech-bubbles represents the support of a statement within the group and the type of smiley assists the reader even more to quickly get an impression of the sentiment and importance of the statements.

The square text-boxes contain a short topic summary of what the statements are about. The core content is embedded in a page template that contains general information about meeting and a legend explaining the meaning of colour, fonts etc.

The first prototype of the system has already been used in a series of meetings in an actual client case and received a very positive reaction.

To be continued on page 3

Issue N° 17 / January 2009
Editor : Valérie Devanthéry, valerie.devanthery@idiap.ch
Newsletter

Automatic storyboard generation for moderated large group chats (continued)

The special format of these multimodal presentations mixing visual and textual information has proven to be a strong means of communication. Synthetron and DFKI are currently finishing up the project and have already begun negotiations for further, commercial development of the system.

For more information about this work, please contact Dr. Tilman Becker at DFKI GmbH, Dept. Intelligent User Interfaces, *becker@dfki.de*

Being a SME, and having followed AMI for several years, we are very enthusiastic with the mini-project experience.

It allowed us to hook on, at the right time and very efficiently, to some of the latest research results of the AMIDA project most relevant for us. We were able to translate these into a practical application with AMIDA researchers in very short time, testing the first "comics format" reports with several of our end clients in less then a month allowing pragmatic and quick cycle time.

By Joanne Celens, CEO of Synthetron N.V.

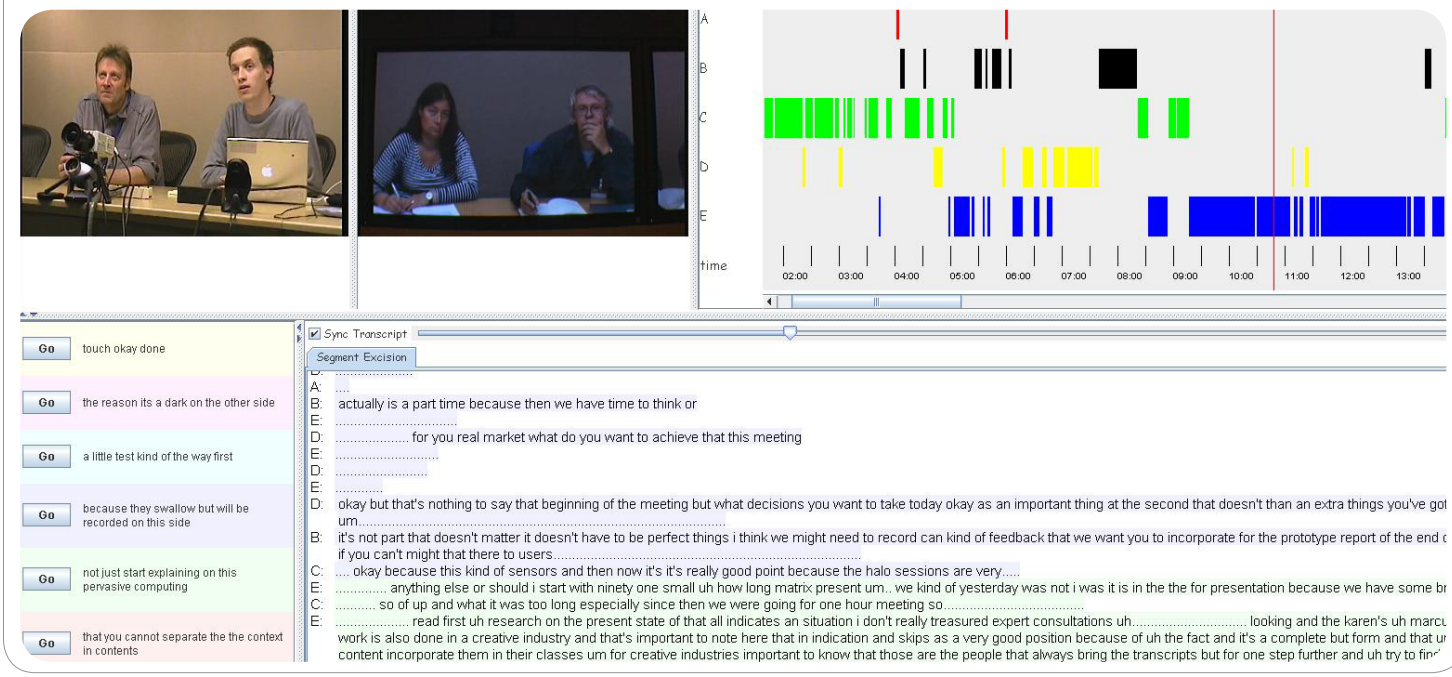
Incorporating and evaluating AMI meeting browsing technology with the HP Halo collaboration system

MINI-PROJECT: UNIVERSITY OF EDINBURGH AND HP LABS BRISTOL

Previous research, such as the Task Based Evaluation (TBE), has shown that AMI developed meeting browser technology is useful for participants in face to face meetings - the browser allowing participants to quickly and accurately extract information from previous recordings to aid in subsequent meetings. While these evaluations are important in terms of evaluating the browsers, it leaves a question over the portability of the analysis and browsing technologies for a number of reasons. Firstly, the meeting content is highly related to that of the AMI scenario meetings, on which many of the algorithms have been trained. Secondly, the meetings are conducted in the constrained environment of the AMI instrumented meeting rooms which provide ideal conditions for capturing media to subsequently display in a browser. Finally, the TBE participants were required to find information which could only be found by browsing the meetings and it is unclear whether more 'normal' meeting users would benefit as greatly from being able to review previous meeting content.

In this mini project, a collaboration between HP, The University of Edinburgh and the University of Sheffield, we aim to address some of these issues by taking the AMI browser technology out of the Instrumented Meeting Room and into the real world. We plan to record a series of meetings being conducted between participants in the HP labs Bristol and HP Amstelveen Halo suites. This data will then be processed using AMI technologies to generate a speech recognition transcript, automatic summary, and speaker and topic segmentation for the meeting. We will investigate methods by which, for instance, the recognition language model and dictionary may be adapted to improve speech recognition performance for a particular series of meetings. The processed meetings will then be supplied to the participants with the Jferret browser to be used in the course of their normal work. We will then obtain feedback from them in the form of interviews and questionnaires relating to their use of the recordings and any improvements they would like to see in the browser layout or content, which would then be implemented. In this way we aim to demonstrate the portability and adaptability of AMI technologies to new domains and recording conditions. In addition, the recordings may also be used by HP to demonstrate potential new technologies which could be implemented for Halo users.

For more information about this work, contact Dr. Mike Lincoln at the University of Edinburgh, *Mlincol1@inf.ed.ac.uk*



AMI c/o Idiap research institute, Centre du Parc, Rue Marconi 19, P.O. Box 592, CH-1920 Martigny,
info@amiproject.org - www.amiproject.org

Issue N° 17 / January 2009
Editor : Valérie Devanthéry, valerie.devanthery@idiap.ch

Newsletter

News

Second Review Meeting

December 4-5, 2008, Edinburgh

AMIDA's second review took place in Edinburgh's new Informatics Forum during the first week of December.

In line with the consortium's increasing focus on knowledge transfer and the generalizability of our technologies, most of the review was structured around a set of demonstrations. First and foremost came our real-time ASR system, since most other demonstrations either already rely on it or will before the end of the project. We showed two major technology concepts that had been subject to the entire design cycle from user requirements capture to demonstration, and will be formally evaluated in the next year.

Both are for use during live meetings. The first listens in to what is happening and retrieves relevant documents and segments of past meetings to aid the participants' memories. The second helps someone who is using desktop conferencing to dial-in to a meeting room to participate more fully, for instance, by making it clearer when people in the room are paying attention to him. In addition, we showed that our technologies can be deployed on consumer platforms using CE-HTML and Near Field Communication to move between devices on the fly. A range of other, more specialised demonstrations, some involving our mini-project partners, showed the breadth of our impact.

Jean Carletta,
University of Edinburgh

Events

Community of Interest Workshop

5-6 February 2009 Edinburgh, Scotland

AMI Consortium partners are preparing for the next important Technology Transfer event—the 2009 Community of Interest (COI) Workshop.

During the COI workshop the latest research results and emerging technologies, as well as our mature technologies, serve as the focal points for in-depth face-to-face meetings between AMI Consortium partners, the Community of Interest members and other interested organizations. Each participant will be provided a customized meeting schedule designed to meet his or her unique needs and interests. This portion of the workshop's agenda is designed to foster collaboration in small groups and provides numerous hands-on demonstrations.

Before, between and following the private meetings, there will be plenary sessions.

In the plenary presentations, workshop participants will discover:

- deeper context for our ongoing work,
- our emerging strategy for commercialization,
- demonstrations of the results of a COI mini-project conducted in 2008 and,
- results of our AMI Market Readiness Assessment survey.

We hope that those interested in learning more about the latest developments in AMI and who seek to integrate AMI Consortium technologies in their future products or services will participate in the workshop.

In order to achieve our objectives, the number of external (COI member and Friends of AMI) participants is limited to 35. Participation will be determined on a first come, first serve basis. Registration will close on January 22, 2009 and there is a 350 € fee.

For further information, visit *http://www.amiproject.org/ami-community-of-interest-workshop-2009*

If you have any questions, please contact Christine Perey at *cperey@perey.com*

Selected publications

Spoken Term Detection System Based on a Combination of LVCSR and Phonetic Search.
I.Szoke, M.Fapso, M.Karafiát, L.Burget, F.Grezl, P.Schwarz, O.Glembek, P.Matejka, J.Kopecky and J.Cernocky
In Machine Learning and Multimodal Interaction, 28.-30.6.2007, pages 1, Brno, CZ, 2007.

The ICSI RT07s Speaker Diarization System.
C. Wooters and M. Huijbregts
In CLEAR, pages 509-519, Springer, Baltimore, 2007.

Time-Compressing Speech: ASR Transcripts are an Effective Way to Support Gist Extraction.
S. Tucker, N. Kyprianou and S. Whittaker
In 5th Joint Workshop on Machine Learning and Multimodal Interaction (MLMI 2008), Utrecht, The Netherlands, 2008.

To separate speech! a system for recognizing simultaneous speech.
J. McDonough, K. Kumatani, T. Gehrig, E.Stoimenov, U.Mayer, S.Schacht, M.Wolfel and D. Klakow
In Proc. 4th Joint Workshop on Machine Learning and Multimodal Interaction, 2007.

Towards an Objective Test for Meeting Browsers: the BET4TQB Pilot Experiment.
A. Popescu-Belis, P. Baudrion, M. Flynn and P. Wellner
Lecture Notes in Computer Science, volume LNCS 4892/2008, pages 108-119, Springer Verlag, ISBN 978-3-540-78154-7, 2008.

Towards Automated Observational Analysis of Leadership in Clinical Networks.
I. McCowan and H. Harden
In Proc. Third International Conference Information Technology in Health Care (ITHC2007): Socio-technical approaches, 2007, pp. 133–141.

Toward Automatic Decision Detection: Empirical, Statistical and Machine Learning Approach
P.-Y. Hsueh
In Proc. MMKM Workshop (Multimedia Knowledge Management): Industry meets academia, 2007,

Unleashing the killer corpus: experiences in creating the multi-everything AMI Meeting Corpus.
J. Carletta
In Language Resources and Evaluation, vol. 41, issue 2, pp. 181–190, 2007.

Unsupervised Speech/Non-speech Detection for Automatic Speech Recognition in Meeting Rooms.
H. K. Maganti, P. Motlicek, D. Gatica-Pere
Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP), 2007.

Using audio and video features to classify the most dominant person in meetings.
H. Hung, D. Jayagopi, C. Yeo, G. Friedland, S. O. Ba, J.-M. Odobez, K. Ramchandran, N.Mirghafori, and D. Gatica-Perez
In Proceedings of ACM Multimedia, Augsburg, Germany, 2007, pp. 835–838.

Video Summarization at Brno University of Technology
V. Beran, A. Herout, M. Hradis, P. Chmelar, I.Potucek, S. Sumec and P. Zemcik
In ACM Multimedia, pages 16-19, Association for Computing Machinery, Augsburg, Bavaria, DE, 2007.

Virtual Meeting Rooms: From Observation to Simulation.
D.Reidsma, H. J. A. op den Akker, R. Rienks, R. Poppe, A. Nijholt, D. Heylen, J. Zwiers
In AI and Society, The Journal of Human-Centred Systems, vol. 22, pp. 133–144, 2007.

AMI c/o Idiap research institute, Centre du Parc, Rue Marconi 19, P.O. Box 592, CH-1920 Martigny,
info@amiproject.org - www.amiproject.org