AMIDA

Augmented Multi-party Interaction with Distance Access

`http://www.amidaproject.org/`

Integrated Project IST–033812

Funded under 6th FWP (Sixth Framework Programme)

Action Line: IST-2005-2.5.7 Multimodal interfaces

Deliverable D8.2 Updated Technology Transfer Materials

Due date: 1/10/2007 **Submission date:** 06/11/2007
Project start date: 1/10/2006 **Duration:** 36 months
Lead Contractor: IDIAP **Revision:** 1

Project co-funded by the European Commission in the 6th Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	✓
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

D8.2 Updated Technology Transfer Materials

Community of Interest Workshop

11-13 September 2007
Amsterdam, NL

Welcome to the
AMI Community of Interest Workshop

It is with great pleasure that the AMI Consortium brings together experts on the subject of business meetings, meeting technologies and meeting participant behaviors from around the world to join us for this workshop. Thank you for responding to our invitation.

During these two days the AMI Consortium will help our Community of Interest members and the Friends of AMI to experience the results of our research to date and to glimpse the technologies which are emerging. We will share with you the details of the exciting new Mini-Projects program and will invest our time and resources to increase our engagement with this important community.

The AMI Consortium partners are eager to watch and hear the reactions, the questions and the recommendations of experts in response to our presentations and demonstrations.

Interact with us! Be curious. Be sincere. Be critical. Be encouraging!

When, at the conclusion of this workshop we part ways, the AMI Consortium partners will seek to integrate the feedback we receive during these short days into our organizations on behalf of the community.

Two days is too short to achieve all that we have the potential to do together. We hope that this workshop is just one of many milestones we will share with the Community of Interest in the months and years ahead. And together we will lead (and observe) changes in how people work between and during meetings.

Your COI Workshop team

Table of contents

Workshop committee

This workshop is the collaborative effort of many people.

Please contact us any time if you have any requests or recommendations.

Christine Perey

Consultant to AMI Consortium
c/o IDIAP Research Institute
1920 Martigny, Switzerland

PEREY Research & Consulting
Route de Chernex 2B
1820 Montreux Switzerland

cperey@perey.com

+ 41 79 436 68 69 mobile (CH)

+ 31 6 24 92 60 68 mobile (NL)

Josie Scarr

TNO
Brassersplein 2
PO Box 5050
2600 GB Delft, The Netherlands

josie.scarr@tno.nl

+ 31 6 22 47 36 15

Ronny Müller

Technische Universität München
Institute for Human-Machine Communication
N0146
Theresienstr. 90
D-80333 Muenchen, Germany

mueller@mmer-systems.eu

+49 89 289 28550

Wilfried Post

TNO
P.O. Box 233769
ZG Soesterberg, The Netherlands

wilfried.post@tno.nl

+ 31 346 356 312

Marcia Bush

International Computer Science Institute (ICSI)
1947 Center Street, Suite 600,
Berkeley, CA 94704-1198 USA

mbush@icsi.berkeley.edu

+ 1 510 666 28 91

List of participants

COI Member	AMI Scientist	Friend of AMI	Workshop staff
Rod BACON	Tilman BECKER	Joanne CELENS	Christine PEREY
Michel BAIS	Hervé BOURLARD	Frank CRITTIN*	Josie SCARR*
Erik GEELHOED	Honza CERNOCKY	Johan DE HEER	
Edward GONEN	Daniel GATICA-PEREZ	Marten DEN UYL	
Jim HOOD	Benedikt HÖRNLER	Keith EDWARDS*	
Imre KISS	Alex JAIMES	Kamil MARES*	
Snorre KJESBU	Adam JANIN	Jean-François RAFFESTIN	
Keith LANTZ	Wessel KRAAIJ	Arjan VAN HESSEN	
Steven LI	Mike LINCOLN		
Gerald MORRISON	Ronald MÜLLER		
Christian RENAUD	Gabriel MURRAY		
Tim SIGLIN	Anton NIJHOLT		
Carlo TARANTOLA	Mannes POEL		
Julia TURKU	Wilfried POST		
	Stephan RAAJMAKERS		
	Steve RENALS		
	Stanislav SUMEC		
	Simon TUCKER		
	David VAN LEEUWEN		
	Alessandro VINCIARELLI		
	Job ZWIERS		

* Technology Transfer from AMI Consortium partner organizations

Wifi instructions

Your computer should detect a wireless network called: **RodeHoed**

It is a secure network. The SSID is: **RodeHoed**

The network authentication is **Open**

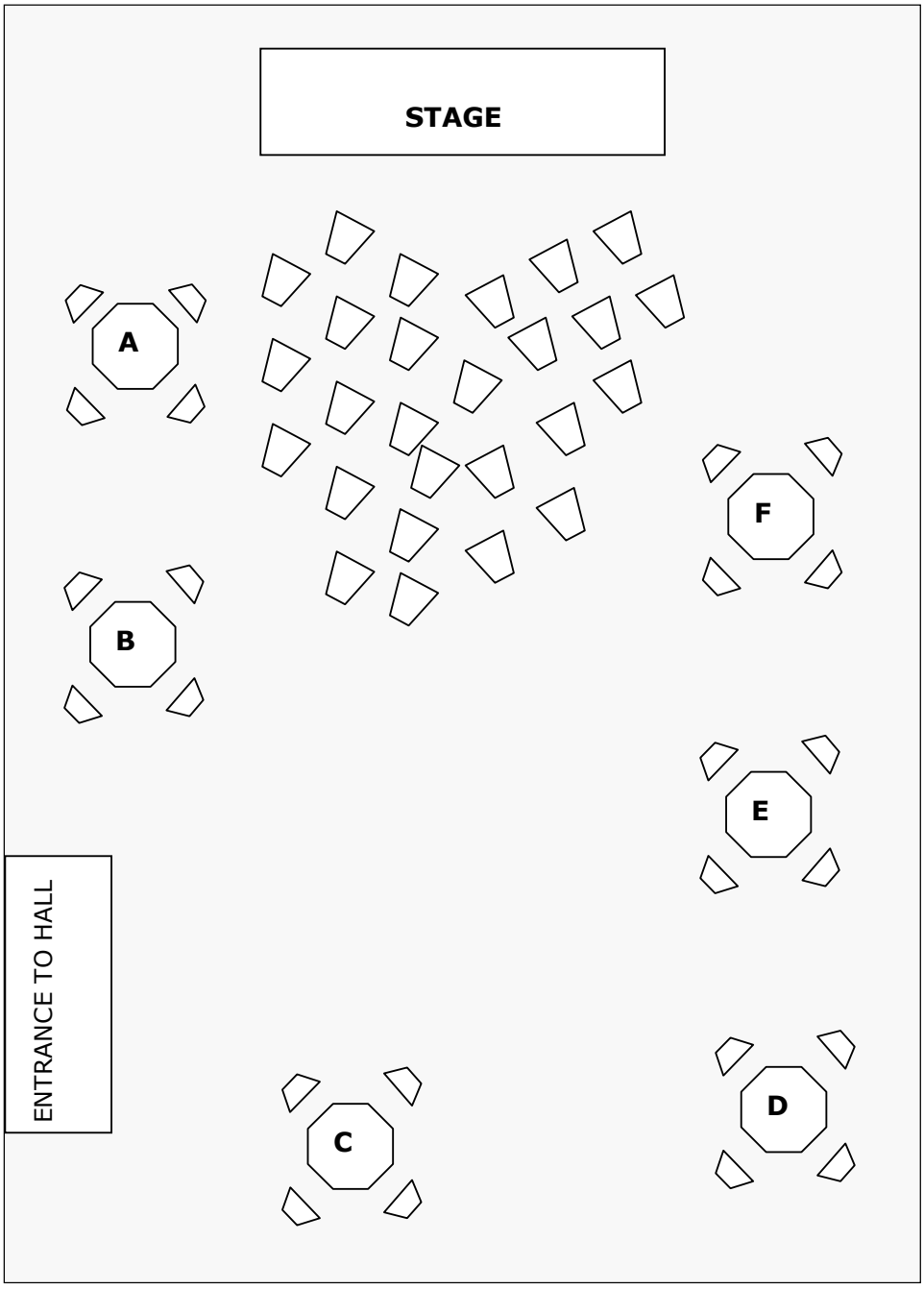
The data encryption is **WEP**

The network key is **102**

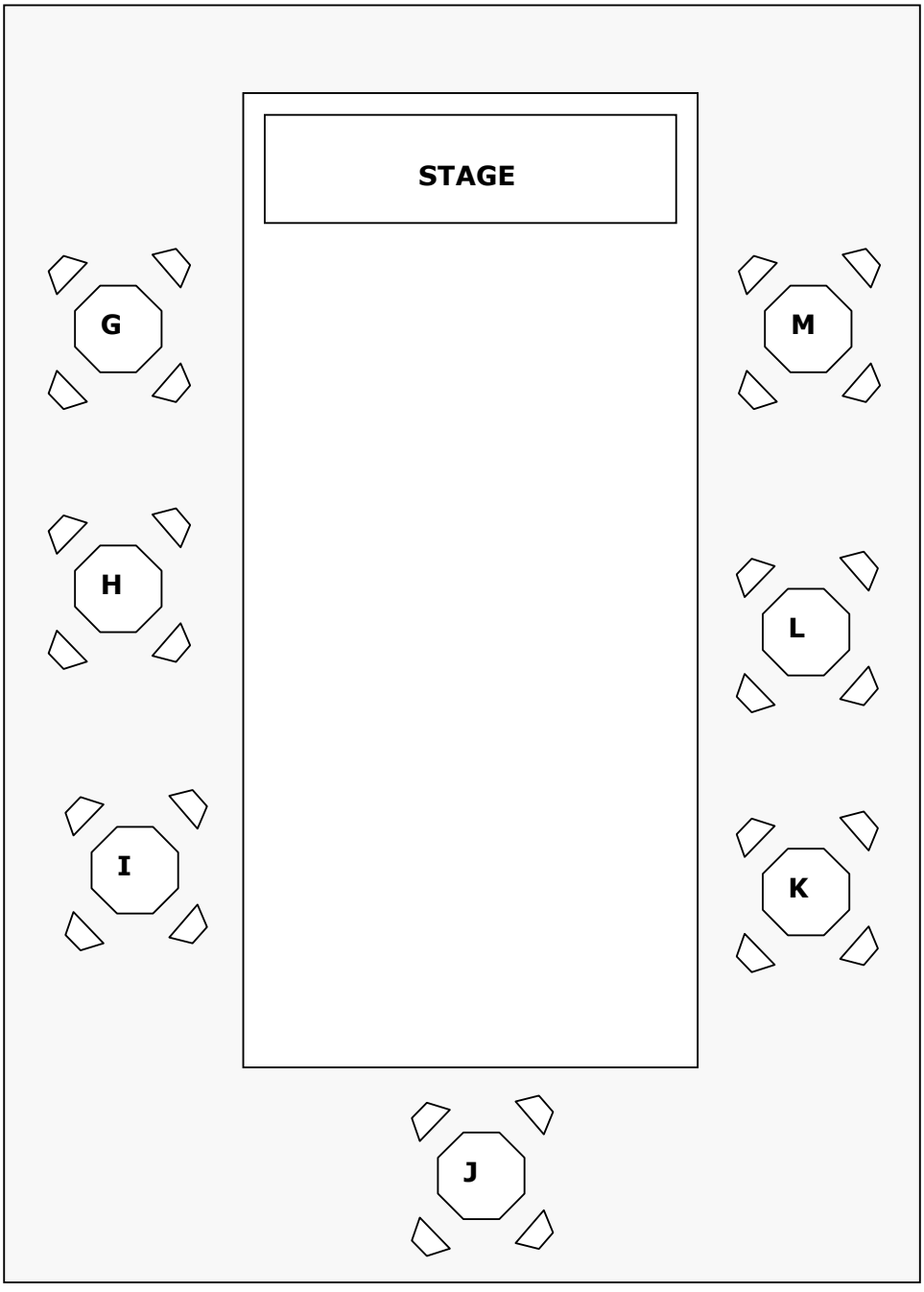

AMI
CONSORTIUM

SEATING AND TABLE LAYOUT

Map of ground floor



Map of first balcony



Club of Amsterdam

The **Club of Amsterdam** is a forum for exploring our future. To develop fresh ideas on how we want our future to look like. Not only on an individual level but certainly on a community level, city level, country level and global level. It addresses important questions about our life in the future: how can we secure sustainable wealth? Economic, social and cultural wealth. How do we balance our lives in such a way that even for 9 billion people (in 2070) can live on our planet in peace and with sufficient resources to share for all?

The Club of Amsterdam is a reflection and result of Amsterdam's achievements as a leading city for cultural, economic, scientific and social development. This translates into a creative, innovative organisation that not only strengthens the identity of the city, but also its global image.

The challenge is to think and work in that direction together. The Club of Amsterdam is connecting people who care for the future, apply creativity and are committed to discover new ways to shape our future.

In November 2002 we started with our first season of monthly events and with building our website as a knowledge resource open for everybody. We are covering issues of relevance to industry, society and government. The topic range reaches - as examples - from nanotechnology to healthcare, philosophy, global economy, robotics, reputation management to ambient intelligence.

In the meantime the Club of Amsterdam successfully organised more than 40 Season Events, which we call our monthly short conferences as well as 4 international events with global reach. We have more than 3.700 members, an online community with more than 4.500 contacts, our website is visited by more than 60.000 unique viewers per month and established Knowledge Partnerships with more than 30 institutes and with leading specialists.

Contact

Club of Amsterdam
Felix Bopp
chairman
Universal Peace Ambassador
P +31-20-615 4487
M +31-653-295700

felix@clubofamsterdam.com
http://www.clubofamsterdam.com

Directions to IGC and meeting points

TNO

TNO, the Netherlands Organisation for Applied Scientific Research, was established by law in 1932 to support companies and governments and public organisations with innovative, practicable knowledge. As a statutory organisation we have an independent position that allows us to give objective, scientifically founded judgements.

The daily work of some 5000 employees is to develop and apply knowledge. We provide contract research and specialist consultancy, as well as granting licenses for patents and specialist software. We test and certify products and services and issue an independent evaluation of quality. And we set up new companies for market innovations. The development and application of innovative knowledge: that's what we're all about.

We hold a prominent position in international science and develop new knowledge together with universities and the technology top institutes. TNO also participates in a large number of projects within the European Union's R&D programme. Thanks to this development of knowledge, TNO is always able to offer its clients the latest high-quality knowledge. Our expertise is applied in the assignments we carry out for our clients.

Developing, integrating and applying knowledge: that is the combination that differentiates us from other knowledge institutions. By encouraging the effective interplay of knowledge areas, we generate creative and practicable innovations: new products, services and processes, fully customised for business and government. Through our core areas; TNO Quality of Life, TNO Defence, Security and Safety, TNO Science and Industry, TNO Built Environment and Geosciences, TNO Information and Communication Technology. Within our five core areas we work on twelve central research themes, such as Public Safety, Work Participation and Ageing, Living with Water and High-tech Systems, Processes and Materials, that concern many issues that affect society.

Program for September 13 Field Trip

08:30 - departure by bus from Amsterdam

09:30 - arrival at TNO Soesterberg for site visit and demonstrations (Operation Cockpit)

Short presentation

Tour of facilities

Regroup/concluding remarks

12:00 - departure from TNO Soesterberg

13:00 - arrival at Amsterdam

Workshop feedback form

Thank you for participating in the AMI COI Workshop. Your comments are extremely important to help us improve our contact with you.

1. On a scale of 1 to 10, how valuable was this workshop for you? .....

2. What did you like MOST?.....

.....

.....

3. What did you specifically NOT like?

.....

.....

.....

4. Please rate the workshop elements (circle our choice)

	Negative Positive									
Pre-workshop experience (web site, registration, payment)	1	2	3	4	5	6	7	8	9	10
Printed materials	1	2	3	4	5	6	7	8	9	10
Presentations	1	2	3	4	5	6	7	8	9	10
Social/networking opportunities	1	2	3	4	5	6	7	8	9	10
Meals	1	2	3	4	5	6	7	8	9	10
Host facility	1	2	3	4	5	6	7	8	9	10
Other	1	2	3	4	5	6	7	8	9	10

5. Please rate the general breakout/demonstration format (circle our choice)

	1	2	3	4	5	6	7	8	9	10

Provide your recommendations for improvements to specific demonstrations

.....

.....

.....

6. Please share any additional comments or suggestions for improvement of AMI technology transfer or this workshop:

.....

.....

.....

.....

.....

7. For COI members: how would you like for AMI technology transfer to keep you up to date on AMI Consortium developments, events and programs

	undesirable									most desired
Individualized briefings with AMI	1	2	3	4	5	6	7	8	9	10
If desired, at what frequency?	Monthly				Quarterly				Annual	
Webcasts (presentations) by AMI Scientists	1	2	3	4	5	6	7	8	9	10
Web site with special information for COI	1	2	3	4	5	6	7	8	9	10
Newsletter (quarterly)	1	2	3	4	5	6	7	8	9	10
Format of Newsletter	Printed and Mailed						Electronic (web, PDF)			
Other:	1	2	3	4	5	6	7	8	9	10

8. For AMI Scientists: how would you like to receive periodic updates about the COI members (new products, events/local resources)

Individualized briefings with COI members	1	2	3	4	5	6	7	8	9	10
If desired, at what frequency?	Monthly				Quarterly				Annual	
Webcasts (presentations) by COI members	1	2	3	4	5	6	7	8	9	10
AMI Wiki page	1	2	3	4	5	6	7	8	9	10
e-mail from AMI Technology Transfer	1	2	3	4	5	6	7	8	9	10

Name (optional): .....

Rules of Engagement

The AMI COI Workshop is a special event designed for AMI scientists to show and speak about what they have been researching and are currently working on, to hear how the COI would envision using this research in the future, and to receive the feedback of experts in the commercial fields nearest to the work of the AMI Consortium.

The Community of Interest representatives, the "friends of AMI" (other participants who are not at present members of the COI but are participating in the workshop) and the AMI scientists will have opportunities to establish meaningful relationships.

The best relationships are built on open dialog conducted in an environment of trust.

This document sets out the rules of engagement which are necessary for all to honor during the workshop in order for trust-based dialog to be fostered and relationships to develop.

1. No non-AMI soliciting permitted. During the workshop plenary and breakout sessions, there will be no demonstrations, marketing or sales pitches of third party (non-AMI) products or services. Informal networking sessions are ideal for these conversations.

2. Confidentiality of AMI information. The AMI Consortium's research and activities are funded by the European Commission, a public source of support. Under these circumstances, all research is performed in an open and public spirit, which cannot be covered by confidentiality disclosures. This said, any AMI Consortium partner may also be conducting research or development activities which are the property of the partner and could be considered sensitive.

3. Confidentiality of non-AMI participant remarks and information. Non-AMI workshop participants (COI members, friends of AMI participants) are encouraged to disclose to AMI participants all frank opinions about/objections to the commercial applications for and feedback on AMI technology demonstrations. The comments or pre-release product information shared by the COI member (or the "friends of AMI") participant with the AMI Consortium scientist and AMI Technology Transfer will be treated as confidential and sensitive, and will be used by the AMI Consortium for internal and non-commercial purposes, unless permission to do otherwise is requested and granted in writing.

4. Post-event information. The AMI Consortium may publicize overall, general results of the COI Workshop, through press release and post-workshop briefings of selected industry analysts and press. This will include general workshop information, names of the companies who sent delegates (not the names of individual delegates). Likewise, the COI participants may publicize their participation in the AMI COI Workshop


AMI
CONSORTIUM

Opening Plenary Presentation

Welcome to
AMI Community of Interest
Workshop

September 11 and 12, 2007

Agenda

- AMI Consortium and Community of Interest
- AMI Technology Transfer
- The workshop
  - What are we going to DO here?
    - Objectives
    - Approach
  - What will we try to AVOID?
- Your binders and other "tools"

Introductions

- Project co-directors

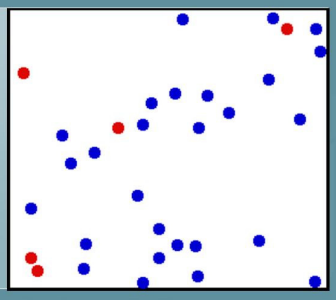
Dr.Steve Renals
Centre for Speech Technology Research
University of Edinburgh

Dr. Hervé Bourlard
IDIAP Research Institute and EPFL

- Workshop staff and supporters
- AMI Scientists
- COI members
- Friends of AMI

Find your place

- Think like a gas molecule
- And mix!



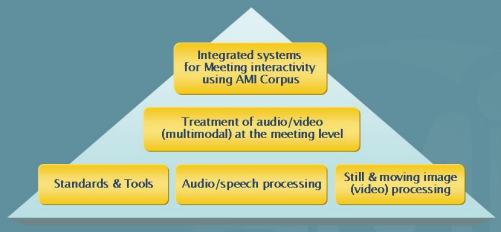
AMI Consortium

- an 11-member multi-disciplinary consortium dedicated to the development of technologies that will enhance multiparty interactions
- Partners in the AMI Consortium conduct research on human-to-human communications, particularly during business meetings between co-located and remote participants

Areas of focus

- Real-time team meeting dynamics
- Automatic meeting content indexing and viewing
- Data collaboration and/or consensus building
- Content management (publishing, indexing and repurposing of pre-recorded meetings)
- Knowledge management (mining/extracting information about and from meetings)
- Consulting about improvements in meetings
- Other

AMI Building Blocks



Core technologies (1)

- Automatic Speech Recognition (LCVSR)
- Keyword Spotting
- Localization and Tracking
  - Speaker Tracking
- Speaker Segmentation
- Gesture and Action Recognition

Core technologies (2)

- Hot spot detection
- Focus of Attention
- Dialogue Act Recognition
- Topic Segmentation
- Summarization
- Participant Influence Levels
- Meeting Browsers

AMI and AMIDA

Area of focus	AMI Project	AMIDA Project
Real-time meetings	Human-human, face-to-face	Same but in remote meetings
Automatic meeting capture	Automatic indexing/tags	Same with remote meetings
Data collaboration	Low emphasis	Higher
Content mgt	High emphasis	Equal
Knowledge mgt	Low emphasis	Higher
Consulting	High emphasis	Equal

What is the COI?

- group of people with
  - shared need to solve common problems, develop skills and share best practices
- contains multiple areas of focus
  - "communities of practice"
- In the case of AMIDA, the COI is
  - people representing their corporate employers or clients
  - at the interface between basic and applied research, and development
  - seek new ways of using technology in meetings
  - seek to overcome the technical and societal obstacles of multimodal communication at a distance

Community of Interest

AMI Technology Transfer

Technology Transfer

- A suite of programs and deliverables
  - increase the project's (project management, project participating members') relevance to target audiences
  - increase the target audiences' (industrial partners and external to project) awareness of and need for the project's results

September 11, 2007 14

Measurable Impacts

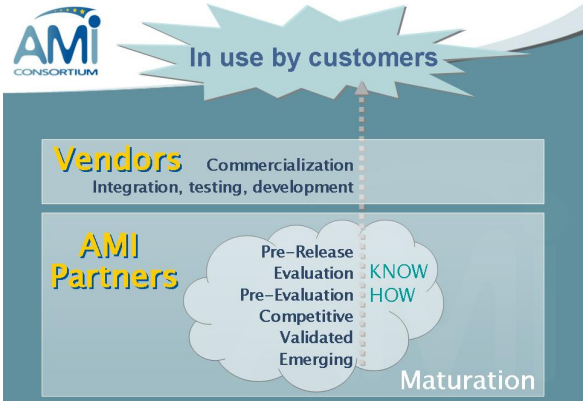
- Contact with target audience
  - Repeat contact
  - Engagement
- Create jobs
  - Students, training program participants
  - Preferably in Europe

September 11, 2007 15

Project TT objectives

- To drive changes in the way people work during and between meetings
  - To contribute to the transformation of meeting content into knowledge
- To encourage and to enable the development of tools for business meetings
- To initiate or contribute to the development of international or industry standards for how meeting archives are indexed and used

September 11, 2007 16



September 11, 2007 17

Stages (1)

- Emerging
  - Clearly identified research activities underway which could become technology of value
- Validated
  - repeatable results using the MMM database or a live media source as input
  - produces outputs which are identifiable as quantitative results under limited research conditions

September 11, 2007 18

Stages (2)

- Competitive
  - alternatives have been studied and, in comparison to other technologies, AMI technologies are believed to have potential value to industry due to a clear differentiator
- Pre-evaluation (legal)
  - IPR ownership established, terms and conditions
  - Assessment of relative contribution to total solution

September 11, 2007 19

Stages (3)

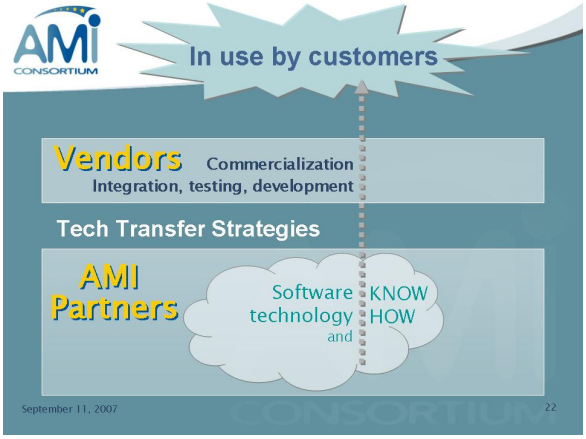
- Evaluation
  - packaged in a form (computer language, written to the operating system of the target vendor, *perhaps in a library*) which can be demonstrated in conditions OTHER than those of the AMI controlled conditions
  - Accepts inputs other than AMI MMM database
- Pre-release
  - supporting documentation (specifications in English)
  - clear IPRs identified
  - technology stable under specified conditions

September 11, 2007 20

The routes to maturity

- Crossing chasm research → development
  - Strong incentives
  - Financial resources
  - Knowledge (experience)
- Investments may be from
  - National or regional funding bodies
  - Corporate R&D

September 11, 2007 21



September 11, 2007 22

Strategies

- Collaboration, adaptation
- Commercial licensing
  - Existing company
  - Formation/transfer to a start up
- Open Source license
- Consulting services ("know how")
- Development of standards

September 11, 2007 23

The COI Workshop

at de Rode Hoed in Amsterdam

Workshop objectives

- establish fruitful, collaborative working relationships between COI members and AMI scientists
- chart the course of future tools, products and meeting processes using AMI Consortium technologies, and
- develop proposals for new joint projects involving COI and AMI Consortium

September 11, 2007 25

September 11

50/50

09:30	Overview
10:30	Coffee break
11:00	3 breakout sessions
12:30	Buffet luncheon
13:30	3 breakout sessions
15:00	Coffee break
15:15	3 breakout sessions
17:45	Closing remarks
18:00	Cocktail reception
20:00	Dinner

September 11, 2007 26

September 12

70/30

09:00	Mini-projects presentation
09:30	3 breakout sessions
11:00	Coffee break
11:30	AMI and COI (separately)
12:30	Buffet luncheon
13:00	Plenary discussion
14:00	"free radicals" (coffee)
15:30	Plenary discussion
16:30	Close of workshop

September 11, 2007 27

Core technologies (1)

Automatic Speech Recognition (LCVSR)	
Keyword Spotting	
Localization and Tracking Speaker Tracking	
Speaker Segmentation	
Gesture and Action Recognition	

September 11, 2007 28

Core technologies (2)

Hot Spot Detection	
Focus of Attention	
Dialog Act Recognition	
Topic Segmentation	
Summarization	
Participant Influence Levels	
Meeting Browsers	

September 11, 2007 29

Breakout sessions

25 min. 3-4 people per table

:00	greetings
:05	demonstration of AMI technology
:10	discussion of applications, issues, concerns Have you seen this or similar technology before? Have you thought of other approaches to this challenge? Do you have applications for this technology today? What value do you think this would bring to your end users?
:20	private time to complete papers (write notes of issues raised, possible next steps, ratings) in individual binders
:25	end

September 11, 2007 30

Breakout session outputs

- Sharing/exchange in real time
- In the place of full recording…
  - handwritten notes in your binders
  - score the likelihood of collaboration
  - sticky memos
    - What you find most valuable
    - Major concerns

Rules of engagement

- We will take good care of you
  - Provide information, binders, signs, seating plan, schedule of meetings
  - Entertain and feed you!

- Be creative, innovative, dare to ask
- Be good listeners → capture feedback
- COI members refrain from sales pitches during the breakouts

Rules of engagement (2)

- AMI Consortium demos and information provided are public
  - AMI partners may also have sensitive information
- COI member remarks *not* under strict non-disclosure
  - unless signature with individual AMI partner
- AMI and COI may publicize high level results

Field Trip to TNO

08:30	Bus departs Hotel Victoria
09:30	Presentation
10:00	Tour of facilities
11:30	Presentation/discussion
12:00	Bus departs Soesterberg
13:00	Arrive Hotel Victoria

Your Binders

Personal guide to your workshop

Questions?

Demonstrations

During the breakout sessions, many small groups will be meeting. The initial focus of your 30 minute meeting is the demonstration prepared by AMI Scientists.

On your individual schedule, COI members and friends of AMI will find the letters assigned by breakout period. The letters correspond to the following table assignments.

You will use the table layout (pages 7-8) to find the meeting location.

A	Meeting Evaluation System
B	Automatic Segmentation of Meetings
C	Visual Focus of Attention Recognizer
D	Extractive Summarization
E	Smart Access to Presentation Content
F	Automatic Video Editing
G	Keyword Spotting
H	JFerret Meeting Browsers
I	Virtual Presence Support
J	Temporal Compression of Meetings
K	Comic-style multi-modal Summarization
L	Meeting Metadata Standardization
M	Face Analysis

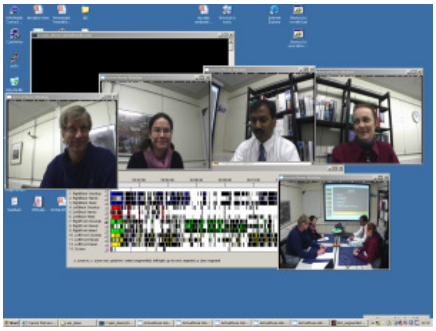
The AMI Scientists have prepared descriptions of their demonstrations including some background and suggested application scenarios. The signs at each table are reproduced in this binder for participant use during the meetings.

Following the demonstration, please discuss any aspect of the technology or potential application scenarios.

Automatic Segmentation of Meetings

Meetings have natural junctures between topics as well as different phases. Software can detect logical segments during a meeting.

Multiple levels of segmentation and indexing



The AMI Consortium is developing technologies to support multilevel segmentation and annotation.

Currently, we focus on the following technologies:

- Topic segmentation: Recognizing topical structure in meeting transcripts
- Sentiment Analysis: Finding emotive speech and hot spots
- Speaker segmentation based on the characteristics of the speaker's voice using non-intrusive far-field microphones. (Who's speaking when?)

Application Scenarios

Scenario 1: building a browsable meeting recording archive

Full-scale, multimodal indexing of meetings constructs a browsable meeting archive that supports many applications. The multimodal indexing techniques in use by the AMI Consortium may include:

- image analysis (e.g. motion zones),
- speech analysis (e.g. diarization, laughter detection)
- transcript analysis (e.g. topic segmentation, summarization, topic classification, sentiment analysis).

Scenario 2: searching/filtering/finding interesting segments

As soon as meeting segments have been detected and annotated, they can be exploited for any search or summary generation application. For example, all positive comments on the concept or idea during the meeting can be retrieved.

Contact

Wessel Kraaij
TNO ICT
wessel.kraaij@tno.nl

JFerret Meeting Browsers

What is JFerret?

Constructing applications to effectively and intuitively browse the contents of a meeting is very important. JFerret is an extremely flexible meeting browser development framework, which presents the multi-media meeting recordings of the AMI project. It can be configured to display many types of data and recognition results, including speaker segmentations, speech transcripts, slides, meeting actions and other annotations, created either manually or automatically. It is used in various guises in the AMI project.

Key Functionalities

- Integrates various AMI technologies
- Intuitive graphical user interface
- Very flexible, simple plug-in integration
- Easily integrated within any Java application

Application Scenarios

Scenario 1: Browsing

A single browsing interface built using JFerret can include any or all of the following components: graphic representation of speaker segmentation (color bars that show who spoke when), textual meeting transcript (from automatic speech recognition or input manually), dominance levels and argumentation graph (extracted automatically), images of slides, multiple video and audio. All components appear synchronized and the text is color coded. The user can click on the timeline, a particular slide, or the transcripts to view the corresponding segments.

Scenario 2: Searching

Text transcripts of the meeting can be searched using keywords. The results are presented with associated metadata using the JFerret architecture.

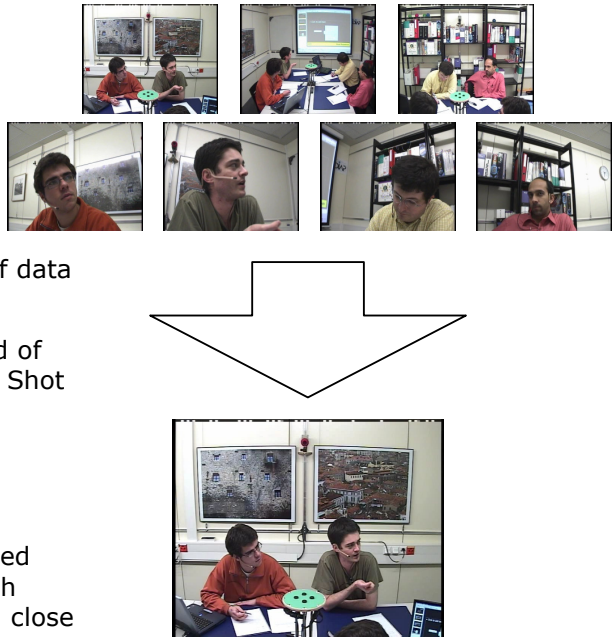
Integration

The JFerret framework can be used to easily integrate any of the above components into any application that uses Java. For example, the player and timeline could be added to another application.

Contact

Alex Jaimes
IDIAP Research Institute
alex.jaimes@idiap.ch

Automatic Video Editing



Multiple Video Sources

Current technology permits recording of live events using several cameras. This produces data for automatic detection, recognition, indexing etc. However, how can be such amount of data presented to the human viewer?

Only the most important shots are selected instead of unnecessary presentation of all the available data. Shot composition satisfies many user requirements and elementary movie maker's standards.

Automatic Video Editing

AMI technology allows editing of the events recorded simultaneously with several cameras. Cameras with different type of view can be utilized (distant view, close view). Virtual camera tool is available so e.g. persons can be tracked on camera with distant view and satisfactory resolution.

Various types of the events can be processed; editing algorithm can be adjusted according to the available event inputs. Important events are preferred in the output videos – shots are selected according to their measure of "importance".

Aesthetical aspects are taken into account during shot composition – some elementary rules from movie makers are included.

Various effects can be included in resulting video – zooming cameras, picture in picture, fade in/out etc. Both real-time and offline applications are possible.

Application Scenarios

Meeting video summary: automatic generation of video summary of meetings, which are recorded with several cameras.

Multi-camera Videoconference system: If more than one camera is used at a location, stream from the location is pre-edited.

Automatic lecture video: Output video will be composed from such shots according to lecturer activity, projected slides etc.

Contact

Stanislav Sumec
Brno University of Technology, Faculty of Information Technology
sumec@fit.vutbr.cz

Face Analysis

A human face provides information about the person: gender, age, ethnic origin, glasses, beards, identity, gaze direction, and mental state.

In human-human interaction the appearance of the face contributes to the efficiency and quality of the communication.

For applications which strive to analyze and improve human-human or human-computer communication with non-verbal cues, analyses of faces is mandatory.

Real-time face analysis with FEASy

FEASy provides:

- localization of head and eye positions
- compact and precise description of the appearance of a person's face
- various recognized properties (e.g. gender, age, facial expression)

Application Scenarios

Scenario 1: Meeting analyses

The FEASy output can be used in automatic analyses of meeting participants and for a comprehensive monitoring, e.g. of the levels of activity, attention, and dominance during any meeting.

Scenario 2: Anonymity in video communication

Video chat rooms with generated characters (avatars) are a novel approach for new online communication platforms. FEASy allows a user to have remote control of the avatar's face and head motions leading to a natural – but anonymous – way of communication. See also remote presence demo.

Scenario 3: Adaptive human computer interfaces

Future HCIs will provide ease of use and be fun to use by way of active adaptation to their user. New HCIs will be found in driver/navigation assistants, remote and local meeting assistants, gaming consoles and a wide array of consumer and business devices.

Contact

Ronald Mueller
MMER-Systems, Technische Universität München
mueller@mmer-systems.eu

Multi-Meeting Keyword Spotting

Searching on the basis of keyword spotting

People want to quickly find the points in a series of meetings where a particular keyword or combination of keywords was used by one or more meeting participants. Users do not want to be limited by a fixed vocabulary of current speech-to-text engines.

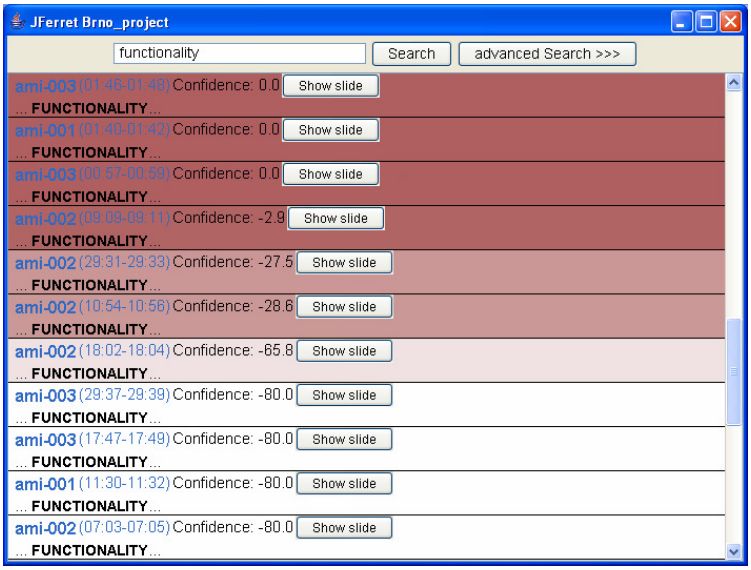
They also want to have a solution with high performance. This solution successfully participated in 2006 Spoken Term Detection Evaluations organized by U.S. NIST.

The research and technology behind the demo

Meetings are off-line recognized by AMIDA large vocabulary continuous speech recognizer (LVCSR) and phoneme recognizer. Their results in forms of graphs (lattices) are indexed and stored in a form suitable for fast access.

When the user enters a query, known words ("in vocabulary") are searched in LVCSR indexes. For out-of-vocabulary words (OOVs), phonetic search is used, so that it is possible to search even for person, company, brand names, etc.

The user interface to multi-meeting search is the popular meeting browser JFerret. All information such as video, slides, etc. is synchronized by simply clicking on the found segment.



Application Scenarios

Scenario1: looking for a keyword or combination of keywords in "Google-like" style.

Scenario 2: verifying if a person really said what you think he/she said.

Scenario 3: quickly browsing around the found segment for important information.

Contacts

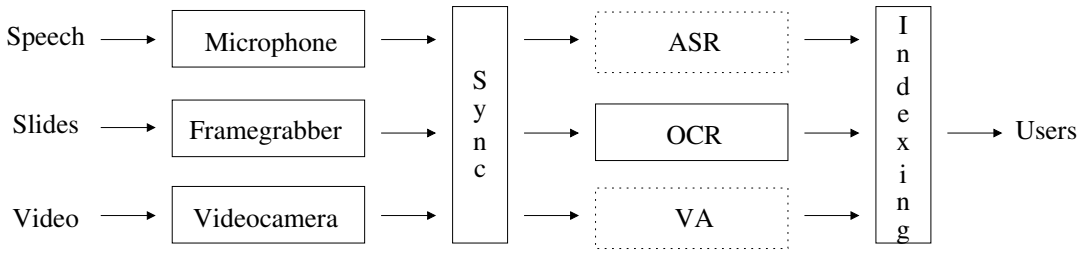
Honza Cernocky
Speech@FIT
Brno University of Technology
cernocky@fit.vutbr.cz

Smart Access to Presentation Content

The Presentation Acquisition System (PAS) performs capture and indexing of oral presentations based on slides. The system is fully transparent and the speakers do not need to change their behaviors in order to make the system work. Moreover, the PAS can be activated and de-activated using a single button and this makes it suitable for use in any real world environment.

The system acquires three channels: audio, video and PC-projector output (i.e. whatever is projected onto the screen). The three channels are synchronized and, at any instant, it is possible to know what is being said, what is projected onto the screen and how the speaker appears (see Figure for the architecture).

The sequence of the slides is extracted from the PC-projector: each slide image is transcribed with an OCR and the resulting text is indexed. The slide transcriptions can be searched like any other text and, by retrieving a slide, it is possible to retrieve audio and video segments corresponding to the moment when the slide was projected.



The acquired presentations are delivered through a Google-like system that retrieves video segments and slides relevant to the submitted query. An online demonstration is available at: http://www.idiap.ch/mmm/talk-webcast/uist-06.

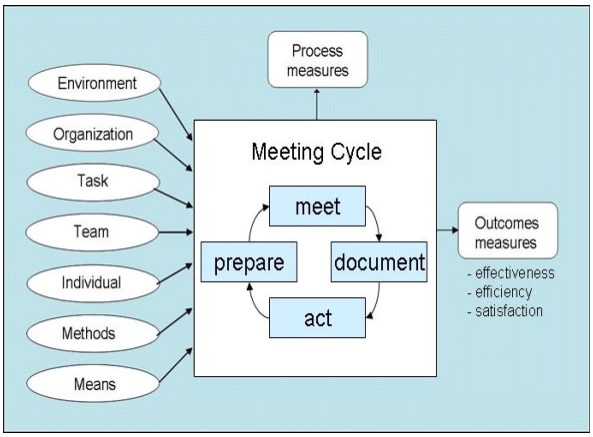
Contact

Alessandro Vinciarelli
IDIAP Research Institute
vincia@idiap.ch

Quantifying the Impact of Meeting Technologies

Organizations need to know the likely impacts of adopting new ways of working.

Within the AMI project, an extensive evaluation environment is available for determining the (potential) effect of either an individual component or a fully integrated meeting support system on a business. Assessment takes place through standardized experimental sessions in which meeting participants will work in a real but controlled task setting, while a large set of measurements are taken, both on the process and the outcome of a (cycle of) meetings. This environment is now available for evaluating your meeting product.



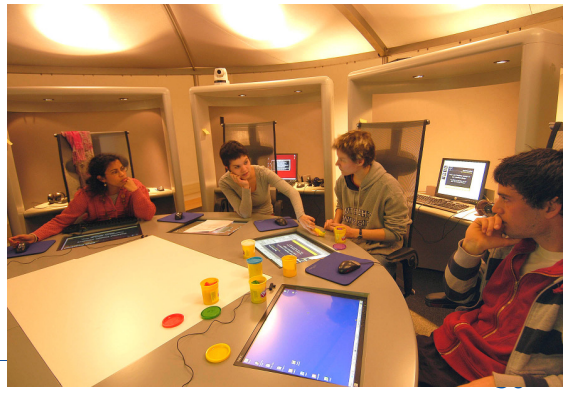
Application Scenarios

Scenario1: Directing your product development

In developing your product, many options need to be considered. Some decisions are hard to take, since their consequences are often unclear. The AMI meeting evaluation paradigm can help you out at and early stage.

Scenario 2: Convincing your client

Clients would usually like to know whether purchasing your product is worth their money. The AMI meeting evaluation paradigm can show the impact of your product in terms of effectiveness, efficiency and user satisfaction for your client's organization.



Contacts

Wilfried Post
TNO Human Factors, the Netherlands
wilfried.post@tno.nl

Extractive Summarization

Hand-written minutes are often unavailable or inaccurate/unreliable. But AMI technology can automatically generate meeting summaries.

What's important in a meeting?

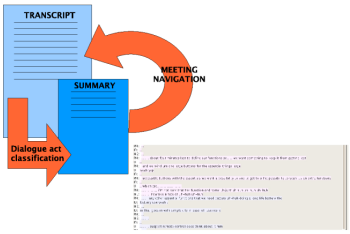
Which topics were discussed?

Which speakers were active? What was decided?

How we automatically create a summary

Using automatically generated text transcript, algorithms extract the most important meeting sentences, using:

- prosodic information, i.e. *how* the participants speak
- linguistic information, e.g. presence of *keywords*
- structural information, e.g. *location* in the meeting



Concatenate these sentences to form a single document.

Application Scenarios

Scenario1: Between Meetings

You are between meetings and would like to revisit the discussion of a previous meeting before attending the next one. Using our meeting browser, you can navigate the content of prior discussions and get up-to-speed. Software can automatically provide a condensed version of the meeting and permit non-linear browsing/topic segment-based navigation.

Scenario 2: Real Time Keyword Spotting

You are not present in the meeting but are monitoring it remotely while addressing other tasks. A running summary can be generated by a meeting assistant who pings when subjects of interest enter the discussion.

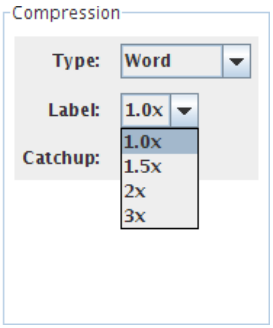
Scenario 3: Meeting acceleration (catch up with meeting in progress)

Want to know what has been discussed so far, but you do not want to interrupt the current discussion? A summary of a meeting allows the late participant to acquire context for a discussion and join the conversation without inconveniencing those already in the meeting.

Contacts

Gabriel Murray
CSTR, University of Edinburgh
gabriel.murray@ed.ac.uk

Temporal Compression of Meeting Audio



Temporal compression is a technique which allows you to spend less time listening to speech recordings but still collect all the useful content. There are a number of different techniques to do this - two of which are present in this demo. We can remove either words or sentences that are less important, or we can speed up the recording but keep the pitch of the speakers constant.

The different techniques have different trade offs at different compression rates.

Application Scenarios

Scenario 1: Mobile Meeting Reminder

This technology is fast and extremely portable and so one possible application is a rapid meeting reminder that can run over a mobile phone or pda. In this scenario you are en route to a meeting but would like to have a rapid reminder of the previous meeting - using temporal compression you can highly compress the meeting and, for example, get a five minute overview of what was said in the previous meeting.

Scenario 2: Meeting Catch Up

You are en route to a meeting (or you have arrived) but are running late. Using temporal compression you can rapidly skim the beginning of the meeting and join the meeting with some knowledge about the subjects and content of the discussion you have missed. This can be tuned to your requirements (e.g. catchup 10 minutes of missed meeting in 5 minutes or catchup 10 minutes of missed meeting in 1 minute).

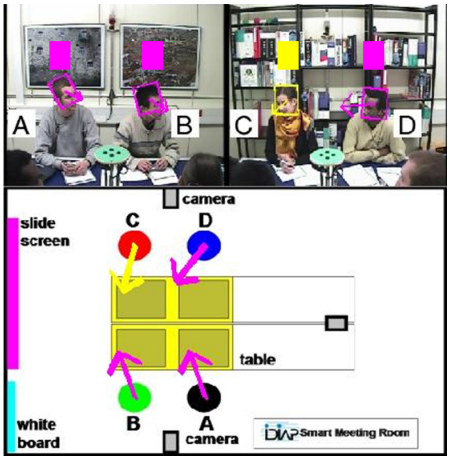
Scenario 3: Rapid Meeting Playback

The temporal compression system can be used as a general playback mechanism to allow you to process meeting recordings much faster. Meetings contain a great deal of irrelevant content (lengthy pauses, um's and err's etc.) - with temporal compression you can automatically skip these portions of the meeting and extract relevant and useful information much faster.

Contact

Simon Tucker
Dept. of Information Studies, University of Sheffield, UK
s.tucker@shef.ac.uk

Visual Focus of Attention (VFOA) Recognizer



Today, when people meet face-to-face, they provide one another visual clues regarding their attention. For example, if a person is looking at their screen, their attention is probably on the contents of the screen, not the person or people speaking. When people are meeting remotely, it is difficult to know what they are focusing on, however, it is equally important, if not more important.

For many future meeting applications, the focus of a person's attention is an important element.

What can technology do?

The AMI Consortium has developed systems that automatically estimate the Visual Focus of Attention (VFOA) of people based the orientation of their head. The **VFOA** of a person at a given time instant, defined from his eye gaze, **indicates who or what a person is looking at**.

What is it?

VFOA is defined as a set of discrete visual targets of interest. For example, in the meeting context, these are: the other participants, the slide screen, the white-board and the table.

Application Scenarios

Meeting Assistants: digital assistants that analyze the social dynamics of non-verbal communication in a group and provide feedback on the group dynamics. This can be for the remote participant to better understand what is happening in the environment where other meeting participants are co-located, thus increasing participant's satisfaction level. An assistant could also judge how "safe" it is to interrupt a person based on what they are focusing on.

Participant Influence Level estimation: In some meetings, it is important to know who has the greatest influence over a decision.

Market Research: the recognizer can be used to assess whether people in a given environment are attracted to visual stimuli. A tool can automatically measure the effectiveness of outdoor advertisements, similarly to the Nielsen ratings that measure media effectiveness.

Contacts

Daniel Gatica-Perez
IDIAP Research Institute
gatica@idiap.ch

Jean-Marc Odobez
IDIAP Research Institute
odobez@idiap.ch

Supporting Virtual Presence

Why is virtual presence important?

In meetings with one or more remote participants the remote persons are in a less favorable situation then the people who meet face-to-face. The remote participants have difficulties with:

- Taking the floor
- Using and interpreting non-verbal signals
- Determining the focus of attention

This may make the remote participants less influential and will disturb the flow of interaction. Virtual presence can reduce these obstacles and improve the interaction flow.

How to support virtual presence

There are several ways to support remote participation. A participant can be present as a virtual head on a screen. Or be represented by a physical robot, such as the iCat. Such devices need to be equipped with intelligent listening behavior, floor taking behavior, all of which support remote presence.

Application Scenario

A person, traveling around, needs to join a meeting, and has an internet connected mobile phone. This phone is running a special meeting application. Our traveler is represented in the common meeting room by means of a "talking head" or avatar. Apart from the normal speech channel, the following options are available for the remote participant:

- "Request the floor". This has the effect that the talking head show non-verbal behavior appropriate for requesting attention.
- "Emotion display". This will cause the taking head to show surprise, joy, agreement or disagreement etcetera.

Contact

Anton Nijholt
Human Media Interaction Group
University of Twente

anijholt@ewi.utwente.nl

Comic-Style Multi-modal Summarization

Meeting Metadata Standardization

The word 'annotation' is used in many ways, along with the word 'metadata'. They both refer to supplementary information added to an initial collection of data (e.g. multimodal meeting recordings) in order to enrich its informational content. This information might represent supplementary knowledge about the data (e.g. entered by users upon recording of the data) or a more abstract or "semantic" representation of the information contained in the data (e.g. computed from the data).

Annotations are crucial for research and development based on machine learning algorithms, as reference annotated data is useful to train recognizers and to evaluate them. Unsupervised learning methods can use non-annotated data.

The AMI Consortium designs tools that assign metadata and annotations automatically to the multimedia recordings that they process, although some metadata elements (such as name and description of participants) could be added manually by a meeting administrator, in relation to a company's own information systems.

Why are standards needed?

Standards are necessary because as soon as annotated meeting archives become available, users will seek to use them on heterogeneous platforms and for different applications (meeting assistants, storage systems or meeting browsers).

Standards are crucial to ensure interoperability between resources used for training and test, and between commercial systems. "Plug-and-play" resources can be developed and shared by the community, thus reducing the annotation effort, and establishing common grounds for benchmarking.

What is the problem?

There are many standards proposed. None are perfectly matched to the needs of this community and the meeting archives of the future.

Contact

Hervé Bourlard and Andrei Popescu-Belis
IDIAP Research Institute
herve.bourlard@idiap.ch

Overview of AMI Partners

The AMI Consortium is comprised of research institutes, universities and an industrial partner (Philips). The consortium partners collaborate on research projects and integrate their results into demonstrations and deliverables.

Partners contribute to work packages in their special areas of expertise as noted in the following pages.

Research Institute

IDIAP Research Institute, CH

AMIDA Coordinator

Prof. Hervé Bourlard

Areas of expertise

Speech processing, computer vision, machine learning, multimodal interaction, smart meeting room, handwriting recognition, multimodal indexing.

Contacts

IDIAP
Prof Hervé Bourlard
Centre du Parc
Av. des Prés-Beudin 20
Case Postale 592
CH-1920 Martigny
Switzerland

herve.bourlard@idiap.ch

http://www.idiap.ch

German Research Centre for Artificial Intelligence (DFKI), DE

Prof. Wolfgang Wahlster and Dr. Tilman Becker

Areas of expertise

Intelligent visualisation and simulation systems, language technology, intelligent user interfaces, summarization.

Contacts

German Research Center for
Artificial Intelligence (DFKI)
Dr. Tilman Becker
Stuhlsatzenhausweg 3
Room C0.12
D-66123 Saarbrücken, Germany

Tilman.Becker@dfki.de

http://www.dfki.de/~becker

www.dfki.uni-kl.de

Research Institute

International Computer Science Institute (ICSI), USA

Prof. Nelson Morgan and Dr. Barbara Peskin

Areas of expertise

Language and dialogue modelling, spoken language processing, speaker modelling, instrumented meeting room design.

Contacts

International Computer Science
Institute (ICSI)
Dr. Adam Janin
1947 Center Street
Suite 600
Berkeley, CA 94704 USA

adam@janin.org

www.icsi.berkeley.edu

Research Institute

Netherlands Organisation for Applied Scientific Research (TNO), NL

Dr. Wessel Kraaij (Delft), Dr. Wilfried Post (Soesterberg)

Areas of expertise

Multimodal summarisation, information retrieval, natural language processing, computer vision, human factors, HCI, audio analysis, user-centered design, usability testing, group decision making, computer supported collaborative work.

Contacts

The Netherlands Organization for
Applied Scientific Research (TNO)
Dr. Wessel Kraaij
Brassersplein 2
Delft NL

wessel.kraaij@tno.nl

Dr. Wilfried Post
Kampweg 5
3769 DE Soesterberg NL

wilfried.post@tno.nl

www.tno.nl

e-Health Research Centre, Australia

Dr Iain McCowan

Areas of expertise

Content analysis of medical multimedia, technological support for clinical teams, sensor signal processing for patient monitoring.

Contacts

www.e-hrc.net/

University of Edinburgh (UEDIN), UK

AMIDA Coordinator

Institute for Communicating and Collaborative Systems

Prof. Johanna Moore, Prof. Steve Renals and Dr. Jean Carletta

Areas of expertise

Dialogue understanding, cognitive engineering, multimodal annotation.

Contacts

University of Edinburgh
Dr. Steve Renals
CSTR
2 Buccleuch Place
Edinburgh EH8 9LW, UK

s.renals@ed.ac.uk

www.iccs.informatics.ed.ac.uk

Sheffield University (USFD), UK

AMIDA Training Coordinator

Computer Science Department, Speech and Hearing Research Group

Prof. Phil Green

Department of Information Studies

Prof. Steve Whittaker

Areas of expertise

Intelligent visualisation and simulation systems, language technology, intelligent user interfaces, summarization.

Contacts

www.dcs.shef.ac.uk/spandh

Brno University of Technology (BUT), Czech Republic

Institute of Computer Graphics and Multimedia

Prof. Honza Cernocky

Areas of expertise

Speech coding, speech recognition, speaker recognition, machine vision.

Contacts

Brno University of Technology
Pavel Zemcik
Bozetechova 2
612 66 Brno, Czech Republic

zemcik@fit.vutbr.cz

Honza Cernocky
Bozetechova 2
612 66 Brno, Czech Republic

cernocky@fit.vutbr.cz

www.fit.vutbr.cz/research/groups/speech

Technische Universität München (TUM), DE

Institute of Human-Machine Communication

Prof. Gerhard Rigoll

Areas of expertise

Multimodal man-machine communication, language engineering, computer vision, gesture recognition.

Contacts

Technische Universität München
Benedikt Hörnler
Institute for Human-Machine Communication
Arcisstraße 21
80333 Munich DE

b@tum.de

www.mmk.ei.tum.de

Academic Partners

University of Twente (UT), NL

Parlevink language engineering group

Computer Science Department, Speech and Hearing Research Group

Prof. Franciska de Jong and Prof. Anton Nijholt

Areas of expertise

Multimodal interaction, multimedia retrieval, virtual reality, agent technology.

Contacts

University of Twente
Dr. Anton Nijholt
Faculty EEMCS, PO Box 217,
7500 AE Enschede NL

anijholt@cs.utwente.nl

http://parlevink.cs.utwente.nl

Industrial Partner

Philips Consumer Electronics BV, NL

Mr. Kees Tuinenbreijer, Bram van der Wal

Areas of expertise

Interactive services and applications on the move.

Overview of AMI Community of Interest

A community of interest is a group of people connected to each other by a shared need to solve common problems, develop skills and share best practices.

In the case of AMIDA, the community of interest is composed of people representing their corporate employers or clients who are frequently at the interface between basic and applied research, and development.

The AMI Consortium works with the AMI Community of Interest to advance consortium technology transfer objectives.

A community of interest may contain subsets of people sharing information within their respective communities of practice.There are at least four communities of practice in the AMI COI: vendors, consultants, research labs, large end user organizations.

We expect the members themselves to seek to experience new ways of using technology in meetings, and to overcome the technical and societal obstacles of multimodal communication at a distance by sharing knowledge and insights in a structured and constructive fashion.

The COI is not a scientific review committee, not a project advisory board and not a decision making body.

Ceannard

COI representative: **Jim Hood**

Company overview

Ceannard provides a software product and professional services by which large enterprises can increase the effectiveness of knowledge access.

It is a small enterprise, based near Glasgow, Scotland

Products

Ceannard developed and current licenses software and systems that help companies "capitalize on knowledge" by capturing meetings and presentations, automatically indexing and archiving them for knowledge mining.

All media (video, audio, slides and transcript) are synchronized in the playback window.

Consulting services

Ceannard consultants focus on the effectiveness of knowledge access as well as knowledge capture (in meetings and seminars), in order to boost individual, team and business performance.

Contacts

Ceannard
Jim Hood
Ulverscroft
Duchal Road
Kilmacolm
PA13 4AY Scotland

jim.hood@ceannard.com

www.ceannard.com

Cisco

COI representatives: **Keith Lantz and Christian Renaud**

Company overview

Cisco is a large enterprise with over 47,000 employees worldwide. It has offices in Europe, with research and development in Rolle, Switzerland.

It is the world's largest provider of IP infrastructure. Cisco funds organic development of new products and services - over $4.07 Billion spent on R&D in Fiscal Year 2006. It is also an active acquirer of/investor in innovative start-ups. Cisco has acquired 114 companies since 1993. In the area of meetings, the company has purchased two leading companies in web conferencing:

- Latitude Nov 2003 $80M
- WebEx March 2007 $3.2B

Products in the domain of Collaboration and conferencing

In the domain of collaboration and conferencing, Cisco also conducts significant research and development internally, and has OEM relationships with suppliers of technology as well. Cisco is the largest OEM customer/partner of RADVISION from which it obtains video gateways and MCUs and resells as part of its enterprise solutions business. The company also resells TANDBERG videoconferencing end points.

In addition, Cisco internally develops IP video and telepresence solutions. The company introduced its TelePresence family of solutions in Oct 2006.

Services

WebEx is the largest provider of virtual meeting services (see WebEx profile below). The company also provides teleconference bridge and MeetingPlace Express hosting services.

Contacts

Cisco Systems Inc.
Dr. Keith Lantz
170 West Tasman Drive
San Jose, CA 95134 USA

klantz@cisco.com

Dr. Christian Renaud
170 West Tasman Drive
San Jose, CA 95134 USA

crenaud@cisco.com

www.cisco.com/telepresence

Codian

COI representative: **Killko Caballero** (not attending the COI Workshop)

Company overview

Founded in 2002, Codian is an SME based in Slough, UK.

Products

Codian provides hardware-accelerated standard definition and high definition videoconferencing products which are deployed in IP infrastructure for service providers and enterprises.

Products include multipoint conferencing servers for SD, HD and mixed meetings, gateways and video meeting recorder/streaming servers.

In addition, the company offers scheduling and network control/monitoring software.

Contacts

Codian
Killko Caballero
1570 The Alameda #216
San Jose, CA 95126

killko@codian.com

www.codian.com

Hewlett Packard

COI representative: **Erik Geelhoed**

Company overview

HP is a technology solutions provider to consumers, businesses and institutions globally. The company's offerings span IT infrastructure, global services, business and home computing, and imaging and printing.

For the four fiscal quarters ended October 31, 2006, HP revenue totaled $91.7 billion making it the largest Information Technology company.

Products

HP has products in:

- Communications, media, entertainment
- Business solutions
- Networks
    - Servers: hardware (Digital, Compaq) and software
- Computers
- Mobile devices
- Imaging and Printing (includes cameras)

In the area of meetings, HP offers Halo Collaboration Studio (announced in December 2005) as well as a growing portfolio of telepresence products and services.

Contacts

Hewlett Packard Labs
Erik Geelhoed
Building 3, Mail Stop 67,
Filton Road
Stoke Gifford, Bristol BS34 8QZ.
U.K.

erik_geelhoed@hp.com

www.hp.com

IBM

COI representative: **Vova Soroka** (not attending the COI Workshop)

Company overview

IBM is a multinational computer technology and consulting corporation headquartered in Armonk, New York, USA. IBM manufactures and sells computer hardware and software, and offers infrastructure services, hosting services, and consulting services in areas ranging from mainframe computers to nanotechnology

Over 600 individuals work at the IBM Haifa Labs; 25 percent of the technical staff have doctorate degrees in computer science, electrical engineering, mathematics, or related fields. Employees are actively involved in teaching at Israeli higher education institutions and supervising post-graduate theses. Many employees have received IBM awards for achievements and excellence.

In the area of meetings, Collaboration Technologies group of IBM Haifa Research Labs focuses on adapting well known collaboration tools and paradigms to the workplace. The group is developing collaboration middleware and tools that allow people to use collaboration more efficiently for their business needs. This group has strong skills both in development of collaborative tools and frameworks and in the social side of technology - User Interfaces, diffusion of technologies and usability.

Contacts

IBM Haifa Research Lab
Vova Soroka
Sr. Manager, Collaboration Technologies
Information and Interaction Technologies Department

vladi@il.ibm.com

/www.haifa.il.ibm.com/

Intel

COI representative: **Cindy Pickering** (not attending the COI Workshop)

Company overview

The world's largest semi-conductor company, based on revenues, Intel has over 94,000 employees. The largest division, generating 56% of revenues in 2006, is the Digital Enterprise Group.

Intel has a significant IT group as well, which is responsible for running the infrastructure and services necessary for the operation of the global company. In the IT group there is research conducted on meetings and meeting technologies.

Products and services

Intel's products include chips, boards, and other semiconductor products that are the building blocks integral to computers, servers, handheld devices, and networking and communications products. Its component-level products consist of integrated circuits used to process information, including microprocessors, chipsets, and flash memory.

In addition, Intel is also engaged in other industries and develops products and services for internal as well as partner uses.

In the area of meetings, Intel has in the past sold various products, including Intel ProShare for desktop and small group videoconferencing. Currently the company has a number of products and technologies under development, such as 3D spaces for sharing information during meetings.

Contacts

Intel
Cindy Pickering
5000 W Chandler Blvd
C2-131
Chandler, AZ 85226 USA

cynthia.k.pickering@intel.com

LifeSize

COI representative: **Casey King** (not attending COI Workshop)

Company overview

Founded in 2003 by videoconferencing industry veterans, LifeSize is an Austin, Texas-based SME. The company was the first commercial provider of enterprise HD videoconferencing end points.

Products

The company provides audio and video conferencing terminals with high definition video for personal, small and large group meetings

In addition, it provides infrastructure hardware and software for multipoint meetings and managing terminals and servers.

Services

WebEx is the largest provider of virtual meeting services (see WebEx profile below). The company also provides Teleconference Bridge and MeetingPlace Express hosting services.

Contacts

LifeSize Communications
Casey King
901 S. Mopac Building 3
Suite 300
Austin, Texas 78746

cking@lifesize.com

www.lifesize.com

Logitech

COI representative: **Bernard Gander** (not attending COI Workshop)

Company overview

Logitech is a large enterprise (approximately 7,400 employees worldwide) with headquarters in US and Switzerland. With fiscal 2007 sales exceeding $2.1B, it is a world leader in personal peripherals, including PC navigation, Internet communications, digital music, home-entertainment control, gaming and wireless devices. Logitech's retail (to consumer) business accounts for 89 percent of its revenue. The company's original presence and growth in retail mice, webcams and speakers was driven by a trend among consumers to enhance their basic desktop PC systems with more fully featured personal peripherals that add functionality and cordless freedom to their desktop.

Products

In the domain of meetings, Logitech is the world's leader in webcam sales. It also develops and distributes third party software for enhancing video communications over the Internet and private corporate networks.

Contacts

Logitech
Pascal Eichenberger
Head of Logitech EPFL Incubator
CH-1015 Lausanne

Pascal_Eichenberger@eu.logitech.com

www.logitech.com

Media Publisher

COI representative: **Rod Bacon**

Company overview

Media Publisher is an Emeryville, California-based SME.

The company provides solutions based on a modular software platform for media capture, management and delivery (publishing). The software platform centrally manages the video communications lifecycle, simplifying how organizations create, publish, manage, report, and deliver thousands of live and on-demand webcasts securely across corporate networks.

Products

Media Publisher has two enterprise product lines:

- Enterprise software solutions/services for content publishing
- Appliances/servers for content publishing.

Services

Media Publisher provides servers and a variety of value added services to telecommunications service providers such as AT&T.

Contacts

Media Publisher
Rod Bacon
2000 Powell Street, Suite 1650
Emeryville, CA 94608 USA

rbacon@mediapublisher.com

www.mediapublisher.com

MG Taylor

COI representative: **Tim Siglin**

Company overview

MG Taylor is a Nashville, TN-based SME providing products and services which improve the performance of groups.

Primarily focusing on large group interactions, MG Taylor offers furniture, custom-designed and modular spaces (architectural services), and integrated technologies which together form the platform for special meeting processes and facilitated group discovery.

Products

In Europe, MG Taylor has clients in Spain, Switzerland and Italy. The UniCredit NavCenter, in Turin, Italy, opened its doors on January 29, 2007.

This project is a benchmark for MG Taylor - perhaps, one of the first NavCenters capable, on a regional scale, of the mission set out by Matt Taylor a quarter of a century ago.

The largest open space can hold 300 people, and the space around it an additional 200, yet a team of five can be comfortable, and work effectively, in this same room without feeling lost. To use this space requires active interaction and participation on the part of the users. This is a designed-in feature of the system. There are 400 lights in the main area - each is individually controlled. There are 40 video displays with up to 8 different content feeds to each one, a sound system that can move the building, sound pickup, remote and hand controlled video cameras and, of course, plugs for electricity, media and computers.

Contacts

MGTaylor
Matt Taylor
me@matttaylor.com

MG Taylor
Tim Siglin
701 Yadkin Street
Kingsport, TN 37660 USA

tsiglin@gmail.com

www.matttaylor.com

Nokia

COI representative: **Imre Kiss, Julia Turku**

Company overview

Nokia is a large enterprise based near Helsinki, Finland. It is currently the world's largest manufacturer of mobile telephones, with a global device market share of approximately 38% in Q2 of 2007. Nokia is by far the largest Finnish company, accounting for about a third of the market capitalization of the Helsinki Stock Exchange (OMX Helsinki).

Products

Nokia produces mobile phones for every major market segment and protocol, including GSM, CDMA, and W-CDMA (UMTS). The corporation also produces telecommunications network equipment for applications such as mobile and fixed-line voice telephony, ISDN, broadband access, voice over IP, and wireless LAN.

Nokia Enterprise Solutions offers businesses, corporations and institutions a broad range of products and solutions, such as enterprise-grade mobile devices, underlying security infrastructure, software and services. Nokia also works with a range of companies to provide network security, bring mobilized corporate e-mail and extend corporate telephone systems to work with Nokia's mobile devices

Contacts

Nokia
Dr. Imre Kiss
Visiokatu 1
33720 Tampere, Finland

imre.kiss@nokia.com

www.nokia.com

Oracle

COI representative: **Carlo Tarantola**

Company overview

Based in Redwood City, California, Oracle is a large enterprise providing enterprise software and services to manage information. Globally, the company offers database management systems (DBMS), tools for database development, middle-tier software, enterprise resource planning software (ERP), customer relationship management software (CRM) and supply chain management (SCM) software.

Products

Numerous Oracle products could be important in the meeting information value chain, especially in the domains of knowledge management.

In the area of meetings, Oracle has Oracle Real Time Collaboration Suite, a set of tools permitting team Web conferencing, calendar and document management for team, group and business processes. The Collaboration Suite includes software for live audio and video web meetings.

In addition, Oracle mobile collaboration solutions permit people on the move to connect with meetings in progress and other enterprise collaboration resources.

Contacts

Oracle
Carlo Tarantola
522 chemin de la Colle
06560 Valbonne FR

carlo.tarantola@oracle.com

www.oracle.com

ParkWood Advisors

COI representative: **John Parkinson** (not attending the COI Workshop)

Company overview

ParkWood Advisors is a small Chicago, IL-based enterprise providing consulting services for innovation and process re-engineering to large enterprises globally.

Products

ParkWood Advisors develops custom products for clients.

John Parkinson is also a public speaker and author of many articles on the topic of emerging technologies. He is the author of Meetings in 2020, distributed at the workshop.

Contacts

ParkWood Advisors, LLC
John Parkinson
1130 E Ashley Road
Lake Forest, IL 60045

john@parkwoodadvisorsonline.com

www.parkwoodadvisorsonline.com

Polycom

COI representative: **Phil Keenan, Ed Gonen**

Company overview

Based in Pleasanton, California, Polycom is a manufacturer of teleconferencing and videoconferencing equipment. It is the videoconferencing market share leader in terms of units sold per quarter, approximately 42% according to a recent survey by Wainhouse Research. The company commands around 75% market share for tabletop voice conferencing devices. It is a medium sized company with approximately $580M in sales annually.

Products and services

ParkWood Advisors develops custom products for clients.

Polycom provides both audio (tele)conferencing and videoconferencing products and platforms for enterprise and service providers.

In the area of videoconferencing, it offers a range of products from personal desktop (software) up to telepresence solutions.

The company also provides videoconferencing servers/infrastructure hardware and software products for the monitoring and management
of networks for video meetings.

Contacts

Polycom
Phil Keenan
9040 Roswell Road
Suite 400
Atlanta, GA 30350 USA

phil.keenan@polycom.com

www.polycom.com

Quindi

COI representative: **Stan Rosenschein** (not attending the COI Workshop)

Company overview

Quindi is a small enterprise based in Los Altos, California. The company is develops and markets collaboration software allowing professionals and organizations to capture and share spoken information, video and data introduced in meetings.

Products

The company's first product, Quindi Meeting Companion, is used to record meetings and to enhance the meeting participants' ability to re-purpose and enrich past meetings

Contacts

Quindi Corporation
Stan Rosenschein
4966 El Camino Real
Suite 210
Los Altos, CA 94022 USA

Stan@quindi.com

www.quindi.com

RADVISION

COI representative: **Eli Doron** (not attending COI Workshop)

Company overview

RADVISION is a medium sized Israeli-based research, US-based management medium size company. The company's products and technologies (developer toolkits) are used for videoconferencing, video telephony, and the development of converged voice, video and data over IP and 3G networks. RADVISION solutions support SIP and H.323, as well as ISDN and 3G network protocols.

Products

The company's product divisions provide servers, gateways, streaming systems, management software for meeting networks, as well as a software application platform for conferencing and collaboration (formerly CUSeeMe, acquired when RADVISION purchased First Virtual Corp).

The company's infrastructure solutions for videoconferencing are used by enterprises and service providers.

Contacts

RADVISION
Eli Doron
24 Raoul Wallenberg St.
Tel Aviv, 69719 Israel

elid@radvision.com

www.radvision.com

SMART Technologies

COI representative: **Gerald Morrison**

Company overview

A medium size company with 800 employees based in Calgary, Alberta, SMART Technologies is market-segment leader in easy-to-use interactive whiteboards and other group collaboration tools.

Products

The company provides a full line of interactive whiteboards and related products such as interactive pen displays and software for group collaboration and management of its devices on networks. The company also offers interactive digital signage, wireless slates and software for desktop audio, video and data conferencing, classroom management and concept mapping.

Contacts

SMART Technologies
Gerald Morrison
1207 11th Ave SW
Suite 300
Calgary, Alberta, Canada

GeraldM@smarttech.com

www.smarttech.com

Spiderphone

COI representative: **Pierre Wellner** (not attending COI Workshop)

Company overview

Spiderphone is an SME based in New York and Martigny, Switzerland. It is a past AMI Consortium member.

Services

Spiderphone provides audio and web conferencing services for enterprises and individuals by way of an easy to use web interface.

The company offers meeting recording and archive management/hosting services as part of its portfolio.

Contacts

Spiderphone.com, Inc.
Pierre Wellner
135 William Street
7th Floor
New York, NY 10038 USA

wellner@spiderphone.com

www.spiderphone.com

TANDBERG

COI representatives: **Snorre Kjesbu and Michel Bais**

Company overview

TANDBERG is a Lysaker, Norway-based medium size company providing several lines of products for videoconferencing. TANDBERG frequently is the market share leader in videoconferencing, in terms of revenues. It generates approximately $345M in sales per year. In October 2005, the company acquired Ectus Ltd, a New Zealand-based software development company specializing in streaming and media archiving software.

Products and services

TANDBERG offers a wide array of terminal products for audio, video and data conferencing ranging from personal desktop (software) up to telepresence rooms (partnership with HP).

The company also provides videoconferencing servers/infrastructure hardware and software products for the monitoring and management of networks for video meetings.

Contacts

TANDBERG
Snorre Kjesbu
Philip Pedersens vei 20
1366 Lysaker, Norway

snorre.kjesbu@tandberg.com

www.tandberg.com

WebEx

COI representative: **Steven Li**

Company overview

Based in Santa Clara, California, WebEx Communications is a Cisco company that provides on-demand collaboration, online meeting, web conferencing and video conferencing applications and hosted services.

Products and services

WebEx provides a suite of applications specifically designed for business processes such as sales, support, training and marketing processes. These include:

WebEx Meeting Center - Recreates face-to-face meetings with real-time data, application, voice and video sharing capabilities.

WebEx Sales Center - Features automatic attention notification to alert sales professionals when they are losing a prospect's attention, branded prospect portals to provide a secure location in which to share information and real-time sales analytics and reporting.

WebEx LiveStream - Specifically designed for large events (100 - 200,000 attendees) which require onsite production, TV quality video, voice and powerpoint.

Designed for small businesses, WebEx WebOffice provides an on-demand collaboration suite including a document manager, group calendar, database manager, task manager and several other collaborative business tools.

MeetMeNow is a web meeting application for individuals.

Contacts

WebEx
Steven Li
3979 Freedom Circle
Santa Clara, CA USA

Steven.li@webex.com

www.webex.com


AMI
CONSORTIUM


MINI
PROJECTS


AMI
CONSORTIUM

AMI Consortium COI Mini-projects

to foster collaboration on applied research and accelerated transfer of AMI technologies

Overview

- Applied research or prototyping focusing on area of mutual interest
- Mini-project partners (minimum)
  - 1 AMI Consortium partner
  - 1 COI member
- Funds
  - AMI funds work of its scientists
  - COI member funds equal effort to be executed by COI staff/sub-contractor
  - Funding only for staff resources

Types of Mini-project

- Applied research
  - Engineering prototypes/concept demonstrators
- Feasibility studies
- User acceptance/usability testing
- Market assessment/business analysis
- Any pre-competitive research, development or marketing related activity of mutual interest

Process

- **Meet and discuss ideas with potential collaborators**
- **Respond to call for proposals (submit)**
- Selection by panel of AMI Consortium members
- Project awarded within 30 days of submission
- Project initiation/execution
- Project results submitted/demonstrated
- Payment awarded at conclusion

Program timeline

Eligibility

- All AMI Consortium partners
- All Community of Interest members
- Third parties – CoI members may outsource their contributions to a Mini-project e.g. engineering/market assessment/user trial facilitation etc.

Criteria for selection

- Topic focus
  - Aligned with scope of the AMIDA project
- Matching funds/effort
  - Demonstrate equal and appropriate investments, appropriate resource/skill levels or access to them
- Technology transfer potential
  - Deemed likely to lead to transfer of AMI technology, if project results are positive

Confidentiality

- Mini-project proposals/reports public domain as with all EU funded research
- AMI Consortium partners share rights to mini-project generated IP for research
- AMI partner must be able to publish results
- COI member & AMI partner can sign NDAs and agree/execute work outside the Mini-project's reportable objectives

IPR *during* Mini-project

- AMI/AMIDA background IPR made available via evaluation license(s)
- work of AMI Consortium partners covered by AMIDA project IPR agreement
- work of 3rd parties owned by 3rd party
- joint work has separately owned background IPR (AMI and COI) with all parties having right to exploit for purpose of mini-project

IPR *after* Mini-project

- Foreground IPR:
  - AMI partner and COI member both have right to exploit foreground IP generated on the Mini-project
- Background IPR:
  - Terms & conditions for COI partner use of AMI background IP to be negotiated with relevant AMI partner's technology transfer team & legal teams

End of project report

- A report must be submitted describing
  - Results of study
  - Demonstration developed
  - Results of evaluation/assessment
- After a mini-project
  - A follow-on AMI Mini-project?
  - A Mini-project in a new area?
  - A COI-funded evaluation/development project
  - Licensing?

Funds

- AMI Consortium contribution
  - At AMI Consortium daily rates
  - 10-30K €/mini-project
- COI member
  - Establish/publish daily rate
  - Actual cost breakdown for project in final report
- Awarded to AMI Consortium partner at conclusion of mini-project
- Awards based on delivery of commitments


Thank you
Any Questions ?
Contact cperey@perey.com


AMI
CONSORTIUM

With the Community of Interest

Call for Proposals

Background

The AMI Consortium's second major collaboration, the AMIDA Project, is a European Union Framework Programme 6 Research Project which seeks to improve the experience of meetings and increase business productivity by using advanced signal processing and machine learning on human-to-human communications during enterprise meetings which involve both remote and local participants.

In order to increase the collaboration between AMI Consortium partners working on AMIDA and the members of the AMI Community of Interest, financial resources have been ear-marked to support collaborative research or proof of concept projects.

This Call for Mini-Project Proposals sets out the guidelines and describes how interested AMI Consortium partners and CoI members can prepare proposals to receive support from AMIDA for technology transfer.

Selection criteria

AMI Mini-projects will be selected on the basis of:

Focus: project proposal should be directly and clearly aligned with the scope of the AMIDA project

Matching funds: the effort contributed by the COI member(s) must be equal to or exceed the AMI Consortium partner(s) effort. We recognize that different daily rates may apply to different expertise/resources contributed.

Technology transfer: deemed likely to lead to transfer of technology or to a new project/service contract if the initial project results are positive

Submission form, review/selection process and acceptance procedure

With the aim at minimizing the proposal submission overhead, resulting in easy and fast selection, a simple submission form is provided to all candidates. Proposers are encouraged to carefully follow the guidelines, avoiding general statements, favoring clear project objectives (well in line with current AMIDA focus) and clear progress evaluation metrics (list of milestones to be achieved).

Submissions will be accepted electronically by filling the form in MS Word or PDF format as an attachment to an e-mail sent to cperey@perey.com. An e-mail acknowledgement of the submission will be received within 48 hours. Project proposals will be evaluated within 15 days of submission by the AMI Mini-project Selection Committee, including possibly end user organizations, appropriate WP heads, and AMI partner representatives.

Although the AMI partner will be the Mini-project manager of record, for EC review and project communications purposes, notification of proposal acceptance or rejection will be sent to the co-proposers by e-mail within 30 days of proposal submission.

Financial and contractual obligations

Grants range and duration

Awarded grants can range from 10,000.- to 30,000.- Euros, covering only AMI human resources. A project shall have a budget for a minimum of 2 man-months, and a maximum of 6 man-months. The project shall be executed within a maximum of 12 calendar months of award of the mini-project.

Matching funds

COI member(s) must contribute value of human resources equal to or exceeding AMI human resources investment. AMI will pay for AMI partner human resources/engineering and research. COI member partner(s) must financially support their human resource investments.

Evaluation License(s)

AMI partner(s) are responsible for identification for the AMI intellectual properties needed for a proposed project. Evaluation license(s) with appropriate AMI partner(s), specifying the mini-project title and estimated duration/scope, must be signed by the COI member(s) prior to start of the project. The evaluation license will only cover the duration of the mini-project. See sample Evaluation License separately. AMI partner evaluation licenses may differ.

Collaboration Agreements

Upon AMI award of the mini-project grant, the partners will sign the mini-project collaboration agreement. This will not in any way address the rights of commercialization upon conclusion of the mini-project.

Eligibility

Mini-projects will only be considered for funding if they involve at least one COI Member and one AMI Consortium partner, with principal investigators/project leaders clearly named. The AMI Consortium partner will be designated the Project Manager.

Confidentiality and non-disclosure clauses

Due to the public nature (European Commission FP6) of the source of AMI funding and to comply with regulations prohibiting unfair competitive advantages, it is not possible for the AMI Consortium to fund research or development of prototypes or to conduct studies which are not subsequently available to the public.

Mini-projects involving confidential or sensitive technology strategies or emerging products may be conducted by AMI Consortium members and COI members on the basis of AMI technologies, but *cannot be funded as part of this program*.

End of Project Reports

At the conclusion of the mini-project, the partners must prepare, in order to receive the funding from AMIDA project:

- a written report (for internal AMI Consortium use), summarizing the min-project results
- a financial report showing partners expenses (human resources expressed in man-days)

Deadlines

The mini-projects can run any time during the life of the AMIDA project (January 2007-June 2009). The call for proposals is open and submissions will be welcomed at any time prior to January 31, 2009. Dates for proposed mini-project start and completion should be part of the included in the description of tasks/milestones.

Contact and information

Questions about the application procedure should be directed to Christine Perey. cperey@amiconsortium.org


AMI
CONSORTIUM

Proposal Submission Form

Contact Info

AMI Consortium Partner Proposer (project leader) Details

Title	
Family Name	
First Name	
Email	
Institution	

Community of Interest Member Co-proposer Details

Title	
Family Name	
First Name	
Email	
Institution	
Position	
Areas of responsibility within company	

Community of Interest Member Co-proposer Details (if applicable)

Title	
Family Name	
First Name	
Email	
Institution	
Position	
Areas of responsibility within company	

Scientific Content

Proposal Title

Abstract (max 1000 Characters, approx. 150 words)

Short, but clear and precise!!! Clear project objectives and progress evaluation measures.

Key Words (open format, max 400 characters, approx. 50 words)

Text of proposal (max 10000 characters, approx 1500 words)

Should be clear and precise, clearly addressing the objectives of the AMIDA project, the mini-project and the proposed progress evaluation measures (milestones).

1. Background: What is the opportunity/problem to be examined? **2. Mini-Project Objectives:** How does your mini-project help clarify this problem or solution? Is this more a User acceptance project, Scientific research project or are you seeking to innovate in new areas? **3. Tasks (milestones you will reach):** Can be written as work packages to be undertaken in parallel or sequence. **4. Deliverables:** What will be the tangible mini-project output (report of results or demonstration)? **5. Technology which could be transferred or service could be provided by AMI partner(s) as a result of the mini-project:**

Estimated Resource Distribution by Task (using tasks listed/described above)

Activity (task list from above)	Which partner?	type of resource	man-days	daily rate (Euro)	subtotal for task
Total	Xxxx	xxxx		xxxx	Euros

Estimated Resource Distribution by Partner

Partner	type resources used	man-days	cost
Total AMI Cost	Xxx		Euros
Total COI member Cost	Xxx		Euros

In 14th Intl. Conference on Image Analysis and Processing (ICIAP 2007) & Workshop on Visual and Multimedia Digital Libraries (VMDL 2007), Modena, Italy, Sept. 10-14, 2007

Recording, Indexing, Summarizing, and Accessing Meeting Videos: An Overview of the AMI Project

Alejandro Jaimes[1], Herve Bourlard[1], Steve Renals[2], and Jean Carletta[2]
[1]*IDIAP Research Institute,* [2]*University of Edinburgh*
alex.jaimes@idiap.ch

Abstract

In this paper we give an overview of the AMI project. AMI developed the following: (1) an infrastructure for recording meetings using multiple microphones and cameras; (2) a one hundred hour, manually annotated meeting corpus; (3) a number of techniques for indexing, and summarizing of meeting videos using automatic speech recognition and computer vision, and (4) an extensible framework for browsing, and searching of meeting videos. We give an overview of the various techniques developed in AMI, their integration into our meeting browser framework, and future plans for AMIDA (Augmented Multiparty Interaction with Distant Access), the follow-up project to AMI.

1. Introduction

Over the last few years research interest in recording, archiving, and retrieving of meeting videos has increased significantly. This is due to major drops in hardware costs, broadband availability (for remote meetings), and concerns by corporations about record keeping (auditing decision-making, corporate memory, and complying with regulatory requirements, etc.).

Meetings play a crucial role in the generation of ideas, documents, relationships, and actions within an organization. The wealth of information exchanged in meetings, however, is often lost because manual creation of meeting minutes is subjective, incomplete, and captures only a fraction of the information. Audio-visual recording of meetings is therefore attractive, but leads to many practical challenges, from the infrastructure to record the meetings to the archival, indexing, and retrieval of relevant meeting segments. Given the number of meetings in most organizations, efficient and effective recording and access to meeting videos is of extreme importance, making research in content-based indexing and retrieval of meeting videos an important research area, not only because of its potential impact, but also because it requires combining research in several disciplines (e.g., speech recognition, computer vision, etc.).

In this paper, we describe the AMI project. AMI deals with meeting videos throughout the media production chain: from modeling of meetings, to recording infrastructure and recording, to multimodal, automatic indexing, retrieval, and browsing of meeting videos. We give a general overview of each of the components above and discuss use of AMI technologies within the framework we have developed for browsing, searching, and summarization of meeting videos. The goal of this paper and its main contribution, therefore, is to give an overview of the technologies developed in the project and their integration within applications for searching and browsing.

Related Work. Meeting room projects focus on portable recorders [10], speech [4], modeling [1], video capture [13], and others. The AMI project's components build on and improve the state-of-the art in many areas, and since this paper gives a general overview we refer interested readers to specific AMI publications [6] for details on how specific techniques developed within AMI differ from related work.

2. Instrumented Meeting Rooms & Ami Corpus

Three meeting rooms were designed and constructed at AMI partners IDIAP, TNO and the University of Edinburgh. These rooms, which were designed for the recording of videos of four person meetings, all contained a set of standardized recording equipment (plus additional cameras, microphone arrays, and binaural manikins):

- 6 cameras: 4 providing close-up views of the participants, 2 providing a view of the whole room;
- 12 microphones: a headset microphone per participant and an 8-element circular microphone array;
- data projector capture (VGA);
- white board capture and digital pen capture.

The meeting rooms were used to record the AMI Meeting Corpus [6], which consists of 100 hours of meeting recordings. The corpus includes manually produced orthographic transcriptions of the spoken dialogues, aligned at the word level with the common time line, and annotations describing participant behavior during the meetings (e.g., dialogue acts; topic segmentation; extractive and abstractive summaries; named entities; limited forms of head gesture, hand gesture and gaze direction; movement around the room; emotional state; head localization, etc.).

The corpus consists of two types of meetings: (1) remote control design scenario (approx. 2/3 of AMI corpus), (2) free topic. In the design scenario, each group of four participants had four meetings and given tasks to complete between meetings (with the final goal of designing a T.V. Remote control). Participant roles were driven in real-time by emails and web information. This control made it easier to understand the content of the meetings, enabled the construction of ontologies, and the building of outcome measures (e.g., preferred design output). The meetings are also replicable, enabling system-level evaluations. Free topic meetings were naturally occurring meetings in a range of domains. The project further developed NXT (NITE XML Toolkit [7]), an open source XML-based infrastructure for the annotation and management of multimodal recordings. NXT consists of libraries from which user interfaces for annotating and searching annotations of multi-modal data sets can be easily built. Within AMI, new tools for annotation were created, for instance for dialogue acts, named entities, topic segmentation, summarization, and a generic time-aligned coder and display.

3. Audio-Visual Processing

AMI work in audio-visual processing was primarily concerned with the development of algorithms that, given raw audio-visual streams, can automatically answer each of the following questions [1]:

- What has been said during the meeting? (Speech recognition)
- What acoustic events and keywords occur in the meeting? (Keyword spotting)
- Who and where are the persons in the meeting? (Localization and tracking)
- Who in the meeting is acting or speaking? (Speaker tracking)
- How do people act in the meeting? (Gesture and action recognition)
- What are the participants' emotions in the meeting? (Emotion)
- Where or what is the focus of attention in meetings? (Focus of attention)

Speech recognition. AMI developed systems for two types of microphone configurations in the instrumented meeting rooms (close-talking headset microphones and tabletop microphone arrays), focusing on the headset microphone conditions to develop core acoustic modeling approaches, but with an overall orientation to tabletop microphone arrays, which are less intrusive [15]. The AMI speech recognition effort addressed several research issues including the following:

- microphone array beamforming: filtering and combining the individual microphone signals to enhance signals coming from a particular location (and suppressing competing locations);
- development of novel acoustic parameterizations, including approaches based on posterior probability estimation;
- automatic construction of domain-specific language models using text extracted from the web;
- acoustic segmentation;
- development of a flexible large vocabulary decoder, based on a weighted finite state transducer formalism.

AMI developed an evaluation framework that is generic, flexible, comparable, and that allows us to conduct research and development in a stable environment. Using this framework, our system obtains exceptionally good results on AMI meeting data; in international technology evaluations organised by NIST, no other system was significantly more accurate than the AMI system on close-talking microphones [16]. This system has been used to decode the complete AMI corpus (using an n-fold cross-validation technique). The transcriptions have been used for tasks such as summarization and topic segmentation.

Keyword spotting. In acoustic keyword spotting (KWS), the goal is to find keywords and their position in speech data. AMI developed three approaches: acoustic, LVCSR, and a hybrid approach [17]. In the *acoustic approach*, a keyword score is obtained by comparing the posterior probability of the keyword phonetic model, with a background model. This is very fast since many of the key parameters may be precomputed. It is relatively precise (the precision increases with the length of the keyword) and any word can be searched provided its phonetic form is available. It is ideal for on-line applications (such as monitoring remote meetings), but it is not suitable for browsing huge archives, as all of the acoustic data must be processed for each search. The *LVCSR lattice approach* locates the keywords in lattices generated by

a large vocabulary continuous speech recognition system. Given the output of the speech recognizer, this approach is very fast, but it is accurate only for frequently occurring words. There is a degradation in performance for less common words, which is a drawback since these words (such as technical terms and proper names) carry most of the information and are likely to be searched by users. Therefore, this approach has to be complemented by a method unconstrained by the recognition vocabulary. The *hybrid phoneme lattice approach* is based on the construction of graphs of phoneme probabilities, from which the phonetic form of the keyword may be extracted. This is a reasonable compromise in terms of accuracy and speed. Currently, AMI work on indexing phoneme lattices using tri-phoneme sequences is advancing and preliminary results show good accuracy/speed trade-off for rare words.

Speaker tracking. The objective of speaker tracking is to segment, cluster and recognize the speakers in a meeting, based on their speech. The first approach developed in AMI uses the acoustic contents of the microphone signal to segment and cluster speakers. In the NIST evaluations this system produced very good results for speech activity detection (the lowest error rate reported) and for speaker diarization (who spoke when). The second approach developed in AMI, based on cross-correlations between microphone signals operates in real time, and has been integrated with the online keyword spotter [11] .

Localization and tracking. Location coordinates of each person in the meeting are an essential input to various meeting analysis tasks, including focus of attention and action recognition. The steps required are identification, localization, and tracking. For identification, generative approaches have proven to be the most robust so in AMI a variety of models with different trade offs between speed and accuracy have been used (e.g., based on Gaussian mixtures and HMMs). The algorithms have been developed as a machine vision package for the open source machine learning library, TORCH (extended within AMI (http://www.torch.ch)). For localization and tracking AMI developed, applied, and evaluated four different methods including approaches based on dynamic Bayesian networks, active shape trackers using particle filters, and face trackers based on skin colour.

Gesture and action recognition. We have defined a set of actions and gestures that are relevant for meetings (e.g., hand, body, and head gestures such as pointing, writing, standing up, or nodding). Special attention has been paid to negative signals, such as a negative response to a yes-no question, usually characterized by a head shake. This kind of gesture contains important information about the decision making in meetings, but can be very subtle and involve little head movement, making automatic detection very difficult. For gesture recognition two methods were applied: *Bayesian Information Criterion* and an *Activity Measure* approach. For each person in the meeting, the 2D location of the head and hands, a set of nine 3D joint locations, and a set of ten joint angles were extracted. In addition classification was performed in the segmented data. Due to the temporal character of gestures the focus was on different HMM methods. Gestures like standing up and important speech supporting gestures produced satisfactory results (100% and 85% recognition rate, respectively). However, the results for the detection of negative signals were not significantly better than guessing. Detecting gestures such as shaking or nodding is challenging and requires disambiguating the meaning of very subtle head movements.

Focus of Attention. Gaze detection requires higher resolution of facial images than what is available in the AMI corpus. As an approximation, we have developed algorithms for tracking the head and estimating its pose, based on a Bayesian filtering framework, which is then solved through sampling techniques. Results (evaluated on 8 minutes of meeting recordings involving a total of 8 people) were good, with a majority of head pan (resp. tilt) angular errors smaller than 10 (resp. 18) degrees. As expected, we found a variation of results among individuals, depending on their resemblance with people in the appearance training set. In addition, we formulated focus of attention (FoA) as a classification task by automatically classifying FoA into one of the following categories: meeting participants, objects in the meeting room, and an “unfocused” location. Experiments using the ground truth head-pose pointing vectors resulted in frame-based classification rate of 68% and 47%, depending on the person's position in the smart meeting room. Accuracy is lower than reported in other works, mainly because of the complexity of the scenes and number of categories. Exploiting other features/modalities (e.g speaking status) in addition to the head pose can be used to disambiguate FoA classification. We found that using the estimated head-pose instead of the ground truth did not degrade the results strongly (about 9% decrease, thus much less than the differences w.r.t. position in the meeting room), which was encouraging given the difficulty of the task. We also found that there was a large variation of recognition amongst individuals, which directly calls for approaches such as

Maximum A Posteriori techniques for the FoA recognition (the topic of current research).

4. Content Extraction

Dialogue acts are labels for utterances which roughly categorize the speaker's intention. They are useful, for example as part of a browser which highlights all points where a suggestion or offer was recognized. Dialogue acts also serve as elementary units, upon which further structuring or discourse processing may be based (e.g., summarization). The following dialog act labels were used:

- Information exchange: giving and eliciting information;
- Possible actions: making or eliciting suggestions or offers;
- Commenting on the discussion: making or eliciting assessments and comments about understanding;
- Social acts: expressing positive or negative feelings towards individuals or the group;
- Other: utterances which convey an intention, but do not fit into the four previous categories;
- Back channel, Stall and Fragment: utterances without content, which allow complete segmentation of the material.

We have used combinations of machine learning based on a multimodal set of features, including a word-based language model, prosodic features (based on duration, energy and intonation), context features (e.g., speaker overlap), and discourse features (history of previously recognized dialogue acts). Using generative models that explicitly take account of the dependence on multiple streams of data (such as dynamic Bayesian networks, factored language models, and hidden event language models) we have obtained state-of-the-art results for dialogue act segmentation. Interestingly, although the best approach to dialogue act segmentation involves jointly segmenting and labeling the dialogue act sequence, we have found that the labeling may be substantially improved by re tagging using discriminative approaches, in particular conditional random fields. Comparing the performance on automatically transcribed speech with human transcribed speech, we find that the performance of dialogue act recognition drops by about 10%.

Topic segmentation. The aim of topic segmentation is to automatically infer the sequential structure of the meeting by topic (and sub-topic); it differs from dialogue act recognition in that the fundamental units (topics) are typically many minutes in duration. We have explored two basic approaches to this task. An unsupervised approach, LCSeg automatically infers (without training) topic boundaries as points where the statistics of text change significantly. The supervised approach, on the other hand, learns topic boundaries based on a hand-annotated training set. An advantage of the supervised approach is that it is possible to use additional features relating to prosody (e.g., pauses) and the structure of the conversation (e.g., speaker overlap). These additional features are also relatively independent of errors in the automatic speech transcription. We have also developed approaches to automatically generate labels for topics, based on the statistics of the automatically transcribed words that make up a topic.

Summarization. We have investigated two distinct ways of constructing summaries of a meeting. Extractive techniques construct summaries by locating the most relevant parts of a meeting and concatenating them together to provide a 'cut-and-paste' summary, which may be textual or multimodal. Abstractive summaries, on the other hand, are similar to what a human summarizer might construct, generating new text to succinctly describe the meeting. Abstractive summarization is more challenging than extractive summarization, and requires relatively deep domain knowledge. Our approach to extractive summarization is based on automatically extracting relevant dialogue acts from a meeting. It thus requires (as a minimum) the automatic speech transcription and the dialogue act segmentation modules described above. Lexical information is clearly extremely important for this task, but we have found it beneficial to augment information derived from the transcription with speaker features (relating to activity, dominance and overlap), structural features (the length and position of dialogue acts), prosody, and discourse cues (phrases which signal likely relevance). All of these features are important to develop accurate methods for extractive summarization. We have also explored reduced dimension representations of text, based on latent semantic analysis, which also add precision to the summarization. Using an evaluation measure referred to as weighted precision, we have discovered that it is possible to reliably extract the most relevant dialogue acts, even in the presence of speech recognition errors. We have explored “dialogue act compression,” in which the extracted dialogue acts are condensed by removing irrelevant portions. Again, taking account of speech features such as the overall intonation contour of the dialogue act helps to improve the overall performance. We have also implemented a prototype abstractive summarization system, based on an ontology of the AMI scenario meetings, together with annotations of propositional content, and the topic structure of the meetings. Given these annotations an

ontological representation is built, which is then passed to a natural language generation component which produces a one paragraph summary of the meeting.

Influence and dominance detection. Person-to-group influence (i.e., influence of a person over the group) is estimated from audio features with a framework based on a two-level Dynamic Bayesian Network, in which an influence distribution is defined as the prior probability of individual state streams contributing to the group state stream. Such a distribution can be automatically estimated from data and was tested on AMI spoke data. Dominance relations between meeting participants has also been inferred. Using SVMs we were able to predict who is more, less or normally dominant in a meeting with an accuracy of 75%.

Video content extraction. We have developed "automatic camera operator" algorithms based on extracted video and audio features to perform this operation. Subjective evaluation with users indicated that the deployed algorithms were functionally acceptable, but were of significantly lower aesthetic quality compared with human production. We have also developed methods for identifying "hot-spots" such as laughter, directly from video features based on things such as motion and texture.

5. AMI Meeting Browsers

Many AMI technologies are demonstrated within a Java-based browsing framework, referred to as JFerret. JFerret is a multimedia browser that is extremely flexible, enabling almost any user interface to be composed, using a combination of plug-in modules. An XML configuration specifies which plug-in components to use, how to arrange them visually, and how they communicate with each other. JFerret comes with a library of pre-defined plugins, for presentation of video, audio, slides, annotation time-lines, controls, and so on, and it is straightforward to write new plugins. This has been the main route to demonstration for many of the technologies described in previous sections. Java allows the application to run cross-platform, either as a applet (inside a web-browser) or as a stand-alone application. An example JFerret configuration, enables browsing via keyword search on the speech-recognized transcript, search within captured slides, and browsing by speaker activity. Time-synchronized recordings that may be browsed include multiple video and audio streams and white board capture. Other semantically rich browser components that have been constructed include direct keyword-spotting, video hot spots, and argumentation.

We have also begun to explore techniques for time-based media compression, since this can clearly contribute to efficient browsing of recorded meetings. Time-based compression can be done in three major ways: 1) speech speedup, 2) excision of less important parts, and 3) simultaneous presentation of speech from two locations. Two interactive prototypes for accelerated listening of recorded speech have been implemented. One prototype provides support for speed controls as well as skipping ahead and back based on speaker segmentations. The other prototype presents two parts of the meeting simultaneously using binaural in two different locations so that the user can listen to one part of the meeting while monitoring another part. We also devised a PDA-based wireless presentation system, including recording of slide presentations, which was integrated with the meeting browser using VNC.

Evaluation. AMI scientists have been closely involved in several international evaluation efforts such as the NIST Meeting Recognition evaluation of speech recognition and speaker diarization in meetings, for which the AMI corpus has been one of the main data sources. AMI has also participated in the CLEAR evaluations of focus of attention and face detection. Additionally, the AMI corpus, together with speech recognition output, has been provided to the Cross Language Evaluation Forum (CLEF) for their 2007 evaluation on cross-lingual question answering. In addition, AMI developed a framework for extrinsic evaluation of browser components, called the Browser Evaluation Test (BET). The BET provides a framework for the comparison of arbitrary meeting browser setups, where setups differ in terms of which content extraction or abstraction components are employed. The BET consists of a set of experiments in which test subjects have to answer true/false questions about observations of interest for a meeting recording. The test subject uses the browser under test to answer these questions, given a time limit (typically half the meeting length). This framework has proven to be a successful way to evaluate browser components.

We have also developed a task-based evaluation that is supported by the design of the AMI corpus (about 70% of corpus meetings are based on a replicable design team scenario). In the task-based evaluation, a new team takes over for the fourth meeting, with access to the previous three meetings. The evaluation compares team performance in the existing case with basic meeting records (including powerpoint files, emails and minutes), with a basic AMI meeting browser, and with a task-based browser. The task-based evaluation is in terms of both objective measures

such as design quality, meeting duration, assessment of outcome, and behaviourial measures of leadership, and subjective measures including browser usability, workload (mental effort), and group process.

6. Conclusions and Future Work

We have provided an overview of the AMI project. The major achievements of AMI are in six areas: Instrumented meeting rooms (development of a recording infrastructure, based on instrumentation of meeting rooms, in which we can capture all aspects of interaction in a meeting, in a time synchronized manner), the AMI Corpus (a 100 hour corpus of recorded meetings, with multiple time synchronized signals across several modalities, annotated at many different levels), audio-video processing (significant advances in several areas including speech recognition, audio-video localization and tracking, and detection of focus of attention), content extraction (new state-of-the-art techniques in several areas such as summarization and dialogue act recognition), integrated demonstrations (AMI developed an integrated browsing framework in which the outputs of multimodal recognition and content extraction modules may be incorporated as plugins or data streams), and evaluation (novel frameworks for system evaluation). For each of the areas described there are many ongoing improvements and plans for future work. In general, improving robustness, speed, and accuracy are important issues, as well as scaling the techniques to deal with larger amounts of data. Within the new AMIDA project [6] we are working on improving many of the techniques, paying particular attention to their integration into a framework of "meeting assistants" that can perform in close-to real-time (i.e., delays of several seconds or even minutes may be acceptable) within applications that integrate these techniques for use during, and between meetings, in remote and co-located settings.

Acknowledgements. This work has been performed by the AMI consortium, which is a 6th Framework Research Programme of the European Union (EU), contract number: IST-033812. The authors would like to thank the EU for the financial support and the partners within the consortium for a fruitful collaboration. Part of the work presented here was also funded the Swiss National Science Foundation, through the National Center of Competence in Research (NCCR) on "Interactive Multimodal Information Management (IM2)", http://www.im2.ch. Special thanks to John Dines for useful comments.

10. References

[1] M. Al-Hames, et. al., "Audio-Visual Processing in Meetings: Seven Questions and Current AMI Answers," in Wksp on Machine Learning for Multimodal Interaction (MLMI), Washington DC, May 2006.

[2] C. Costa, P. Antunes, and J. Dias, "A Model for Organizational Integration of Meeting Outcomes," in Contemporary Trends in Systems Development, M.K. Sein et. al., eds, Kluwer Plenum, 2001.

[3] R. Cutler, et. al., "Distributed Meetings: A Meeting Capture and Broadcasting System," Proc. ACM Multimedia 2002, Juan Les Pines, France, 2002.

[4] R. Dhillon, S. Bhagat, H. Carvey, and E. Shriberg, "Meeting Recorder Project: Dialog Act Labeling Guide," ICSI Technical Report TR-04-002, 2004.

[5] W. Geyer, H. Ritcher, and G. Abowd, "Making Multimedia Meeting Records More Meaningful," IEEE ICME 2003, Baltimore, MD, July 2003.

[6] www.amiproject.org; http://corpus.amiproject.org

[7] http://sourceforge.net/projects/nite/

[8] D. Hillard, M. Ostendorf, and E. Shriberg, "Detection Of Agreement vs. Disagreement In Meetings: Training With Unlabeled Data," Proc. HLT-NAACL Conference, Edmonton, Canada, May 2003

[9] A. Janin, J. Ang, S. Bhagat, R. Dhillon, J. Edwards, J. Macias-Guarasa, N. Morgan, B. Peskin, E. Shriberg, A. Stolcke, C. Wooters, B. Wrede. "The ICSI Meeting Project: Resources and Research," NIST ICASSP 2004 Meeting Recognition Workshop, Montreal, May 2004.

[10] D.-S. Lee, B. Erol, J. Graham, H.J. Hull, and N. Murata, "Portable Meeting Recorder," in proc. ACM Multimedia 2002, Juan Les Pines, France, 2002.

[11] P. Motlicek, L. Burget, and J. Cernocky. Non-parametric speaker turn segmentation of meeting data. In Proceedings Eurospeech, 2005.

[12] I. McCowan, S. Bengio, D. Gatica-Perez, G. Lathoud, F. Monay, D. Moore, P. Wellner, and H. Bourlard. Modeling human interaction in meetings. In IEEE ICASSP, 2003.

[13] Y. Rui, A. Gupta, J. Grudin and L. He, "Automating lecture capture and broadcast: technology and videography," ACM Multimedia Systems Journal, 3-15, Springer V., 2004.

[14] B. Wrede and E. Shriberg, "Spotting Hot Spots in Meetings: Human Judgements and Prosodic Cues," in EUROSPEECH 2003, Geneva, September 2003.

[15] T. Hain, L. Burget, J. Dines, G. Garau, V. Wan, M. Karafiat, J. Vepa and M. Lincoln, "The AMI system for the transcription of speech in meetings", Proc. ICASSP Honolulu, Hawaii, 2007.

[16] S. Renals,, S. Bengio, J. Fiskus (Eds). Machine Learning for Multimodal Interaction: 3rd Intl. Workshop, MLMI'06. Springer Lecture Notes in Computer Science, Bethesda, MD, USA, 2006,

[17] I. Szöke, P. Schwarz, L. Burget, M. Fapso, M. Karafiát, J. Cernocký, P. Matejka. "Comparison of Keyword Spotting Approaches for Informal Continuous Speech", Proc. Interspeech '05 (Eurospeech), Lisabon, Portugal, 2005.

THE FUTURE OF BUSINESS MEETINGS

APPLICATIONS FOR AMI TECHNOLOGIES

TABLE OF CONTENTS

WELCOME TO THE FUTURE

Processes and technologies used by business meeting participants are continuously changing in order to increase efficiency in the workplace, or enhance meeting productivity. How can more technology help more than it hurts?

The goal of this white paper is to take what is currently known about meetings and to overlay a vision of the future, to see how the addition of new technologies based on advanced signal processing and information analysis can have positive impacts on meetings.

The reader will also learn about the AMI Project and explore how moving beyond the analysis of simple verbal communications—adding non-verbal communications—can reveal deeper trends and patterns.

Applications using AMI technology could give people the ability:

- to prepare better for upcoming meetings,
- to review parts of meetings in progress or past meetings missed,
- to analyze behaviors and positions taken by individuals or groups, and
- to attend multiple meetings without missing critical elements in any.

At a management level, having technologies which analyze verbal and non-verbal content and communications could be integrated with other enterprise managements systems to:

- be the basis of meeting behavior/methods training programs, even permitting self-analysis by participants,
- improve team construction based on team members' past meeting behaviors,
- reduce risk of disclosures and delays caused by underlying conflicts, and
- recommend strategies for human resource utilization across projects and teams.

THE AUGMENTED MULTI-PARTY INTERACTION PROJECT

The AMI Project, an EU-funded research project involving dozens of scientists across a fifteen-member consortium, focuses on meetings in order to develop intelligent software algorithms and systems. The algorithms and related technologies under development can become core building blocks on which products and services may emerge for use by people in and between meetings.

Scientists in the AMI Project bring expertise from many disciplines. They include world-renowned experts in the fields of speech processing, video/vision processing and human-computer interfaces, as well as sociology, psychology and linguistics. The focus of their research is on the human-to-human communication which occurs between people during product design meetings. The design of the research should permit expansion of the scope to include many more types of meetings and team processes.

Statistical machine learning is used by the AMI Project in the context of improving our understanding of business meetings. Once machine learning is successful software building blocks can be developed. The process of developing these core technologies involves extracting information from a large number of multimedia meeting recordings (specially developed recordings which form the AMI Meeting Corpus). All the information of interest is labeled. Based on the labels, models are developed ("trained") to recognize events, words and other patterns of interest. Then, once the models are able to reliably recognize information from sources on which they were trained, a system deployed in a meeting environment can automatically recognize patterns based on new multimedia meeting data it receives.

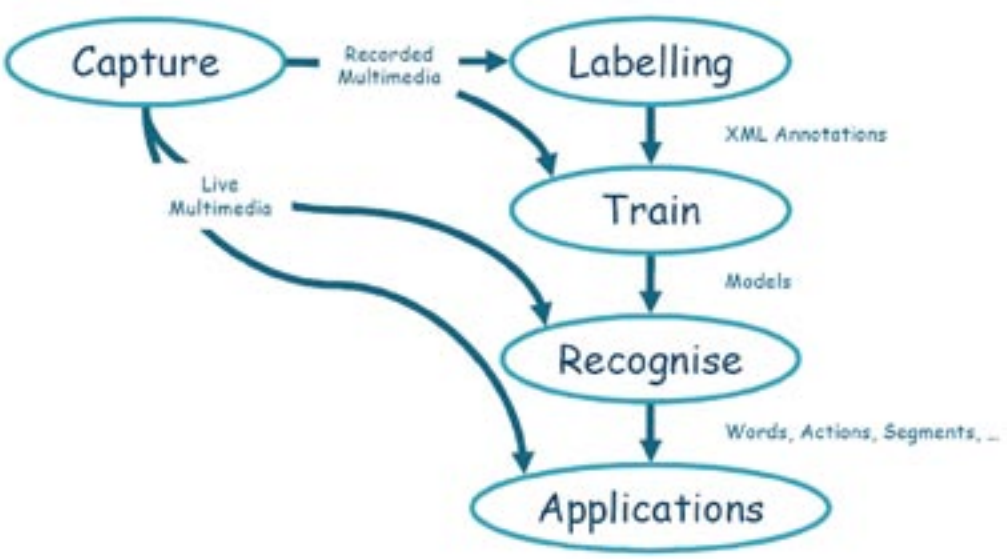


Figure 1: Statistical machine learning is used by the AMI Project to develop algorithms and technologies which can recognize meeting elements automatically.

APPLICATIONS FOR AMI TECHNOLOGIES

AMI technologies can add value to participants between and during meetings.

Between Meetings

In the future, the knowledge worker will have the need—and, using AMI technologies, will possess the tools—to access multimedia assets from the corporate knowledge base for work between meetings with fewer of the issues we encounter today.

Anyone can imagine situations in which the archives of past meetings (in which the searcher herself participated or meetings conducted by others) in a corporation would be useful to search for particular key attributes or content. The problem is that the key attributes are "lost" among the un-interesting meeting segments.

Meetings could be more effective if, prior to entering, each participant were better prepared. AMI technology can be integrated into tools that help people prepare for their meetings.

Between meetings a user of the multimedia meeting archive can:

- Review a summary of one or any number of past meetings,
- Search one or more past meetings to answer specific questions,
- Browse one or more past meetings to answer specific or general questions,
- View the entire meeting (or multiple meetings) in faster-than-real time, and
- Detect patterns exhibited by groups or individuals during past meetings which may provide insights into the upcoming meeting.

Imagine what it would be like if knowledge workers who are unable to attend a meeting to which they would have added value or from which they could benefit had access to recordings and the functionalities above. Wouldn't the process and reliability of "catching up" be different?

Better Summaries are Crucial

Today if a person misses a meeting, they must go to others who attended and ask for a summary. Meeting summaries used in business today are verbal, contain the biases of a participant's point of view and are not searchable. Sometimes summaries can be in the form of notes or minutes but most meeting participants do not have the time to formalize their conclusions in a form that is useful to others in a project team.

A summary should capture the essence of the content of a meeting. There are as many summary formats as types of meetings. One can imagine options such as:

- Bullet summary,
- Paragraph summary,
- Summary in audio, and
- Summary in audio, video and with supporting media introduced during the meeting.

Regardless of their presentation media or their depth, those who rely on them need the content of summaries to be linked to the detailed contents (the multimedia record) of the meeting. In much the way one navigates a web site or any interactive application, a summary statement should be a "window" into the meeting at the particular time when an issue is discussed or a decision made. The idea of an intelligent meeting database architecture which would be able to produce summaries of multiple meetings is also part of the AMI vision. From the summary, the user of meeting archives must also be able to search, browse and have flexible ways of accessing the contents of the meeting or multiple meetings in a database.

Searching, Browsing and Skimming Archives

When unstructured media files from a meeting archive are indexed and stored in an appropriate repository, their contents are temporally associated with structured data, consisting of other relevant information in the database (time stamps, text transcripts of speech and all written additions or information projected on the screen, names of people in a meeting, the subject of the meeting, the agenda of the meeting, and any files introduced during the meeting).

A user interface for a multimedia meeting repository provides a search function. One can imagine a dialog box in which an inquiry is entered by the user (it might be typed in using a keyboard but in the future, the user might speak or point to designate what aspects are sought). A pop-up menu might have the most frequently used queries.

Questions which the AMI Project is using machine learning to answer on its database of meetings include:

- Who is in the meeting?
- What are the participants saying?
- When and how do they communicate?
- What are they doing?
- What are their emotional states?
- What are they looking at (focus of attention)?

Based on the details, higher order questions are asked, such as:

- What topics are discussed and when?
- What decisions are made and by whom?
- What roles do the participants play?
- What positions do they take on issues?
- What activities are completed?
- What tasks are assigned or reported done?

In some cases, the person using a meeting archive may not know exactly what they are looking for. This requires a different type of interaction with the archive or the repository permits "skimming" in a linear fashion as well as non-linear browsing (through text).

JFERRET

The AMI Project architecture for developing user interfaces which can be used to navigate and browse multimedia meetings is called JFerret. It is a Java-based framework which supports the visualization and interaction with meta-data and raw data in a unified view on the user's screen. The visualization is specified through easily configured plug-in components, with XML. Developers can modify the specifications and extend the plug-ins.

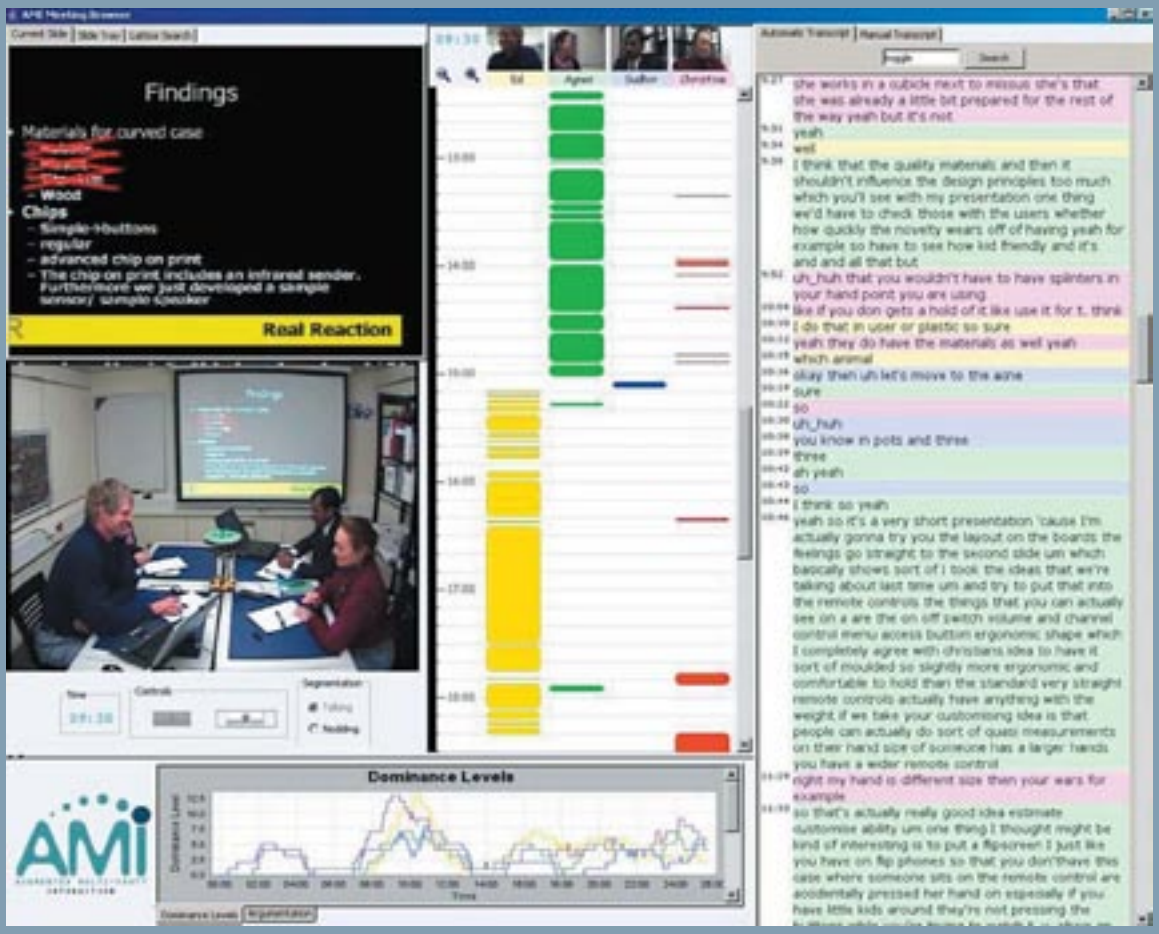


Figure 2: JFerret is a Java-based framework for creating user interfaces which support navigating and browsing meeting archives. Source: AMI Project

Accelerating Meeting Playback

The user may also seek tools to experience the meeting in less time than it took to conduct the meeting. As illustrated in Figure 3, the user can accelerate the playback of a telephone conference by only asking to hear or "see" those sections attributable to a particular meeting participant. Or can adjust the speed of the playback of all the meeting media. This is a use of AMI technologies between meetings.

Imagine being in a meeting and suddenly needing to step out to attend to an emergency or arriving after a meeting has already begun. Prior to entering or returning to the meeting, the essence of the segment missed could be obtained and permit continuing the meeting without interruption or loss of context. Figure 3 illustrates how the user would control the playback of a meeting in progress.

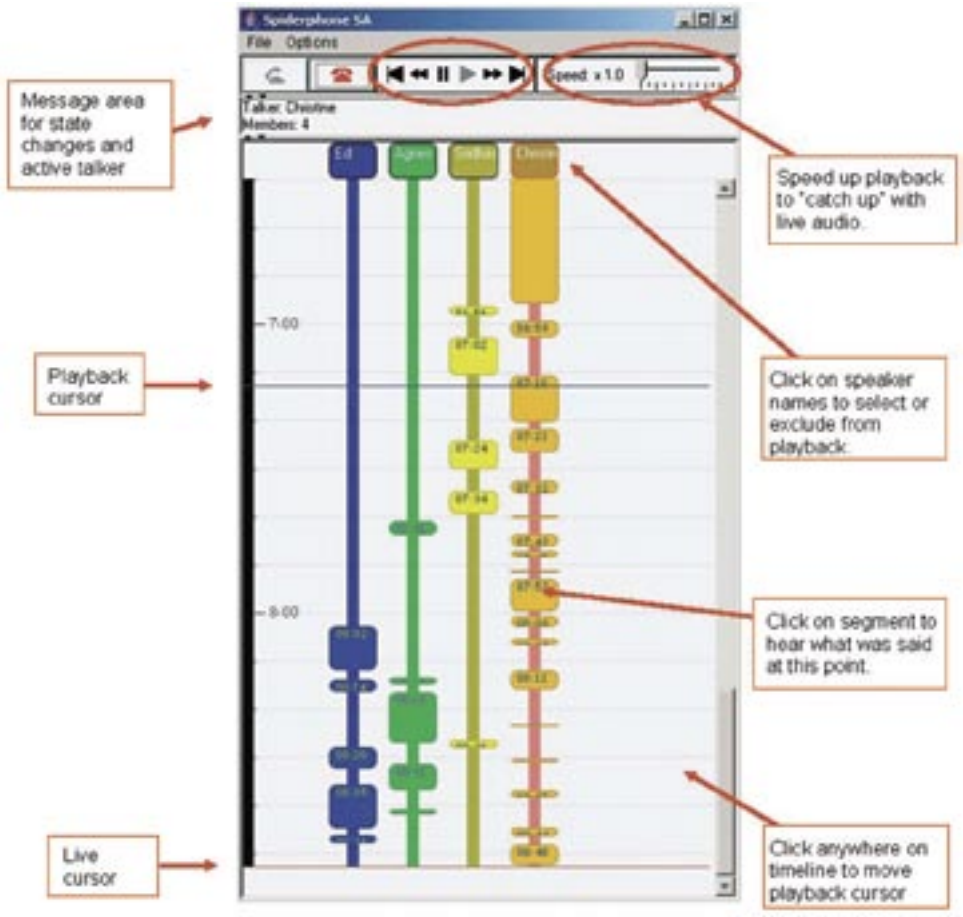


Figure 3: An AMI-assisted playback user interface developed using JFerret permits the user to control the variables and speed up the experience of a meeting archive or a meeting in progress. Source: AMI Project

Detecting Patterns

Summaries are, in some ways, the detection and compression of patterns into smaller, more accessible chunks. Patterns can come in any shape and size. They may consist of the utilization of a word or expression, a gesture or non-verbal type of communication such as nodding to indicate agreement or nodding when a person is drowsy. These are subtle differences which the human brain can distinguish and, in time, the algorithms on which AMI is working will also be able to detect and flag or enter in the database for use by meeting applications.
In some scenarios for AMI technology use, a meeting participant's gestures or position relative to others can be the cue which causes a response in a virtual representation of a remote participant. For example, as illustrated in figure 4, when all the participants in the meeting are turned towards a white board, the virtual participant is expected to turn similarly.

Detecting patterns could also help decisions in rendering agent actions (body language). If during a meeting everyone has their arms folded, would the remote participant also seek to assume this posture as well? These are other examples of how using AMI technology to detect patterns will be potentially valuable during meetings.

Figure 4: In the AMI-assisted Virtual Meeting Room, the focus of attention of the meeting participants is detected and helps an agent to behave according to meeting norms. Source: AMI Project

Support During Meetings

In much the same manner as archives can be resources to people between meetings, or that AMI can help the late meeting participant to "catch up," the recordings of past meetings should also serve as resources to participants during a meeting. Suppose participants in a meeting wish to answer a question about a previous meeting. Features similar to those accessible between meetings should be available but would also take into account the participants of the live meeting and the sensitivity of the sources or contents of past meetings.

Improving Meeting Management and Progress

There are many scenarios for improving the flow and dynamics of communication during a meeting. Since the AMI project technologies are able to measure the interactions and participation of people in a meeting, analyses could be summarized and presented to a chairperson during a meeting. Imagine a system which compares a proposed (ideal?) agenda with the progress of an actual meeting and alerts the participants about deviations from their goals.

In some more futuristic AMI application scenarios, the directives or opinions of leaders or behaviors of participants in past meetings could be privately or publicly compared with the real time progress. The comparison could be used by the meeting chair to re-orient discussions to key issues which are known to cause delays in a project, for example.

A meeting and the life cycle of a project can be shortened if known obstacles are anticipated and addressed. Imagine a meeting in which an action item is being taken to prepare an analysis of a risk. If, by accessing past meeting repositories, or having an agent which automatically compares new action items with the past, a knowledge worker can be notified that such an analysis already exists, the meeting chairperson can introduce the relevant conclusions and accelerate the project.

An AMI-based technology could help the moderator of the meeting follow how long the monologue has been in progress and intervene to involve more participants in a discussion. Metrics such as time spoken, the number of times a participant has successfully "grabbed" the floor, the number of people who are paying attention to a participant (regardless of their having the floor or not) all help to manage the process of a meeting or a project's outcome more effectively. People who repeatedly grab the floor could receive automatically generated notifications that others are finding their input valuable or irritating and permit the participant to adapt behaviors in real time given the conditions.

Meeting Agents

Frequently it is necessary for the success of multiple projects for a person to be assigned responsibilities with overlapping time requirements. Another scenario for AMI technologies includes a system which helps knowledge workers "attend" two or more meetings simultaneously. The individual may participate in one meeting in person or by telephone and request to have an agent monitor one or more meetings. Provided participants in another meeting agree, the monitoring agent can be configured to detect real time events such as changes in the agenda, discussion of a particular item on the agenda which concerns the employee directly, a new person entering the meeting or someone who is known to be important leaving a meeting. This could optimize the use of limited human resources.

In the AMI demonstration of this scenario, the Remote Meeting Assistant (RMA) will detect events (e.g., keywords, entry or exit, change in dialog, debates) which it has been configured to monitor and alert the user. These could be real time alerts (via a pop up or toaster like an Instant Message) and they could be compiled for later review. Taking action based on information provided by a RMA would require first gaining the context for the alert, perhaps by way of an accelerated playback of recent remarks or discussion.

Figure 5: The AMI technology-based remote meeting assistant helps those who cannot attend a meeting to send their agents to monitor a meeting in progress. Source: AMI Project

BRINGING SCIENCE IN CONTACT WITH BUSINESS, ALIGNING VISION WITH REAL WORLD LIMITS

In the business of meetings, it is crucial for those in the trenches, those who are managing the technologies for enterprise meetings to expand their frameworks, vocabularies and working models about business process in order to envision more efficient workforces and accelerated, highly-informed decision making. At the same time, the research community will be working on machine learning and the development of algorithms which process multi-modal signals with ever increasing accuracy, regardless of the use for these systems in enterprise. Where possible the research and business communities can nourish one another and better the world.

It is important, regardless of the scope of ones vision or on which side of the science-business divide one stands, to understand the practical limits of technology as well as the ability for business and humans to change. Some of the applications are challenging to implement in real time products and will not be realized for decades due to the processing complexity.

Only time will tell how large an impact the AMI Project will have on people, processes and technology of the future. Without very strong incentives or a clear demonstration that the risk is worth the reward, humans resist changing their behaviors. Cost of technology requiring investments greater than the foreseen return may also be an obstacle. Other scenarios explored in this article are difficult to implement for reasons other than known technology limitations. The only thing we can be certain of is that in the future there will be more business meetings, and they can be dramatically improved using new technologies.

Transfer of AMI technology into mainstream products and services is a crucial aspect of the project. Vendors developing meeting technologies must experience the research underway in the AMI Project to assess opportunities for future features and functionality. For information on the AMI research approach and to see demonstrations and screen movies of the technology prototypes in action, visit www.ami-consortium.org

AMI c/o IDIAP Research Institute
P.O. Box 592
Rue du Simplon 4
1920 Martigny - Switzerland
info@ami-consortium.org
www.ami-consortium.org

Tel +41 27 721 77 11
Fax +41 27 721 77 12

AMI CONSORTIUM

ParkWood Advisors, LLC
Business and Technology Strategy

MEETINGS IN 2020

AMI Consortium
and the Club of Amsterdam

IGC, Amsterdam
September 11, 2007

Credits

This white paper is provided to the participants of the Club of Amsterdam September 11 2007 meeting and the members of the Community of Interest by the AMI Consortium as part of an ongoing initiative to increase global study and understanding of the human-to-human communications and the future of technology-assisted meetings using automation and intelligent agents in an environment of virtually unlimited processing and bandwidth resources.

This white paper is designed to provide accurate and authoritative information in regard of the future of meetings. It is made available by the AMI Consortium, with the permission of ParkWood Advisors LLC, with the understanding that the intent is not to render legal, investment, accounting or other professional advisory services.

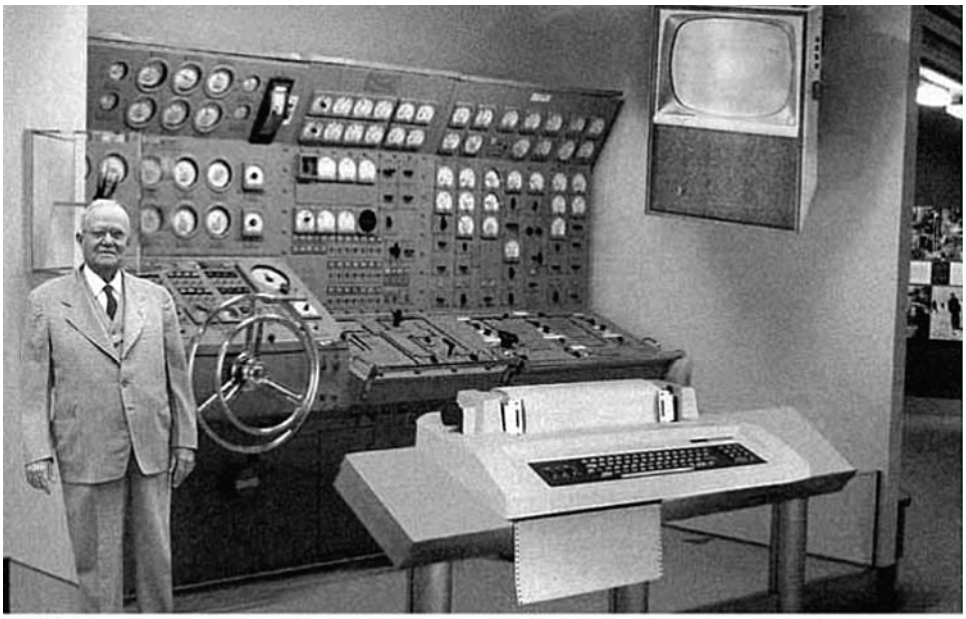
If investment advice or other expert assistance is required, the services of a competent professional person should be sought. Requests for permission to reuse the contents of this document or for further information about its contents should be addressed to John Parkinson at ParkWood Advisors LLC.

PREFACE

Speaking and digitally publishing about a subject are two very, very different things. In the case of predicting the future, however, the tangible results may be the same.

When speaking about meetings in 2020, a presenter has a lot of liberty because chances are relatively high that no one in the audience will remember what he or she said about the subject by 2020. And, in contrast with what you might expect, a digitally published/stored archive of the same concepts will probably also be "lost" for all intents and purposes. For John Parkinson, Chairman and Managing Partner of ParkWood Advisors LLC, the risk of his words fading and disappearing long before the accuracy of his predictions are tested just comes with the territory.

Parkinson introduced his talk, a keynote address at the Wainhouse Research European Forum 2006 in Berlin entitled "Meetings in 2020," with a touching reminder that predictions of the future—even the future of technology over the past 50 years—have more frequently been wrong than right. Looking back at the predictions made by the RAND Corporation in the early fifties, a satirist depicted what might have been contained in an RAND forecast.



Scientists from the RAND Corporation have created this model to illustrate how a "home computer" could look like in the year 2004. However the needed technology will not be economically feasible for the average home. Also the scientists readily admit that the computer will require not yet invented technology to actually work, but 50 years from now scientific progress is expected to solve these problems. With teletype interface and the Fortran language, the computer will be easy to use.

Figure 1: This figure illustrates how difficult it is, even for experts, to get the future right

All predictions/forecasts and recommendations made in this paper are the rights and responsibilities of ParkWood Advisors LLC.

Only time will tell how well the words of 2006 will fit the future.

Christine Perey, AMI Consortium Technology Transfer

THE FUTURE IS FUZZY

One of the hardest things for a futurist to do is to take people out of the familiar context of today, and yet, it is absolutely necessary to break free of the present if one is to catch a glimpse of the future.

Humans have a perception of change over time which is distorted—most interpret change as linear. Unfortunately, ten years in the future is about as different from today as twenty-five to thirty years in the past. And this departure from linearity is accelerating. In order to understand how different the world will be in 2020, comparing today with an equal time in the past (14 years at the time of this writing) would be misleading. Rather, one needs to compare what is going on today with what was status quo in 1965. No cell phones. No personal or portable computers. No Internet. Much greater obstacles to travel.

In order to get a glimpse of the future of meetings, we need to build scenarios on which we can project trends. But scenarios do not make predictions. They only provide a framework in which to position a few assumptions and to examine the consequences if the assumptions, in fact, are true.

There are several rational scenarios in which it will become either very expensive or relatively dangerous to travel to routine meetings. In one scenario energy is very expensive. The second scenario examines what could be if the first scenario wasn't true, if travel was easy and cheap but that concerns about global disease spread would reduce travel.

In both these scenarios people don't like to travel and people don't like to meet strangers. Instead, because meetings are necessary for operation of the world's economy, people develop alternatives to face-to-face meetings. The alternatives are extensions of the advancement of core technologies.

The purpose of the first section of this paper is to extrapolate some basic parameters into the future in order to form a foundation for the predictions. Next, we briefly look at how human activity and behavior will change in response to the core technological advances. Finally, we will focus on the behavior in meetings and types of meetings possible in 2020.

WHAT IS SCIENCE DOING FOR US ?

A lot.

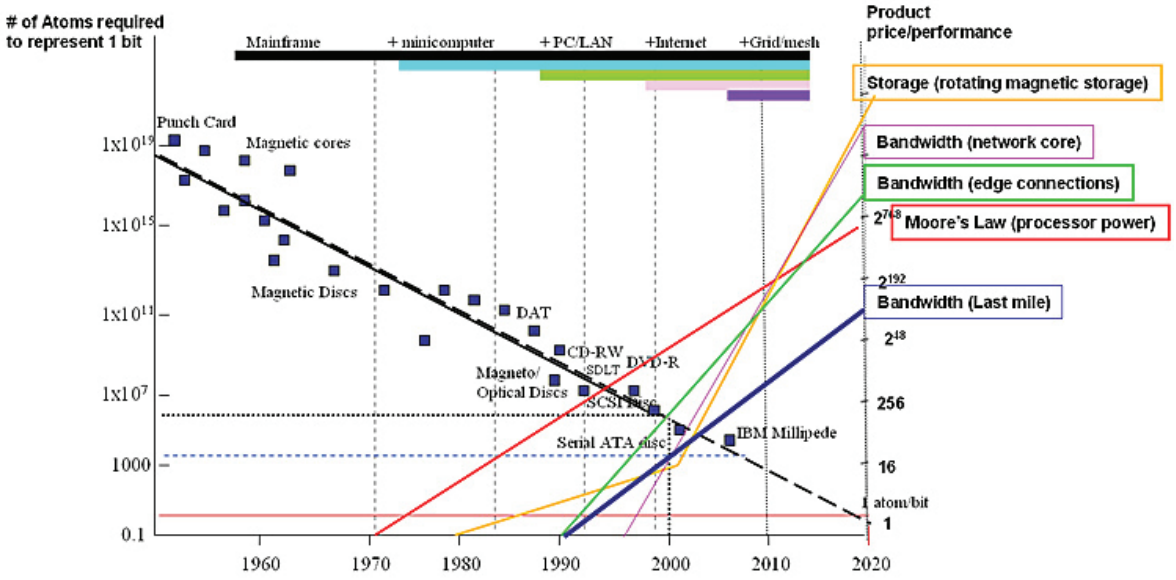


Figure 2: atomic requirements for storing one bit of data. Source: IBM Research and ParkWood Advisors LLC

Figure 2 is based on research from IBM and expanded upon by ParkWood Advisors with newer technologies and the price performance lines for processor power, storage and connectivity. This shows that in 1950, punch cards could store 80 bytes of data and it has been improving ever since. Very clearly there's been a log linear correlation to improvement in storage density such that sometime around 2020, we will be able to store one bit for every atom. This is good news for the people who build storage, but for anyone else it is disastrous because if you look at the trend in data storage requirements, by 2020, we need more atoms than we have. In reality, the problem is even greater because to be secure you must keep at least one copy of all information. And even though compression will help the situation, to keep the calculation simple, it is fair to predict that most information will be duplicated.

By 2020, there will be so much digital information that there will begin to be doubts about how to store it, but in fact, that will be heavily influenced by the stability and reliability of software. The problem of error-proof software is far from being solved. However, many engineers are and will continue to work to address the challenge and the present analyses assume that software reliability will be common place.

The other resource which the scenarios assume is abundant is bandwidth. The global bandwidth consumption up to 2005 is depicted in Figure 3.

As long as these measurements have been used for predicting the future bandwidth requirements, there have been errors. Regardless of this poor track record, however, the scenarios and models on which this paper is based assume that in 2020 growth in data traffic will exceed growth in economic activity. It is, in effect, Metcalf's law essentially applied to macroeconomics. The more humans connect things together, the more they will need to connect things together, and information needs to travel faster than atoms. So, although economists expect robust global GDP in the next decade and a half, the rates of increase in information traffic as a consequence will be much greater. In 2020, peak packet loading and the core Internet will be reaching around ten zitabytes per second.

As are result of these core technologies improving, we have lots of bandwidth, lots of reliable software running on low cost processing sites and we're getting all these new capabilities as a result: For example, network-connected digital sensors are everywhere.

Sensors are capturing information and software is processing its contents, continuously. Surveillance is going to become ubiquitous over the next decade and a half so in 2020 all public spaces are continuously monitored for safety and security reasons, a lot of private spaces probably as well—work spaces? Certainly. Domestic spaces? a little harder to judge, but very probably, especially if you want insurance at a reasonable price.

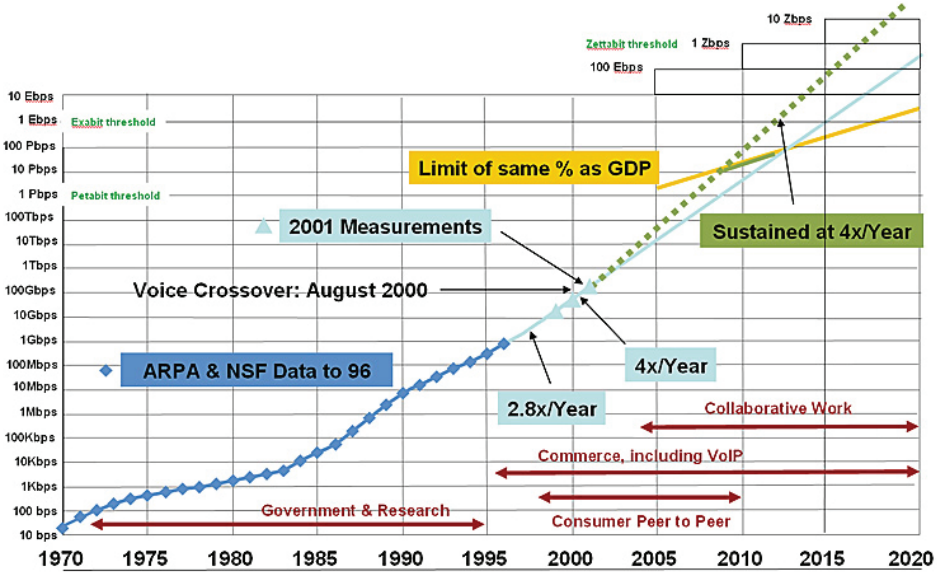


Figure 3: Global bandwidth consumption forecasts. Source: ParkWood Advisor LLC . Based on original graphic by Telegeography.inc, using sources such as ARPA and NSF up until 1996. Projections beyond 1996 by ParkWood based on IETF and ICANN estimates.

LEVARAGING AUTOMATION

Reliable software will be working overtime for us in 2020. Lightening the "weight" of information processing on the human, making more information processing digital, is one of the logical consequences of advances in core technologies which we have examined above.

Most people in 2020 receive so much information that they can't process it. They don't need (are not expected) to process it because software reads (or processes) the low priority pieces and files them or discards them based on user-defined settings. This is an example of what a semi-autonomous agent can do for its user. Another example is responding to inquiries, such as a ringing telephone. How often does the phone ring in 2020 and the person's agent answers? Frequently. As IP telephony progresses, a number is no longer tied to physical locality. Already today the most commonly asked question when speaking with someone by phone is "where are you?" but in 2020 the systems permit us to know more (and care about certain details less) in less time.

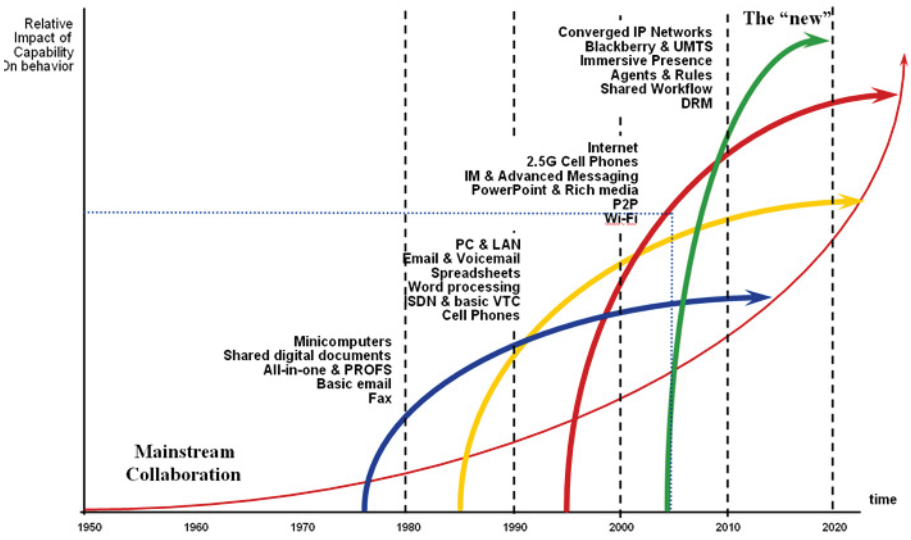


Figure 4: New Technologies change behavior in unpredictable ways

We have not been good predictors of how new technologies change human behavior, but we can hope that our ability to focus and collaborate will improve in the future.

One of the results of increased automation in the environment surrounding people is that the average person will focus on other tasks than the mundane.

No one will remember phone numbers or addresses. The phone, or whatever the personal computation and communication device will be called in 2020, is an extension of the owner/user's memory. In fact, the reliance on digital processing over bio-processing—having the brain store the information and do the work—will be such that the very nature of what we do will be completely different.

So this brings up some new challenges and opportunities. First, if we have a completely monitored world we can take the media from surveillance capture devices and we can use algorithms to "see"—to interpret—what is happening in the present or what happened in the past. Let's take an easy example.

Imagine algorithms that detect whether a person is being ambiguous or lying on purpose, like a lie detector works today. Some justice systems already have this technology and it will probably be in use everywhere in law enforcement before the end of this decade. Variants of these systems will be used in business settings as well. As a result of algorithms detecting non-verbal human communications—something we do subjectively already, will be performed objectively, based on statistics and machine learning, in the future—a lot of explanations and misunderstandings can be avoided. This will be particularly valuable in cross-cultural settings. In meetings involving people who say "yes" when they mean "no" and the opposite, many meeting "assistants" will in effect translate the original intended meaning in the appropriate signals for the user.

The second thing we can do with an abundance of processing cycles is to automatically simulate different future scenarios. There will be servers who will be dedicated to simulating everything that we are speaking about or thinking of doing in models. These servers will then compare the outcome of the models with what is actually going on and, by flagging discrepancies automatically, permit individuals, teams and companies to see whether the future is diverging from their expectations in fundamental ways.

Another manifestation of automatic, real-time processing is the emergence of virtual offices. The third fundamental transformation enabled by the surplus of processing and bandwidth is an improved geographical distribution of resources, human resources, in particular. Already today, as a result of corporate videoconferencing, people will be more likely to live where and how they want while maintaining or exceeding the levels of productivity they would have if sharing the same building or office. With High Definition videoconferencing this trend will become more and more attractive. Immersive environments, such as telepresence installations today, will be accessible to more people.

From the standpoint of risk management, distributing key resources geographically is a positive consequence of the emerging communications technologies. It also means that a company will have better access to local resources, filling needs, in real time, in local markets.

Automation will also reduce decision-making responsibilities of certain people on certain tasks. We can stop having to decide everything for ourselves. We can assign certain decisions to the software which is based on past preferences, as well as the forecasting abilities mentioned above.

PAYING ATTENTION

One of the consequences of automation is that humans will be able to focus greater attention on things that cannot be delegated to software. Some view what we currently do in meetings as augmented attention. People sit in meetings, on conference calls, with IM windows open, doing e-mail, surfing the Internet and whatever else people choose to do when others can't see them and think they have the attention of others. More and more people realize that they are distributing their attention across many more diverse objects or conversations. Figure 5 shows what a future social network might appear when plotted.

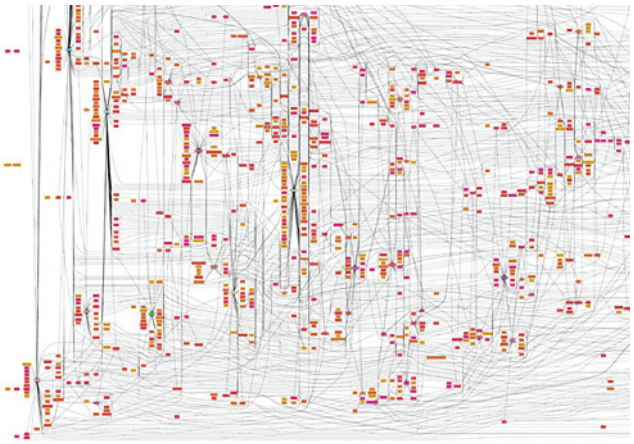


Figure 5. Social Networking Map in 2020

Youth in 2008 are already doing this as a part of their natural social existence. In the future, attention will be one of the most valuable human resources available and people with the skills to become the focus of attention, or to distribute attention very well will be recognized for their achievements.

Instead of paying two seconds of undivided human attention to determining which room is available for a meeting, a software application will designate a meeting place convenient for all those in the same building. For those who are participating in remote meetings—the majority of people in certain domains—the challenge will be designating or designing the best virtual meeting spaces.

A significant amount of attention could be given to the selection of virtual meeting "places." Imagine a meeting place that would look exactly like a normal meeting room with tables and chairs, but everyone could feel as if they are sitting in the front row. In the acoustic domain, the audio level of a space could be made uniform so everybody can hear the speaker all the time. This is already possible in some high-end meeting facilities. Now let's apply automation to this meeting and perhaps the participants can speak in their native language and hear others in any language they choose, making it easier to pay attention.

MEETING IN 2020

Given all of this background and these processing and bandwidth resources, what will meetings be like?

The first attribute which we can be relatively certain about is the physicality of people in a meeting. It would cost too much to get all the participants in a meeting together in the same place under the scenarios used for this analysis of the future. It would be too expensive in terms of fuel or it would be dangerous. So very, very few people will meet in person—in vivo—in 2020.

Most meetings will be conducted in virtual. The environments could still resemble one another, to reduce the distractions but the participants will also have the possibility of sending digital representations of themselves. A whole new set of rules around how to prepare for meetings and participate virtually in meetings will need to be developed.

Let's examine the flexible parameters around future meeting "places." Some meetings will be conducted on the dark side of the moon. All we need is the telemetry data to build the model. Some might want to meet on the death star from StarWars. All they need is George Lucas's permission to use the intellectual property that the representation of that space requires. This raises some interesting issues of intellectual property. Who owns the representation of the Taj Mahal or the White House? If a model can be generated of an existing physical space, and if that model can be digitally represent and sold as a place to meet, then who owns it?

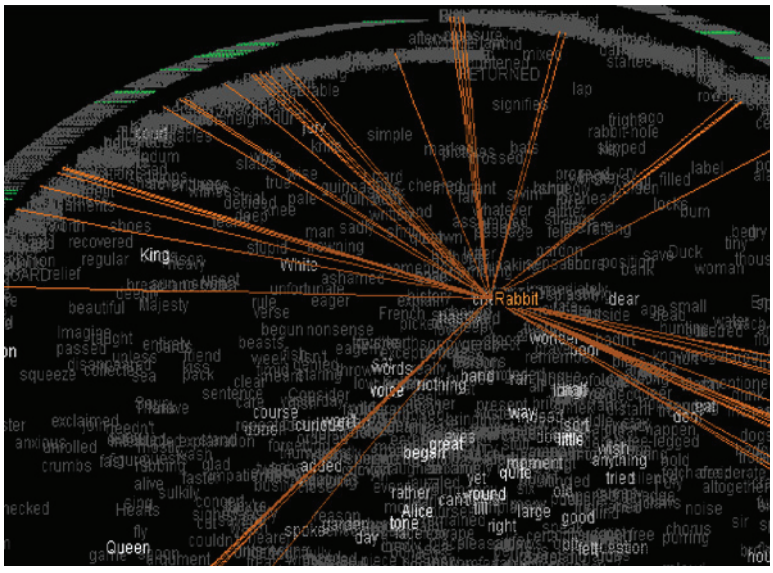


Figure 6. Lexical Analysis of Alice in Wonderland.

Similarly, when people will attend meetings in 2020 they might choose a persona different from their own. A persona can be constructed digitally that would represent a participant and behave as if they were that person. And one of the problems that arise from this advanced capability is being able to authenticate the people who join a meeting. This is not a matter of a user ID and password. How will people know who they are meeting with, their level of attention and their rights and responsibilities?

People who appear to be participating in a meeting and paying attention, might in fact be listening to their digital music and doing IM and whatever else they do in meetings even though they will appear fully engaged. Conversely, if an agent is attending a meeting on behalf of a participant who is completely absent, how much should other people know about that person's background, past positions on issues and how far from the originally planned agenda can the meeting diverge? When a meeting participant assumes an action item, or influences a decision, who is responsible, the person or the company who provided the agent?

Another necessary precursor to a meeting in which important decisions will be made is to make sure the right people participate. Clearly, new authentication etiquette will need to be developed to clarify these grey areas of who is participating fully, partially via an avatar and who has only sent an agent. There will need to be some very strong out-of-band, key exchange protocols to have the encryption that is going to make these kinds of meetings possible.

After a meeting participants will be able to replay all or parts of a meeting, at least until the planet runs out of digital storage capacity. Everything that goes on in meetings will be subject to post-meeting analysis, post real-time meeting progress analysis, and a much closer scrutiny than we have today.

For example, there will be software to measure the conversation that is happening in the meeting. An example is the lexical analyzer in Figure 6. It is showing the relative strength of association between the keywords that are being used during the course of the meeting and then mapping these out in a hyperbolic tree that shows how far away from the decision words the people's words are.

Meetings in 2020 will not be low cost even though they will become virtual because they consume processing and bandwidth resources. In order to optimize resources in 2020 people will also do different things between meetings.

Meetings themselves will, especially with important people and on important topics, be prefaced by a great deal of preparation. Simulations, data gathering and organization, agenda management, profiling of the other participants, expected outcome analysis, interaction analysis models that will all be preloaded so that during the meeting, the participants are attentive but in the background there is software comparing what is happening during the meeting with the expectations. For example, participants will be able to compare their actual amount of time on a topic with their projected or model agenda.

Figure 7. Visualizing information sets, decisions and outcomes.

MEETING PATTERNS

It's already pretty clear, with the behavioral analysis that has been done on meetings, that most people don't meet very effectively. Humans can concentrate on something for about twenty minutes and then whether they like it or not, their attention drops off. Even in life threatening situations, that is a pretty good generalization.

Meetings could be designed and managed to meet this human limitation. People may concentrate all their attention and work hard for a short period, then slow down and exit a meeting, then they might return to a higher state of attention and focus, then slow down; this behavior is sometimes called "sprint and glide" by psychologists. There are other names for this pattern in various meeting formats but regardless of the name, such entry and exit patterns will be more common place when people can also "catch up" with a meeting in progress. Being an intermittent meeting participant may not seem productive today, however, the 2020 some form of sprint and glide will be the norm.

One variation on this theme is that of having different meeting types. There are different interactions modalities that are associated with those different purposes and it's not clear if people are most productive when these are mixed in one meeting. It is not difficult to justify having one meeting to organize the "confirming" activities and separate out the "deciding" activities for another meeting.

This also aligns well the change in attention focus. In 2020 people will be more prone to getting interrupted because meeting assistants or agents are continuously introducing new information to individual participants in the background; what is the protocol for interrupting someone when they are in a virtual meeting via presence: does one replace use of an avatar? How do participants make it clear that they aren't paying attention anymore? Then workplace psychologists will have to reinvent all of the rules and ritual of meetings because we aren't doing them the way that humans evolved to do them anymore. It's going to take a while to work these new rituals out and it will go far beyond what we can imagine today.

Here is an approach to representing complex data structural navigation problems. So if you're trying to figure out how to do a merger between two companies you want to know where the information environments, the social, and asset environments mesh together. How do you do that? That is an impossibly complex representation problem in data visualization today, but it won't be in 2020. We will have routine tools that we will have learned to interpret with the help of software to make that work.

Figure 8 is a model of how a meeting is supposed to go: it's a plot of what the sub-themes of the meeting are supposed to look like and a trace of how the timeline is moving through those themes: are you getting anything done? That is what it is there to detect. So we will have a lot of these tools and they'll all be around for everyone to use.

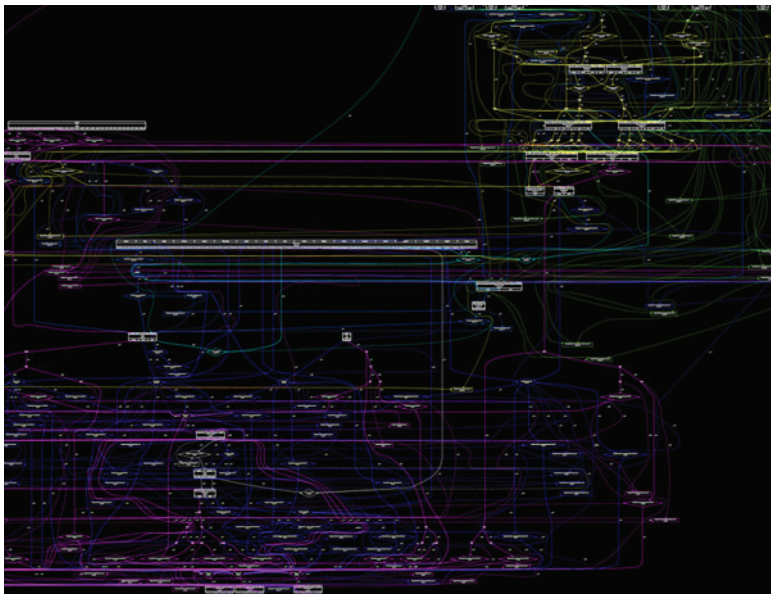


Figure 8: Managing Meeting Models, Time Lines and Interaction Effectiveness

BEYOND 2020

Work on most of the challenges we have discussed so far began in the 20th century and is being stabilized and commercialized in the first two decades of this century. Looking out ahead, what will be the 21st century's biggest contribution to the human productivity? One area we have highlighted here and which will certainly merit a great deal of investment is attention management.

How do you make me able to function more than 24 hours a day, more than seven days a week, more than 365 days a year? Humans still only have 24 hours a day no matter how many time zones they cross physically and virtually. If in the second half of this century we can make more tools that permit people to function at more than a one to one ratio with real time, human productivity will be measured in completely different metrics. These are the metrics futurists will discuss in presentations and papers prepared in 2020. Between now and then we will need to focus our attention on what we already have but have yet to use!

LINKS

For more information, visit the AMI Consortium web site:
www.ami-consortium.org

For a white paper describing the many applications for AMI, download and read:
www.amiproject.org/pdf/Applications-for-AMI-Technologies.pdf

For an overview of the future research directions, download and read:
www.amiproject.org/pdf/AMI-overview-prospects-for-future-research-Jan2006.pdf

For technical backgrounders about AMIDA research areas, download and read:
www.amiproject.org/pdf/SOTA-Annotation-and-Query-Jan2006.pdf
www.amiproject.org/pdf/SOTA-Conversational-multiparty-ASR-using-remote-mics.pdf
www.amiproject.org/pdf/SOTA-Focus-of-Attention-Jan2006.pdf
www.amiproject.org/pdf/SOTA-Localization-and-Tracking-Jan2006.pdf

AMI c/o IDIAP Research Institute
Centre du Parc
Av. des Prés-Beudin 20
P.O. Box 592
1920 Martigny - Switzerland
info@ami-consortium.org
www.ami-consortium.org

Tel +41 27 721 77 11
Fax +41 27 721 77 12

Project directors
Dr. Hervé Bourlard
IDIAP Research Institute
herve.bourlard@idiap.ch

Dr. Steve Renals
University of Edinburgh
s.renals@ed.ac.uk

Technology Transfer
Christine Perey
c/o IDIAP Research Institute
cperey@ami-consortium.org

The New Era of Business Meetings

Making time
between and during meetings
more effective

Agenda

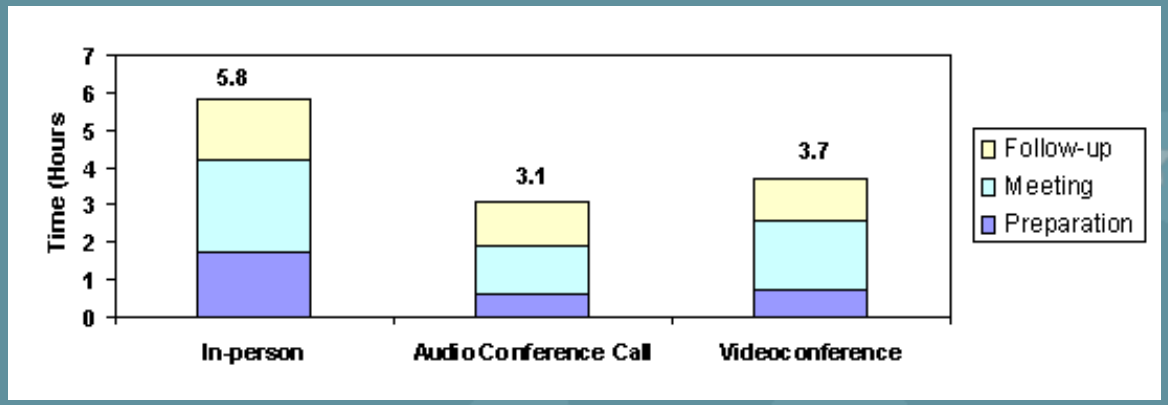
- Business meetings
- Our vision
- AMI-based features (modules)
- AMI Partnerships
- Conclusion

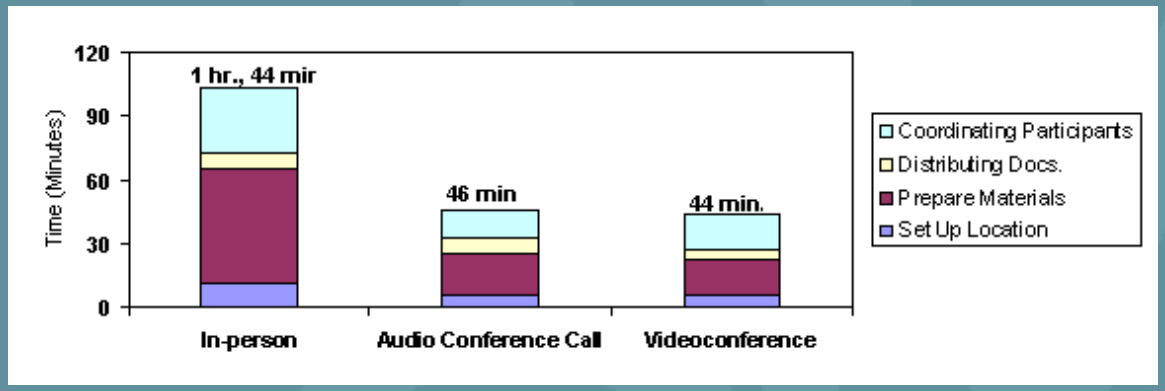
Business Meeting

- >12 million business meetings daily
- 37% of employee time spent in meetings
- 62 meetings/month
  - **15% involving travel, audio conferences, or videoconferences**
  - **85% internal or local face-to-face meetings**
- Can't attend all the meetings scheduled

http://e-meetings.mci.com/meetingsinamerica/uswhitepaper.php

Time and Meetings





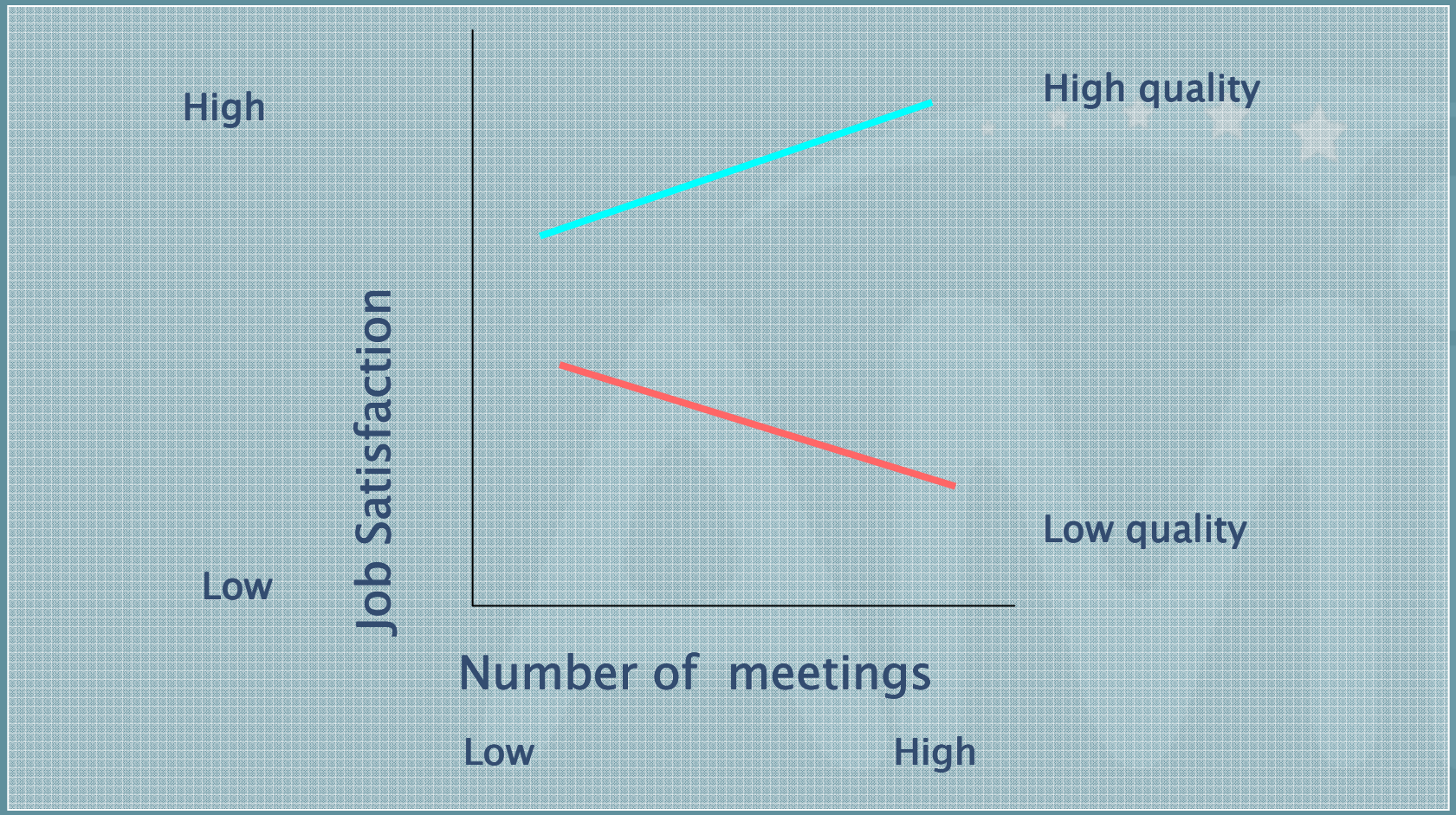
http://e-meetings.mci.com/meetingsinamerica/uswhitepaper.php

Participant's Suggestions

- Provide an agenda; improve organization/structure of meetings
- Set goals/objectives to achieve
- Attend only if information is relevant to the person
- Pay attention to timing: time limit; make meetings shorter; start/end on time
- Reduce the number of meetings held; meet only when necessary
- Use a facilitator (chair); rotate the facilitator
- Make meeting environment more comfortable

- Stayed focused on the topic
- Prioritize items and allow time for each
- Record and distribute minutes
- Clarify plan of action and anticipated outcomes; set deadlines for assigned tasks
- Delegate responsibilities
- Follow up with proposed solutions
- Provide training on how to conduct meetings
- Allow time to prepare for meetings

Meeting Quality



What is AMI Consortium

- an 11-member multi-disciplinary consortium dedicated to the development of technologies that will enhance multiparty interactions
- Partners in the AMI Consortium conduct research on human-to-human communications, particularly during business meetings between co-located and remote participants

AMI Consortium Projects

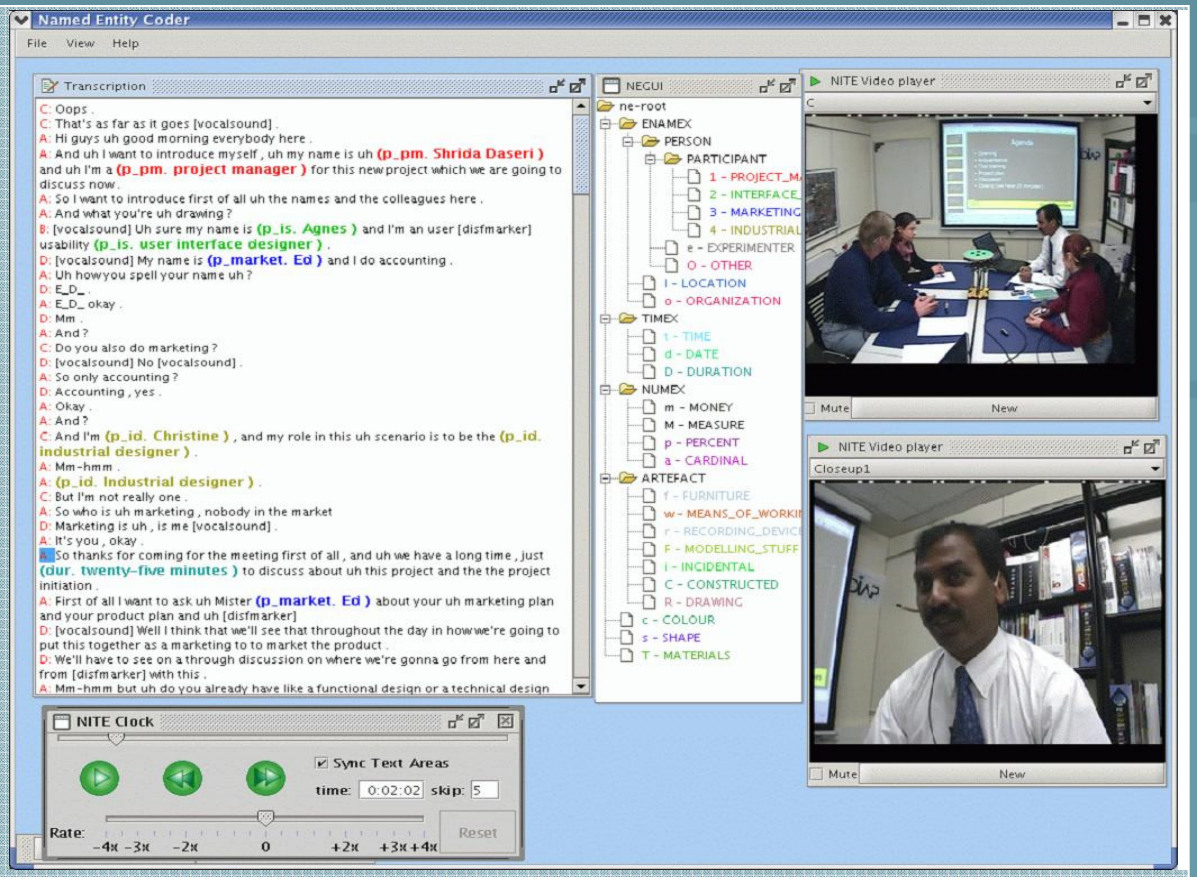
- Augmented Multiparty Interaction (AMI)
  - Jan 2004–Dec 2006
  - Achieved all objectives
  - MMM Database
- AMIDA= AMI+Distance Access
  - Jan 2007–June 2009

Capture for Research

- Instrumented Meeting Rooms at IDIAP, UEDIN and TNO
  - Multiple synchronized media
    - Close-up and room view cameras
    - Close-talking and far-field audio
    - Whiteboard, PowerPoint and I/O digital pen interaction


Annotate

- NITE XML Toolkit



Processing technology

Visual Tracking

Face orientation

Visual Tracking

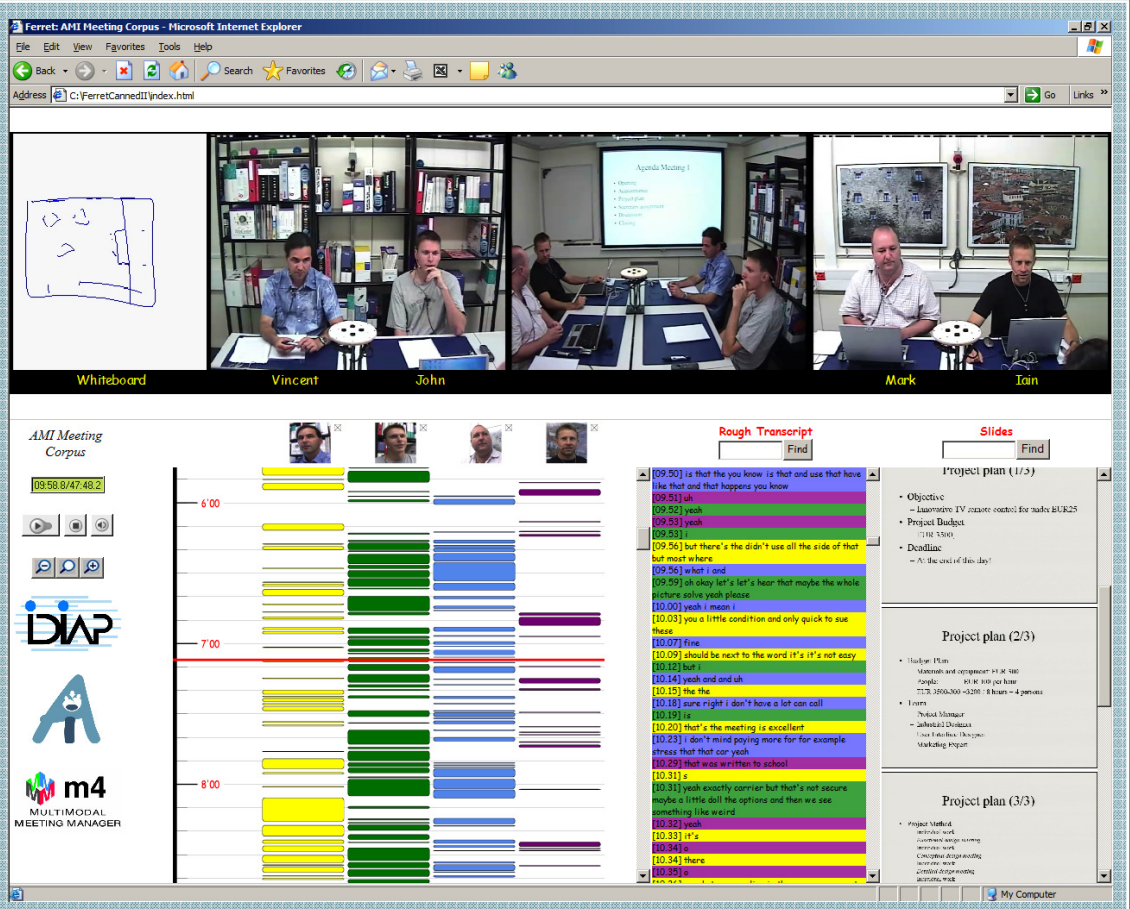
Look who is talking

Content, events and patterns

- Who is in the meeting?
- What do the participants say?
- When and how do they communicate?
- What are they doing?
- What are they looking at?

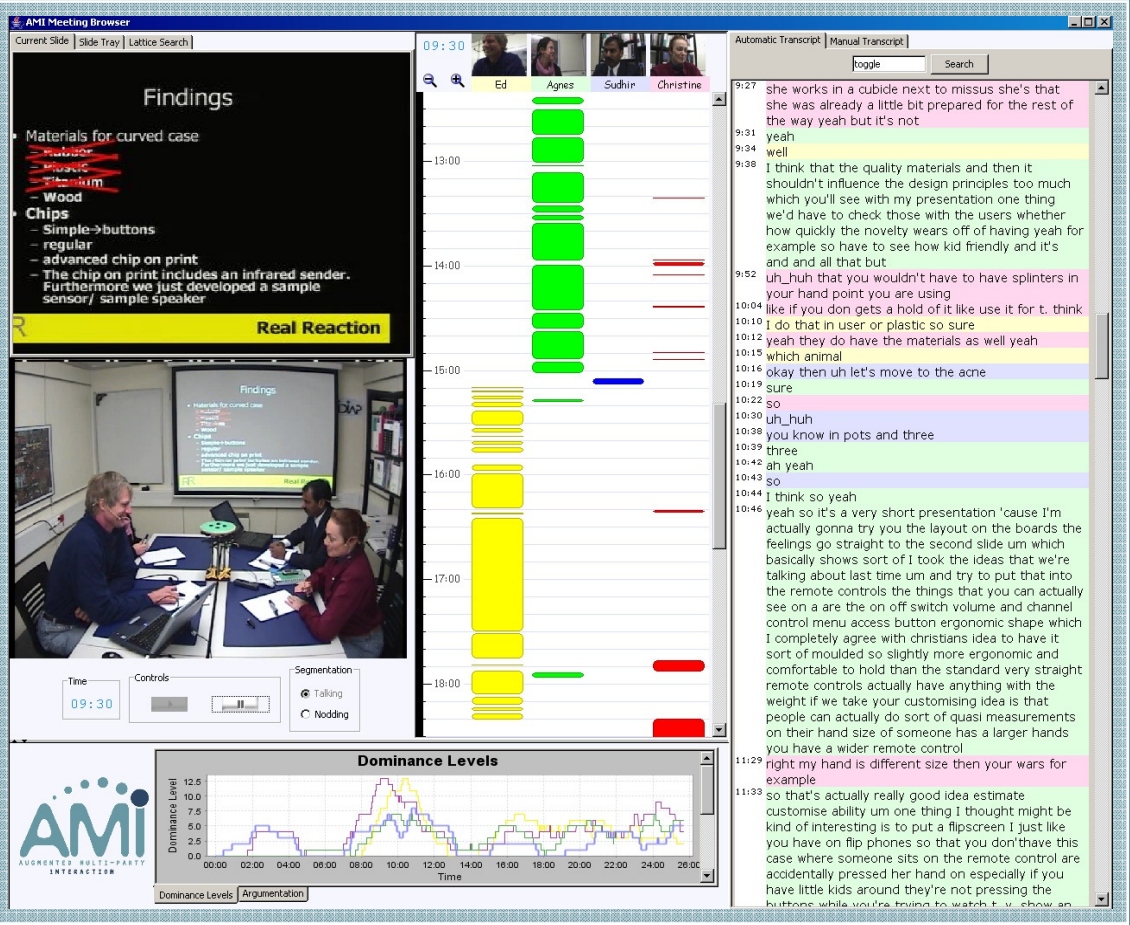
Browsing meetings

- Navigate
- Search
- Skip



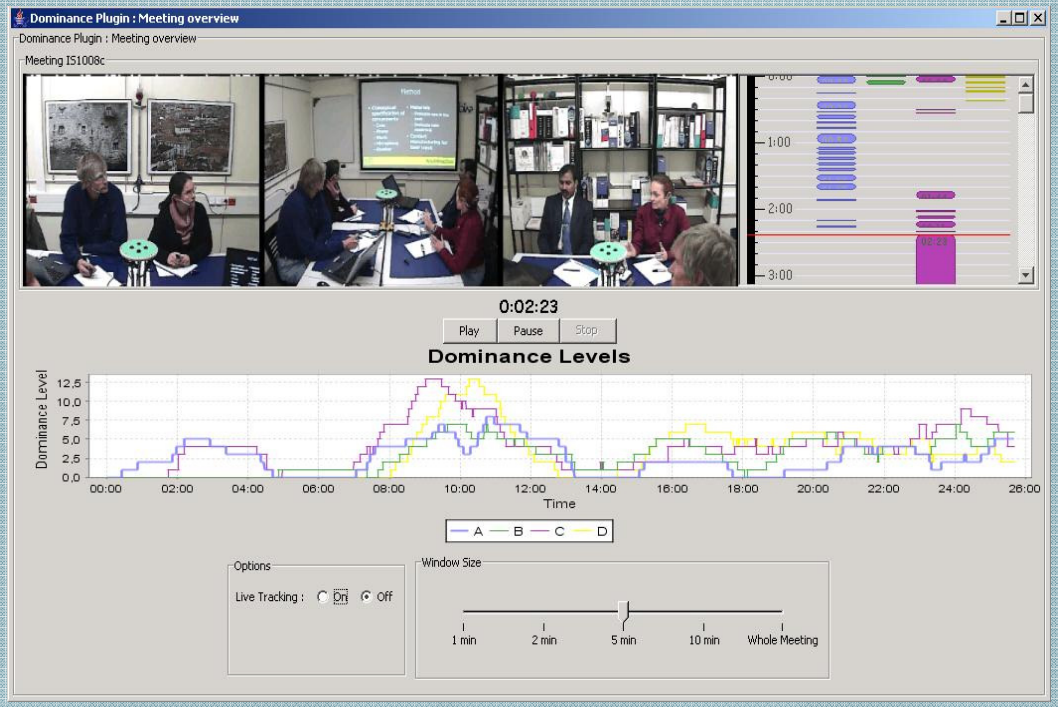
Browsing meetings

- Show
  - Slides
  - Media
  - Transcript
  - Analytics



Recognize non-verbal

- Emotions
- Gestures
- Actions
- Dominance



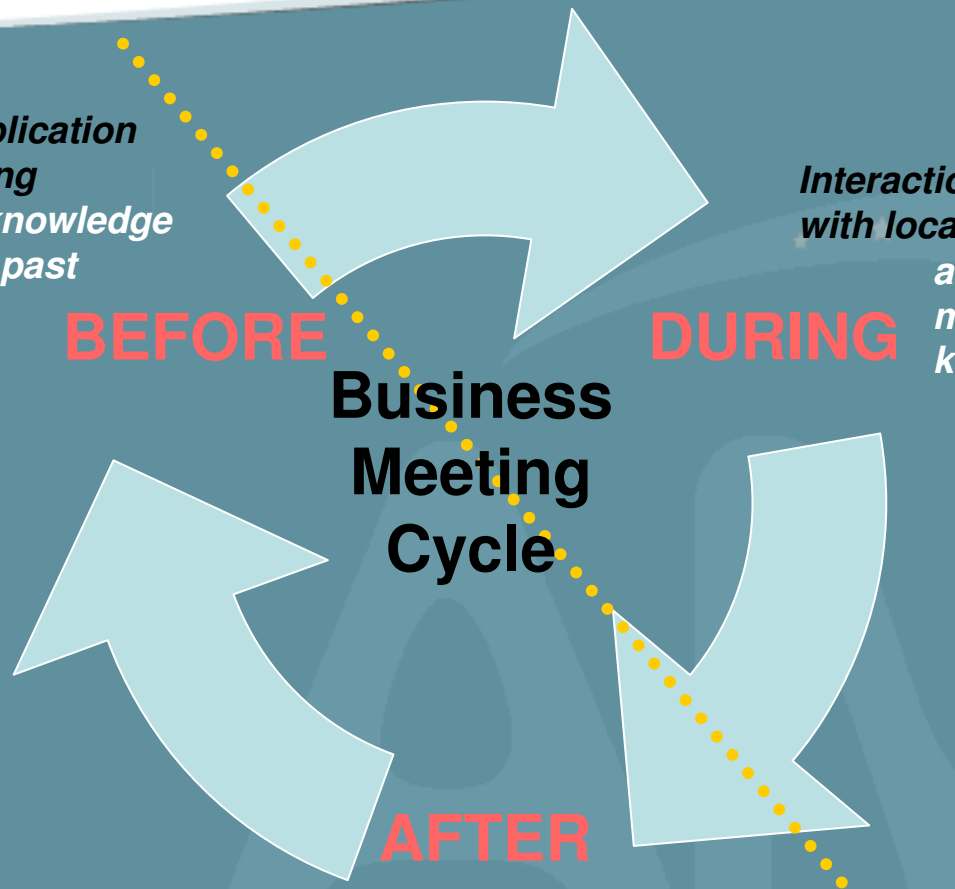
Syntheses & Analyses

- What topics are discussed and when?
- What decisions are made and by whom?
- What roles do the participants play?
- What positions do they take on issues?
- What activities are completed?
- What tasks are assigned or reported done?

Part II

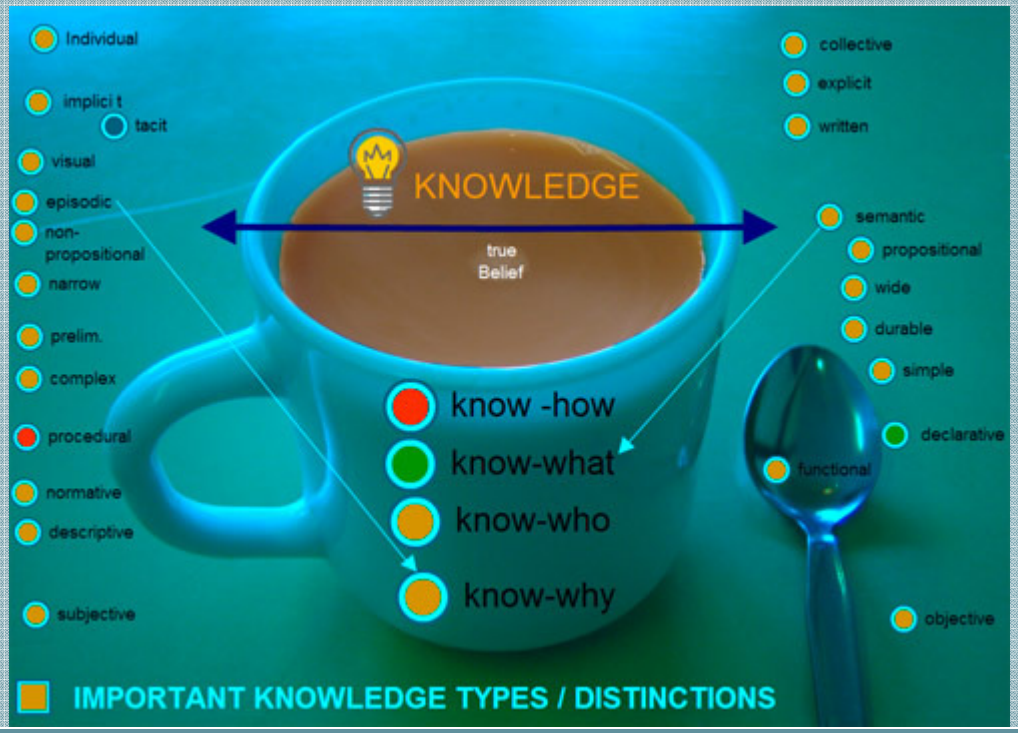
AMI Vision for Meetings

Ami Vision for Meetings



Vision

The Content of Meetings is part of the Corporate Knowledge repository



Architectural Scenario

Financial Services Scenario

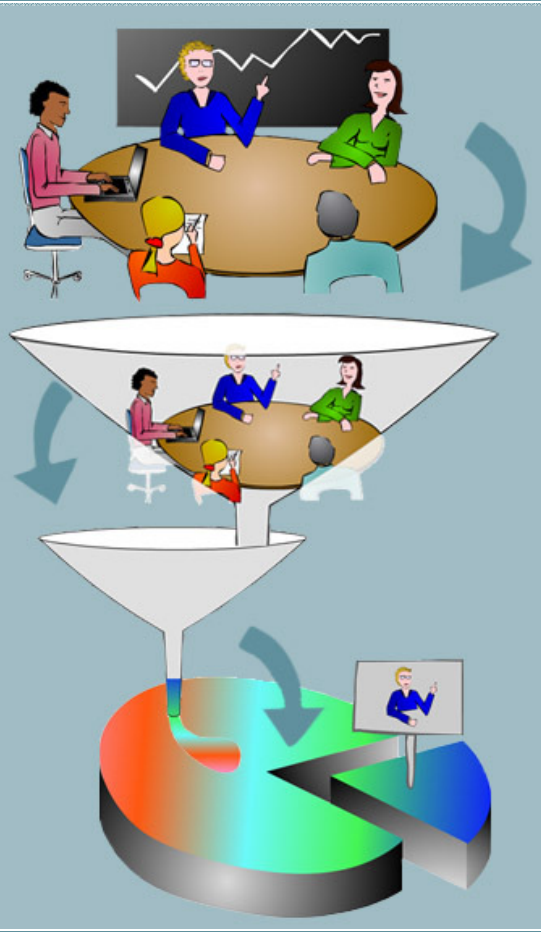
What if?

Part III

Integrated AMI-enabled systems

AMI Meeting Summarizer

- Between Meetings
  - Understand meetings missed
  - Find patterns in meetings
  - What people ask for most often
- During Meeting
  - Find relevant info
  - Manage an on-going meeting
  - Catch up

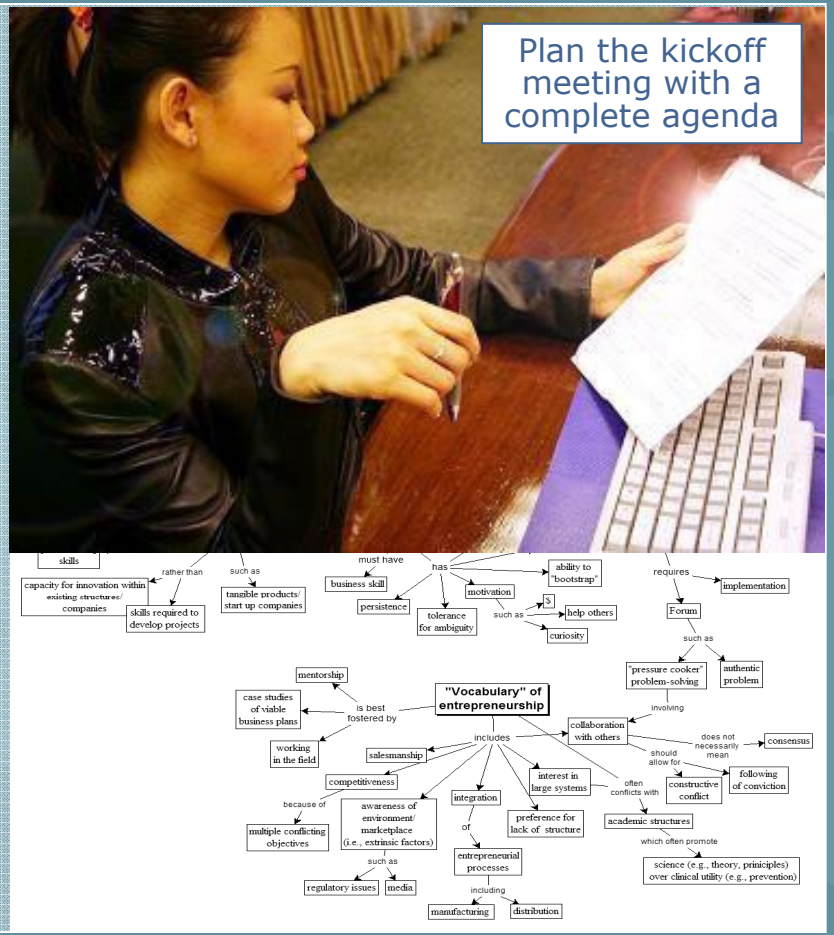


AMI Meeting Accelerator

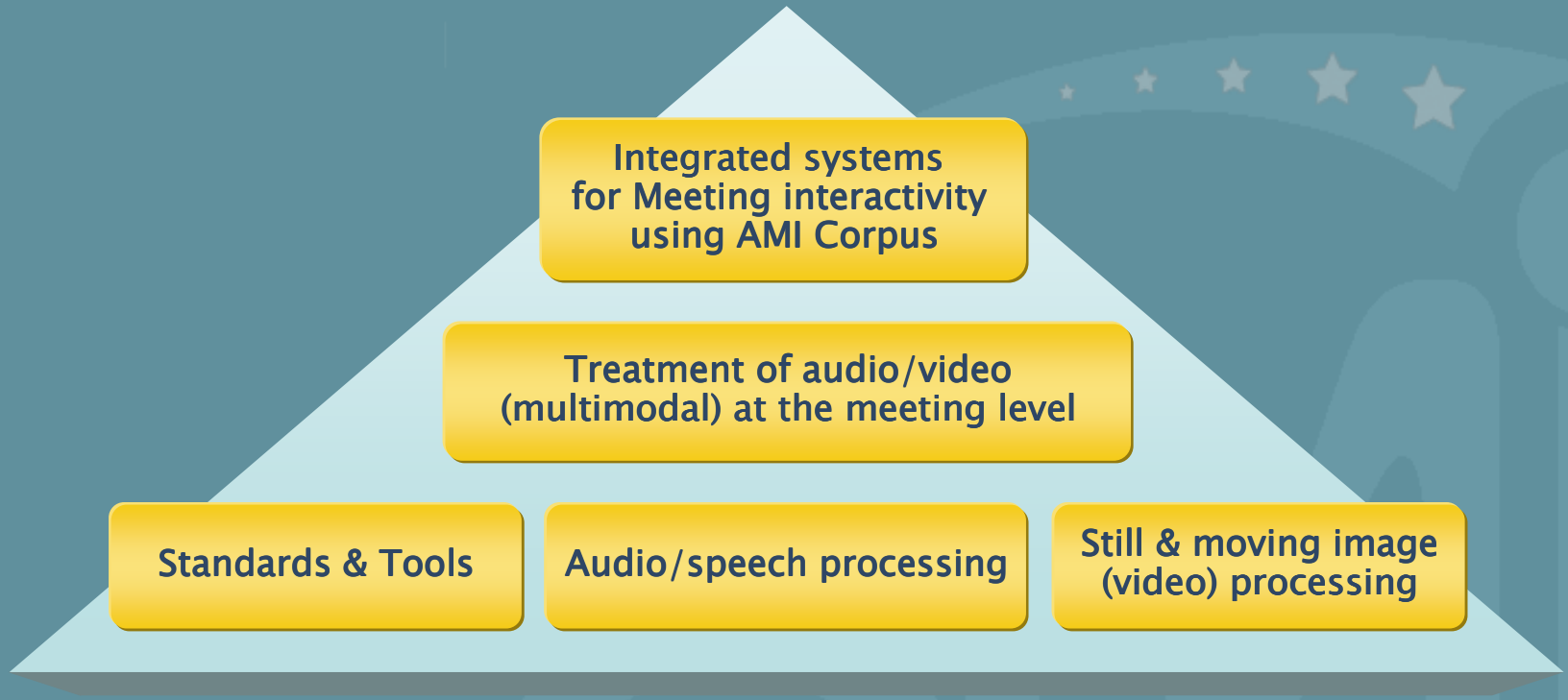
- Between Meetings, it accelerates playback
  - To get all the nuances
  - To review a segment of a meeting
- All or a subset of meeting participants
- Skip to next

AMI Agenda Manager

- Between Meetings
  - Compare upcoming with past
  - Be realistic
- During Meeting
  - Stick to the agenda
  - Receive reminders
  - Finish the agenda



AMI Building Blocks



Core Technologies

Part IV

Partnerships

AMI Consortium partners








Community of Interest

COI Workshop

- Two day face-to-face event
- Purposes
  - **Establish bonds**
  - **Listen to industry (vendor) needs**
  - **Launch the Mini-Projects program**
- Involve 50 people
- September 11-13, 2007

COI Mini-Projects

- Applied research or prototyping focusing on area of mutual interest
- Project partners
  - 1 AMI Consortium partner
  - 1 COI member
- Funds
  - AMI funds work of its scientists
  - COI member funds equal effort to be executed by COI R&D staff

AMI Showrooms

- Complete suite of demonstrations
- Interactive, linked to AMI MM database
- Permit visualization of AMI research
- Work in progress, reflecting evolution

Thank you

Any Questions ?