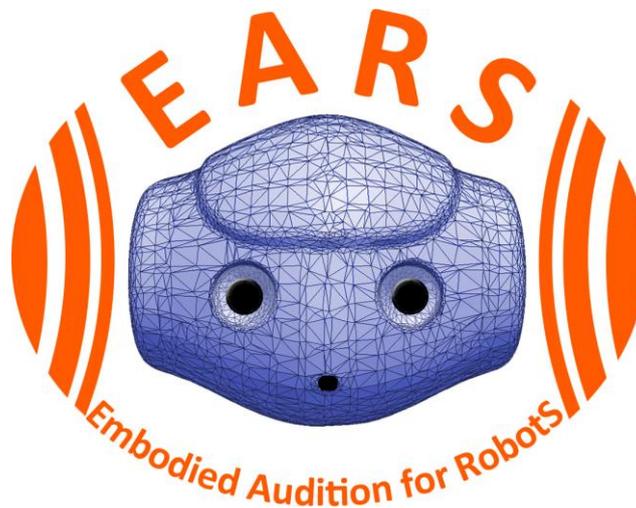




# EARS



## Minutes of SB Meeting June 2014

Collaborative Project - EARS			
Editor(s)	Heinrich Löllmann, Heidi Zinser, Walter Kellermann		
Responsible Partner	Friedrich-Alexander-Universität		
Status-Version:	Revision B – final		
Date:	17/07/2014		
EC Distribution:	Confidential		
Document History			
Rev.	Issue Date	Description of Change	Author
A	08/07/2014	Draft - First Issue	H. Löllmann
B		Final version	H. Löllmann, H. Zinser, C. Evers, W. Kellermann
C	Date	Short description	

Revision	Summary of Changes	
	Reference	Description
A	All sections	Draft
B	All sections	Final version (inclusion of amendments and corrections)

Grant Agreement Number:	609465
Project Full Title:	Embodied Audition for RobotS
Project Acronym:	EARS
Call (part) identifier	FP7-ICT-2013-10

Type of Document:	Minutes of EARS SB Meeting June 2014
Date of Delivery to the EC:	-

Work package responsible for the document:	WP7 - Management
Editor(s):	Henrich Löllmann, Heidi Zinser and Walter Kellermann
Contributor(s):	All Partners
Reviewer(s):	All Partners
Approved by:	All Partners

**Venue:**

The Gabor Seminar Room (Level 6)  
Department of Electrical and Electronic Engineering  
Imperial College  
Exhibition Road  
London SW7 2AZ

**Participants:** (see Appendix No.1)

Walter Kellermann  
Heinrich Loellmann  
Heidi Zinser  
Antoine Deleforge  
Patrick Naylor  
Christine Evers  
Alastair Moore  
Boaz Rafaely  
Jonathan Scheaffer  
Vladimir Tourbabin  
Verena Hafner  
Guido Schillaci  
Sasa Bodiroza  
Radu Horaud  
Israel Gebru  
Dionyssos Kounades-Bastian  
Israel Dejene Gebru

**Thursday, June 26<sup>th</sup>, 2014****Opening** (9:00 – 9:15)

A short welcome speech was held by the host Patrick Naylor (IMPERIAL) and the Project Coordinator Walter Kellermann (FAU).

**Report and Discussion on WP 2: Acoustic Scene Analysis** (9:15-10:45)

Presentation on WP 2 was given by Patrick Naylor, Christine Evers, Alastair Moore and Antoine Deleforge (see Appendix No.2 and 3).

Regarding Tasks and MS3, Patrick will prepare an internal report.

A joint paper of IMPERIAL and FAU is planned for the next EUSIPCO.

A common data structure for acoustic maps, which may also include video data, is needed for T2.1. The development of this data structure should be done by a working group including Christine Evers, Jonathan Sheaffer, Antoine Deleforge, and Hendrik Barfuss.

Alaistar Moore provided a demo with the Nao robot and an Eigenmic where the robot localises a person speaking to him and turns its head towards the speaker.

**Report and Discussion on WP 1: Embodied Acoustic Sensing for Real-world Environments**  
(10:45 – 12:20; Tea break 11:05-11:25)

Presentation on WP 1 was given by Vladimir Tourbabin, Jonathan Sheaffer and Boaz Rafaely (see Appendix No.4).

Discussion about the head microphones (cf., slide no. 11): The new Nao head contains two additional mics, where the installation of more mics is expensive since the electronic system must be modified considerably. However, for a final EARS prototype, the number of microphones should not be limited to that of a direct "marketable" prototype.

The work of FAU conducted in WP 1 was presented by Antoine Deleforge.

FAU discovered by means of measurements with chirp signals severe nonlinearities for the loudspeakers of Nao and artefacts for the recordings with the head microphones. ALD was not aware of these problems and will investigate these effects.

Vladimir presented the results of his paper published at the last HSCMA (see Appendix No.2).

Boaz presented the results of the collaboration between BGU and INRIA on source localization (**Presentation missing?**).

The simulated HRTFs of BGU should be used mainly for the microphone array design and to guide the microphone placing.

An interesting aspect is to substantiate the correlation between the effective rank of the GHRTF matrix and the localisation performance by means of real measurement data.

Radu gave a presentation on sound source localisation for audio-visual data alignment.

### **Report and Discussion on WP 3: Audio-Visual Disambiguation (12:20-12:50)**

Israel Dejene Gebru gave a technical presentation about the work conducted in WP 3 (**see Appendix No.5**).

### **Lunch break (13:00-14:10)**

### **Report and Discussion on WP 4: Robot Embodied Cognition and Interaction (14:10 – 15:50)**

A presentation on WP 4 was given by Verena Hafner, Guido Schillaci and Sasa Bodiroza (see Appendix No.6).

The representation of acoustic maps by means of MFCCs was identified as an interesting aspect for further work.

The robot gesturing is important to support the performance of robot audition. The robot might indicate that the ASR has failed by gestures motivating the speaker to repeat his commands or to speak louder.

Gregory Rump gave a presentation on the dialog engine of Nao (see Appendix No.7). He also gave a demo of the Nao with Choregraph exemplifying some programmed dialogue scenarios.

The incorporation of other speech dialogue data bases was identified as an interesting aspect. This could avoid the redoing of some work on developing and defining speech dialogues on one hand and show the generality of the developed approach on the other hand. Rodolphe pointed out that the presented system uses knowledge on speech dialogues developed by ALD and is not from a publicly available domain.

### **SB Meeting (SB members only) with report on WP 7: Management (16:15-18:50)**

Heidi Zinser gave a presentation on WP 7 (see Appendix No.8).

Proposal for a 1<sup>st</sup> Supplementary Agreement to the Consortium Agreement was distributed as handout to all Partners. All Partners agreed to this amendment subject to the approval by their respective legal departments.

The Project Officer will invite the Partners to a first project technical review meeting in March/April 2015. Half of the day will be reserved for presentations and the other half may be reserved for demos.

The Partners wish to have a meeting early in March 2015. The meeting should take place at INRIA in Grenoble since demos with the Nao should be shown.

The Coordinator may propose technical reviewers to the EU. The following persons were suggested:

- Peter Svensson (NTNU, Trondheim)
- Gordon Cheng (TU Munich)
- Tomohiro Nakatani (NTT Labs Japan) who might be too costly for the EU.
- Gerhard Sagerer (Universtiy of Bielefeld)
- Gerhard Schmidt (CA University of Kiel).

It was agreed to finalize this list at the next SB meeting in July. In preparation of that meeting, further proposals should be exchanged and discussed to obtain a more substantiated list. Suggestions of who to invite as speaker (preferably a member of the SUAB) at the annual project meeting in Erlangen (Dec. 10-12, 2014) were discussed and the following names were mentioned

- Gary Elko (mh acoustics)
- Tim Haulick (Nuance)
- Gerhard Sagerer (if not named as reviewer).

It was agreed to collect more proposals where it has to be taken into account that there is no travel budget for any reimbursements of costs.

Heidi Zinser reminded the Partners that costs cannot be claimed beyond the project duration and the budget for personal costs cannot be transferred to the budget for travel. It was pointed out that according to the new guidelines from the EC, restrictions have been implemented on extension of the duration of grant agreements.

It was agreed that the project manager will no longer collect and control the Partner's time-sheets. Instead, each Partner must report all of their personnel efforts against the various deliverables worked on, per WP, by each person on a biannual basis.

It was pointed out that attendance of conferences/meetings can only be considered under dissemination activities if the presented poster / paper clearly acknowledges the EARS project and EC funding under the 7<sup>th</sup> Framework Programme (using the mandatory acknowledgement statement as given in the QM-Plan).

A bi-annual SB progress report has to be provided to the EU. A template for that will be send to all Partners by FAU. Each Partner will be required to complete their sections of the report. The content of these biannual reports shall be later merged into the annual periodic report to avoid unnecessary additional work.

It was agreed to promote the exchange of PhD students and post-docs, e.g., for small workshops ("code camps").

Heidi Zinser presented a list with open issues that were raised in the previous kick-off meeting and SB meetings, respectively. This document will be sent to all Partners with the request to check and address the open points.

---

---

## **Friday, June 27th, 2014**

### **Report and Discussion on WP 5: System Integration and Validation (9:00-10:30)**

Presentation of Rodolphe Gelin gave a presentation on WP 5 and Radu Horaud gave a presentation on speech dialogue systems (see Appendix No.9, 10 and 11).

The Nao is able to perform language identification. In addition, it would be beneficial if the ASR could also provide a confidence measure for the recognition of phrases.

A keyword spotting should be provided by Nao to enable source localization and “selective attention”.

Gregory Rump gave a presentation on the Modularity framework (see Appendix No.12).

An important open issue is the synchronisation of microphone and loudspeaker signals. The internal microphone signals and loudspeaker signals are synchronised. However, these signals are not synchronised with the video data. A second problem is that the internal microphone signals need to be synchronised with external microphone signals.

One possible solution might be to provide all analogue signals to an external multi-channel sound card, which performs all A/D and D/A conversions. This would also allow to perform all audio signal processing on an external PC.

The issue of signal synchronisation will be tackled by a group consisting of Gregory Rump, Antoine Deleforge, Jonathan Sheaffer, Alastair Moore and Soraya Arias. A solution should be implemented until mid-October at the latest. For this, Skype meetings and workshops should take place to address this issue. New algorithms for the Nao should be developed primarily on an external computer and not on the Nao processor. The Modularity framework allows that the digital signal processing is performed by an external PC or laptop. ALD (Gregory Rump) will provide a test framework as well as a tutorial for this.

### **Report and Discussion on WP 6: Dissemination and Exploitation (10:30-11:10)**

Heiner Loellmann gave a presentation on the EARS homepage and the newly implemented Redmine software and document exchange platform (see Appendix No.13). A first completed version of the EARS homepage and Redmine platform has been finished as part of MS2 due at the end of June.

All EARS co-workers were asked to provide input for the homepage like papers and presentation published within the EARS project etc.

Radu pointed out the possibility to place published and unpublished papers on an archive (namely, arxiv.org, Cornell University Library).

(Coffee Break & discussion time: 11:00 – 11:45)

### **Technical Presentations (11:50-14:15, including a lunch break of 45min)**

Antoine Deleforge gave a presentation about the recordings performed with the Nao robot (see Appendix No.14). The measured data and its documentation will be soon made available by the EARS homepage to all EARS co-workers.

Jonathan Sheaffer gave a presentation on the Matlab EARO format to store the position of objects. An interesting aspect, raised by Alastair, would be to extend this data structure to account for varying source positions, e.g., by a string of EARO data cells (see Appendix No.15)

Verena Hafner led a discussion about robot behaviours, which was supported by a presentation of Sasa Bodiroza (see Appendix No.6).

The gesturing of the robot is important to convey a natural (human) robot behaviour. In addition, the gesturing can indicate to the interacting person that the robot with whom it is speaking was not able to understand certain phrases properly. However, this requires that the ASR of Nuance provides a confidence measure about the detected phrases. Patrick will contact Nuance in Ulm to discuss this issue.

If the ASR is not able to provide this confidence measure, the reverberation time or SNR might be taken as indicator. For this, it might be possible that the robot emits some dedicated sound sequences in an idle phase.

### SB meeting (14:15-14:55)

Measures to conduct the monthly SB meetings more efficiently were discussed. It was suggested to shift purely technical discussions, which are not of interest to all Partners, to bilateral meetings and to discuss in the SB meeting only issues which are concerned with the timely delivery of deliverables, the achievement of milestones, the discussion of overarching technical problems (like the synchronisation issue), other risk issues, etc.

### Meeting of Map Working Group (14:15 – 14:55)

The meeting, led and recorded by Christine Evers, has achieved the following results.

- The EARO (Matlab) object defines inputs to subsystems
- Maps define outputs / results of subsystems
  - o Maps are objects containing deterministic representation of belief as their properties
  - o Map objects contain the *common* information shared by all subsystems that is required to interpret data
    - Additional fields can be added locally but does not need to be available globally
    - E.g., audio-visual processing may require fields internally that aren't relevant to pure audio localisation algorithms. Internally create a subclass AV\_maps that inherits from general map class and adds required fields
  - o AL Memory provides API / utilities for frame conversions (camera frame, torso, joint, world centric)

Objects within the map should have the following properties:

- Unique ID
- Position + covariance of detected objects in the environment
- Walls: boundaries, feature points
- Flag for active / passive sensing
- Flag for time-varying/moving or stationary
- Sensor type
- Obscuration periods / silent periods
- ID of EARO object that was last used to modify the object
- Pointer to sound file?

In addition, the following operations should be performed with map objects:

- reading / writing to/from maps (XML)
- get/set data
- 

The following questions remained open:

- How to best represent extended objects and walls in an acoustic sense

- Spatial representation: From the discussion it became clear that each group may use different spatial representations (e.g., pixel grids, egosphere). Should projections (e.g. to/from egosphere) be part of the map methods or should they be specific to internal subclasses only? Which ones are common?
- Sound data should be stored somewhere and EARO as well as map objects should contain links to the specific wav files.

In a next step, an initial map object with properties and methods as discussed before should be generated and circulated for review to the working group members.

### **Closing Session** (15:00-15:10)

Walter Kellermann thanked Patrick Naylor and his team for being a great host and all the participants for their presentations and efforts, and the many fruitful discussions.

---

## Appendix

All presentations and documents of this meeting can be found on the EARS homepage under the category [Meetings](#).

- Appendix 1: [EARS\\_Meeting\\_London\\_List\\_of\\_Participants.pdf](#)
- Appendix 2: [EARS\\_Meeting\\_London\\_Presentation\\_WP2\\_IMPERIAL.pdf](#)
- Appendix 3: [EARS\\_Meeting\\_London\\_Presentation\\_WP2\\_FAU.pdf](#)
- Appendix 4: [EARS\\_Meeting\\_London\\_Presentation\\_WP1\\_BGU.pdf](#)
- Appendix 5: [\(Presentation not yet submitted.\)](#)
- Appendix 6: [EARS\\_Meeting\\_London\\_Presentation\\_WP4\\_UBER.pdf](#)
- Appendix 7: [EARS\\_Meeting\\_London\\_Presentation\\_Dialog\\_Engine\\_ALD.pdf](#)
- Appendix 8: [EARS\\_Meeting\\_London\\_Presentation\\_WP7\\_FAU.pdf](#)
- Appendix 9: [EARS\\_Meeting\\_London\\_Presentation\\_WP5\\_Overview\\_ALD.pdf](#)
- Appendix 10: [EARS\\_Meeting\\_London\\_Presentation\\_WP5\\_Details\\_ALD.pdf](#)
- Appendix 11: [\(Presentation not yet submitted.\)](#)
- Appendix 12: [EARS\\_Meeting\\_London\\_Presentation\\_Modularity\\_Framework\\_ALD.pdf](#)
- Appendix 13: [EARS\\_Meeting\\_London\\_Presentation\\_WP6\\_FAU.pdf](#)
- Appendix 14: [EARS\\_Meeting\\_London\\_Presentation\\_Recordings\\_with\\_the\\_Nao\\_FAU.pdf](#)
- Appendix 15: [\(Presentation not yet submitted.\)](#)