



D3.5 Final version of audio-visual analysis components



Grant Agreement nr	611057
Project acronym	EUMSSI
Start date of project (dur.)	December 1st 2013 (36 months)
Document due Date :	July 31st 2016 (32 months)
Actual date of delivery	September 12th 2016
Leader	IDIAP
Reply to	odobez@idiap.ch
Document status	Submitted

Project co-funded by ICT-7th Framework Programme from the European Commission

Project ref. no.	611057
Project acronym	EUMSSI
Project full title	Event Understanding through Multimodal Social Stream Interpretation
Document name	EUMSSI_D3.5_AV_FinalVersion.pdf
Security (distribution level)	PU – Public
Contractual date of delivery	July 31st 2016
Actual date of delivery	September 12th 2016
Deliverable name	Final version of audio-visula processing modules, including speech recognition and person identification
Type	P – Prototype
Status	Submitted
Version number	V1
Number of pages	
WP /Task responsible	WP3: T3.1-LIUM, T3.2-LIUM, T3.3-IDIAP, T3.4-IDIAP, T3.5-IDIAP
Author(s)	Jean-Marc Odobez (IDIAP)
Other contributors	
EC Project Officer	Mrs. Alina Lupu Alina.LUPU@ec.europa.eu
Abstract	The deliverables should report on the audio-visual modules developed within the project, including those used in the demonstrator. For a more comprehensive report, the corresponding content will be described along with D36 on System integration and evaluation report of the audiovisual analysis component.
Keywords	Speech recognition, Speaker Diarization, face Diarization, Audio-visual person recognition, OCR
Circulated to partners	Yes
Peer review completed	Yes
Peer-reviewed by	Jens Grivolla (UPF)
Coordinator approval	Yes

1. SUMMARY

The deliverable D3.5 should report on the latest developments made within the EUMSSI project on all tasks of workpackage WP3 on audio-visual analysis, which includes speech recognition as well as person segmentation and identification in videos. In particular it should provide the description of the modules integrated in the final version of the demonstrator.

For a more comprehensive report, we propose to provide the corresponding content along with deliverable D36 on System integration and evaluation report of the audiovisual analysis component, since:

- D36 would have to remind a similar content;
- Final integration and data processing is still going on until the user day evaluation on the 22nd of September;
- Evaluation of some modules (person identification) is done in the context of the Person Discovery challenge of the 2016 MedialEval initiative. This challenge is ongoing, and final results and evaluation will be provided at the time of the workshop (21st of October). Note that EUMSSI has contributed in several ways to this challenge, in particular by providing a significant amount (around 50 hours) of videos from the Deutsche Welle partner.

In this view, the deliverable D36 will be delivered at the end of October (planned date in Annex B: 30th of September).