



 chist-era project **JOKER** 

ERA-Net CHIST-ERA: European Coordinated Research on Long-term Challenges in Information and Communication Sciences & Technologies

JOke and Empathy of a Robot/ECA: Towards social and affective relations with a robot

Seminar CHIST-ERA Istanbul : 4 March 2014

Kick-off meeting : **27 January 2014** (call IUI 2012)

<http://www.chistera.eu/projects/joker>

Laurence Devillers (devil@limsi.fr – LIMSI/SLP)



Seminar CHIST-ERA Istanbul 2014



**KOÇ
ÜNİVERSİTESİ**



Partners background

JOKER♥



KOÇ
ÜNİVERSİTESİ



UMONS
Université de Mons

- **affective and social dimensions in spoken interaction**, emotion and affect bursts detection, user models, IHRobot, **dialogue, generation** (Laurence Devillers, Sophie Rosset)
- **social interaction, multimodal interaction**, collection of data, affect bursts detection and generation (Nick Campbell)
- **user detection using visual cues and dialog, visual interpretation** (eyes tracking, face, gesture), affect bursts detection, temporal model of gesture, gaze and speech
•(Metin Sezgin)
- **speech recognition (Kaldi/Sphinx), models of humor** (Yannick Estève, Daniel Luzzati)
- **speech synthesis, affect bursts detection and generation** (laugh, breath, sigh, throat, etc.) (Stéphane Dupont)



Objectives **JOKER**❤️

JOKe and **E**mpathy of a **R**obot/ECA

create a generic intelligent user interface providing a multimodal dialogue system with social communication skills including humor, empathy, compassion and other informal socially-oriented behavior.

fuse the verbal and non verbal cues (audio, eye-gaze, gestures) including affect bursts for social and emotional processes both in perception and generation

build rich user profiles taking into account user's personality, interactional behavior

explore advanced dialogues involving complex social behaviors in order to create a long-term social relationship

react in real-time



Main challenges

JOKER ❤️

Social interactions require social intelligence and “understanding” for dealing with news circumstances by anticipating the mental state of another person.

JOKER will investigate humor in human-machine interaction. Humor can trigger surprise, amusement, or irritation if it doesn't match the user's expectations.

JOKER will explore two social behaviors: expressing empathy and exchanging chat with the interlocutor as a way to build a deeper relationship.

Implementing empathy or humor in a companion-machine requires

- that the emotional expression and intention of the user should be detected,
- that the context should be understood,
- that the system should have a memory,
- and that the system is able to express an emotional/expressive response comprehensible by the user.



Use-case

JOKER ❤️

Application prototype in a laboratory cafeteria with regular participants (students, staff, visitors...)

Social interactions in cafeteria beside coffee machine both in Ireland and France (2 languages) with different devices (robot or ECA)

2 situations will be studied

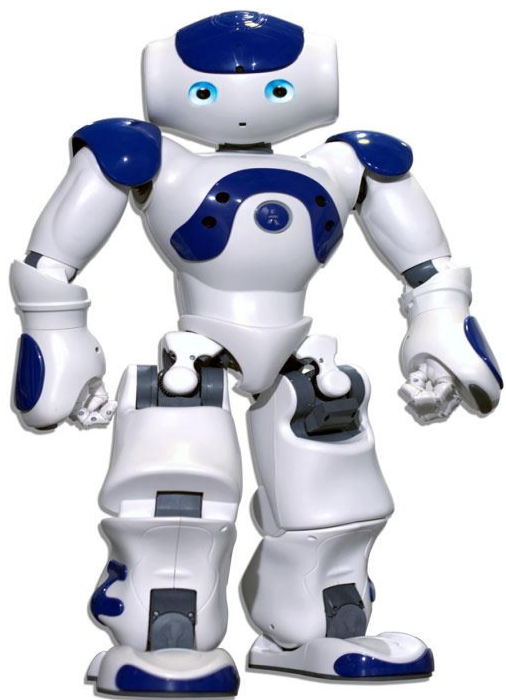
- one-on-one : human-robot/ECA**
- robot/ECA with multiple people**

We will build specific scenarii for engaging people in a conversation with the robot

Our results/platforms will be useful for designing robot for other applications such as for elderly people



Robot/ECAas **JOKER**❤️





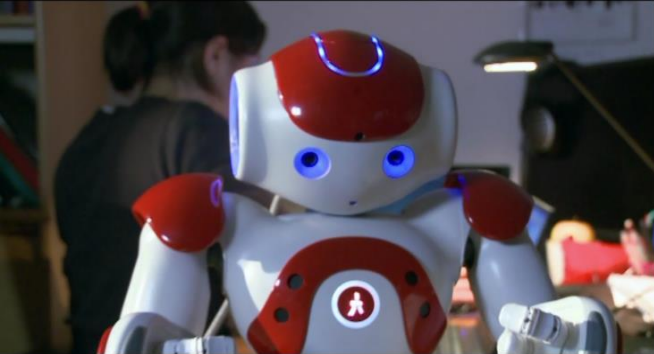
WPs **JOKER**♥

JOKER will react in real-time with

- a **robust perception module (WP3)** (sensing user's facial expressions, gaze, voice, audio and speech style and content),
- a **social interaction module modelling user and context, with long-term memories (WP4)**,
- and a **generation and synthesis module for maintaining social engagement with the user (WP5)**.

The research will provide

- a **collection of multimodal data with different socially-oriented behavior scenarios in two languages (French and English) (WP2)**
- and an **evaluation protocol for such systems.(WP6)**



Partners and WPs **JOKER**♥



Start : January 2014

WP1 (LIMSI): Management

WP2 (TCD): Domain and Databases of interactive speech

WP3 (KOC) : Perception modules

WP4 (LIMSI) : Dialogue and decision modules

WP5 (UMONS): Generation and synthesis modules

WP6 (LIUM): Evaluation



WP 1 JOKER♥

LIMSI, LIUM, TCD, UMONS, KOC

D1.1	M1	Kick-off meeting (done)
D1.2	M3	Web-site (in progress)
D1.3	M8	Consortium agreement (first version sent)
D1.4	M12	Annual report
D1.5	M18	Intermediate report
D1.6	M24	Annual workshop
D1.7	M42	Final public workshop
D1.8	M42	Final report



WP 2 JOKER♥

TCD, LIMSI, LIUM

Deliverable	Month of delivery	Title of deliverable
D2.1	M6 – M24	Domain definition and scenarios (v1 and v2) (in progress – collective work)
D2.2	M12 M24 M36	Data collection tool with real system (v1, v2 and v3) (in progress) – first tests
D2.3	M12 M24 - M36	Data collection (dataset 1 , 2 , 3)
D2.4	M15 M27 - M39	Data annotation (dataset 1 , 2 , 3) (annotation protocol will be a collective work)
D2.5	M42	Study of cultural aspects of social interaction

- Main idea : use as soon as possible an automatic system instead of a WoZ for the data collection → bootstrapping procedure



WP 2 JOKER

TCD, LIMSI, LIUM

- TCD will lend expertise in domain specification and initial data collection for training and building the conversational dialogue system (Herme project).
- Example of LIMSI background work (experiences with WoZ, ex: elderly people)



WP 3 JOKER♥

KOC, LIMSI, LIUM, TCD, UMONS

Deliverable	Month of delivery	Title of deliverable
D3.1	M10- M22- M34	User detection using visual cues - Real-time emotion and social behaviors detection using visual cues – - deliverable software (v1, v2 and v3)
D3.2	M10- M22- M34	Real-time emotion and social behaviors detection using audio cues - deliverable software (v1, v2 and v3)
D3.3	M10- M22- M34	Real-time emotion and social behaviors detection using affect bursts - deliverable software (v1, v2 and v3)
D3.4	M10- M22- M32	Automatic speech recognition - deliverable software (v1, v2 and v3)
D3.5	M10- M22- M34	Names entities, Topics detection - deliverable software (v1, v2 and v3)
D3.6	M12- M24 M36	Integration and Fusion of linguistic and multi-modal cues for emotion and social behavior detection - deliverable software (v1, v2 and v3)



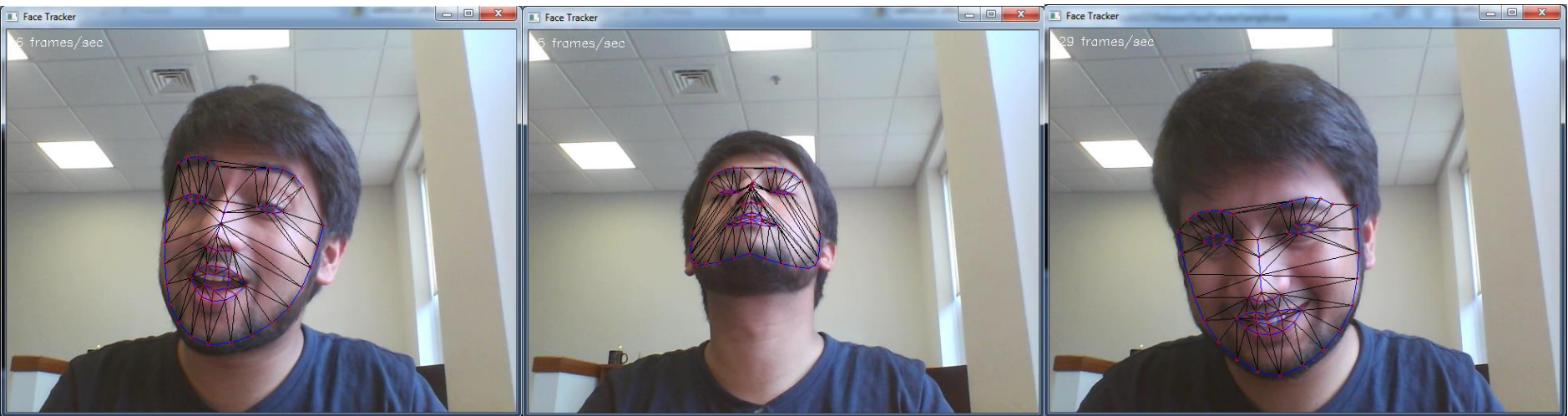
WP 3 JOKER♥

KOC, LIMSI, LIUM, TCD, UMONS

- **Preliminary Work – Facial Tracking (KOC)**
- **Preliminary Work – Emotion detection from speech with NAO (LIMSI)**
LivingWithRobot 2012: <http://www.youtube.com/watch?v=p1ID-gvUnWs>

Preliminary Work – Facial Tracking (KOC)

- Attempts of making our own dataset using Kinect to learn about affect bursts and emotions since Kinect is ideal device to capture multimodal data (facial tracks, body gesture and sound)
- Non rigid facial tracking using Jason Saaragih's implementation*.



Figure(1): Sample outputs of the facial tracker

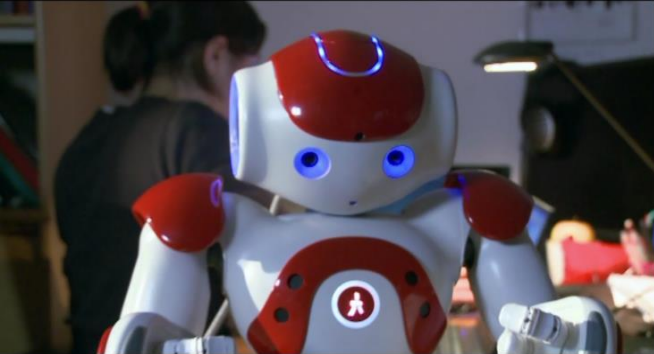
* J. Saragih, S. Lucey and J. Cohn, "**Deformable Model Fitting by Regularized Landmark Mean-Shifts**", *International Journal of Computer Vision (IJCV)*, 2010



WP 4 JOKER♥

LIMSI, LIUM, TCD

Deliverable	Month of delivery	Title of deliverable
D4.1	M3 - M6	Semantic representation (deliverable) - Dialog platform (software)
D4.2	M12- M24 M36	Dynamic emotional profile of the user (software and deliverable) - (v1, v2 and v3)
D4.3	M12- M24 M36	Ontology, history of the dialog, anticipation and memorization modules - blackboard (software and deliverable) (v1, v2 and v3)
D4.4	M12- M24 M36	Intuitive decision way : Dialog strategies using synchrony, mimics (software and deliverable) (v1, v2 and v3)
D4.5	M12- M24 M36	Cognitive decision way (software and deliverable) (v1, v2 and v3)



WP 5 JOKER♥

UMONS, LIMSI, TCD

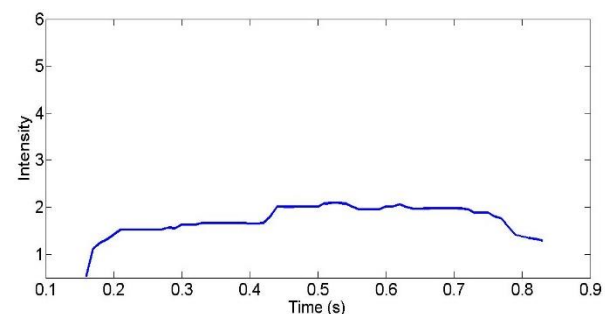
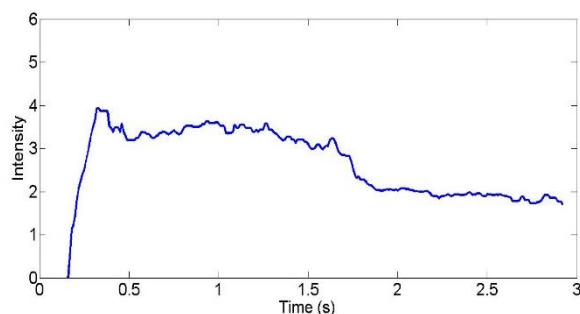
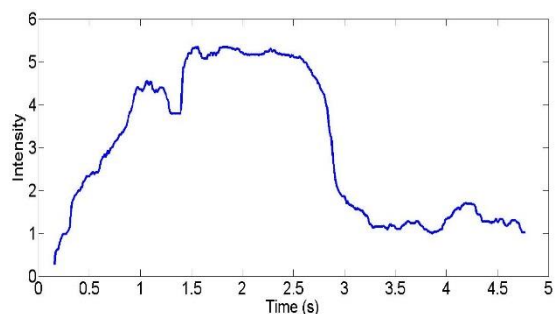
Deliverable	Month of delivery	Title of deliverable
D5.1	M12- M24 M36	Generation (v1, v2 and v3)
D5.2	M12- M24 M36	Speech Synthesis (v1, v2 and v3)
D5.3	M12- M24 M36	Affect bursts generation (v1, v2 and v3)
D5.4	M12- M24 M36	Multimodal generation and synthesis (v1, v2 and v3)



WP 5 JOKER♥

UMONS, LIMSI, TCD

- Example of background work – laughter synthesis:
 - From laughter intensity curves
 - To laughter “phonetic” transcriptions
 - And laughter audio and audiovisual generation





WP 6 JOKER♥

LIUM, UMONS, LIMSI, TCD, KOC

Deliverable	Month of delivery	Title of deliverable
D6.1	M12- M24 – M36	Protocol and Metrics (engagement measures : verbal and non verbal such as laugh, smile, eye-tracking, interaction duration)
D6.2	M15 M27 - M39	Individual components evaluation (3 times evaluation during the project)
D6.3	M39	Final Evaluation – One use-case in cafeteria
D6.4	M42	Impact of the companion (Robots, ECAs)
D6.5	M42	Dissemination (final workshop)



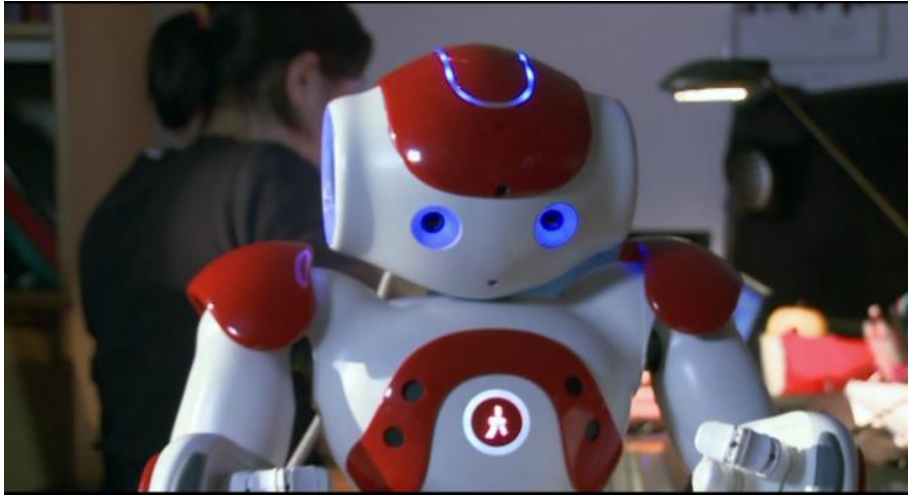
Work in progress

JOKER ❤️

Our JOKER system will be tested in two different languages with at least two different platforms (Robot and ECA). – first version in M12

Main results:

- Collaboration between complementary European teams on perception, dialogue and generation modules for HRI,**
- Multimodal original corpora available for the community,**
- Longitudinal experiment of people engagement with a social robot**
- Impact of the humor in the social interaction with a machine.... and a study of laugh and humor across languages in French and in English.**



Thanks for your attention
JOKER♥ team