KRIStina
A Knowledge-Based Virtual Conversation Agent
– with social and cultural competence –

http://kristina-project.eu/

Leo Wanner & Consortium
leo.wanner@upf.edu
In Europe, the number of people who feel marginalized or excluded from the society they live in increases every year.

- Two groups are particularly affected:
  - **Elderly**: often miss active social and public life, the feeling being cared for and being needed, ...
  - **Migrants**: suffer from cultural and language barriers; are not acquainted with the administration, ...

There is thus a need for intelligent conversation agents as social companions that are able to entertain, converse, coach, etc.
... that is, there is a need for KRISTINA

with a focus on

- geriatric
- basic healthcare

contexts
A sample dialogue with KRISTINA

A: You look downhearted today. What is wrong?
U: I feel sad. Because of my eyes, I even can’t read the newspaper anymore.
A: Shall I read the newspaper aloud for you?
U: Yes, this would be great!
A: You certainly can still read the headings of the articles. Just tell me which one I shall read. ...

...
KRISTINA’s Settings

• Target roles
  ▪ **Social companion** in geriatric contexts (small talk, daily routines, cultural and physical activities, ...)
  ▪ **(Basic care) assistant** (preferences of the care taker, care advices)
  ▪ **(Health care) adviser** (info on health care system, body care advices, ...) 

• Target users
  ▪ **Elderly** (care recipients) with German and Turkish background
  ▪ **Care givers** with Polish background
  ▪ **Migrants** in need of health care information with North African background
KRISTINA’s Technologies

• Know how and when to interact
  ▪ Dialogue management

• Listen and watch
  ▪ (Multilingual) Vocal analysis
    ▪ ASR
    ▪ Language analysis
  ▪ Emotion analysis
    ▪ Facial expression analysis
    ▪ Gesture analysis
    ▪ Paralinguistic cue analysis

• Think and learn
  ▪ Knowledge Processing
    ▪ Ontology building
    ▪ Knowledge integration
    ▪ Reason
  ▪ Web knowledge Retrieval
    ▪ Retrieval of background info
    ▪ Retrieval of cultural info

• Speak and act
  ▪ Modality selection
  ▪ Expressive speech generation
    ▪ Language generation
    ▪ Prosody generation
  ▪ Embodied character creation
Where do we stand?

- Third (and last) year
- Working on the Final Demonstrator
- The coverage of the number of conversation topics is (still) limited, but can be extended as needed
- Language coverage: German, Polish, Spanish, Turkish

Come to see us at our stand!
Thank you!
KRIStINA Architecture

ASR

Language Analysis

Facial Expr. Recognition

Gesture Analysis

Paraling. Cue Recognition

Multimodal Fusion

Interaction Manager (VSM)

Dialogue Manager

Turn-taking & Idle Behavior Manager

Multimodal Fission & Discourse Planning

Facial Expr. Determination

Gesture Determination

Language Generation

Mode Sync. & Real.

Knowledge Base

Targeted IR

SSI

Facial Expr. Recognition

Gesture Analysis

Paraling. Cue Recognition

Multimodal Fusion

Dialogue Manager

Turn-taking & Idle Behavior Manager

Multimodal Fission & Discourse Planning

Facial Expr. Determination

Gesture Determination

Language Generation

Mode Sync. & Real.

Knowledge Base

Targeted IR

SSI
KRISTINA’s Knowledge Model

- Represent user profiles, etc.
- Cover domain models for both fusion and interpretation (“understanding”) of verbal and non-verbal aspects of user turns and generation of the turns of the agent

I feel sad

+ Valence / Arousal
KRISTINA’s Knowledge Model

- Incorporate retrieved information
- Support dialogue management
  - Plan the possible next moves of the agent
  - Model the DA typology
Multimodal Interaction

Multimodal communication analysis

- Speech recognition (using adapted off-the-shelf technology)
- Language analysis
  - Speech act
  - Content (Surface syntax -> Deep syntax -> Semantics -> Concepts -> OWL)

Surface-syntactic representation

Deep-syntactic representation

P
SBJ
PRD

"I"
"feel"
"sad"

"feel"
"sad"
Multimodal Interaction

Multimodal communication analysis

• Speech recognition (using adapted off-the-shelf technology)
• Language analysis

• Multimodal affective cue analysis
  ▪ Facial cues (automatically determining Action Units and mapping them into the valence /arousal space)
  ▪ Paralinguistic cues (using lexical and intonation (acoustic) features)

• Event-driven and semantic fusion of the modalities
Multimodal Interaction

Multimodal communication generation

• Agent move planning (Dialogue Management)

• Modality fission
  ▪ Modality determination (rule-based)
  ▪ Modality instantiation (using NLG, mimicon and gesticon)
Multimodal Interaction

Multimodal communication generation

- Modality fission
- Modality realization
  - verbal

Conceptual representation

Deep-syntactic representation

Predicate-argument representation

Surface-syntactic representation
Multimodal Interaction

Multimodal communication generation

- Modality fission
- Modality realization
  - Non-verbal
## Preliminary evaluation

<table>
<thead>
<tr>
<th>Evaluation statement</th>
<th>Likert scale value (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>It is clear what KRISTINA wants to communicate</td>
<td>3.23 (± 1.42)</td>
</tr>
<tr>
<td>KRISTINA does not provide the right amount of information</td>
<td>2.73 (± 1.10)</td>
</tr>
<tr>
<td>The conversation with KRISTINA is confusing</td>
<td>2.84 (± 1.27)</td>
</tr>
<tr>
<td>KRISTINA behaved as expected</td>
<td>3.0 (± 1.21)</td>
</tr>
<tr>
<td>KRISTINA acted on own initiative</td>
<td>3.25 (± 1.29)</td>
</tr>
</tbody>
</table>
Thank you!