*Actes de CORIA-TALN 2023*
*Paris, France, 5 au 9 juin 2023*

# µDialBot: MUlti-party perceptually-active situated DIALog for human-roBOT interaction

Fabrice Lefèvre, Timothée Dhaussy, Ahmed Ndouop Njifenjou,
Virgile Sucal, Bassam Jabaian
LIA, Avignon Université, France
{fabrice.lefevre}@univ-avignon.fr

## RÉSUMÉ

Dans muDialBot (ANR-20-CE33-0008-01), notre ambition est d'incorporer pro-activement des traits de comportements humains dans la communication parlée. Nous projetons d'atteindre une nouvelle étape de l'exploitation de l'information riche fournie par les flux de données audio et visuelles venant des humains. En particulier en extraire des événements verbaux et non-verbaux devra permettre d'accroître les capacités de décision des robots afin de gérer les tours de parole plus naturellement et aussi de pouvoir basculer d'interactions de groupe à des dialogues en face-à-face selon la situation. Récemment on a vu croître l'intérêt pour les robots compagnons capable d'assister les individus dans leur vie quotidienne et de communiquer efficacement avec eux. Ces robots sont perçus comme des entités sociales et leur pertinence pour la santé et le bien-être psychologique a été mise en avant dans des études. Les patients, leurs familles et les professionels de santé pourront mieux apprécier le potentiel de ces robots, dans la mesure où certaines limites seront rapidement affranchies, telles leur capacité de mouvement, vision et écoute afin de communiquer naturellement avec les humains, aù-délà de ce que permettent déjà les écrans tactiles et les commandes vocales. Les résultats scientifiques et technologiques du projet seront implémentés sur un robot social commercial et seront testés et validés avec plusieurs cas d'usage dans le contexte d'une unité d'hôpital de jour.

## ABSTRACT

In muDialBot (ANR-20-CE33-0008-01) our ambition is to actively incorporate human-behavior cues in spoken human-robot communication. We intend to reach a new level in the exploitation of the rich information available with audio and visual data flowing from humans when interacting with robots. In particular, extracting highly informative verbal and non-verbal perceptive features will enhance the robot's decision-making ability such that it can take speech turns more naturally and switch between multi-party/group interactions and face-to-face dialogues where required. Recently there has been an increasing interest in companion robots that are able to assist people in their everyday life and to communicate with them. These robots are perceived as social entities and their utility for healthcare and psychological well being for the elderly has been acknowledged by several recent studies. Patients, their families and medical professionals appreciate the potential of robots, provided that several technological barriers would be overcome in the near future, most notably the ability to move, see and hear in order to naturally communicate with people, well beyond touch screens and voice commands. The scientific and technological results of the project will be implemented onto a commercially available social robot and they will be tested and validated with several use cases in a day-care hospital unit.

MOTS-CLÉS : projet ANR, interaction humain-robot, pro-activité, multi-partie.

KEYWORDS: ANR project, human-robot interaction, pro-activity, multi-party.

Actes de la 30e Conférence sur le Traitement Automatique des Langues Naturelles (TALN), volume 6 : projets, pages 26–29.

# 1 Motivations

Companion robots able to assist people in their everyday life and to communicate with them :

**social assistive robots**

**How to improve acceptance ?**

Several technological barriers to overcome :

- better ability to move, see and hear in order to naturally communicate with people in various configurations (face-to-face, multi-party, close, distant etc)
- pro-active use of perceptions
- audio-visual strategies to handle interactions

# 2 Ambitions

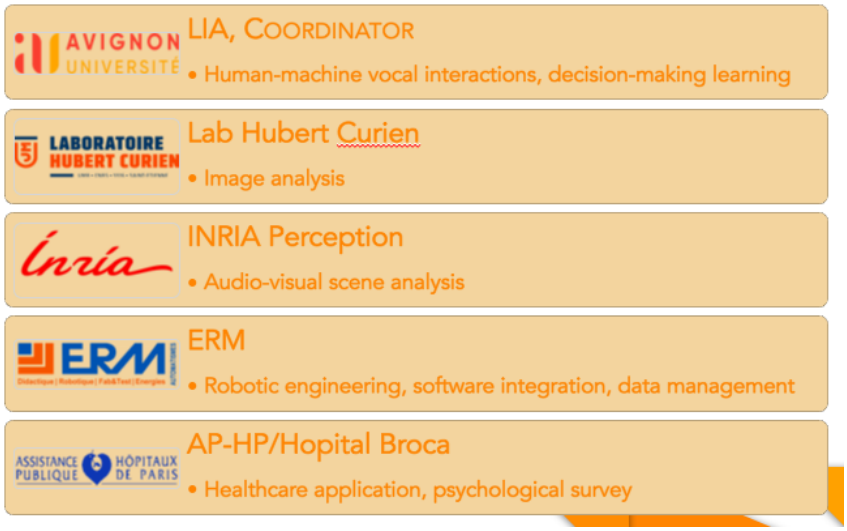
New level in exploitation of rich information available with audio and visual data flowing from humans when interacting with robots :

- fusing highly informative **verbal and non-verbal perceptive features** to enhance the robot's decision-making ability such that it can
- switch between **multi-party/group interactions and face-to-face dialogues** where required, and
- take speech turns **more naturally**


Tested and validated with several **use cases in a day-care hospital** unit. Large-scale data collection, complement in-situ tests, to fuel further researches.

Project mainly **funded by ANR and labeled by the Pole SCS**. More information on : https ://www.pole-scs.org/projets/mudialbot/

# 3 Partners




# 4 Organization

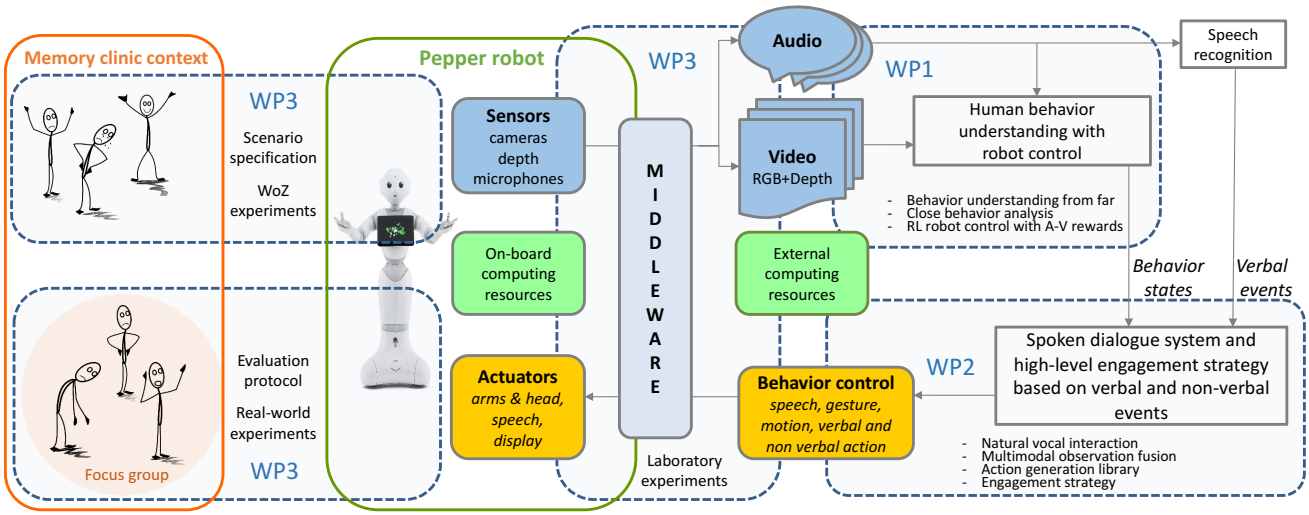
Methodology followed in µDialBot at the same time mainstream, in line with the classical methodology to handle **new machine learning-based approaches**, and pioneering, with **novel online learning techniques** to reduce the requirement for initial data collection to its minimum. Relying on deep learning techniques which have shown their efficiency in other domains, a new formalism dedicated to **pro-active perceptually-based control of a conversational robot** will be proposed.

General methodology of the project consists of two operational building blocks for :

— the estimation of non-verbal states and

— the learning of an event-guided strategy.

Integrated on the robotic platform through a software abstraction layer. A first round of WoZ experiments to test the proposed models, then the complete system gradually introduced in the true clinical context, using a well-defined protocol.




3 main work-packages implemented :

— ***WP1 Human behavior understanding with robot control***
**Objectives** : to develop methods and algorithms to extract HBU cues from audio and visual data.

— ***WP2 Spoken dialogue system and high-level engagement strategy based on verbal and non-verbal events***
**Objectives** : to develop the natural vocal interaction ability of the robot.

— ***WP3 Specifications, integration on robotic platform, iterative and final evaluation of the human-robot interactions in a memory clinic***
**Objectives** : to define the experimental protocol, specify laboratory and real-world experimental protocols, and conduct “Wizard of Oz” (WoZ) experiments.