Social Robotics

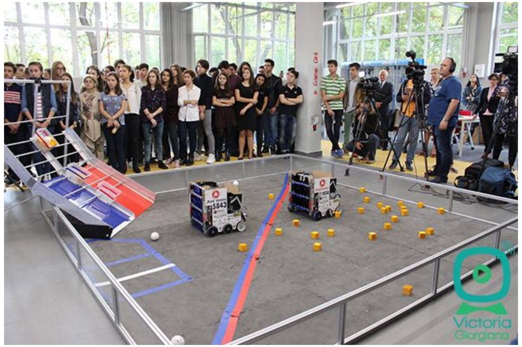
Challenges and Applications

Alexandru Sorici

Robotics (the usual thought)

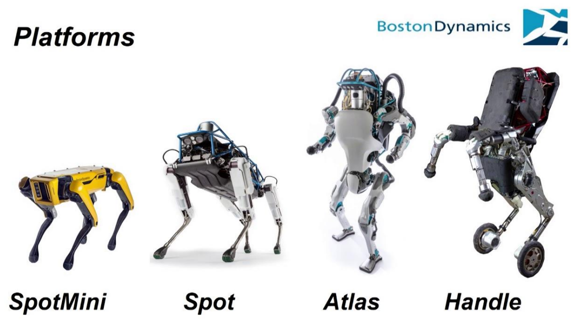
Photo Credit: Disruption Hub

Amazon Warehouse Robots Photo Credit: gifs.com



First Tech Challenge Robot Competition
Photo Credit: EditiaDeDimineata.ro

Da Vinci Surgical Robot
Photo Credit: Jewish General Hospital



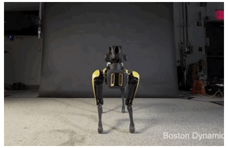


Photo Credit: Boston Dynamics

Robotics (the usual thought)

ISO definition: *“reprogrammable, multifunctional manipulator designed to move material, parts, tools or specialized devices through variable programmed motions for performance of a variety of tasks.”*

Wikipedia definition: “A **robot is a machine**—especially one programmable by a computer—**capable of carrying out a complex series of actions automatically**. A robot can be guided by an external control device, or the control may be embedded within. **Robots may be constructed to evoke human form, but most robots are task-performing machines**, designed with an **emphasis on stark functionality**, rather than expressive aesthetics.”

Assistive Robotics - an umbrella term

Defining characteristics of *assistive robots*:

- Work **alongside / in assistance of** humans => direct interaction with humans
- Can perceive their environment and **other individuals** using sensors and intelligent algorithms
- Can **communicate with people** multimodally
- **Can have a degree of autonomy** - for navigation, decision making
- Have a strong focus on **safety of the interaction**

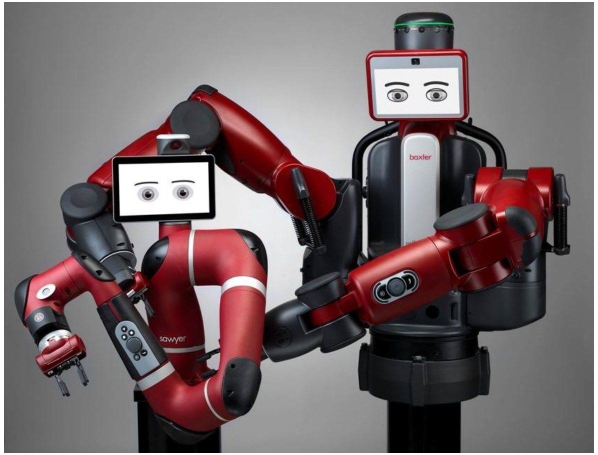
Assistive Robotics - when viewed as *collaboration*

COBOTS (collaborative robots)

Wikipedia definition: *"A cobot, or collaborative robot, is a robot intended for direct human robot interaction within a shared space, or where humans and robots are in close proximity. Cobot applications contrast with traditional industrial robot applications in which robots are isolated from human contact."*

Assistive Robotics - when viewed as *collaboration*

Kuka Robotic Arm
Photo Credit: RobotWorx - Collaborative robot safety



Baxter Robot
Photo Credit: NS Medical Devices

Assistive Robotics - the more common understanding

- An **assistive robot** performs a **physical task** for the **well-being** of a **senior person / person with disabilities.** The **task** is usually in the context of **Activities of Daily Living**.

- The **person** is in control of the robot (=> no autonomy)

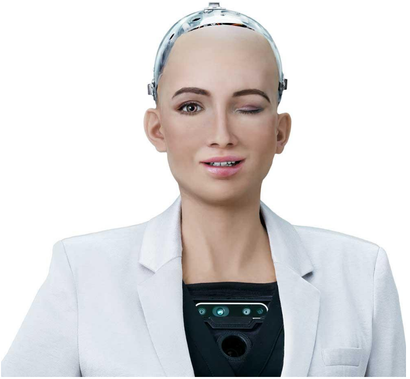
Assistive Robotics - the more common understanding

HAL (Hybrid Assistive Limb) by Cyberdyne
Photo Credit: roboticsfinder.com

JACO Assistive Robot Arm by Kinova Robotics
Photo Credit: kinovarobotics.com

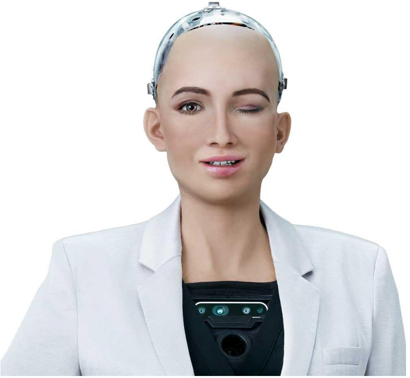
Social Robotics

Social Robotics



Sophia by Hanson Robotics
Photo Credit: hansonrobotics.com

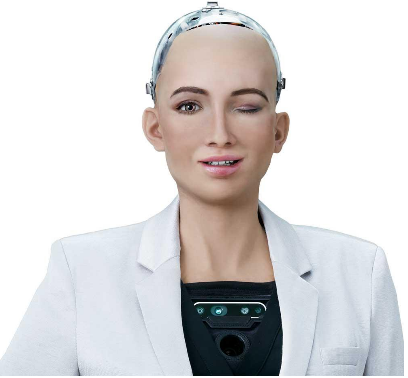
Social Robotics



Sophia by Hanson Robotics
Photo Credit: hansonrobotics.com

Paro by Paro Robotics
Photo Credit: parorobot.com

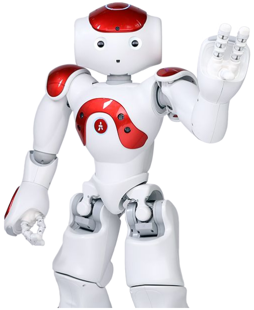
Social Robotics



Sophia by Hanson Robotics
Photo Credit: hansonrobotics.com

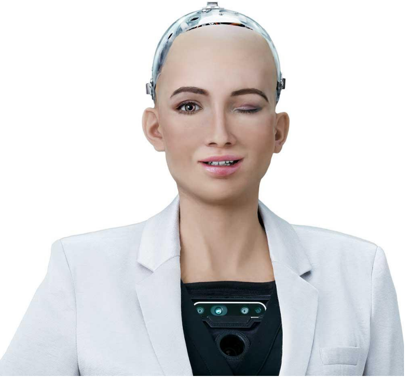
Paro by Paro Robotics
Photo Credit: parorobot.com

Pepper by Softbank
Photo Credit: softbankrobotics.com



Nao by Softbank
Photo Credit: softbankrobotics.com

Social Robotics



Sophia by Hanson Robotics
Photo Credit: hansonrobotics.com

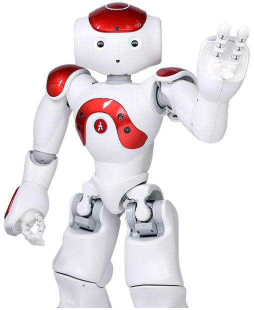
Paro by Paro Robotics
Photo Credit: parorobot.com

~~**Amazon Echo**~~

~~**Google Home**~~

Digital Assistants (robots require embodiment)

Pepper by Softbank
Photo Credit: softbankrobotics.com



Nao by Softbank
Photo Credit: softbankrobotics.com

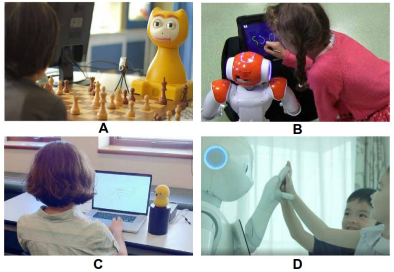
Social Robotics - Domains of Activity

- **Healthcare and Active and Assisted Living**
  - Assist aging or disabled individuals who are in need of ***supervision*** (but not active care)
    - This includes companion robots with manipulation capabilities, but is mostly focused on **communication capabilities** and facilitation of tele interactions
  - Emotional assistance
    - Mostly "pet" robots → based on pet-therapy in hospitals or care facilities
  - Therapy for people with Autism Spectrum Disorder
    - Robots have non-humanoid, humanoid (e.g. Nao) or animal-like form
    - Robots help in addressing and practicing social behaviors (e.g. eye contact, touch, liking), language development, stereotyped behaviors

- Why use robots in Healthcare and Assisted Living?
  - Many elderly users live alone
  - Time constraints for quality care on both formal and informal caregivers
  - Cost-savings (given the lack of sufficient care workers and growing aging population)

Social Robotics - Domains of Activity

- **Education**
    - Focused mostly on tutoring and teaching for children (age groups 3-12)
    - Research focuses on both *cognitive (e.g. learning gains, improved test completion times)* and *affective* learning outcomes (e.g. improving attention, fatigue measurement, engagement measurement, anxiety reduction)
    - Examples:
        - learning a game (e.g. chess)
        - learning a foreign language (e.g. english for japanese students - including ***robot as novice*** setup)
        - tutoring during puzzle games
        - Handwriting improvement (*teachable robot*)

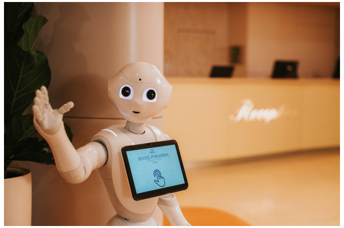


T. Belpaeme et al.
"Social robots for education: A review."
Science robotics 3, no. 21 (2018)

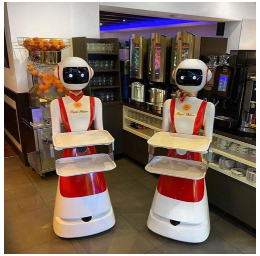
- Why use robots in Education?
    - High availability + easy to provide fact-based knowledge
    - Increased acceptance by young users
    - Studies begin showing positive effects when compared to just computer-based tutors (effect of “embodiment”)

Social Robotics - Domains of Activity

- **Public Assistance and Entertainment**
    - Social Robots as “public info points”, “waiters” and “performers” at various events and public interest institutions / venues (hotels, business centers, banks, restaurants)
    - Includes entertainment in social care scenarios (e.g. games / quizzes for the elderly)



- Why use robots in Entertainment?
    - High availability
    - Novelty factor



Social Robotics - how researchers define it

“**Social robots** are ***physically embodied agents*** that have ***some (or full) autonomy*** and engage in ***social interactions with humans***, by ***communicating***, ***cooperating***, and ***making decisions***. These *behaviours* are then *interpreted* by human onlookers as *‘social’*, according to current norms and conventions.” [1]

[1] Sarrica M, Brondi S, Fortunati L. How many facets does a “social robot” have? Inf Technol People. 2019

Social Robotics - what users expect of it

Being accepted as a **social entity** in a user home requires [2]:

1. **Two-way interaction** (robot has to respond to a human in a *human manner*)
2. Display **thoughts and feelings**
3. Be **socially aware** of their environment
4. Provide **social support** (be ***there*** for a person, like a friend)
5. Demonstrate **autonomy**

[2] de Graaf MMA, Ben Allouch S, van Dijk JAGM. What makes robots social?: a user's perspective on characteristics for social human-robot interaction. Social Robotics, 2015

Social Robotics - some distilled requirements

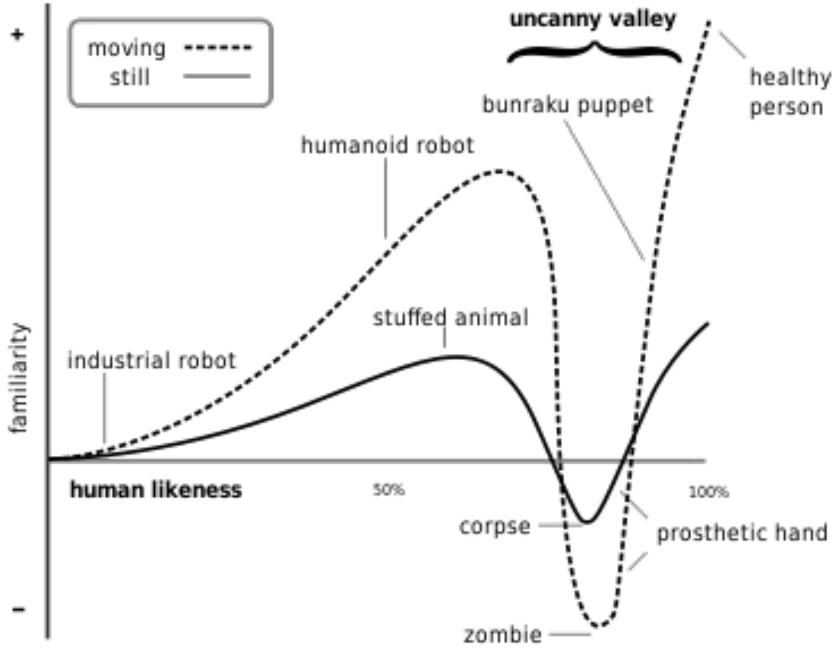
- Embodied ***agents*** that are part of a ***heterogenous group: humans and robots***
- Robots must **perceive and interpret the world**, creating their **own history**
- Engage in **social interactions / communicate** with humans (and other robots) following **behavioral norms**
- **Emotion modeling** (through speech, facial expressions, body language)

Social Robotics - Social Challenges :-)

- Researchers focus on *general* (navigation, people and environment perception) and *communication capabilities* (multimodal human-robot interaction)
- Users **expect** to **relate** to a social robot as they would to a **human friend**
  - The current limited social capabilities of a social robots lead humans to view them as **household servants** (rather than companions)
  - Users quickly lose interest if their **"social capabilities"** (e.g. reciprocal conversation) expectations are unmet

- Current research is **still** mostly focused on the technical challenges (**it still has to be**). The *social sciences* (psychology, user studies, behavioral science) are not yet in focus.

Social Robotics - Social Challenges :-)



The Uncanny Valley effect
Photo Credit: Wikipedia

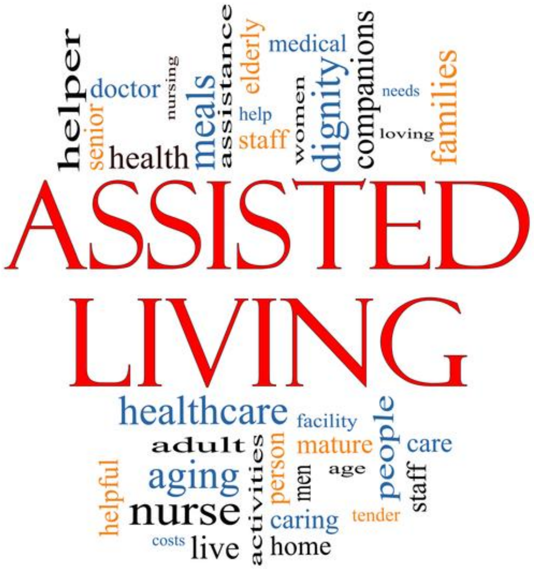
Social Robotics - Technical Challenges

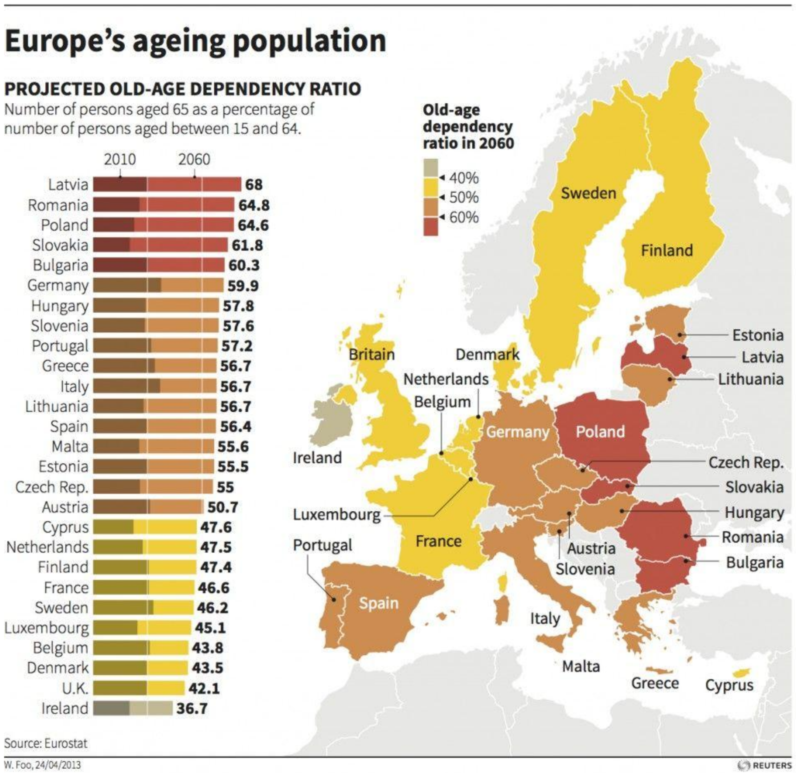
- **Human-oriented perception**:
  - People tracking, face detection/recognition, gesture recognition, action recognition, facial expression classification
- **Environment perception**:
  - Navigation/Exploration, object detection/recognition, lifelong SLAM
- **Interaction/Planning**:
  - Realistic Dialogue Management, People Modeling, Lifelong Behavior Management

Social Robotics @ AI-MAS

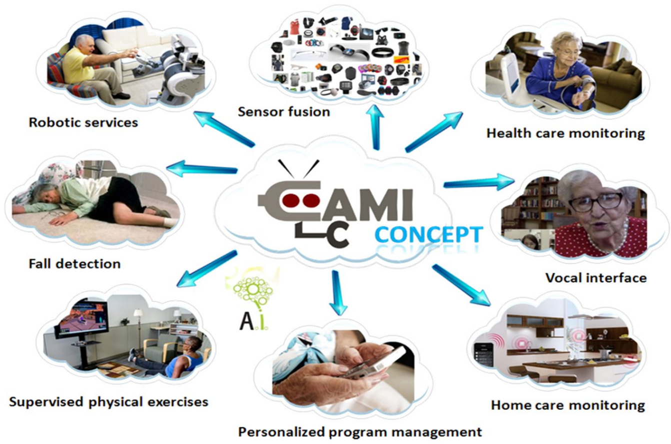
Our Experience with the Pepper Robot

Active and Assisted Living



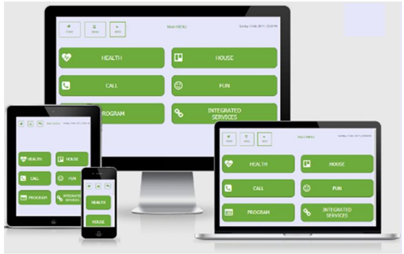


CAMI Project



CAMI Ecosystem

- Integrated solution to support elderly needs
- Functionalities:
  - Health data monitoring and sharing
  - Home monitoring
  - Supervised Physical Exercises
  - Intelligent Reminders and Planning
  - Multimodal Interactions



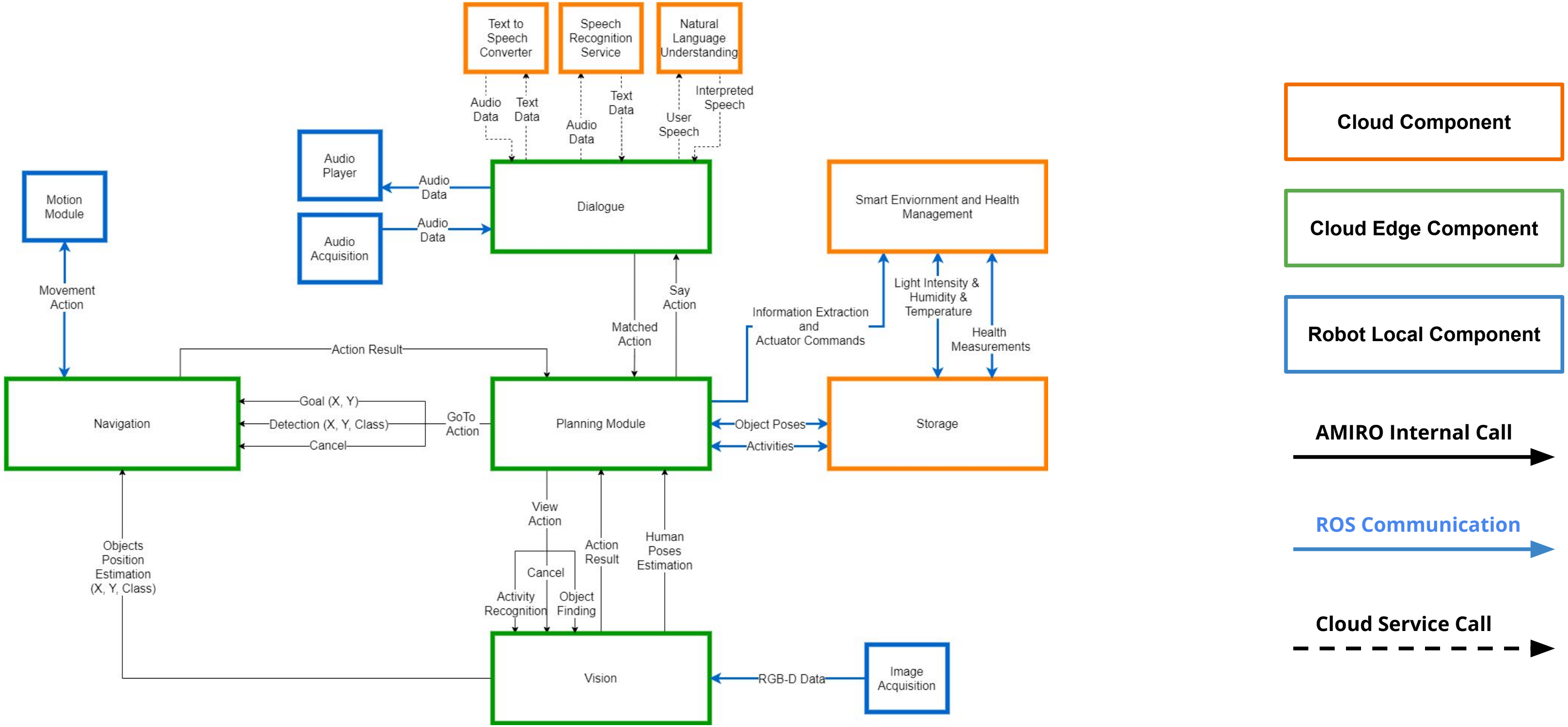
CAMI Multi-Modal Interface

- Set user preferences
- Status of the user (medical condition + reminders)
- Environment condition
- System Config

www.camiproject.eu

Our Work with the Pepper Robot: the AMIRO framework

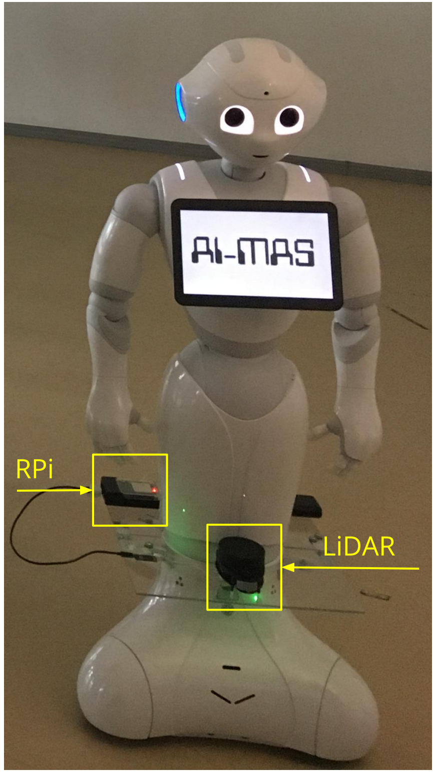
Our Work with the Pepper Robot: the AMIRO framework



S. Ghiță et al. "The AMIRO Social Robotics Framework: Deployment and Evaluation on the Pepper Robot." *Sensors* 20, no. 24 (2020)

Environment Perception in AMIRO: Navigation Module

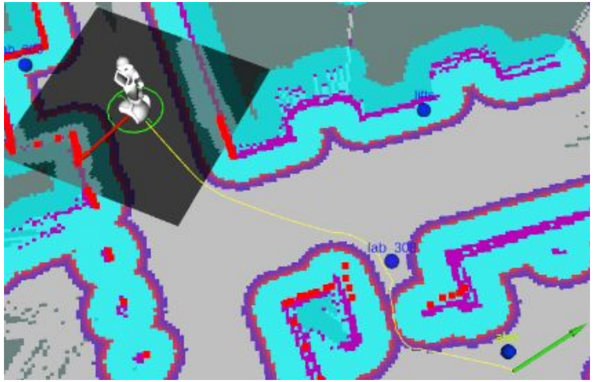
Environment Perception in AMIRO: Navigation Module

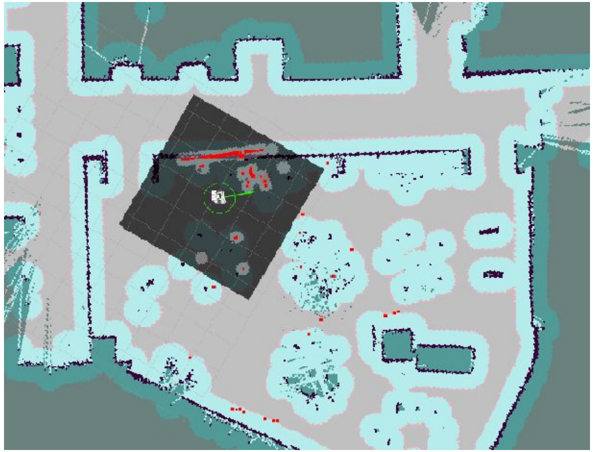


Localization / Map construction

- Requires use of an external 2D LiDAR (360° RP1 Lidar) + Raspberry PI3 board for LiDAR data acquisition

Environment Perception in AMIRO: Navigation Module





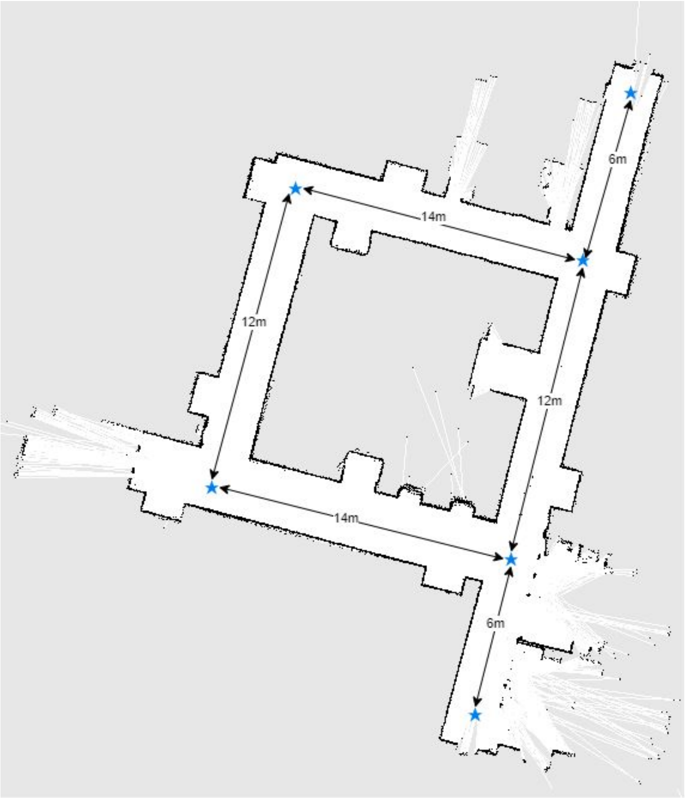
Localization / Map construction

- Requires use of an external 2D LiDAR (360° RP1 Lidar) + Raspberry PI3 board for LiDAR data acquisition

- Hector SLAM [3] for environment mapping
- **Fine-tuned** *amcl* ROS module [4] for localization (fine tuning looks at the physical / geometrical constraints of the robot - e.g. maximum translation and turn speed, move base footprint, safety margins)

[3] http://wiki.ros.org/hector_slam
[4] http://wiki.ros.org/amcl

Environment Perception in AMIRO: Navigation Module



Localization / Map construction

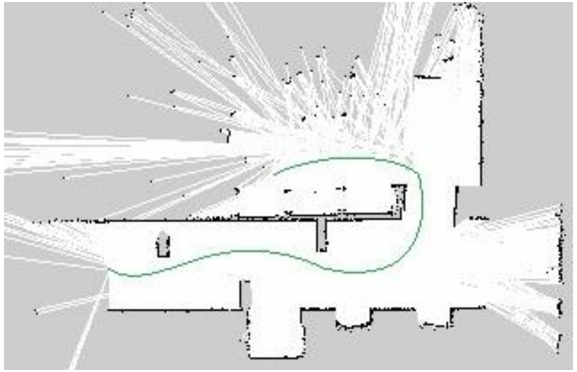
- Requires use of an external 2D LiDAR (360° RP1 Lidar) + Raspberry PI3 board for LiDAR data acquisition

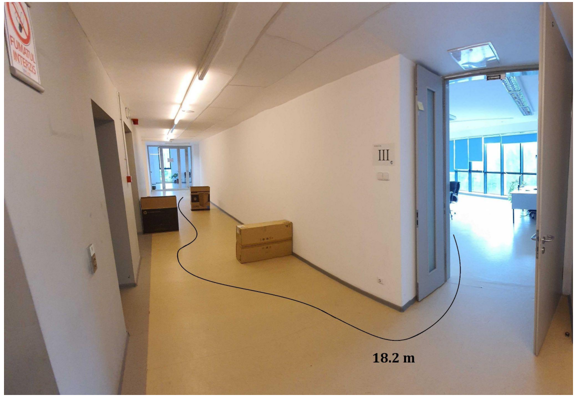
- Hector SLAM [3] for environment mapping
- **Fine-tuned** *amcl* ROS module [4] for localization

- External LiDAR required for large scale SLAM-based mapping (e.g. lab + floor) - range sensors of Pepper robot have too limited a range

[3] http://wiki.ros.org/hector_slam
[4] http://wiki.ros.org/amcl

Environment Perception in AMIRO: Navigation Module





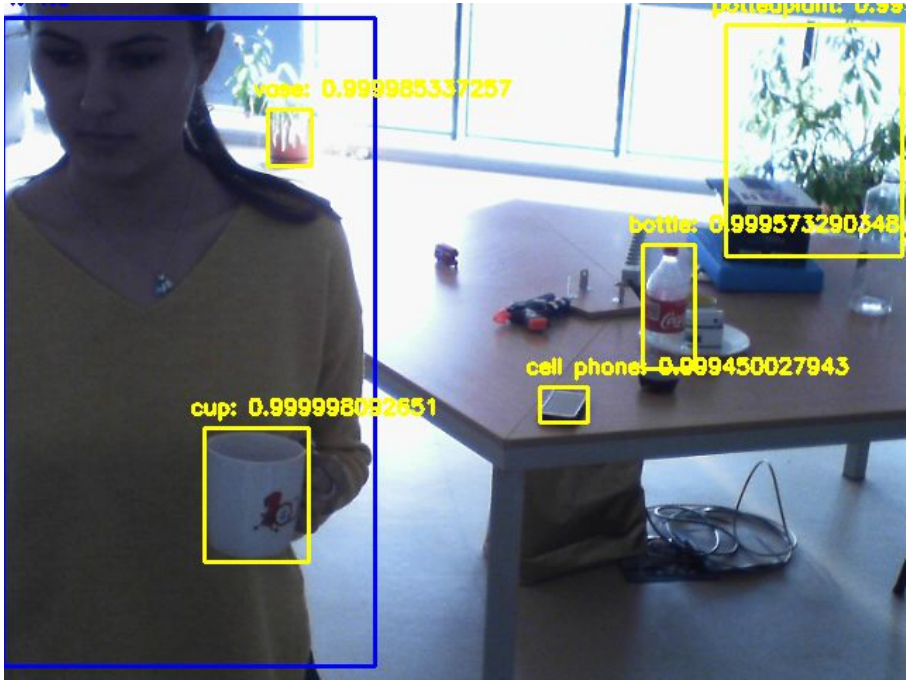
Navigation using existing map

- Uses the *move_base* ROS module, using DWA local planner and A* global planner

- Pepper robot has large orientation errors on rotation (10° on every 360° turn) + drift on translation (1° on every meter forward) => external LiDAR still required for correct localization

- Global Planner configured to plan for 5m movements at a time: e.g. navigation from lab to hallway for 18.2m takes 58.11s to execute

Environment Perception in AMIRO: Object Detection

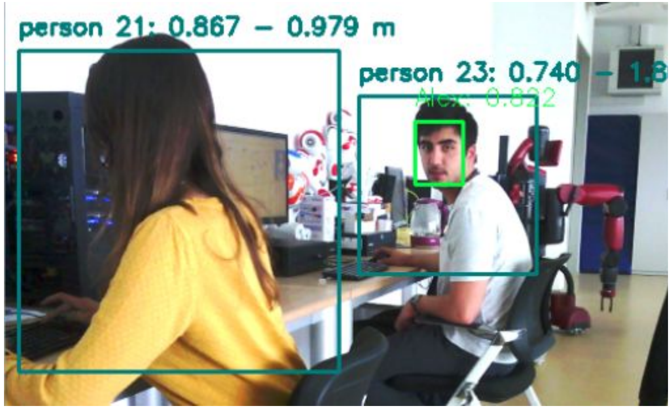


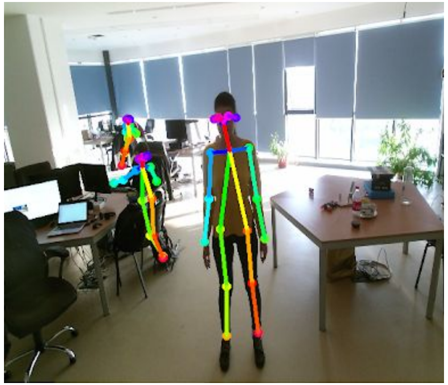
- Use of YOLOv3 [5] model for object detection
- Tracking of objects (including identified people) using the SORT [6] algorithm
- Use of an object segmentation algorithm [7] to *align* object pixels with robot depth-map => can estimate **distance to objects**

[5] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." (2018).
[6] A. Bewley et al., “Simple online and realtime tracking,” ICIP 2016
[7] D. Pakhomov et al., “Deep residual learning for instrument segmentation in robotic surgery”, 2017

Human-Centered Perception in AMIRO

Human-Centered Perception in AMIRO

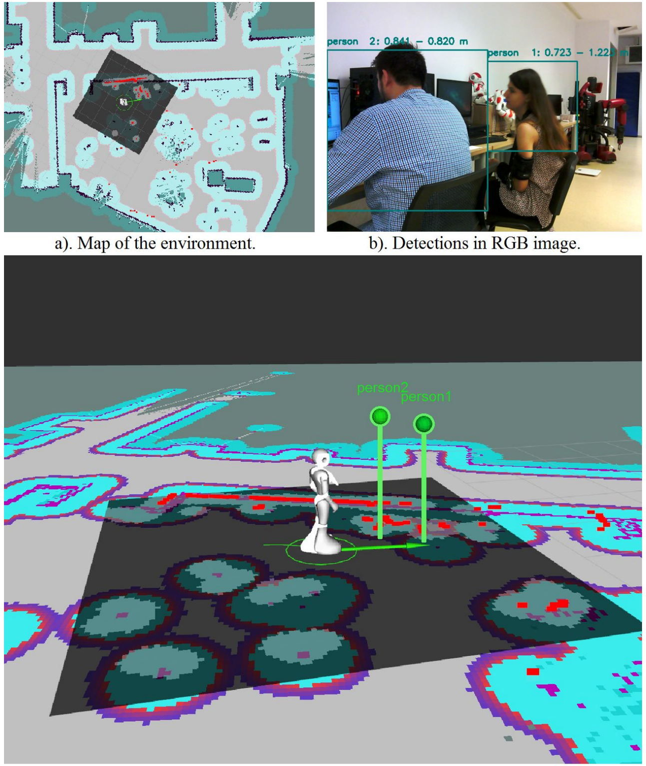




- **Person *Detection*** using YOLOv3 [5]
- Tracking of people using the SORT [6] algorithm
- **Person *Recognition*** using FaceNet [8]
- Pose Recognition using Openpose [9]

[5] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." (2018).
[6] A. Bewley et al., “Simple online and realtime tracking,” ICIP 2016
[8] F. Schroff et al.,, “Facenet: A unified embedding for face recognition and clustering”, CoRR, 2015
[9] Z. Cao et al. "Realtime multi-person 2d pose estimation using part affinity fields.", ICCV, 2017.

Human-Centered Perception: People Localization



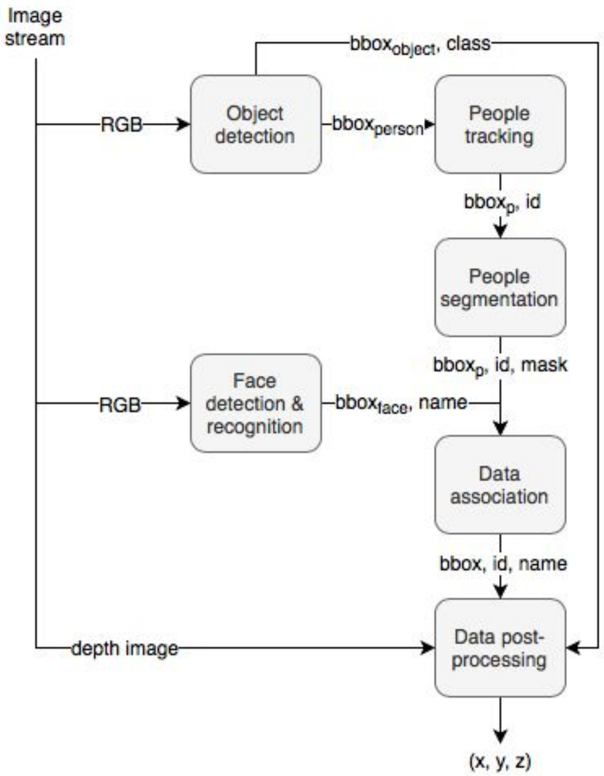
a). Map of the environment.

b). Detections in RGB image.

c). Detected people placed on the map.

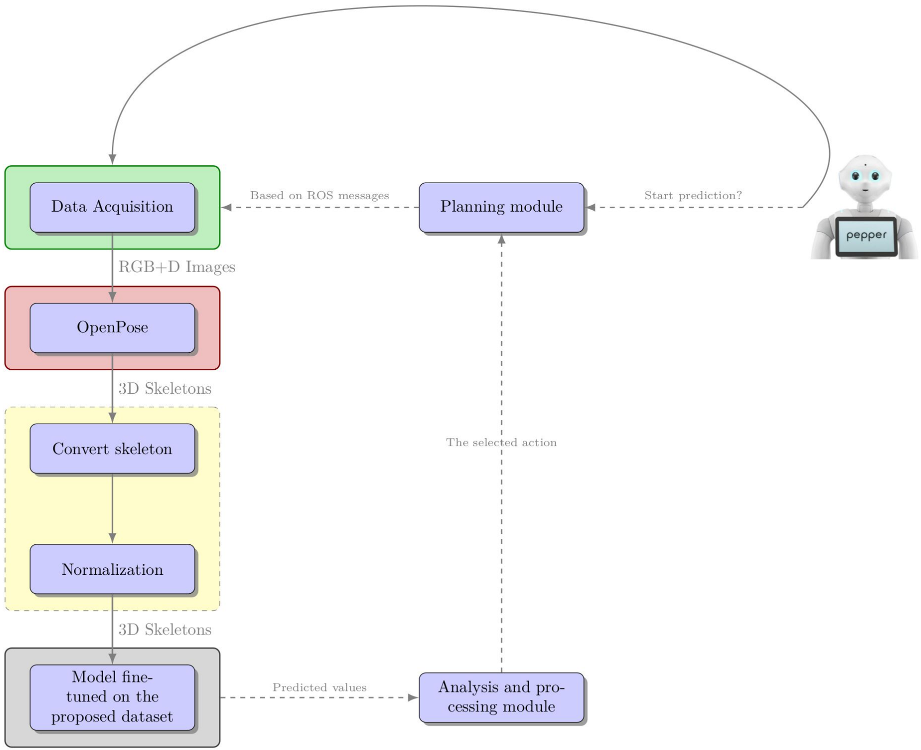
- AMIRO Framework combines information from Person Identification, Depth Mapping + amcl Localization services to save *last seen* position of recognized people on the map

Human-Centered Perception in AMIRO - Vision Pipeline



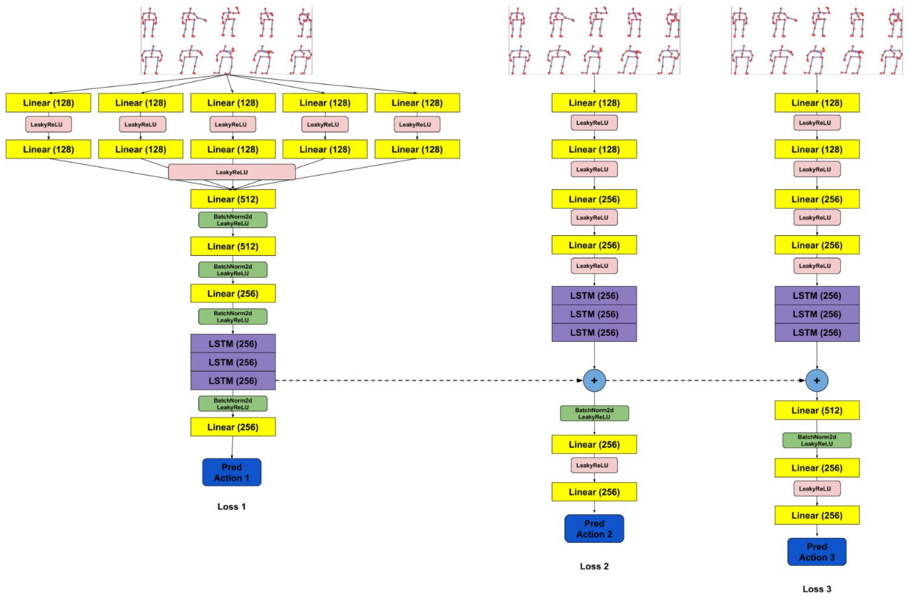
- **Entire processing pipeline** (object det. + people det. + people/object tracking + people rec. + pose estimation) yields a 3 fps throughput

Human-Centered Perception: Action Recognition



- The Action Recognition module allows recognizing common human ADL actions such as: walking, standing, sitting, drinking, typing, pointing etc.

Human-Centered Perception: Action Recognition



- The Action Recognition module allows recognizing common human ADL actions such as: walking, standing, sitting, drinking, typing, pointing etc.
- Module uses an in-house action recognition model based on human *"skeleton"* data (model pre-trained on NTU RGB+D dataset [9]

[9] https://rose1.ntu.edu.sg/dataset/actionRecognition/

Interaction / Planning in AMIRO: Dialogue Management

NLU Input (user's command ASR result): Display my blood pressure.

NLU Output (partial):

```
{
    "intent": "get_health",
     "entities":
     {
          "health_entity": "blood pressure",
          " output_entity ": "display"
     }
  ....
}
```

Story 1 *(story name)*

* greet *(recognized intent)*
 - utter_greet *(system answer/action)*

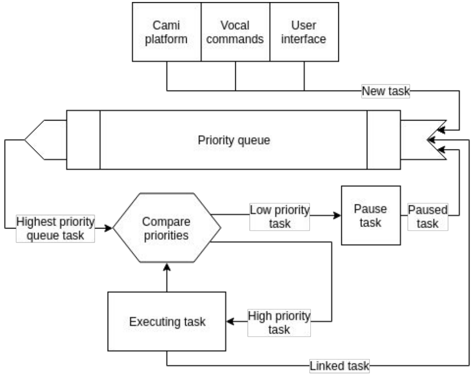
* get_health *(recognized intent)*
 - utter_health *(system answer/action)*

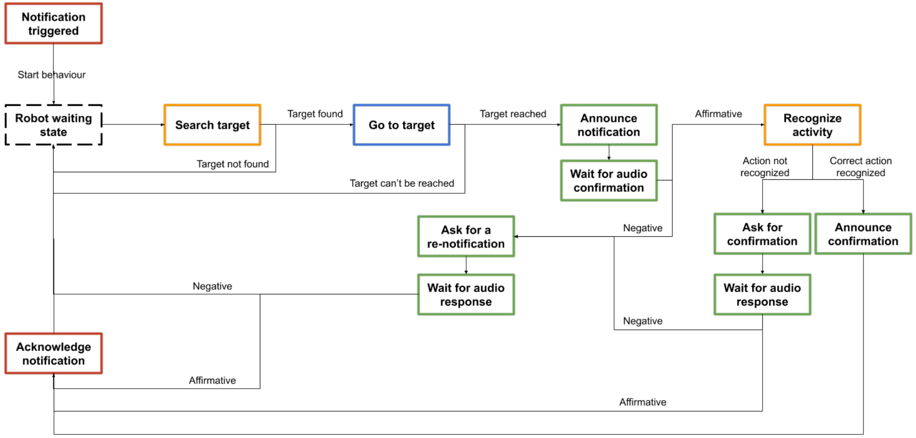
* goodbye *(recognized intent)*
 - utter_goodbye *(system answer/action)*

- Dialogue Management in AMIRO uses services for each step in a conversation:
  - Local processing to **detect utterance** (active listening on microphone)
  - Google Cloud Speech Recognition API [10] for **Speech-to-Text** (works for English and Romanian)
  - Wit.ai [11] to perform **Natural Language Understanding (NLU)** → recognize speaker **intent**
  - RASA [12] for **conversation management**

[10] https://cloud.google.com/speech-to-text
[11] https://wit.ai/
[12] https://rasa.com/

Interaction / Planning in AMIRO: Behavior Management

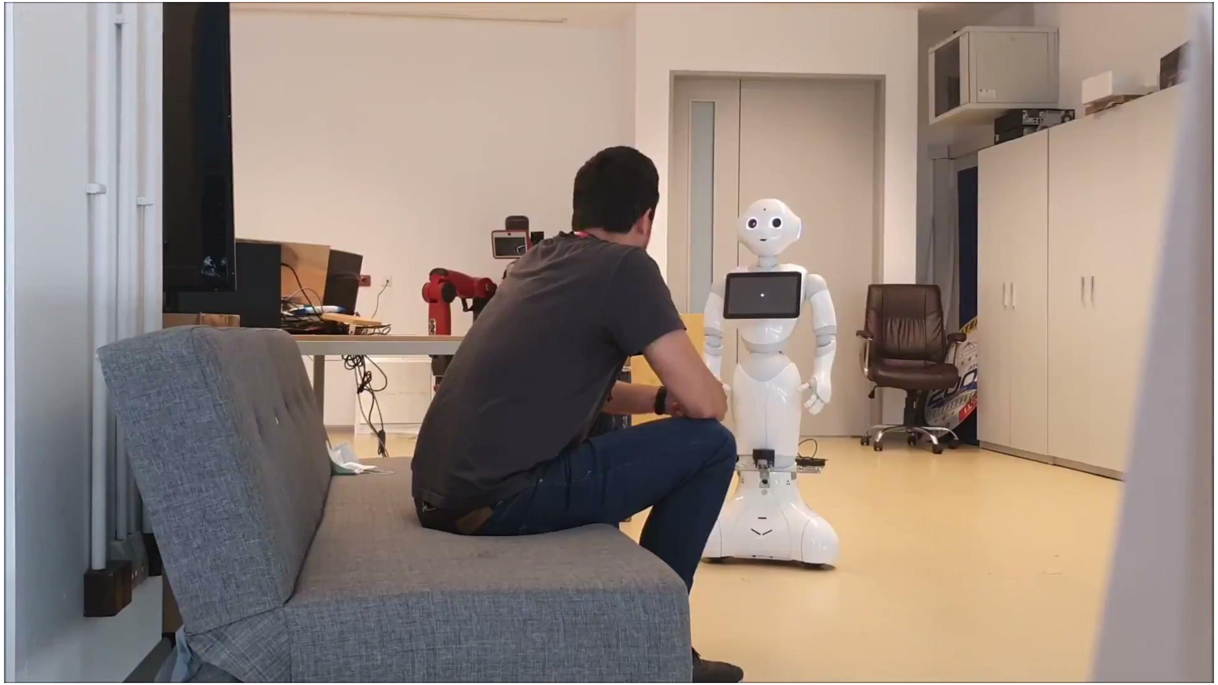




- Behavior Management is implemented using a priority-queue based task manager (works like a preemptable state machine)
  - Tasks have *success* and *failure* continuations
  - Tasks can be paused (when a higher priority task is inserted into the queue)

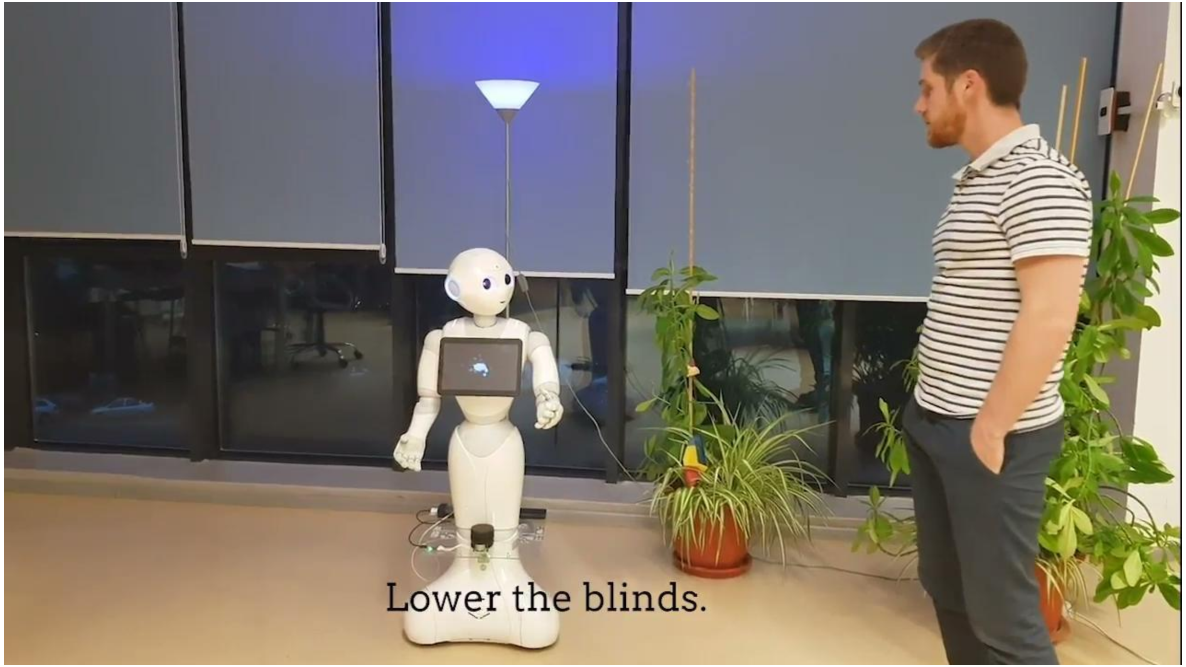
AMIRO in Action

AMIRO in Action: Find a Person + Recognize Action



AMIRO in Action: Notification Reminders

AMIRO in Action: Putting it all together



AMIRO on Pepper: the challenges behind the scenes

- Vision Modules (people detection, action recognition) **highly** dependent on lighting conditions and camera position
- Noisy robot microphone; it's a challenge to have a group of people interact with the robot
- Continual life cycle still an ongoing challenge
  - AMIRO allows chaining together interaction episodes, but a "global" robot behavior is still missing
- Robot navigation requires "special" arrangements (sufficient distance from obstacles - more than 30cm, wide open doors etc.)
- Network Bandwidth was an unexpected bottleneck :-)

AMIRO: going forward - it's still technical

- Train vision and action recognition models on a much wider set of use cases and environment conditions
- Work on implementing *single user* vs *group* interaction modes
- Augment robot navigation with lifelong SLAM and *semantic mapping*
- Extend the behavior management module to have a *default proactive* state
- More research into local vs edge vs cloud module deployments to increase re

Social Robotics

The Summary

Social Robotics: The Summary

- Social Robotics Research is most active in Healthcare/Assistance, Education and Entertainment
- Social Robotics Research is *still* dealing with the technical challenges and awaits a look into the *social* aspects
  - Social robotics requires believable capabilities of navigation, communication, environment perception, long-lived interaction to be accepted by humans
  - Currently limited to individual scenarios that can be handled well (e.g. in retail, healthcare, education)
- The large amount of required ML/AI models that need to be integrated give rise to new deployment models such as Cloud Robotics

Thank you!

alexandru.sorici@upb.ro
AI-MAS Lab