

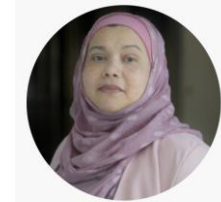


SLES Synergy Project, IIS 2417003; 09/2024 – 08/2028

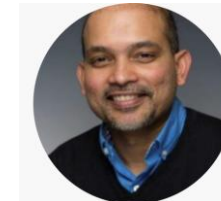


# Learning Safe Policy from Human Demonstrations to Support Robot-Assisted Aging-in-Place

PI: Momotaz Begum, Robotics-LfD



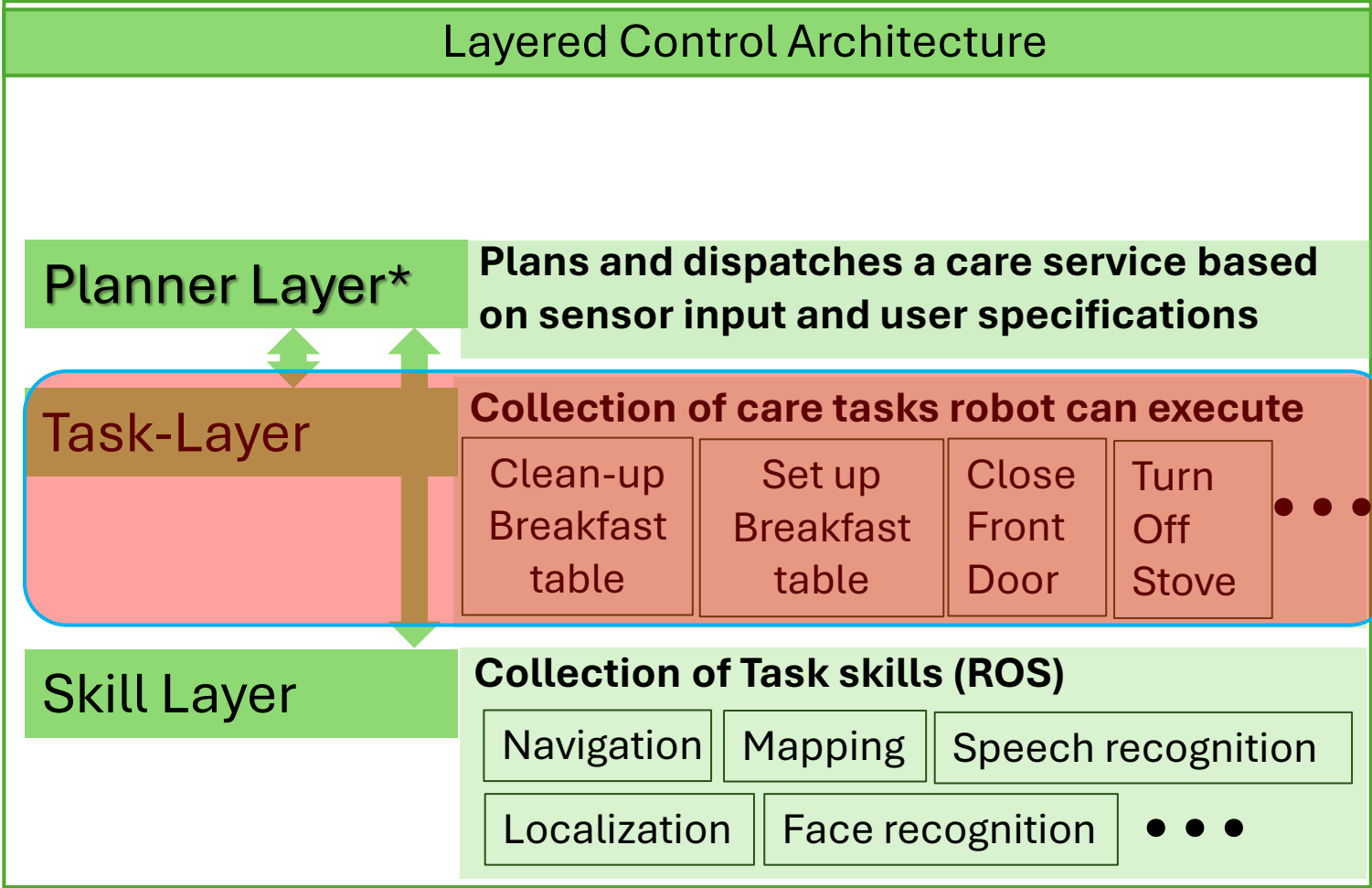
Co-PI: Sajay Arthanat, Gerontechnology



Co-PI: Dain LaRoche, Exercise Science







**Practical use cases for LfD/Imitation Learning**

**IL policy has no notion of safety!**

\*Good, old STRIP-style planning; Hardcoded PDDL

## Layered Control Architecture for a SLES

**Safe-LfD**

Learns from user how to execute a task safely

Planner Layer

Plans and dispatches a care service based on sensor input and user specifications

Task-Layer

Collection of care tasks robot can execute

Clean-up  
Breakfast  
table

Set up  
Breakfast  
table

Close  
Front  
Door

Turn  
Off  
Stove

Skill Layer

Collection of Task skills (ROS)

Navigation

Mapping

Speech recognition

Localization

Face recognition

End-user training

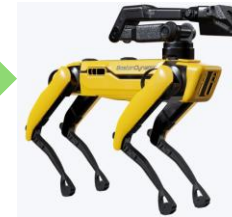


**Safe-LfD:**  
Cognizant, Informed

Caregiver Demonstration

A run-time  
scenario

Hey! it's  
breakfast  
time



*Field Imitation Learning*

## *What are the constraints in Field Imitation Learning?*

- Demonstrations are limited, even a dozen may be too many
- There is no simulator in the field

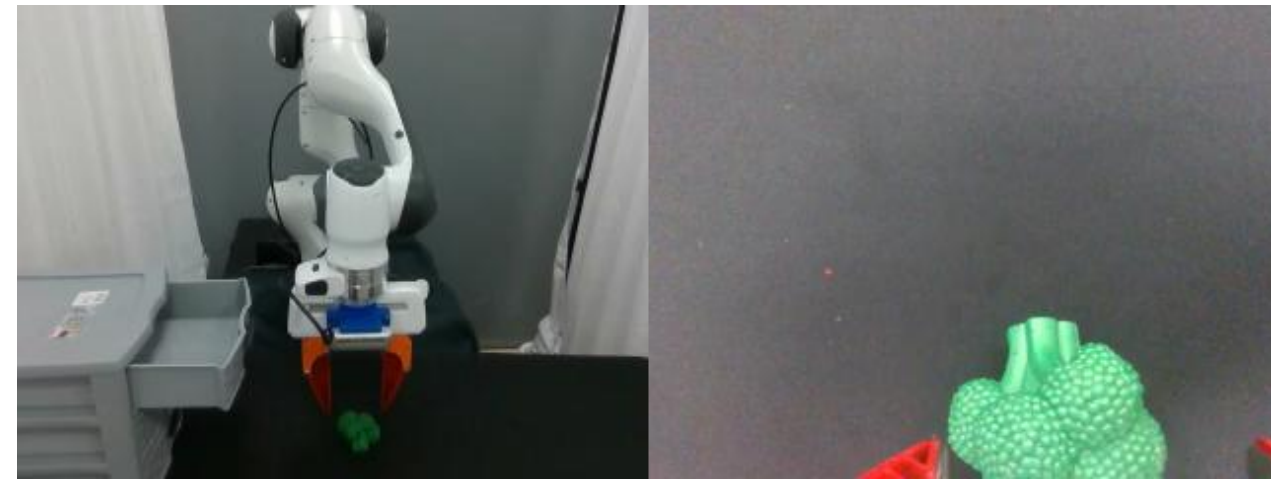
# *What are the Safety concerns in Field Imitation Learning?*

## *Things that we know..*

- End-users make mistakes as they give demonstrations; Policies learned from those demonstrations may be unsafe



End-user demonstrations



Learned diffusion policy (80+20 demo)

# *What are the Safety concerns in Field Imitation Learning?*

*Therefore, the things that we can do..*

- Identify bad demo/bad segments in a demo and discard those before policy learning



Learned diffusion policy after data curation  
(Under review, RSS 2025)

Learned diffusion policy (80+20 demo)

# *What are the Safety concerns in Field Imitation Learning?*

*However, we still don't know how to..*

- Learn the **best** policy while discarding the least amount of data from a demonstration set automatically
  - The best policy:
    - perceived safety
    - task success
    - resilience to run-time changes (OOD)
- Learn policies for long-horizon tasks from a small # of demo



# *What are the Safety concerns in Field Imitation Learning?*

## *Things that we know..*

- There is a zero-tolerance for policy error; don't act if you are not sure. Don't harm the property, don't hurt yourself



Learned BC-RNN policy (IROS 2024)



## *What are the Safety concerns in Field Imitation Learning?*

*Therefore, the things that we can do..*

- Offer guarantee that the policy can reach the (sub)goal location
  - Synthesizing a maximum control invariant set from a set of demonstration

## *What are the Safety concerns in Field Imitation Learning?*

*However, we still don't know how to..*

- Offer guarantee that a manipulator will correctly execute the task after reaching the goal location.

# *What are the Safety concerns in Field Imitation Learning?*

*And we kind of don't know..*

- What else makes a policy safe to end-users?
- How can we extract safety constraints that humans naturally pass into the demonstrations.

# *What are the Safety concerns in Field Imitation Learning?*

*And we kind of don't know..*

- What else makes a policy safe to end-users?
- How can we extract safety constraints that humans naturally pass into the demonstrations.
- What is better? -- learning a policy considering safety constraints vs employing a safety filter to prevent unsafe actions

*And the last but not the least..*

- Field deployment

