Performance of **Autonomy** and **Identity** for Trust- and Workload-Sensitive Interaction with Distributed Autonomous Systems

Dr. Thomas E. Williams
Colorado School of Mines

Research Objectives

Objective 1: Enable optimal *performance of autonomy* to manage trust and workload. The state of the art in robot autonomy differentiates between *adaptive autonomy* and *mixed initiative interaction* approaches. We bridge these perspectives through an approach grounded in Speech Act Theory, in which agents to make mixed-initiative dialogue moves that **selectively perform low autonomy to influence trust and workload.**

Approach

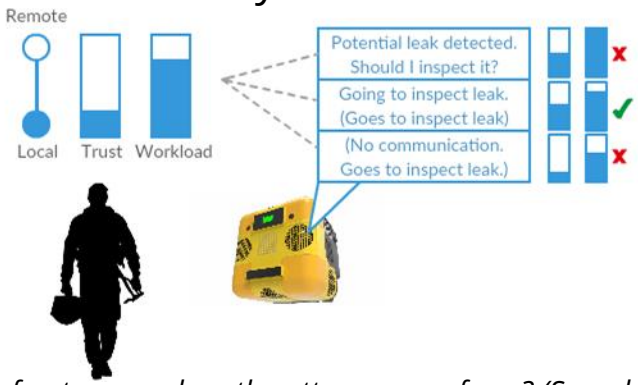
Thrust 1: Autonomy Performance

- Phase 1: **Empirically investigate** autonomy performance strategies for proactive explanation under differing (simulated) spatiotemporal distance.
- Phase 2: **Model** autonomy performance strategy selection of using probabilistic graphical models.
- Phase 3: **Evaluate** autonomy performance model in a NASA-informed experimental context.

Thrust 2: Identity Performance

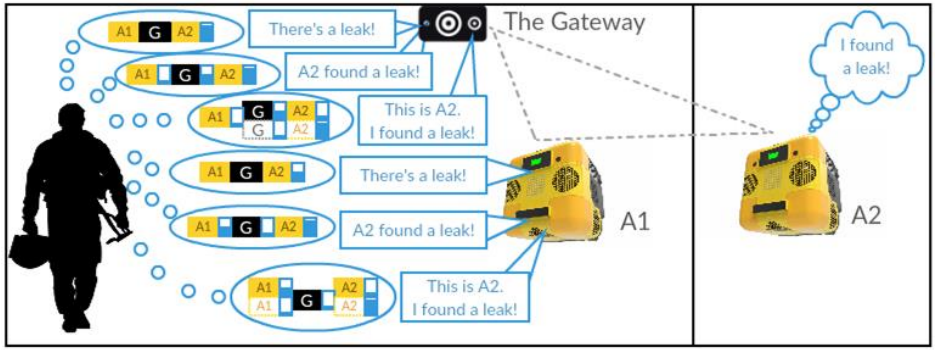
- Phase 1: **Prototype** identity performance strategies using Utterance representations that distinguish between speaker body and identity.
- Phase 2: **Empirically investigate** how these strategies impact trust resilience.
- Phase 3: **Analyze** our results to model *trust localization, dissociation, and fragmentation* and produce identity performance *design guidelines*.

Autonomy Performance



What level of autonomy does the utterance perform? (Speech bubbles) Does this increase or decrease trust and workload (bars on right) given current trust, workload, and distance (bars on left)?

Identity Performance



What body and identity are used to communicate? (Speech bubbles) Is trust placed in the group (Utterances 1,4) or individuals (2,3,5,6)? Do body- and identity-based trust dissociate (solid/dashed boxes in 3,6)? How does this fragmentation affect trust overall?

Objective 2: Enable optimal *performance of identity* to promote resilient trust

The state of the art in modeling human-robot trust assumes body-identity alignment. Because distributed, integrated, autonomous systems are comprised of multiple bodies with a single "mind", dialogue strategies may dissociate body and identity. We propose to study the influence of **identity-performing dialogue policies** using an innovative model of trust that accounts for *trust localization, dissociation, and fragmentation*, which may occur when body-identity alignment fails.

Because this research will develop fundamental new theories, it will take place at TRL1. By prototyping strategies and computational models, our goal is to transition to TRL2 by project completion.

Potential Impact

This research will enable **distributed, integrated, autonomous systems** to communicate with local teammates (e.g., astronauts) and remote teammates (e.g., ground control workers) in a way that better...

- **promotes appropriate levels of trust, low cognitive workload, and strong human-machine social relationships.**
- **maintains sensitivity to spatiotemporal distance**
- **avoids uncanny valley effects**