Human Autonomy Teaming Architecture (Philosophy)

Jay Shively

NASA Ames Research Center

Wednesday, July 28, 2021

Outline

- Human Autonomy Teaming (HAT)
  - What is it?
  - Why?
  - How?

- Principles
  - Bi-directional communication
  - Transparency
  - Operator directed interface
  - Shared situation awareness
  - Meaningful Human Control (MHC)

- Applications
  - Reduced Crew Operations
  - Multi-Vehicle Control
  - High Density Vertiplex
  - Wildfire Mitigation

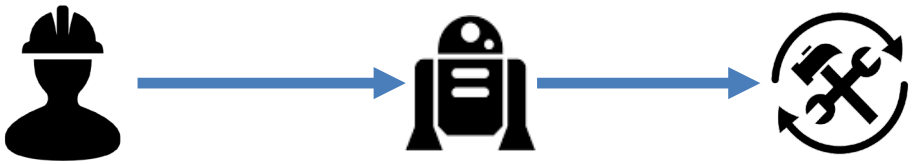
System Configurations

Human & System

Human interacts directly with the work system.

Human & AI & System

AI is between the human and the work system



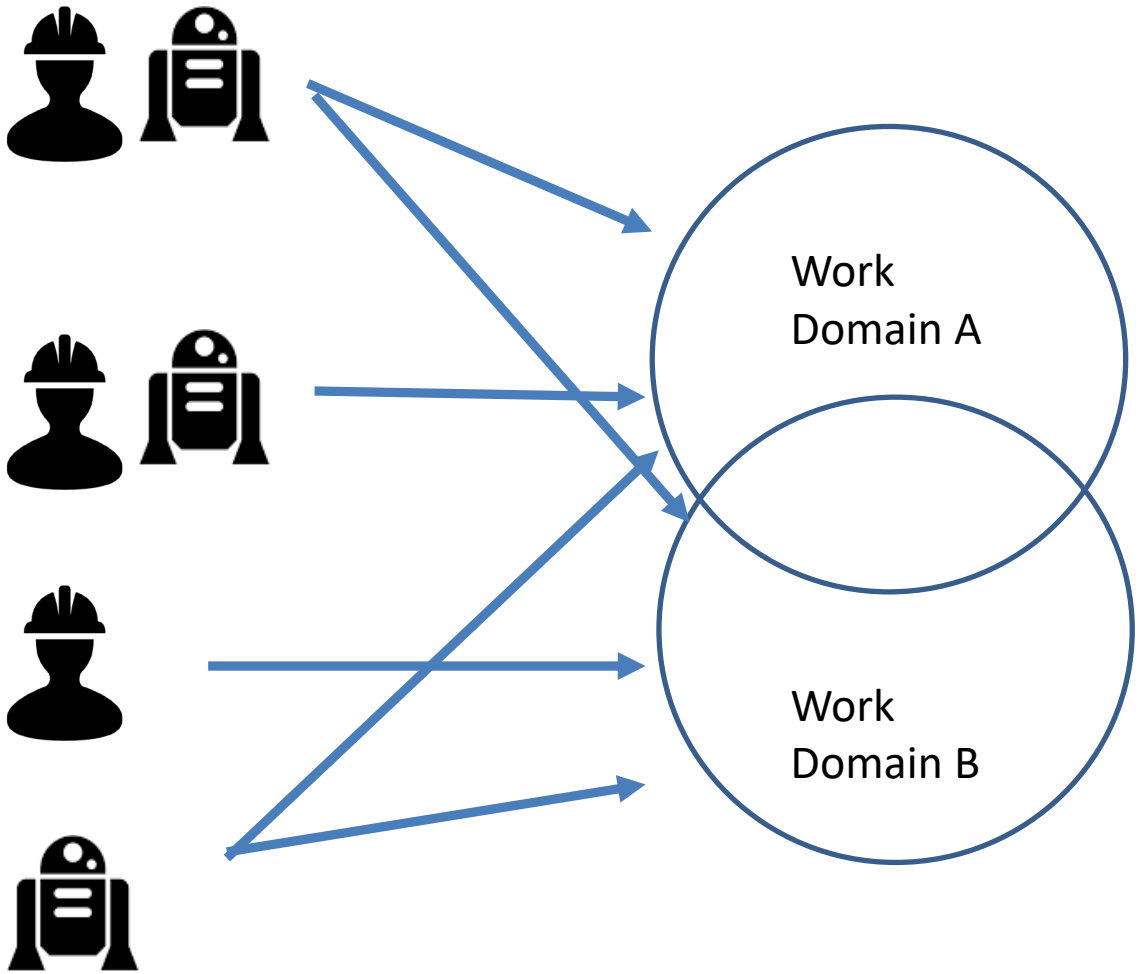
Problems with Serial Arrangement

- Brittle
  - Automation often operates well for a range of situations but requires human intervention to handle boundary conditions (Woods & Cook, 2006)
- Opaque
  - Automation interfaces often do not facilitate understanding or tracking of the system (Lyons, 2013)
- Mis-calibrated Trust
  - Disuse and misuse of automation have led to real-world mishaps and tragedies (Lee & See, 2004; Lyons & Stokes, 2012)
- Out–of-the-Loop Loss of Situation Awareness
  - Trade-off: automation helps manual performance and workload but recovering from automation failure is often worse (Endsley, 2016; Onnasch, Wickens, Li, Manzey, 2014)

Human-AI Teaming (Simple Concept)

Human and AI “team” to interact with the work system.

Real World System of Systems



Principles (Not Exhaustive)

- Bi-directional Communication
- Transparency
- Shared situation awareness (mental model)
- Pilot directed interface
- Meaningful Human Control

Bi-Directional Communication

- Humans and AI should have a channel to communicate
- Human can ask about reasoning:
  - Why did you choose A?
  - What about B?
- About confidence:
  - How sure are you of A?
  - How close was B?
- Human can input information that the AI doesn't have
  - e.g., alternative C doesn't have good emergency medical response

Bi-Directional Communication (cont.)

- Human should be able to bounce ideas off the AI (and vice versa) like a TEAMMATE
- These capabilities do levy requirements on the AI
  - Common/understandable comm mode (e.g., cockpit control language)
  - Natural language interface (desired but not required)
  - Explainability
  - Self-awareness? Level of confidence

Transparency

- Human should be able to understand the reasoning for AI recommendations/decisions
  - Level 1: What is the system doing?
  - Level 2: Why is the system doing that?
  - Level 3: What can I expect to happen next?
- Critical component in fostering **calibrated trust** of AI
- Calibrated trust fosters appropriate utilization/reliance behaviors between teammates

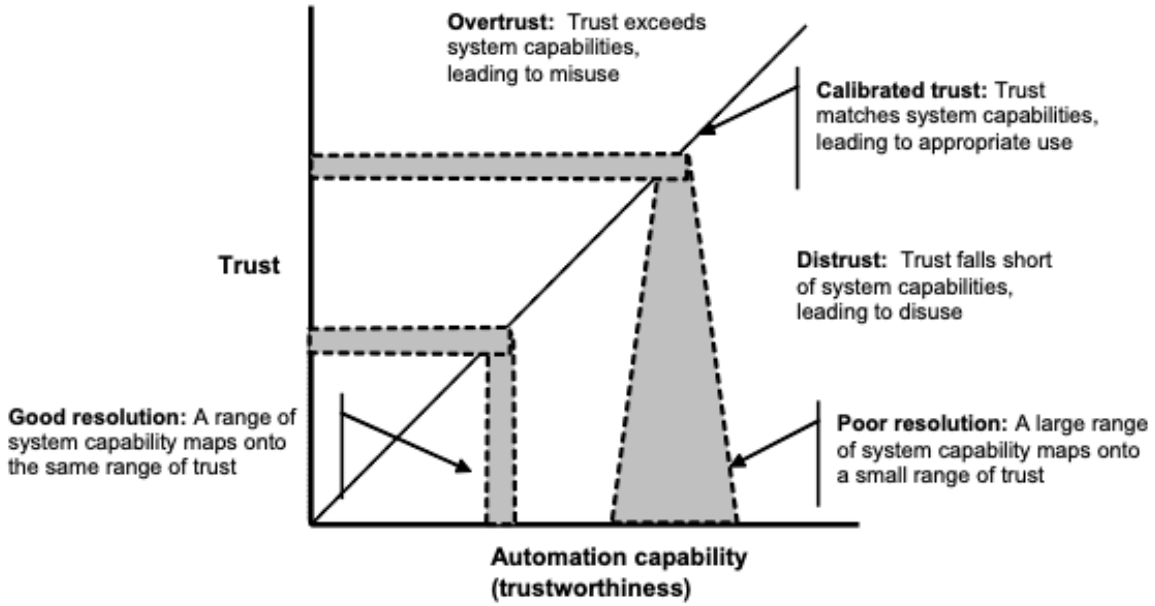


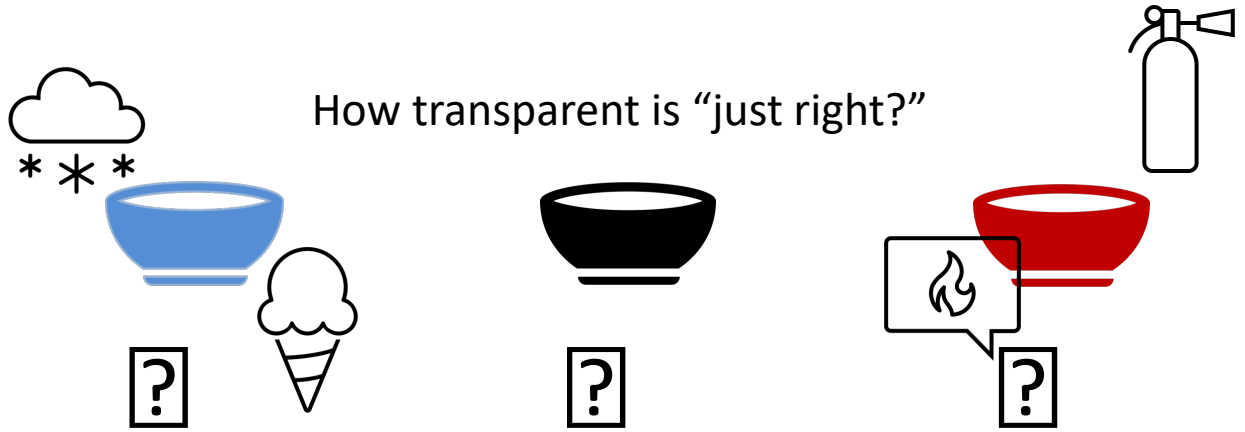
Figure from: Lee, J., See, K. (2004) Trust in Automation Designing for Appropriate Reliance. *Human Factors, 46*(1), 50-80.

Transparency: Issues

- **Granularity:** how much detail is enough but not too much?
- **Time pressure:** processing information takes time; how much time is available?

Transparency Goldilocks:

How transparent is "just right?"



TOO COLD

Too Opaque

↓

Confused, uninformed, misinformed pilot

TOO HOT

Too Transparent

↓

Information Overload

Operator Directed Interfaces

- No set “roles and responsibilities”
  - Default assignments for nominal operations
- Fluid assignment set by operator actions
  - Take control of stick and fly aircraft
- Requires some degree of ”Intent Inferencing” – can be problematic
- Can be difficult to implement in highly automated systems

Shared Situation Awareness

- Humans and AI should have:
  - A shared model of system operation
  - A shared understanding of external influences to the system
  - A shared understanding of the current situation
    - State of the system
    - State of the environment
  - How do you ensure?
  - How do you transfer info, if not available?
  - How do you measure (continually)?

Meaningful Human Control

- Humans should have the ability to exert “meaningful” control
  - Not just hit the red button
- Take advantage of human decision making and analytical skills
- Ethical, moral, humanistic reasons

MHC

- How do you define?
- How do you measure?
- Where in the design-operation cycle?
- Human limitations

Humans have the ability to make informed choices in sufficient time to influence automation-based systems; these can enable a desired effect or to prevent an undesired immediate or future effect on the environment.

Please write an 'X' in the box that best represents your opinion in each of the dimensions below:

Range of Options: *Did you have the range of response options required to respond as needed?*

1----------2----------3----------4----------5----------6----------7

Not at all Moderate Absolutely

Temporal Availability: *Did you have the time to assess the situation and respond as required?*

1----------2----------3----------4----------5----------6----------7

Not at all Moderate Absolutely

Interface Layout: *Did interface elements support an efficient and effective workflow?*

1----------2----------3----------4----------5----------6----------7

Not at all Moderate Absolutely

Information Availability: *Was the information that you needed to respond available?*

1----------2----------3----------4----------5----------6----------7

Not at all Moderate Absolutely

Workload: *Was your workload low enough for you to respond appropriately?*

1----------2----------3----------4----------5----------6----------7

Not at all Moderate Absolutely

Overall MHC: *Did you feel you were able to exert meaningful human control?*

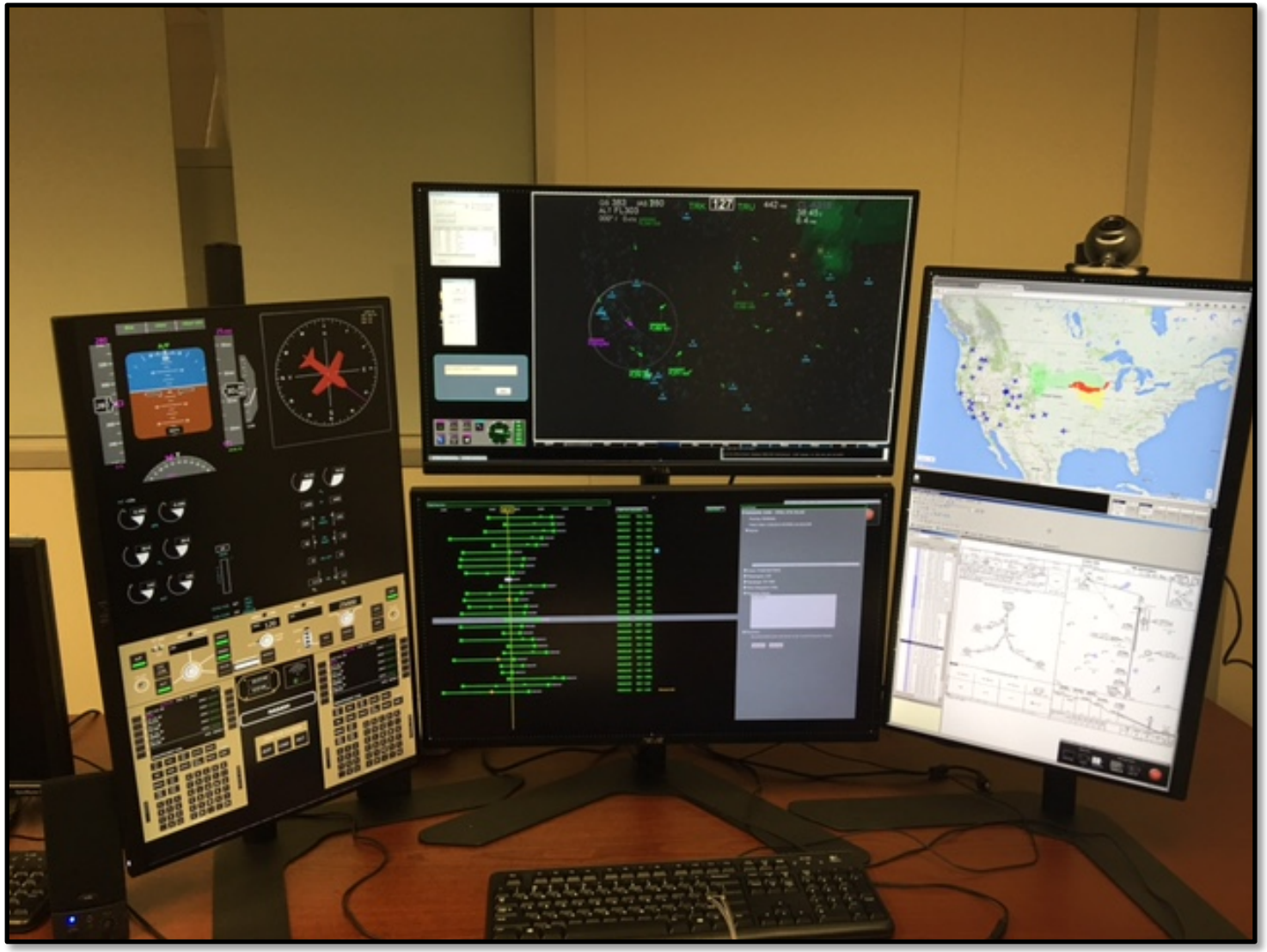
1----------2----------3----------4----------5----------6----------7

Not at all Moderate Absolutely

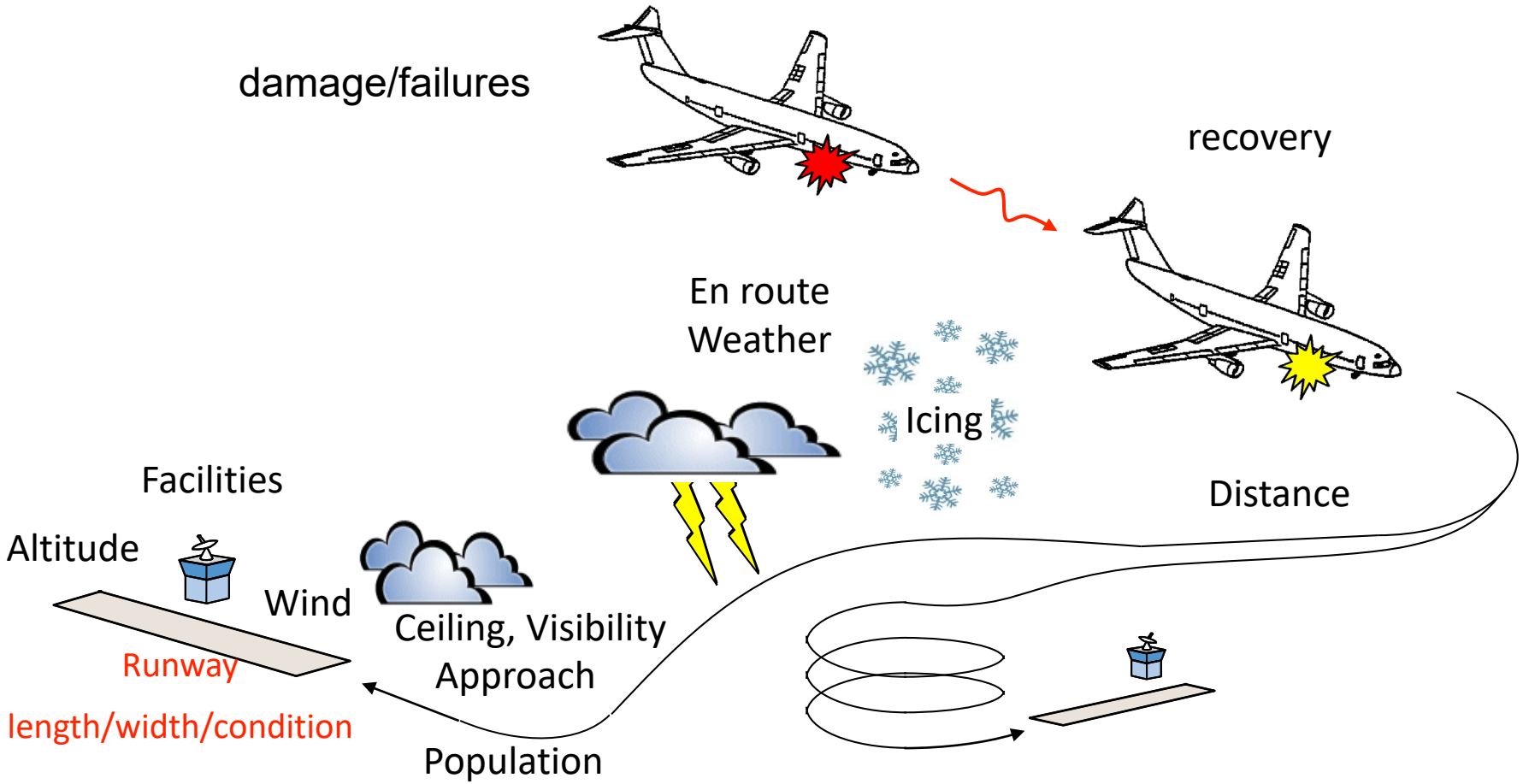
Applications

- Reduced Crew Operations
  - Reduced Costs
  - Pilot Shortage
- m:N Vehicle Control
  - Commercial viability
  - Multiple domains (UAM, drone delivery, High Altitude Pseudo Satellite, infrastructure inspection, auto cargo)
- Advanced Air Mobility (AAM)
  - High Density Vertiplex
- Wildfire Mitigation
  - Human failsafe

Reduced Crew Operations



Emergency Landing Planner Objective

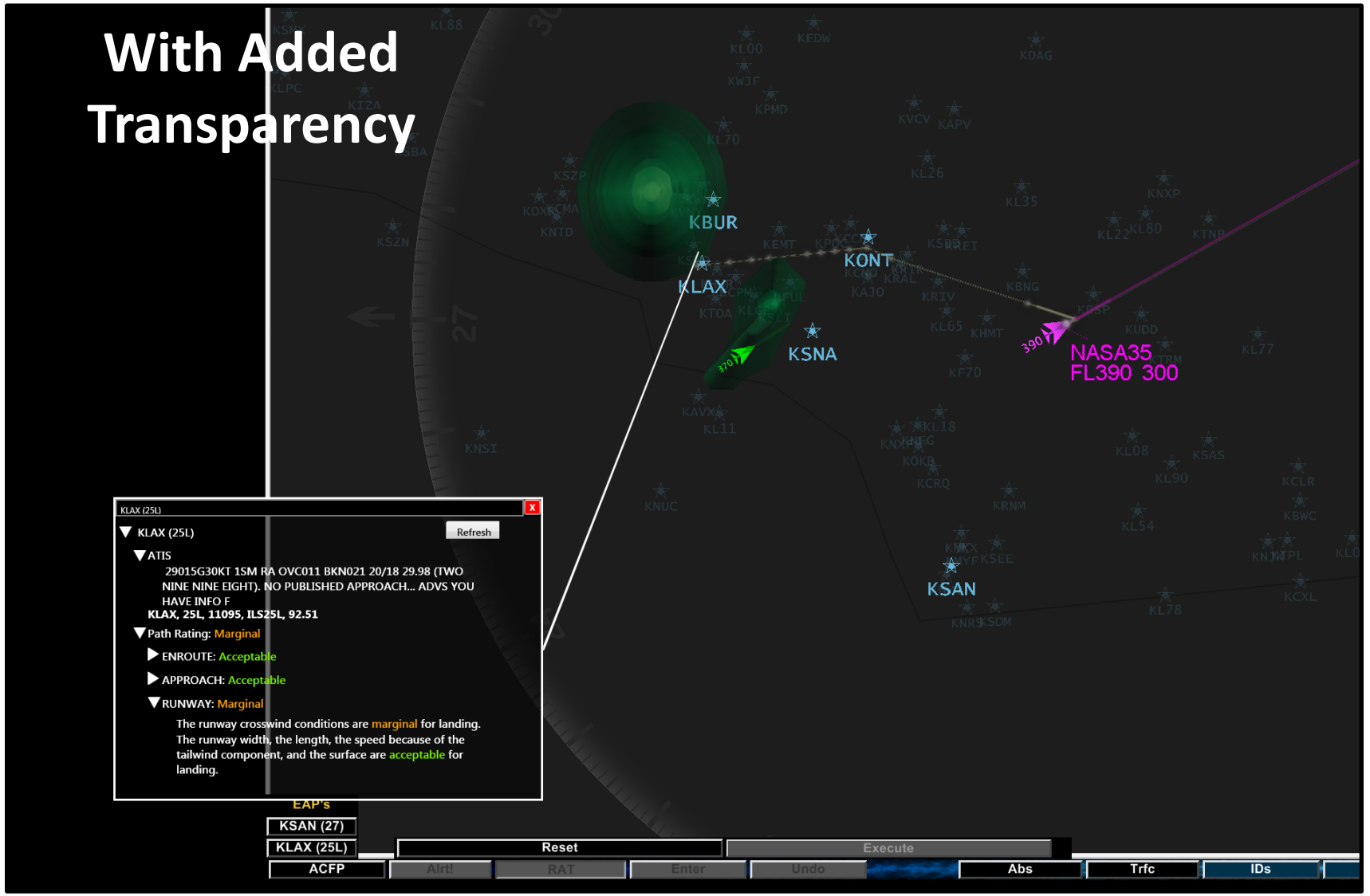


Find the best landing sites and routes for the aircraft

Before HAT

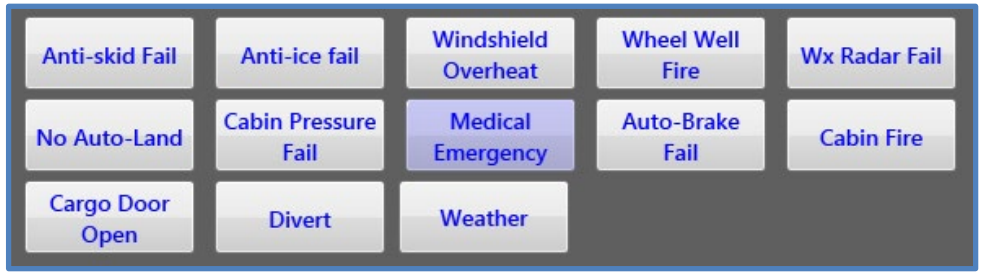
Recommended airports - rank ordered.

Adding HAT Principles to the Ground Station

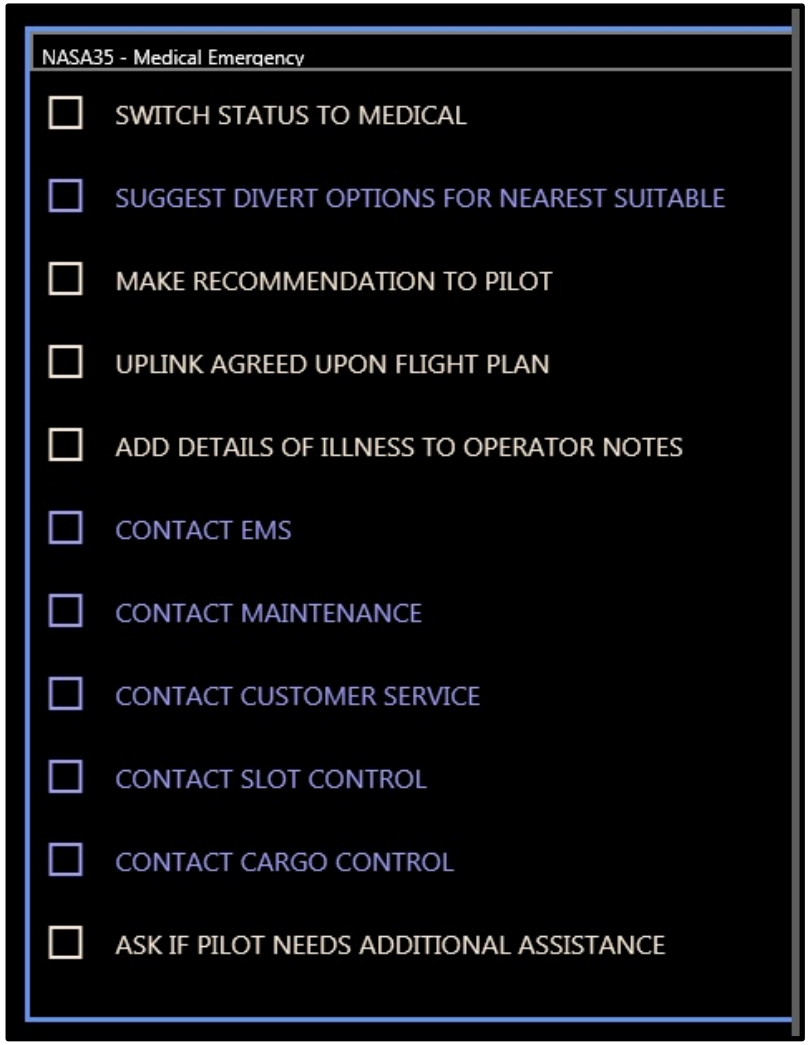


Adding HAT Principles to the Ground Station

- Human-Directed: Operator calls “Plays” to delegate authority to automation

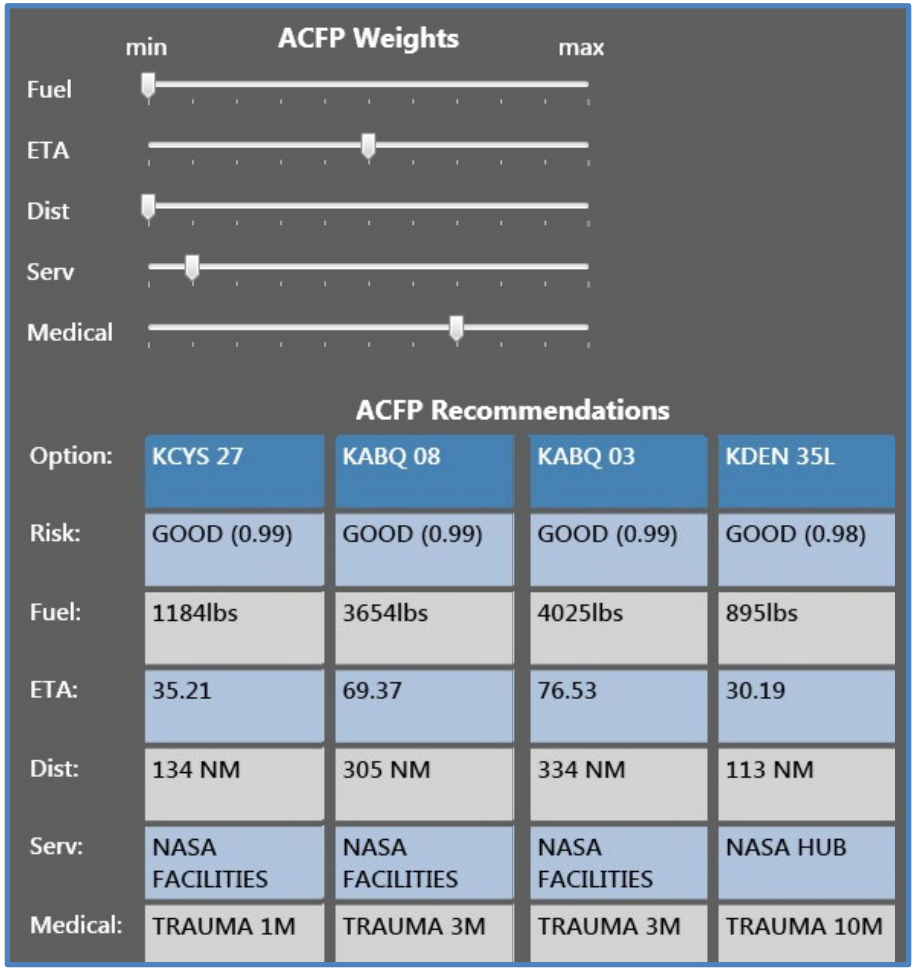


A play encapsulates a plan for achieving a goal.
It includes roles and responsibilities
what is the automation going to do
what is the operator going to do



Adding HAT Principles to the Ground Station

- Transparency: Divert reasoning and factor weights are displayed.
- Negotiation/Dialog: Operators can change factor weights to match their priorities.
- Shared Language/Communication: Numeric output from planner was found to be misleading by pilots. Display now uses English categorical descriptions.



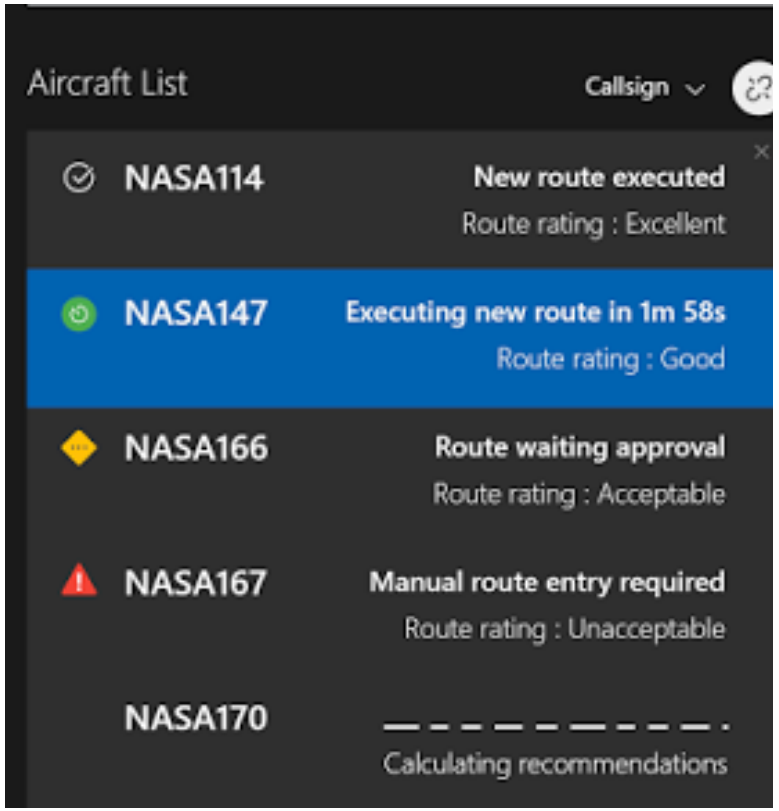
ACFP Recommendations

Option:	KCYS 27	KABQ 08	KABQ 03	KDEN 35L
Risk:	GOOD (0.99)	GOOD (0.99)	GOOD (0.99)	GOOD (0.98)
Fuel:	1184lbs	3654lbs	4025lbs	895lbs
ETA:	35.21	69.37	76.53	30.19
Dist:	134 NM	305 NM	334 NM	113 NM
Serv:	NASA FACILITIES	NASA FACILITIES	NASA FACILITIES	NASA HUB
Medical:	TRAUMA 1M	TRAUMA 3M	TRAUMA 3M	TRAUMA 10M

ALTA Action Phase

Based on working agreements, the Agent will do one of the following:

- [Auto] autonomously executes and informs operator
- [Veto] presents a solution which will be autonomously executed unless the operator intervenes
- [Select] presents multiple options for operator selection



Working Agreements: A Path to Full Autonomy

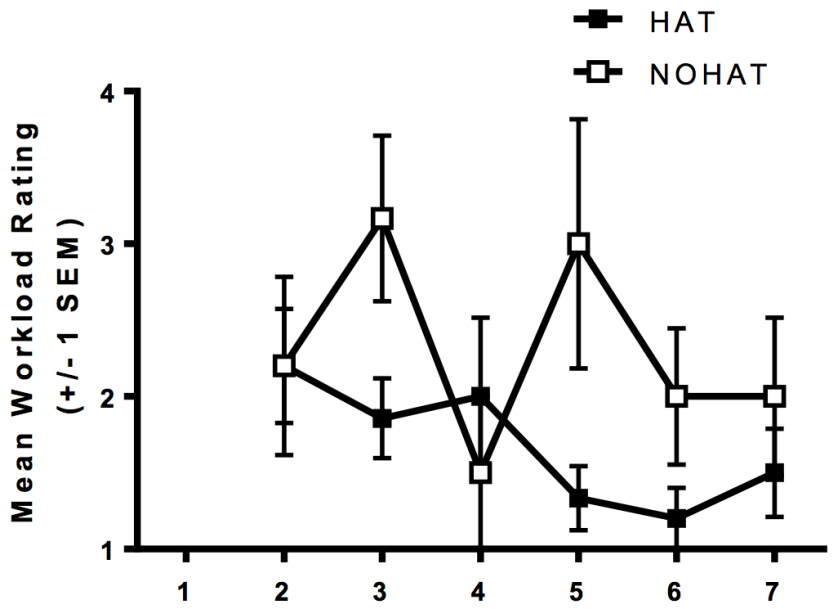
- Codified levels of automation
- Working agreements allow the task structure to remain the same while the involvement of human operators decreases due to improvements in the automation and increases in trust

HAT Simulation: Results

- Participants preferred the HAT condition overall (rated 8.5 out of 9).
- HAT displays and automation preferred for keeping up with operationally important issues (rated 8.67 out of 9)
- HAT displays and automation provided enough situational awareness to complete the task (rated 8.67 out of 9)
- HAT displays and automation reduced the workload relative to no HAT (rated 8.33 out of 9)

HAT Simulation: Results

HAT workload reduction was marginally significant (HAT mean 1.7; No HAT mean 2.3, p = .07)



HAT Simulation: Debrief

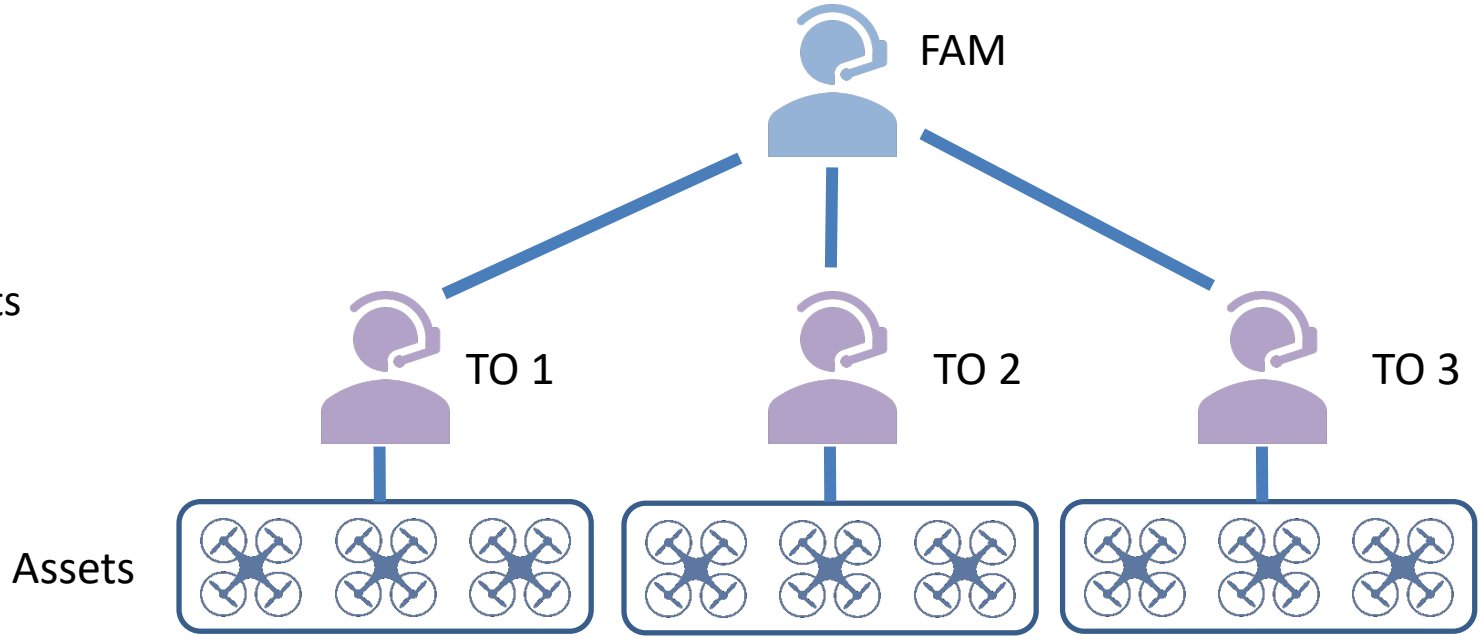
- Transparency
  - "This [the recommendations table] is wonderful…. You would not find a dispatcher who would just be comfortable with making a decision without knowing why."
- Negotiation
  - "The sliders was [sic] awesome, especially because you can customize the route…. I am able to see what the difference was between my decision and [the computer's decision]."
- Human-Directed Plays/Shared Plans
  - "Sometimes [without HAT] I even took my own decisions and forgot to look at the [paper checklist] because I was very busy, but that didn't happen when I had the HAT."

m:N Aircraft Operations

M:N Configuration

- m:N Paradigm
  - Multiple operators to multiple vehicles, where:
    - m = # of operators
    - N = # of vehicles

- Roles in m:N operations:
  - Tactical Operator (TO)
    - Pilot-in-Command (PIC) of multiple assets
    - Monitors airspace and hazards
    - Attends to assets
    - **<u>Participant's role</u>** in the study
  - Fleet Area Manager (FAM)
    - Responsible for overseeing a select number of TOs in a specific area
    - Monitors fleet health and system health
    - Evaluates TO performance and workload online
    - Played by researcher confederate



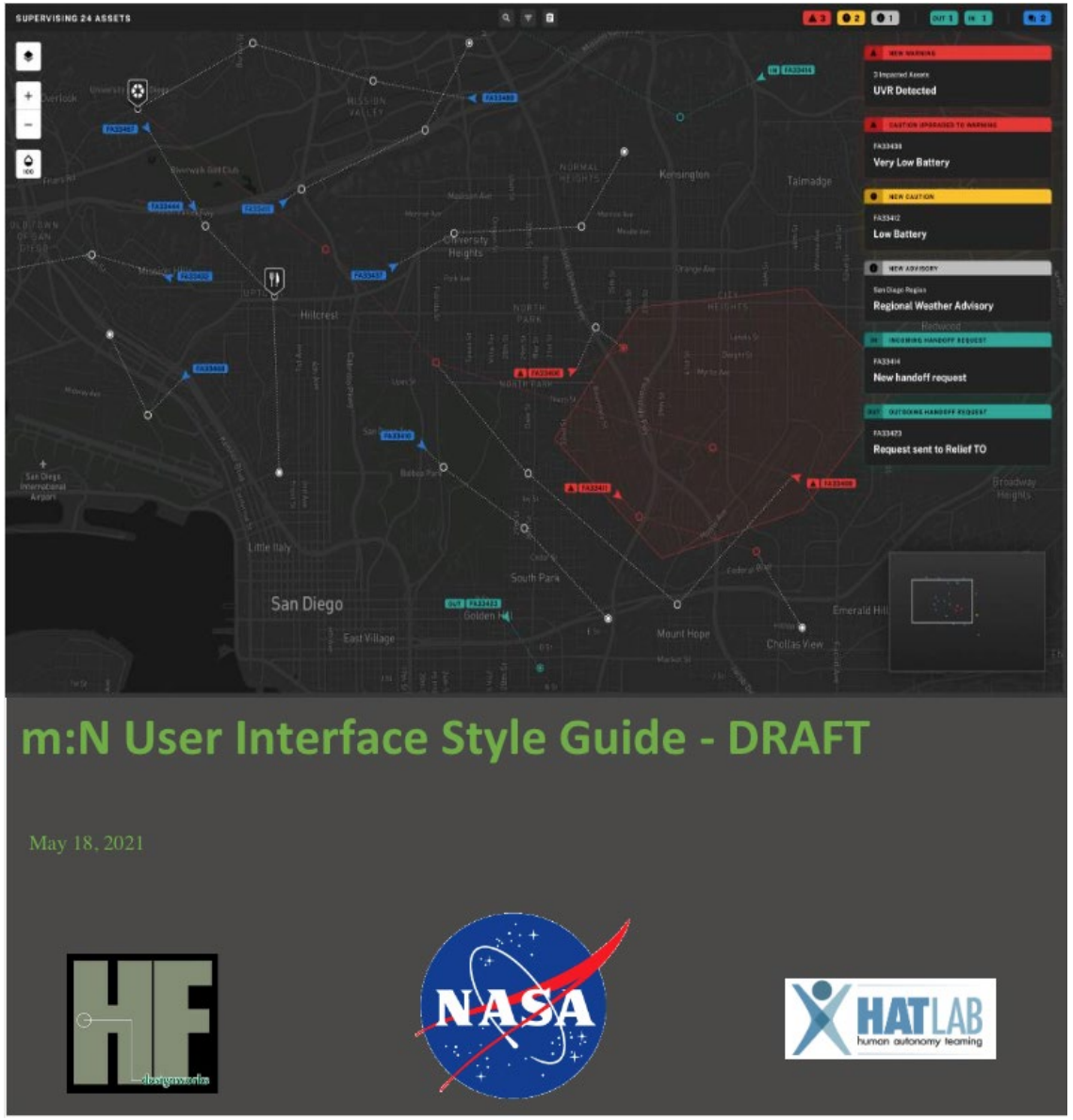
m:N GCS Design

- HAT GCS interfaces for m:N applications currently in design

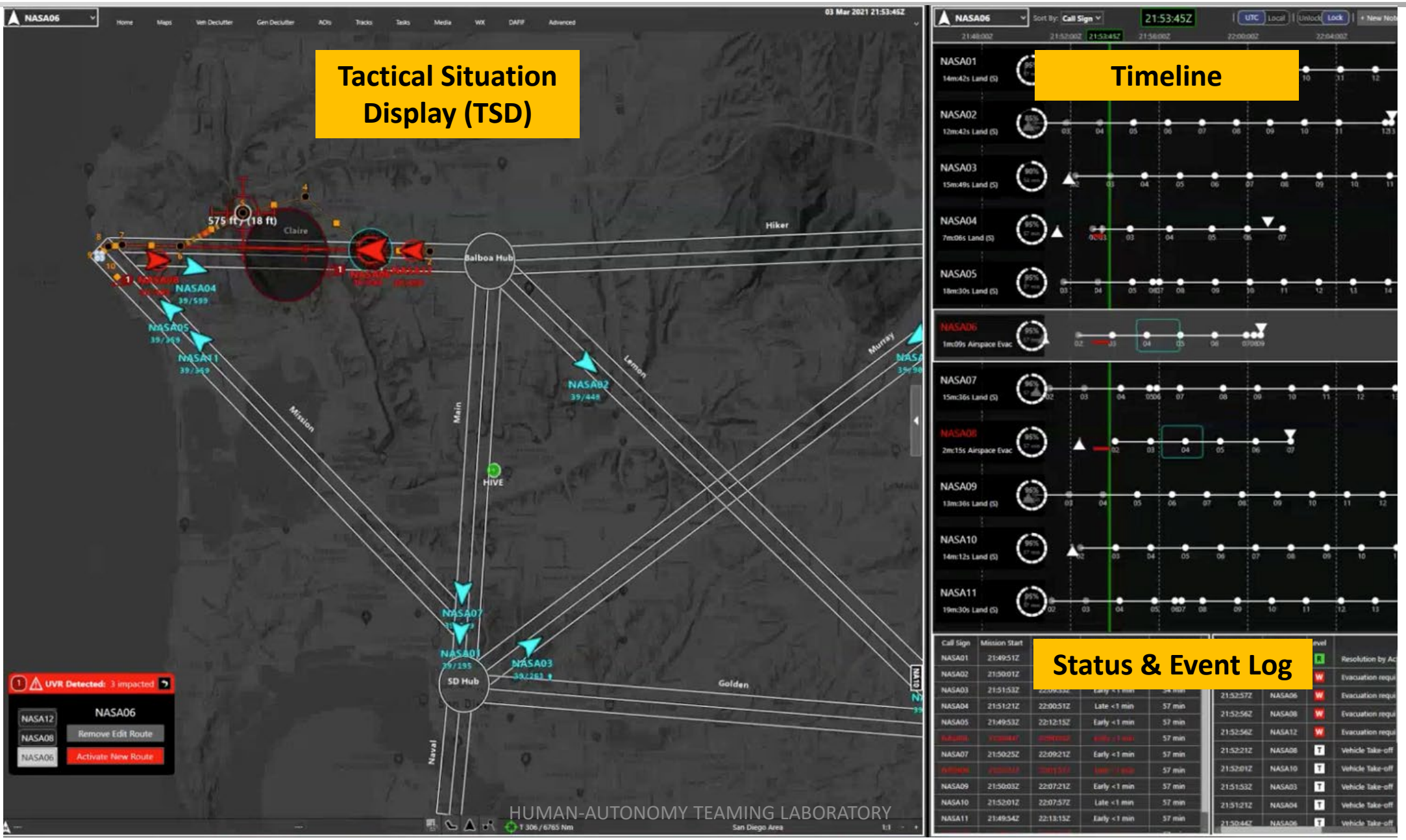
- Design aims to be applicable to multiple m:N domains
  - Small UAS
  - Large UAS
  - UAM
  - HAPS
  - Swarms

- Utilizes HAT principles
  - Bi-directional Communication
  - Transparency
  - MHC (hopefully)

So far…



m:N HITL Work: ConOps/R&R Simulation



Conclusions: m:N ConOps/RR

- Pilots found the task of controlling 12 aircraft as the Tactical Operator manageable
  - Largely were able to avoid airspace violations and terrain collisions
  - ConOps and TO role suitable for the task
    - However, the system was doing a lot of the work: *"Realistically, I was not controlling 12 vehicles; it was 12 autonomous vehicles with me intervening with one vehicle at a time."*

- Suggested reroutes provided in the Automation condition improved performance
  - Decreased Service Time, UVR Resolution Time, initial and final upload times
  - Resulted in shorter routes around UVR area

- Automation condition decreased workload
  - Temporal and Effort TLX scores decreased the most
  - Composite TLX score decreased 34%

- Another outcome of the HITL:
  - Simulation environment shown capable for m:N applications, new (to us) airspaces
  - Single-monitor GCS, compact display
  - Demonstrated subject-surrogate capability

Summary

- HAT Philosophy
  - Not just interface design
  - Not just HCI
  - A paradigm change in the way we view the interaction
- Defined Basic Principles
  - Need to be refined and added to
- Applied to our projects
  - Future of our lab is tied to this philosophy
  - Perhaps automation, not AI, but building the architecture & methodology
- As with most things, simple in principle, harder in practice