34th Annual INCOSE international symposium

hybrid event

Dublin, Ireland
July 2 - 6, 2024

Presented by: Amir Abrari

Co-Authors: Dr. Steven Dam (SPEC), Michael Jordan (SPEC), Lilleigh Stevie (SPEC), Dominic Galarza (SPEC), Dr. Lance Sherry (GMU), Jomana Bashatah (GMU)

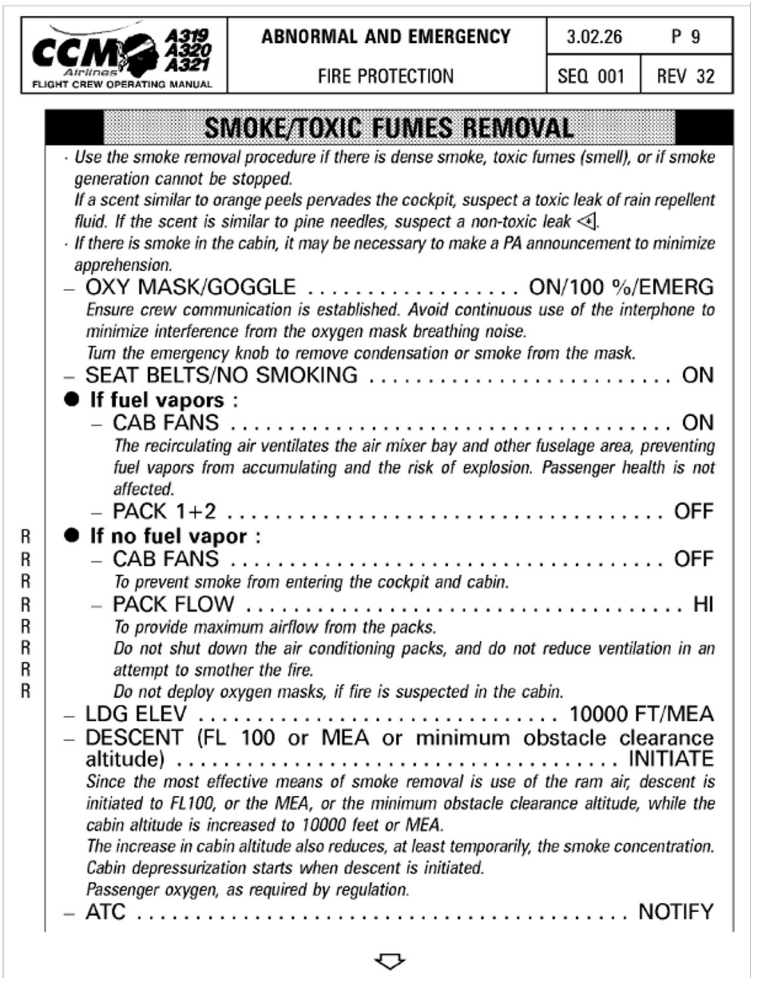
Enabling MBSE for SOPs through Generative AI

Center for Air Transportation Systems Research at George Mason University

Standard Operating Procedures (SOP)

Standard Operating Procedures (SOPs)

- Set of formal guidelines that define the execution of routine and emergency activities
  - Heavily used in safety-critical domains
- Procedural Steps
  - **Define the particular order** of actions to meet the end goal of the procedure
  - **Specify Who** is responsible for executing the step
  - **Specify How** the step is performed
  - Define the cognitive and motor aspects
- Development and revision process is currently document-based



CCM Airlines A319 A320 A321 FLIGHT CREW OPERATING MANUAL | ABNORMAL AND EMERGENCY | 3.02.26 | P 9 | FIRE PROTECTION | SEQ 001 | REV 32

SMOKE/TOXIC FUMES REMOVAL

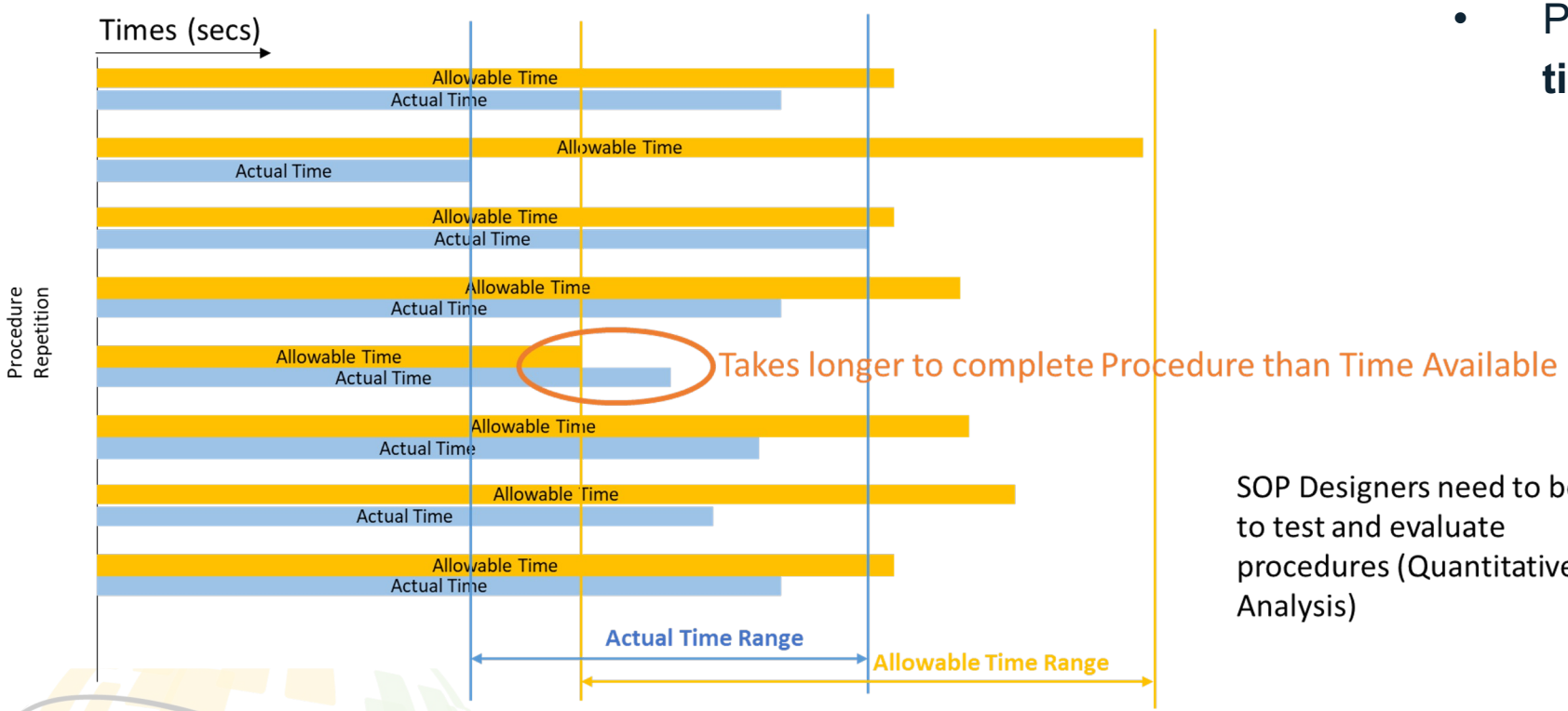
· *Use the smoke removal procedure if there is dense smoke, toxic fumes (smell), or if smoke generation cannot be stopped.*
If a scent similar to orange peels pervades the cockpit, suspect a toxic leak of rain repellent fluid. If the scent is similar to pine needles, suspect a non-toxic leak ◁.
· *If there is smoke in the cabin, it may be necessary to make a PA announcement to minimize apprehension.*

– OXY MASK/GOGGLE . . . . . ON/100 %/EMERG
Ensure crew communication is established. Avoid continuous use of the interphone to minimize interference from the oxygen mask breathing noise.
Turn the emergency knob to remove condensation or smoke from the mask.
– SEAT BELTS/NO SMOKING . . . . . ON
● **If fuel vapors :**
– CAB FANS . . . . . ON
The recirculating air ventilates the air mixer bay and other fuselage area, preventing fuel vapors from accumulating and the risk of explosion. Passenger health is not affected.
– PACK 1+2 . . . . . OFF
R ● **If no fuel vapor :**
R – CAB FANS . . . . . OFF
R *To prevent smoke from entering the cockpit and cabin.*
R – PACK FLOW . . . . . HI
R *To provide maximum airflow from the packs.*
R *Do not shut down the air conditioning packs, and do not reduce ventilation in an*
R *attempt to smother the fire.*
R *Do not deploy oxygen masks, if fire is suspected in the cabin.*
– LDG ELEV . . . . . 10000 FT/MEA
– DESCENT (FL 100 or MEA or minimum obstacle clearance altitude) . . . . . INITIATE
Since the most effective means of smoke removal is use of the ram air, descent is initiated to FL100, or the MEA, or the minimum obstacle clearance altitude, while the cabin altitude is increased to 10000 feet or MEA.
The increase in cabin altitude also reduces, at least temporarily, the smoke concentration.
Cabin depressurization starts when descent is initiated.
Passenger oxygen, as required by regulation.
– ATC . . . . . NOTIFY

Importance of SOPs

- Capture best practices in Critical Operations
- Standardize Activities
- Improve compliance, situation awareness, and decision making
- Minimize possibility for error, miscommunication, and misconceptions
- Critical for safe and efficient operation of systems
- Applications
  - Systems with Human & Automation Teaming
  - Training
  - Domains with multiple user operators

Need for MBSE & Analysis of SOPs



SOP Designers need to be able to test and evaluate procedures (Quantitative Analysis)

*Same Crew, Same Procedure, Different Circumstances

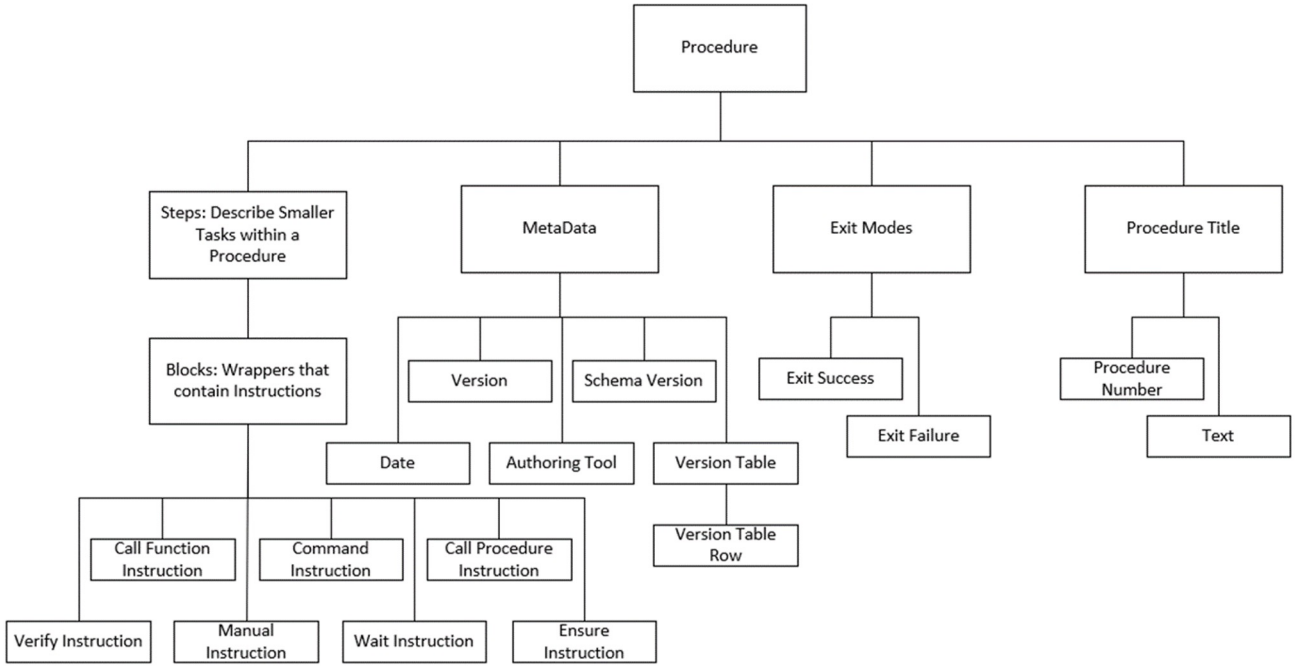
- **Specify WHAT is done, but not WHEN**
- Quality and performance **not quantified**
- Do not account for **variability** in operator and/or operational environment
- Procedures **cannot be completed in time**
  - Case of Swissair 111
    - Time to complete both required checklists is 20-30 minutes.
    - Time for a piece of equipment to progress from smoke to fire:
      - Extremely variable
      - Ranges from 6 minutes to 10+ minutes

Objective

- To improve the **development**, **revision**, and **analysis** of Standard Operating Procedures (SOPs) by applying MBSE, enabled by Gen-AI:
  - Digitize SOPs in a Safe & Cloud Database
  - Perform Modeling, Simulation, and Analysis of SOPs to gather qualitative data
  - Verify and Validate each SOP can be performed within time constraints

Procedure Representation Language (PRL)

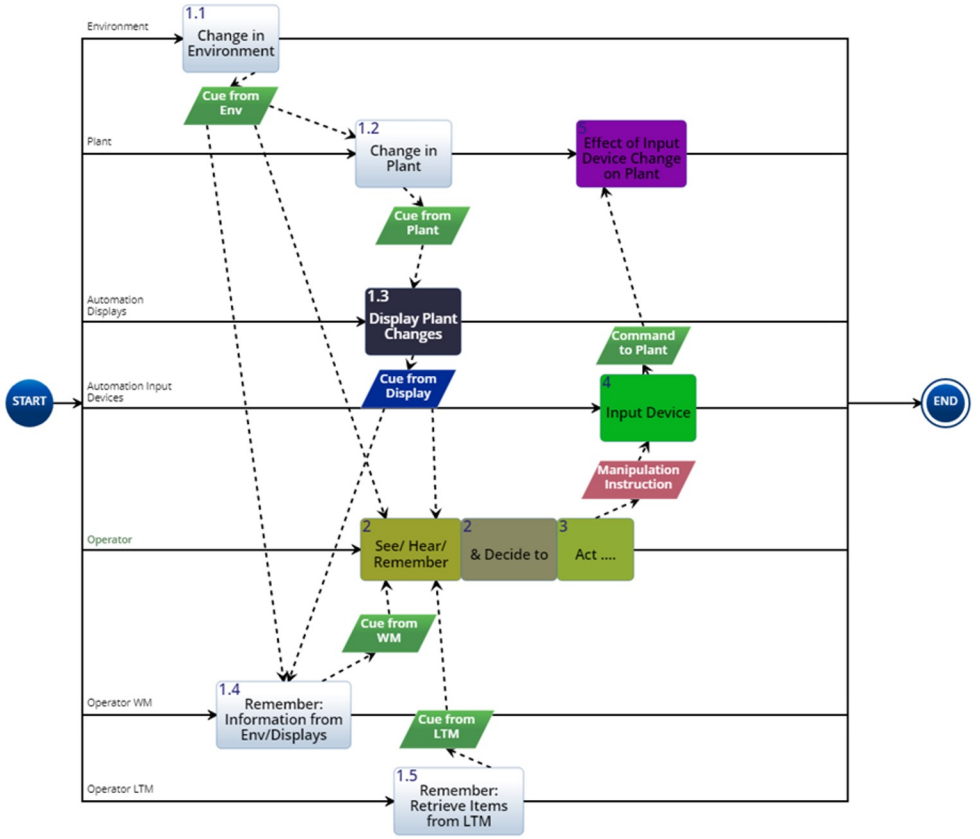
- Formal Representation of Procedures
- Designed for humans and automated systems
- Defined as an XML Schema
- Instructions
  - Lowest level of abstraction
- Manual Instruction
  - Used for commands that must be completed by the human operator
- No underlying structure to the manual instruction



Kortenkamp, David et al. 2007

SOP Step Canonical Structure

- SOP steps are governed by the SOP Step Canonical Structure
- Reflects steps involved in operator cognition to Observe, Orient, Decide, and Act
  - **Observe:** Operator senses triggering information, such as changes in the environment
  - **Orient:** Information is coupled with pre-existing knowledge and mental models
  - **Decide:** Formulate an appropriate response or decision
  - **Act:** Perform an action or response that was decided

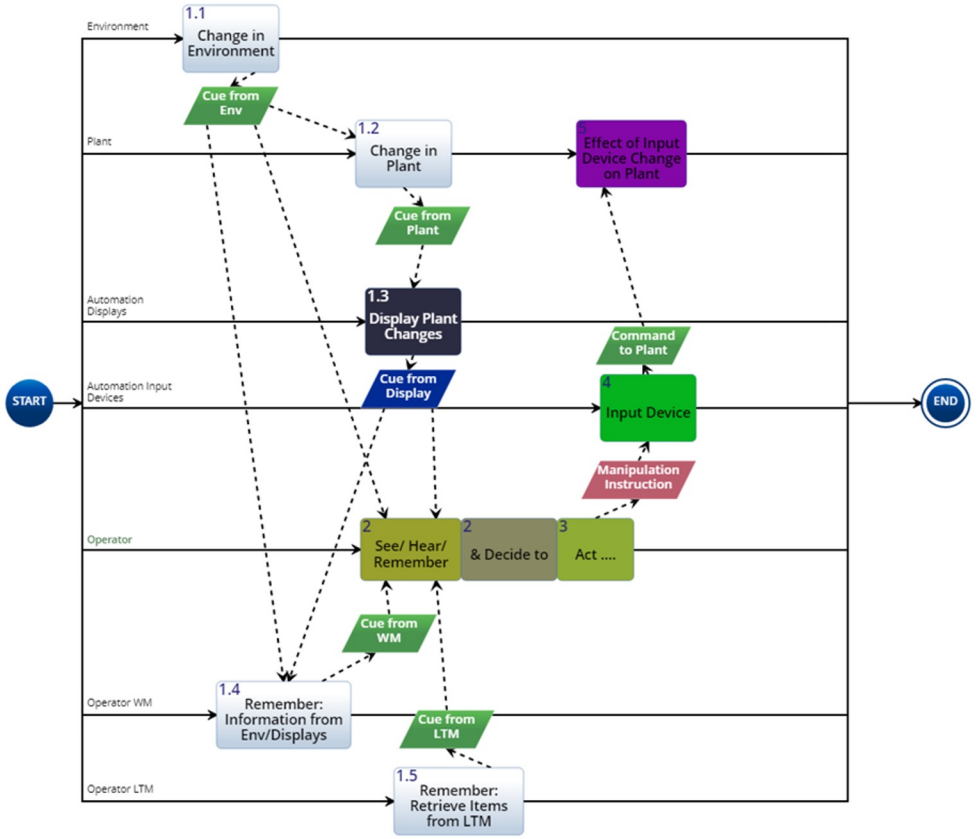


Bashatah & Sherry (2023)

[Actor]: [1. Trigger] [2. decide] then [3. Action] <<Then wait for [4. Waiting/Timing], and [5. Verification after Waiting]>>

SOP Step Canonical Structure (Cont.)

- Each SOP step has the components:
  1. Trigger (perceptual)
  2. Decision (cognitive)
  3. Action (motor)
  4. Waiting/timing (perceptual)
  5. Verification after waiting (perceptual)

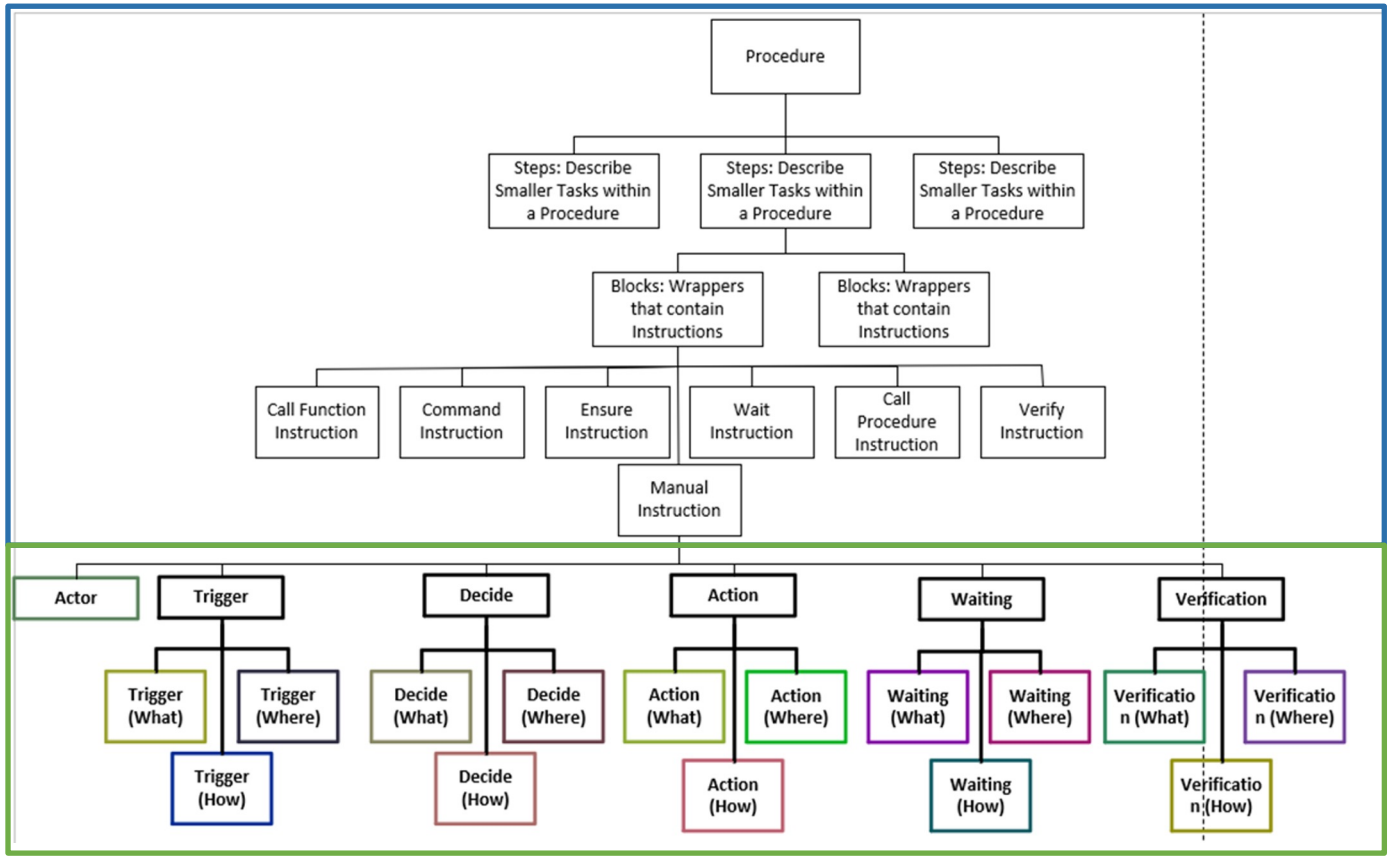


Bashatah & Sherry (2023)

Extended PRL (e-PRL)

PRL
Kortenkamp et al, (2007)

- Extension to the PRL
- Extends the Manual Instruction
- Standardized format mapping to how Humans operate
- Yields Four SOP step types:
  1. Action-Only
  2. Decision-Action
  3. Action with Waiting and Verification
  4. Decision-Action with Waiting and Verification



e-PRL Example with Trigger - Action - Waiting - Verification:

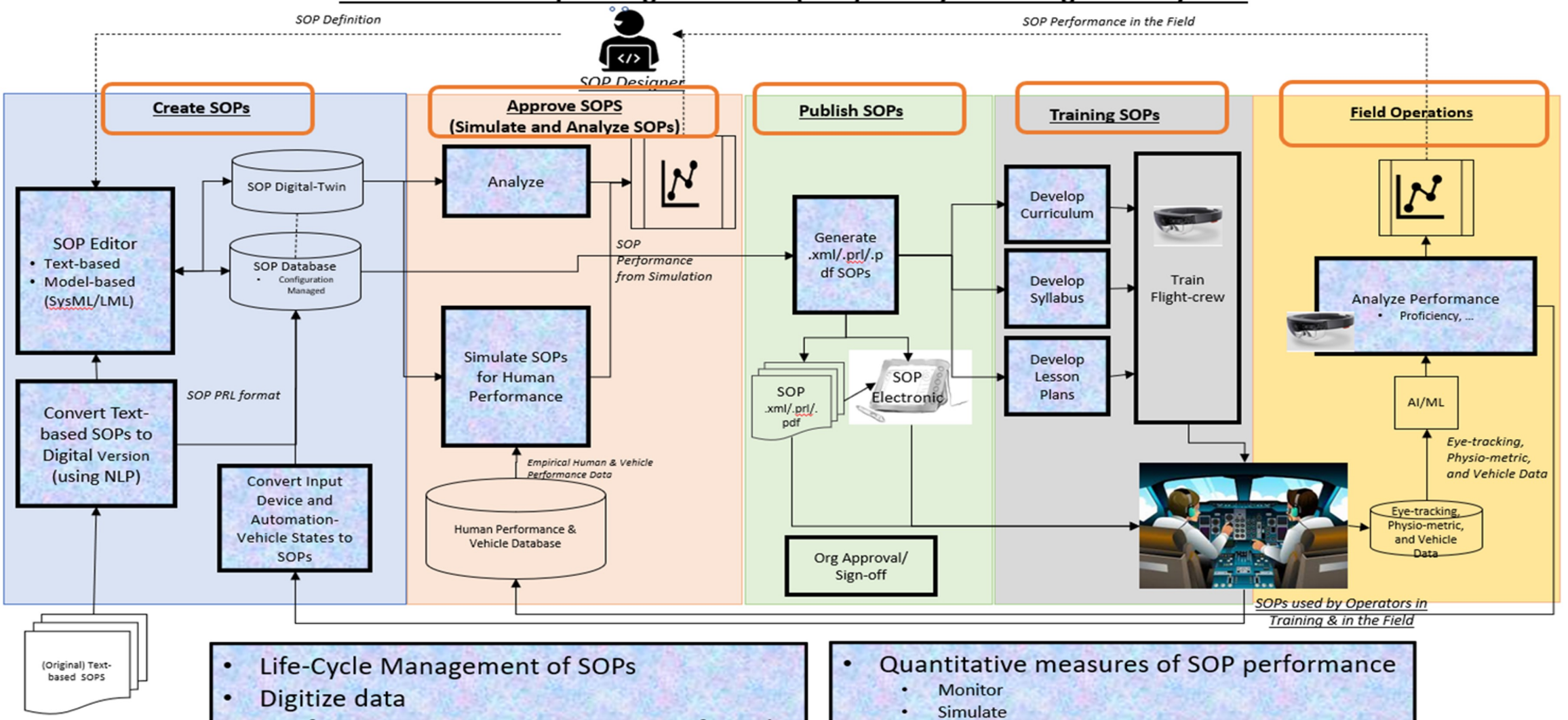
[*(After completion of previous step)*], [STANDBY POWER selector …… BAT]. [Verify EICAS advisory messages BAT DISCH MAIN and BAT DISCH APU display]. [Messages may take up to 3 minutes to display].

e-PRL
Bashatah & Sherry (2023)

Large Language Models (LLMs)

- Artificial Intelligence capable of understanding and generating text, images, etc.
- Trained using massive corpora of text data to understand syntax, semantics, and language concepts
  - LLMs have a deep understanding of human behavior and emotions
  - Can understand and generate contextually relevant responses
- Excels at various Natural Language Processing (NLP) tasks
- Notable LLM Chatbots:
  - OpenAI ChatGPT 4o
  - Anthropic Claude.ai 3.5

TO-BE Standard Operating Procedure (SOP) Life-Cycle Management System



- Life-Cycle Management of SOPs
- Digitize data
- Configuration Management *across* Life-cycle

- Quantitative measures of SOP performance
  - Monitor
  - Simulate
  - Analyze

Center for Air Transportation Systems Research at George Mason University

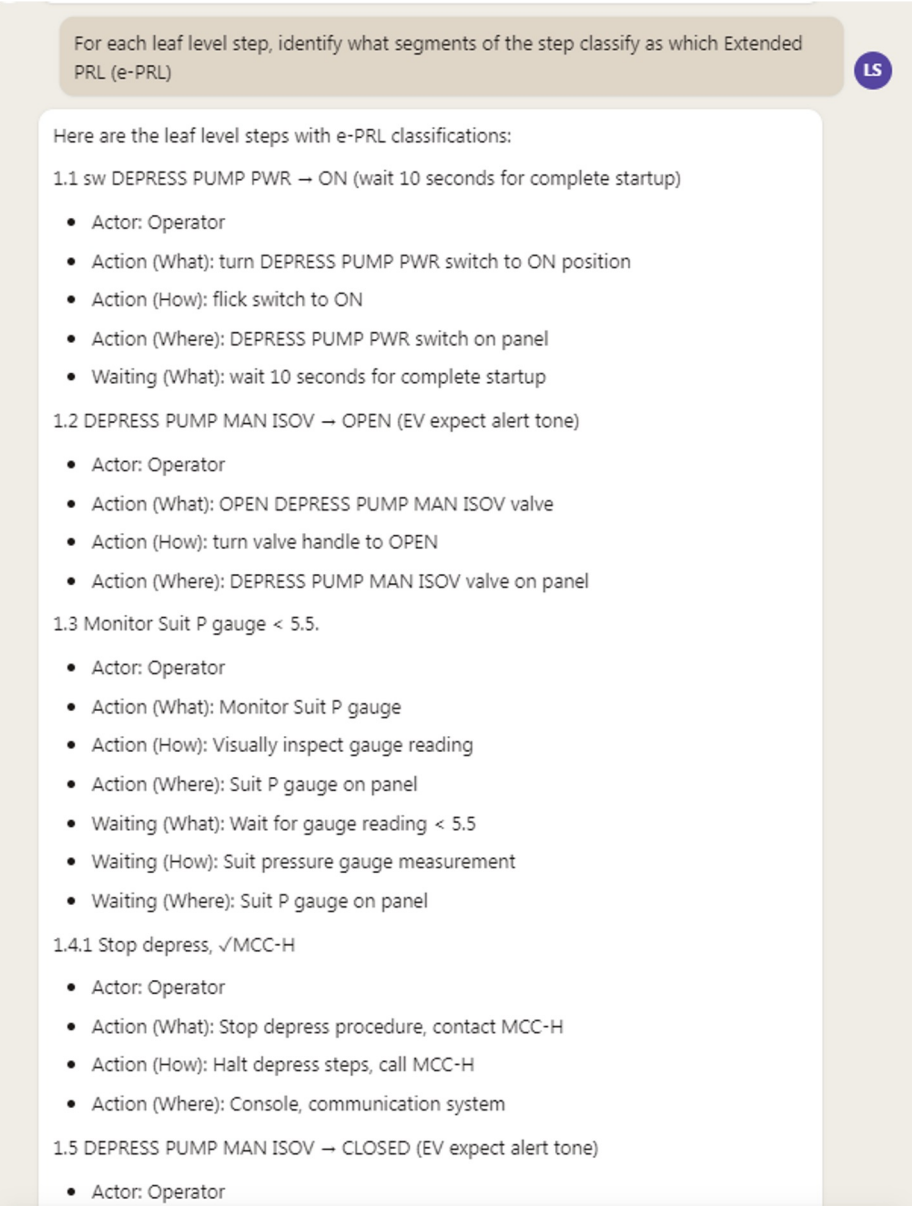
Digitize SOP

Digitize SOP Documents

- **Challenge:** Each SOP has different formatting and structure
  - Long time to annotate the SOPs manually into the e-PRL elements
  - NLP is inconsistent as human behavior can be vastly different and complex
- Each step can be classified into its e-PRL components
  - Each e-PRL component becomes a data element
  - Adds functionality
  - Enables MBSE of the SOP
- Explored using LLM-based chatbots to classify SOP steps
  - Understand the context and know the contents of the whole text
  - Understand human language and its nuances
  - Can be trained to specialize in performing the tasks needed

Classifying SOP Steps

- Used Claude.ai LLM to annotate SOPs
  - Supplied rules and assumptions to identify e-PRL elements in an SOP step
  - Supplied the SOP text and extracted the procedure steps generated
  - Classified the bottom-level steps into e-PRL elements
- Required trial and error to develop a predictable process



Pros and Cons

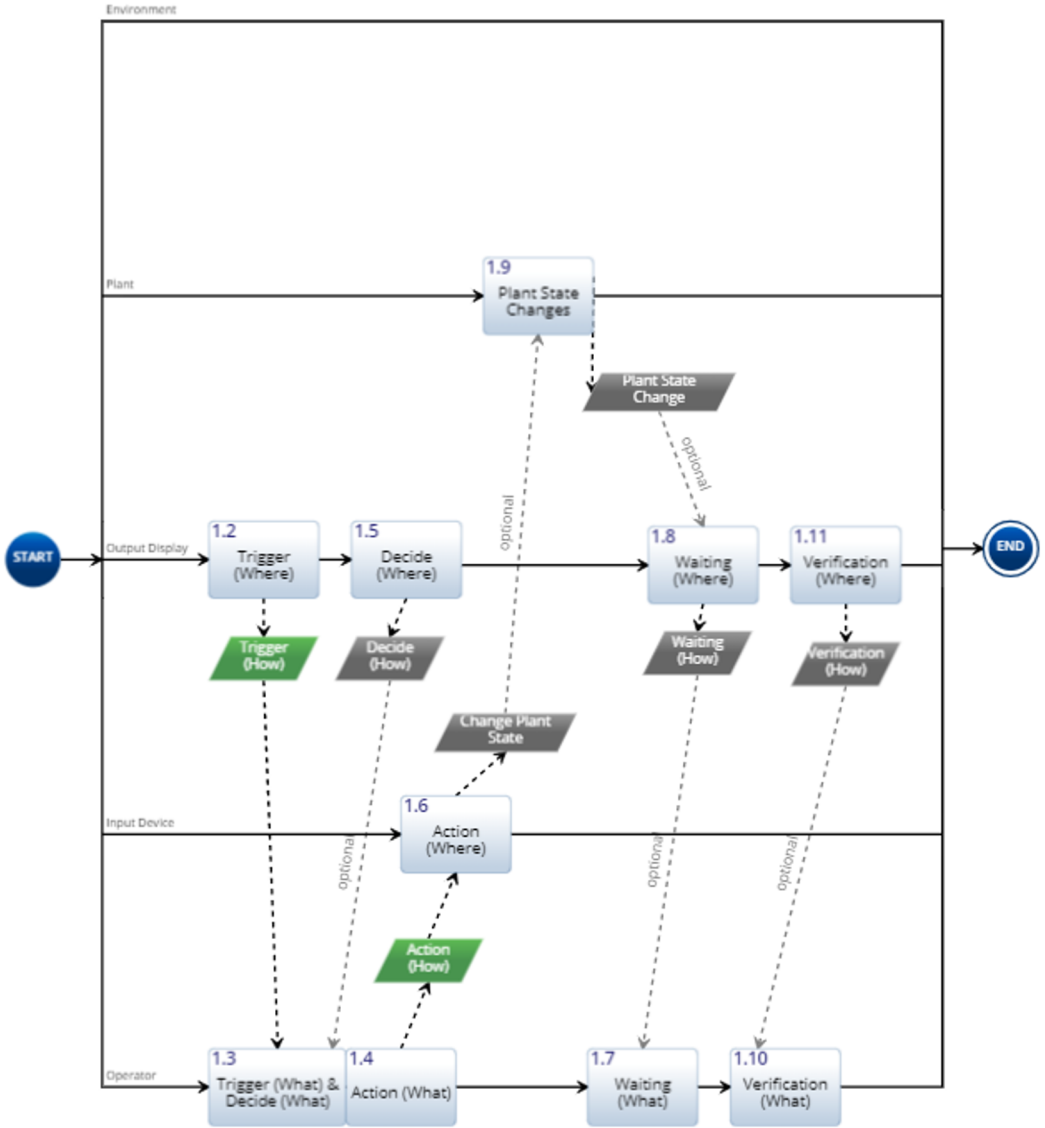
- **Pros**
  - Understood context from the beginning of the SOP, but not in the Step
  - Made corrections/adjustments when informed
  - Improved time to annotate SOPs

- **Cons**
  - Doesn’t return exact wording
  - Adds or omits information
  - Extensive prompting required to get consistent results

Simulate and Analyze SOP

Modeling SOPs using e-PRL

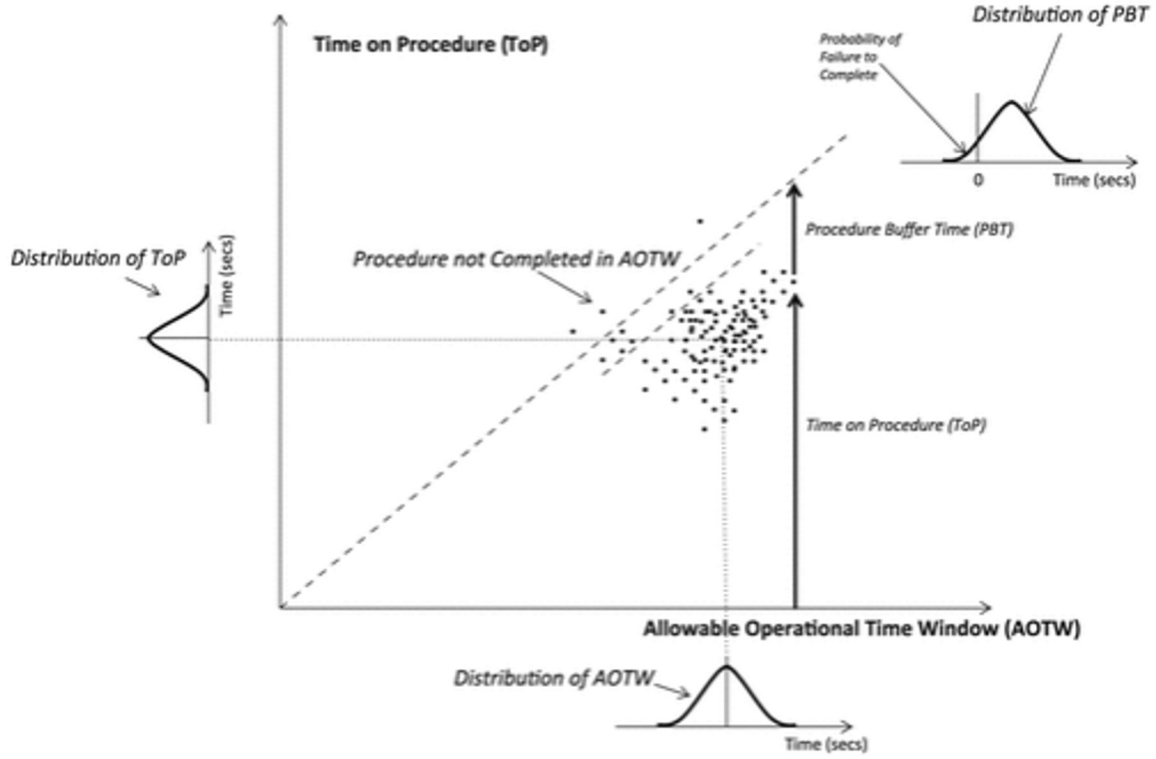
- Each e-PRL component can be modeled in an Action/Activity diagram
- Asset Entities
  - **Actors**
- Action/Activity Entities
  - **Where** components
  - **What** components
- Input/Output Entities
  - **How** components
- Enables the identification of missing information or context



SOP Time Analysis Metrics

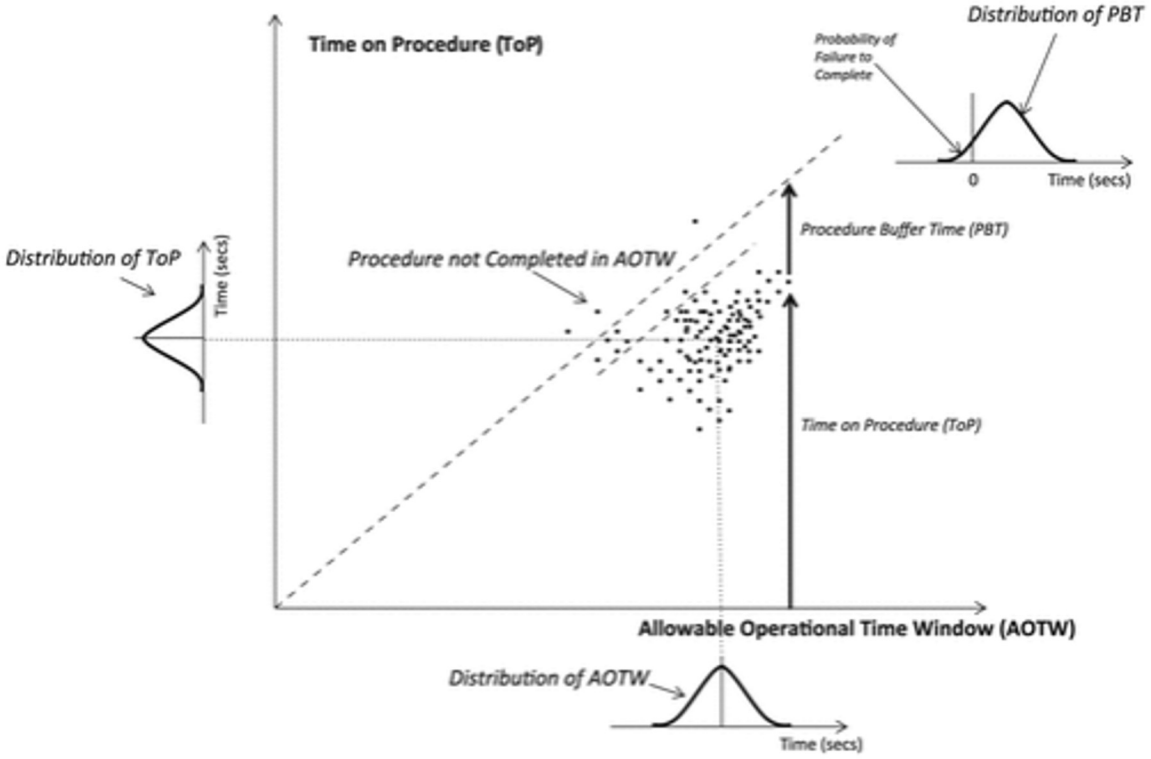
Time metrics help provide data on the operator's performance in the SOP

- Allowable Operational Time Window (AOTW)
  - Time that the SOP must be completed to avoid any hazards
- Time of Procedure (ToP)
  - Time it takes to complete the procedure
- Procedure Buffer Time (PBT)
  - Difference between AOTW & ToP
  - Value < 0, then the Procedure is not feasible



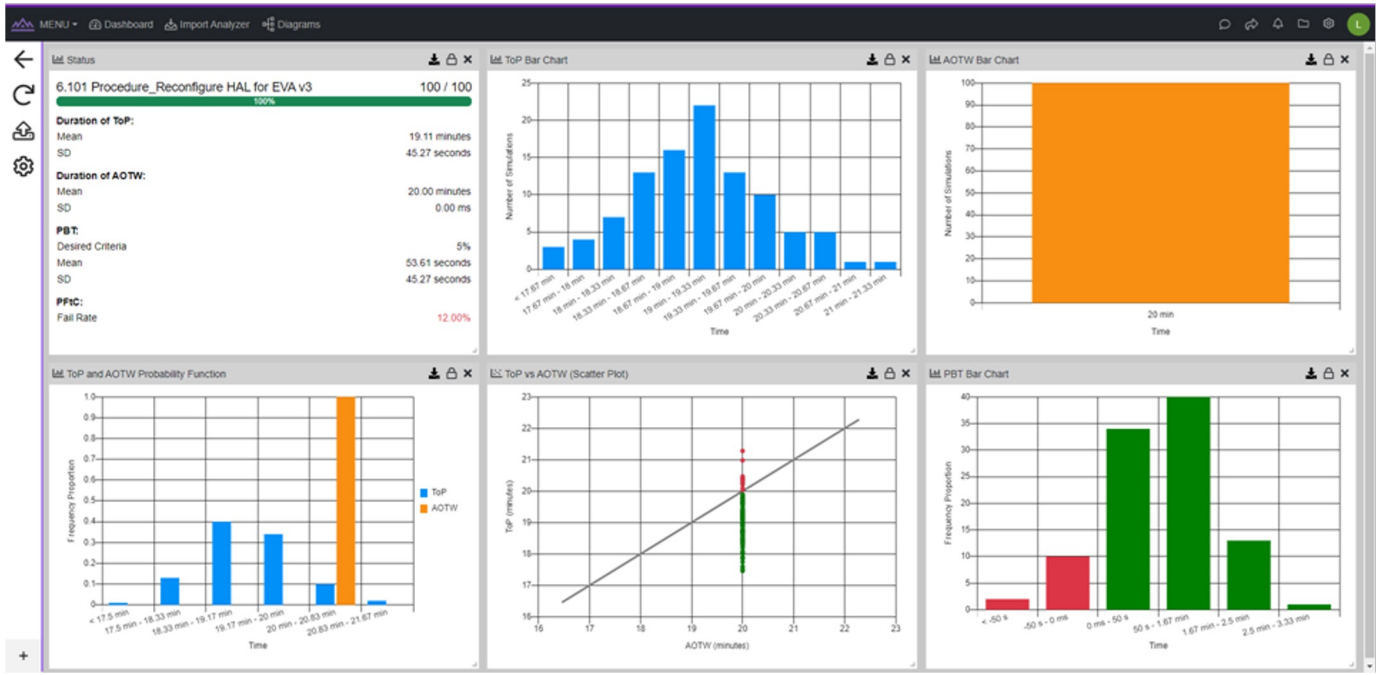
SOP Time Metrics (Cont.)

- Time metrics based on time distribution data when feasible
  - LLM chatbots can extract time data from an SOP text and provide estimated time distribution data for each step in the SOP text
- Time metrics will provide a high-level view to find the operator's performance gaps



Monte Carlo Simulation

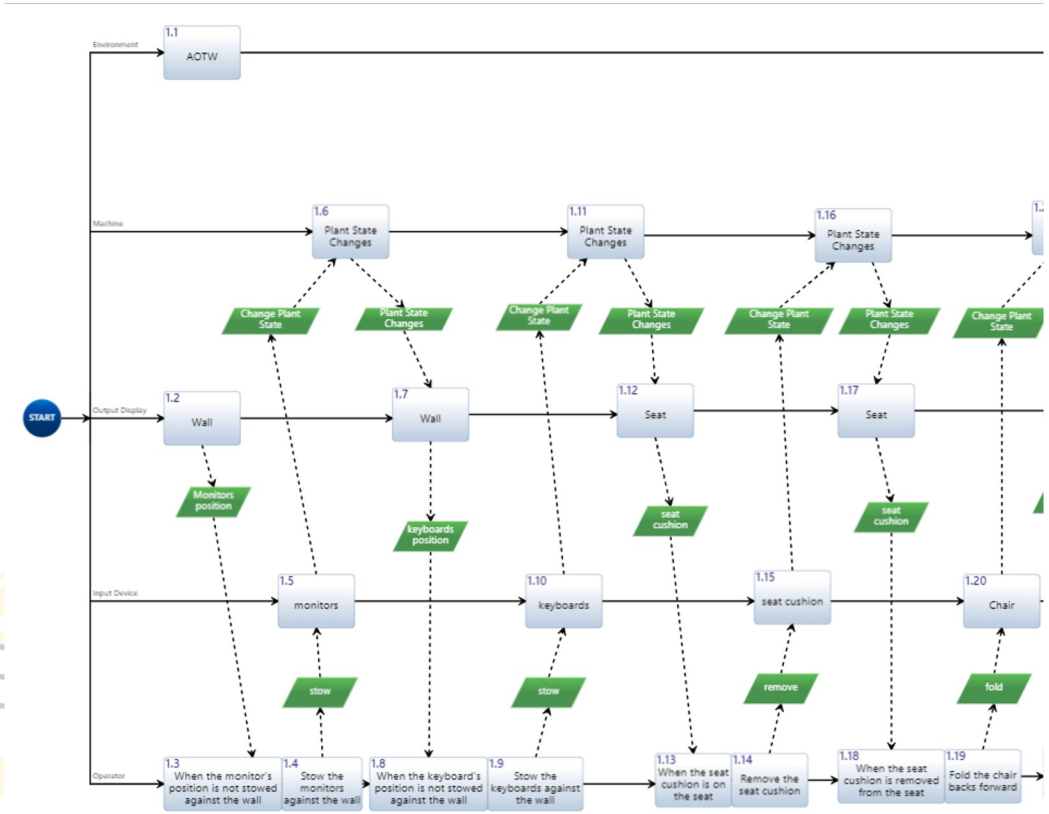
- Time metrics data added to Action/Activity can be used to perform a Monte Carlo Simulation for the diagrammed SOP
- Runs multiple iterations to find results, outliers, and performance gaps
- Tracks the time metrics (AOTW, ToP, PBT) to evaluate an operator's performance in the SOP
- Verifies and Validates if the SOP can be completed in the given time to reduce the risk of failure
- Provides insight into which steps contribute to good/bad performances



Generate SOP

Generate SOPs from Models

- e-PRL components need to be put into the expected format of an SOP
- Generative AI can be used to take the visual fragments of the e-PRL Action diagram to generate imperative procedure sentences and documentation.



6.101 - Procedure Reconfigure HAL for EVA v3

Objective:

Configure the habitable airlock for EVA by removing EVA related items from PRS drawers and moving them to their designated location on the suit or in the habitable area. Non-EVA related items that cannot go to vacuum and are not essential for daily operations will replace the EVA items in the PRS drawers.

DURATION: 20 Minutes

Procedure:

1. Preparing HAL Cabin for EVA
   - 1.1. Reconfigure the Command and Control Station
     - 1.1.1.Stow monitors against the wall.
     - 1.1.2.Stow keyboards against the wall.
     - 1.1.3.Remove seat cushion.
     - 1.1.4.Fold the chairbacks forward.
   - 1.2. Detach crew hygiene kit from the aft transfer port hatches.
   - 1.3. Stow the crew hygiene kits in Lockers SA-1 and PA-1.
   - 1.4. Remove hatch cargo nets from lockers SA-1 and PA-1.
   - 1.5. Secure hatch cargo nets at the starboard and port hatch openings to 3 of the 4 D-rings. Note: Leave the forward, bottom D-ring unhooked in order to stow
   - 1.6. Remove IVA Common Tool Kit from PM-5.
   - 1.7. Temp Stow IVA Common Tool Kit behind the Port Hatch Opening.
2. Removing EVA items from PRS, and stowing in HAL Cabin
   - 2.1. Open PRRS Hatch
     - 2.1.1.PRESS "Enable" on PRS switch panel. Verify "Enable" button light up on PRS switch panel.
     - 2.1.2.PRESS "Door Open" on PRS switch panel.
   - 2.2. MOVE PRS to Drawer 1, and PULL it into the HAL Cabin
     - 2.2.1.PRESS "Enable". Verify button light up on PRS switch panel.
     - 2.2.2.PRESS "Next Drawer" or "Prev Drawer". Wait until, Drawer 1 is positioned on in front of transfer port door.

Conclusion

- Integrating Generative AI and MBSE enables significant enhancements in the **Development**, **Revision**, and **Analysis** of Standard Operating Procedures (SOPs)
- Generative AI introduces potential bias and security risks, but they can be mitigated through careful training and review processes
- A new approach that offers:
  - Enhanced effectiveness, efficiency, and safety of critical procedures
  - Ensures SOPs contain necessary content and align with human cognitive processes

Papers Published

- Bashatah, Jomana A., and Lance Sherry. “A Model-Based Approach for the Qualification of Standard Operating Procedures.” In *2021 Integrated Communications Navigation and Surveillance Conference (ICNS)*, 1–10, (2021). https://doi.org/10.1109/ICNS52807.2021.9441587.
- Bashatah, Jomana A., Lance Sherry, Steve Dam, Lauren Flenniken, Patrick Hartmann, and Tom Harold. “Analyzing Standard Operating Procedures Using Model-Based Systems Engineering Diagrams.” *INCOSE International Symposium* 31, no. 1 (2021): 1130–44. https://doi.org/10.1002/j.2334-5837.2021.00891.x.
- Jomana Bashatah and Lance Sherry. “Model-Based Analysis of Standard Operating Procedures’ Role in Abnormal and Emergency Events.” *INCOSE International Symposium*, (2022).
- Bashatah, Jomana, and Lance Sherry. “Lessons Learned From Human Operator Intervention for Ai Navigation and Flight Management Systems.” In *2022 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, 1–15, (2022). https://doi.org/10.1109/ICNS54818.2022.9771518.
- Bashatah, Jomana, and Lance Sherry. “Usability Analysis of an MBSE Model of Standard Operating Procedures.” In *INCOSE International Symposium*, 32:38–47, (2022). https://doi.org/10.1002/iis2.12870.
- Bashatah, Jomana, and Lance Sherry. “Method for Formal Analysis of the Type and Content of Airline Standard Operating Procedures.” In *2023 Integrated Communication, Navigation and Surveillance Conference (ICNS)*, 1–8, (2023). https://doi.org/10.1109/ICNS58246.2023.10124313.
- Bashatah, Jomana, and Lance Sherry. “Prompt Engineering to Classify Components of Standard Operating Procedure Steps Using Large Language Model (LLM)-Based Chatbots'.” In *2024 Integrated Communication, Navigation and Surveillance Conference (ICNS),* (2024).
- Bashatah, Jomana, and Lance Sherry, “Prompt Engineering for Classifying the Components of SOP Steps to Enable SOP Simulation and Analysis” submitted to Int J. Human Factors Modelling and Simulation (2024)

Questions?

34th Annual INCOSE international symposium

hybrid event

Dublin, Ireland
July 2 - 6, 2024