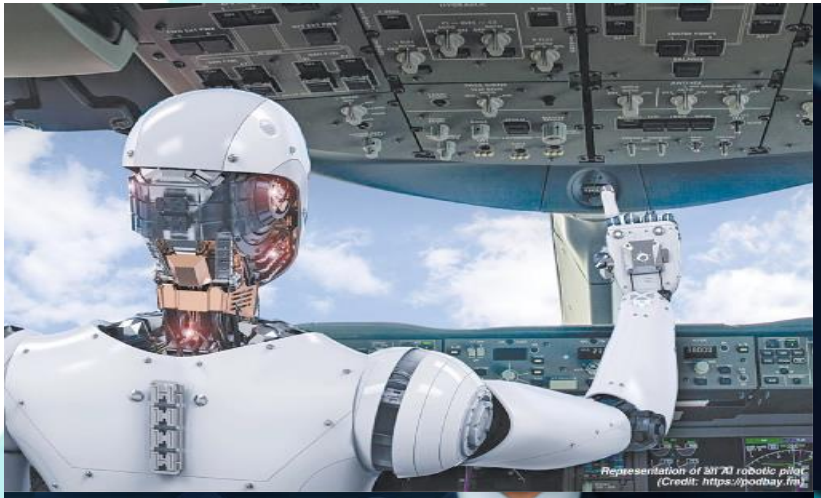
Human Factors Assessment of Future Aviation Intelligent Assistants

Presentation at CRA Forum, 20-21 September, Birmingham UK

Barry Kirwan, EUROCONTROL

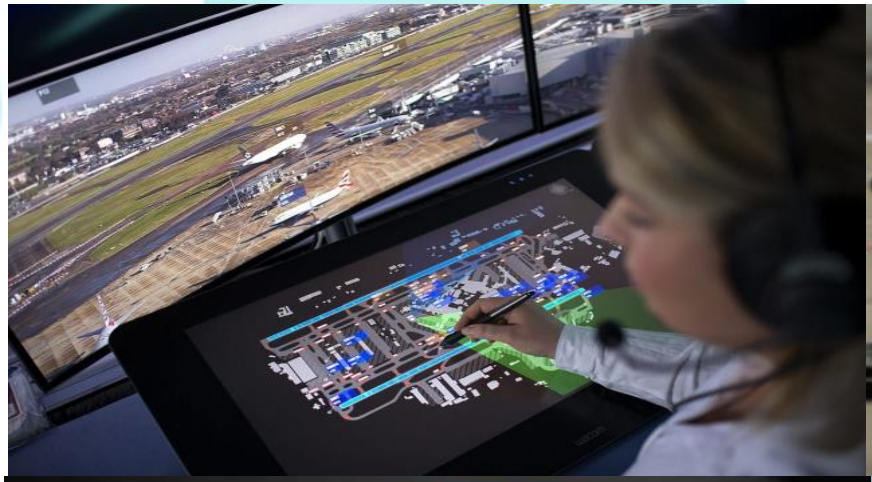
This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

Visions of AI in Aviation…



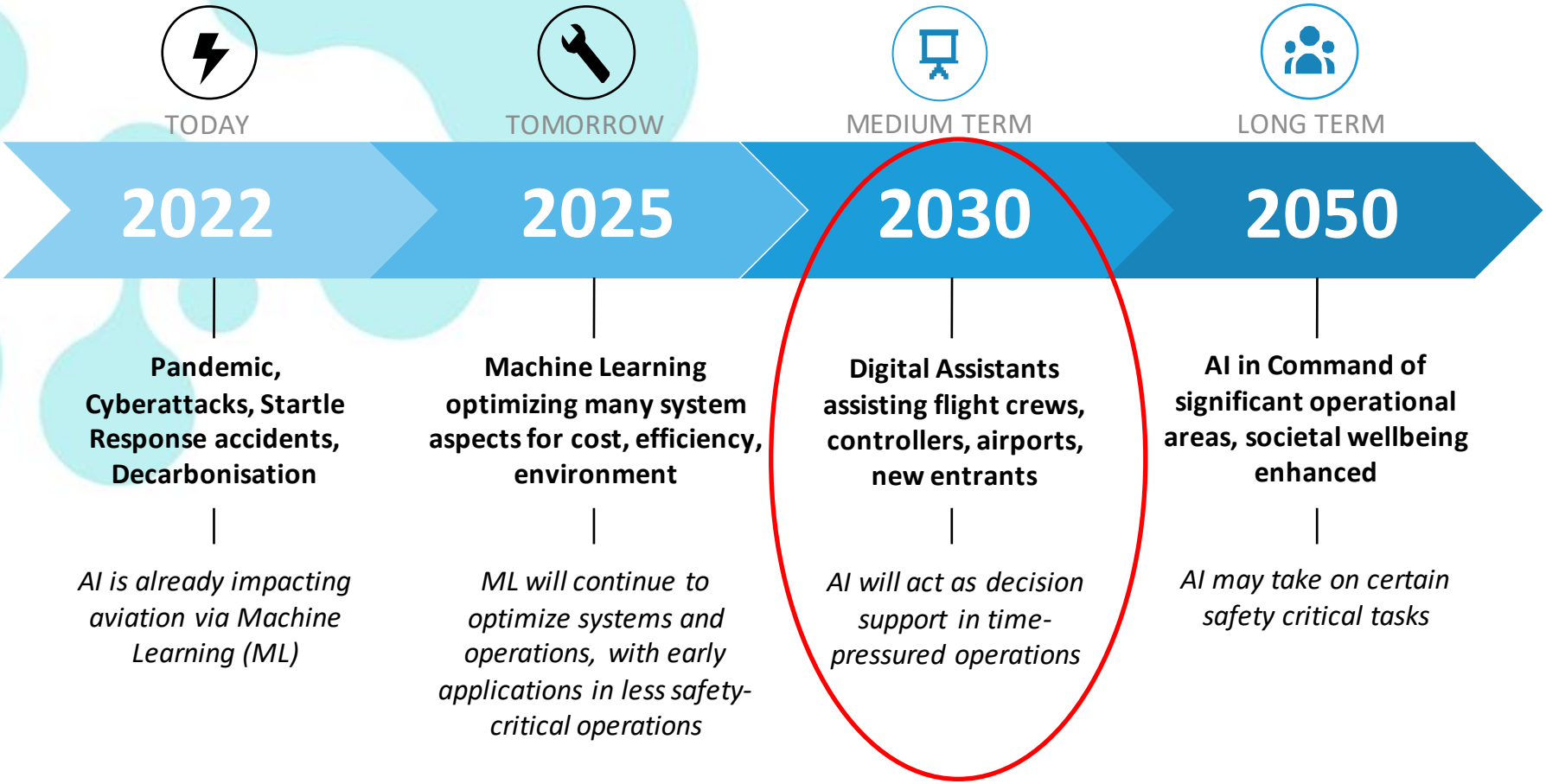
This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

...in diverse aviation operations



This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

HAIKU looks far ahead...



This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

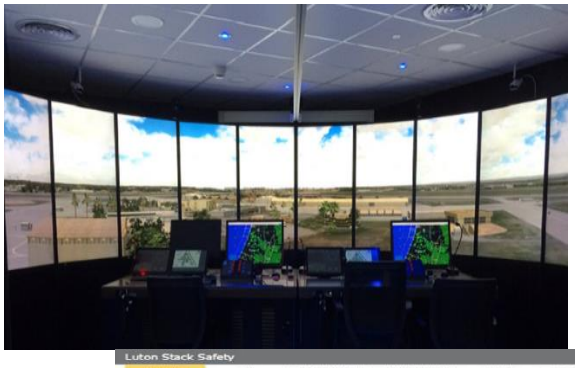
Our goal >>>

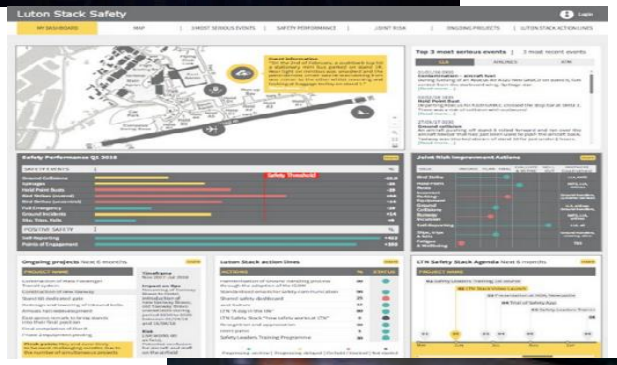
is to pave the way for **human-centric-AI** via the exploration of interactive **AI prototypes** in a **wide range of aviation contexts**

Our challenge

is to deliver **truly *human-centric* Digital Assistants**, capable to **'fit' the way humans work**.

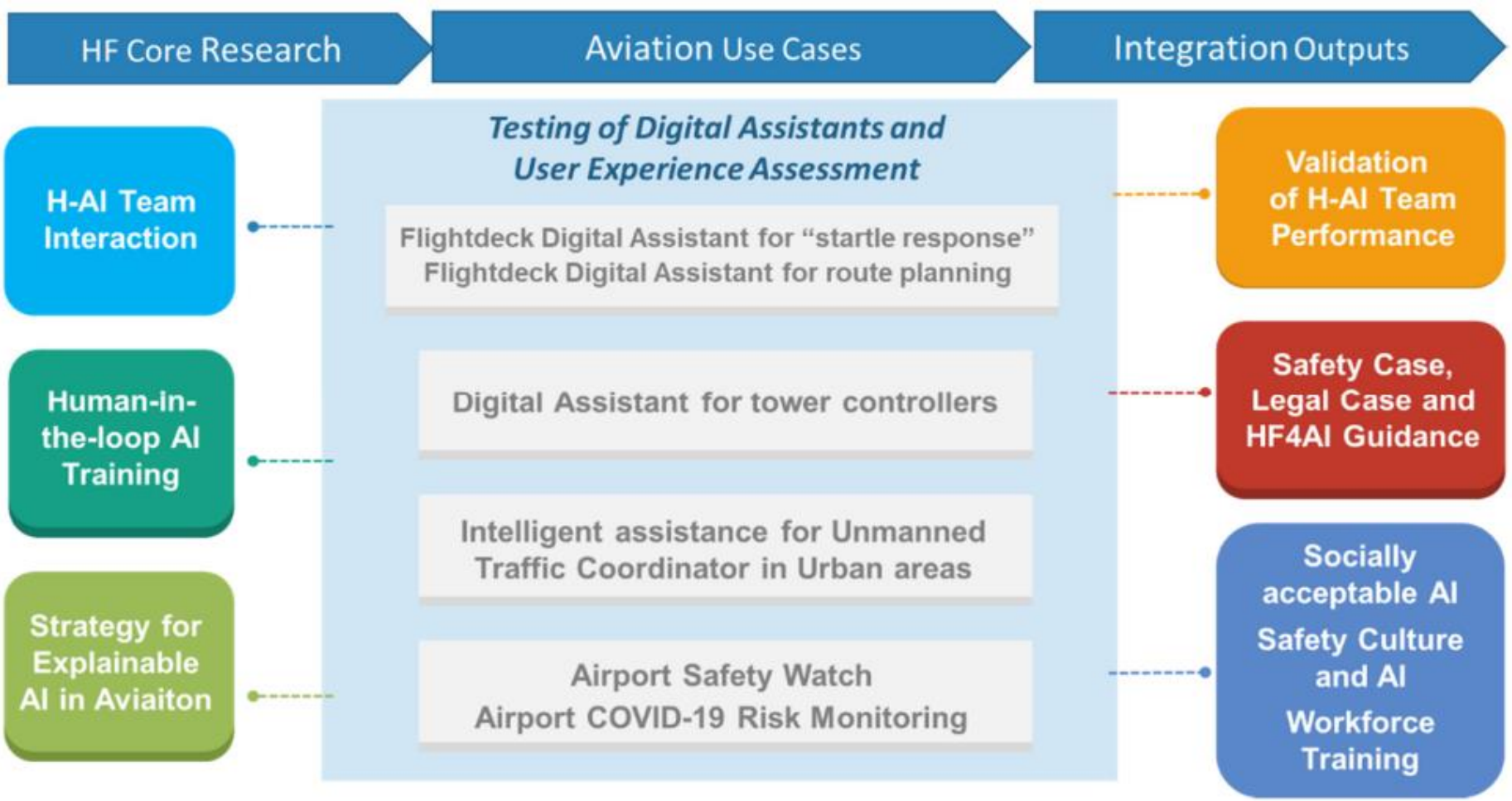
This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332





This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

Our Strategy



This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

EASA's AI Classification System

Level 1 AI : assistance to human	Level 2 AI : human/machine teaming	Level 3 AI : more autonomous machine
• Level 1A: Human augmentation • Level 1B: Human cognitive assistance in decision and action selection	• Level 2A: Human and AI-based system cooperation • Level 2B: Human and AI-based system collaboration	• Level 3A: The AI-based system performs decisions and actions, overridable by the human. • Level 3B: The AI-based system performs non-overridable decisions and actions.

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

Key Human Factors Aspects with Human-AI Teaming

'Internal' HF

- Situation Awareness
- Workload
- Mental Model
- Causal thinking & biases
- Trust
- Motivation
- Autonomy
- Engagement
- Startle / Surprise
- Experience / Expertise
- Values & Ethics
- Fast decision-making
- Multi-tasking
- In-and-out of loop
- Social interaction / teaming
- Wellbeing / Mental Health
- Culture

AI-related

- AI Model / SA
- Data biases
- Self-checking, monitoring and confidence level assessment
- Monitoring of user's cognitive state
- [Operational] Explainability
- Dialogue capability
- Rational argumentation with user
- Temporal sensitivity
- Task flexibility
- Edge / Corner cases / Hallucinations
- Trade-offs & core values
- Supervised / unsupervised learning
- Failure modes
- Detection of poor user strategies
- Customisation to user
- Digital twins

System-Related

- Distributed SA
- HMI intelligibility / usability
- Communications
- Authority Gradient
- Task Allocation
- Teamwork
- Training & Selection
- Procedures
- Leadership
- Error / Failure Management
- HAIRM' (HAT-CRM)
- Fatigue Management
- Safety culture
- Organisational Culture

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

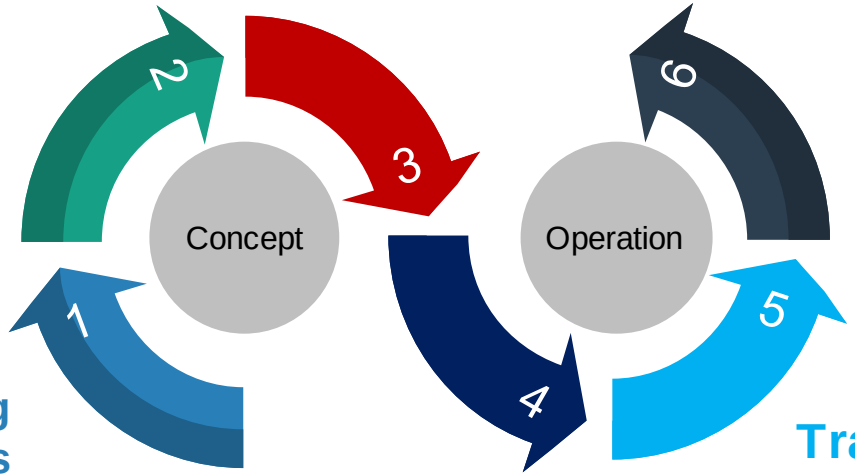
Human-AI Teaming Human Factors Assurance Process

Human-AI HAZOP

Considering what could go wrong and identifying safeguards and mitigations

Task Analysis

How the system should work, including all team players, using scenarios and timelines, tracking the evolution of human and AI situation awareness



Monitoring, Adapting, Learning

Human-AI systems may evolve rapidly once implemented. Impacts on safety performance, crew wellbeing, and safety culture, must be closely monitored and managed.

Human AI Teaming (HAT) Requirements

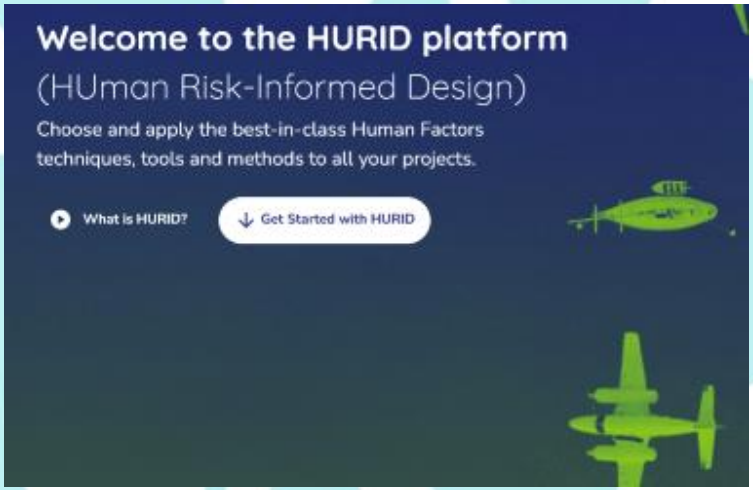
From maintaining skill sets and situation awareness, to managing workload in emergencies, to dialogue modalities to reciprocal error checking

Human-in-the-loop simulations

Human-AI simulations where HAT requirements can be evaluated using Human Factors tools and techniques.

Training & Operational Readiness Testing

Training of staff in new roles, motivating the workforce, fine-tuning of explainability etc. for a wider range of situations, development of failure management and fallback systems.



Task Analysis (OSD) & HAZOP

OSD (Operations Sequence Diagram) and Timeline Analysis

Operations Sequence Diagram | Design | Human Factors | Data Representation

Putting the people pieces together to achieve effective and robust task performances

Human Hazard and Operability Study (HAZOP)

safety | HumanFactors | analysis | safeflix

What-if? approach, using experts to determine what could go wrong and how to prevent it.

On Operational Sequence Diagram for HAT

What does an OSD include?

- Time
- The actual system state
- The goal that needs to be achieved
- The people involved & location
- Key information sources
- Decisions, actions, communications
- What they may think the system state is, if different from actual
- Equipment resources issues
- Potential bottlenecks that may block goal achievement
- Secondary tasks / distractions

What else does a HAT OSD need?

- What the AI believes/predicts to be the system state
- The AI recommendation
- Closeness of advice to performance edge / distance from operator's understanding
- The AI rationale (explainability) whether before, during or after the event
- HMI for HAT communication: signals, understanding, predictive, uncertainty
- Human-AI 'dialogue' including cross-checking and queries
- The authority gradient ('who's in charge')
- HF Impact: trust, surprise/startle, workload, engagement (out of loop); competence

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

HAT OSD Unpacked

Time	Actual System State	Goal	Human1	Human2	Human3	Info sources (non-AI)	Operator believed system state	AI believed system state	AI solution
14:00 UTC	Weather and traffic configuration will lead to increased risk of incorrect taxiway selection	Alert operational units to heightened risk and employ counter-measures	Safety Watch Supervisor (located in the Operational Control Centre) receives the AI advice and wants to verify it.			Traffic arrival / departure monitoring and schedules; weather monitors; surface monitoring	Okay now, could get difficult later.	increased taxiway selection error risk	Increase monitoring of conformance or reduce traffic capacity. The advice will be time-framed, e.g. from 15:00-17:30 UTC.

AI HMI	AI Rationale (XAI)	Closeness of solution to performance edge	H-AI Dialogue	Authority gradient	Decision / Action	HF Impact: trust, SA; startle / surprise; workload; engagement; competence	Comments & Observations
The AI HMI can present graphical display tracking current and predicted parameters including risk of 3 error types, as well as threshold for alert. LTN Airport map display highlighting taxiway sections and junctions at risk.	The Safety Watch IA will be able to cite historical evidence, or show the clusters of factors that predict increased risk. It should show which parameters are involved, as well as any that are not indicated. Ideally it should give information on uncertainty / confidence level.	Unless the parameters are wildly off, this would be within expected parameters of the Safety Watch tool.	The Safety Watch Supervisor queries whether Hold-Point Busts will also be an issue, as a number of the indicators for this error are also present in the prediction. The supervisor can do this simply by selecting Hold-Point Busts Prediction'. It shows that while indeed several of the indicators are preent, they are just below the statistical threshold.	The human is in control. This is HAT category 1B.	The supervisor decides to alert for bothtaxiway selection error and holdpoint busts.	No negative impact on Human Factors in this scenario.The AI is simpy augmenting human performance and system safety.	

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

Traditional HAZOP Guidewords

- No, not done
- Other than
- Sooner / Later
- Less / More
- As well as / Part of
- Reverse

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

How HAZOP Works

Step	Guide Word	Hazard	Cause(s)	Consequence	Existing Safeguards	Recommendation
AI sends alert warning	NONE	Alert not triggered by AI	AI algorithms/data not sensitive enough to error causes	No alert raised, yet incorrect taxiway selected by an aircraft.	Normal ('as-is') detection and recovery processes will operate to prevent taxiway collision	Refine data-set / algorithms to increase sensitivity to actual causal patterns
	NONE	Supervisor judges situation ok and decides not to issue warning	(1) Supervisor experience differs from full dataset. (2) Previous false alarms (3) 'Pushback' from airlines or ATC.	No alert raised, yet incorrect taxiway selected by an aircraft.	Normal ('as-is') detection and recovery processes will operate to prevent taxiway collision	Training review for supervisor. Discussion at Safety Stack over the threshold for triggering the alert.
	PART OF	Sup issues alert but message fails to reach everyone	(1) Comms difficulties (2) Community App problems (3) Internal channels fail to reach all parties	Insufficient reaction to alert; incorrect taxiway may be selected.	Normal ('as-is') detection and recovery processes will operate to prevent taxiway collision	Review of comms links and processes to ensure 100% coverage next time.

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

What we aim to deliver (2 more years)

- How to develop trustworthy AI systems that humans can work with, demonstrated by TRL6 case studies
  - Cockpit, virtual tower, airport safety, UTM
- HF Guidance
  - How to capture HF4AI requirements
  - How to evolve and validate the concept via requirements, task analysis, HAZOP and simulations
  - HMI and two-way communication design
  - Reciprocal explainability
  - Guidance on current & future workforce requirements & safety culture
  - Societal acceptance, regulatory & organizational considerations, and HF capability needed to ensure safe entry of AI systems into operation.
- Safety, Security, Human Factors, & Validation Approaches

This project has received funding by the European Union's Horizon Europe research and innovation programme HORIZON-CL5-2021-D6-01-13 under Grant Agreement no 101075332

We kind of need to get Human AI Teaming right…

“Open the pod bay doors, HAL.”

“I'm sorry, Dave. I'm afraid I can't do that.”

Haiku
Human AI teaming Knowledge and Understanding for aviation safety

Thanks for your attention

Barry Kirwan

barry.kirwan@eurocontrol.int

https://haikuproject.eu/