Supporting European Aviation

EUROCONTROL

How will AI affect safety culture in future aviation?

Part 1: AI in the aviation context
Part 2: AI and aviation safety culture

Barry Kirwan, EUROCONTROL

SAFETY TEAM 38, Brussels, 26th June 2024

barry.kirwan@eurocontrol.int

CONSORTIUM COORDINATOR sesar DEPLOYMENT MANAGER

FOUNDING MEMBER sesar JOINT UNDERTAKING

NETWORK MANAGER

What is Artificial Intelligence (AI)?

Narrow AI

Focusing on a particular domain, e.g. ATM. From Machine Learning to Cognitive Assistants

Number-crunching & Algorithms

Generative AI

E.g., Large Language Models such as Chat GPT, which can 'seem' human.

Learn, mimic, create…

Artificial General Intelligence

True 'thinking machines', which will (one day) surpass human cognition. They do not exist yet.

Transformative, potentially sentient

EU Act on AI:

‘Artificial Intelligence system’ (AI system) means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments.

- UNACCEPTABLE RISK (Banned) [+6 months]
  - Use of A.I. systems for predictive policing
  - Biometric identification systems in real time
  - Targeted reading of facial images from the Internet or from video surveillance systems to create facial recognition databases
  - Conclusions about people's emotions in the workplace
- HIGH RISK (Stringent requirements) [24-36 months]
  - **Risk Management / Data Governance / Human Oversight**
  - **Cybersecurity / Accuracy & Robustness / Ops Monitoring**
  - **LLMs: Model evaluation / Adversarial Testing / Reporting**
- LIMITED RISK
  - if you deploy, provide or use an AI system in this category (e.g. a chatbot), you must inform users that they are interacting with an AI system and also label all audio, video and photo recordings as being generated by AI.

EASA Human-AI Teaming Classification

As interpreted by HAIKU

1A
Human Augmentation

‘As-is’ today

Enhanced information. No significant impact on the operation or the end user.

1B
Human Assistance

Decision Support

The AI is a cognitive assistant that helps optimise decision-making.

2A
Human-AI Teaming Cooperation

Task Owner

Tasks allocated to the AI. Explainability (XAI) and some human-AI communication are required.

2B
Human-AI Teaming Collaboration

Goal-Sharing

The AI negotiates with human towards common goals. More explainability and dialogue required.

3A
More Autonomous AI

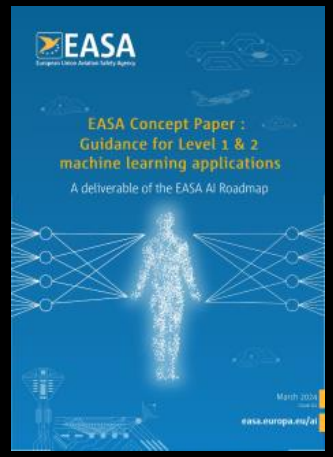
System Control

The AI is ‘running the show’ unsupervised, human intervenes if required.

3B
Fully Autonomous AI

AI Alone

The AI is running everything with no human intervention except maintenance and shutdown.



What we're trying to do…

- **Humane AI**. (https://www.humane-ai.eu/). The core of the humane AI vision is as follows:
- *"The overall vision is to facilitate AI systems that **enhance human capabilities** and **empower individuals** and society as a whole while respecting **human autonomy** and **self-determination**."*
- The EU Act on AI aims to retain ***human agency***.
- Okay. Nice. But policy is not practice. In the real (commercial) world, we need to inform design… ***We need tools that maintain human values (including safety), that designers and developers can work with***…

Haiku
Human AI teaming Knowledge and Understanding for aviation safety

Home About Use Cases Products News & Events Contact Us

https://haikuproject.eu/

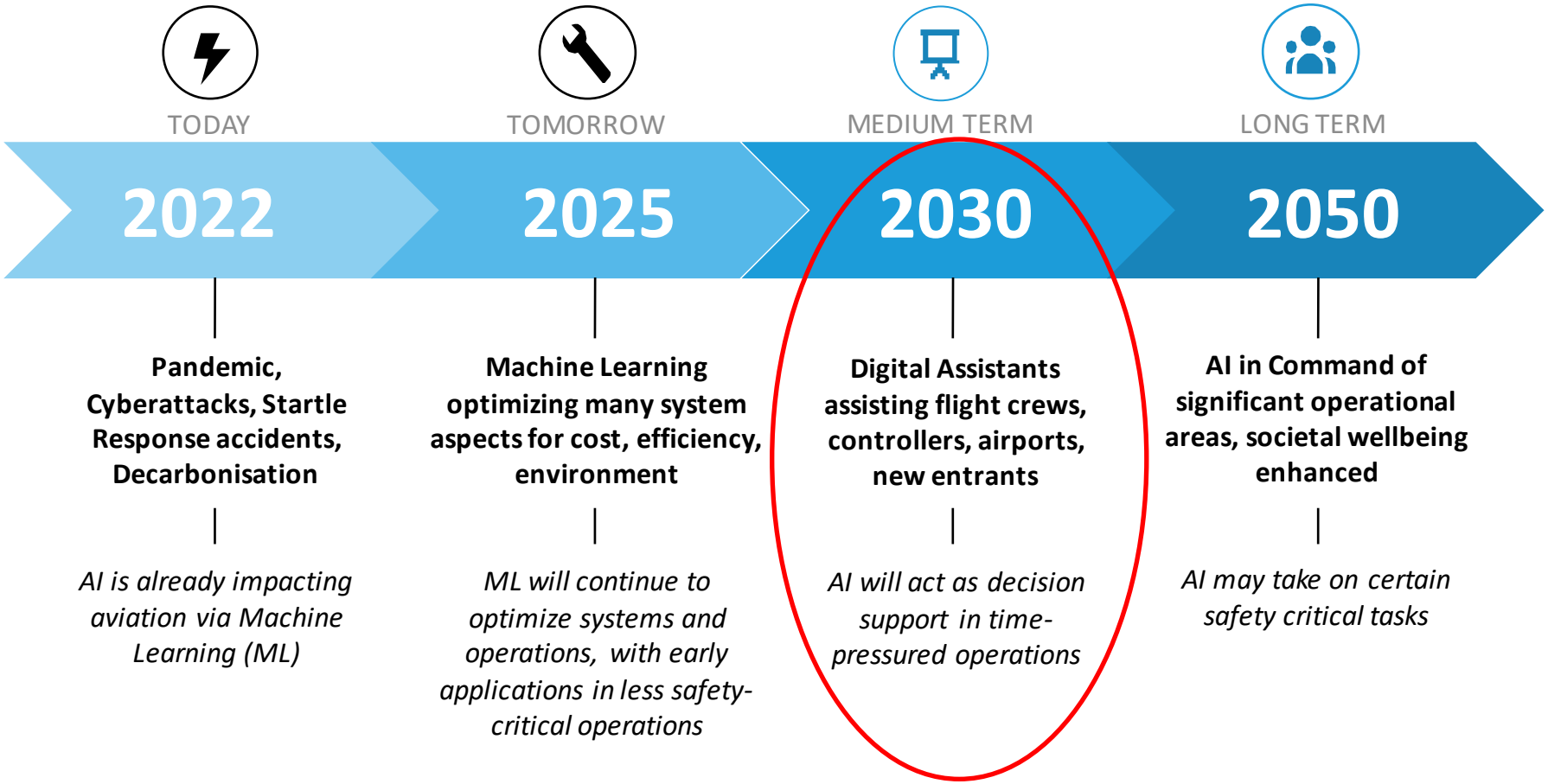
Our goal

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**.

HAIKU looks far ahead...



HAIKU Use Cases

Intelligent Assistant in the **cockpit** to assist in "**startle response**" adverse events

Led by ENAC

Intelligent Assistant for **tower controllers** to assist in **routine and repetitive tasks for aircraft on approach**

Led by SkyWAY

Intelligent Assistant in the **cockpit** to assist in **route planning/re-planning**

Led by TAVS

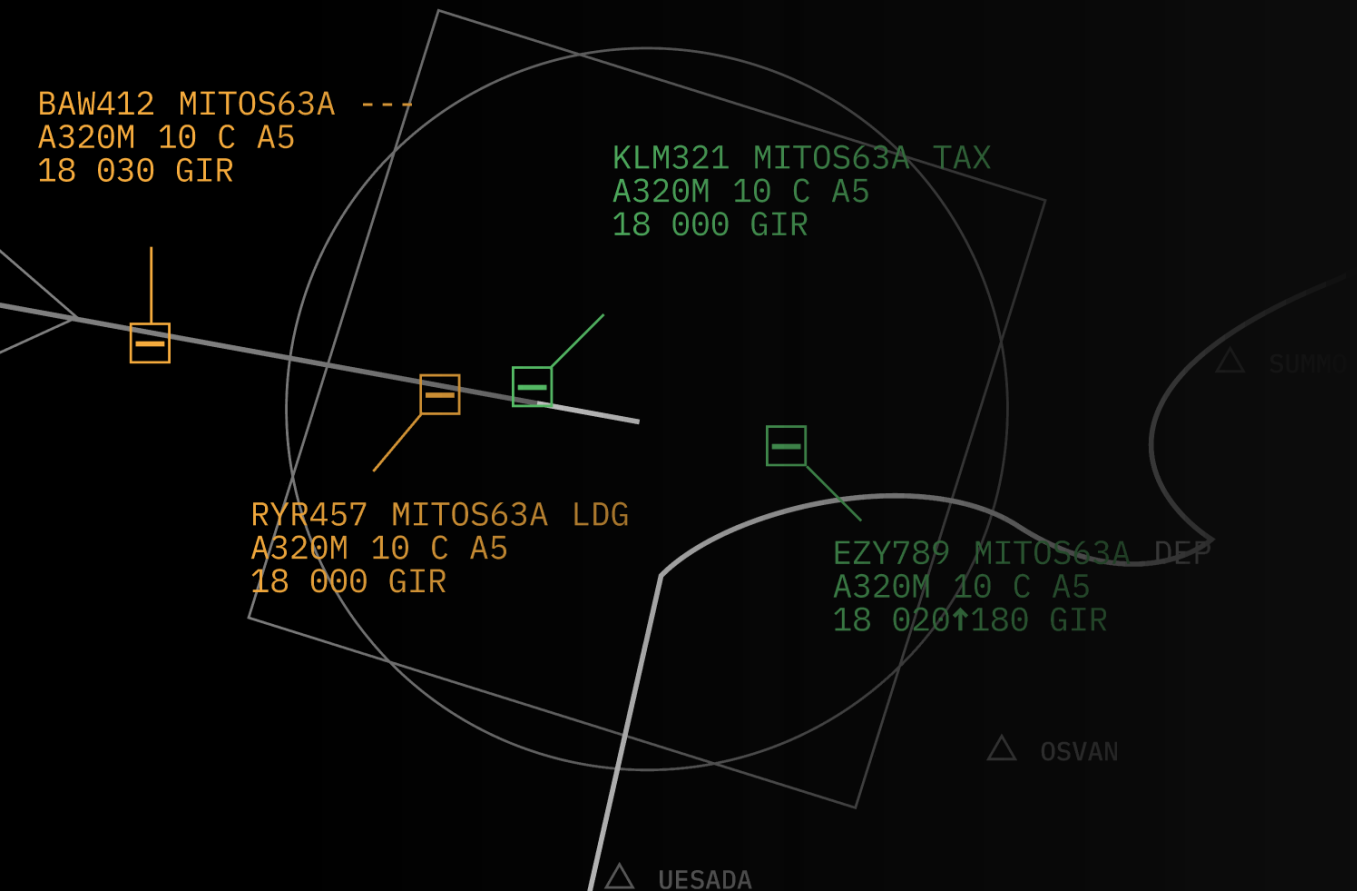
Intelligent Assistant for **Urban Air Mobility** coordinator to assist in **traffic management**

Led by LiU & LFV

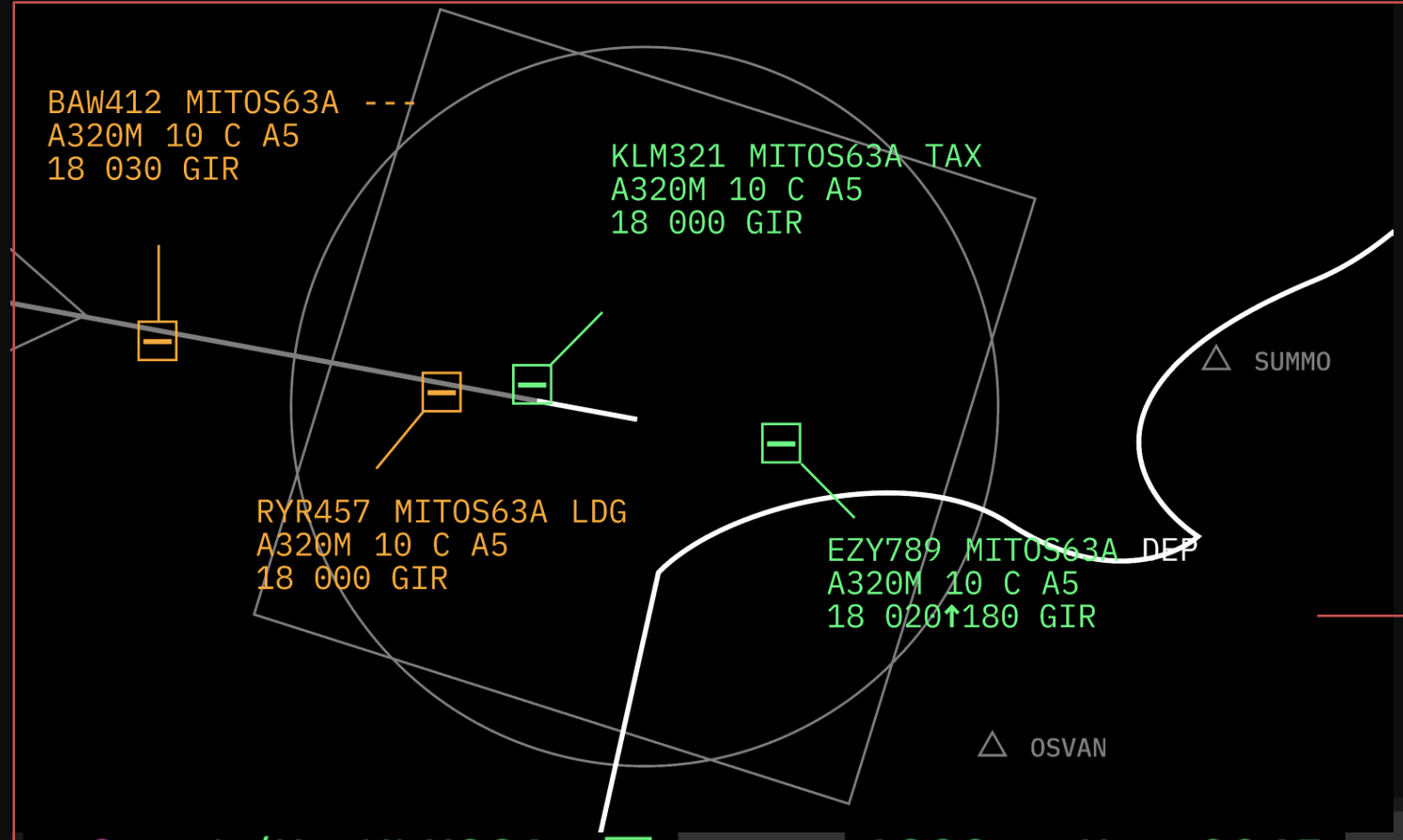
AI in the aviation workspace

Ask the right questions

HAIKU Use Case 4



- Intelligent Assistant for **tower controllers** to assist in **routine and repetitive tasks for aircraft on approach**
  - *Led by SkyWAY*



Intelligent Sequence Assistant (ISA)

2	L/U	KLM321	☑		A320	M	0945		10	45		MITOS3A	I

Seq. change: window for take/off of KLM321 too small

2 ↑	---	BAW412	☑		A320	M	0952

HAZOP Guidewords: No, Not Done, Too Late, Too Early, Other Than, Too Much, Too Little, More Than, Less Than, Reverse, Part of, As Well As

Human AI HAZOP

An App for Evaluating Human-AI Teaming systems

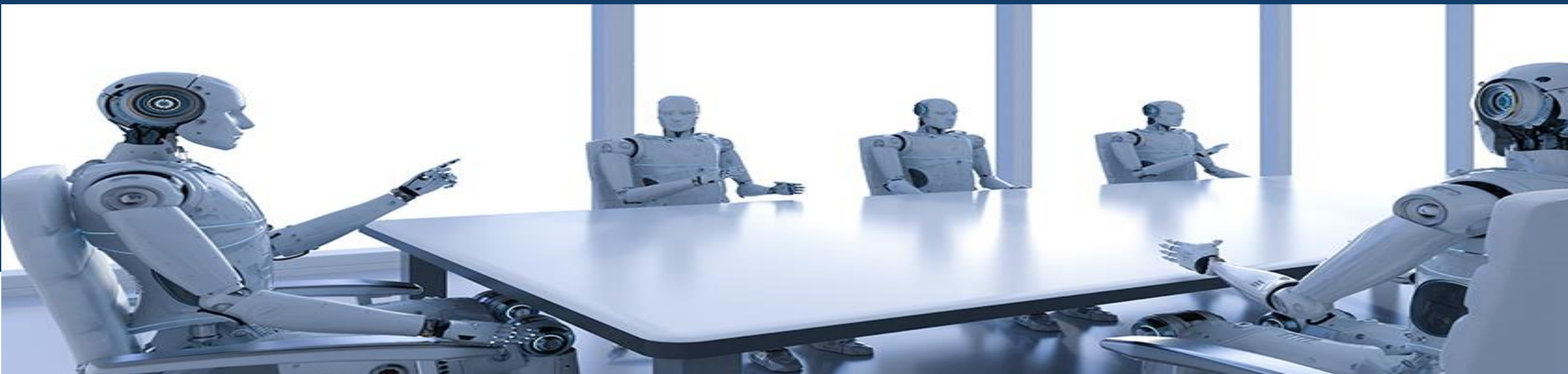
- 170 Guidelines
- Includes EASA HF, XAI and Ethics Guidelines
- Builds on SESAR Human Performance Assessment Process
- Requirements already tested on 2 HAIKU Use Cases

Supporting
European
Aviation

EUROCONTROL

PART 2:

How can we harness the power of AI to enhance safety culture?



Intelligent Assistants (IAs) & Safety Culture

- How will working with IAs impact the safety culture of pilots, controllers, ground staff, engineers, and management?
- If IAs perform really well, might aviation staff become complacent about safety?
- Equally, are there areas where Intelligent Assistants might actually improve safety culture?

future transportation

MDPI

Article

The Impact of Artificial Intelligence on Future Aviation Safety Culture

Barry Kirwan

EUROCONTROL, Centre du Bois des Bordes, 91222 Brétigny sur Orge CEDEX, France; barry.kirwan@eurocontrol.int

Abstract: Artificial intelligence is developing at a rapid pace, with examples of machine learning already being used in aviation to improve efficiency. In the coming decade, it is likely that intelligent assistants (IAs) will be deployed to assist aviation personnel in the cockpit, the air traffic control center, and in airports. This will be a game-changer and may herald the way forward for single-pilot operations and AI-based air traffic management. Yet in aviation there is a core underlying tenet that 'people create safety' and keep the skies and passengers safe, based on a robust industry-wide safety culture. Introducing IAs into aviation might therefore undermine aviation's hard-won track record in this area. Three experts in safety culture and human-AI teaming used a validated safety culture tool to explore the potential impacts of introducing IAs into aviation. The results suggest that there are indeed potential negative outcomes, but also possible safety affordances wherein AI could strengthen safety culture. Safeguards and mitigations are suggested for the key risk owners in aviation organizations, from CEOs to middle managers, to safety departments and frontline staff. Such safeguards will help ensure safety remains a priority across the industry.

Keywords: aviation; artificial intelligence; safety culture

check for updates

Citation: Kirwan, B. The Impact of Artificial Intelligence on Future Aviation Safety Culture. *Future Transp.* 2024, *4*, 349–379. https://doi.org/10.3390/futuretransp4020018

Academic Editor: Lynnette Dray

Received: 15 January 2024
Revised: 2 April 2024
Accepted: 3 April 2024
Published: 9 April 2024

1. Introduction

1.1. Overview of Paper

Currently, aviation is seen as a very safe mode of transport, and this is in part due to its safety culture. The question raised in this paper is about the potential impact of Artificial Intelligence (AI) on aviation safety culture. Although machine learning has already integrated AI into various aviation sectors, this paper specifically examines the prospects of more advanced AI systems. These systems may include intelligent assistants that have the potential to function semi-autonomously, or even autonomously, in collaboration with human crews and teams.

The paper begins by briefly outlining safety culture in aviation today, including how it is evaluated. The fast-developing area of AI itself is then outlined, focusing on different 'levels' of AI autonomy and the concept of human-AI teaming. This wide-ranging exploration of AI is necessary to envision how human crews and Intelligent Assistants (IAs) might work together in a range of future AI settings (e.g., cockpit, air traffic tower and operations room, airports). The application of an aviation safety culture method is then analyzed in relation to future human-AI teaming scenarios to assess potential safety culture outcomes. The paper concludes by noting the most serious threats to safety culture posed by AI, and how to safeguard against them, as well as suggesting ways forward to harness the potential safety culture benefits from human-AI teaming.

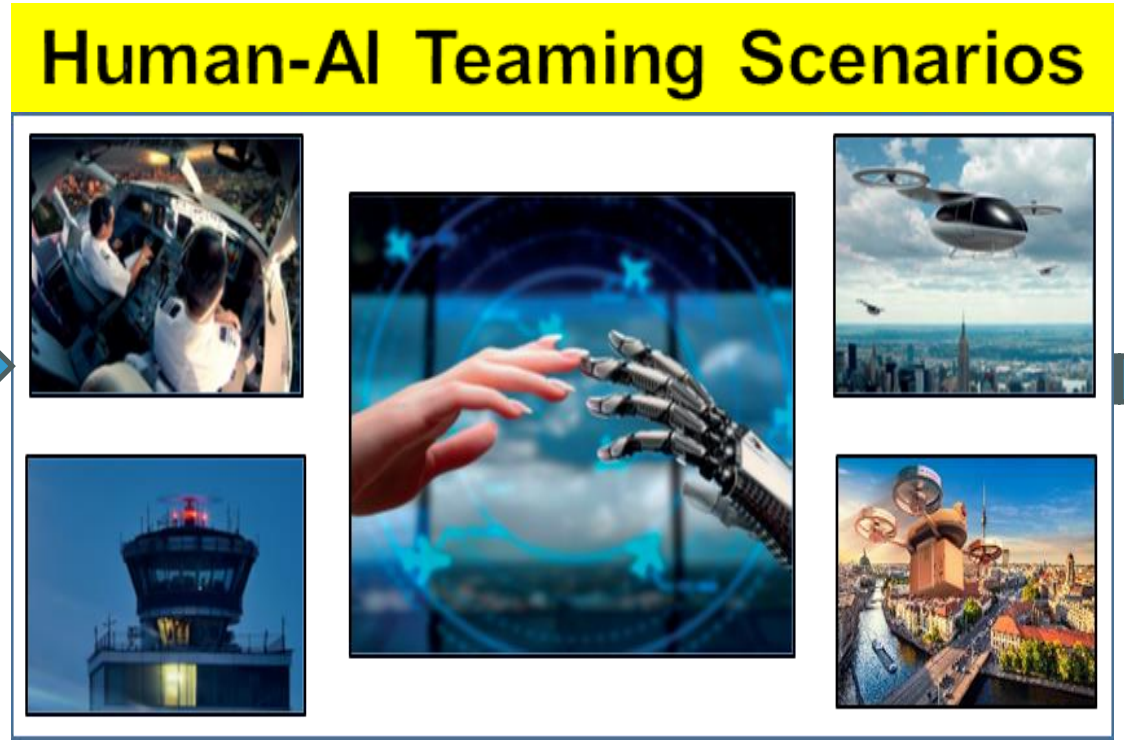
1.2. Safety Culture—An Essential Ingredient of Aviation Safety

In European commercial civil aviation today, safety in terms of low accident rates is very strong, with no major accidents involving EU-registered aircraft in commercial air transport over the past seven years [1], although there have still been fatal air crashes

https://www.mdpi.com/2673-7590/4/2/18

Analytic Approach

Safety Culture Questionnaire



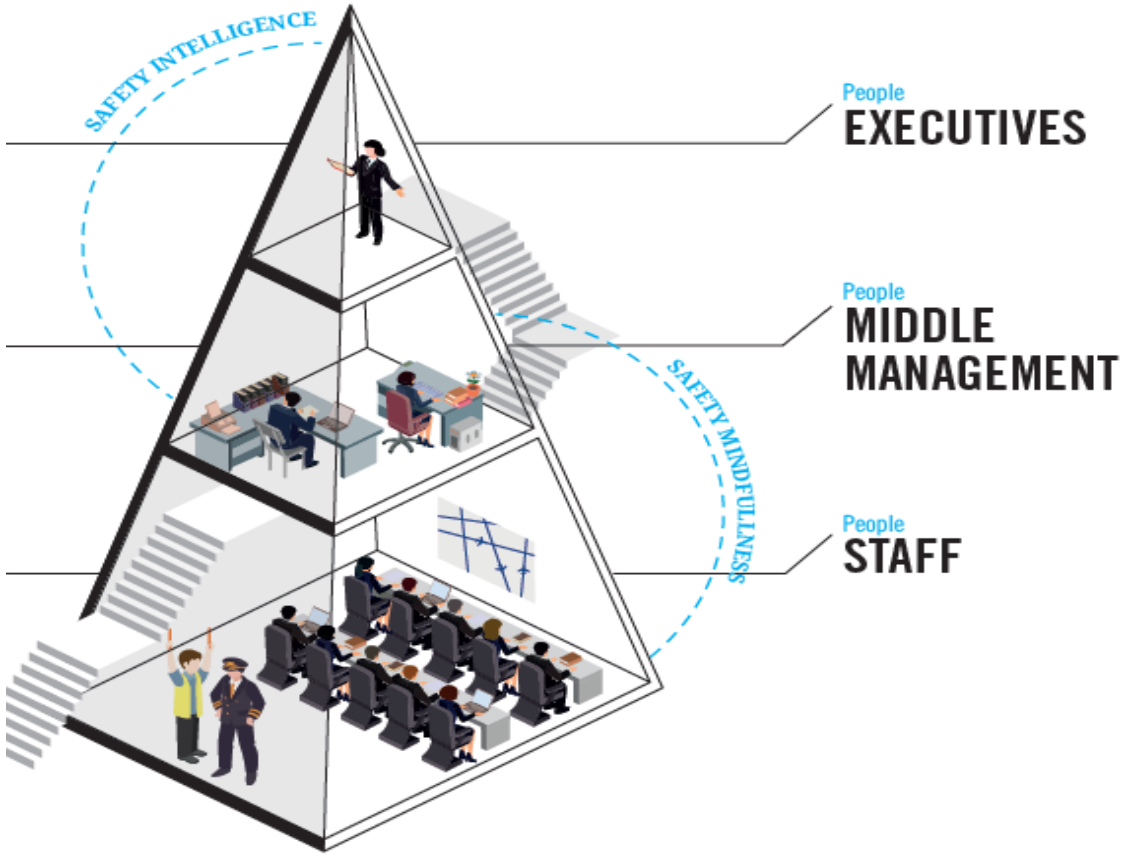
Concerns

Affordances

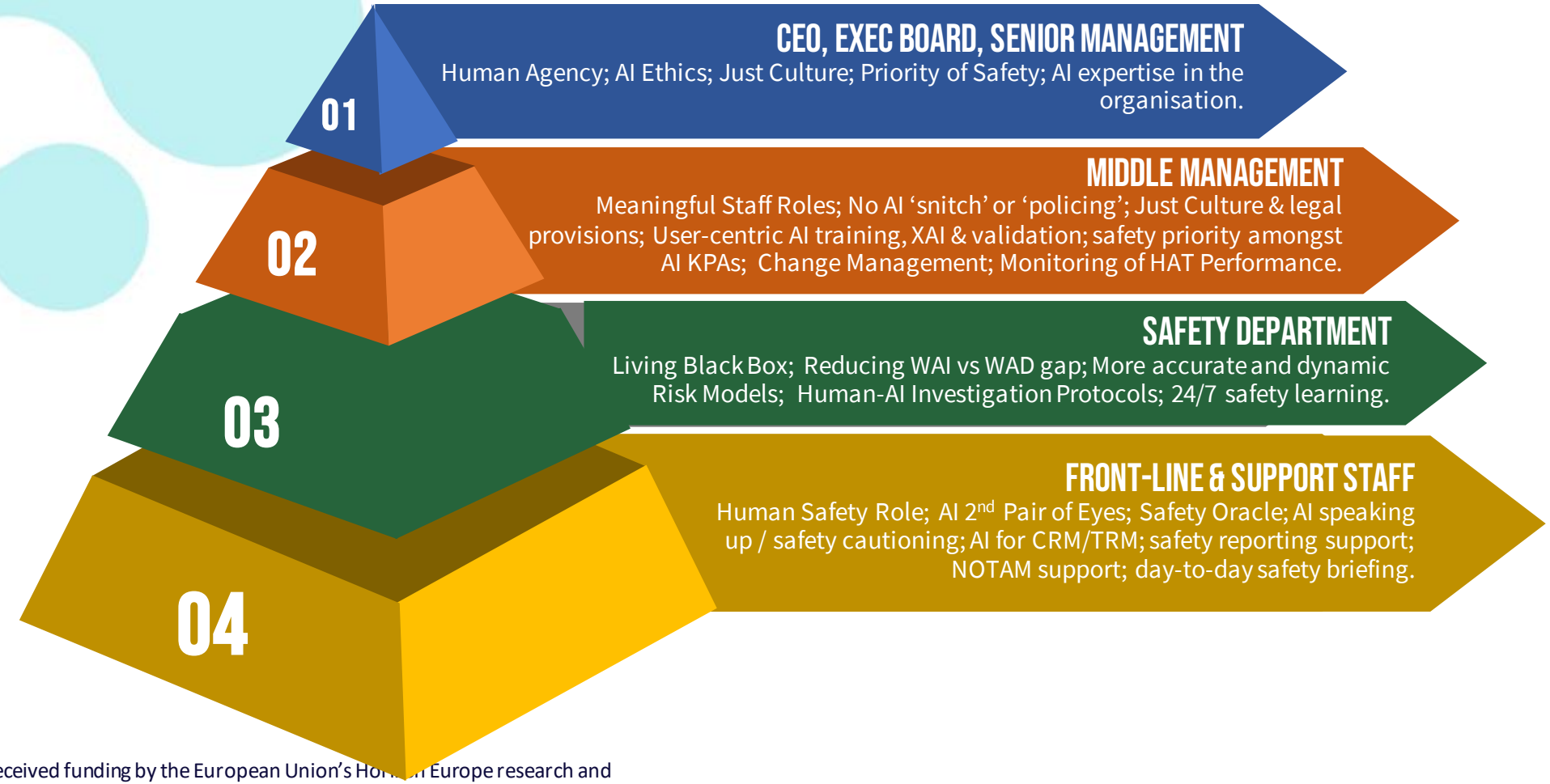
Concerns	Affordances
Humans may become less concerned with safety if the IA is seen as handling safety aspects. This is an extension of the 'complacency' issue with automation, and may be expected to increase as the IA's autonomy increases.	The IA could 'speak up' if it assesses a human course of action as unsafe.
Humans may perceive a double-bind: if they follow 'bad' IA advice or fail to follow 'good' advice, and there are adverse consequences, they might find themselves being prosecuted. This will lead to lack of trust in the IA.	The IA could be integrated into Crew Resource Management practices, helping decision-making and post-event review in the cockpit or air traffic Ops Room.
If the IA reports on human error or human risk-taking or other 'non-nominal behaviour' it could be considered a 'snitch' for management, and may not be trusted.	The IA could serve as a living black box recorder, recording more of decision-making strategies than is the case today.
If IA recordings are used by incident and accident investigators, Just Culture policies will need to address such usage both for ethical reasons and to the satisfaction of the human	If the IA is able to collect and analyse day-to-day safety occurrence information it may be seen as adding objective (dispassionate) evidence and a more balanced sassessment of

EUROCONTROL

Who are the 'Risk Owners'?



Safeguards for Future Safety Culture Assurance



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

04

FRONT-LINE & SUPPORT STAFF

Human Safety Role; AI 2nd Pair of Eyes; Safety Oracle; AI speaking up / safety cautioning; AI for CRM/TRM; safety reporting support; NOTAM support; day-to-day safety briefing.

- People should have meaningful jobs and safety roles, supported by AI
- Develop a 'safety oracle' for operational personnel, for advice on safety considerations
- AI can act as a *second pair of eyes*, aiding in an emergency or noting a safety issue or deviation or risky course of action by the human operator
- Human-AI Resource Management (HAIRM) focused on safe Human-AI Teamworking, including *speaking up for safety* (whether AI or human)
- An aid to safety reporting for operational staff
- Support with NOTAMs and daily briefing material via AI-assisted personalization, also acting as a reminder in case anything is forgotten or overlooked.

03

SAFETY DEPARTMENT

Living Black Box; Reducing WAI vs WAD gap; More accurate and dynamic Risk Models; Human-AI Investigation Protocols; 24/7 safety learning.

- Explore the potential for using the AI as a ‘Living Black Box’
- Collect evidence on actual operational practices vs. procedures (WAD vs. WAI)
- Develop Incident Investigation Protocols for incidents involving AI
- Enhance risk models with AI interactions and increased operational realism, to be used dynamically for day-to-day hazard management
- Use AI to record ‘when things go right’ in order to promulgate good practices
- Move towards a 24/7 safety learning system

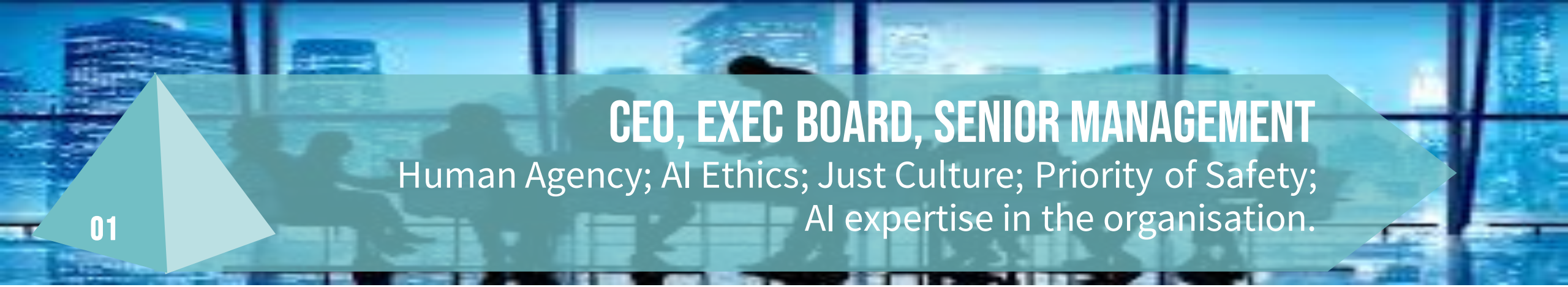
02

MIDDLE MANAGEMENT

Meaningful Staff Roles; No AI ‘snitch’ or ‘policing’; Just Culture & legal provisions; User-centric AI training, XAI & validation; safety priority amongst AI KPAs; Change Management; Monitoring of Human-AI Teaming Performance.

MM must prepare staff for AI introduction, and manage & monitor its deployment

- MM need to ensure that
  - Staff have meaningful jobs, are trained on AI basics, and can recognize if AI goes wrong
  - AI does not act as ‘snitches’ on staff, or police them
  - Just Culture ideals can be translated into effective / trusted principles and practices
  - AI systems are user-centric, designed with user input, and validated with users in simulations
  - The AI’s advice or decisions makes sense to the humans concerned (explainability)
  - The AI’s trade-offs between KPAs do not jeopardise safety



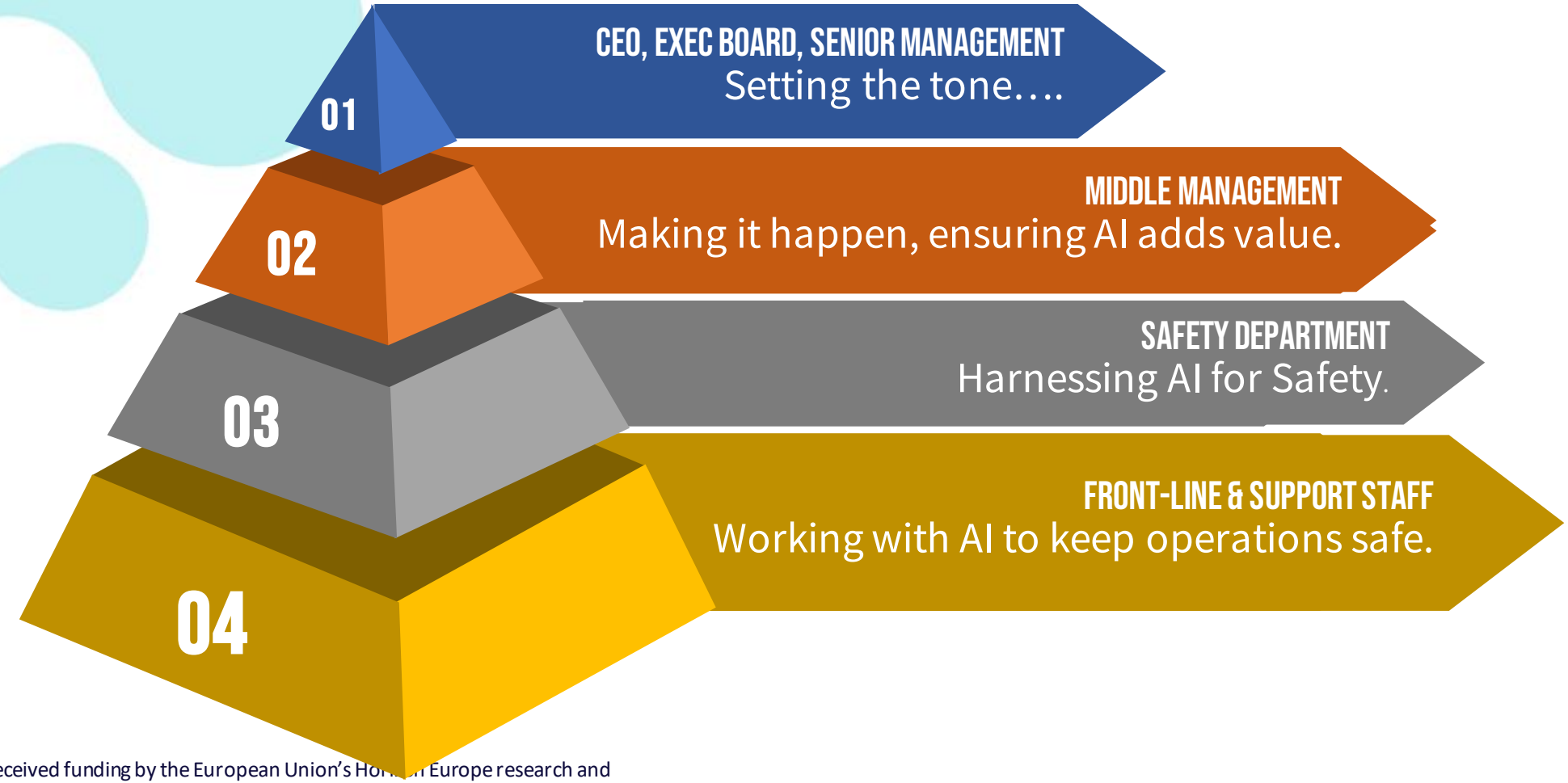
CEO, EXEC BOARD, SENIOR MANAGEMENT

01

Human Agency; AI Ethics; Just Culture; Priority of Safety;
AI expertise in the organisation.

- ✓ An authentic message that safety is the priority, and that '*people create safety*'.
- ✓ A *code of ethics* related to the use of AI in the organization.
  - *Respect for human autonomy*: **AI systems *should not*** subordinate, coerce, deceive, manipulate, condition or herd humans. **AI systems *should*** augment, complement and empower human cognitive, social and cultural skills, leave opportunity for human choice and secure human oversight over work processes, and support the creation of meaningful work.
  - *Prevention of harm*: AI must not cause harm or adversely affect humans, and must protect human dignity.
- ✓ A *Just Culture policy and framework* which deals with AI accountability in the case of an accident.
- ✓ Strategic decision-making on *Internal AI expertise* in the organization

Safeguards for Future Safety Culture Assurance



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

3 Key Development Areas for ST

Just Culture – develop principles then test in ‘legal sandboxes’

SMS – not just old wine in new bottles (same for SMS SoE)

Human Factors will be a key enabler

We 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."

SUPPORTING EUROPEAN AVIATION