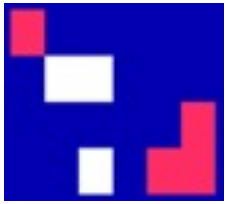
MAI4CAREU | Master programmes in Artificial Intelligence 4 Careers in Europe



Responsibility and automation in Socio-technical systems

The case of air traffic management

Giuseppe Contissa

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Responsibility and automation

- How do we allocate responsibilities among the various participants in complex socio-technical organisations?
- In particular, what is the role of humans interacting with highly automated systems?
- Who is responsible for accidents in highly automated systems?

“responsibility”

As captain of the ship, X was **responsible** for the safety of his passengers and crew. But on his last voyage he got drunk every night and was **responsible** for the loss of the ship with all aboard.

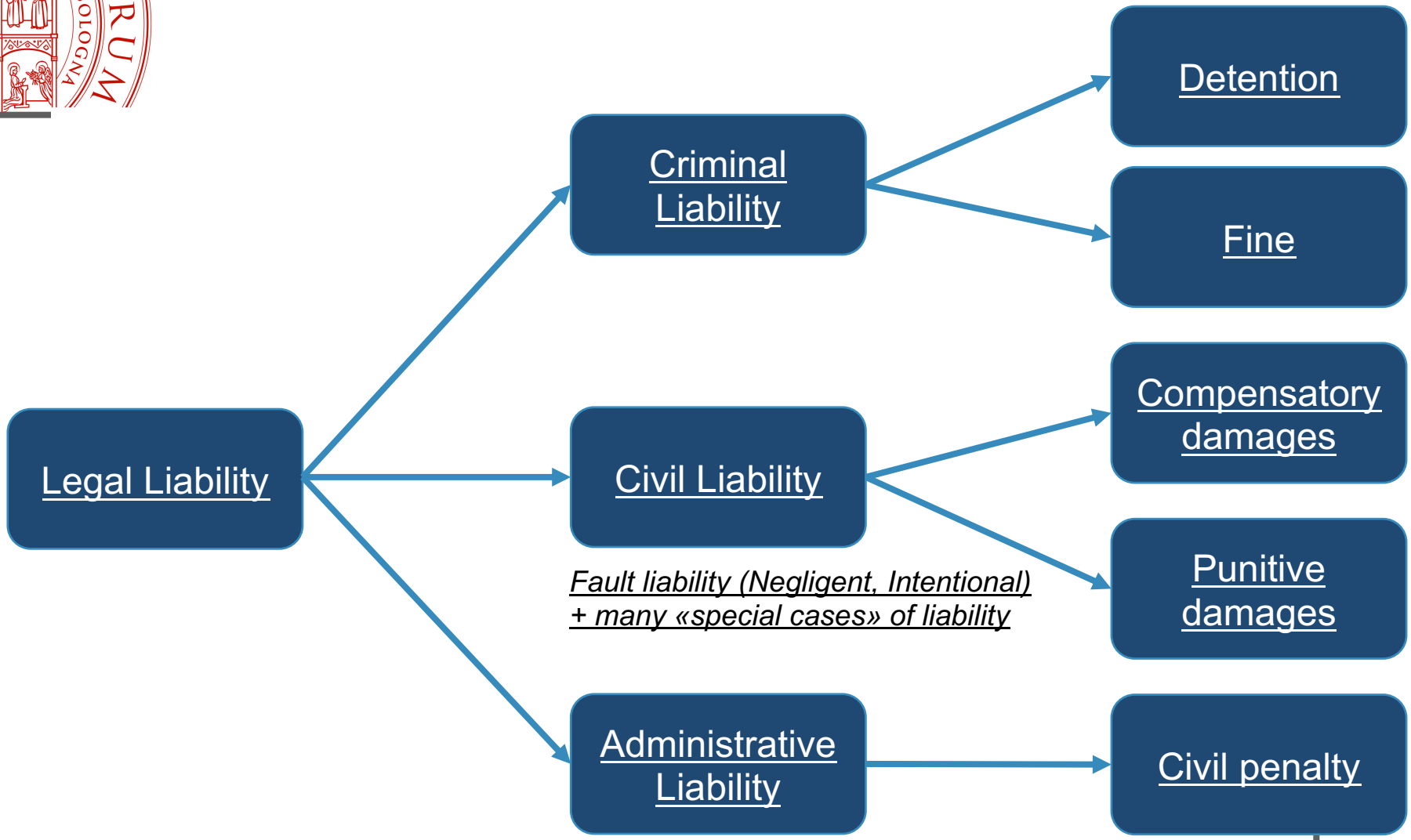
It was rumoured that he was insane, but the doctors considered that he was **responsible** for his actions. Through out the voyage he behaved quite **irresponsibly**, and various incidents in his career showed that he was not a **responsible** person.

He always maintained that the exceptional winter storms were **responsible** for the loss of the ship, but in the legal proceedings brought against him he was found criminally **responsible** for his negligent conduct, and in separate civil proceedings he was held legally **responsible** for the loss of life and property.

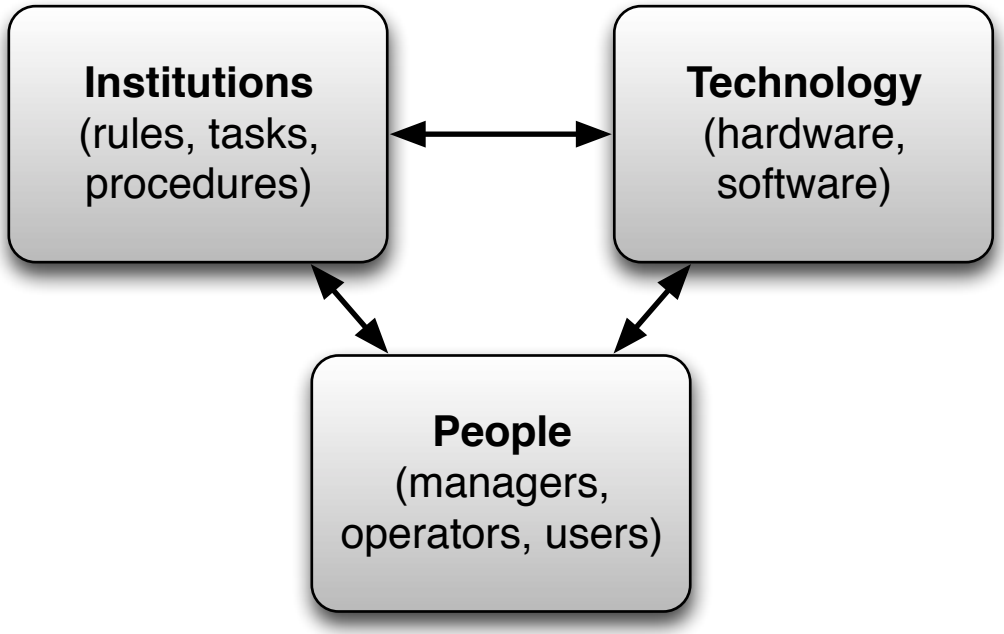
He is still alive and he is morally **responsible** for the deaths of many women and children.

(Hart,H.L.A.,Punishment and Responsibility: Essays in the Philosophy of Law, 1970)

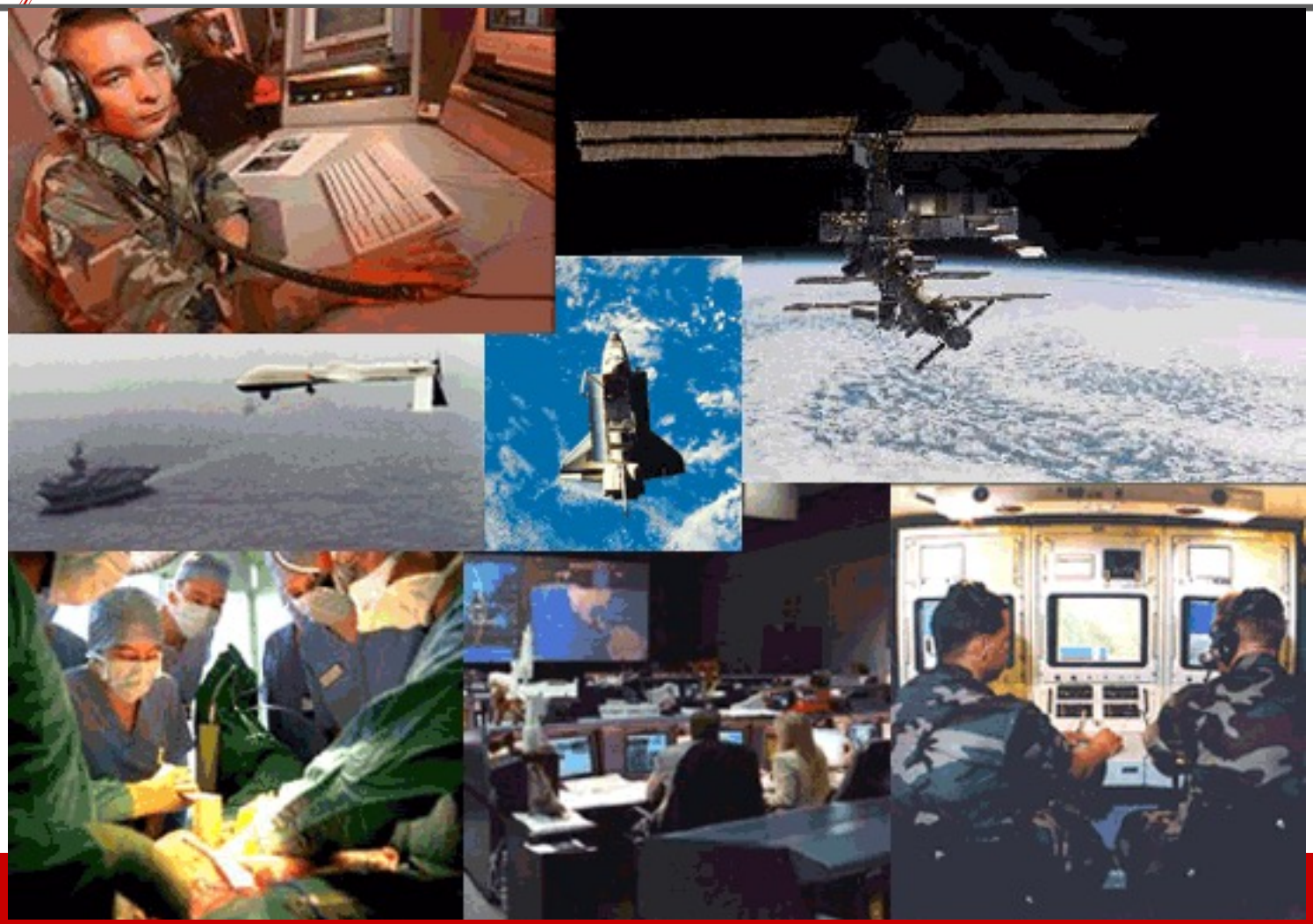
Liability (legal responsibility)



Socio-technical systems: basic structure



Socio-technical systems: examples



The future of ATM

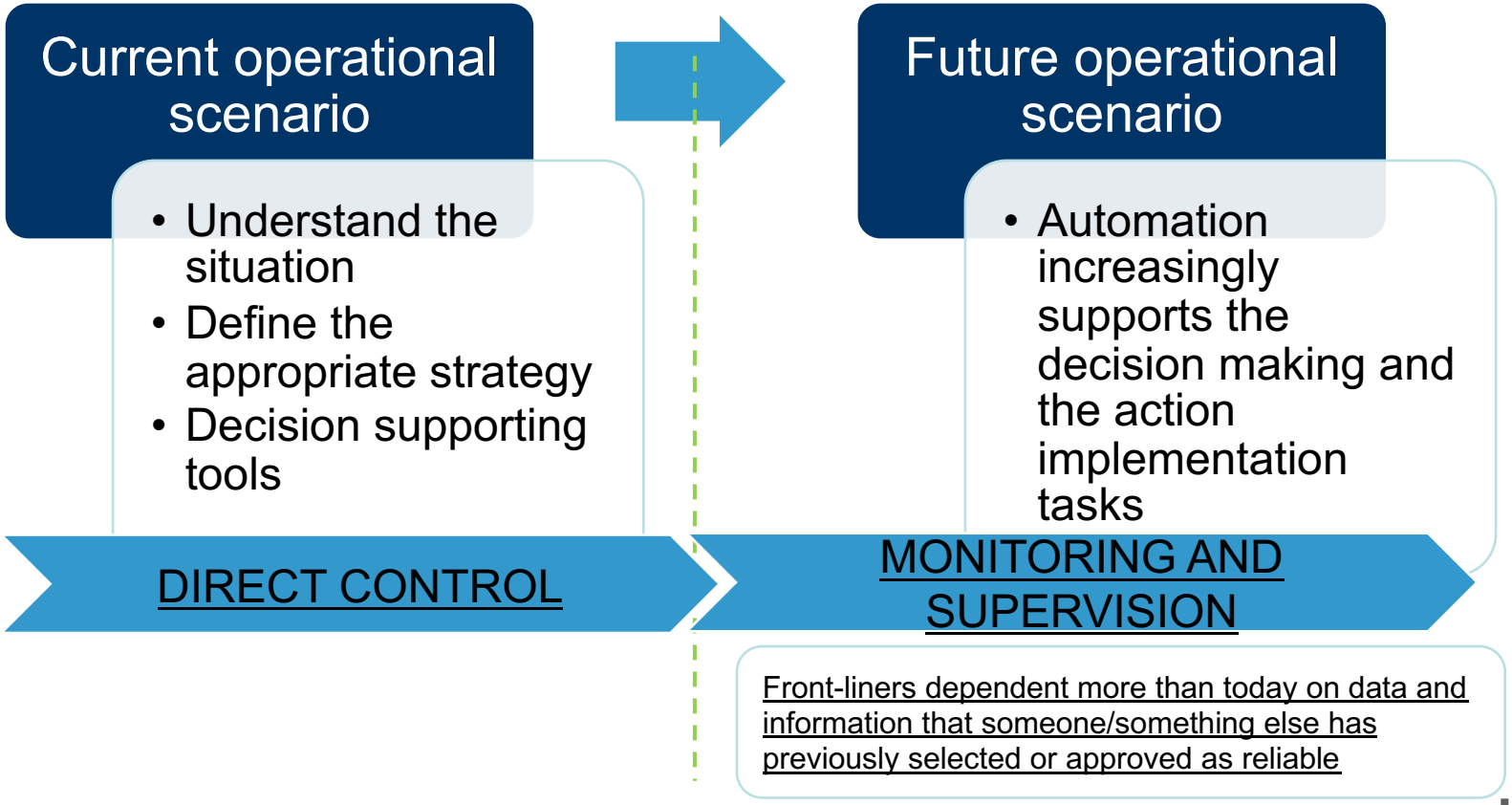
- In the time horizon of SESAR, that is over the next 30 years, a new generation of air traffic management systems will be developed.

- Such systems will be highly automated. They will make choices and engage in actions with some level of human supervision, or even without any such supervision.

Automation and the future ATM scenario

- New generation of ATM systems to increase capacity, safety, efficiency and sustainability
- Higher levels of automation

AUTOMATION SUPPORT



Implications of automation

- Delegation of task from operators to technology
- Humans as controllers and supervisors
- Hybrid agency (symbiosis/coagency → joint cognitive systems)
- Machine intelligence and autonomy (= independence + cognitive skills)
- The challenge of complexity (technological, “many hands”)

Automation: not all or nothing

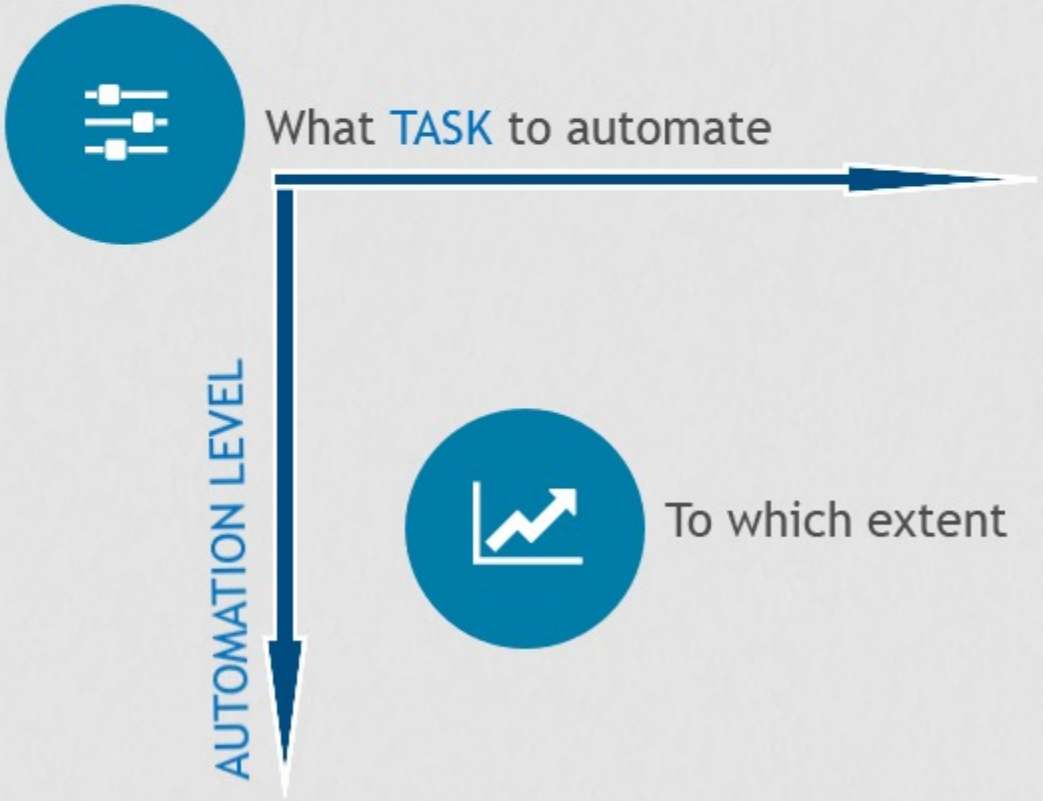
- Not just **substitution of a human operator**
- Support to human capabilities

in performing tasks

- Some degree of cooperation is usually required

Automation: not all the same

Different tasks involve different **psychomotor** and **cognitive** functions, which in turn implies the adoption of different automation solutions.



The level of automation taxonomy (SESAR 1)

From INFORMATION to ACTION →

INCREASING AUTOMATION ↓

A INFORMATION ACQUISITION	B INFORMATION ANALYSIS	C DECISION AND ACTION SELECTION	D ACTION IMPLEMENTATION
A0 Manual Information Acquisition	**B0** Working memory based Information Analysis	**C0** Human Decision Making	**D0** Manual Action and Control
A1 Artefact-Supported Information Acquisition	**B1** Artefact-Supported Information Analysis	**C1** Artefact-Supported Decision Making	**D1** Artefact-Supported Action Implementation
A2 Low-Level Automation Support of Information Acquisition	**B2** Low-Level Automation Support of Information Analysis	**C2** Automated **Decision Support**	**D2** Step-by-Step Action Support
A3 Medium-Level Automation Support of Information Acquisition	**B3** Medium-Level Automation Support of Information Analysis	**C3** Rigid Automated **Decision Support**	**D3** Slow-Level **Support** of Action Sequence Execution
A4 High-Level Automation Support of Information Acquisition	**B4** High-Level Automation Support of Information Analysis	**C4** Low-Level Automatic **Decision Making**	**D4** High-Level **Support** of Action Sequence Execution
A5 Full Automation Support of Information Acquisition	**B5** Full Automation Support of Information Analysis	**C5** High-Level Automatic **Decision Making**	**D5** Low-Level **Automation** of Action Sequence Execution
		C6 Fulll Automatic **Decision Making**	**D6** Medium-Level **Automation** of Action Sequence Execution
			D7 High-Level **Automation** of Action Sequence Execution
			D8 Full **Automation** of Action Sequence Execution

A condensed version of the LOAT matrix

ROT / Use of video cameras in the control tower

A INFORMATION ACQUISITION	
A0	Manual Information Acquisition
A1	Artefact Supported Information Acquisition
A2	Low Level Automation Support of Info Acquisition
A3	Med. Level Automation Support of Info Acquisition
A4	High Level Automation Support of Info Acquisition
A5	Full Automation Support of Info Acquisition

The system **supports** the human in **acquiring information** on the process s/he is following. Filtering and/or highlighting of the most relevant information are **up to the human**.

Activation of speed vectors by controllers

B
INFORMATION ANALYSIS

- **B0** Working-memory based Information Analysis
- **B1** Artefact Supported Information Analysis
- **B2** Low Level Automation Support of Info Analysis
- **B3** Med. Level Automation Support of Info Analysis
- **B4** High Level Automation Support of Info Analysis
- **B5** Full Automation Support of Info Analysis

Based on user's request, the system **helps** the human in comparing, combining and analysing different information items regarding the status of the process being followed.

AMAN sequence of landing aircraft

C
DECISION AND ACTION SELECTION

- **C0** Human Decision Making
- **C1** Artefact Supported Decision Making
- **C2** Automated Decision Support
- **C3** Rigid Automated Decision Support
- **C4** Low Level Automatic Decision Making
- **C5** High Level Automatic Decision Making
- **C6** Full Automatic Decision Making

https://www.eurocontrol.int/sites/default/files/article/content/documents/nm/fasti-aman-status-review-2010.pdf, page 16

The system **proposes one or more decision alternatives** to the human, leaving freedom to the human to generate alternative options. The human can **select** one of the alternatives proposed by the system **or** her/his own one.

Autopilot

D ACTION IMPLEMENTATION
D0 Manual Action and Control
D1 Artefact Supported Action Implementation
D2 Step by step Action Support
D3 Low Level Support of Action Sequence Execut.
D4 High Level Support of Action Sequence Execut.
D5 Low Level Automation of Action Sequence Exec
D6 Medium Level Automat. of Action Seq. Execut.
D7 High Level Automation of Action Seq. Execut.
D8 Full Automation of Action Sequence Exec

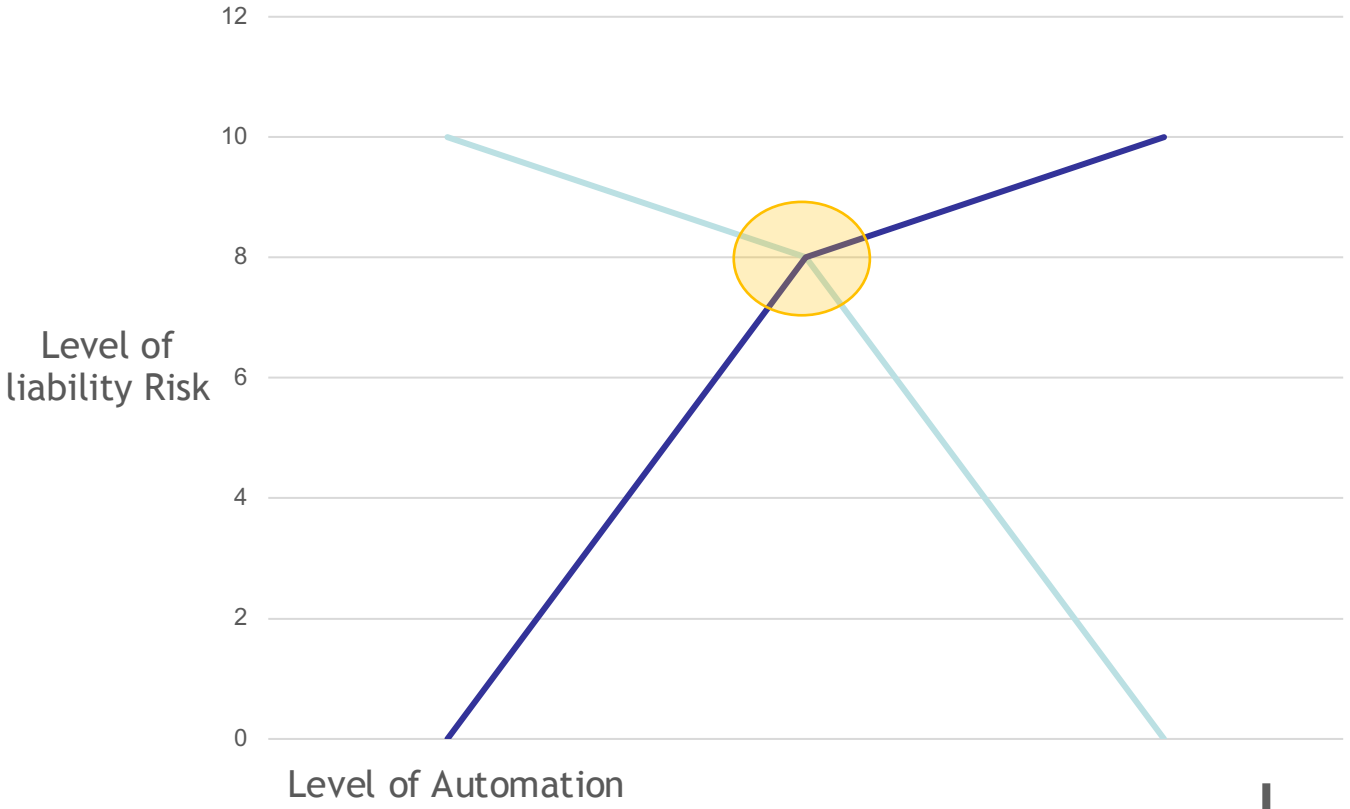
The system performs automatically a sequence of actions **after activation by the human**. The human can **monitor** all the sequence and can **interrupt** it during its execution.

Some questions

- How automation transforms operators' roles and tasks? What impact on their responsibilities?
- Who is responsible for the behaviour of systems that humans cannot fully monitor and control?
- Who is responsible for information supplied by automated systems that the human cannot verify?

Level of automation and liability risk

- Increasing the level of automation will proportionally increase the liability risk for the **technology provider** and decrease the liability risks for the human operator.
- However, the employment of technologies with **intermediate levels of automation** may result in a high liability risk both for the technology provider and the human operator



Fragmentation of tasks and liability

The **fragmentation of tasks** may results in uncertainity and complexity of procedures

Human operator

- difficult to asses how and who should carry out each task
- high liability risk for negligence

Technology provider

- difficult to design HMI to adequately support decision making and/or to provide exahustive information
- high product liability risk, caused by design and information defects.

Highly automated systems/AI systems: liability shift

Liability for injury/harm caused by technological failure **gradually** transferred to the **organisation(s) developing / using/ maintaining** the technology

Grounds for the attribution:

- Product liability (no-fault liability, grounded on **defectiveness**, in particular wrt design defects and warning defects)
- Organisational / no-fault liability: generation of risks and ability to prevent them (and possibility to distribute losses)
- Vicarious liability (for faults of employees, residual)
- In the future: Liability for failing to deploy automated/AI systems?

– Liability assessment should be carried out as soon as possible in the life cycle of technology.
– Liability allocation related to level of automation of technology, in particular to the cognitive functions of the automated/AI technology and on the H-M interaction
– To be assessed in relation to role of technology in accidents

Highly automated systems/AI systems: liability shift /2

Individual liability (*criminal/civil*, fault liability) would persist

- only when the human acted with an intention to cause harm or with recklessness (e.g. Just Culture)?... Or..
- always, human as «moral crumple zone» (Elish 2018)?

- What about decisions taken by humans when interacting with automated/AI systems?

Other important issues on liability

- **Liability and standards/certification**
  - Liability shield for the producer?
  - “Legitimate” expectation for the user/operator?
  - Liability of certificators / standard setters

- **Right of recourse. Who will pay in the end?**
  - In complex systems, the *law may channel liability* towards one actor (e.g. in ATM, the air carrier), *but recourse against the one who had control* over the malfunctioning component of the system

- **The role of insurance**
  - Mandatory insurance for producers/manufacturers?
  - Specific issues of highly automated /AI systems (cyber risk, wilful misconduct)

- **International context: “forum shopping”**

Open issue: Decision making authority

- ***Effective* decision-making authority in socio-technical systems**
  - **Joint cognitive systems?**
  - **The model described (or prescribed) by laws, regulations, procedures:**

  - Right not to be subject to (fully) automated individual decision-making (Art 22 **GDPR)**: “[oversight of the decision] should be carried out by someone who has the authority and competence to change the decision” (Art29WP)
  - **Aviation**: ICAO Annex 2, sec. 2.3.1 Responsibility of pilot-in-command (ultimate authority, ultimate responsibility)
  - Vienna Convention on **Road Traffic**, Art. 1(v) "Driver" means any person who drives a motor vehicle or other vehicle (but amendments for ADS)
  - Art 14 new **AI ACT** proposal, human oversight for high-risk AI systems

EFFECTIVE DECISION-MAKING AUTHORITY

What about decisions to be taken jointly with AI, in conditions of limited resources – time, information, explanations? E.g.:

- **Medical diagnosis** assisted by AI (Lagioia, Contissa 2020)
- **Frontex border** controls: «12 seconds to decide»

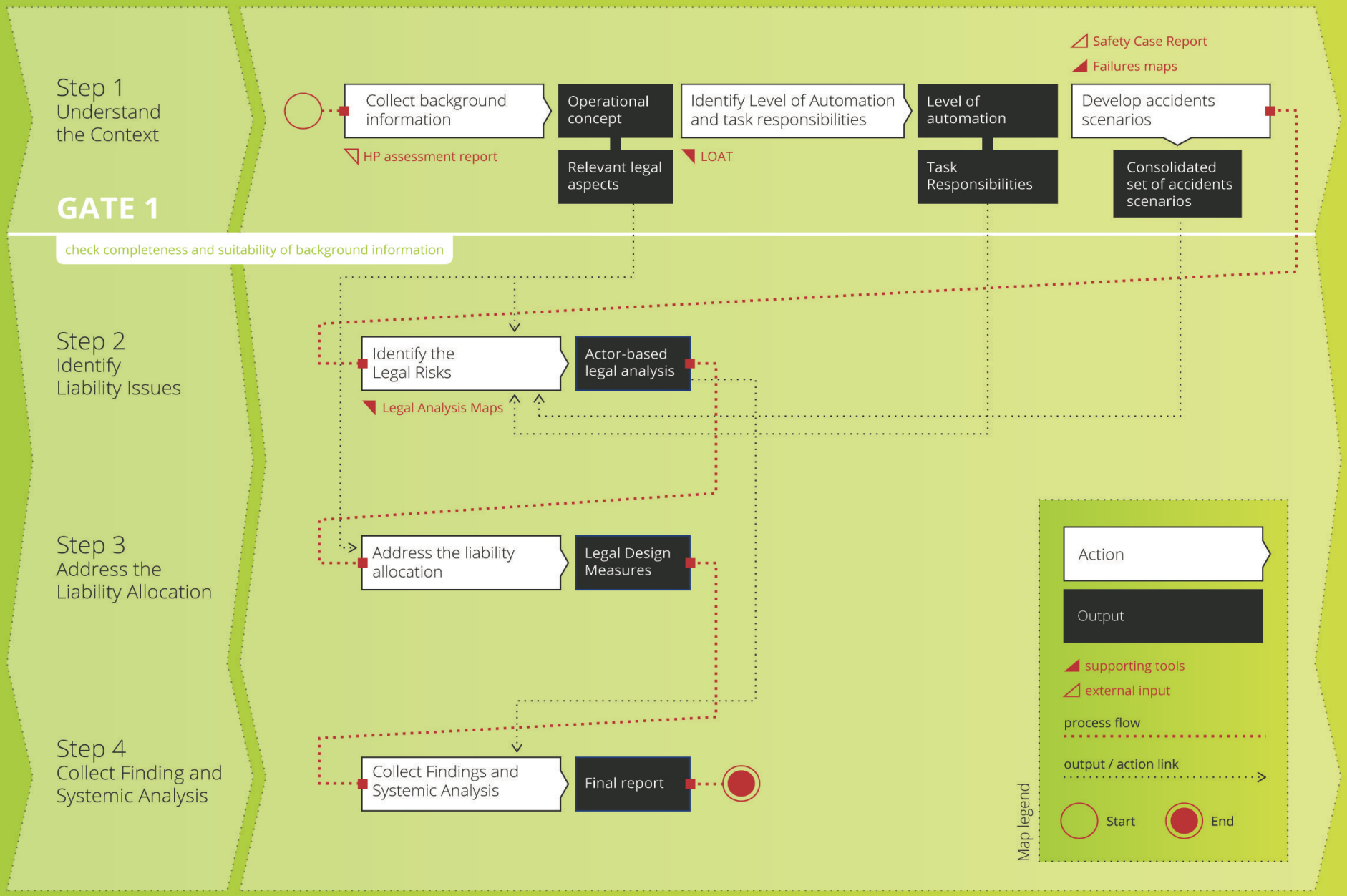
*Machine intelligence is fundamentally **alien**, and often, the entire purpose of an AI system is to learn to do or see things in ways humans cannot[..]*

*Ultimately, the **lack of a principled basis to contradict AI predictions implies that the reasonableness of an action in individual cases must be tied to the decision to use AI as a general matter.** (Selbst 2019)*

*Owing to the **evidence** in their favor (stipulated by definition), it is more appropriate to think of **expert robots as above average in their ability to make decisions that will produce desirable outcomes […]***

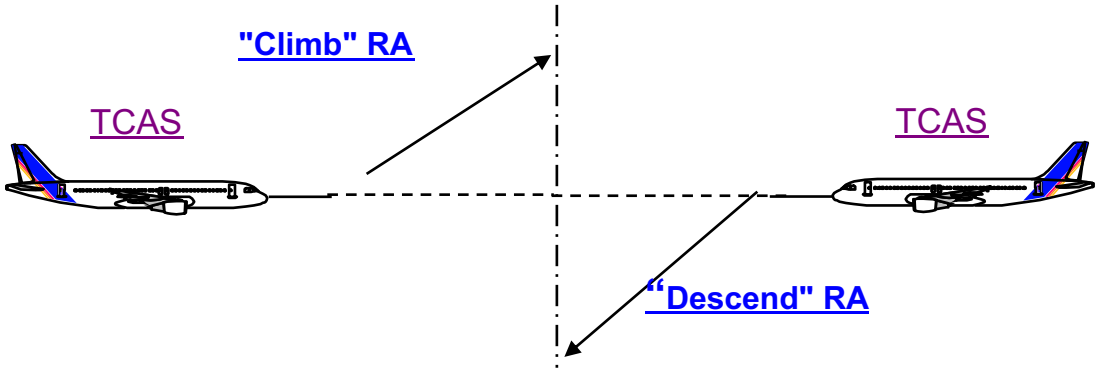
*This fact suggests that **granting a general decision-making authority to human experts will be problematic once expert robots are properly on the scene**.*
(Millar, Kerr 2018)

Legal Case


Step 1
Understand the Context
Collect background information
HP assessment report
Operational concept
Relevant legal aspects
Identify Level of Automation and task responsibilities
LOAT
Level of automation
Task Responsibilities
Develop accidents scenarios
Safety Case Report
Failures maps
Consolidated set of accidents scenarios
GATE 1
check completeness and suitability of background information
Step 2
Identify Liability Issues
Identify the Legal Risks
Legal Analysis Maps
Actor-based legal analysis
Step 3
Address the Liability Allocation
Address the liability allocation
Legal Design Measures
Step 4
Collect Finding and Systemic Analysis
Collect Findings and Systemic Analysis
Final report
Map legend
Action
Output
supporting tools
external input
process flow
output / action link
Start
End

ACAS/TCAS II (TRAFFIC COLLISION AVOIDANCE SYSTEM)

http://www.skybrary.aero/index.php/TCAS



- Visual and aural advices
- 2 types of advisories: **TA (Traffic Advisory)** and **RA (Resolution Advisory)**

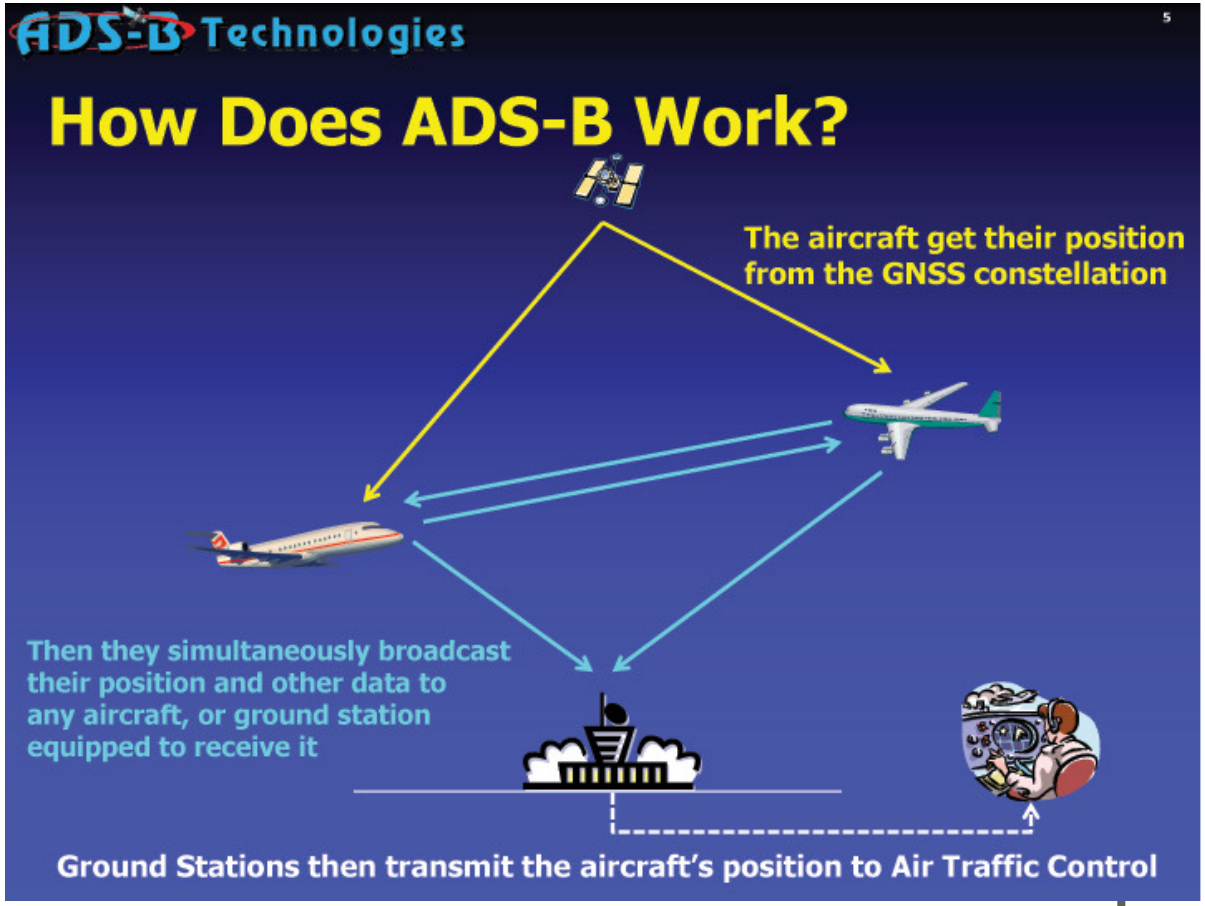
- RA shall be executed by the crew; The system decides the best option and informs the human
- During the execution by the pilot the system provides guidance through continuous visual and aural feedback

ACAS X and ADS-B

ACAS X will replace the current generation of systems ACAS/TCAS II

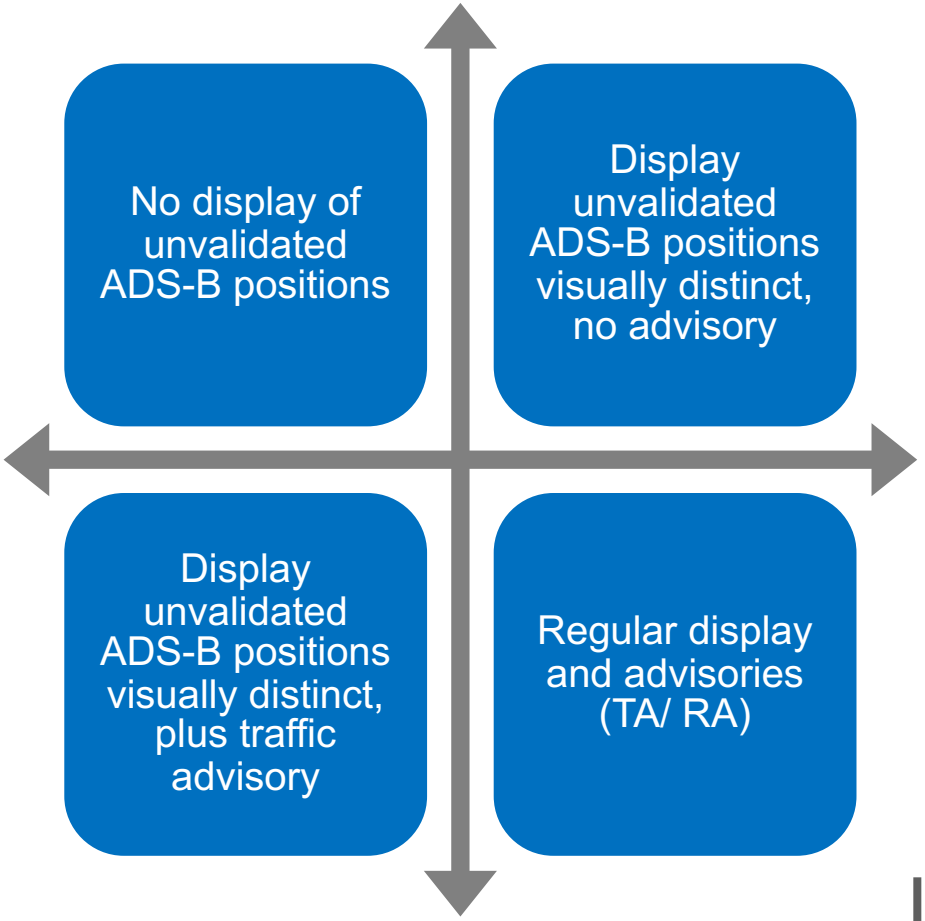
It will use new sources of surveillance data, including ADS-B (Automatic Dependent Surveillance Broadcast)

ADS-B is an enabler for the change from radar based towards satellite based aircraft location systems.

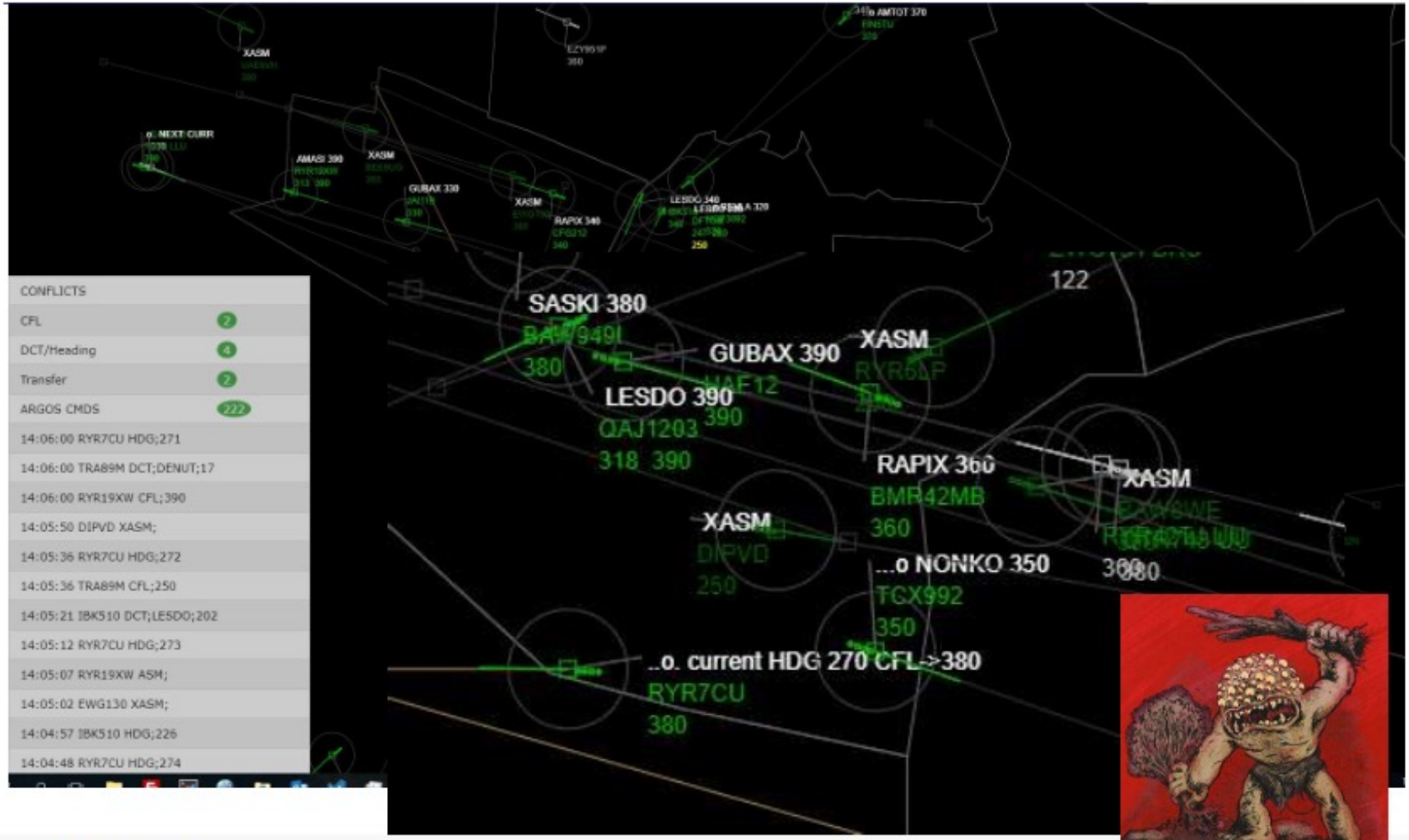


FOCUS OF THE CASE STUDY: ACAS X AND UNVALIDATED ADS-B POSITIONS

- The treatment of **unvalidated ADS-B positions** by ACAS X emerged as one of the controversial design issues with respect to liability.
- 'Unvalidated' refers to positions which are solely based on ADS-B data, not validated through other surveillance data sources.
- 4 design options debated by EUROCAE as in the diagram.



ARGOS V0.1



ARGOS modes of operations

L3

ARGOS AS A DECISION SUPPORT TOOL

For all flights, ARGOS displays the best plan. The ATCO can approve the plan, impose a constraint to let ARGOS revise the plan, or come up with his/her own plan. For CPDLC flights, ARGOS executes the plan. For non-CPDLC flights, the ATCO is reminded and the plan is the default selection in the menus.

L5

ARGOS MANAGES A SUBSET OF FLIGHTS

ARGOS manages certain flights (for each flight, a plan is presented and executed). The ATCO monitors and can take flights away from ARGOS. The ATCO controls all non-ARGOS flights.

L8

ARGOS MANAGES ALL FLIGHTS

ARGOS manages all flights (for each flight, a plan is presented and executed). The ATCO is alerted by ARGOS when monitoring is required: ARGOS still manages the situation but outside its normal comfort zone (i.e. conflict-free look-ahead time is reduced). The ATCO monitors as requested (i.e. stays in L8). The ATCO can take flights away from ARGOS (i.e. revert to L5).

SAFELAND

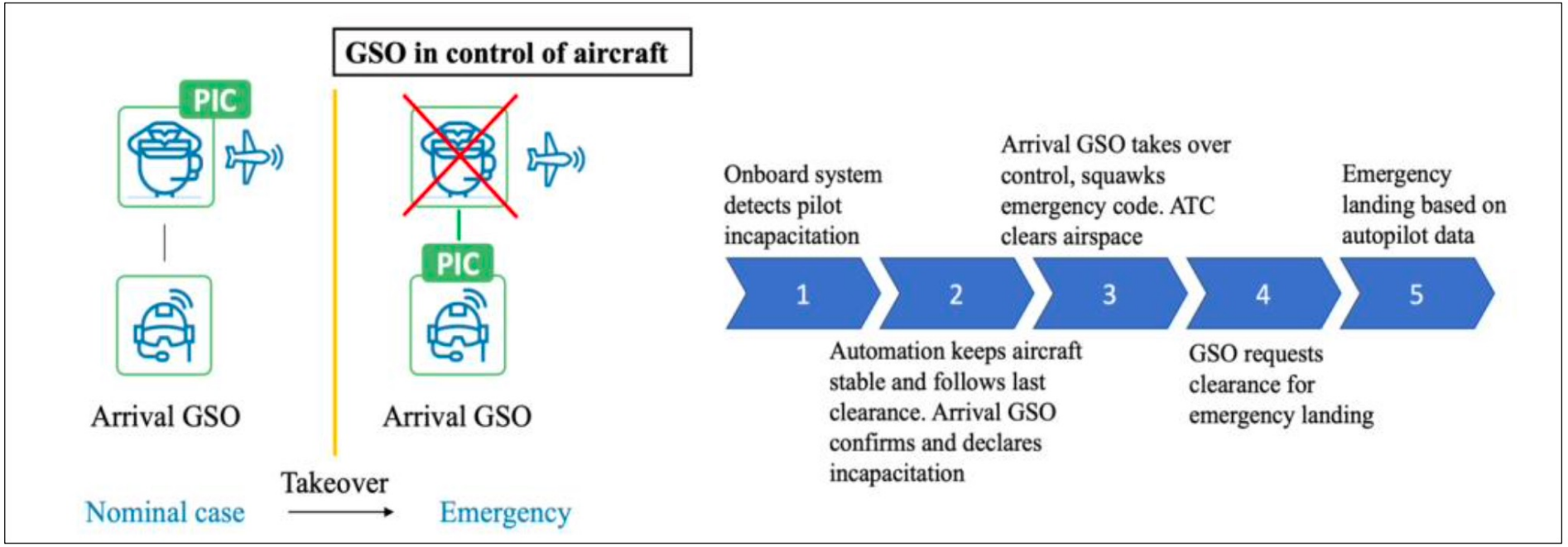
- Supporting flight and landing of aircraft operated by a single pilot, in case of partial or total incapacitation of the pilot.
- **Three implementation options:**

ATCO Focused: most of the single pilot tasks are assigned to the Air traffic controller

GSO Focused: most of the single pilot tasks are assigned to the Ground Station Operator

Automation Focused: most of the single pilot tasks are assigned to the cockpit automation

- Selected solution: GSO + Automation



Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

MAI4CAREU Master programmes in Artificial Intelligence 4 Careers in Europe