



# Explainability and Trust: Unravelling the Black Box of AI in ATM

Jose Manuel Cordero

# SESAR 2020 SHOWCASE

Would you fly in a fully automated aircraft?

Would you fly if the ATC service were provided by an AI?

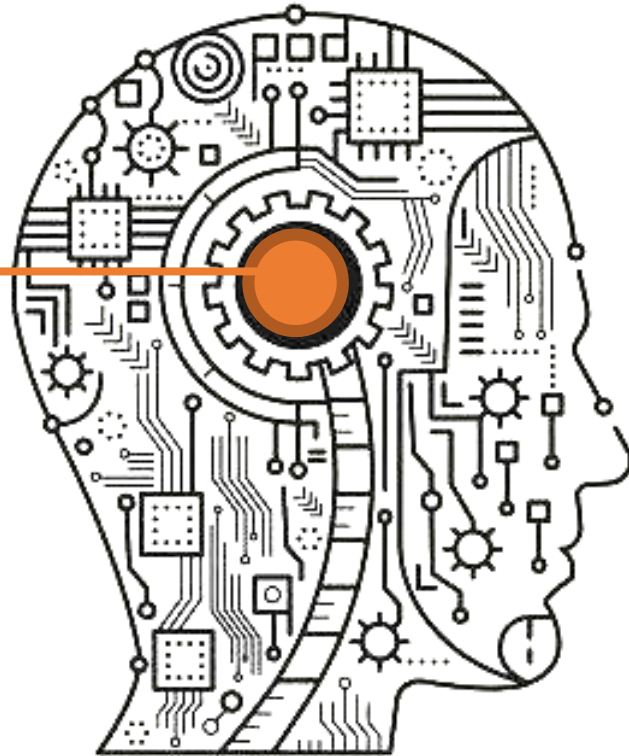
If you were an ATCO, would you accept decisions being taken by an AI?



# The challenges of AI in a Human-Centric ATM

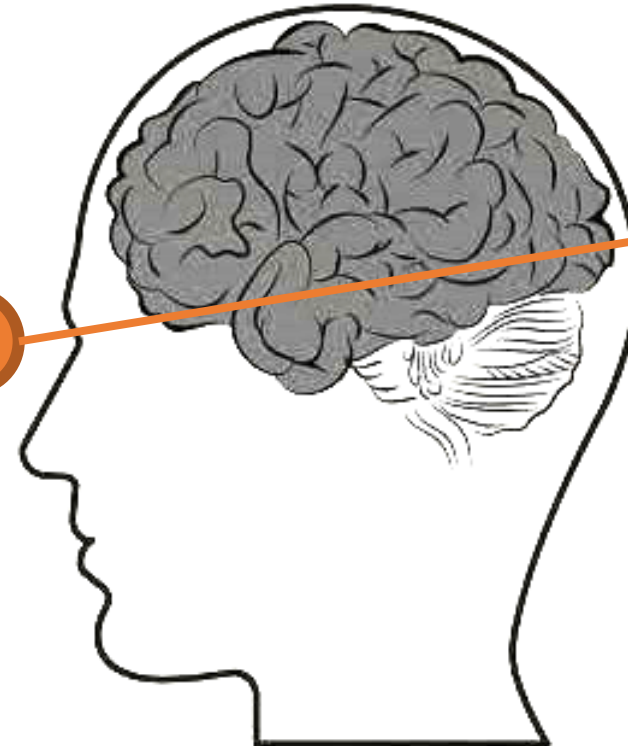
## Generalisation

*Up to which point a system trained with a dataset remains valid in a different context?*  
*Applicability*



Long, monotonous tasks  
Assembling information to reach decisions  
Finding patterns and trends  
Fast reaction

**Better together...  
if combined in a  
right way**



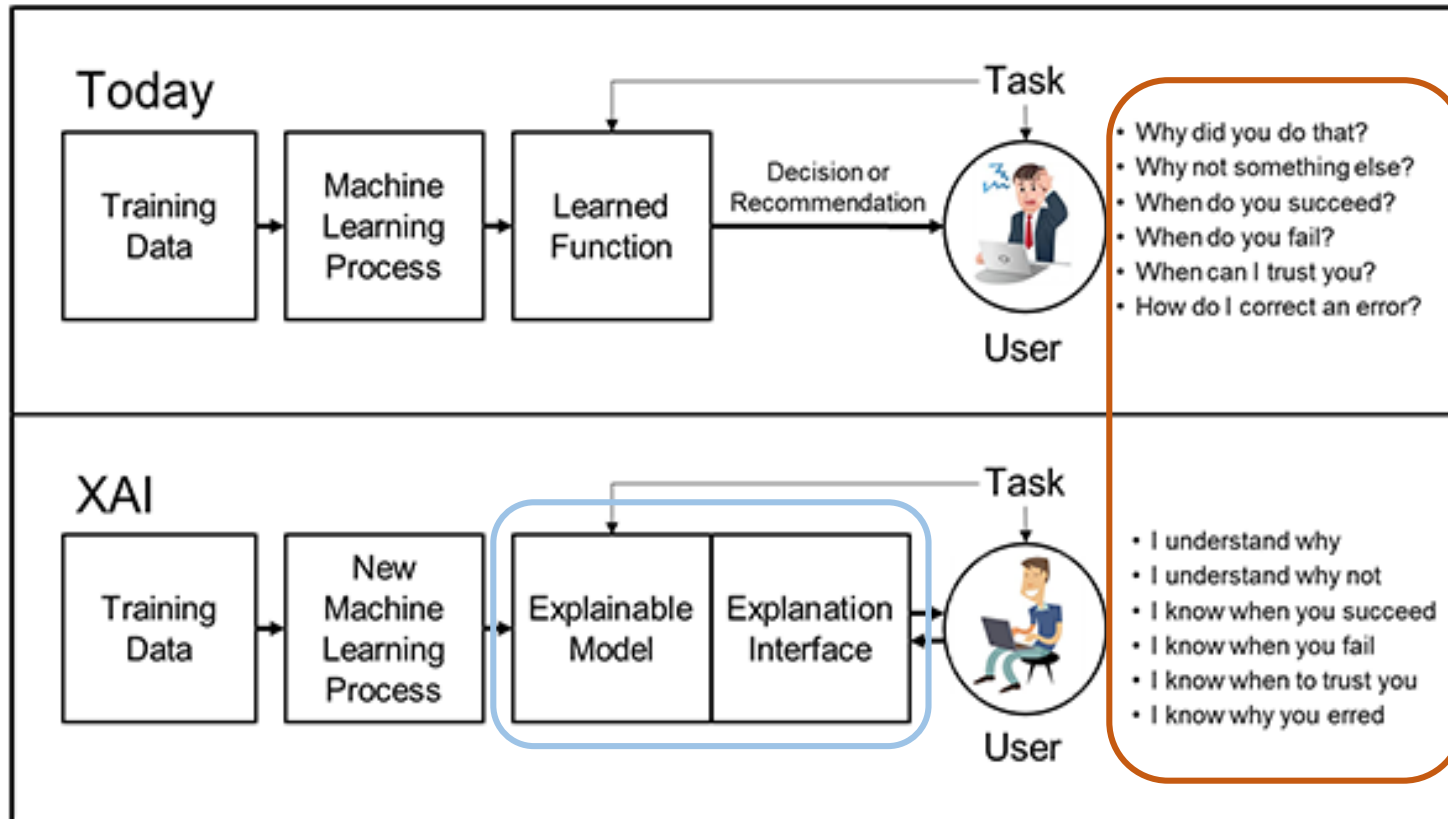
## Explainability

*How much the humans are able to understand the AI decisions?*  
*Trust*

Holistic perspective  
Thinking outside the box  
Creative solutions



# Managing Explainability



# Automation in ATM

Today  
(Baseline)

Research  
target

Definition			Definition of Level of Automation per Task				Automation Level Targets per MP phase (A, B-C, D)	
			Information Acquisition and Exchange	Information Analysis	Decision and Action Selection	Action Implementation	Autonomy	Air Traffic Control
Action can only be initiated by Human	Level 0 Low Automation	Automation supports the human operator in information acquisition and exchange and information analysis	Dark Blue	Medium Blue	Light Blue	Light Blue	Light Blue	Light Blue
	Level 1 Decision Support	Automation supports the human operator in information acquisition and exchange, information analysis and action selection for some tasks/functions	Dark Blue	Dark Blue	Medium Blue	Light Blue	Light Blue	Light Blue
Actions can be initiated by Automation	Level 2 Task Execution Support	Automation supports the human operator in information acquisition and exchange, information analysis, action selection and action implementation for some tasks/functions. Actions are always initiated by Human Operator. Adaptable/adaptive automation concepts support optimal socio-technical system performance.	Dark Blue	Dark Blue	Dark Blue	Light Blue	Light Blue	Light Blue
	Level 3 Conditional Automation	Automation supports the human operator in information acquisition and exchange, information analysis, action selection and action implementation for most tasks/functions. Automation can initiate actions for some tasks. Adaptable/adaptive automation concepts support optimal socio-technical system performance.	Dark Blue	Dark Blue	Dark Blue	Medium Blue	Medium Blue	Medium Blue
Action can be initiated by Automation	Level 4 High Automation	Automation supports the human operator in information acquisition and exchange, information analysis, action selection and action implementation for all tasks/functions. Automation can initiate action for most tasks. Adaptable/adaptive automation concepts support optimal socio-technical system performance.	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Dark Blue
	Level 5 Full Automation	Automation performs all tasks/functions in all conditions. There is no human operator.	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Dark Blue	Dark Blue

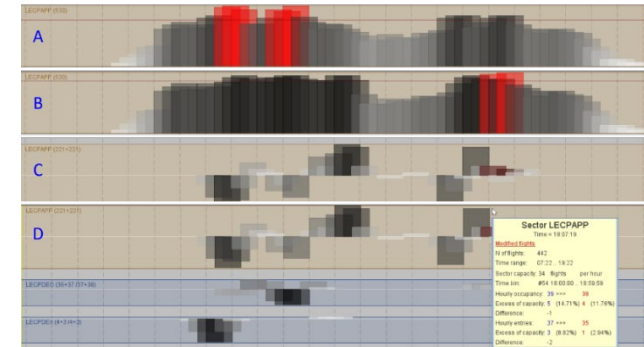
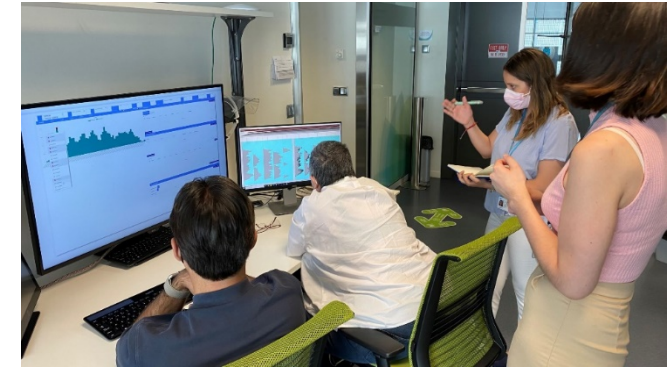
Automation level 3 includes automated action decision and implementation by AI system

Degree of automation support for each type of task




# The Research

- Two different use cases tested: ATFCM and CD&R
- XAI prototypes developed for Automation levels 2 and 3
- Integration with Visual Analytics tools to facilitate the provision of explanations on AI decisions
- Real Time Simulations in ATC platforms involving experts ATCOs
- Principles for Transparency in AI/ML applications in ATM



# Main conclusions

- Trust is more important for the operator than explanations
  - Some explanations may be helpful, but often decisions are self-explanatory or too complex for real time
- But trust needs to be built before... through explanations
  - Mainly during certification and training phase, where explanations on the behaviour of the system are required
  - And is very easy to lose it!
  - Explainability requirements should be considered since design phase
- Higher automation levels imply new challenges
  - Operation before human capabilities imply safe back-to-normal modes



# Check more results on the website



<https://tapas-atm.eu/>

