Master I livello
"Airline Management and Regulation
of the Aviation Industry"
a.a. 2022/2023

UNIVERSITA' DEGLI STUDI DI MESSINA
DIPARTIMENTO DI GIURISPRUDENZA

MASTER DI I LIVELLO IN
"AIRLINE MANAGEMENT AND REGULATION OF THE AVIATION INDUSTRY"
A.A. 2022/2023

From AI to Sky:
Navigating the Path to Compliance
with the Upcoming EU AI Act and EASA Legislation.

Tesi di

Matteo MARCOZZI

Relatori

Antonio LICU - ***EUROCONTROL***

Benjamin CRAMET - ***EUROCONTROL***

Anno Accademico 2022-2023

Table of Contents

1. Introduction

The advent of Artificial Intelligence (AI) marks a turning point in the annals of human civilization, redefining the contours of our social fabric, economic paradigms, and the very essence of human interaction. This profound influence of AI has catalyzed the European Union to adopt a vanguard role, setting the stage to become the world's first entity to introduce a holistic regulatory framework through the so-called 'AI Act'.

This legislative endeavor seeks to create a harmonized environment for the development, deployment, and application of AI technologies across its Member States, striking a delicate balance between encouraging innovation and protecting fundamental human rights and safety.

In the aviation sector, known for its strict safety standards and complex operational requirements, the emergence of AI signals a transformative shift. The sector stands on the precipice of evolution, with AI integration poised to significantly enhance operational efficiency, safety, and environmental sustainability. From the realms of predictive maintenance to the sophistication of advanced Air Traffic Management Systems, AI promises to redefine aviation.

However, this promising horizon comes with its own set of challenges, primarily navigating the uncharted future shaped by the AI Act and its impact on the regulatory framework that supports aviation.

As we teeter on the edge of this new epoch, it is imperative for the aviation community to proactively engage with the shifting regulatory landscape. The AI Act emerges as a pivotal moment, laying the groundwork for a future where AI and aviation ascend to unprecedented heights, steered by the principles of safety, efficiency, and ethical stewardship.

This intricate dance between innovation and regulation is set to chart the course of aviation for the foreseeable future, providing a model for the seamless integration of cutting-edge technologies into the sector's DNA, while ensuring that the skies remain a bastion of safety and that the industry's forward march is sustainable, adhering to global standards and expectations.

In the culmination of an enriching traineeship period at Eurocontrol, this work aims not only to delve into these transformative trends and in the Regulatory framework, but also to ground the discussion in the practicalities of AI's application within the aviation sector.

Specifically, it will shed light on two pioneering AI applications poised for deployment at the Eurocontrol Network Manager Operations Centre (NMOC).

Through this exploration, the thesis endeavors to offer insights into the practical implications of AI within the realm of aviation, serving as a testament to the sector's adaptability and forward-thinking approach in the face of emerging technological landscapes.

This narrative, rooted in the synthesis of innovation and regulation, aims to captivate and inform, guiding the reader through the intricate interplay of AI and aviation as we venture into an era marked by profound change and boundless potential.

2. Unveiling Artificial Intelligence: Definitions and Insights

2.1 The Dawn of AI: An Historical Exploration

The origins of artificial intelligence (AI) can be traced back to several key moments in the history of technology and computational thinking. One such moment concerns the '**Bombe Machine**' developed during World War II to decipher the 'Enigma' code used by the Germans. This device, conceived by Alan Turing and others, was essentially an electromechanical machine that automated the process of codebreaking, significantly speeding up the Allies' ability to read enemy communications. However, this machine was not an example of artificial intelligence in the modern sense; it was a calculation tool that performed specific tasks based on preset configurations without learning, adapting, or operating outside its initial programming parameters, *running the numbers,* instead of thinking.

The work of Alan Turing, especially his concept of the 'imitation game' (also known as the ***Turing Test***), proposed in his 1950 paper 'Computing Machinery and Intelligence' is unanimously considered a fundamental pillar in the history of the development of artificial intelligence. Turing suggested that a machine could be considered intelligent if it could mimic human behavior to the extent that a human interrogator could not distinguish whether the response came from the machine or a human being. This test laid the philosophical and conceptual groundwork for the field of artificial intelligence.

Transitioning from Alan Turing and his seminal Turing Test, let's introduce **John McCarthy**, a pivotal figure in the realm of artificial intelligence. Born in 1927, McCarthy was a mathematician and computer scientist who significantly contributed to the development of AI. McCarthy's academic career was marked by his tenure at institutions like Stanford University and MIT, where he made lasting contributions to the field of computer science.

In **1956**, **McCarthy coined the term 'artificial intelligence'**, defining it as 'the science and engineering of making intelligent machines', and played a key role in organizing the **Dartmouth Conference**. This event is widely recognized as **the birth of AI as a distinct field of study**. His vision for AI was comprehensive, focusing on creating machines capable of performing tasks that, when performed by humans, require intelligence. This includes the ability to improve autonomously, which laid the groundwork for the field's future research directions.

The Dartmouth Conference, under McCarthy's guidance, brought together researchers across disciplines to discuss and explore the potential of intelligent machines. This gathering set the stage for AI to emerge as an autonomous discipline, prioritizing the creation of systems that could learn and solve problems without direct human intervention.

John McCarthy is also renowned for creating Lisp, a programming language integral to robotics and scientific applications. Lisp's utility in early computing endeavors, such as chess with IBM machines, underscores its significance. McCarthy's innovations didn't stop there; he pioneered the concept of time-sharing, foreseeing a future where computing resources could be shared, with a super central computer that could accommodate multiple simultaneous connections, laying groundwork for the internet and cloud computing. At Stanford, his AI lab's explorations included early autonomous vehicles and the quest for machine-based common sense, highlighting his foresight in computing's evolution.

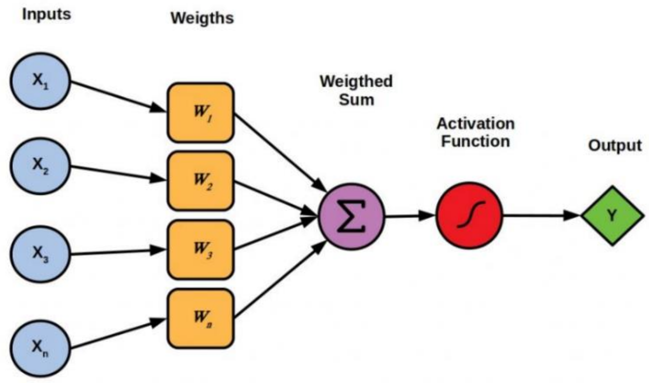
Comparing the contributions of Turing and McCarthy to AI, we see two foundational but distinct approaches. Turing, with his Turing Test, provided a conceptual framework, positing that a machine's intelligence could be gauged by its ability to imitate human behavior indistinguishably. McCarthy, on the other hand, offered the first formal definition of AI and established it as an independent field of study. While Turing's contributions laid the philosophical and experimental groundwork for AI, McCarthy's efforts were instrumental in defining the field's scope and objectives, emphasizing the scientific and engineering aspects of creating intelligent machines.

Despite the initial enthusiasm, progress in AI faced significant periods of stagnation and reduced investment, known as '***AI winter***', especially after the 1980s. These periods were characterized by inflated expectations followed by disappointments over the real capabilities of AI technologies of the time, leading to cuts in funding from both government and private sectors. Challenges included limitations in computing power, a lack of data, and insufficiently advanced algorithms.

The overcoming of these challenges truly began with the development and application of ***machine learning*** (ML) concepts, a subset of AI that allows computers to learn from data and improve their performance over time without being explicitly programmed for each task.

Actually, one must go back to the '50s again to find one of the early historical examples of machine learning, called ***Perceptron***, invented by Frank Rosenblatt in 1958. The Perceptron stands as a seminal breakthrough in the realm of artificial intelligence, heralding a novel paradigm for computational processes. This innovative mechanism, modeled on the biological neuron, unveiled the potential for machines not merely to execute predefined tasks but to learn from their interactions with data. Distinct from the rigid, rule-based algorithms that dominated the early computing era, the Perceptron introduced an element of adaptability and learning, adjusting its internal parameters (or weights) in response to the information it processed. This capacity to learn from experience and modify its operations accordingly was a radical departure from previous computational models. As such, the Perceptron laid the foundational stones for the evolution of neural networks and machine learning, cornerstones of contemporary artificial intelligence systems. Its inception is celebrated as a milestone in AI, spotlighting the shift towards machines that can learn, adapt, and potentially think, echoing the cognitive functions of the human mind.

The following figure shows this learning process schematically.



Performing simple binary classifications, which means it could decide whether an input belonged to one category, or another based on learned examples, it could differentiate between two different shapes or recognize simple patterns. One practical example of the Perceptron's use was in image recognition tasks, such as distinguishing between photos with cats versus those without. This early form of machine learning was capable of analyzing an array of pixels in an image, learning to recognize patterns that denote the presence of a cat. Though rudimentary by today's standards, these experiments were groundbreaking at the time, showcasing the potential of machines to learn from data and make decisions based on their learning.

Later, the progress in neural networks, combined with an increase in data availability and computing power, led to significant improvements in the field, ultimately paving the way for the development of modern AI systems capable of learning from vast amounts of data and improving over time. A notable example of AI evolution is the progression from **programmed** to **learning** systems in the domain of chess. Initially, the challenge to human chess champions came from powerful, but explicitly programmed, computers like ***IBM's Deep Blue***, which defeated the undefeated world champion Garry Kasparov in 1997. In terms of computational power, IBM's Deep Blue was not an example of artificial intelligence in the way we understand it today; rather, it represented a pinnacle of brute-force computing power and human-engineered chess strategy. Deep Blue's system could evaluate 200 million chess positions per second, a staggering figure that allowed it to explore numerous possible moves and outcomes deep into the future of a game. However, this capability stemmed from the system's ability to rapidly process a vast database of chess games and strategies that had been pre-loaded by human experts, rather than the machine learning or adaptive algorithms that characterize contemporary AI.

Deep Blue's achievement in defeating world champion Garry Kasparov in **1997** is nonetheless considered a milestone in the history of computing and artificial intelligence. This is because it demonstrated that a **machine could outperform a human in a complex intellectual task**, provided it was equipped with sufficient processing power and a comprehensive database of human knowledge. Deep Blue's victory marked the first time a computer had beaten a reigning world champion in a match under standard chess tournament conditions, challenging previous notions of human dominance in strategic thought and opening the door for future advancements in AI, particularly in the areas of machine learning and decision-making. Despite its reliance on human-preloaded strategies rather than genuine AI, Deep Blue's success laid the groundwork for the exploration of AI's potential in solving complex problems and performing tasks requiring deep analytical capabilities.

The more recent challenge came from a less powerful computer that was not programmed with specific chess knowledge but was capable of learning. DeepMind’s **AlphaGo** and its successor, **AlphaZero**, demonstrated this shift.

AlphaZero marks another significant milestone in the evolution of artificial intelligence by showcasing an advanced form of machine learning known as **reinforcement learning**. Unlike its predecessor, AlphaGo, which was trained on a large dataset of human played ‘Go’ games, AlphaZero learned chess, Go, and Shogi from scratch, without any prior knowledge except for the game rules.

AlphaZero managed to defeat *Stockfish*, one of the world's strongest chess engines that relies on extensive opening books and endgame tablebases, after just 24 hours of self-training.

Here again, two different approaches:

Stockfish operates on **brute-force calculation**, enhanced by heuristics, opening books, and endgame tablebases, meticulously crafted by human experts. It evaluates millions of positions per second, using a vast database of past games and predefined strategies to choose its moves.

AlphaZero, on the other hand, adopts a **machine learning approach** based on reinforcement learning from self-play, without relying on human game data. Starting with a blank slate, it plays games against itself, learning from each victory and defeat. This process enables AlphaZero to develop its own understanding of chess, discovering strategies and tactics independently.

This victory was not just a testament to its computational power but rather to its ability to learn and adapt strategies that are not pre-programmed by humans.

This paradigm of learning, starting from zero knowledge and achieving expertise through self-play and continuous self-improvement, demonstrates a shift towards AI systems capable of autonomous learning. Such systems do not just follow explicit programming but can discover complex strategies and adapt to new challenges on their own, showcasing the potential for AI to reach beyond human-taught knowledge and strategies.

Following the milestone achievement of AlphaZero by DeepMind in 2017, the journey of artificial intelligence (AI) has continued to evolve, marked by several key moments that have pushed the boundaries of what AI can achieve. These moments not only showcase the rapid advancements in AI technologies but also highlight the growing integration of AI into various sectors, from healthcare to creative arts, further blurring the lines between human and machine capabilities.

One of the key moments that made it possible for everyone to understand what AI is truly capable of doing is the advent of OpenAI's **ChatGPT**, an advanced chatbot designed to generate human-like responses based on the input it receives. By utilizing neural networks to develop a language system that closely mimics the natural language used by humans, known as **Natural Language Processing**, it has enabled an extensive segment of the population to truly understand what artificial intelligence can achieve.

The mere ability to converse with machine using natural language and receive immediate, rational and logical responses in virtually every field of human knowledge has astonished everyone, getting really closer to 'Turing's imitation game'. Even many of those who were studying the field did not anticipate such an outcome, both in terms of the product offered and the timing. This already prompts reflection on what can be achieved in increasingly shorter times in the future.

2.2 Defining Artificial Intelligence: A Comprehensive Overview and Technical Insights

A generally recognized definition by the academic world[1] of **Artificial Intelligence** is the following:

"*Artificial intelligence is the theory and development of computer systems able to perform tasks normally requiring human intelligence*.".

This definition of Artificial Intelligence (AI) covers the theoretical and practical sides of crafting intelligent systems. It mentions that AI's '*theory and development of computer systems'* points to the scholarly and technical endeavors in devising software and hardware for intelligent actions. This includes exploring intelligence's nature (the theory) and using this insight for practical applications (the development).

The goal of AI, as highlighted, is for machines to execute tasks previously thought to require human cognitive capabilities, like understanding languages, recognizing patterns, and learning from experiences. Essentially, AI strives to emulate human intelligence to a degree, enabling machines to perform activities relying on human intellect, such as reasoning and problem-solving.

This definition emphasizes AI as a discipline aimed at creating autonomous systems capable of functions that traditionally required human thought and intervention, merging theoretical intelligence insights with their practical use.

Having established a general definition of AI, it's crucial to delve into its subsets and techniques to gain a fuller understanding.

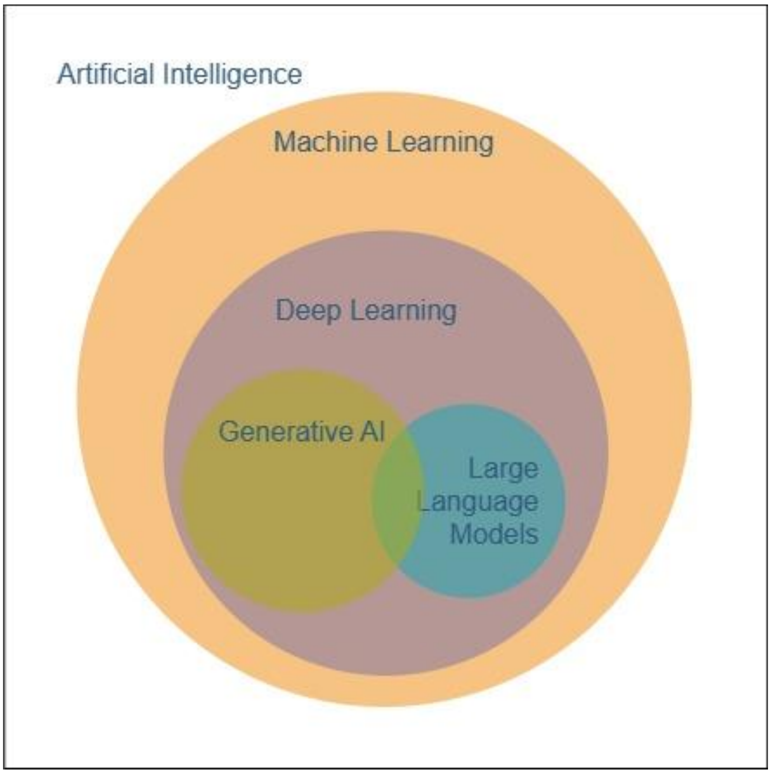
Machine Learning (ML) stands out as a primary subset, which further encompasses **Deep Learning** (DL) as a significant subcategory. ML can be divided into supervised learning, where

[1] Oxford Reference.

models are trained on labeled data, and unsupervised learning, which deals with raw, unlabeled data.

Deep Learning, leveraging **neural networks**, excels in processing more complex patterns, effectively utilizing both labeled and unlabeled data to enhance learning capabilities.

In the intricate landscape of Deep Learning, **Generative AI** stands as a specialized subset, distinct yet interconnected with **Large Language Models** (LLMs). Generative AI focuses on the creation of new data, closely mimicking the distribution of the training set it was exposed to. This encompasses a broad range of applications, from predicting the next word in a sentence to generating entirely new images or music that resemble human-made artworks. Its capability to generate novel outputs is fundamentally different from discriminative models, which are primarily concerned with categorizing input data based on learned relationships from labeled datasets.



Generative AI serves as the backbone for **foundation models**, which are highly versatile and capable of generating text, code, images, and audio. These models have fundamentally transformed how machines understand and create content, enabling a level of creativity and efficiency previously unattainable. For instance, generative models within the visual arts can produce artworks that are indistinguishable from those created by humans, while in the domain of text, they can write essays, poems, or code snippets that feel genuinely human-crafted.

Large Language Models (LLMs) have revolutionized natural language understanding and generation operating by digesting vast amounts of text data, learning complex patterns, and applying this knowledge to predict the most likely subsequent text based on the input they receive. This ability makes them incredibly powerful tools for tasks ranging from completing sentences in a conversation to generating articles on given topics. The depth and breadth of their training data allow LLMs to produce outputs that are not only coherent and contextually relevant but also surprisingly insightful and creative.

2.3 The OECD's Vision of AI: Defining the Boundaries

The OECD (Organisation for Economic Co-operation and Development) has also engaged in the field of Artificial Intelligence. In line with its goals of fostering economic growth, enhancing employment, improving living standards, maintaining financial stability, aiding economic development in other countries, and contributing to the growth of world trade, the OECD has taken a proactive stance on AI.

In 2019, a significant work was published - the '*2019 OECD Recommendation on Artificial Intelligence" (AI Principles)*. This set of guidelines is aimed at fostering responsible innovation and use of artificial intelligence, adopted by OECD member countries[2]. These principles encourage practices that uphold human rights, democracy, and ethical values in AI deployment, ensuring that AI development and use are inclusive and sustainable. They cover various aspects, including transparency, safety, privacy, accountability, and international collaboration, with the goal of guiding governments, the private sector, and civil society. The initiative promotes innovation that benefits all while minimizing the risks associated with AI, providing a foundation for further global discussions and regulatory developments regarding AI ethics and policy.

The **definition** contained in the work mentioned above, published in 2019, is as follows:

'*An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy*'.

Following[3], the AIGO[4] delegates' discussions with technical experts, and after the session, the Commission agreed to forward the modified version to the OECD Council for adoption on November 8, 2023.

[2] A total of 38 states, European states plus the United States and Canada.
[3] On 9 November 2022, 20 April 2023, 2 July 2023, 11 September 2023 and 6 October 2023.

In March 2024, the OECD published a specific document containing clarifications to the definitions of an AI system contained in the motioned document, in order to support their continued relevance and technical soundness, providing a comprehensive overview of the revised definition of an AI system.

Here **the latest definition** of AI system provided by OECD:

'*An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment*'.

The **OECD definition aims to be flexible** by reflecting a broad understanding of AI, and actors using this definition are encouraged to exercise judgement on its relevant scope, depending on the context it is being used in.

The aim of OECD about the actual application of the updated definition of the AI system, even wider, is clearly written in the document when it says that the update definition of AI is '***inclusive*** *and encompasses systems ranging from simple to complex'* and AI '*represents a set of technologies and techniques applicable to many different situations*'. Given that many specific techniques (such as machine learning) may require particular and dedicated considerations by policymakers in relation to, for example, biases and explainability, when applied in practice '*additional criteria may be needed to narrow or otherwise tailor the definition when used in a specific context'.*

[4] AIGO delegates are part of the OECD Working Party on Artificial Intelligence Governance (AIGO), which oversees the OECD's work on artificial intelligence (AI) policy.

2.4 Defining AI within the AI Act: A Close Look at the European Legislative Perspective

Regarding the European legislator's approach to artificial intelligence, it is important to note the process during the approval of the proposal made by the European Commission for the '*regulation of the European Parliament and of the Council on laying down harmonized rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts*', better known as the **AI Act**.

This process, not yet concluded at the date of this work, has been long and winding, as evidenced by the <u>changes to the definition of Artificial Intelligence</u> throughout the various stages of the legislative process.

The importance of a definition of AI provided by the legislator cannot be overstated, as it delineates the scope of the regulation itself, what falls within this definition is subject to regulation, while what does not, is exempt. A definition that is too precise and detailed runs the risk of excluding technologies not anticipated by the definition, particularly in a field as rapidly evolving as technology, whereas a definition that is too broad faces the opposite risk of regulating as AI what essentially may not be AI. The implications of these definitional boundaries and the balance between precision and inclusivity will be further explored in this chapter.

Refer to the accompanying communication sent on January 26, 2024, by the Presidency of the European Council to the members of the Permanent Commission, to outline the <u>main compromise</u> elements achieved during the <u>fifth</u> and last (troubled) <u>trilogue</u> held between December 6 and 8, 2023.

At Chapter II, point. 2, named '***Definition of an AI System***' it is stated that '*The definition of an AI system in **Article 3(1)** <u>has been modified</u> to align it more closely with the work of international organizations working on artificial intelligence, notably the **OECD**. Moreover, the* *corresponding* ***Recital 6*** [Ed. Note: Recital 12 of the adopted text] *details further the* ***key characteristics*** *of the definition and clarifies that the definition is* ***not*** *intended to cover* ***simpler traditional software*** *systems or* ***programming*** *approaches, which are based on the* ***rules defined solely by natural persons*** *to automatically execute operations* […]'.

When it is mentioned that the definition has been modified, refer to the table below, which displays in various columns the changes made from the Commission's proposal to the Parliament's mandate, to the Council's mandate, up to the latest draft.

	Commission Proposal	EP Mandate	Council Mandate	Draft Agreement
	Definitions	Definitions	Definitions	
Article 3, first paragraph				
127	For the purpose of this Regulation, the following definitions apply:	For the purpose of this Regulation, the following definitions apply:	For the purpose of this Regulation, the following definitions apply:	
Article 3, first paragraph, point (1)				
128	(1) 'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with;	(1) 'artificial intelligence system' (AI system) means ~~software~~***a machine-based system*** that is ~~developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined~~***designed to operate with varying levels of autonomy and that can, for explicit or implicit*** objectives, generate outputs such as ~~content,~~ predictions, recommendations, or decisions***, that influence physical or virtual environments*** ~~influencing the environments they interact with~~;	(1) 'artificial intelligence system' (AI system) means ~~software~~***a system*** that is ~~developed with one or more of the techniques and approaches listed in Annex I and can, for~~***designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve*** a given set of ~~human-defined objectives,~~ ~~generate~~***objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated*** outputs such as content ***(generative AI systems)***, predictions, recommendations~~,~~ or decisions***,*** influencing the environments ~~they interact with~~***with which the AI system interacts***;	(1) ~~'artificial intelligence~~***An AI*** system~~' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives~~ ***is a machine-based system designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives, infers, from the input it receives***, ***how to*** generate outputs such as ~~content, predictions~~***predictions, content***, recommendations, or decisions ~~influencing the~~***that can influence physical or virtual*** environments***. .*** ~~they interact with;~~

The **final version**, adopted by the European Parliament on March 13, 2024, is as follows:

'*AI system' means a machine-based system designed to operate with varying levels of autonomy, that may exhibit adaptiveness after deployment and that, for explicit or implicit objectives,* ***infers****, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments*.

It is noted that there is almost a perfect match with the last definition provided by the **OECD**, which is:

'An AI system is a machine-based system that, for explicit or implicit objectives, ***infers****, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment.'*

Moreover, the **Recital 12)** of the AI Act, elucidates, very clearly, what are the **characteristics** of the AI systems, emphasizing alignment with international standards for clarity and global acceptance. It highlights AI's distinct features from traditional software '*based on the rules defined by solely natural persons to automatically execute operations*', focusing on the capability for **inference**, which involves generating outputs like predictions or decisions based on data or encoded knowledge.

The **key points** that summarize the criteria defined by the legislator for the **AI system** are as follows:

- Alignment with international standards: This criterion ensures that the AI system's framework is clear and has global acceptance, facilitating a broader understanding and adoption across different jurisdictions.
- Distinct from traditional software: The AI system is characterized by its focus on inference capabilities, setting it apart from traditional software through its ability to draw new conclusions based on data analysis.
- Inference process: A key feature of AI systems is the inference process, which enables them to produce outputs that have tangible effects on their environments, demonstrating their ability to interact dynamically with the surrounding world.
- Techniques: The techniques employed by AI systems include both machine learning, where the system improves its performance on tasks through exposure to data, and logic/knowledge-based approaches that rely on explicit rules and knowledge representation.

- 'Machine-based': AI systems operate on machines, highlighting the technological aspect of AI where physical hardware is a critical component of AI functionality.
- Objectives: AI systems can have objectives that are either explicit, clearly defined and designed for a specific purpose, or implicit, where the outcomes may diverge from the system's original intended purpose. ('*may be different from the intended purpose of the Ai system*')
- Environments: This refers to the various contexts and operational domains where AI systems are deployed, underlining the importance of considering the system's interaction with its operational environment.
- Outputs: AI systems generate outputs such as predictions, content, recommendations, or decisions, showcasing the broad range of actions and influences AI can have in different sectors.
- Autonomy levels: AI systems exhibit some degree of autonomy, characterized by their ability to operate with minimal or no human intervention, highlighting the systems' capability to make independent decisions or actions. ('*they have some degree of independence of actions from human involvement and of capabilities to operate without human intervention*').
- Adaptiveness: This trait refers to the AI system's ability to learn from its environment and experiences, and to adapt its operations post-deployment, ensuring continuous improvement and relevance. ('*refers to self-learning capabilities, allowing the system to change while in use*').
- Integration: AI systems can either be integrated into products, enhancing their functionality, or exist as stand-alone solutions, demonstrating the versatility and wide applicability of AI technologies.

Thus, it is well understood that the European legislator's intention with the AI Act is to carefully define an AI system in a way that distinguishes it from traditional software systems, by highlighting its unique ability to infer from data, affect both physical and virtual environments, and evolve through learning. This definition strives for legal clarity and international alignment, all while preserving the adaptability needed to keep pace with fast-moving technological advances. By focusing on AI's capability to generate models, content, recommendations, or decisions through sophisticated machine learning and logic-based methods, and excluding systems based solely on human-defined rules, the legislator aims to ensure that the regulation is both precise and broad enough to remain effective and relevant amidst the ever-changing landscape of AI technology.

2.4.1 Analyzing changes to the AI Definition throughout the legislative process

As observed, this broader, risk-based definition was arrived at after several stages and modifications. By examining the definition initially proposed in the first draft by the EC, we can understand the contradictions it contained, and the potential risks had it not been revised.

As can be seen, the original definition consisted of two parts: the first being functional in nature and the second more prescriptive.: '*artificial intelligence system' (AI system) means software that is developed with* ***one or more of the techniques and approaches listed in Annex I*** [prescriptive part] *and can,* ***for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with*** [generic part]'.

In the '*explanatory memorandum'* accompanying the European Commission's original proposal, we see a concerted effort to set a harmonized framework for AI systems within the Union. The proposal emphasizes '*harmonized rules for the development, placement on the market and use of AI systems in the Union following a proportionate risk-based approach*.' It ambitiously seeks a '*single* ***future-proof definition of AI,***' reflecting a balance between innovation and regulatory oversight.

Feedback from **stakeholders** highlighted a desire for a '*narrow, clear and precise definition for AI*.' They stressed not just the importance of clarifying what constitutes AI, but also the need to define related concepts such as 'risk', 'high-risk', 'low-risk', 'remote biometric identification', and 'harm'. This request for a "specific" definition underscores the community's call for clarity and precision to navigate the complex landscape of AI technologies and their applications.

Title I of the memorandum outlines the regulation's subject matter and its scope, touching on the marketing, deployment, and use of AI systems. It strives for a definition of AI that is '*technology neutral and future proof,*' mindful of rapid advancements in AI technology and market dynamics. The pursuit of legal certainty is matched with a detailed enumeration in Annex I[5], listing the approaches and techniques for AI development, adjustable by the Commission to reflect new technological trends. It delineates key stakeholders in the AI value chain, such as providers and users, covering both public and private entities to ensure fairness. This approach embodies a vision that is 'technologically neutral – future proof,' yet detailed for legal certainty, balancing the need for a comprehensive framework with the flexibility to adapt to future innovations.

Also, the Recital 6 further elaborates on this vision, advocating for a definition of AI that ensures '*legal certainty, while providing the flexibility to accommodate future technological developments*.' It proposes a definition grounded in the '*key functional characteristics of the software*,' especially its capability to produce outcomes like content, predictions, recommendations, or decisions impacting the system's interaction environment, whether physical or digital. The Recital acknowledges AI systems' potential for varying autonomy levels and their application either as standalone solutions or product components, highlighting the necessity of a supportive list of specific development techniques and approaches that remains abreast of technological and market progress. This duality aims to capture '*legal certainty and flexibility*,'

[5] List has been removed in the final version of the AI Act.

ensuring the framework remains relevant and adaptive to evolving AI technologies and their societal impacts.

Through these sections, the memorandum delineates a path forward that seeks to balance the intricacies of defining AI with the pragmatic needs of regulation, aiming for a policy that is both inclusive and precise, ensuring safety and innovation go hand in hand.

It thus becomes clear that the European Commission's legislator aimed to concurrently pursue the two mentioned objectives: legal certainty on one side and technological flexibility on the other. However, it's essential to assess whether this approach was suitable for achieving these goals.

The decision to specify a detailed list of techniques[6] and approaches[7] in the prescriptive part of the definition, defining the scope of systems included in Annex I, is, in the author's view, highly debatable. This approach seems to starkly contrast with the legislator's declared intention to adopt a *'risk-based'* and technology-neutral model (*'future-proof'*). Moreover, the chosen method of identifying the three mentioned macro-areas runs the risk of excluding technologies that pose significant risks but are not listed.

For instance, consider two automation systems that have sparked significant legal controversies: Deliveroo's Frank system and the Parcoursup system. The Frank system[8] optimizes deliveries using algorithms to manage order assignments to couriers, while Parcoursup aids in university enrollments in France by matching students and courses through specific algorithms[9].

Both rely on automated systems that, despite being advanced, would not fall under the category of artificial intelligence according to the 'typified' definition used by the legislator.

[6] Machine learning.
[7] Logic and knowledge-based approaches and statistical models.
[8] The ruling of the Tribunal of Bologna, Labor Section, on December 31, 2020, highlighted how Deliveroo's "Frank" algorithm was discriminatory towards riders, failing to distinguish between justified and unjustified absences and negatively affecting their reliability score. This case underscores the risk of biases and errors in algorithms and the importance of a data protection impact assessment to prevent privacy and discrimination issues, highlighting the need to involve experts and worker representatives in the design and evaluation phase of digital tools.
[9] 'Il Regolamento Europeo sull'intelligenza artificiale – Analisi informatico-giuridica', Giuseppe Contissa, Federico Galli, Francesco Godano, Galileo Sartor. Dicembre 2021.

Therefore, the adoption of selective criteria in the definition of artificial intelligence carries the risk of excluding from the AI Act's application highly risky systems, such as the ones mentioned, while at the same time it could include systems that, despite utilizing the described techniques, do not necessarily present significant risks.

To pursue the repeatedly stated goal of adopting a risk-based and technology-neutral approach, subsequent amendments made during the legislative process have addressed this issue, as previously clarified.

2.5 Aviation domain

2.5.1 Artificial Intelligence as Defined by ICAO

To conclude our analysis on the definition of AI (or AI system), we now delve into the world of aviation, tracing the various phases in which the key players have approached AI, aiming to outline its boundaries with a definition.

Starting from the working paper made by ICAO in 2019, titled "*Artificial Intelligence and digitalization in aviation*", the document combines insights on AI, noting its lack of a universal definition but recognizing it as a blend of machine computational and cognitive skills, enabling high-level task execution. It spans various AI technologies, like machine learning and neural networks, emphasizing their potential to outperform human task efficiency.

This foundational understanding serves as a backdrop for discussing digitalization and AI's role in enhancing aviation through safety and efficiency. The transition to operationalizing these technologies demands collaboration among ICAO, national governments, and the industry to establish frameworks for training, certification, and data sharing, ensuring the advancement of aviation on multiple fronts.

The cited paper also provides a sort of classification of AI, through with '*AI should be categorized into the number of stages depending on the application and level of autonomy*'.

After the first stage for certification and qualification procedures based on AI, according to different research agencies, there are four stages, or 'waves', of AI development:

1. First wave: Rule-based systems adhering to human-defined rules.
2. Second wave: Systems gain intelligence through statistical methods.
3. Third wave: Contextual adaptation by AI.
4. Fourth wave: Fully autonomous AI integrating data across systems for effective environment sensing and response (e.g., UAV swarms or ATC data exchange).

In this case too, as one might expect, the definition is very broad and generic. It refers to "technologies" in a general way, using terms like "machines" and "cognitive power," and concludes by focusing on the outcome of the process, namely the ability to perform tasks possibly even better than humans.

It will be seen later that as one moves to a more detailed level in terms of legislative production and guidance material, the definitions will take on a different level of detail.

2.5.2 Clarifying Artificial Intelligence: The EASA's Approach to Defining AI

EASA too has defined AI, but with a more pragmatic approach than what ICAO did. As is well known, and as will be detailed in a dedicated chapter, EASA operates as a legislative drafter in the fields of civil aviation and environmental compatibility, also making autonomous decisions regarding its assigned aspects (like as: Certification Specifications; Acceptable Means of Compliance; Guidance Material; etc.).

Therefore, it too has had to confront the advent of artificial intelligence on multiple occasions and at various stages.

Since October 2018, the Agency had set up an internal task force on AI, whose work resulted in the publication of the '*EASA AI Roadmap 1.0'* in February 2020, that shows the EASA vision for addressing the challenges and opportunities of AI in aviation. This initial plan generated two major deliverables:

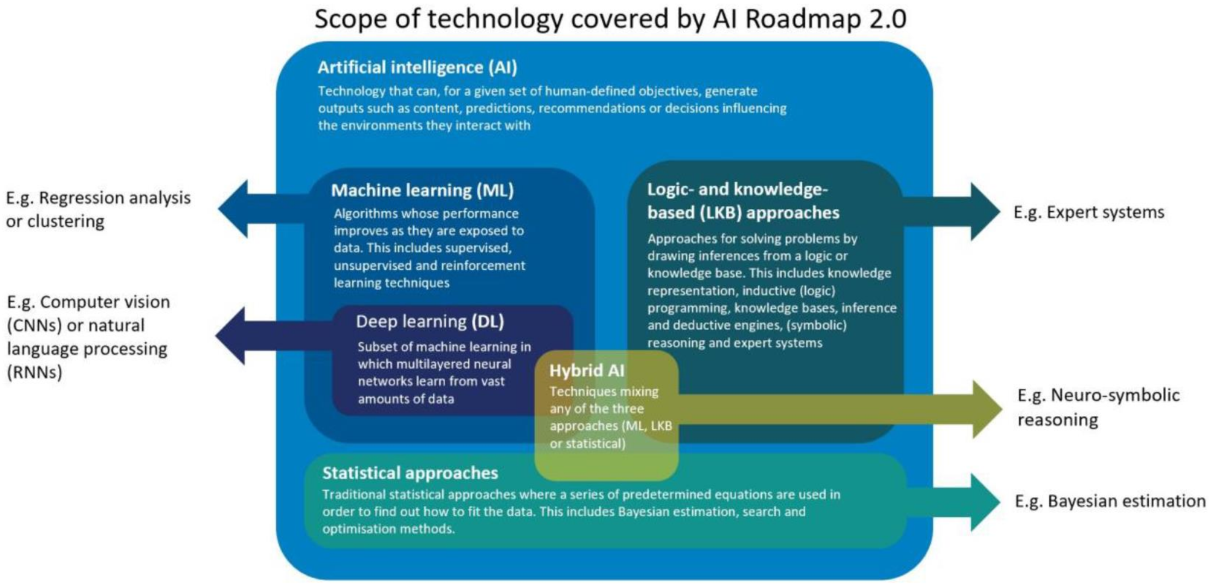
- The '*First usable guidance for Level 1 ML applications*', in December 2021.
- The '*EASA Concept Paper: First usable guidance for Level 1 & 2 ML applications'*, public consultation in February 2023.

The analysis of the EASA strategy on AI and the related regulatory projections will be discussed in the related Chapter of this work, but, for now, let's focus on the definition of AI provided by EASA in those documents.

From the updated version of the EASA Roadmap, the 2.0, in which the activity on AI evolved to a cross-domain program, AI is '*a broad term and its definition has evolved as technology developed*. […] *EASA has moved to the even wider-spectrum definition from the […] EU Artificial Intelligence Act, which is: "Technology that can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with"*'.

As can be seen, this definition mostly coincides with the one initially proposed by the European Commission in its proposal for the AI Act.

From such a broad definition, when drafting guide material that serves as a guideline for a sort of anticipated compliance with a regulation that, in fact, does not yet exist, but of which only the regulatory drafts are known, EASA proceeds by detailing more specifically what the techniques used in the field of artificial intelligence are, as can be seen from the image shown below, even though it is clarified that at the current state the **guidance material applies only to machine learning (ML) and its subset deep learning (DL)**:



The evolution from the 2019 ICAO document to the 2024 AI Act and the subsequent OECD update reflects the rapidly advancing field of AI and the need for definitions that encompass its growing complexity and capabilities.

The progression shows an increasing recognition of AI systems' ability to learn and adapt, moving away from solely rule-based systems to more autonomous entities that <u>can infer</u> and act upon both explicit and implicit objectives. This shift acknowledges the nuanced and dynamic nature of AI technologies, ensuring regulatory frameworks remain relevant and effective in promoting safe and beneficial AI development and use.

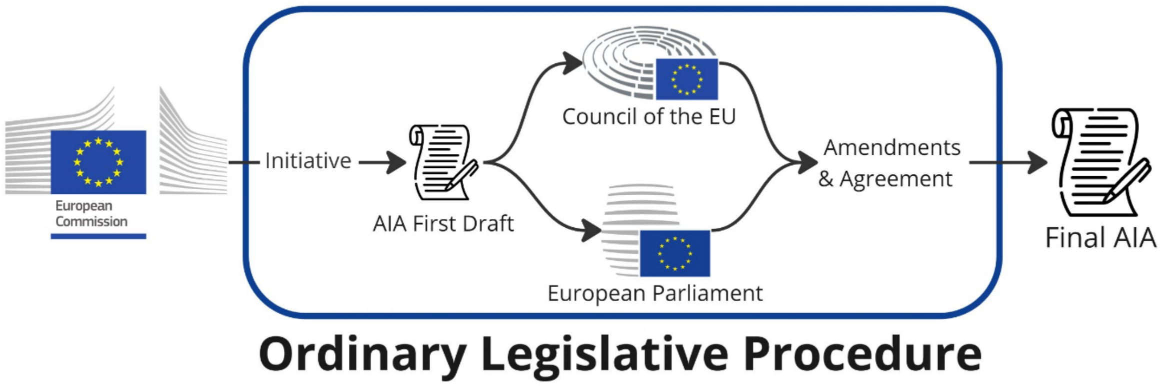
The reason for these evolving definitions is <u>twofold</u>: firstly, the technological advancements in AI necessitate a broader and more nuanced understanding of what AI systems are capable of, including learning, adapting, and making decisions with varying levels of autonomy; secondly, as AI becomes more integrated into various sectors, including aviation, there is a need for regulatory frameworks that accurately reflect the capabilities and potential impacts of AI systems. These evolving definitions help to ensure that AI is developed and used responsibly, with considerations for safety, ethics, and societal impacts at the forefront.

3. AI and the Law: Deciphering How Artificial Intelligence is Regulated

3.1 Understanding EU Rulemaking Processes

It might be useful do briefly illustrate the multifaceted nature of law-making within the European Union (EU) and its specific application to the aviation sector. The governance framework of the EU is characterized by the synergistic interaction of the **'Institutional Triangle'**, comprising the Parliament, the Council, and the Commission. This tripartite structure is pivotal in the creation of laws, with the Treaties providing the foundational rules for decision-making across the EU's broad spectrum of activities. Within this legal ecosystem, EU laws manifest in three principal categories: Regulations, Directives, and Decisions, each with its unique scope and impact.

Below an explanatory image of the ordinary process[10] (of the AI Act):



Within this architecture of the legislative process, several key actors play pivotal roles, each contributing uniquely to the fabric of EU governance. The **European Commission**, comprising 27 Commissioners nominated by member states and ratified by the European Parliament, serves as the executive arm of the EU. This body is tasked with initiating legislation, implementing

[10] Source: Future of Life institute - https://artificialintelligenceact.eu/context/

decisions, and upholding EU treaties, thus acting as a principal driver of policy formulation and execution within the Union.

Parallel to the Commission's executive function, the **European Parliament** stands as the voice of the EU's citizenry, with Members of the European Parliament (MEPs) directly elected to represent the populace. The distribution of seats within the Parliament is proportionate to the population of each member state, ensuring a balanced representation of the diverse demographic landscape of the EU. Within the Parliament, specialized committees take the lead on matters that are part of their area of expertise. Notably, in the context of the Artificial Intelligence Act (AI Act), the committees on Civil Liberties, Justice and Home Affairs (LIBE) and Internal Market and Consumer Protection (IMCO) are at the forefront. Each committee designates a "rapporteur", a MEP tasked with leading the committee's discussions on a legislative proposal.

The **Council of the European Union**, or "The Council," introduces a distinct dimension to the legislative dialogue, convening government ministers from each member state to deliberate on issues at hand. The AI Act, for instance, falls under the purview of telecommunications ministers, who gather within the 'Working Party on Telecommunications and Information Society (WP TELECOM)'. This configuration allows for a detailed consideration of legislation, reflecting the multifaceted interests and priorities of the EU's member states. The presidency of the Council, rotating biannually among member states, plays a strategic role in steering the Council's agenda and embodies the Council's interactions with other EU institutions.

This dynamic interplay among the European Commission, Parliament, and Council reflects the collaborative and complex nature of law-making in the EU, underscoring the importance of each actor in shaping legislation.

At the heart of this meticulously structured legislative process is the European Commission, which holds the exclusive '**right of initiative**'. This prerogative allows the Commission to craft the initial draft of any legislative proposal, setting the stage for the legislative journey.

Upon the Commission's presentation of a draft, the legislative task is passed to the Parliament and the Council. What follows is a dynamic exchange of the draft legislation, a process marked by the addition of amendments and adjustments. This iterative negotiation aims to forge a consensus between the two legislative bodies. The mechanism of '**trilogues**' (informal dialogues involving the Council, Parliament, and the Commission) plays a critical role during this phase. These discussions, although informal, have become the crucible for legislative agreement, often facilitating consensus in the initial exchange.

This collaborative process continues until either an agreement is reached, allowing the legislation to be passed, or an impasse is encountered, indicated by the failure to achieve consensus along the entire process. In the latter case, the process is concluded without the adoption of legislation.

Following the achievement of mutual consent between Parliament and the Council, the final step in the legislative journey is the formal approval and publication of the text in the Official Journal of the European Union (OJEU). This publication acts as the official birth certificate of the legislation, specifying the timeline for its entry into force.

3.2 Navigating the AI Act: A Legislative Odyssey

Having provided this overview of the ordinary legislative procedure in the EU, let's now move on to illustrate the tumultuous journey that the AI Act has undergone up to today.

From the European Commission's proposal in April 2021 to its adoption by the Parliament in March 2024, nearly 3 years have passed, filled with ongoing discussions, amendments, disputes, trilogues, etc.

Below is the timeline:

2021

- On 21 April, the **Commission** unveiled its **proposal** for regulating artificial intelligence across the EU.
- The Slovenian Council Presidency hosted an online conference discussing AI regulation, ethics, and fundamental rights on 20 July.
- The **European Commission concluded its public consultation** on the AI Act on 6 August, receiving feedback through 304 submissions.
- A **study** commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs on the ethical and legal aspects of biometric techniques was released on 6 August.
- A **preliminary compromise text on the AI Act draft**, highlighting significant modifications in social scoring, biometric recognition, and high-risk applications, was shared by the EU Council's rotating presidency on 29 November.
- The **AI Act negotiations** were to be co-led by the European Parliament's internal market and civil liberties committees, with Brando Benifei and Dragoş Tudorache taking the lead, as decided on 1 December.

2022

- The AI Act's initial joint discussion was held by the **European Parliament's Internal Market and Civil Liberties committees** on 25 January.

- A **new Standardization Strategy** was presented by the **European Commission** on 2 February, detailing their approach to Single Market and global standards.
- On 3 February, the **French Council Presidency** circulated a compromise proposal for Articles 16-29 of the AI Act, addressing high-risk systems' user and provider obligations, followed by another text for Articles 40-52 on harmonized standards and transparency obligations.
- Amendments on the AI Act were published by the European Parliament's Committee on Legal Affairs (JURI) on 2 March, with the Committee on Industry, Research, and Energy (ITRE) releasing their draft opinion the following day.
- Brando Benifei and Dragoș Tudorache, MEPs spearheading the AI Act in the IMCO and LIBE committees, disclosed their draft report on 20 April.
- **A proposal** to regulate **general-purpose AI systems** was published by the **French Council Presidency** on 13 May.
- Political groups in the European Parliament were given until 1 June to propose amendments to the AI Act, resulting in thousands of submissions.
- The **final compromise text** before the Czech Republic assumed the Council presidency was circulated by the **French Presidency** on 15 June.
- The Czech Council Presidency shared a discussion paper outlining the AI Act's main priorities with EU governments on 17 June.
- The Committee on Legal Affairs (JURI) of the European Parliament endorsed their viewpoint on the AI Act on 5 September, marking the last committee to do so.
- Proposals for a targeted harmonization of national AI liability rules were made by the European Commission on 28 September, aiming to support the AI Act.
- The **EU Council adopted its unified stance** ('general approach') on the AI Act on 6 December.

2023

- The **European Parliament set its negotiation stance** on the AI Act with 499 votes in favor, 28 against, and 93 abstentions on 14 June.
- A **provisional consensus** on the AI Act was reached by the **Parliament** and the **Council** on 9 December.

2024

- **Unanimous endorsement** of the AI Act by the EU's **27 member states** was confirmed, solidifying the political accord achieved in December.
- On 13 February, the **AI Act was approved by the Internal Market and Civil Liberties Committees** with a 71-8 vote (7 abstentions) following negotiations with the member states.
- The establishment of the **European Artificial Intelligence Office within the Commission**, under the Directorate-General for Communication Networks, Content, and Technology, was announced on 21 February, aiming to facilitate the AI Act's implementation, particularly concerning general-purpose AI.
- **Lawmakers in the European Parliament approved the AI Act** on March 13th, rules aimed at regulating AI according to a risk-based approach with an overwhelming majority. The law passed with 523 votes in favour, 46 against and 49 abstentions.

As I write, we are waiting for the last formalities, including translations and corrections, before proceeding to the publication in the Official Journal of the European Union.

There will then be a differentiated application depending on the subject concerned, as will be illustrated later.

3.3 The AI Act Decoded: Key Principles Shaping Artificial Intelligence Regulation

The European AI Act represents a groundbreaking regulatory approach designed to manage the use and development of artificial intelligence throughout the European Union. Its main objective is to create a safe AI environment that **respects legal standards and fundamental rights** across the EU.

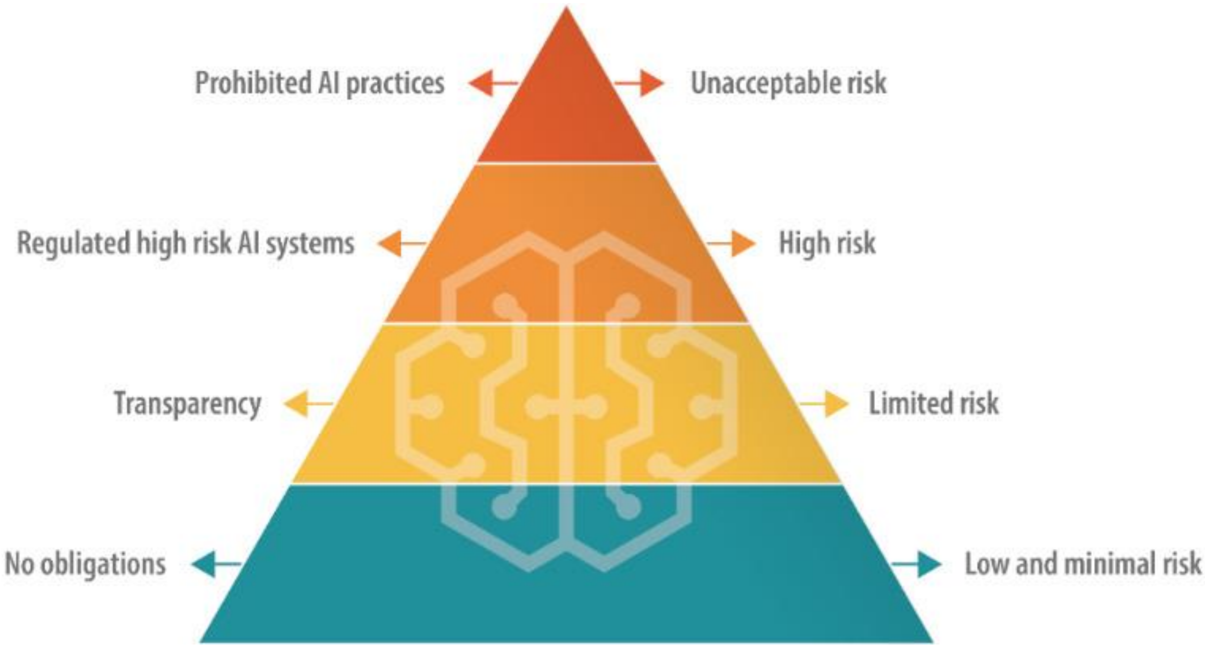
By categorizing AI technologies into four risk levels, from minimal to completely unacceptable, it outlines specific regulations for each. For instance, AI practices deemed to pose an unacceptable risk, including social scoring and certain biometric identification methods, because seen as posing an unacceptable risk to European values, are outright banned. Meanwhile, high-risk AI technologies are subject to strict regulations, whereas those considered to pose minimal risk are less regulated.

At its core, the Act aims to protect essential rights, enhance AI's safety and transparency, and stimulate innovation and competitive spirit within the EU's tech market. It sets clear responsibilities for both AI providers and users, focusing on responsibility, data management, and the necessity for human oversight.

This risk-based approach produces the following **classification** of the AI-based system:

- **Unacceptable risk** is prohibited (e.g. social scoring systems and manipulative AI).
- **High-risk** is permitted but subject to compliance with AI requirements and ex-ante conformity assessment.
- **Limited risk** is permitted but subject to information – transparency obligations (e.g. chatbots and deepfakes. Developers and deployers must ensure that end-users are aware that they are interacting with AI).
- **Minimal or no risk** is permitted and with no restrictions (e.g. spam filters).

* **General Purpose AI model** is permitted but with specific obligations (e.g. providers must provide technical documentation, instructions for use, comply with the Copyright Directive, and publish a summary about the content used for training, etc.).



Source: European Parliamentary Research Service-

As we will explore in more detail later in this document, **high-risk** systems are identified by a criterion designed to pinpoint on one hand AI systems that serve as safety components within regulated products (such as medical devices and machinery), which undergo assessment by third parties under the relevant sector-specific legislation. On the other hand, it also identifies certain stand-alone systems used in specifically **designated fields**, identified as high-risk due to the potential impact of AI systems on citizens' safety, rights, and freedoms.

These sectors (high risk) include:

* **Biometric Identification and Categorization**: The use of AI for real-time biometric identification and categorization in publicly accessible spaces.

- **Management and Operation of Critical Infrastructure**: AI applications involved in the control of critical infrastructure (also critical digital infrastructure) that could put the life and health of citizens at risk in case of failure, such as in transportation and supply of water and energy.
- **Education and Vocational Training**: AI systems intended to be used for determining access to education and assessing students in educational and vocational training institutions.
- **Employment, Workers Management, and Access to Self-Employment**: AI systems used for recruitment processes, making decisions on promotion and termination, and managing workers, including monitoring and tracking.
- **Access to and Enjoyment of Essential Private Services and Public Services and Benefits**: AI systems that are used to evaluate eligibility for public assistance programs and access to essential private services like banking and insurance.
- **Law Enforcement**: AI systems used by law enforcement agencies that could affect individuals' fundamental rights, including for predicting policing, profiling, and surveillance activities.
- **Migration, Asylum, and Border Control Management**: AI applications used in the management of migration, asylum requests, and border control operations.
- **Administration of Justice and Democratic Processes**: AI systems used in the administration of justice and democratic processes, including the application of law and legal decision-making.

Continuing the discussion on risk-based classification, it's also important to understand the responsibilities outlined under these regulations, especially for high-risk AI systems. The majority of these obligations are shouldered by the providers, who are essentially the developers of these systems. This applies to those who plan to market or implement high-risk AI systems within the EU, regardless of whether they are EU-based or located in a third country. Additionally, providers from third countries must comply if the outputs of their high-risk AI systems are utilized within the EU.

On the other hand, the users of these high-risk AI systems also bear certain responsibilities, albeit to a lesser extent than the providers. These users are typically natural or legal persons who deploy an AI system in a professional capacity. It's crucial to note that these obligations extend not just to users within the EU, but also to those in third countries, provided the AI system's output is employed within the EU.

Turning our attention to General Purpose AI (GPAI), the regulations stipulate that all providers of GPAI models must furnish technical documentation and instructions for use. They must also adhere to the Copyright Directive and publish a summary detailing the content used for training the systems. Providers who offer GPAI models under a free and open license are required to comply only with copyright laws and publish a summary of the training data, unless their models pose a systemic risk.

Furthermore, providers of any GPAI models that present a systemic risk, whether open or proprietary, must undertake model evaluations and adversarial testing. They are also obligated to track and report any serious incidents and ensure robust cybersecurity protections.

This comprehensive framework aims to ensure that the deployment of AI systems, especially those classified as high-risk or GPAI, is managed with a high degree of responsibility and transparency.

The AI Act also imposes significant penalties on businesses that breach its established regulations. Specifically, violations involving prohibited practices or non-compliance with data requirements could result in fines of up to € 35 million or 7% of the company's total global annual turnover for the previous financial year, whichever is higher. For other breaches, including failure to meet various regulatory requirements or violations concerning general-purpose AI models, penalties can reach up to € 15 million or 3% of the annual global turnover. Additionally, providing inaccurate, incomplete, or misleading information to notified bodies and national authorities in response to a request could attract fines up to € 7.5 million or 1.5% of the total global annual turnover.

Crucially, the level of these sanctions is an effective deterrent, and it is inversely proportional to the size of the company. For SMEs and startups, the lesser of the two amounts is applied, reflecting a scaled approach that considers the financial capabilities and impact on smaller entities. This approach ensures that while the penalties are stringent enough to enforce compliance and maintain high ethical and safety standards, they are also fair and considerate of the challenges faced by smaller businesses.

This tiered penalty system highlights the EU's commitment to responsible AI use across all business scales, ensuring that larger firms face heftier fines, which in turn acts as a powerful deterrent, whereas smaller entities face lower thresholds, easing their burden and encouraging compliance.

The regulation is set to take effect 20 days following its publication in the Official Journal of the EU. However, in a manner similar to the implementation seen with the GDPR, most of its provisions will not be enforceable until 24 months have elapsed. This period is intentionally provided as a grace period to give businesses and stakeholders adequate time to make the necessary adjustments to comply with the new regulatory framework. Nonetheless, there are notable exceptions to this phased implementation: prohibitions on certain banned practices will take effect just six months after the regulation becomes active. Furthermore, codes of conduct are expected to be implemented nine months post-activation, and the rules concerning general-purpose AI systems, which include aspects of governance, will be enforceable after one year. For high-risk systems, the obligations will come into force 36 months after the regulation's effective date.

During this interim, companies will have the opportunity to analyze their systems to determine if they incorporate AI techniques and, if so, under which categories they fall. This analysis will help establish whether actions are needed to comply with the stipulations of the AI Act.

For systems posing limited risk, businesses may engage in self-regulation to adopt standards they deem most suitable for adhering to codes of conduct.

The situation is different for high-risk systems, which must undergo a conformity assessment to verify compliance with the regulation's requirements, ensuring the AI is reliable. This process will involve evaluating the quality of data, documentation, accuracy, and traceability, as well as assessing the system's transparency, cybersecurity, and robustness.

3.5 Taking Off: The Impact of the AI Act on Aviation

3.5.1 Foundations of Aviation: A Quick Overview of Regulations

Transitioning to the domain of aviation, the International Civil Aviation Organization (ICAO), under the auspices of the Chicago Convention of 1944, plays a crucial role. It prescribes Standard and Recommended Practices (SARPs), encapsulated in 19 Annexes, covering various aspects of aviation such as Meteorological Services, Air Traffic Services, Aeronautical Information Services, Environmental Protection, and more. However, it's pivotal to recognize that these ICAO Annexes wield no legal force unless incorporated into the national legislation of the contracting States.

Furthermore, the European Aviation Safety Agency (EASA) represents a cornerstone of aviation regulation within the EU. Established by the European Parliament and Council Regulation 1592/2002 and later superseded by Regulation (EU) 2018/1139 (the Basic Regulation), EASA is endowed with a distinct legal personality. It is tasked with regulatory and executive duties in the realm of civil aviation safety and environmental protection. The Basic Regulation outlines common requirements for safety and environmental sustainability in civil aviation, bestowing upon the European Commission the authority to enact detailed rules for its implementation.

This authority is elaborated through Powers and Recitals within the Basic Regulation. For instance, Recital (75) emphasizes the necessity of uniform implementation conditions, granting the Commission implementing powers, predominantly exercised in accordance with Regulation (EU) No 182/2011. Recital (77) allows for the delegation of powers to the Commission for amending or supplementing airworthiness, flight time limitations, and other significant areas, ensuring adaptability to technical, scientific, or safety needs.

EASA's responsibilities are extensive, ranging from drafting safety legislation to ensuring uniform application of EU aviation safety laws across Member States through inspection, training, and standardization programs. Moreover, recent expansions of EASA's role encompass drones, environmental protection, and cybersecurity in aviation, reflecting the evolving challenges within the aviation sector.

European aviation regulation is further delineated into Essential Requirements (ER) and Implementing Rules (IR), with the ER established by the European Parliament and Council to define objectives and essential requirements, and the IR developed and adopted by the Commission to operationalize the Basic Regulation's mandates. EASA contributes significantly to the production of ER and IR pertaining to civil aviation and environmental compatibility, advising the European Commission on technical matters within its expertise. Additionally, EASA autonomously adopts Certification Specifications (CS), Acceptable Means of Compliance (AMC), and Guidance Material (GM), providing clarity and guidance on meeting the regulatory requirements, albeit in a non-binding manner.

This comprehensive overview underscores the intricate interplay between various EU institutions and agencies in shaping aviation law, highlighting the critical role of EASA in promoting safety, environmental sustainability, and regulatory compliance within the European aviation sector.

3.5.2 The AI Act's Regulatory Impact on Aviation

The AI Act, in principle, applies to all domains, without any distinction. However, upon a closer examination of the text, specific exemptions or limitations to the rule's application can be identified for certain cases whose regulation is entrusted to other so-called harmonized laws.

Regarding the aviation sector, it is necessary to analyze various articles to understand how the legislator intended to regulate specific areas and to try to comprehend the *ratio* behind such choices.

For example, there are two important references to the EASA Regulations:

1. Annex I, section B.
2. Article 108.

1. Starting with the first, **Annex I, section B**, lists a series of Regulations in a '*Union harmonization legislation*', including: aviation security; rail system; market surveillance of motor vehicles; marine equipment; etc. Among these is the reference to the so-called **EASA Basic Regulation** (Reg. (EU) 2018/1139), specifically concerning the production and marketing of aircraft, unmanned aircrafts, etc.

   This Annex is referenced by **Article 2, paragraph 2** of the AI Act (from now on, if there is no reference to the regulation, it will be understood that it refers to the AI Act), which states that: '*For AI systems classified as* ***high-risk AI systems*** *in accordance with Article 6(****1****) and (****2****) related to products covered by the Union harmonization legislation listed in* ***section B of Annex I****,* ***only Article 112 applies***[11]'.

   So, **two conditions** must both be met (par. 1) in order to obtain the 'exemption':

   1. the AI System must be considered **high-risk under Article 6 (1)** if:

      *a. it is designed as a safety component of a product or is a product subject to EU harmonization legislation (****Annex I****)*

[11] We could call it a specific **exemption** to the entire AI Act.

b. *and requires a **third-party conformity assessment** for market placement or service according to that legislation (**Annex I**)*

Article 6 (2), states that, **in addition** to the points mentioned above (therefore, it is not an additional condition to be met, but a separate case in itself), also the '*AI systems referred to in **Annex III*** [this Annex concerns a list of areas, such as the use of biometric systems, rather than education, training, critical digital infrastructure, supply of water, etc.] *shall be considered to be **high-risk'***.

and

2. the related products must be covered by the cited harmonized legislation (**Annex I, section B**).

In very brief terms, this complex web of references between regulations helps us understand that if the system is to be considered **high-risk** (as classified above by Art. 6, par. 1) and par. 2)) and at the same time the related **products are governed by** harmonized legislation listed in *Annex I, Section B* (as far as we are concerned, the Basic Regulation 1139) then **only** Article 112 of the entire AI Act needs to be followed (which essentially requires monitoring for any changes to the regulation, thus staying updated).

2. The second reference mentioned is the **Art. 108**. This article provides for a **direct amendment to Regulation (EU) 2018/1139** (hereinafter also referred to simply as the Basic Regulation or BR).

This article focuses on introducing an obligation for those (we will find out who later) who are intending to adopt delegated or implementing acts, dealing with systems that are safety components in the indicated domains, must consider ('to *take into account'*) the requirements specified by the AI Act for high-risk systems.

In order to try to understand the reasons that led the legislator to envisage such amendments to the BR and only for some domains, it is necessary to make a brief introduction to the legislative process related of the EASA Regulation and on the relative powers.

The EASA Regulation is structured to provide a '*lex superior*', a higher level of regulation (**BR**) that establishes **common requirements** for regulating safety and environmental sustainability in the field of civil aviation and, at the lower level, the adoption of detailed rules in the Regulation's implementation (the so called Implemented and Delegated Regulations (hereinafter also referred to DRs and IRs).

In addition to setting out common requirements, the **BR also grants** the **European Commission** the **power to legislate** in relation to the DR and IRs[12]. After common requirements are introduced through regulations approved by the Parliament and the European Council, and the power for implementation is delegated, it will be enough for the European Commission to legislate on the implementation (based on the drafting by EASA).

Therefore, had this intervention not occurred at the BR's hierarchical level, any changes to the EASA regulations aimed at introducing new elements not included in the common requirements (as the new AI elements are) would have required the full legislative process, involving a vote from both the Parliament and the Council.

That said, one may legitimately ask why the legislator of the AI Act wanted to intervene directly in amending the BR, which, it is reminded, from a hierarchical point of view, is at the same level as the AI Act (both are EU Regulations, therefore voted on by both the Parliament and the Council).

The most logical answer seems to be that the legislator of the AI Act, wanting to **ensure that** for **certain domains there will be proper legislation**, including delegated legislation (as are the IRs), that considered those domains as deserving of special protection, or, as in this case, there

[12] In this sense, see recital no. 75 of the BR: '*In order to ensure uniform conditions for the implementation of this Regulation, implementing powers should be conferred on the Commission*. […]'.

was certainty that the domains referred to in the amendments[13] were considered high-risk and therefore should follow the related requirements specified, if the referred IRs were to be amended.

To confirm this viewpoint, see all the other titles of the **rules amended** in the same way (imposing the use of the requirements provided for high-risk systems[14]) in the Chapter XIII called '**Final Provisions**' from **Art. 102** to **Art. 110**:

- **Art. 102**: Regulation (EC) No 300/2008 of the European Parliament and of the Council of 11 March 2008 on common rules in the field of civil aviation **security** […].
- **Art. 103**: Regulation (EU) No 167/2013 of the European Parliament and of the Council of 5 February 2013 on the approval and market surveillance of agricultural and forestry vehicles […].
- **Art. 104**: Regulation (EU) No 168/2013 of the European Parliament and of the Council of 15 January 2013 on the approval and market surveillance of two- or three-wheel vehicles and quadricycles […].
- **Art. 105**: Directive 2014/90/EU of the European Parliament and of the Council of 23 July 2014 on marine equipment […].
- **Art. 106**: Directive (EU) 2016/797 of the European Parliament and of the Council of 11 May 2016 on the interoperability of the rail system within the European Union […].
- **Art. 107**: Regulation (EU) 2018/858 of the European Parliament and of the Council of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles […].

[13] [13] It concerns the following domains, with the respective article of the BR amended: AIRWORTHINESS (Articles 17 and 19); ATM/ANS (Articles 43 and 47); UNMANNED Aircrafts (Articles 57 and 58).

[14] All amendments provide the same wording: '*When adopting delegated acts* [or implementing acts], […] *which are safety components* […], *the requirements set out in* [the high-risk part of the AI Act] *shall be taken into account*.

- **Art. 108**: Regulation (EU) **2018/1139** of the European Parliament and of the Council of 4 July 2018 on **common rules in the field of civil aviation** and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU and 2014/53/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91.
- **Art. 109**: Regulation (EU) 2019/2144 of the European Parliament and of the Council of 27 November 2019 on type-approval requirements for motor vehicles and their trailers, and systems, components and separate technical units intended for such vehicles, as regards their general safety and the protection of vehicle occupants and vulnerable road users […].
- **Art. 110**: Directive (EU) 2020/1828 of the European Parliament and of the Council of 25 November 2020 on representative actions for the protection of the collective interests of consumers […][15].

The list highlighted above is exactly the same as the one mentioned in **Annex I, Section B**.

As further confirmation of the *ratio legis*, see **Recital no. 49** of the AI Act:

'*As regards high-risk AI systems that are safety components of products or systems, or which are themselves products or systems falling within the scope of […] Regulation (EU) 2018/**1139** of the European Parliament and of the Council31[16], it is appropriate to amend those acts to **ensure that the Commission takes into account**, on the basis of the technical and regulatory specificities of each sector, and without interfering with existing governance, conformity*

[15] This amendment has a different wording.

[16] Here the text of the note n. 31: *Regulation (EU) 2018/1139 of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU and 2014/53/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91 (OJ L 212, 22.8.2018, p. 1).*

assessment and enforcement mechanisms and authorities established therein, the mandatory requirements for high-risk AI systems laid down in this Regulation when adopting any relevant delegated or implementing acts on the basis of those acts.'.

So, the main points are:

- Ensure that the **Commission takes into account** the mandatory **requirements for high-risk systems** (the 'old' Title III, Chapter 2, now the Chapter III, Section 2 of the AI Act) when adopting any relevant delegated or implementing acts

  and doing this
- **Without interfering with existing** governance, **conformity assessment** and enforcement mechanism and authorities.

Given that the legislator wants to ensure that in certain areas and in the presence of AI safety components, the delegated legislator (EU Commission and EASA as drafting) takes into account the requirements provided for high-risk applications considered by the AI Act, one can question the methods and related timings with which it intends to intervene on the existing legislation.

By examining the diverse regulatory provisions to gain a comprehensive understanding, it seems reasonably clear that the legislator of the AI Act aimed to modify the existing regulations in various ways and at different times, as follows:

- **directly** to the safety product legislation under the **new legislative framework**[17] (machinery, toys, radio equipment, medical devices directive, etc.).
- only **when** the "old" safety legislation[18] is **revised** (automotive, aviation, etc.).

This would explain the reason why the list of "old" legislation is listed in Annex I, section B, and the exact same list is amended in Chapter XIII[19], to oblige the legislator to take into account the

[17] That is the list of the legislations provided by the Annex I, Section **A**.
[18] That is the list of the legislations provided by the Annex I, Section **B**.

requirements related to high-risk systems, and, also, the “exemption” for the same legislation provided by the Article 2, par. 2.

[19] Articles from 102 to 110.

4. The Flight Path Forward: EASA's Regulatory Vision for Artificial Intelligence

After defining the scope of the field of AI, starting from the definitions progressively modified by various bodies and entities, and after understanding how the European legislator intends to regulate AI for all sectors, we delve more into detail in our area of expertise, civil aviation, showing how EASA intends to act in the near future to adapt its regulations, as a drafter for the European Commission, to the ongoing evolution.

The strategy is clearly set out in the documents produced by EASA:

- The '**European Plan for Aviation Safety** (**EPAS**)[20]'.
- The '**Artificial Intelligence Roadmap. Human-centric approach to AI in aviation**' ((hereinafter also simply referred to as **EASA AI Roadmap**).
- The '**EASA Concept Paper**: Guidance for Level 1& 2 machine learning applications. A deliverable of the EASA AI Roadmap' (hereinafter also simply referred to as '**Concept paper**').

As explained earlier, the AI Act directly intervenes in the civil aviation sector through a specific amendment of the Basic Regulation that obligates the delegated legislator (the European Commission) to take into account the requirements set out by the AI Act for the high-risk AI-based systems, in certain cases and for certain domains.

The reason why this was done has already been explained in the dedicated chapter of this work, but what is important to highlight here is what EASA's approach will be in implementing the regulations on the use of AI both in the domains indicated by the above amendment and in other sectors.

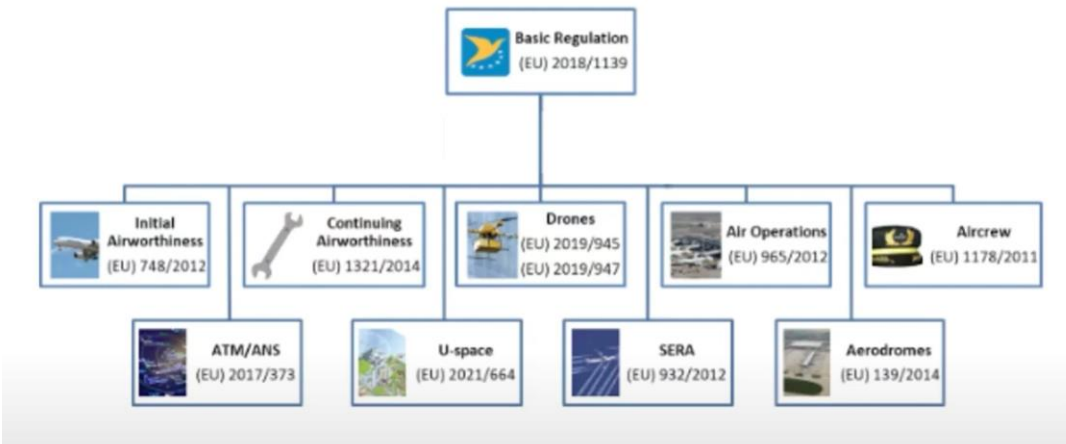
[20] In the EASA Basic Regulation (Regulation (EU) 2018/1139), Article 6 directly references the **European Plan for Aviation Safety (EPAS)**. According to Article 6, the European Aviation Safety Programme shall describe the process for the development, adoption, update, and implementation of the European Plan for Aviation Safety (EPAS). This plan is to be closely developed involving Member States and relevant stakeholders, indicating the EPAS's pivotal role in aligning national and European aviation safety efforts.

Starting from the *EASA AI Roadmap (2.0)*, where it's said that '*the analysis of the anticipated impact of AI on the various domains* […] *has been initiated in the EASA AI Concept Papers and indicates that the statement of issue is largely shared across domains. However, the first guidance development shows in addition the need to account for certain domain specificities.'*.

This means, essentially, that while being aware of the fact that it is desirable to approach the regulation of AI in a general manner, by producing regulations that uniformly govern all domains, it is nevertheless important to take into account some specificities of the individual sector concerned, and for this reason, EASA has decided to use a *'mixed rulemaking approach, involving on the one hand cross-domain rules (horizontal) and, on the other hand, domain-specific rules (vertical)*'.

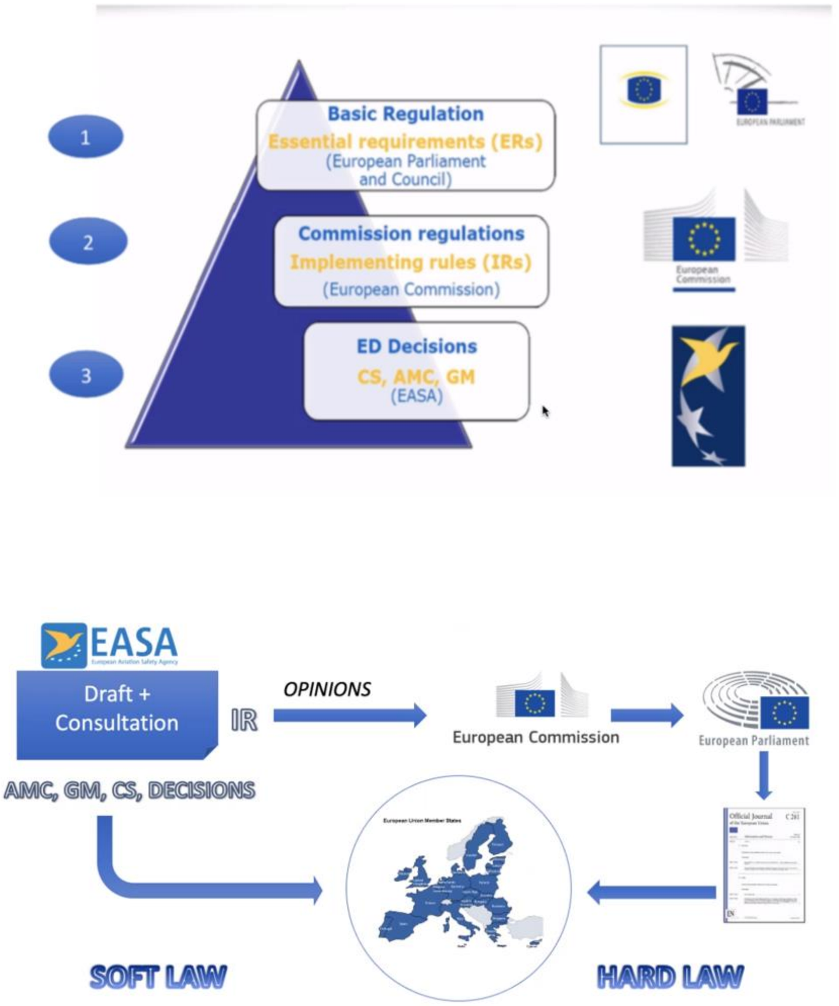
This is confirmed in the *Concept paper* where it states that '*Considering the potential application of AI/ML solutions in all the domains under the remit of the Agency, in line with the Rulemaking concept described in EASA AI Roadmap 2.0, EASA intends to define a common policy that can be applied to the whole of the EU civil aviation regulatory framework, while accounting for certain domain specificities*.'.

Therefore, to better understand how this mixed approach can be carried forward, let's do a brief refresh on how civil aviation is regulated today.



As can be seen at the highest level there is the Basic Regulation which set the substantive requirements on civil aviation and, also, the empowering provisions on the European

Commission to develop rules for the harmonized implementation[21] of these substantive requirements.



EASA Regulation structure

Under the empowering provisions in place since 2002, the Commission has developed these implementing rules. This hierarchy, visible beneath the Basic Regulation, outlines the necessary requirements for applying the substantial mandates of the Basic Regulation.

The historical basis for our current regulatory framework, structured by domain, stems from the member states entrusting the European Union with the authority to regulate and manage civil aviation safety effectively. This transfer of power began in 2002-2003 with competencies for

[21] This is what we call today 'Implementing Rules', 'Implementing / Delegated Act'.

Airworthiness, expanded in 2008 to include Air operations and Flight Crew Licensing, and further in 2009 to encompass Air traffic management and Aerodromes. The responsibility for drones was added in 2018.

This gradual delegation of authority let EASA to the development of implementing rules in a structured and reliable manner:

- From 2002 to 2004, the development of the IRs for Airworthiness.
- From 2008 to 2012, the development of the IRs for Air operations.
- From 2009 to 2014, development of the IRs for aerodromes.

This domain-specific approach might not be entirely suitable for topics that are transversal or horizontal, covering multiple areas. The question then arises: how can EASA most effectively address these horizontal topics when they impact all domains?

There are already examples of cross-domain applications in recent rulemaking activities. For example, the integration or transposition of the ICAO Standards and Recommended Practices (SARPs) on *Safety Management Systems* (SMS) was executed effectively. This success was achievable because it happened in parallel with the development of EASA's domain-specific rules architecture.

Additionally, across all these regulations, there are other similar elements that are horizontal very often, within the so-called *Authority requirements* and the *Organization requirements*. These elements are identical or closely similar across all domains.

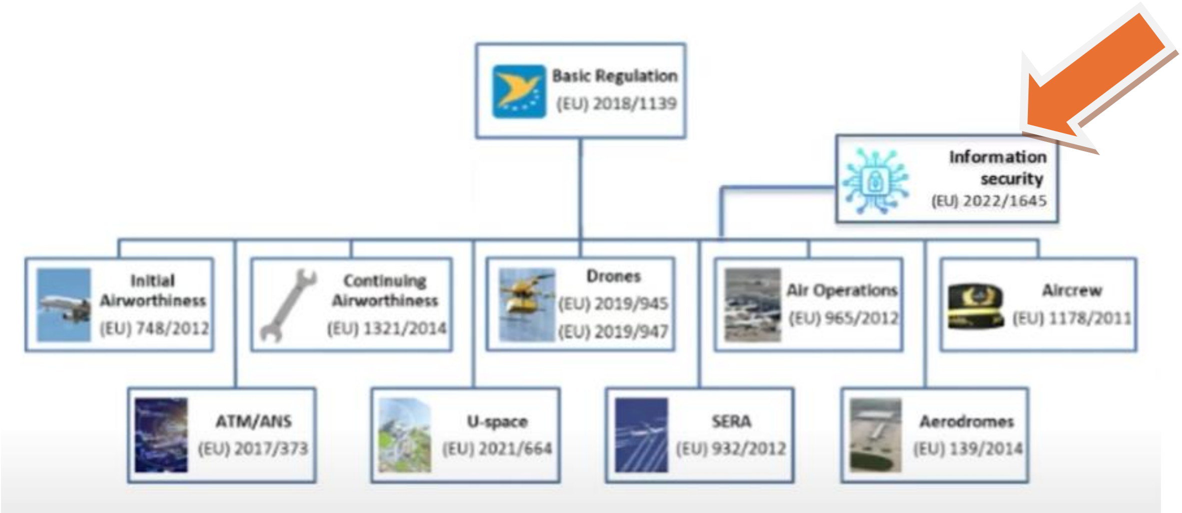
This approach undeniably offers a significant benefit: it enables users to easily identify which regulation to consult for specific requirements. For instance, to understand the requirements on Air operations, one would turn to IR 965/2012.

However, this method also presents clear disadvantages, including major ones. When developing requirements intended to be common, similar, or identical across all domains, and starting with a particular domain, let's say Air operations, EASA engages in consultation and discussion with the Air operations' community before enacting them in the Air operations Regulation. Then, intending to apply the same principles to the Airworthiness community, EASA faces the challenge of ensuring the Airworthiness community's perspectives are acknowledged and integrated. This requires balancing feedback from consultations to reflect these views accurately. Failure to do so means remain stuck with identical regulations across all domains.

This approach presents real difficulties, especially in maintaining consistency in rules across domains. The moment updates or changes are implemented in one domain, achieving a harmonized approach becomes increasingly complex, threatening the consistency of the regulatory framework across different areas of aviation.

So, the approach has changed in the recent years because it is much easier to have common rules, common requirements regulated in a specific and dedicated act and in order to address the potential inconvenience for users, who may no longer find rules within their familiar, domain-specific contexts, EASA has innovated by digitalizing the regulations. This led to the creation of the so called 'e-rules' (easy access rules), a system that enables EASA to flexibly organize and present every rule as needed. This solution not only makes it easier for users to navigate the regulations but also allows for a more dynamic and responsive regulatory environment.

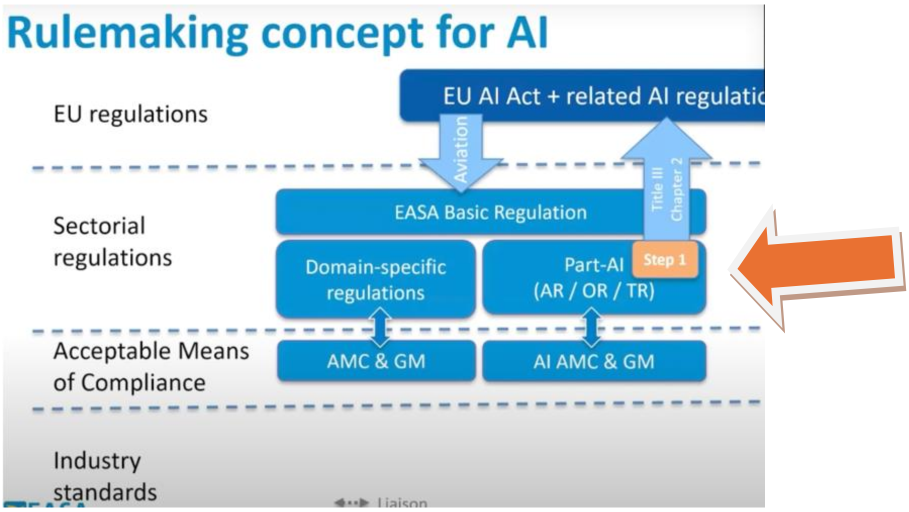
For the first time, EASA has moved away from the traditional method of regulating topics vertically within each domain's specific rules architecture. Instead, it has addressed a cross-disciplinary, transversal topic through a dedicated regulation. This groundbreaking approach has been applied in the creation of Regulations on Information Security.



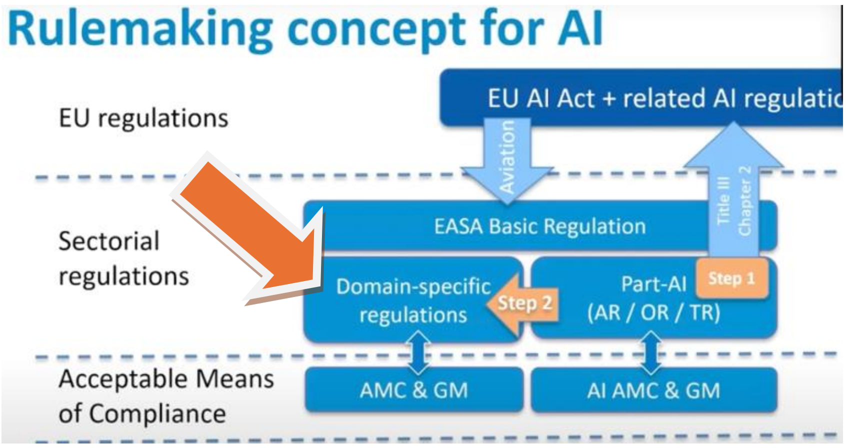
AI represents a horizontal topic too, as it impacts multiple domains. This raises an important question: How can EASA integrate AI into its regulatory framework in the future?

To answer this question, let's go back to what EASA says in its document **'EASA AI Roadmap 2.0'** about the 'mixed approach' that will be developed in two phases:

- ***Step 1****: development of a* ***transversal Part-AI*** *to contain the three major provisions anticipated in the EASA Concept paper guidance:* ***requirements for authorities*** *(Part-AI.AR),* ***requirements for organisations*** *(Part-AI.OR) and* ***requirements linked to AI trustworthiness*** *(Part-AI.TR)* […]. *In addition, some acceptable means of compliance (AMC) and guidance material (GM) are anticipated to account for the 'anticipated MOC' currently captured in the EASA AI Concept Papers, to reference or complement where necessary applicable industry standards.*



- ***Step 2:*** *analysis, per domain, of those* ***requirements that are domain-specific*** *and those that need to be complemented to provide an adequate regulatory basis for deploying the new Part-AI.*



The EASA idea is to transpose what already done in the Guidance Materials, the Concept papers, into Regulations, taking into account the requirements set by the AI Act when amending the existing domain-specific Regulations, evaluating the necessity for modifications, and simultaneously developing common requirements pertinent to all domains (part AI, that will be performance based, with: Trustworthiness requirements AI.TR; Authority requirements AI.AR; Organization requirements AI.OR). The approach mirrors the methodology applied in the recent

Information Security Regulation, aiming to replicate its concept for broader applicability and coherence across different aviation sectors.

This priority is confirmed by **EPAS vol 1**, where it states that '*in February 2022 EASA established the* ***EASA Artificial Intelligence Programme****. The Programme revolves around five objectives'*:

- Develop a human-centric AI trustworthiness framework.
- Make EASA a leading certification authority for AI.
- Support European aviation leadership in AI.
- Contribute to an efficient European AI research agenda.
- Contribute actively to EU strategy and initiatives.

And most importantly, regarding **rulemaking**:

- transforming the AI Concept Papers' guidance into a generic set of acceptable means of compliance.
- referencing this guidance in the aviation regulations of the domains impacted by AI; and
- preparing for an EU regulation laying down harmonised rules on AI.

And to address the rulemaking plan, the reference is in the **EPAS vol 2, RMT.0742** called '*Artificial intelligence trustworthiness'* which provides the following.

Development of an ***implementing or delegated regulation*** *for the* ***AI trustworthiness framework*** *in response to the future EU AI Act Title III Chapter 2, development of associated set of generic AI-related AMC and GM, and implementation of the necessary* ***adaptations to domain-specific***

***regulatory** material for aviation **domains identified** in future the **EU AI Act Article 81**[22] (Amendment to Regulation (EU) 2018/1139 for the domains of airworthiness (Articles 17 and 19), ATM/ANS (Articles 43 and 47) and unmanned aircraft), **as well as for other affected domains** (e.g. aerodromes).*

The affected Regulations will be: (EU) No 748/2012 (Certification of Aircraft), (EU) 2019/945 and (EU) 2019/947 (UAS design, production, rules, procedures, etc.), (EU) No 1321/2014 (Continuing Airworthiness of Aircraft), (EU) No 965/2012 (Air Operations), (EU) No 1178/2011 (Aircrew), (EU) 2015/340 (Technical Requirements and Administrative Procedures Relating to Air Traffic Controllers' Licenses), (EU) 2017/373 (Requirements for Providers of Air Traffic Management/Air Navigation Services and Other Air Traffic Management Network Functions), (EU) 2021/664 (Establishing a Framework for the Management of Network Functions and Amending Regulations (EU) No 551/2004 and (EU) No 552/2004), (EU) No 139/2014 (Aerodromes).

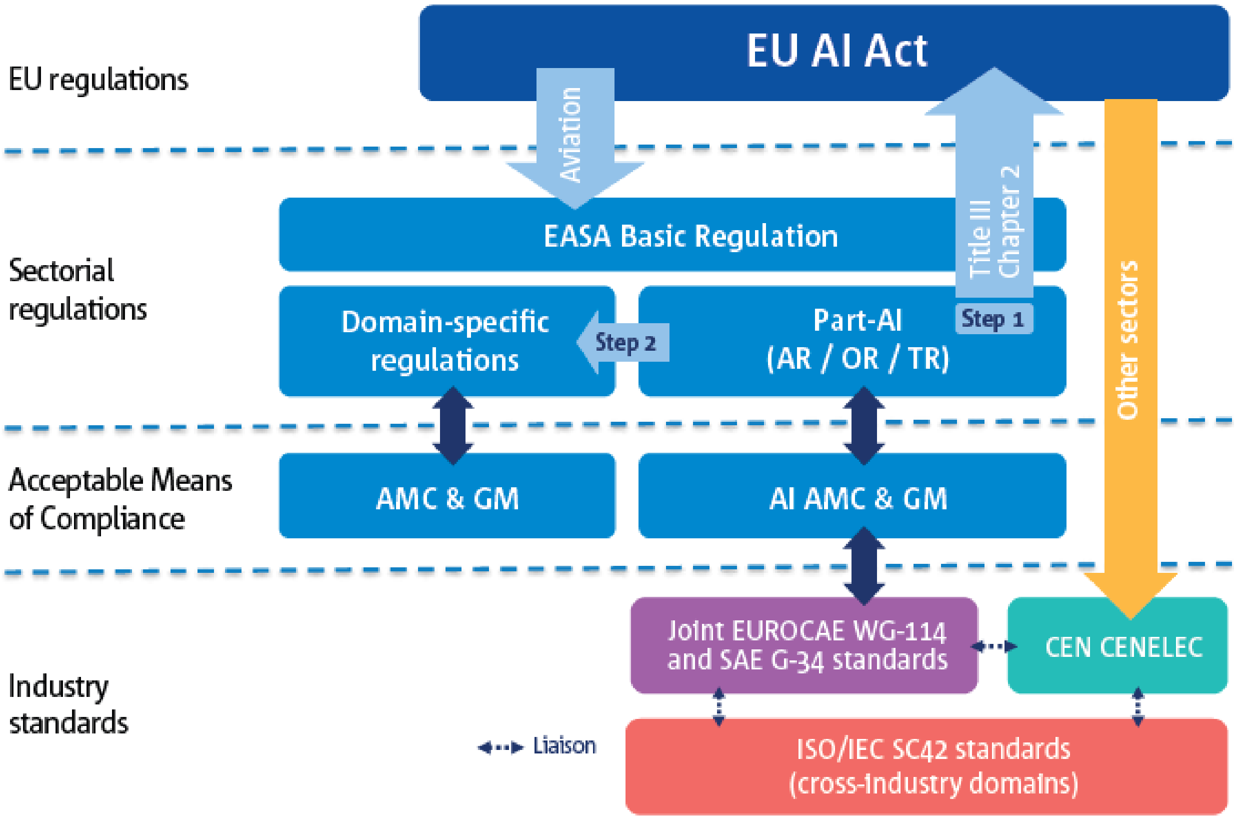
Even for the timing, the planning milestones are the following:

PLANNING MILESTONES					
SubT	Initiation	Consultation	Opinion	Commission IR	Decision
1	2024-Q1	2025	2026	2026	n/a
2		2025	n/a	n/a	2026
3		2025	2026	2027	2027
4		2027	2027	2027	n/a

To finalize the process, industry standards play a crucial role. EASA participates in the EUROCAE Working Group 114 and the SAE G34. It's critical that all components of the industry standards, including those from ISO and IEC, are well-coordinated.

In the Figure below, the hole anticipated regulatory structure for AI.

[22] The latest version of the AI Act, voted by the European Parliament, has changed numbering, so the article referred to now is Article 108.



One last consideration about the integration and the compatibility between the AI Act and the future part AI of the future EASA Regulations, is that the objective-based / performance-based future EASA part AI Regulation actually reflects what is outlined in the high-risk Requirements of the AI Act, covering aspects such as data governance, technical robustness, safety and security, transparency, human oversight, etc.

It can be reasonably stated that given the timelines for the AI Act's entry into force and the development of the new EASA regulatory framework, there will be a uniform and coherent application of the respective regulations. This is made possible by EASA producing its guidance material and carrying out regulatory drafting in parallel and in continuous consultation with the overarching regulation of AI through the AI Act, thereby ensuring alignment and consistency.

The key consideration moving forward is determining the appropriate actions for AI application developers within the aviation sector during the interim period leading up to the implementation of the new regulatory framework. At present, adhering to EASA's guidance materials is the advised course of action.

5. EUROCONTROL

5.1 Eurocontrol Unveiled: The Heartbeat of European Skies

For those unfamiliar with EUROCONTROL (from now also just ECTL), a brief introduction is in order to elucidate what this organization represents and its pivotal role within the European aviation sector. Established by the International Convention relating to Cooperation for the Safety of Air Navigation on December 13, 1960, and coming into operational effect on March 1, 1963, ECTL is a central, pan-European, civil-military body committed to enhancing aviation safety, efficiency, cost-effectiveness, and environmental sustainability across Europe.

This organization stands at the forefront of advancing the European Union's Single European Sky vision, aiming to unify and streamline air traffic management (ATM) across the continent. ECTL's remit spans a broad array of functions, from operational management and service provision to concept development, research, and the implementation of pan-European projects. It focuses on raising performance standards, coordinating with various aviation stakeholders, and supporting the sector's future strategic evolution.

A testament to its central role and expertise, ECTL was initially appointed as the Network Manager with a Commission Decision on July 7, 2011. This mandate was to run until the end of the Performance Scheme's second Reference Period, which was December 31, 2019. Following this, ECTL was reappointed by the European Commission as the Network Manager for the years 2020-2029, based on Commission Implementing Decision C(2019)709 dated May 6, 2019. This role involves a wide array of responsibilities, from operational management of air traffic to extensive collaboration with civil and military stakeholders, aimed at enhancing the safety, efficiency, and sustainability of European aviation.

The organization's value lies in its neutrality, allowing it to bring together the diverse aviation community for the collective benefit of the European ATM network. With a deep understanding of the entire spectrum of ATM technology evolution, from initial concepts to detailed standards and operational practices, ECTL facilitates the seamless integration of civil and military airspace usage, thus enhancing the network's overall performance.

Moreover, ECTL's expertise and longstanding knowledge in air traffic management empower it to develop innovative solutions and technologies. These initiatives aim to improve safety and efficiency levels through forward-thinking operational procedures. By collecting and disseminating extensive data, including statistics and forecasts, ECTL provides stakeholders with a comprehensive view of European aviation performance, enabling informed decision-making for the future.

Governance at ECTL is structured around the Permanent Commission and the Provisional Council, with the Agency serving as the executive arm. This setup ensures the involvement of a wide range of stakeholders in guiding the organization's efforts, reflecting the interests of all parties involved in aviation, including air navigation service providers, airspace users, airports, and aircraft/equipment manufacturers. Through its concerted efforts, ECTL is instrumental in addressing the industry's challenges, promoting collaboration, and steering European aviation toward a safer, more efficient future.

ECTL also operates the Experimental Centre, which serves as its principal laboratory and research facility for air traffic management (ATM). Located in Brétigny-sur-Orge, France, this centre is a leading facility in Europe dedicated to the development, experimentation, and evaluation of ATM concepts, technologies, and procedures.

The Experimental Centre plays a crucial role in advancing ECTL's research and development efforts. It enables the testing of innovative ATM solutions in a controlled environment before their deployment across the European network. This facility is equipped with advanced simulation and modeling tools that allow ECTL and its partners to assess the performance and safety implications of new ATM strategies, supporting the organization's mission to improve the overall efficiency and sustainability of air traffic in Europe.

5.2 Navigating New Heights: The Role of AI in Eurocontrol's Evolution

As we can see in this work, it's essential to highlight that ECTL is actively engaged in leveraging Artificial Intelligence within its technological sectors, especially through the Tec Division's 'NM-lab' for the NMOC. This involvement is aimed at developing AI-driven applications with several critical objectives, including the prediction of cyber-attacks to enhance operational stability by forecasting various air traffic elements (such as turnaround times, taxi durations, arrivals, diversions, and delays). Additionally, there's a focus on monitoring CNS infrastructures to ensure adherence to the NM mandate, forecasting airport operations to assess their network impact, and automating repetitive tasks for the NMOC, like prioritizing NOTAMs and managing regulations.

ECTL is now setting its sights on the forthcoming 'iNM Wave 3,' marking a pivotal shift towards an AI-enabled and data-driven paradigm. The core aims of this transition involve a seamless migration to an upgraded AI platform, adherence to regulatory standards, and the successful incorporation of AI within iNM. By adopting AI, ECTL aims to achieve significant reductions in delays, thereby generating savings for various aviation stakeholders, decreasing infrastructure costs, enhancing network efficiency, and facilitating the integration of existing AI technologies within iNM.

5.3 My Traineeship

During my internship at EC, my main focus was on understanding how to be compliant with regulations, which were mostly still in draft form or in guidance material, in the development of artificial intelligence-based applications. EC, as mentioned, is developing many AI-based applications, but it also plans to use applications not related to the aviation domain in the future, including those not developed in-house.

Therefore, it has been very relevant considering the current historical period, characterized by the imminent enactment of the world's first AI regulation covering all sectors, and by the concurrent efforts of EASA in producing guidance material, anticipated MOCs, and legislative drafting.

The work unfolded in collaboration among various EC departments interested in the project: the LAB, Safety Department, and Infrastructure Department, aiming to coordinate and assess all related aspects. In addition to collaborating with a ‘working group’ established to create a framework on the current situation for the 'Legal service' (not only for the aviation sector), I worked closely particularly with the ‘Lab team’ and the ‘Safety Department’ and I was assigned two internally developed applications in advanced stages of development: NOTAM AI and ISOBAR.

To understand how to practically engage in the various stages of development and deployment of applications, in compliance with current and forthcoming regulations, all relevant aspects were analyzed, from technical to legal considerations.

The following chapters will detail the framework for the two mentioned applications more thoroughly.

6. USE CASES

6.1 NOTAM AI in NMOC with Machine Learning

As known to everyone involved with the aviation world, both professionally and at an amateur level, '*A Notice to Airmen (NOTAM), also known as Notice to Air Mission (FAA definition), is a notice containing information concerning the establishment, condition or change in any aeronautical facility, service, procedure or hazard, the timely knowledge of which is essential to personnel concerned with flight operations*'[23]. In essence, they are notices primarily aimed at informing pilots about potential imminent dangers they might encounter during flight.

It's interesting to note that the acronym NOTAM came into common use following the ratification of the CICA in 1947, and they are heavily criticized for many reasons, including the presence of a free text field. To address this, the concept of a digital NOTAM (DIGITAM) was introduced, which would replace the text with a series of 'structured facts', but its development is currently on hold.

What is less known is who deals with these notices on a daily basis and how they make them available to all industry operators.

NOTAMs are primarily issued by national air traffic control authorities or specific airport operators[24] once there's a need to communicate a notice, or a modification of a previous notice, relevant to flight safety.

Once drafted, the NOTAM is sent to EC[25], which acts as a central hub for the subsequent distribution of flight information in Europe.

[23] ICAO Annex 11, Air Traffic Services.

[24] Airspace management bodies; airport operators; Entities in charge of critical infrastructure maintenance.

[25] Before being sent to Eurocontrol, the NOTAM undergoes a preliminary check by the national NOTAM office, in Italy by Enav, which then proceeds with its subsequent publication in the national and international system, thereby allowing its distribution beyond national borders.

EC is then tasked with collecting all these notices, verifying their compliance with international standards, and entering them into the centralized EAD INO system[26], thereby making them available for consultation by all involved parties.

As a result, EC receives daily between 600 to 1000 NOTAMs, which are monitored by the NMOC[27], and only a small percentage of them require intervention, the so-called DMRs[28]. However, to identify these, it's necessary to read all received NOTAMs, without an order of importance.

This is where the idea of automating processes to prioritize DMRs, considered more important, over other NOTAMs was born, and it was decided to resort to AI for this purpose. To better understand this innovation, it's necessary to discuss the structure of a NOTAM and how AI, through its predictive capabilities, comes into play.

[26] EAD INO: European AIS Database - International NOTAM operations. service helps data providers create, maintain and distribute NOTAM, SNOWTAM and ASHTAM for their NOTAM office in accordance with ICAO Standards and Recommended Practices (SARPS) and EUROCONTROL Operating Procedures for AIS Dynamic Data (OPADD).

[27] Network Manager Operations Centre optimises traffic flows by constantly balancing capacity and demand, while helping to ensure the safe and efficient operation of flights going to and over Europe.

[28] Data Modification Requests.

6.1.1 NOTAMs, DMRs and Prioritization

A NOTAM is structured as a document with multiple sections, starting with its identifier at the top (typically the first line). Following this, each section is denoted by a letter followed by a closed parenthesis, such as ‘Q)’. These sections may include Q, A, B, C, D, E, F, G, H, with some being optional. The 'E' section allows for **free text** and may contain snippets of other NOTAMs. Despite efforts for standardization, deviations occur, such as non-English texts within the free text area. The European AIS Database (EAD), which supplies NOTAMs to ECTL, is on a continuous path to enhance their precision, though it's only permitted to amend a specific subset of NOTAMs.

Here an example of NOTAM with the explication:

NOTAM (B5678/23)

Q) LIRR/QFAXX/IV/NBO/A/000/999/4154N01229E005

A) LIRF

B) 2302100000

C) 2302102359

D) 0000-2359

E) COVID-19: PASSENGER RESTRICTIONS APPLY.

F) GND

G) 999FT

Let's explain each component of this NOTAM:

NOTAM (B5678/23): The reference or identifier of the NOTAM. 'B5678' is the **NOTAM number**, and '23' indicates the **year of issue**, 2023.

Q) LIRR/QFAXX/IV/NBO/A/000/999/4154N01229E005: The qualifier section detailing the **purpose and scope**.

- LIRR: The FIR (Flight Information Region) for Rome, indicating the NOTAM's applicable airspace.
- QFAXX: The NOTAM code, signifying the nature of the notice. 'QFA' relates to aerodrome facilities, and 'XX' is a placeholder.

- IV/NBO/A: Classification codes - 'IV' signifies IFR (Instrument Flight Rules), 'NBO' for the Notifying Body, and 'A' for Aerodrome.
- 000/999: The altitude range affected, from the ground to 999 feet.
- 4154N01229E005: Geographic coordinates with a radius, indicating the notice is relevant within 5 nautical miles of the point 41°54'N 12°29'E, which is near Rome's airport.

A) LIRF: The **ICAO code** for the affected location, Leonardo da Vinci–Fiumicino Airport in Rome.

B) 2302100000: The **start date and time** of the NOTAM's validity, here February 10, 2023, at 00:00 UTC.

C) 2302102359: The **end date and time** of the NOTAM's validity, here February 10, 2023, at 23:59 UTC.

D) 0000-2359: The **daily operational time range**, indicating the NOTAM is in effect for the entire day.

E) COVID-19: passenger restrictions apply: The description of the condition, including actions, precautions, or instructions. **This field contains free text** with specific information about passenger restrictions due to COVID-19 at the airport.

F) GND: The **lower limit of the affected airspace**, indicating the ground level.

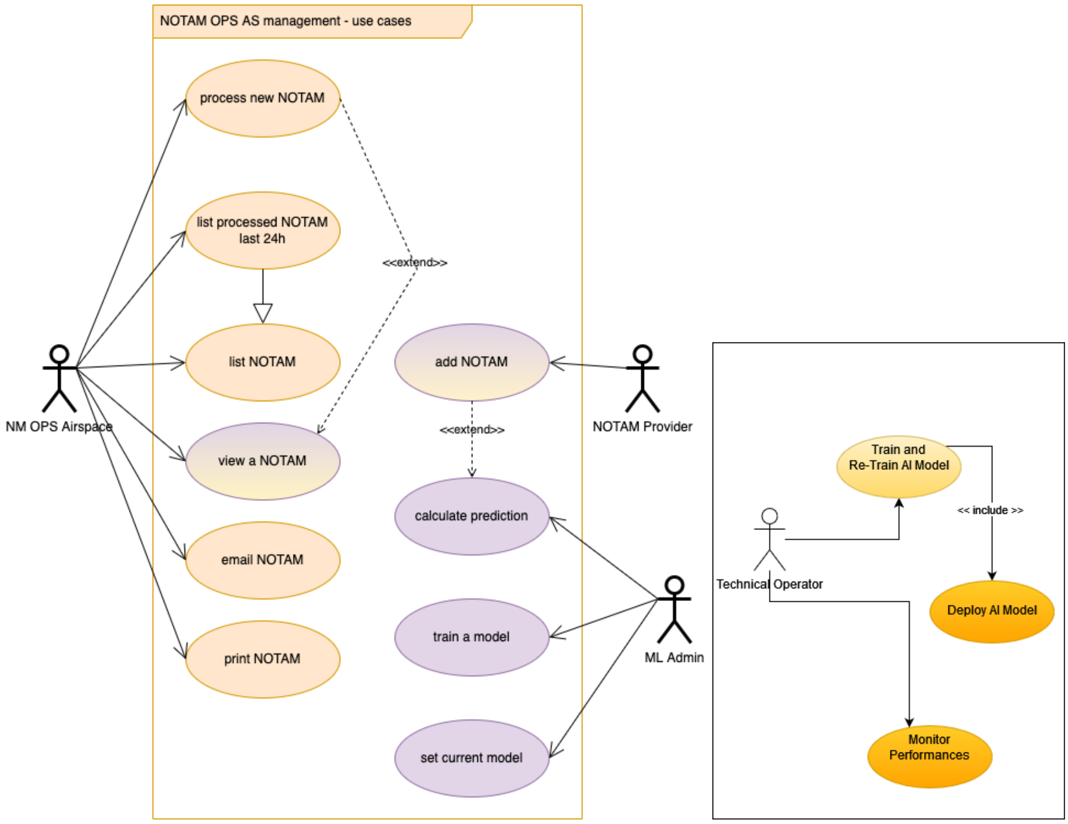
G) 999FT: The **upper limit of the affected airspace**, up to 999 feet.

Now, imagine the aforementioned NOTAM is incorrect because it contains a wrong date. In such a case, it's crucial that it be corrected as soon as possible, since incorrect information poses an even greater danger than the absence of information or a delay in providing information. To do this, a DMR (Data Modification Request) needs to be issued.

The operator at the NMOC will receive this modification request not through a separate information channel, but in the same information flow as all the other NOTAMs received during the day, with these DMRs representing about 1.5% of the total. It thus becomes evident that searching for such messages, considered important, among a thousand others deemed less important, is practically impossible. The only way is to scroll through the messages chronologically, as they come in. Therefore, within EC, the idea was born to use a technological system capable of assigning a priority, a sort of **score**, to the various types of NOTAMs.

6.1.2 How AI can help. Introducing Prediction.

Given the vast amount of historical data in the EC database regarding NOTAMs and their related DMRs, the immediate thought was to use a machine learning (ML) system for making '**predictions**', as it is particularly effective with a significant dataset for learning. The technological choice was to use a hybrid system combining standard algorithms, ML, and Natural Language Processing (NLP) to understand the parts of the NOTAM written by the user. Below is a brief description of the technical workings.



In the Figure above, it can be seen that in addition to the standard use case, new use case (colored in purple) is added in order to administrate the model (training the model, generating the prediction, evaluating the quality of the model).

The prediction is calculated through 5 steps:

- parsing,
- keyword filter,
- fixed filter,

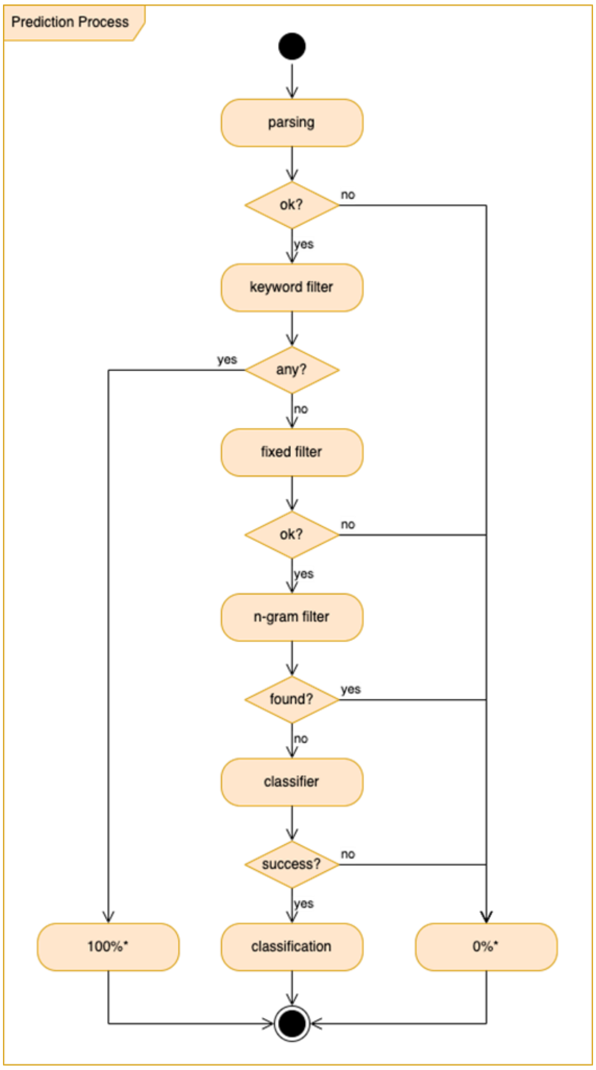
- n-gram filter,
- classifier.

The first four steps are filters that are looking for NOTAM's properties which are linked either to:

- skipping the NOTAM – because of the 0% probability – relevance of DMR – prediction '0'.
- DMR generation – because of the 100% probability – relevance of DMR.
- O

Once the two certain cases have been identified, it is necessary to apply ML to the NOTAMs still under scrutiny.

Below is the flowchart of the 'prediction process'.



- **Parsing**: Break down the NOTAM text into its separate components, which isn't always feasible due to the occasional poorly formatted NOTAMs. These issues mostly occurred before 2014, before the EAD started improving some NOTAMs' quality.
- **Keyword Filter:** An optional tool to aid OPS operators. If keyword input is enabled, the system will look for these keywords in the NOTAM's **E field**. Discovering a keyword guarantees a 100% match.
- **Fixed Filters**: Created by analyzing historical data to identify combinations that **never** resulted in a **DMR**. Over eight years of NOTAMs managed by NOMC were examined, finding certain category combinations never led to a DMR. These

combinations, mainly based on the **Q code field**, are deemed fixed, automatically assigning a **0% DMR** relevance to the NOTAMs.

- **N-Gram Filter**: Identifies word sequences (N-Grams) previously used and their effect on DMR issuance. It attributes a 0% DMR relevance to NOTAMs containing frequently used sequences that never resulted in a DMR.
- **Classifier**: A machine learning model that predicts the likelihood of a NOTAM resulting in a DMR, using data from the **Q code** and **E field**. It has tested two algorithms: Logistic Regression and Neural Network.

The reasoning for using **negative filters** in the system is threefold. First, they significantly reduce the number of negative examples in what is an unbalanced dataset, thereby improving the dataset's balance and making the classifier's job easier. Second, they clear the dataset of any non-standard usages, which could potentially confuse the classifier, a point that holds particular importance for the fixed filters. Lastly, they offer NMOC users clear explanations for the low scores of the NOTAMs they filter out, which in turn, improves the software's ability to explain its decisions.

The acceptance criteria for the model are mainly that at least 85% of NOTAM with DMR shall be classified (by the neural network') in 'score A' (A is the most susceptible to producing DMR, F is the lowest possible score).

As seen, the difficulties of the classification are basically in which features give to the algorithm to make the choices. Some are very simple (like NOTAMs with 'New' or 'Replace' or 'Cancel') and clearly referred to a DMR, but others are much more difficult to engineer, such analyzing the human-written texts via NLP.

6.1.3 Results

In conclusion, instead of navigating through NOTAMs that arrive in a random manner, operational staff will first look for flagged with an A score or both A and B scores. This focused method ensures that up to 92% of NOTAMs requiring a change in NM systems are identified.

The goal is also to minimize operational mistakes by giving operators the opportunity to spend more time on the most pertinent NOTAMs. However, with the introduction of AI, there's the risk of introducing a new error type known as automation bias. This happens when users excessively trust the technological system, leading them to overlook important, contradicting information. Nevertheless, the impact of automation bias can be lessened by simplifying the presentation of information, like changing probabilities into letter grades, or providing additional context rather than direct instructions.

This tool won't eliminate the need for NMOC to review the entire daily influx of NOTAMs, as overlooking even a single notice is not an option and it's a given that the software will occasionally make errors, that's unavoidable. However, this tool will enable operational staff to manage their workload more effectively by prioritizing based on the scores assigned to NOTAMs.

Consequently, staff can more quickly address NOTAMs requiring DMRs, which will be highlighted at the top of their list and choose quieter moments of the day to review the remaining ones.

Regarding NOTAM AI, we will not proceed with detailing the various assessments in this work, both because such work is still in progress for NOTAM AI and because the assessments carried out for ISOBAR, which will be illustrated in the following chapter, are fundamentally similar in principle as a *modus operandi*.

6.2. ISOBAR

The ISOBAR initiative has effectively blended accurate and probabilistic forecasts of convective weather into air traffic flow management (ATFM) processes. This integration aims to enhance predictions of demand and capacity, enabling a more accurate prediction and management of variances, thereby enhancing safety, efficiency, and overall capacity.

The project leads to a flexible network plan that adapts based on demand, capacity, and weather changes.

The project has shown that its AI-driven methods for managing weather events represent a significant advancement in managing network performance in challenging conditions. This is supported by both quantitative and qualitative validation outcomes.

Machine learning plays a crucial role in forecasting convective weather, offering reliable predictions that aid spatial and temporal decision-making in ATFM.

Operational evaluations have demonstrated the potential of machine learning in predicting how traffic and capacity might respond to adverse weather, leading to highly accepted and practical concepts for managing such situations.

Although there is a limitation regarding the availability of certain data, such as specific Airspace Users' preferences, the project has applied statistical methods to assess route choices, underlining the importance of incorporating Airspace Users' preferences into mitigation measures.

The project has developed automated tools that surpass traditional methods in reducing delays and the number of flights delayed.

The project's success in operational validation exercises points to a need for further development and refinement towards higher maturity levels and performance optimization.

Air Traffic Flow and Capacity Management (ATFCM) within Air Traffic Management (ATM) is critically dependent on accurate forecasting of capacity and demand, and the system's ability to respond accordingly. This setup is particularly vulnerable to weather unpredictabilities and their forecast inaccuracies. Moreover, current ATFCM practices lack a systematic evaluation, heavily relying on the expertise of human operators for assessing forecasts and making decisions.

The project ISOBAR aims to enhance both the predictability of situations and the effectiveness of responses to imbalances between demand and capacity. It proposes an AI-driven approach to Network Operations Planning (NOP), encompassing:

- Improved weather forecasts tailored to ATFCM needs, emphasizing not just prediction accuracy but also the presentation of these forecasts to users and decision-makers.
- Integration of ATM and weather data, managing various formats and granularities to enable collaborative component functioning and seamless user experience.
- AI-facilitated automation for identifying and addressing demand and capacity imbalances, leveraging AI for precision in forecasting imbalances and devising complex mitigation strategies—ensuring human oversight remains central to reviewing AI-generated solutions, blending advanced cognitive skills and invaluable experience.

The ISOBAR research holds particular significance for pre-tactical ATFCM, executed by flow managers at the local (Area Control Centre (ACC) level) and by the Network Manager on a European scale. It focuses on en-route operations, yet the methodologies and concepts it introduces are adaptable for use in Terminal Manoeuvring Areas (TMA).

The Figure below[29] illustrates the ISOBAR overall vision, representing in blue the core processes and operational parts, whereas in green the data provisions. The figure introduces the process to create an adaptive NOP depending on unfolding weather conditions, within a range between -24H to operations.

[29] Source: ISOBAR, Final Project Results Report, Sesar Joint Undertaking.

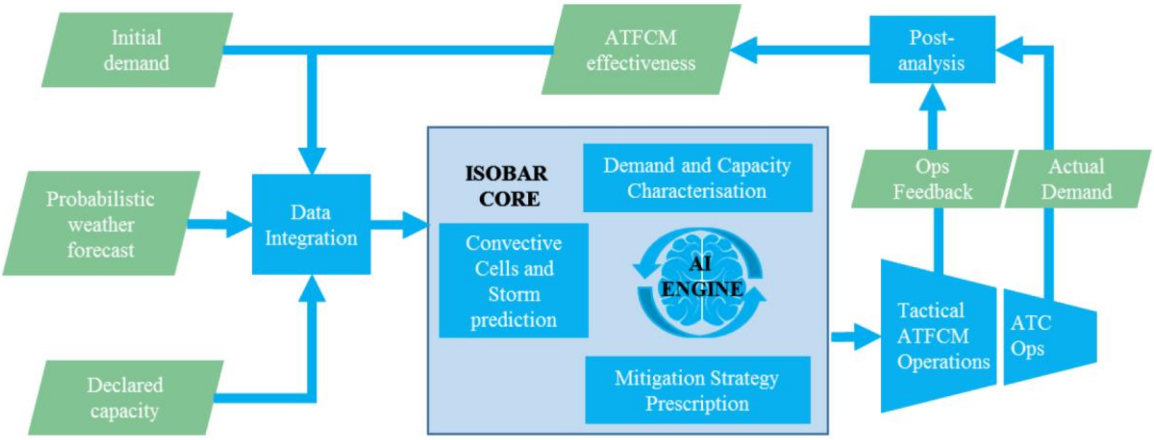


Figure 1: ISOBAR Enhanced ATFCM Vision

6.2.1 ISOBAR's first Assessments

After understanding what kind of application ISOBAR is with this overview, it is necessary to analyze all the existing regulations and those about to be published, as well as all the existing guidance material to understand how to proceed for a possible use of the app.

Assuming that it is not necessary to apply the prerogatives contained in the AI Act at the current state, it is necessary to apply the guidance material provided by EASA, namely the '*Concept Paper*' already mentioned previously.

The scope of the entire guidance (the *Concept Paper*) is anticipating future EASA guidance and/or requirements to be complied with by safety-related ML applications, providing applicants with a first framework to help in taking choices in the development phase for ML solutions.

The **guidance applies to**:

- any system that incorporates one or more **ML models**,

  and

- are intended for use in **safety-related applications** (or for applications related to environmental protection) **covered by the Basic Regulation**[30].

For the applications subject of this work, it is necessary to focus primarily on the ATM/ANS (Air Traffic Management/Air Navigation Services) domain, for which it is important to specify the following.

The process of managing changes to the functional system, including hardware, software, procedures, and personnel, follows a structured approach under Regulation (EU) 2017/**373**:

- Change Management Procedures: Changes are managed as part of the air navigation service provider's change management process.

[30] In particular for the following domains: Initial and continuing airworthiness; Air operations; ATM/ANS; Maintenance; Training; Aerodromes; Environmental protection.

- Competent Authority Approval: Approval is required for the complete change introduced.
- **Safety Assessments**:
  - Air Traffic Service (ATS) Providers: Required to perform a safety assessment as part of the change management process.
  - **Non-ATS Providers** (e.g., CNS): Required to perform a **safety support assessment**. This assessment ensures that, after the introduction of the change, the associated services will operate as specified and continue to do so[31].

Furthermore, **new Regulations** have been adopted in **support** of the **conformity assessment** framework in the **ATM/ANS domain**:

- Delegated Regulation (EU) 2023/**1768**: Specifies detailed **rules for the certification** and declaration of ATM/ANS systems and constituents.
- Implementing Regulation (EU) 2023/1769: Sets out technical requirements and administrative procedures for the approval of organizations involved in the design or production of ATM/ANS systems and constituents.

The Conformity Assessment Framework now also includes AMC (Acceptable Means of Compliance), GM (Guidance Material), and DSs (Design Specifications) for the certification, declaration of conformity, or statement of design compliance of ATM/ANS equipment.

[31] As will be seen later, this will be the type of assessment that will be carried out for ISOBAR.

6.2.2 The EASA's building blocks

The EASA AI Roadmap 1.0 introduced **four critical 'building blocks'** essential for crafting a trustworthy AI framework and preparing for AI/ML use in aviation. However, with advancements and insights from the document and EASA AI Roadmap 2.0, the **scope** of two original building blocks has been expanded:

1. **AI Trustworthiness Analysis building block**

- Interfaces with the EU Ethical Guidelines by the EU Commission.
- Begins with **AI application characterisation**, including **ethics**, **safety assessment,** and **security assessment**.

These assessments are prerequisites for the development and approval of AI/ML systems in aviation, adapting existing mandatory safety and security practices for AI specifics.

2. **AI Assurance building block**

Addresses AI-specific guidance for AI-based systems, covering:

- **Learning Assurance**: Shift from traditional programming to learning, requiring new assurance methods for AI/ML learning processes.
- **Development & Post-Operations Explainability**: Ensures users receive understandable and relevant information on AI/ML application results.
- **Data Recording Capabilities**: For continuous safety monitoring and incident investigation support.

3. **Human Factors for AI building block**

Provides guidance on human factors needs with AI introduction, focusing on:

- **AI Operational Explainability**: Gives end users detailed and timely information on AI/ML outputs.
- **Human-AI Teaming**: Promotes effective cooperation between users and AI systems for achieving objectives.

4. **AI Safety Risk Mitigation building block**

Acknowledges the challenge of fully opening the 'AI black box' and the need to address residual risks due to AI's inherent uncertainty.

These four building blocks are foundational in establishing trust in AI/ML applications' trustworthiness, reflecting the shift towards integrating **ethical, safety, security**, and **human-centric** considerations in the development and deployment of AI technologies in aviation.

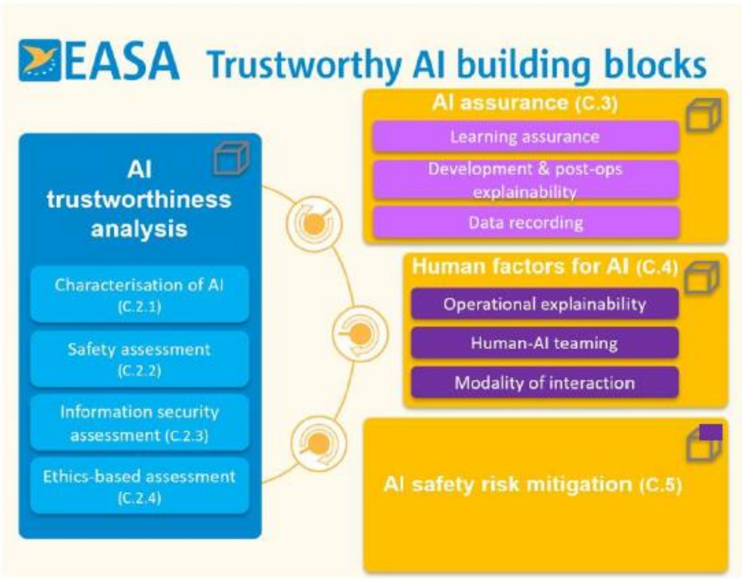


Figure: EASA AI trustworthiness building blocks[32]

Once the concept of *'trustworthy building blocks'* and their purpose has been clarified, it is important to understand two further aspects that will influence the depth of the entire assessment process: the '**criticality**' and the '**classification**' of AI applications.

The '**criticality'** of AI applications in aviation directly affects their assurance levels, which are allocated based on the safety criticality and the specific aviation domain. These levels, such as the development assurance level (DAL) for airworthiness and air operations or the software

[32] Source: EASA Concept Paper: guidance for Level 1 & 2 machine learning applications - Issue 02.

assurance level (SWAL)[33] for air traffic management/navigation services, determine the depth of assessment required.

[33] The SWAL, or Software Assurance Level, is a classification used within the aviation industry to indicate the safety criticality of software components, especially in the context of air traffic management and air navigation services (ATM/ANS). It serves as a measure of the risk associated with a software failure and dictates the rigor of the development and verification processes required to ensure the software's safety and reliability. SWAL categories range from low criticality to high criticality levels, with SWAL 1 indicating the highest level of criticality, requiring the most stringent development standards, and SWAL 4 representing a lower level of criticality, with less stringent requirements. The purpose of assigning a SWAL to a software component is to ensure that the software development process is appropriate for the level of risk involved, thereby enhancing the safety and reliability of aviation systems.

6.2.3 Classification of the AI-based system

The **classification** framework for AI applications is structured around three primary levels, reflecting industry projections on the expected functionalities of AI-based systems. This layered framework is designed to align with the incremental adoption strategy envisaged by many in the industry, beginning with AI that provides **assistance to humans** (Level 1 AI), progressing to systems that facilitate a deeper partnership between humans and AI (Level 2 AI: **human-AI teaming**), and ultimately aiming for a higher degree of machine autonomy (Level 3 AI: **advanced automation**). Additionally, the roadmap introduces a nuanced subdivision within Level 2 AI, drawing on the latest human factors guidance. This approach to classification underscores a systematic escalation in the sophistication and independence of AI applications, from supporting roles to full autonomy.

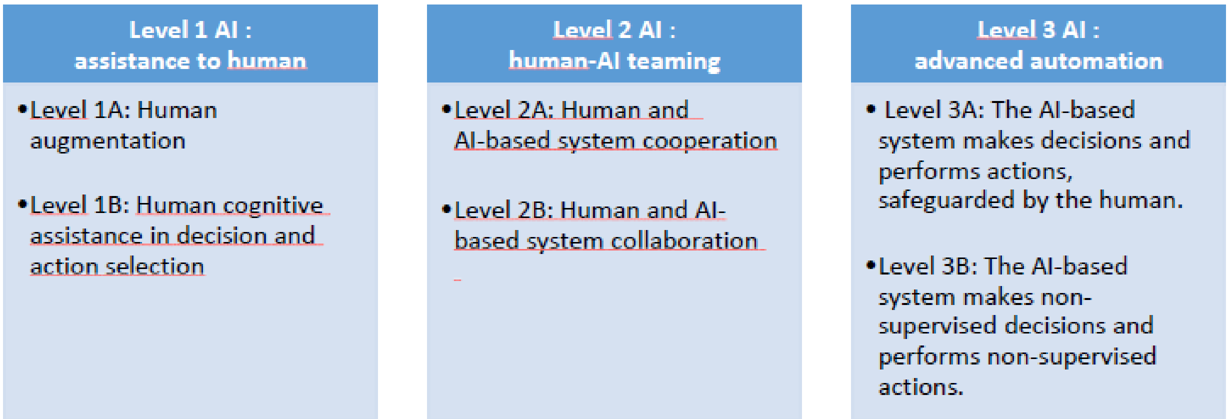


Figure: Classification of AI applications[34]

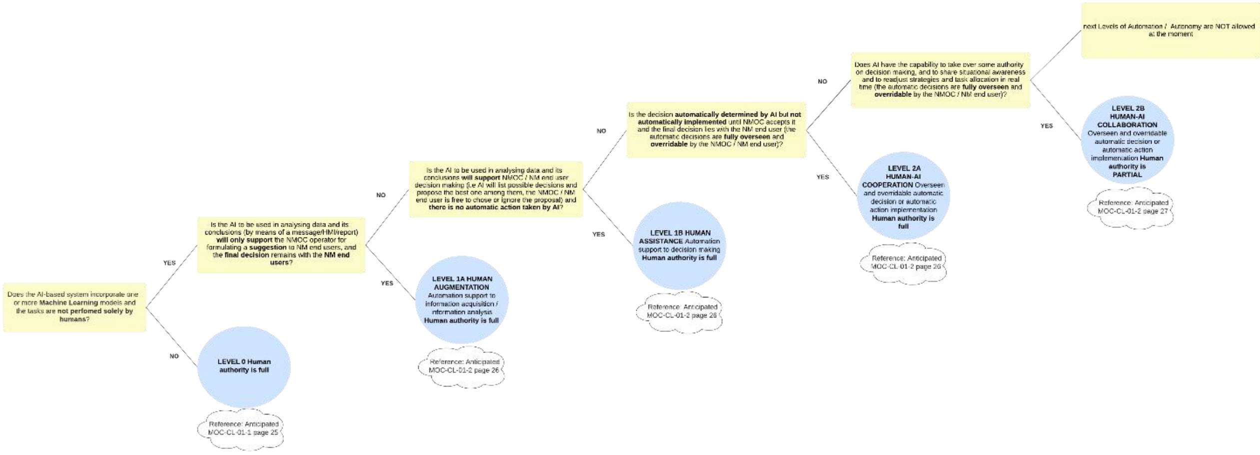
The applicability of the guidelines we are analyzing here, is limited as follows:

- covering Level 1 and Level 2 AI applications, but not covering yet Level 3 AI applications.
- covering supervised learning or unsupervised learning, but not other types of learning such as reinforcement learning.

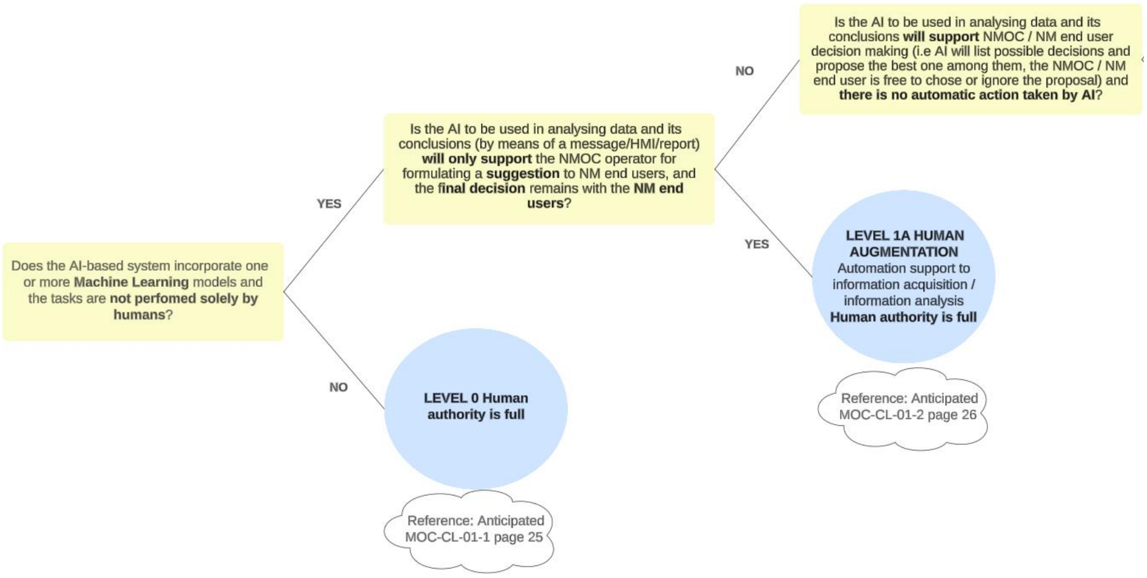
[34] Source: EASA Concept Paper: guidance for Level 1 & 2 machine learning applications - Issue 02.

- covering offline learning processes where the model is ‘frozen’ at the time of approval, but not online learning processes.

Below you can see a flowchart that I created with the support of the safety department, which helps to proceed with the classification specifically concerning applications in use in ECTL.



With a zoom on the part that is currently of most interest, as will be seen later, in reference to the use cases analyzed in this work.



6.2.4 Criticality, Trustworthiness Analysis and Proportionality

In order to identify the level of *'safety criticality'*, it was decided to determine the SWAL level that would be applied to this type of software in the case it was directly employed in ATM/ANS, precisely because this guide does not provide the possibility of not applying this parameter in order to determine the number of objectives to pursue (as will be seen later). Indeed, being not directly employed in ATM/ANS, it would not be necessary, but the lowest level has been assigned anyway, that is **SWAL 4**.

Once defined, through classification, the **'AI level'** and identified the **level of 'criticality'**, we can proceed in detail with the 'trustworthiness analysis' with all the elements useful to be able to define precisely the objectives indicated in the concept paper to pursue.

In fact, the guide provides for a very high number of objectives, as can be seen from the figure below, but, at the same time, it envisages a proportionality in the application of these objectives depending on the 'AI level' and 'criticality'[35].

[35] EASA will initially accept only applications where AI/ML constituents do not include IDAL A or B / SWAL 1 or 2 / AL 1, 2 or 3 items.

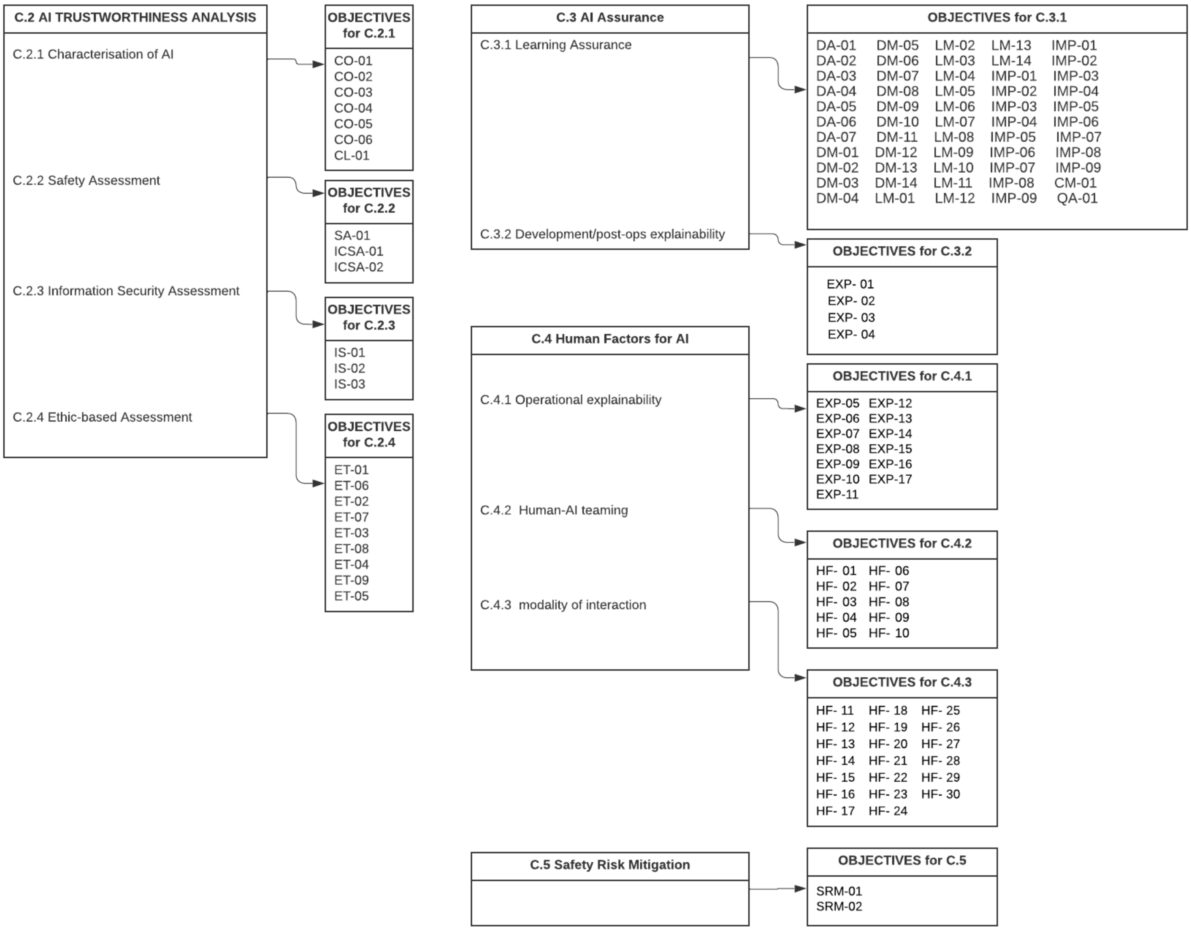


Figure: All the objectives for all the building blocks[36].

To better understand this concept, see below the diagram included in the concept paper of the objectives to be met.

[36] I created this diagram before the latest version of the concept paper was released, so many references to the objectives have changed. What matters here is to express the concept, so do not refer to the specific objective indicated.

Building block	Objectives	Assurance Level				
		AL 1 DAL A SWAL1	AL 2 DAL B -	AL 3 DAL C SWAL2	AL 4 - SWAL3	AL 5 DAL D SWAL4
Trustworthiness analysis	**ET-01:** The applicant should perform an ethics-based trustworthiness assessment for any AI-based system developed using ML techniques or incorporating ML models .	○	○	○	○	○
	ET-02: The applicant should ensure that the AI-based system bears no risk of creating over-reliance, attachment, stimulating addictive behaviour, or manipulating the end user's behaviour.	○	○	○	○	○
	ET-03: The applicant should comply with national and EU data protection regulations (e.g. GDPR), i.e. involve their Data Protection Officer (DPO), consult with their National Data Protection Authority, etc.	○	○	○	○	○
	ET-04: The applicant should ensure that the creation or reinforcement unfair bias in the AI-based system, regarding both the data sets and the trained models, is avoided, as far as such unfair bias could have a negative impact on performance and safety.	○	○	○	○	○
	ET-05: The applicant should ensure that end users are made aware of the fact that they interact with an AI-based system, and, if applicable, whether some personal data is recorded by the system .	○	○	○	○	○
	ET-06: The applicant should perform an environmental impact analysis, identifying and assessing potential negative impacts of the AI-based system on the environment and human health throughout its life cycle (development, deployment, use, end of life), and define measures to reduce or mitigate these impacts.	○	○	○	○	○
	ET-07: The applicant should identify the need for new competencies for users and end users to interact with and operate the AI-based system, and mitigate possible training gaps.	○	○	○	○	○
	ET-08: The applicant should perform an assessment of the risk of de-skilling of the users and end users and mitigate the identified risk through a training needs analysis and a consequent training activity.	○	○	○	○	○
AI assurance	**DA-01:** The applicant should describe the proposed learning assurance process, taking into account each of the steps described in Sections C.3.1.2 to C.3.1.14, as well as the interface and compatibility with development assurance processes.	○	○	○	○	○
	DA-02: Based on (sub)system requirements that have been allocated to the AI/ML constituent, the applicant should capture the following minimum requirements for the AI/ML constituent: — safety requirements allocated to the AI/ML constituent; — information security requirements allocated to the AI/ML constituent; — functional requirements allocated to the AI/ML constituent; — operational requirements allocated to the AI/ML constituent, including AI/ML constituent ODD monitoring and performance monitoring, detection of OoD input data and data-recording requirements; — other non-functional requirements allocated to the AI/ML constituent; and interface requirements.	○	○	○	○	○
AI	**DA-03:** The applicant should define the set of parameters pertaining to the AI/ML constituent ODD, and trace them	○	○	○	○	○

As can be observed, the first column represents the building block, while the second one contains the specific objectives of that building block. Then, two things can be noted: the color and the column, which is in turn divided into 5 sub-columns, related to the Assurance Level (the latter represents the level of 'criticality').

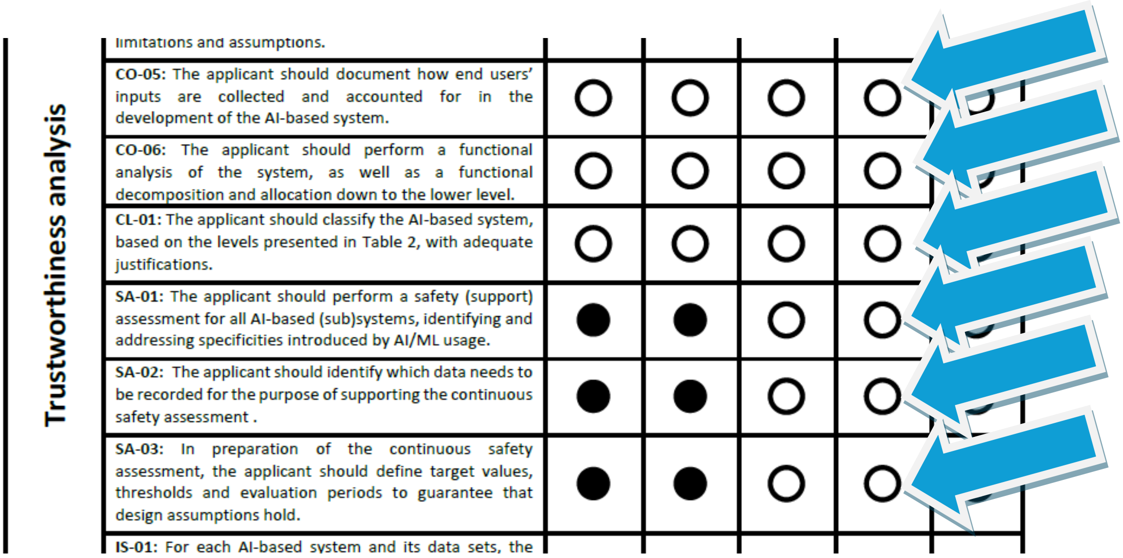
Thus, to define which objectives need to be met based on the classification made, see the diagram illustrated below.

2. Risk-based levelling of objectives

Applicability by Assurance Level	
●	The objective should be satisfied with independence.
○	The objective should be satisfied.
	The satisfaction of the objective is at the applicant's discretion.

Applicability by AI Level	
	The objective should be satisfied for AI level 1A, 1B, 2A and 2B.
	The objective should be satisfied for AI level 1B, 2A and 2B.
	The objective should be satisfied for AI level 2A and 2B.
	The objective should be satisfied for AI level 2B.

Therefore, in the case of AI level 1 and SWAL 4, only look at the last column on the right and the color white and the circle. In the table shown previously, it can be noted that in the part related to the building block 'Trustworthiness analysis', the objectives indicated with the color orange (ET stands for 'ethics-based assessment') must be met for AI levels higher than 1B. Thus, in the cited example, only the objectives that are indicated in the table with a white background, with a circle, and in the last column on the right will be pursued (as will be seen, they are, in any case, very numerous).



Trustworthiness analysis	limitations and assumptions.					
	CO-05: The applicant should document how end users' inputs are collected and accounted for in the development of the AI-based system.	○	○	○	○	[illegible]
	CO-06: The applicant should perform a functional analysis of the system, as well as a functional decomposition and allocation down to the lower level.	○	○	○	○	[illegible]
	CL-01: The applicant should classify the AI-based system, based on the levels presented in Table 2, with adequate justifications.	○	○	○	○	[illegible]
	SA-01: The applicant should perform a safety (support) assessment for all AI-based (sub)systems, identifying and addressing specificities introduced by AI/ML usage.	●	●	○	○	[illegible]
	SA-02: The applicant should identify which data needs to be recorded for the purpose of supporting the continuous safety assessment .	●	●	○	○	[illegible]
	SA-03: In preparation of the continuous safety assessment, the applicant should define target values, thresholds and evaluation periods to guarantee that design assumptions hold.	●	●	○	○	[illegible]
	IS-01: For each AI-based system and its data sets, the					

Let's then proceed with the trustworthiness analysis for ISOBAR.

As mentioned earlier, the application provides, using Machine Learning models:

- **forecasts** that can assist air traffic management (ATM) systems and operators in making informed decisions about capacity regulation and allocation.
- a regulation prediction analysis tool (**RPA**) that allows users to monitor and evaluate the performance of the ML models and the quality of the weather data.

6.2.5 First building block: AI Trustworthiness Analysis

The first building block, ‘trustworthy analyses’ as mentioned earlier, contains one characterisation and three assessments, so, let’s start with the characterisation, indicating only the objectives that we considered relevant, based on the considerations made before.

6.2.5.1 Characterisation of the AI Application.

The portrayal of the AI application offers a detailed and thorough explanation of the ISOBAR project (the AI-based system), outlining its goals, extent, and limits. It further establishes the duties and obligations of the stakeholders participating in the creation, implementation, and management of the AI system.

In order to characterising the AI-based system, the first step consists in identifying the **list of end users** [CO-01[37]] intended to interact with the AI-based system, the associated **high-level tasks** [CO-02] and the **AI-based system definition** [CO-03]:

1) [CO-01] **list of end users** that are intended to interact with the AI-based system, together with their **roles**, their **responsibilities** and expected **expertise:**
   a. **Deputy Operation Manager (DOM)**
      i. Roles: managing, streamlining and improving air traffic operations in Europe.
      ii. Responsibilities: In addition to the usual responsibilities expected for the DOM, they will have the responsibility of making informed decisions about capacity regulation and allocation. There won't be any level of 'teaming' with the AI-based system, but just an increase in information sufficient to improve the awareness of the situation and the possible actions to be taken.
      iii. Expected expertise: no specific additional expertise is required beyond what is already possessed by the DOM. A brief training is recommended to become familiar with the new tasks.
2) [CO-02] For each end user, identify which **goals** and associated **high-level tasks** are intended to be performed in interaction with the AI-based system:

[37] These acronyms will be used from now on in the context of assessments in this chapter to represent the objectives that need to be pursued. CO stands for ConOps; SA for Safety Assessment; IS for Information Security, etc.

a. **DOMs**:

i. **Goals**:

1. Make informed decisions (suggestions/advice to the FMPs?) about capacity regulation and allocation due to the weather.

   a. Associated **high level tasks** in interaction with ISOBAR tools:

      i. To have access and analyze the predictions of the capacity impact of weather events on traffic volumes (TVs) or airspace sectors (AS) in the European airspace, based on two different data sources: EUMETNET and ICONv, provided by the ISOBAR's tools.

      ii. To have access and analyze the regulation prediction analysis (RPA) tool that allows users to monitor and evaluate the performance of the ML models and the quality of the weather data.

3) [CO-03] **AI-based system definition,** determine the AI-based system taking into account domain-specific definitions of 'system'[38].

   The AI-based system of the ISOBAR project consists of two sets of two ML models[39]:

   1. The first set of these models is trained and inferred using EUMETNET data.
   2. The second set of these models is trained on ICONv data.

   These models serve two main functions:

   - classification (determining whether a regulation is necessary), and

[38] for the ATM/ANS domain (ATS and non-ATS), Regulation (EU) 2017/373 defines a functional system as 'a combination of procedures, human resources and equipment, including hardware and software, organised to perform a function within the context of ATM/ANS and other ATM network functions.

[39] The combination of these models allows for more accurate predictions based on their respective strengths in handling different types of weather-related events.

- regression (predicting the capacity reduction).

All these models produce their respective forecasts which are made available to the operators indicated as support for making informed decisions regarding capacity regulation and allocation, but the second set, in addition to providing real-time predictions, serves for monitoring purposes.

The AI-based system is integrated with iFlow[40], which will query the models via API to obtain their respective forecasts.

[40] iFLOW, a new tool, not yet operational, developed by Eurocontrol, that will help the ECTL Network Manager better manage operations by providing greater digital coordination, improving awareness and decision-making, while sharing between ATM stakeholders real-time information and coordinate their responses to disruptions and delays.

6.2.5.1.1 Concept of operations for the AI application (CO-04)

To support compliance with the AI trustworthiness guidelines' objectives, a detailed ConOps describing precisely how the system will be operated is expected to be established, including the task allocation pattern between the end users and the AI-based system with a focus on the definition of the OD and on the capture of specific operational limitations and assumptions.

The concept of operations for the ISOBAR project describes the operational environment, the operational scenarios, and the operational procedures for the AI application. It also defines the roles and responsibilities of the end-users, referring back to what was already mentioned in the previous objective CO-02, and the system operators.

Operational environment

The operational environment of the ISOBAR project is the European airspace, where weather events can cause disruptions and delays in air traffic flow and capacity. The ISOBAR project is integrated into the existing Cross Border Weather procedure, which aims to reduce weather-related delays in air traffic flow and capacity management (ATFCM).

The Cross Border Weather procedure encompasses these phases:

- Weather information sharing: Air traffic control centres exchange current weather details with adjacent nations, enabling them to make informed decisions about ATM and minimize delays caused by adverse weather conditions.
- Coordinated decision-making: Controllers from different countries work together to coordinate their responses (regulation) to weather-related disruptions.

Operational scenarios

There are two operational scenarios of the ISOBAR project, depending on how and from what the prediction is received:

1. The new weather prediction for a TV is received **from iFlow** via API.
2. The new weather prediction for a TV is received **from EUMETMET** or **ICONv**.

Scenario 1 breakdown for enhanced clarity:

- Initial weather prediction reception:
  - Origin: Received via API from iFlow.
  - Content: New weather forecast for a Target Volume (TV).
- Forecast generation by ISOBAR models:
  - Process: Both ISOBAR model sets are activated to produce their respective forecasts regarding capacity impact and regulation decisions.
  - Display: Forecasts are presented on iFlow, replacing the traditional static maps previously made in PowerPoint format.
- Monitoring and comparison by Deputy Operation Manager (DOM):
  - Location: Network Manager Operational Centre (NMOC).
  - Responsibilities: DOM observes the forecasts and benchmarks them against EUMETNET data for accuracy.
  - Tools: he can also utilize the RPA tool for evaluating model performance and weather data quality.
- Communication to Flow Management Position (FMP):
  - Receiver: FMP at the local control centre.
  - Information Use: FMP leverages the forecasts for informed decision-making on capacity regulation and allocation.

Scenario 2 breakdown for enhanced clarity:

- Receipt of new weather prediction:
  - Source: Received from EUMETNET or ICONv.
  - Content: New weather forecast for a Target Volume (TV).

- Activation of the 2nd ISOBAR Model set:
    - Action: Triggered to update forecasts incorporating the new data.
- Storage and monitoring of updated predictions:
    - Storage: Updated forecasts are archived for monitoring.
    - Access: Predictions can be reviewed using the RPA tool.

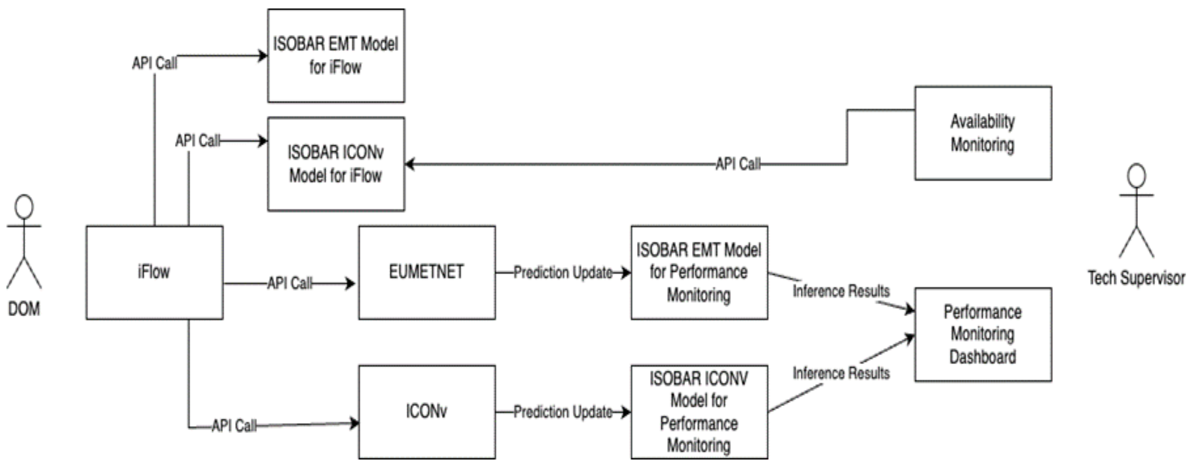
Operational procedures

The ISOBAR project operational procedures encapsulate a streamlined approach to integrating and utilizing weather predictions for air traffic management.

This schema illustrates the key steps involved, from initial weather forecast reception to decision-making and model updates, ensuring clarity and efficiency in operations.

- **Step 1: Weather prediction reception and analysis.**
    - Action: The Deputy Operation Manager (DOM) receives a new weather prediction for a Target Volume (TV) via iFlow.
    - Process:
        - The DOM reviews predictions from both ISOBAR model sets.
        - Compares them with EUMETNET data.
        - Utilizes the RPA tool for model performance analysis and weather data quality assessment.
    - Communication: The DOM forwards the predictions to the Flow Management Position (FMP) at the local control centre.
- **Step 2: Decision-making and implementation.**
    - Reception: The FMP receives predictions from the DOM.
    - Decision-Making Process, the FMP:

        - Uses the predictions for informed capacity regulation and allocation decisions.
        - May consult EUMETNET data and the RPA tool for additional insights.
    - Implementation: the FMP executes the regulation decision and informs relevant stakeholders.
- **Step 3: Model update and monitoring.**
    - Data Reception: The second set of ISOBAR models gets a new weather prediction for a TV from EUMETNET or ICONv.
    - Update Process:
        - Updates forecasts based on new data.
        - Stores them for monitoring.
    - Monitoring and Feedback:
        - The RPA tool accesses updated predictions, and
        - provides feedback to users.



This flowchart illustrates the operational procedures for weather prediction and performance monitoring within the ISOBAR project.

Here's a step-by-step breakdown of the flowchart:

- **DOM interaction with iFlow**:
    - The Deputy Operation Manager (DOM) uses iFlow to initiate two API calls:
        - One to the ISOBAR EMT Model for iFlow.
        - Another to the ISOBAR ICONv Model for iFlow.
- **Data gathering from weather services**:
    - iFlow makes further API calls to EUMETNET and ICONv to gather weather predictions.
- **Prediction updates**:
    - The ISOBAR EMT and ICONv Models for iFlow update their predictions based on the latest data received from EUMETNET and ICONv.
- **Performance monitoring**:
    - Both models also send data to their respective modules for Performance Monitoring.
    - The results of these inferences (or predictions) are then displayed on a Performance Monitoring Dashboard.
- **Tech supervisor role**:
    - A technical supervisor oversees Availability Monitoring.
    - The system operators are also involved in reviewing the inference results displayed on the Performance Monitoring Dashboard.

These operational procedures are designed to ensure that the most current and accurate weather data is used to inform traffic flow and capacity management decisions, while also providing a system for continuous performance evaluation and system availability.

Integrating what was stated earlier, in objectives CO-01 and CO-02, which articulated the roles, responsibilities, and expected expertise, as well as high-level tasks, we must also include 'system operators', defining their role and responsibilities within this context.

System operators, comprising data scientists and engineers, hold pivotal responsibilities. They are the backbone of the RPA tool's lifecycle, overseeing its creation, implementation, and consistent upkeep. These operators handle data from EUMETNET and ICONv, leveraging it alongside end-user feedback to enhance the ISOBAR models continually.

System operators, through iterative processes of training, testing, and refining, ensure the ISOBAR models are progressively improved to meet user requirements.

Moreover, they undertake the critical task of monitoring, scrutinizing the performance, and ensuring the quality of both the ISOBAR models and the RPA tool remains at its peak. This careful supervision is fundamental to maintaining the integrity and reliability of the ISOBAR project's outputs.

To satisfy the objective **CO-05** of the Concept paper, which requires to <u>document how end users' inputs are collected and accounted for in the development of the AI-based system</u>, it has been thoroughly documented which subjects are involved (stakeholders, persons), the impact on them and on the project developers, the key elements for them and the modality on how they can contribute. For confidentiality reasons, I don't report here these elements.

6.2.5.1.2 Functional analysis of the AI-based system (CO-06].

[CO-06] The functional analysis of the AI-based system provides a detailed description of the functions and features of **ISOBAR models** and the **RPA tool**. It also defines the inputs, outputs, and interfaces of the AI-based system.

Below are the **functions** and **features** of:

- **ISOBAR models**.
    - **Clipper functionality**: The ISOBAR models employ a clipping tool to identify Target Volumes (TVs) affected by weather conditions. This tool processes geographic information from weather sources (EUMETNET and ICONv), along with airspace data from the NM Environment database, as input, to output TVs impacted by predicted weather conditions.
    - **Classification capability**: Utilizing weather forecasts and capacity data, the ISOBAR models are equipped to classify a TV as either subject to regulation or not. This classification assists end-users in determining the necessity of imposing regulations on a specific TV.
    - **Regression analysis**: The models are designed to predict the extent of capacity reduction for a regulated TV, drawing from weather forecasts and capacity information. This regression analysis aids end-users in establishing the most effective regulation level for each TV.
    - **Monitoring and updating**: With the capability to refresh forecasts based on the latest weather data from EUMETNET or ICONv, the ISOBAR models ensure their predictions remain current and precise. This monitoring feature supports system operators in maintaining the models' accuracy and reliability.
- **RPA tool**:
    - **Performance evaluation function**: This feature of the RPA tool allows for a comprehensive assessment of the ISOBAR models' performance using a variety

of metrics, including accuracy, precision, recall, root mean squared error, and mean absolute error. Such an evaluation aids both end-users and system operators in judging the models' efficiency and dependability.

  - **Data quality assessment**: The tool can evaluate the quality of weather data obtained from EUMETNET and ICONv, applying criteria like completeness, timeliness, consistency, and validity. This assessment function is crucial for end-users and system operators to pinpoint and address any discrepancies or inaccuracies in the weather data.
  - **Visualization features**: The RPA tool presents an ability to graphically display the predictions made by the ISOBAR models, alongside weather data from EUMETNET and ICONv, within iFlow. It also visualizes performance and quality metrics through an interactive, user-friendly dashboard. This visualization aids in making the data and forecasts more accessible and understandable to both end-users and system operators, facilitating an intuitive and in-depth analysis.

The **inputs**, **outputs**, and **interfaces** of the **AI-based system** are detailed as follows:

- **Inputs**:
  - **Weather predictions**: Sourced from EUMETNET and ICONv, these predictions are structured as forecast events with areas delineated by polygons.
  - **Capacity data**: Derived from body and perm tables, presented as capacity schedules where each Target Volume (TV)'s capacity is specified in time slots.
  - **Reception method**: Inputs are acquired via API connections from iFlow, EUMETNET, and ICONv.
- **Outputs**:
  - **Capacity impact and Regulation decisions**: For each TV, predictions include the impact on capacity and the decisions on regulation, alongside metrics for model and data quality and performance.

- **Format**: Predictions and metrics are formatted as JSON objects (detailing TV ID, time slot, weather features, capacity, regulation decision, and capacity reduction) and as numerical values, charts, and tables.
- **Display Platforms**: Outputs are visualized on iFlow and the RPA tool.

- **Interfaces**:
  - **iFlow**:
    - **Function**: Acts as the primary interface, facilitating communication between ISOBAR models, end-users, and ATM systems.
    - **Data Reception**: Receives weather predictions and capacity data from EUMETNET and ICONv and receives the predictions from the ISOBAR models.
    - **Data Transmission**: Forwards this data to ISOBAR models via API.
    - **Prediction Display**: Displays ISOBAR models' predictions to end-users and ATM systems.
  - **EUMETNET and ICONv**:
    - **Role:** Serve as data providers, supplying weather predictions.
    - **Interaction with iFlow and ISOBAR models**: Provide weather predictions to iFlow and to ISOBAR models.
    - **Feedback loop:** Receive updated predictions from the second set of ISOBAR models for monitoring and analysis purposes.

6.2.5.1.3 Classification of the AI-based system (CL-01].

As previously discussed in the dedicated chapter the Concept Paper offers invaluable guidance for accurately classifying AI-based systems based on the level of AI autonomy in relation to the human user into one of the three broad categories identified: Assistance to Human; Human-AI Teaming; and Advanced Automation. It should be noted, however, that the current state of this guide does not cover the third area, Advanced Automation.

The classification process for ISOBAR was relatively straightforward, as the distinctions between Level 1 and Level 2 were quite clear.

The primary difference between Level 1 and Level 2 AI applications centers on the implementation of decisions. In Level 1 AI, the decisions are made by the end user, who is supported by the AI-based system; all subsequent actions are also carried out by the end user.

Conversely, Level 2 AI-based systems are capable of automatically selecting and executing actions. However, the end user retains complete oversight and the ability to override the actions of the AI-based system at any time. At this level, decisions can be made either by the end user or automatically by the AI-based system, but always under the user's direction and supervision.

Below is the table provided by EASA for classifying AI-based systems, which is based on the function allocated to the system to contribute to the high-level task and on the level of authority of the final human user.

AI level	Function allocated to the system to contribute to the high-level task	Authority of the end user
Level 1A Human augmentation	Automation support to information acquisition	Full
	Automation support to information analysis	Full
Level 1B Human assistance	Automation support to decision-making	Full
Level 2A Human-AI cooperation	Directed decision and automatic action implementation	Full
Level 2B Human-AI collaboration	Supervised automatic decision and action implementation	Partial
Level 3A Safeguarded advanced automation	Safeguarded automatic decision and action implementation	Limited, upon alerting
Level 3B Non-supervised advanced automation	Non-supervised automatic decision and action implementation	Not applicable

Table 2 — EASA AI levels

As previously highlighted, it is evident that in all tasks, ISOBAR does not possess any level of authority since it merely provides supplementary information to that already held by various users involved in the process. The authority always rests solely with the final user (the FMP). It remains to be explored, once it has been confirmed that the classification level is 'L1', what exactly distinguishes Level 1A from Level 1B.

It is well explained in the Concept Paper that the distinction between Level 1A and Level 1B centers on the extent of decision-making support provided by the AI/ML system.

Level 1A involves enhancing the information available to the end user. This can range from organizing incoming data according to certain criteria to predicting outcomes (either through interpolation or extrapolation) or integrating information to augment the end user's perception and cognitive processing.

Level 1B, on the other hand, specifically supports the decision-making process by helping the end user select from among several possible actions. While the AI-based system at this level may present multiple alternatives, it might sometimes offer only a subset of all possible options. This would still fall under Level 1B. Occasionally, the choice might be as simple as deciding between two alternatives, such as approving a radio-frequency suggestion or modifying an entry proposed by the AI.

It's important to note that in both Level 1A and Level 1B, the AI-based system **does not make decisions**; it merely aids in decision-making. The ultimate decision and action implementation are always carried out by the end user, not the AI system. This framework ensures that while the AI provides significant support, the responsibility and authority remain with the human operator.

As previously discussed, all tasks within ISOBAR are solely aimed at enhancing the level of information available to the end user and providing forecasts that support what the end user already possesses. The integration of this information is designed exclusively to improve the end user's perception of the situation and assist in the cognitive phase, without in any way limiting the number of possible actions to be taken, which would be characteristic of Level 1B. For these reasons, the classification assigned to ISOBAR is **Level 1A**.

6.2.5.2 Safety assessment of the ML Application.

Continuing with the app evaluation within the first building block, the 'trustworthy analyses.', the second point is represented by the safety assessment.

The aim of a safety assessment is to verify that a system meets established safety standards. There's an expected inverse correlation between the likelihood of a failure's occurrence and its impact's severity. In traditional non-AI aviation systems, safety methods might differ by application area, typically assuming hardware as the primary failure source, without quantifying software reliability directly. For example, in airborne systems, using proven development assurance processes is believed to significantly reduce software error-related failures, providing a high level of confidence. Development mistakes, seen as potential common failure causes, are mitigated through system design and analyzed alongside other errors using methods like common mode analysis. Risk assessments often confine digital components' impact to input parameter reliability and the hardware's execution reliability.

Machine learning (ML) introduces new challenges in predictability and uncertainty due to its statistical nature and complex modeling.

The Concept paper aims to help demonstrate that AI/ML-integrated systems are at least as safe as their conventional counterparts, ensuring AI technologies do not increase risk beyond what's acceptable for traditional systems. It strives to align with current aviation safety assessment practices to minimize procedural disruptions. AI/ML applications, capable of estimating output uncertainties, offer insights for safety monitoring and continuous safety evaluation, reinforcing the safety case.

A satisfactory safety standard must be upheld across the entire lifecycle of a product, facilitated by:

- An **initial safety assessment** during the design phase. This evaluation assesses the impact of integrating AI/ML components on system reliability and includes specific design considerations for incorporating AI. This is then followed by

- A **continuous safety assessment**, implementing a data-driven AI safety risk analysis utilizing operational data and recorded incidents. This continuous evaluation of service events might utilize established processes from the domains outlined in these guidelines, though these processes will require modifications to accommodate AI technology. It's acknowledged that the extent and nature of the activities required for EASA certification can greatly differ depending on the specific domain.

As already mentioned earlier, in the case of the ISOBAR project, it is necessary to analyze which type of safety assessment is envisaged by the concept paper, knowing that safety assessment should be understood as '**safety assessment of the functional system'** when it applies to **ATS providers** in the ATM/ANS domain, and should be understood as a system safety assessment in the airworthiness domain and '**safety support assessment of the functional system'** applies to **non-ATS providers.**

So, it is clear that ISOBAR project follows under the 'non-ATS providers' and needs to be assessed with the 'safety support assessment' and there are two phases: the **initial safety support analyses** and the **continuous safety assessment**[41]

The ISOBAR project is anchored by **specific safety objectives** and mandates as follows:

Safety objective 1: Deliver dependable predictions.

Aim to ensure the predictions regarding capacity impact and regulatory decisions for each TV are both accurate and reliable, drawing from weather forecasts and capacity figures.

- Safety requirement 1.1: The models must reach predefined benchmarks for accuracy, precision, recall, and minimal errors, adhering to established performance indicators.

[41] For ATS and non-ATS providers, the notion of 'continuous safety assessment' should be understood as the 'Safety performance monitoring and measurement' for ATS providers, or simply the **'Performance monitoring and measurement' for non-ATS providers**.

- Safety requirement 1.2: Ongoing updates and retraining of the ISOBAR models are essential, incorporating fresh data and insights.
- Safety requirement 1.3: Users should leverage EUMETNET and ICONv data, alongside the RPA tool, for supplementary checks and confirmation before basing decisions on model forecasts.

Safety objective 2: Leverage quality weather data

Utilize high-quality weather information from EUMETNET and ICONv for generating forecasts.

- Safety requirement 2.1: Weather inputs from these sources must fulfill specified quality standards, outlined by set criteria.
- Safety requirement 2.2: Prior validation and verification of weather data are required before their application in models.
- Safety requirement 2.3: The RPA tool should facilitate an evaluation of data integrity, offering user feedback on data quality.

Safety objective 3: Uphold model safety and reliability.

Maintain the integrity and trustworthiness of ISOBAR model predictions.

- Safety requirement 3.1: Models are subject to thorough validation and verification before their initial use and following any updates or training adjustments.
- Safety requirement 3.3: Models need to offer estimates of uncertainty and confidence in their predictions, providing insight into their reasoning and foundational assumptions.
- Safety requirement 3.4: Continuous monitoring and auditing of the models are critical to identifying any discrepancies, mistakes, or biases in their outputs or operations.

Continuous safety assessment.

Continuous safety assessment is checking how well the ISOBAR models meet the safety goals, reported before, using the data from the system's inputs and outputs.

Depending on the aviation domains, different approaches exist to ensure that systems are in a condition for safe operation, at any time in their operating life.

Within the **ATM/ANS domain**, specific requirements have been established to ensure safety performance monitoring and measurement.

Air Traffic Services (ATS) providers are mandated to include in their safety evaluations the definition of monitoring standards necessary to affirm that the service offered by any updated functional systems adheres to established safety benchmarks (ATS.OR.205(b)(6)).

Similarly, also**, non-ATS providers** are required to ensure their **safety support assessments** comprises specification of the monitoring criteria necessary to demonstrate that the service delivered by the changed functional system will continue to behave only as specified in the specified context (ATM/ANS.OR.C.005(b)(2)).

These instances from airworthiness and ATM/ANS areas indicate that certain aviation regulations already foresee the necessity of documenting specific information to preserve safety levels throughout an aviation product's lifespan.

To assure the safe usage of AI-based systems throughout their operational life, objectives aimed at addressing the unique aspects of AI/ML are identified in the Concept paper, thereby augmenting the existing regulatory framework, as follows:

Objective SA-02: The applicant should identify which data needs to be recorded for the purpose of supporting the continuous safety assessment.

To support the continuous safety assessment, the system data needs to be recorded and stored securely.

The data includes:

- The **inputs** from EUROMETNET and ICONv, which are the forecasts of weather parameters such as wind speed, temperature, and precipitation.
- The **inputs** from iFLOW, which are the calls for emergency response services in case of severe weather events.
- The **outputs** from the ISOBAR models, which are the predictions of the risk of flooding and the optimal allocation of resources for mitigation and prevention.

[SA-03] The performance indicators of the ISOBAR models, which measure how accurate and reliable they are compared to the ground truth and the user expectations. The performance indicators are calculated on a daily basis and compared to the target values and thresholds that were determined in the summer 2024 evaluation.

- The uncertainty estimates and confidence levels of the ISOBAR models, which indicate how confident they are in their predictions and how much variability they expect in the outcomes.
- The feature importance of the ISOBAR models, which provide the rationale and the assumptions behind their predictions and suggest possible alternatives or actions for the users.
- The feedback and correction mechanisms of the ISOBAR models, which allow the users to provide comments, suggestions, or complaints about the system and to modify or override the predictions based on their own judgment or experience.

The recorded data serves a dual purpose: it not only facilitates the regular monitoring and auditing of the ISOBAR models to identify any deviations, inaccuracies, or prejudices in their function and conduct but also enhances and refines these models by incorporating feedback from users and operational insights. Through continuous safety evaluation, this approach guarantees that the ISOBAR models adhere to established safety standards and deliver a service of superior quality and reliability to both users and stakeholders.

6.2.5.3 Information Security Assessment

As we delve into the realm of ISOBAR, an AI system pivotal for air traffic management, it becomes crucial to underscore the significance of safeguarding its informational assets. ISOBAR, with its reliance on external datasets to forecast air traffic conditions, sits at the nexus of technology and security.

Here's a closer look at the potential **cyber security risks** [IS-01[42]] that ISOBAR might encounter:

- **Data Integrity Concerns**: The integrity of weather data sourced from EUMETNET and ICONv is paramount. However, this data could fall prey to manipulation by nefarious entities, skewing predictions and potentially compromising the safety and operational efficiency of air traffic management.
- **Unauthorized System Access**: The ISOBAR API stands as a critical juncture that, if breached, could allow adversaries to either siphon off sensitive predictive data or inject false information, leading to potential disruptions or inefficiencies in air traffic coordination.
- **Service Disruption Risks**: A concerted cyberattack could aim to incapacitate the ISOBAR system, either by overwhelming its computational resources or outright disabling its capabilities. Such an event would severely hamper the system's ability to deliver timely predictions, impacting decision-making processes and operational awareness.
- **Software Integrity and Confidentiality**: The threat of malware or viruses introduces a spectrum of risks, from corrupting ISOBAR's operational software to compromising data integrity and privacy. These malicious interventions could lead to inaccurate predictions or unauthorized data exposure.

[42] Objective IS-01: For each AI-based (sub)system and its data sets, the applicant should identify those information security risks with an impact on safety, identifying and addressing specific threats introduced by AI/ML usage.

In response to these cyber security risks, it is imperative that ISOBAR adopts a robust security posture through several **mitigative strategies** [IS-02; IS-03[43]]:

- **Ensuring Data Authenticity**: By implementing rigorous data validation protocols, ISOBAR can authenticate the veracity of information received from EUMETNET and ICONv, thereby averting the risks associated with data tampering.
- **Fortifying Access Controls**: ISOBAR's approach to safeguarding its API involves stringent access controls, including secure authentication mechanisms to deter unauthorized access and comprehensive logging to monitor and audit system interactions.
- **Promoting System Resilience**: Emphasizing the importance of human oversight, ISOBAR ensures that operators are proficient in manual operational protocols, thereby mitigating reliance on the system and maintaining operational continuity in the face of technological disruptions.
- **Software Integrity Assurance**: Notably, the architectural design of ISOBAR inherently minimizes the risks associated with malware and viruses, highlighting the system's inherent resilience to such threats.

Protecting ISOBAR from cyber threats transcends the mere deployment of technical safeguards; it involves cultivating a culture of vigilance and resilience, adept at navigating the complexities of the cyber landscape. This ethos of security is integrated into the development lifecycle of the app, with the effectiveness of security controls being rigorously evaluated at each verification stage, tailored to address the nuanced spectrum of anticipated threats. This holistic approach ensures that as the app evolves, so too does its capacity to counteract emerging cyber risks, ensuring the system's integrity and the safety of air traffic operations remain uncompromised.

[43] Objective IS-02: The applicant should document a mitigation approach to address the identified AI/ML-specific information security risk. Objective IS-03: The applicant should validate and verify the effectiveness of the security controls introduced to mitigate the identified AI/ML-specific information security risks to an acceptable level.

6.2.5.5 Ethics-based Assessment

Referencing the EU Commission's AI High-Level Expert Group (HLEG) from 2019, it is articulated that the trustworthiness of AI systems hinges on three fundamental expectations: adherence to ethical principles, technical robustness, and respect for fundamental rights. This framework is grounded in four ethical imperatives: respect for human autonomy, prevention of harm, fairness, and explainability.

The HLEG expanded these concepts into a detailed framework, introducing **seven key areas**[44], each with its specific focuses, to guide the ethical development and deployment of AI:

- **Human Agency and Oversight**: Ensuring AI supports human autonomy while providing necessary oversight mechanisms.
- **Technical Robustness and Safety**: Highlighting the importance of reliability and safety in AI operations.
- **Privacy and Data Governance**: Emphasizing the protection and ethical handling of data.
- **Transparency**: Advocating for openness in AI functionalities and decisions.
- **Diversity, Non-discrimination, and Fairness**: Promoting inclusivity and equity in AI outcomes.
- **Societal and Environmental Well-being**: Recognizing the impact of AI on broader societal and environmental contexts.
- **Accountability**: Ensuring mechanisms are in place for responsibility and response in AI actions.

o assist in the self-evaluation and guidance for developers, the HLEG, in 2020, introduced the Assessment List for Trustworthy AI (ALTAI), a tool comprising numerous questions and explanations designed to navigate the ethical deployment of AI technologies. This initiative

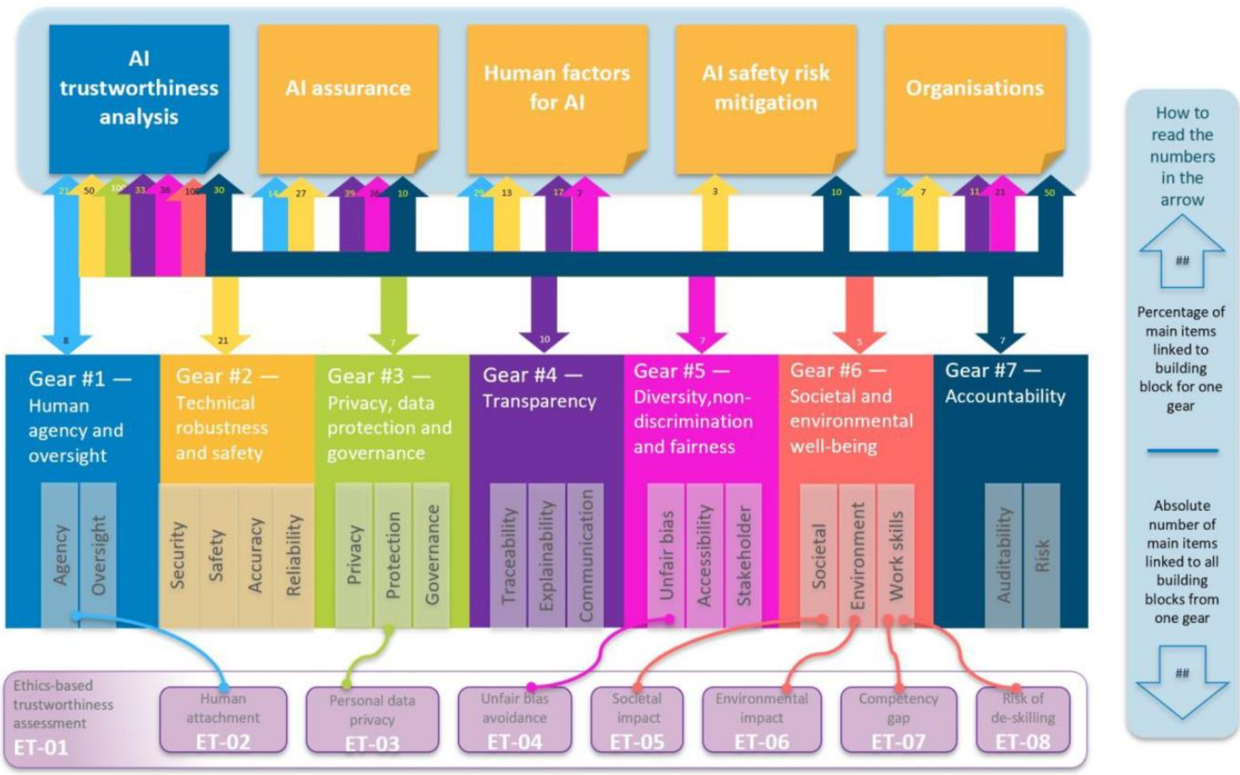
[44] Set by the so-called ALTAI (Assessment List for Trustworthy AI).

underscores the commitment to integrating ethical considerations throughout the AI development lifecycle, ensuring that AI systems are not only technologically advanced but also ethically grounded and socially responsible.

Aware that the 2019/2020 Commission's approach serves as a non-binding recommendation, the EASA guidelines take inspiration from this framework, aiming to refine and adapt the HLEG's concepts specifically for the aviation sector and its stakeholders within the EASA's scope.

When performing the ethics-based trustworthiness assessment [ET-01] it is suggested to take into account the seven points of the Assessment List for Trustworthy AI, showed before.

The following figure, from the Concept Paper, provides an overview of the distribution of the '*ethical gears'* over the '*AI trustworthiness building blocks'*:



And the following figure, also from the Concept Paper, provides an overview of the ALTAI items **requiring additional oversight from authorities other than EASA**:

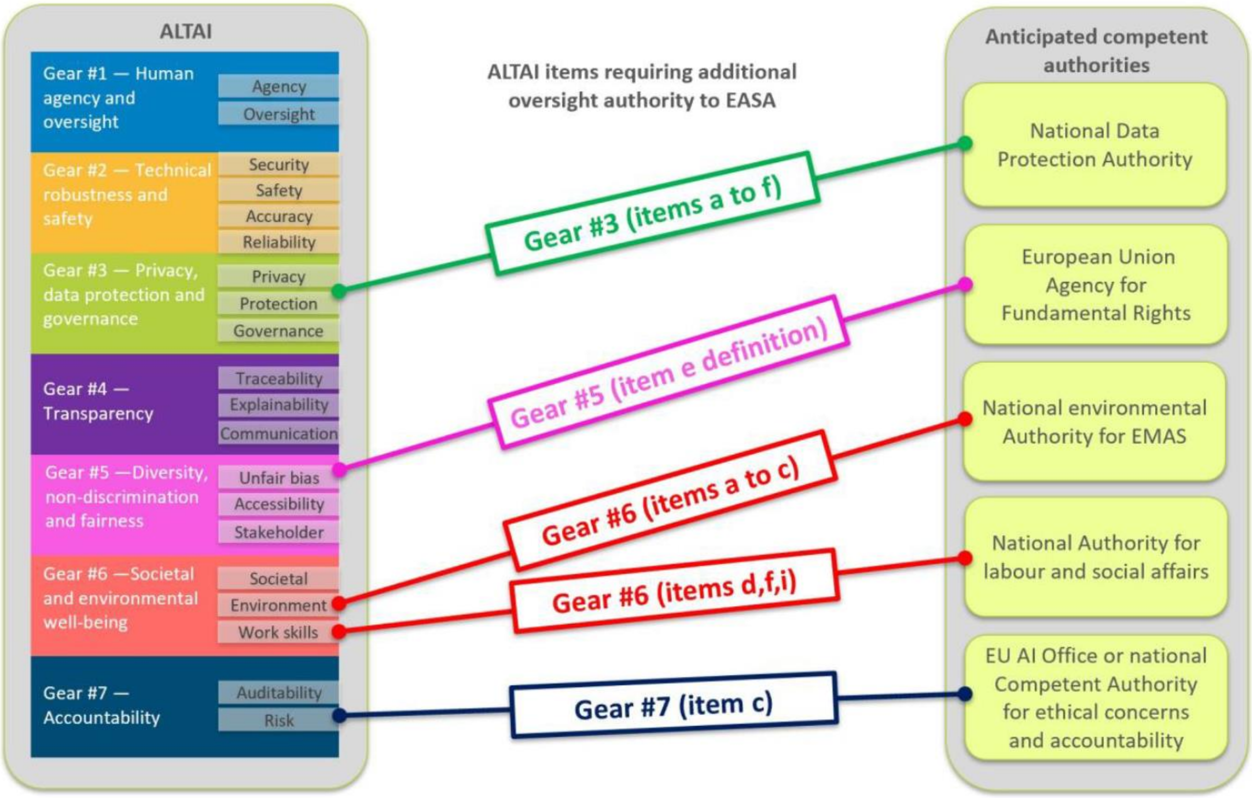


Figure 10 — Anticipation of mapping of outstanding items to other competent authorities

ISOBAR, enhancing air traffic management with weather forecasts and alerts, leverages data from EUMETNET and ICONv, processing this information to provide outputs via an API. The system is specifically crafted for <u>use in professional settings during operational shifts</u>, ensuring it does not engage with or generate personal or person-specific data. Importantly, ISOBAR's decision-making process is structured to <u>avoid the introduction of unfair bias</u>, as **it neither makes autonomous decisions nor profiles user behavior** [ET-01; ET-02; ET-03; ET-04[45]].

[45] Objective ET-01: The applicant should perform an ethics-based trustworthiness assessment for any AI-based system developed using ML techniques or incorporating ML models. Objective ET-02: The applicant should ensure that the AI-based system bears no risk of creating overreliance, attachment, stimulating addictive behavior, or manipulating the end user's behavior. Objective ET-03: The applicant should comply with national and EU data protection regulations (e.g. GDPR), i.e. involve their Data Protection Officer, consult with their National Data Protection Authority, etc. Objective ET-04: The applicant should ensure that the creation or reinforcement of unfair bias in the AI-based system, regarding both the data sets and the trained models, is avoided, as far as such unfair bias could have a negative impact on performance and safety.

Below is a succinct overview of these objectives ET-05; ET-06; ET-07; ET-08[46].

- **Operator Training and Communication. Objective ET-05**: Operators have been extensively briefed on ISOBAR's AI capabilities, particularly its role within RPA. This effort has been bolstered through repeated communications to ensure operators have a solid grasp of the system's functions. The aim is to make end users aware they are interacting with an AI system, potentially highlighting the handling of personal data.
- **Environmental Impact and Sustainability. Objective ET-06**: An in-depth environmental impact assessment has highlighted ISOBAR's role in significantly lowering aviation emissions and fuel consumption. By providing accurate weather forecasts and alerts, ISOBAR enhances the stability of air traffic networks and facilitates the optimization of flight paths under adverse weather conditions. This not only contributes to reduced fuel consumption and greenhouse gas emissions but also supports broader safety and efficiency objectives. ISOBAR's commitment to sustainability is further underscored by its use of renewable energy sources and stringent adherence to data protection and security measures. The ethics-based assessment requested aims to identify and mitigate any potential environmental or human health impacts throughout the system's lifecycle.
- **Integration into Existing Procedures and Impact on Work and Skills. Objective ET-07**: Integrating ISOBAR into existing air traffic management protocols requires no alteration to current practices for utilizing weather information. ISOBAR's delivery of

[46] Objective ET-05: The applicant should ensure that end users are made aware of the fact that they interact with an AI-based system, and, if applicable, whether some personal data is recorded by the system. Objective ET-06: The applicant should perform an environmental impact analysis, identifying and assessing potential negative impacts of the AI-based system on the environment and human health throughout its life cycle (development, deployment, use, end of life), and define measures to reduce or mitigate these impacts. Objective ET-07: The applicant should identify the need for new skills for users and end users to interact with and operate the AI-based system and mitigate possible training gaps (link to Provision ORG-07, Provision ORG-08). Objective ET-08: The applicant should perform an assessment of the risk of de-skilling of the users and end users and mitigate the identified risk through a training needs analysis and a consequent training activity

more accurate and reliable weather data through established interfaces ensures that no new skills are needed for its operation.

- **Human Operator Roles, Responsibilities, and Societal Impact. Objective ET-08**: ISOBAR does not alter the roles or responsibilities of human operators. It serves as a tool that improves situational awareness and supports the decision-making process of the users, without substituting or undermining their authority or judgment. Users of ISOBAR are tasked with confirming the accuracy and relevance of the weather information provided by the system and applying it in accordance with the relevant regulations. Additionally, ISOBAR does not interfere with the communication or coordination among the various participants in the air traffic management system, such as air traffic controllers, pilots, network managers, or meteorological services.

6.2.6 Second building block: AI Assurance. Third building block: Human Factors for AI. Forth building block: AI Safety Risk Mitigation.

At the time of writing this document, the analysis has reached this stage, and the AI assurance building block is being finalized.

The key concepts related to the missing building blocks can be derived from the Concept Paper and are outlined below.

The **AI assurance building block** focuses on three main areas specific to AI-based systems:

- **Learning Assurance**: This area emphasizes the shift from traditional programming to AI/ML learning processes, noting that existing development assurance methods do not fully accommodate the unique characteristics of AI/ML learning.
- **Explainability in Development & Post-Operations**: It highlights the importance of providing users with clear, dependable, and pertinent explanations about how AI/ML applications generate their outcomes, ensuring transparency and understanding at an appropriate level of detail.
- **Data Recording Capabilities**: This aspect addresses the dual needs for operational and post-operational analysis, including ongoing safety monitoring of the AI system and facilitating investigations in the event of incidents or accidents.

The **Human Factors for AI building block** encompasses guidance to address human factors considerations arising from AI integration. It focuses on:

- **AI Operational Explainability**: Ensuring end-users receive clear, reliable, and relevant explanations about the workings and outcomes of AI/ML applications, with proper detail and timing.
- **Human-AI Teaming**: Introducing strategies for effective collaboration between humans and AI systems to optimize performance and ensure safety.

The **AI Safety Risk Mitigation building block** acknowledges the challenges in fully deciphering the '*AI black box*' to meet all objectives of the '*AI Assurance*' and '*Human Factors for AI*' building blocks. It emphasizes the necessity to manage the residual risk associated with AI's inherent uncertainty.

All the four building blocks play a crucial role in establishing trust in AI/ML applications. Each block, whether focused on AI assurance, human factors, explainability, or safety risk mitigation, contributes uniquely to the broader goal of ensuring AI systems are reliable, understandable, and safe. Together, they form a comprehensive framework that addresses various aspects of AI functionality and interaction with human users. By meticulously addressing these areas, stakeholders can foster a higher level of confidence in AI applications, ensuring they not only perform as intended but also operate in a manner that is transparent and aligned with ethical standards.

7. Conclusions

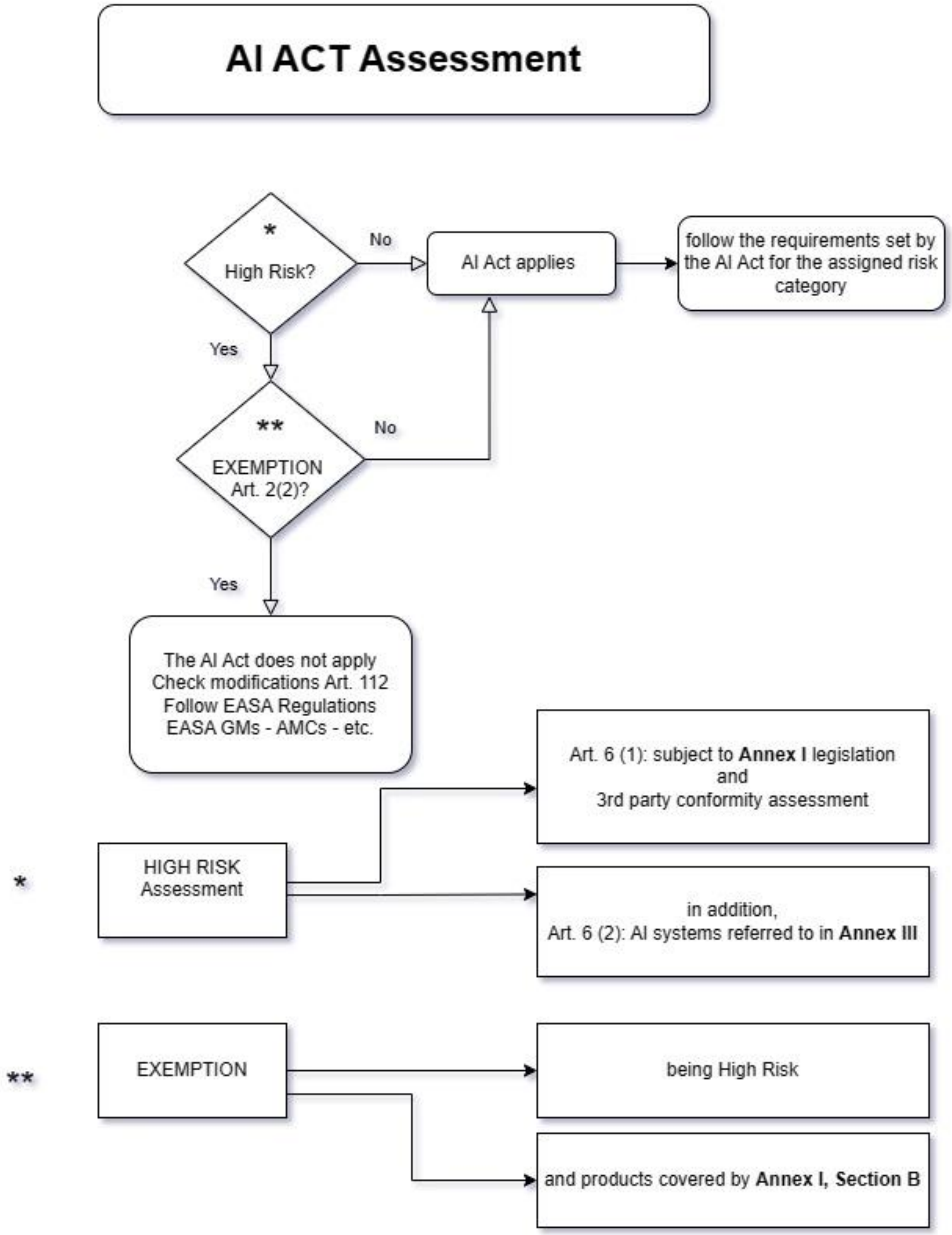
As has been understood, I found myself addressing a particularly innovative theme within an incomplete and constantly changing legal context. At the time of writing, the AI Act has not yet been published in the official journal, and EASA legislation has not yet been implemented to account for AI. As extensively described earlier, the enforcement of the regulations will be delayed, based on the different risk of the AI application.

Through this thesis, I have aimed to elucidate the journey undertaken, from understanding the context through a historical overview to defining the regulatory framework for application, using the far-from-simple legal definition of AI. Then, I sought to clarify how current civil aviation legislation accounts for AI, using guidance material and how it will be amended in the future, in an innovative way, compared to the past. Lastly, I concretely presented two use cases I collaborated on, offering the reader an insight into the various steps / assessments needed to comply with the current guidance received from EASA.

Understanding how to interact with technology that is already available, yet in the absence of binding regulations that unequivocally dictate its usage from development to deployment, has not been straightforward but has certainly been challenging and stimulating. Efforts were made to navigate through non-binding guidance material while attempting to anticipate the future binding legislation, of which only drafts or texts progressing through the legislative pathway are known, providing concrete directions to developers and various entities responsible for safety.

I hereby repropose, through a flowchart, the logical path that results from all the studies and reasoning I've done so far, which should guide the developer of AI apps in Eurocontrol to be considered compliant. Firstly, it should be remembered that there is currently no legislation in force on AI and that the guidance material is in no way binding.

Assuming we are at a point where the AI Act and EASA legislation have come into force, let's say we are about to develop an application based on ML and is 'safety related' (EASA guidance currently applies only to these cases). Not knowing, of course, the text of the regulations EASA will produce in the future (as drafter), let's assume for now that it fully reflects what is provided in the guidance material.



Commenting on the AI Act Assessment flowchart.

The first step is to understand that the AI Act will apply to all domains, thus, the first thing to do is proceed with the Assessment required by the AI Act. As mentioned in the chapter dedicated, the regulation adopts a '*risk-based*' approach with the application of requirements proportional to the risk level. The higher the risk, the more stringent the requirements to be applied, and *vice versa*, excluding activities explicitly indicated as prohibited, for which no requirements apply but are entirely banned, preventing their dissemination and use.

Therefore, the first question to ask is whether our application should be considered *'high risk'* according to the AI Act, as it's important to understand if there's a kind of exemption from applying the AI Act to applications, specifically high risk, whose products fall within EASA regulations, as may happen in our case.

This first assessment is very important as it allows us to understand whether we need to continue considering the AI Act with its requirements, for example, as a *low-risk* application, or as a *high-risk* application but not subject to specific exemption.

It is then necessary to proceed with verifying the satisfaction of what is provided by *Article 6*, to understand if we are in a *high-risk* context, and *Article 2*, to verify the possible exemption.

This exemption, as explained in the dedicated chapter on the AI Act, is counterbalanced by the provision of *Article 108*, which obliges future EASA legislator (for simplicity, it is not remembered that the legislation must be promulgated by the European Commission) to consider the requirements provided by the AI Act for *high-risk* applications, as for this type of regulation, it is desired to avoid the AI Act interfering with existing conformity assessments, rather than with the competent authorities, etc.

Therefore, on one hand, the exemption is provided, and on the other, it is expected that, in any case, when the legislator intends to amend specific regulations, it must do so in line with what the AI Act legislator has envisaged for what he considers *high-risk* applications.

Hence, if the application is not considered *high risk* according to the AI Act parameters, there will be no exemption, but the AI Act will apply even if the applications fall within the EASA regulatory field (assuming and not conceding that one can be considered *low risk* despite falling within EASA regulations, something yet to be demonstrated, but theoretically possible).

However, it should be understood that if, on the one hand, the outcome of the assessment according to the AI Act is *low risk*, as seen, the requirements will be equally lenient (primarily *transparency* needs) and, on the other hand, that EASA regulations can also cover such cases (as mentioned in the dedicated chapter), as long as the requirements set forth do not conflict with those stipulated by the AI Act (which, it is remembered, is a European Regulation that hierarchically is placed at the same level as the Basic Regulation, therefore at a higher level compared to the Implementing Rules).

Therefore, two situations can occur in the presence of applications at *high risk*, according to the AI Act:

- the first in which the exemption provided by *Article 2* cannot be obtained (for example, because it does not fall into one of the sectors indicated by *Annex III*) and for which, as in the previous case (*low risk*), the AI Act will continue to apply, and
- the other in which the exemption from applying the AI Act provisions is obtained, and there will be only the obligation to check any future amendments of the AI Act provisions, as provided by Article 112.

Let's assume, to continue the analysis of the flowchart, that we are in the situation where the exemption provided by *Article 2 (2)* has been obtained: here we are in the presence of an AI application that is *high risk*, according to the AI Act, but for which we are essentially indicated to continue to follow the sectoral legislation (EASA) and verify any changes to the AI Act.

Therefore, EASA legislation comes into play, which we have assumed to be entirely assimilable to the existing guidance material (in reality, as we said earlier, it is known that EASA intends to provide many more requirements, similar to the AI Act – high-risk ones, and indications in the future 'AI-EASA regulations'), which provides, as seen, another type of assessment.

The first assessment to be made is therefore whether the application should be considered based on *Machine Learning* models and *'safety related'*, to then apply all the objectives indicated in the guidance material (the '*Concept Paper'*), modulated according to the proportionality that has been provided to take into account the riskiness of the application; riskiness that can derive from two aspects: the *AI Level* (which, essentially, is parameterized to the degree of AI autonomy with respect to humans) and the *criticality / or Assurance level* (the SWAL) of the item containing the *Machine Learning* model.

In fact, this is what was done in the two analyzed use cases, for which the indications provided by the guidance material were followed pedantically, even in the awareness that things might change in the future.

Some questions remain, and only with the verification of the regulatory evolution will certain answers be obtained, including, for example, how to behave in the presence of an AI application based on *Machine Learning* that falls within the EASA field and that has obtained the exemption from applying the AI Act, but is not considered *'safety related'* (as it is considered, according to safety parameters, that it does not have a possible impact).

In the NM context in Eurocontrol, AI, even if it is not suitable to bring a safety risk, could bring other risks in terms of compliance, violation of human rights, performance, etc. and that should in any case be addressed not only in terms of risk management but also regulated, as EASA is imagined doing, in the future.

In all these cases, not yet regulated by current legislations, it would be advisable to apply in any case an *internal code of conduct* to the organization to protect the principles, widely shareable, found in the AI Act and that would apply to all AI applications developed and/or used within the same organization.

As a final consideration, I would like to make a personal remark about this wonderful experience.

If the theme was undoubtedly fascinating as it combined my legal skills with my passions (aviation and technology), the human factor was crucial in making this experience unforgettable.

I was able to meet and work with wonderful people from different places, with different cultures and languages, but who, also because of this, make a decisive contribution to the teamwork. Teamwork that is the basis of everything. Whether in person, remotely, mixed, it doesn't matter, we organize ourselves and carry out work, for objectives, together in an atmosphere of sincere and effective collaboration.

To convey the atmosphere I breathed in Eurocontrol, I quote the words of Tony Licu (whom I can never thank enough, along with Benjamin Cramet, for how they behaved humanly before professionally towards me) about *Just Culture*, in which, in trying to explain in substance, as he loves to do, what it is, namely, simply, ***'being fair with people and doing the right thing*** […] ***creating the right environment***' foreseeing a way of working that **'*make easy to do the right thing and hard the wrong thing*'**.

I can assure that those who work in Eurocontrol have succeeded in this intent and that this way of dealing with people and working together I will carry with me forever, both professionally and humanly.

Bibliography

- **Contissa, Giuseppe; Galli, Federico; Godano, Francesco; Sartor, Galileo.** *Il Regolamento Europeo sull'intelligenza artificiale – Analisi informatico-giuridica.* December 2021.
- **EASA.** *Artificial Intelligence Roadmap 2.0. Human-centric approach to AI in aviation.* May 2023, Version 2.0.
- **EASA.** *EASA Concept Paper: Guidance for Level 1 & 2 machine learning applications. A deliverable of the EASA AI Roadmap.* March 2024, Issue 2.
- **EASA.** *EASA Concept Paper: Guidance for Level 1 & 2 machine learning applications. A deliverable of the EASA AI Roadmap.* February 2023.
- **EASA.** *European Plan for Aviation Safety (EPAS) – Vol. II – EPAS actions.* 2024 Edition.
- **EASA.** *European Plan for Aviation Safety (EPAS) 2023-2025 – Vol. I – Strategic priorities.*
- **EASA.** *Machine Learning Application Approval - RESEARCH PROJECT EASA.2021.C38, MLEAP UNIFIED DELIVERABLE PHASE 2.* 11 May 2023.
- **European Aviation Artificial Intelligence High Level Group.** *The FLY AI Report - Demystifying and Accelerating AI in Aviation/ATM.* 5th March 2020.
- **European Central Bank.** *OPINION OF THE EUROPEAN CENTRAL BANK of 29 December 2021 on a proposal for a regulation laying down harmonised rules on artificial intelligence.*
- **European Commission.** *COMMUNICATION TO THE COMMISSION - Artificial Intelligence in the European Commission (AI@EC) - A strategic vision to foster the*

development and use of lawful, safe and trustworthy Artificial Intelligence systems in the European Commission. 24.01.2024.

- **European Commission.** *Study to Support an Impact Assessment of Regulatory Requirements for Artificial Intelligence in Europe FINAL REPORT (D5).* 2021.
- **Future of Life Institute.** *Website on EU Artificial Intelligence Act.* 2024.
- **International Civil Aviation Organization.** *Working Paper* - *ASSEMBLY — 40TH SESSION EXECUTIVE COMMITTEE Agenda Item 26: Other high-level policy issues to be considered by the Executive Committee* - *ARTIFICIAL INTELLIGENCE AND DIGITALIZATION IN AVIATION.*
- **International Standard.** *ISO/IEC 42001* - *Information technology — Artificial intelligence — Management system.*
- **Madiega, Tambiama.** *European Parliamentary Research Service* - *Members' Research Service* - *PE 698.792.* June 2023.
- **OECD.** *Explanatory memorandum on the updated OECD definition of an AI system – OECD Artificial Intelligence papers.* March 2024, No. 8.
- **SESAR Joint Undertaking.** *ISOBAR Final Project Result Report.* 7.11.2022.