Universitat Autònoma de Barcelona

Responsabilidad, Human Oversight y Certificación de Sistemas de IA Basados en ML en ATC

Memoria del Trabajo Fin de Grado en Gestión Aeronáutica

realizado por

Adriana Vilamitjana Sagués

y dirigido por

Tomás De Urrengoechea Cantavenera

Escuela de Ingeniería

Sabadell, Junio de 2026

AGRADECIMIENTOS

Este Trabajo de Fin de Grado es el resultado de mucho más que horas de estudio y trabajo académico. Es el reflejo de un camino lleno de retos, dudas y aprendizajes, pero también de apoyo, paciencia y confianza por parte de muchas personas que me han acompañado en todo momento. A lo largo de estos meses he aprendido que un proyecto como este no se sostiene únicamente sobre el esfuerzo individual, sino sobre todas las personas que, de un modo u otro, lo hacen posible desde la sombra. Sin su ayuda, sus consejos y su ánimo constante, este trabajo no habría llegado a buen puerto. Por eso quiero dedicar estas palabras a quienes han estado a mi lado durante este proceso.

A quien más tengo que agradecer es a mi familia. A mis padres, por su apoyo incondicional, por haberme dado la oportunidad de llegar hasta aquí y por no dejar de creer en mí ni un solo momento, incluso cuando yo misma dudaba. Gracias por escucharme, por animarme y por aguantarme en los días en los que parecía que no veía la luz al final del túnel. Por entender mis ausencias durante las épocas de más trabajo, por celebrar conmigo cada pequeño avance y por recordarme, una y otra vez, que era capaz de terminar lo que había empezado. Todo lo que soy y todo lo que he conseguido se lo debo en gran parte a ellos.

En definitiva, a toda mi familia: gracias por ser mi refugio en los momentos de cansancio e incertidumbre y el motor que me ha ayudado a no rendirme. Sois mi apoyo más firme, y nada de esto habría sido posible sin vosotros.

También quiero dar las gracias a mis amigos y compañeros de carrera, con quienes he compartido nervios, dudas y largas horas de trabajo a lo largo de estos años. Gracias por hacer este camino mucho más llevadero, por las tardes de biblioteca, por los apuntes compartidos y por recordarme, en los momentos más difíciles, que no estaba sola. Esta etapa no habría sido la misma sin ellos, y me llevo de aquí no solo un título, sino también amistades que espero conservar mucho más allá de la universidad.

Por último, quiero agradecer a mi tutor, Tomás De Urrengoechea, su guía y sus orientaciones a lo largo de este trabajo, así como a todos los profesores y profesoras que he tenido durante la carrera, por transmitirme no solo conocimientos, sino también valores y espíritu crítico.

ÍNDICE

1. Introducción 6

2. Marco Conceptual de la Inteligencia Artificial 8

2.1. IA, ML y sistemas basados en datos 8

2.2. Diferencias con el software tradicional 9

2.3. Sistemas críticos en aviación 10

2.4. Enfoque basado en riesgo 10

3. Inteligencia Artificial en ATM 12

3.1. Aplicaciones actuales 12

3.2. Aplicaciones futuras 15

Visión estratégica europea 15

Aplicaciones futuras concretas 15

Horizonte temporal por niveles 16

3.3. Beneficios operacionales 18

4. Riesgos de la IA en entornos safety 19

4.1. Naturaleza de los riesgos 19

4.2. Problemas de opacidad 19

4.3. Limitaciones del Machine Learning 21

5. Marco Regulatorio y Certificación de Sistemas ATM 22

5.1 Marco normativo europeo 22

5.2 Estándares aplicables a software aeronáutico 23

5.3 Concepto de certificación en aviación y principio de seguridad operacional ... 24

5.4 La IA como sistema de alto riesgos 26

5.5 Retos de certificación específicos del Machine Learning 27

Los estándares existentes no son suficientes 27

El proceso en W y el concepto de Learning Assurance 27

El proyecto MLEAP y sus conclusiones 28

6. Human Oversight en Entornos ATC Asistidos por IA 29

6.1. Concepto de Human Oversight 29

6.2. Interacción humano-máquina 29

6.3.El rol del controlador aéreo en sistemas altamente automatizados 31

6.4. Riesgos del sesgo de automatización 32

Definición y origen del concepto 32

Manifestaciones concretas en entornos ATC 32

Factores que agravan el riesgo en sistemas basados en ML 33

Reconocimiento regulatorio del sesgo de automatización 33

6.5.Degradación de habilidades y pérdida de conciencia situacional 34
Pérdida de conciencia situacional 34
Degradación de habilidades 36
El círculo vicioso 36

6.6.Requisitos para una supervisión humana efectiva 37
Transparencia y explicabilidad del sistema 37
Mantenimiento de la implicación cognitiva 37
Formación y mantenimiento de competencias 38

7. Responsabilidad Jurídica en Caso de Fallo de un Sistema IA en ATC 38

7.1 Principios generales de responsabilidad en aviación 38

7.2 Responsabilidad del proveedor de servicios de navegación aérea (ANSP) 40

7.3 Responsabilidad del fabricante o desarrollador del sistema 41

7.4 Posible responsabilidad del controlador aéreo 42

7.5 Responsabilidad del Estado como garante del sistema 43

7.6 Problemas de atribución de culpa en sistemas basados en ML 44

7.7 Escenarios hipotéticos de fallo y análisis jurídico 47
Escenario 1 — Fallo por data drift no detectado 47
Escenario 2 — Fallo en zona fuera del ODD 48
Escenario 3 — Sesgo de automatización demostrable 49

8. Hacia un Marco Integrado: Certificación, Supervisión y Responsabilidad ..51

8.1 Necesidad de adaptación del modelo regulatorio actual 51

8.2 Certificación continua y monitorización post-despliegue 52

8.3 Redefinición del rol del Human Oversight 53

8.4 Reparto equilibrado de la responsabilidad 56

9. Conclusiones 57

Bibliografía: 61

TABLAS

Tabla 1. Software tradicional vs IA / ML. *Pag. 9*

TABLA 2 Distribución de la responsabilidad jurídica entre los actores implicados ante el fallo de un sistema de IA en ATC. Elaboración propia Pag 40

FIGURAS

Figura 1. Clasificación de los sistemas de IA por nivel de riesgo según el AI Act, con la posición de los sistemas de ATM/ATC. Elaboración propia a partir del Reglamento (UE) 2024/1689. *Pag. 11*

Figura 2. Concept image del proyecto JARVIS con los asistentes digitales (Airborne, ATC, Airport, Safety y Operational Efficiency). Fuente: SESAR Joint Undertaking (2024). Pag. 13

Figura 3. Niveles de automatización de la IA en ATM y su horizonte temporal. Elaboración propia a partir de EASA AI Roadmap 2.0 (2023). *Pag 17*

Figura 4. Las técnicas de XAI (LIME y SHAP) y su límite en aviación. Elaboración propia a partir de Adadi y Berrada (2018) y EASA AI Roadmap 2.0 (2023). Pag 21

Figura 5. Comparativa de la Table A-1 (Software Planning Process) del DO-178C y del DO-278A. Fuente: RTCA/EUROCAE, DO-178C (2012) y DO-278A (2011) Pag 24

Figura 6. Fases del proceso de certificación de software y Stages of Involvement (SOI). Fuente: RTCA/EUROCAE, DO-178C (2012). Pag 25

Figura 7. Proceso de learning assurance en forma de W. Fuente: EASA (2024). Pag 28

Figura 8. Manifestaciones del sesgo de automatización en ATC: errores de omisión y de comisión. Elaboración propia a partir de Cummings (2004) y EASA (2023). Pag 34

Figura 9. Los tres niveles de la conciencia situacional aplicados al control de tránsito aéreo. Elaboración propia a partir de Endsley (1995). Pag 35

Figura 10. El círculo vicioso de la automatización en ATC. Elaboración propia a partir de Endsley y Kiris (1995), EUROCONTROL (2015) y EUROCONTROL & SESAR JU (2020). Pag 36

Figura 11. El vacío de causalidad en la responsabilidad por IA: la Directiva (UE) 2024/2853 permite probar el defecto, pero la propuesta de Directiva sobre responsabilidad en materia de IA (2022), que cubría la causalidad, fue retirada en octubre de 2025. Elaboración propia. Pag 46

Figura 12 Tres escenarios hipotéticos de fallo de un sistema de IA en ATC y reparto del peso de la responsabilidad entre los actores implicados. Elaboración propia. Pag 50

Figura 13. Las tres condiciones de un Human Oversight efectivo. Elaboración propia a partir del art. 14 del AI Act, EASA AI Roadmap 2.0 (2023), EUROCONTROL & SESAR JU (2020) y Endsley y Kiris (1995). Pag 55

1. Introducción

La gestión del tráfico aéreo (ATC) opera en uno de los entornos más exigentes desde el punto de vista de la seguridad: cualquier fallo en sus sistemas puede comprometer la integridad de múltiples aeronaves y de las personas que las ocupan. Durante décadas, este entorno se ha sustentado en una combinación de procedimientos estandarizados, software determinista y la supervisión directa de controladores aéreos altamente cualificados. Sin embargo, el crecimiento sostenido del tráfico aéreo está poniendo a prueba los límites de ese modelo. Según las previsiones de EUROCONTROL[1], el tráfico aéreo europeo alcanzará entre 11 y 12 millones de vuelos anuales en 2030, lo que representa un aumento de entre el 15 y el 20 % respecto a los niveles actuales. Gestionar ese volumen con los métodos tradicionales, sin deteriorar los niveles de seguridad, resulta cada vez más difícil.

En este contexto, la inteligencia artificial (IA) y, en particular, los sistemas basados en machine learning (ML), están emergiendo como herramientas de apoyo para mejorar la capacidad, la eficiencia y la toma de decisiones en ATM. No se trata de una prospectiva lejana: proyectos como JARVIS de SESAR JU, el ATC-TBO de EUROCONTROL o el sistema U-space de ENAIRE representan aplicaciones que ya están en distintas fases de desarrollo u operación[2]. La IA ha dejado de ser un horizonte conceptual para convertirse en un reto regulatorio, técnico y jurídico del presente.

El problema es que la introducción de estos sistemas en entornos safety-critical como el ATC genera tensiones que los marcos normativos vigentes no estaban diseñados para resolver. Por un lado, los estándares de certificación aeronáuticos, el DO-278A[3] el DO-178C[4] y el ARP4754A[5] fueron pensados para un software determinista, cuyo comportamiento puede verificarse de forma exhaustiva a partir del código fuente. Los sistemas de ML no funcionan así: su comportamiento emerge del proceso de entrenamiento con datos, no de una lógica programada explícitamente, lo que hace imposible aplicarles directamente los métodos de verificación y validación tradicionales. Por otro lado, el nuevo Reglamento (UE) 2024/1689[6], conocido como AI Act (Parlamento Europeo y Consejo de la Unión Europea, 2024), introduce una capa regulatoria horizontal que se superpone a la normativa aeronáutica existente sin que

---

[1] EUROCONTROL. (2024). *EUROCONTROL forecast update 2024–2030 (Autumn 2024)*. EUROCONTROL. https://www.eurocontrol.int/sites/default/files/2024-10/eurocontrol-seven-year-forecast-2024-2030-october-2024.pdf

[2] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[3] RTCA/EUROCAE. (2011). *DO-278A/ED-109A: Software integrity assurance considerations for communication, navigation, surveillance and air traffic management (CNS/ATM) systems*. RTCA.

[4] RTCA/EUROCAE. (2012). *DO-178C/ED-12C: Software considerations in airborne systems and equipment certification*. RTCA.

[5] SAE International. (2010). *ARP4754A: Guidelines for development of civil aircraft and systems*. SAE International. https://www.sae.org/standards/content/arp4754a/

[6] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

existan todavía mecanismos claros de unión entre ambas. Esta fragmentación regulatoria constituye el vacío que este trabajo pretende abordar.

A esta dificultad técnica y regulatoria se suman dos problemas adicionales. El primero es la opacidad de los modelos de ML, el fenómeno conocido como black-box, que dificulta tanto la certificación de los sistemas como la supervisión efectiva por parte de los controladores. El segundo es el riesgo de que la automatización progresiva deteriore las competencias y la conciencia situacional de los propios operadores, generando una dependencia que puede revelarse crítica precisamente en los momentos de fallo. A ello se añade la pregunta de quién asume la responsabilidad jurídica cuando un sistema de IA causa o contribuye a un incidente en el espacio aéreo, cuestión que el marco normativo actual no resuelve con claridad.

Objetivos

El objetivo general de este trabajo es analizar los retos que plantea la integración de sistemas de IA basados en ML en entornos ATC, con especial atención a tres dimensiones inseparables: la certificación de los sistemas, la supervisión humana efectiva y la atribución de responsabilidad jurídica en caso de fallo.

De este objetivo general se derivan los siguientes objetivos específicos:

1. Identificar en qué medida los estándares de certificación aeronáutica vigentes son adecuados para sistemas basados en ML, y qué enfoques alternativos están siendo desarrollados.

2. Analizar cómo el AI Act clasifica y regula los sistemas de IA en ATM, y cómo interactúa esta regulación con el marco normativo aeronáutico preexistente.

3. Examinar los riesgos que la automatización creciente plantea para la supervisión humana efectiva en entornos ATC, con especial atención al sesgo de automatización y la degradación de habilidades.

4. Determinar cómo se distribuye la responsabilidad jurídica entre los distintos actores (ANSP, fabricante, controlador y Estado) cuando un sistema de IA falla en un entorno ATC.

Metodología

El trabajo se sustenta en una metodología de revisión documental y análisis normativo. Las fuentes primarias incluyen reglamentos europeos de aplicación directa, directivas y propuestas legislativas, estándares técnicos aeronáuticos, y documentos institucionales de EASA(Agencia Europea de Seguridad Aérea), EUROCONTROL, SESAR JU e ICAO(Organización de Aviación Civil Internacional). Las fuentes secundarias comprenden literatura académica en los ámbitos de los factores humanos, la explicabilidad de la IA y la responsabilidad civil por productos defectuosos.

Estructura del documento

El trabajo se organiza en nueve capítulos. Los capítulos 2, 3 y 4 establecen el marco conceptual: definición de los sistemas de IA y ML, sus diferencias con el software tradicional, sus aplicaciones en ATM y los riesgos que introducen en entornos safety-critical. Los capítulos 5 y 6 analizan el marco regulatorio y los retos de certificación específicos del ML, y los requisitos

para una supervisión humana efectiva. El capítulo 7 aborda la responsabilidad jurídica distinguiendo entre los distintos actores implicados. El capítulo 8 propone un marco integrado que incluye certificación, supervisión y responsabilidad. El capítulo 9 presenta las conclusiones dando respuesta explícita a los objetivos planteados.

2. Marco Conceptual de la Inteligencia Artificial

2.1. IA, ML y sistemas basados en datos

La inteligencia artificial (IA) es una tecnología que ha evolucionado de forma considerable en las últimas décadas, abarcando un amplio espectro de técnicas y aplicaciones. Para delimitar su alcance en este trabajo, se utilizarán las definiciones recogidas por los principales organismos reguladores europeos, que constituyen el marco de referencia aplicable en el ámbito de la aviación.

La Organización para la Cooperación y el Desarrollo Económicos define un sistema de IA como "un sistema basado en máquinas que, para objetivos explícitos o implícitos, infiere, a partir de la información que recibe, cómo generar resultados tales como predicciones, contenidos, recomendaciones o decisiones que pueden influir en entornos físicos o virtuales"[7].

El *machine learning* (ML), por su parte, es la rama de la IA dedicada al desarrollo de algoritmos de aprendizaje que permiten a los sistemas evolucionar su comportamiento a partir de la observación de datos y la extracción de inferencias, sin necesidad de ser programados de forma explícita para cada tarea. En este sentido, los sistemas basados en ML son también conocidos como sistemas *data-driven*, es decir, un enfoque centrado en construir modelos que aprenden a partir de haber sido entrenados con un gran volumen de ejemplos[8].

Esta naturaleza condiciona sus características fundamentales:

- El rendimiento del sistema depende directamente de la calidad y representatividad de los datos de entrenamiento.
- Su comportamiento no es fijo ni determinista: a diferencia del software tradicional, donde entradas idénticas producen siempre salidas idénticas, los sistemas ML pueden generar resultados variables.
- En determinadas configuraciones, el sistema puede continuar adaptándose una vez puesto en funcionamiento (*online learning*).

[7] *" OCDE. (2024). Recommendation of the Council on Artificial Intelligence (actualizada mayo 2024). OECD. https://oecd.ai/en/ai-principles*

[8] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

2.2. Diferencias con el software tradicional[9]

Los sistemas de IA, especialmente los basados en ML, presentan grandes diferencias respecto al software tradicional.

Una de las principales diferencias es que el software tradicional se basa en una lógica muy definida por los programadores. El sistema es determinista ya que las reglas están codificadas en el mismo sistema por lo que siempre que introduzcas la misma entrada vas a obtener la misma salida.

En cambio, los sistemas de IA:

- Aprenden a partir de los datos
- Crean modelos internos que no siempre se pueden entender fácilmente
- No funcionan de forma fija, sino con cierta incertidumbre
- Su rendimiento depende de cómo se entrenan, no solo del código

Según la Agencia de la Unión Europea para la Seguridad Aérea los sistemas de aprendizaje automático tienen como principal problema la dificultad para explicar cómo funcionan, su gran dependencia a los datos y lo difícil que es comprobar su comportamiento en cada uno de los casos posibles.

Esto no sucede en los sistemas tradicionales, tal y como hemos comentado estos siempre tienen la misma salida si les pones la misma entrada.

Principales diferencias:

Característica	**Software tradicional**	**Sistemas IA/ML**
Lógica	Programada explícitamente	Aprendida de datos
Comportamiento	Determinista	Probabilístico
Validación	Basada en código	Basada en datos + modelo
Explicabilidad	Alta	Limitada (black-box)[10]
Evolución	Cambios manuales	Puede adaptarse

TABLA 1. SOFTWARE TRADICIONAL VS IA/ML

Resumidamente, el software tradicional se basa en una lógica programada de forma explícita por los programadores, por otro lado, los ML aprenden la lógica a partir de datos. Por este motivo el sistema tradicional responde siempre de la misma manera, pero los basados en datos tienen un comportamiento probabilístico por lo que puede variar.

Otra gran diferencia es la validación, en un software tradicional se comprueba directamente sobre el código por donde se puede comprobar y asegurar que funciona según lo que se ha

[9] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[10] Adadi, A., y Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

programado. Por otro lado, en los sistemas de ML la validación depende de los datos y de como se ha entrenado el modelo por lo que tiene una mayor complejidad.

Una diferencia importante es la explicabilidad, es decir, como ha llegado a esa decisión. En este sentido, en el software tradicional se sabe como y porque ha tomado las decisiones. Por lo contrario, los sistemas de IA son más complejos de entender ya que no se sabe cómo a llegado a la conclusión paso a paso ya que sus relaciones internas son muy complejas.

La ultima gran diferencia es que el sistema tradicional evoluciona a través de cambios manuales en los códigos mientras que los otros se pueden ir adaptando con el tiempo si se vuelven a entrenar y reciben nuevos datos.

Estas diferencias son especialmente críticas en entornos safety, donde la previsibilidad y verificabilidad son requisitos fundamentales.

2.3. Sistemas críticos en aviación[11]

Los sistemas safety-critical son aquellos cuyo fallo en el ámbito aeronáutico puede tener graves consecuencias para la seguridad.

Según la Organización de Aviación Civil Internacional (OACI), los sistemas críticos para la seguridad son aquellos cuyo fallo puede comprometer directamente la seguridad operacional

En el ámbito de la aviación y la gestión del tráfico aéreo (ATM), existen diversos sistemas considerados críticos para la seguridad. Entre ellos se encuentran los sistemas de control del tráfico aéreo, los sistemas de navegación, los sistemas de vigilancia y los sistemas de apoyo a la toma de decisiones de los controladores.

Debido a su importancia, estos sistemas deben cumplir requisitos muy estrictos, entre los que destacan la fiabilidad, la trazabilidad, así como procesos rigurosos de validación, verificación y certificación.

Introducir la IA en estos sistemas le suma unos retos complejos ya que los modelos ML pueden afectar a los principios de seguridad.

Eurocontrol subraya que la integración de la IA en ATM debe garantizar que los sistemas mantengan niveles de seguridad equivalentes o superiores a los actuales.

2.4. Enfoque basado en riesgo[12]

En Europa se regula la IA a través de un enfoque basado en el riesgo, este depende del impacto que puede tener el sistema.

El Reglamento de IA de la UE (AI Act), propuesto por la Comisión Europea, clasifica los sistemas en diferentes niveles según el riesgo en función de su impacto y necesidad de control. Los

[11] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[12] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

riesgos inaceptables son esos en los que su uso está totalmente prohibido ya que supone una amenaza clara a la seguridad. Seguidamente se encuentran aquellos sistemas de alto riesgo pero que están permitidos bajo unos requisitos. Su nivel inferior son aquellos sistemas que no necesitan una regulación tan contundente, pero sí que tienen determinadas obligaciones. Finalmente, los sistemas de riesgo mínimo no están sujetos a ningún tipo de regulación ya que su impacto es bajo y no suponen un gran riesgo.

Los sistemas utilizados en aviación y ATM se consideran generalmente de alto riesgo por su posible gran impacto en la seguridad. Por esta razón hay establecidos unos requisitos que pueden garantizar la seguridad, la fiabilidad y el control. Para empezar, deben tener sistemas de gestión de riesgos para prevenir y mitigar posibles fallos en el sistema. Además, también deben asegurar que los datos utilizados son de calidad ya que el rendimiento del sistema se basa en ellos. Otro aspecto importante es la supervisión humana (*human oversight*) gracias a esto las decisiones que toma el sistema pueden ser controladas o corregidas por personas. Finalmente, también es importante destacar que deben tener un gran nivel de resistencia delante de fallos, errores o ataques externos, es decir, ciberseguridad.

El enfoque basado en el riesgo garantiza que los requisitos regulatorios sean proporcionales al impacto potencial de los sistemas de IA.

Este enfoque es coherente con la tradición de la aviación, donde la seguridad ya se gestiona mediante modelos de análisis de riesgo[13]

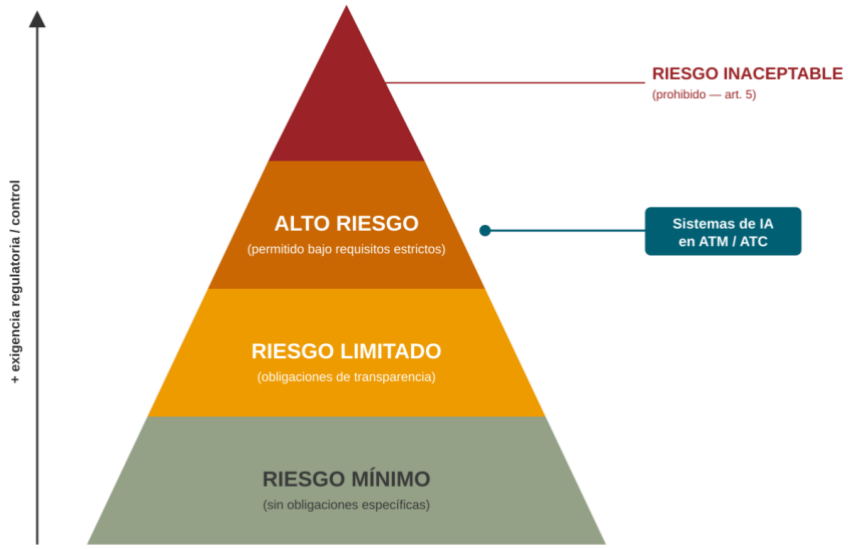


FIGURA 1. CLASIFICACIÓN DE LOS SISTEMAS DE IA POR NIVEL DE RIESGO SEGÚN EL AI ACT, CON LA POSICIÓN DE LOS SISTEMAS DE ATM/ATC. ELABORACIÓN PROPIA A PARTIR DEL REGLAMENTO (UE) 2024/1689

[13] ICAO. (2018). *Safety management manual* (Doc 9859, 4.ª ed.). International Civil Aviation Organization. https://www.icao.int/safety/SafetyManagement/Pages/GuidanceandTools.aspx

3. Inteligencia Artificial en ATM

La aplicación de la inteligencia artificial en el Air Traffic Management (ATM) se entiende como un proceso de transformación digital. No es solo introducir nuevos métodos y herramientas, sino que es una nueva manera de procesar grandes volúmenes de datos, anticipar posibles problemas, apoyar a la toma de decisiones para mejorar la seguridad, la capacidad y la eficiencia, las cuales deben estar perfectamente equilibradas. [14]

Según la Agencia de la Unión Europea para la Seguridad Aérea (EASA) la IA acabará teniendo un papel muy importante en todo lo que comprende la aviación. Como hemos comentado anteriormente, prevén que será un gran mecanismo para conseguir que el sector sea más seguro y sostenible.

Para concretar, en el ATM esta evolución puede tener una gran importancia ya que el tránsito aéreo es un sistema muy exigente, no puede tener fallos y cada vez es más complicado que sea así a causa del gran aumento del tráfico. Las operaciones de tráfico aéreo se están volviendo cada vez más complicadas a causa también del gran aumento de actores, nuevos usuarios del espacio (drones) y la subida de la cota. Por estas razones este sector entiende la IA como una tecnología de apoyo para mejorar la previsión, la toma de decisiones, evitar problemas y la optimización de recursos. Aun así, no es un simple sustituto del controlador aéreo, es solo un apoyo.

3.1. Aplicaciones actuales

Actualmente, la aplicación de la IA en procesos ATM no se debe entender como una automatización completa del control del tráfico aéreo, sino que es una herramienta que en algunos casos se utiliza para ayudar en procedimientos muy concretos en plena fase de desarrollo.[15]

Para este 2026, ENAIRE planea automatizar servicios dentro del marco U-space para drones. El U-space es el conjunto de servicios y reglas que la UE ha definido para integrar drones en el espacio aéreo de baja cota; la IA se incorpora como herramienta para agilizar tareas repetitivas, como la concesión de autorizaciones inmediatas utilizando esta herramienta ya que es un proceso muy repetitivo. El sistema no funcionará solo, es decir, tendrá apoyo humano pero este recurso será inferior al que hay actualmente. Gracias a esto se le podrá dar más apoyo a todo el entorno de ATM el cual necesitará mucho estos recursos ya que se prevé un aumento del tráfico aéreo hasta el 2030.[16]

Nos podemos encontrar con otros casos de la aplicación de la IA, por ejemplo, en 2023 se puso en marcha el proyecto *CIA3 – Clasificación de fiabilidad de Información Aeronáutica con*

[14] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

[15] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

[16] E ENAIRE. (2026). *Plan anual de actuación 2026*. ENAIRE. https://www.enaire.es/docs/inline/es_ES/plan_anual_de_actuación/Plan_anual_actuacion_2026.pdf

Inteligencia Artificial, impulsado por ENAIRE. Esta implementación ayuda a otras inteligencias artificiales a clasificar y dar fiabilidad a la información para que no haya desviaciones en los resultados. La implementación del sistema refuerza la seguridad en la toma de decisiones ya sea a través de otra IA o para los propios humanos. [17]

El proyecto JARVIS[18] **(**Just A Rather Very Intelligent System) parte de la transformación de la digitalización progresiva en la aviación. Lo que se propone es desarrollar asistentes digitales para que colaboren con los operadores humanos, ya sean controladores aéreos, pilotos o personal aeroportuario. Este programa les da apoyo en aquellas tareas que tengan una compleja operación y así mejorar su seguridad y eficiencia

JARVIS aporta cinco asistentes digitales:

- Airborne Assistant: apoya las operaciones en cabina para evolucionar hacia mayores niveles de automatización.
- ATC Assistant: diseñado para apoyar a los controladores en la toma de decisiones, reduciendo la carga de trabajo y mejorando la gestión del tráfico.
- Airport Assistant: optimiza las operaciones en tierra mediante la mejora de la previsibilidad y la gestión de riesgos.
- Safety Assistant: orientado a la detección y gestión proactiva de riesgos operacionales.
- Operational Efficiency Assistant: focalizado en la optimización del rendimiento global del sistema ATM.

El proyecto es un claro ejemplo de un programa que no sustituye al operador humano sino que se usa como un sistema de apoyo para la toma de decisiones. Refuerza la conectividad de sistemas que cada vez están más interconectados.

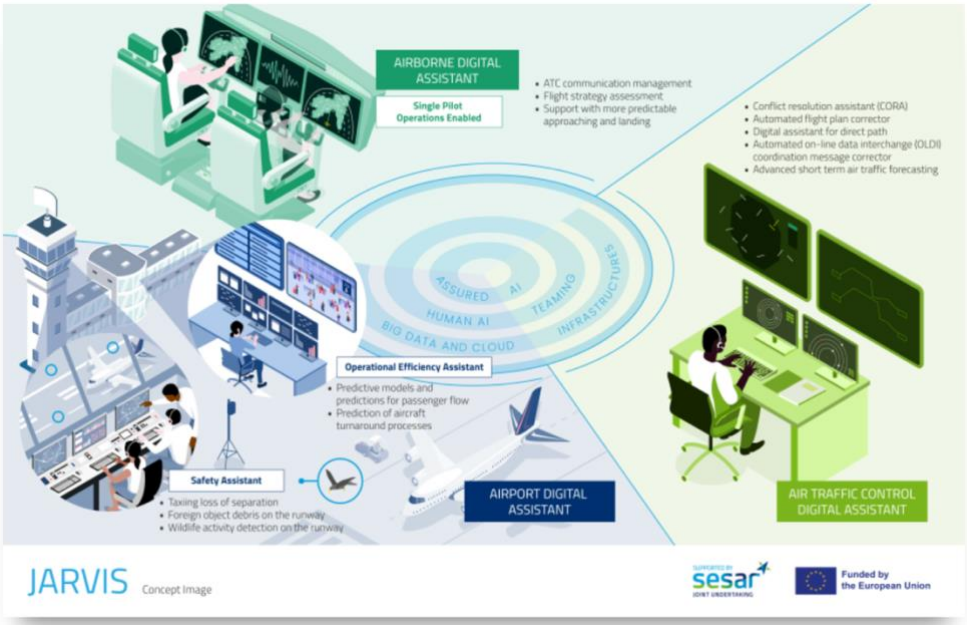


FIGURA 2. CONCEPT IMAGE DEL PROYECTO JARVIS CON LOS ASISTENTES DIGITALES (AIRBORNE, ATC, AIRPORT, SAFETY Y OPERATIONAL EFFICIENCY). FUENTE: SESAR JOINT UNDERTAKING (2024).

[17] ENAIRE. (2023, 20 de junio). *ENAIRE premia ideas de negocio y retos tecnológicos en gestión del tráfico aéreo*.
ENAIRE. https://www.enaire.es/es_ES/2023_06_20/ndp_premio_ideas_negocio_crida
[18] SESAR Joint Undertaking. (s.f.). *JARVIS: Just a rather very intelligent system*. SESAR JU. https://www.sesarju.eu/projects/JARVIS

El proyecto ATC-TBO[19] está impulsado por EUROCONTROL y se basa en la implementación del concepto Trajectory Based Operations (TBO) en la gestión del tráfico aéreo (ATM). Se concibe como una gran evolución ya que sitúa la trayectoria de vuelo en cuatro dimensiones, añadiendo el espacio y el tiempo. Este conjunto de parámetros se une para mejorar la planificación y gestión del tráfico.

El objetivo principal de este proyecto es mejorar la precisión, previsibilidad y eficiencia de todas las operaciones aéreas con el soporte de datos compartidos y actualizados a tiempo real entre todos los actores que participan en el sistema. En lugar de reaccionar a los problemas que ocurres, el sistema permite anticipar estos problemas, optimizar rutas y ajustar la capacidad del espacio aéreo para que todo se una de manera coordinada.

Para hacer esto posible, el proyecto une tecnologías avanzadas en análisis de datos y la inteligencia artificial para dar soporte a la toma de decisiones. Lo que aportan estas herramientas es poder procesar un gran volumen de información y genera recomendaciones que facilitan la gestión de flujo de tráfico, reducen retrasos y mejoran el uso del espacio aéreo.

La función de este sistema es dar apoyo a las operaciones humanas por lo que se ha determinado que está en un punto de transición hacia un modela más colaborativo, predictivo y automatizado. Este punto es muy importante para poder desarrollar el Cielo Único Europeo y así poder modernizar el sistema a escala continental.

El Network Management[20] europeo, desarrollado por EUROCONTROL, constituye uno de los sistemas operativos más relevantes en la gestión del tráfico aéreo en Europa. A través del denominado Integrated Network Management (INM), se coordina de forma centralizada el funcionamiento de toda la red aérea europea, integrando información procedente de múltiples actores: proveedores de servicios de navegación aérea, aerolíneas, aeropuertos y gestores de espacio aéreo.

El objetivo principal de este sistema es optimizar el rendimiento global de la red, para garantizar un equilibrio entre capacidad y demanda. Para ello, el Network Manager recopila y analiza continuamente datos operacionales (planes de vuelo, condiciones meteorológicas, disponibilidad del espacio aéreo o incidencias operativas), lo que permite anticipar situaciones de congestión y aplicar medidas de regulación del tráfico de forma coordinada a nivel europeo.

A diferencia de enfoques tradicionales más fragmentados, el Integrated Network Management introduce una visión sistémica y colaborativa del ATM, donde las decisiones no se toman únicamente a nivel local, sino considerando el impacto en toda la red. En este contexto, la digitalización y el análisis avanzado de datos permiten mejorar la predicción de la demanda, la asignación de capacidad y la eficiencia operativa ante perturbaciones.

Este sistema representa un ejemplo claro de aplicación real de tecnologías digitales en ATM, donde la automatización no sustituye al operador humano, sino que proporciona herramientas de apoyo a la toma de decisiones a gran escala.

[19] EUROCONTROL. (s.f.a). *ATC-TBO project*. https://www.eurocontrol.int/project/atc-tbo
[20] EUROCONTROL. (s.f.b). *Integrated network management*. https://www.eurocontrol.int/project/integrated-network-management

3.2. Aplicaciones futuras

Visión estratégica europea

La integración de la IA en el ATM no es algo que esté pasando de forma espontánea, sino que responde a una visión institucional muy clara a escala europea. Dos documentos son los que mejor reflejan esta dirección.

El primero es el FLY AI Report de EUROCONTROL[21], que fue el primer intento serio de analizar de forma sistemática qué papel podría jugar la IA en la aviación europea. Lo más relevante de este informe es que no habla de la IA como una solución puntual sino como algo que va a afectar a todo el sistema, y en especial al ATM. Según el documento, la IA va a ser clave para mejorar la gestión de la capacidad, hacer predicciones más precisas y optimizar el tráfico ante un crecimiento que ya no para. La formación de los operadores y el mantenimiento de sus competencias no pueden quedar en segundo plano, sino que deben tratarse como un requisito operacional igual de importante que cualquier especificación técnica.

El segundo documento es el EASA AI Roadmap 2.0,[22] publicado en 2023, que aborda la cuestión desde un enfoque más regulatorio y estratégico. A diferencia del FLY AI Report, no describe proyectos concretos, sino que trata de establecer las bases y normativas para que la IA pueda integrarse de forma segura en todos los ámbitos de la aviación. El concepto que más repite es el de *trustworthiness*, que vendría a ser la combinación de seguridad técnica, robustez, transparencia y supervisión humana efectiva. Lo interesante del Roadmap es que reconoce abiertamente que el marco regulatorio actual no está preparado para dar respuesta a todos los retos que plantea la IA, y que hace falta construir uno nuevo de forma progresiva. En definitiva, los dos documentos apuntan en la misma dirección: la IA como apoyo al operador, nunca como sustituto, y siempre de forma gradual y controlada.

Aplicaciones futuras concretas [23] [24]

Más allá de la visión estratégica, lo que resulta interesante es ver qué aplicaciones concretas se están planificando para los próximos años. Los documentos de SESAR JU y EUROCONTROL identifican varias líneas de desarrollo que permiten hacerse una idea bastante clara de hacia dónde va el sector, y que conviene agrupar por ámbito para entenderlas mejor.

En lo que respecta a la gestión de conflictos, se está trabajando en sistemas capaces de detectar conflictos de separación con entre 10 y 20 minutos de antelación, lo que supone una mejora considerable respecto a los 5-8 minutos que ofrecen los sistemas actuales. Disponer

---

[21] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

[22] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

[23] SESAR Joint Undertaking. (2023). *European ATM master plan: Automation roadmap* (ed. 2023). SESAR JU. https://www.sesarju.eu/node/4820

[24] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

de más tiempo para reaccionar cambia completamente la naturaleza de la tarea del controlador, que puede pasar de gestionar situaciones urgentes a planificar con mayor anticipación. Yendo un paso más allá, en sectores de baja densidad de tráfico se plantea que el sistema no solo detecte el conflicto, sino que también proponga y ejecute la maniobra de separación de forma autónoma, aunque siempre con un operador humano supervisando y con capacidad de intervenir en cualquier momento.

En cuanto a la gestión del espacio aéreo, una de las aplicaciones más prometedoras es la sectorización dinámica, que permitiría adaptar los límites de los sectores de control en tiempo real según la demanda real de tráfico. Hoy en día los sectores tienen unos límites bastante fijos que no siempre reflejan bien cómo se distribuye realmente el tráfico en cada momento, lo que genera tanto situaciones de congestión como de capacidad. La sectorización dinámica resolvería este problema, pero para que funcione bien necesita modelos predictivos capaces de anticipar cómo va a evolucionar el tráfico con suficiente antelación, lo que a su vez depende de la calidad y representatividad de los datos disponibles.

La predicción de demanda y meteorología es otra área donde el ML tiene mucho potencial y donde los beneficios serían bastante visibles para el sistema en su conjunto. La idea es integrar datos meteorológicos de alta resolución con información de tráfico en tiempo real para generar predicciones mucho más precisas que las actuales. Esto permitiría anticipar mejor los escenarios de saturación y reducir las medidas de regulación reactivas que tan frecuentemente generan retrasos en cadena en toda la red europea.

Finalmente, hay dos conceptos operativos que podrían cambiar de forma bastante profunda la forma de trabajar en ATC. El primero es el *Single Controller Operations* (SCO), que plantea que un único controlador pueda gestionar sectores que hoy requieren dos personas, apoyándose en sistemas de IA que asumen las tareas de monitorización rutinaria y liberan al operador para concentrarse en las decisiones más complejas. El segundo es la evolución de las torres de control remotas hacia un modelo en el que un operador pueda supervisar varios aeropuertos simultáneamente con asistencia de IA, lo que tendría implicaciones muy importantes tanto para la estructura de los servicios ATC como para los requisitos de formación y certificación de los controladores. Ambos conceptos están todavía en fases de desarrollo y validación, pero representan un cambio estructural que el sector ya está tomando en serio.

Horizonte temporal por niveles

Para entender cuándo puede esperarse que todas estas aplicaciones sean una realidad operativa, el EASA AI Roadmap 2.0[25] propone una evolución en tres niveles que resulta muy útil para situar cada desarrollo en el tiempo y entender por qué algunos proyectos avanzan más rápido que otros.

El Nivel 1 es el punto de partida y el que mejor describe la situación actual del sector. En este nivel la IA asiste al controlador, genera recomendaciones, alertas y predicciones, pero quien toma todas las decisiones operacionales sigue siendo el humano. El sistema informa, el controlador decide. La mayoría de los proyectos que se están desarrollando ahora mismo se encuentran en este nivel, incluyendo los sistemas de detección de conflictos y los asistentes

[25] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

de predicción de demanda. El marco regulatorio para este nivel está bastante más avanzado que para los otros dos, en parte porque el papel del humano en la cadena de decisión es más claro y fácil de certificar. Aun así, todavía tiene limitaciones importantes cuando se trata de sistemas basados en ML, especialmente en lo que respecta a la verificación del comportamiento del modelo.

El Nivel 2 implica un paso importante: el sistema ya no solo recomienda, sino que puede ejecutar ciertas acciones de forma autónoma dentro de unos parámetros definidos, mientras el controlador supervisa y puede intervenir en cualquier momento. La relación entre el humano y la máquina se vuelve más colaborativa, pero también más compleja, porque el operador necesita entender qué está haciendo el sistema y por qué para poder supervisarlo con eficacia. Aplicaciones como la resolución autónoma de conflictos en sectores de baja densidad o la sectorización dinámica automatizada pertenecen a este nivel. Su implantación generalizada no se espera antes de la segunda mitad de la década de 2030, y está condicionada al desarrollo de medios de cumplimiento específicos para ML que hoy todavía no existen y que son uno de los principales retos regulatorios del momento.[26]

El Nivel 3 es el horizonte más lejano y también el más incierto. Aquí el sistema operaría de forma autónoma en un dominio operacional bien definido, reservando la intervención humana únicamente para situaciones excepcionales fuera de ese dominio. Es el nivel que más preguntas abre, no solo desde el punto de vista técnico sino también desde el regulatorio y el jurídico: si el sistema opera de forma casi independiente. No hay todavía una hoja de ruta regulatoria clara para este nivel, y su viabilidad real en entornos ATC depende de avances muy significativos tanto en la certificación del ML como en la definición del marco de responsabilidad.

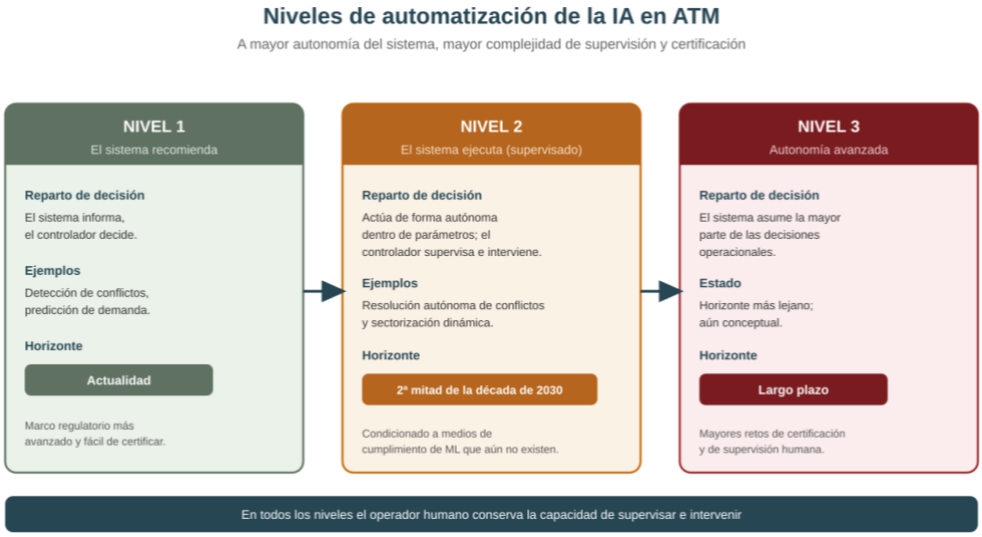


FIGURA 3. NIVELES DE AUTOMATIZACIÓN DE LA IA EN ATM Y SU HORIZONTE TEMPORAL. ELABORACIÓN PROPIA A PARTIR DE EASA AI ROADMAP 2.0 (2023).

[26] EASA. (2024a). *Concept paper: Guidance for level 1 & 2 machine learning applications — Issue 02*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-concept-paper-guidance-level-1-2-machine-learning

3.3. Beneficios operacionales[27]

Se determinan los beneficios de la inteligencia artificial en base a la seguridad, eficiencia, capacidad operativa y sostenibilidad. Esto se cumple gracias a que EASA determina que la IA consigue extraer grandes cuantidades de información, detecta patrones para anticipar riesgos y optimizar procesos por lo que ha determinado que el impacto es positivo en todos los ámbitos de la aviación.

Cuando hablamos de seguridad operacional no nos referimos a que el sistema de IA sea seguro en sí mismo, sino que ayuda a prevenir errores por lo que aporta un incremento en la seguridad de las operaciones. Lo que nos ayuda a que esto sea así es la fiabilidad de la información, la mejora de las predicciones meteorológicas o la detección de conflictos[28]. Es todo este conjunto el que determina que la toma de decisiones se hace de manera segura por un humano con la ayuda de estos sistemas de apoyo.

La eficiencia de las operaciones se basa en la automatización de tareas repetitivas, la reducción del aspecto administrativo y fidelidad en la información para ajustar los recursos. Para ser específicos, en el ámbito del ATM la mejora de la eficiencia se centra en la buena previsión de la demanda y así poder asignar mejor la capacidad e incluso poder dar una respuesta más rápida a los problemas de último momento. La estrategia de ENAIRE es básicamente mejorar esta eficiencia operativa y de respuesta de los agentes de tráfico aéreo.

Mejorar la capacidad del sistema ATM es complicado por qué nos encontramos en un momento de crecimiento en el tráfico aéreo por lo que automatizar estas operaciones sin bajar la seguridad es complicado. Para que esto sea posible es imprescindible mejorar la anticipación de la demanda y equilibrar la relación entre el tráfico, la meteorología y los recursos. Por eso ENAIRE[29] determina que la mejora se encuentra en la relación de la capacidad y la demanda.

Estos tres puntos están todos relacionados con la mejora de la sostenibilidad. La reducción de las emisiones depende de muchos factores, pero si se consigue una mejor eficiencia de los recursos implica una disminución de emisiones. Para que esto ocurra necesitamos un sistema con menos ineficiencias operativas.

Estos beneficios no los vamos a ver de manera inmediata junto con la aplicación del sistema. Es decir, en el ATM solo da valor cuando los resultados son fiables y cumplen con los requisitos de seguridad y supervisión humana.

---

[27] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[28] ICAO. (2025). *Prácticas actuales y mejoras de inteligencia artificial para la aviación civil internacional* (Nota de trabajo A42-WP/401 EX/176). Organización de Aviación Civil Internacional. https://www.icao.int/sites/default/files/Meetings/a42/Documents/WP/wp_401_es.pdf

[29] E ENAIRE. (2026). *Plan anual de actuación 2026*. ENAIRE. https://www.enaire.es/docs/inline/es_ES/plan_anual_de_actuación/Plan_anual_actuacion_2026.pdf

4. Riesgos de la IA en entornos safety

La aplicación de la Inteligencia Artificial (IA) y sobre todo los basados en el *machine learning* (ML), suponen un cambio muy importante y significativo teniendo en cuenta como son los sistemas tradicionales. Los entornos *safety-critical* como puede ser la gestión del tráfico aéreo, necesitan que los sistemas sean de alta fiabilidad. Con estos nuevos sistemas de IA, se introducen nuevos riesgos que los softwares tradicionales no tenían.

Estos modelos suponen más riesgo que los otros por qué su funcionamiento no depende únicamente del código, sino que también de los datos que se le han adjuntado y la manera de entrenar al sistema. Por esta razón es imprescindible que también se actualicen los sistemas de validación de los sistemas tradicionales, esto es muy importante por qué no siempre es posible saber la respuesta del sistema ante una operación.

4.1. Naturaleza de los riesgos[30]

Uno de los principales riesgos viene de la propia naturaleza del sistema, su dependencia a los datos. Si estos no se entrenan correctamente el software puede dar resultados incorrectos o poco fiables. Por esta razón es un aspecto muy crítico de la aviación ya que cada escenario es diferente.

Además, utilizan un método probabilístico por lo que no siempre ofrece la misma respuesta para una misma entrada. Esto provoca mucha incertidumbre de los resultados y dificulta predecir cual será el comportamiento o respuesta para todos los casos. Por otro lado, la validación del sistema resulta muy complicada, el sistema es tan completo que resulta inviable comprobar todas las combinaciones posibles entre las entradas y las salidas del software.

Otro gran riesgo se encuentra en aquellas situaciones en las que el sistema no ha sido entrenado por lo que da un comportamiento inesperado. Estos escenarios extraordinarios provocan respuestas imprevistas que pueden llegar a provocar problemas a nivel de seguridad. Por esta razón, EASA enfoca el problema basándose en el riesgo, los requisitos de validación se determinan en base al impacto que puede generar a la seguridad.

4.2. Problemas de opacidad

Uno de los problemas más importantes de los sistemas de ML es que no podemos saber exactamente cómo llegan a sus conclusiones. Este fenómeno se conoce como *black-box*[31]: sabemos qué entra y qué sale, pero no qué pasa en medio. No es un fallo puntual, sino una característica propia de modelos complejos como las redes neuronales, donde el número de parámetros y relaciones internas hace imposible seguir el razonamiento paso a paso.

Se distinguen dos tipos de opacidad. La primera es la opacidad estructural: el modelo tiene tantos parámetros internos que ni sus propios creadores pueden explicar cómo funciona. La segunda es la opacidad epistémica: aunque conozcas el modelo, no puedes reconstruir por

[30] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[31] Páez, A. (2025). Transparencia, explicabilidad y confianza en los sistemas de aprendizaje automático. En J. D. Gutiérrez y R. Manrique (Eds.), *IA en acción: Desarrollo, impacto e implicaciones éticas en la sociedad*. Ediciones Uniandes. https://philarchive.org/archive/PEZTEY

qué tomó una decisión concreta en un momento dado. En la práctica, ambas suelen aparecer juntas en los modelos más avanzados, que son precisamente los que mejor rendimiento tienen.

En un entorno como el ATM esto es un problema serio por tres razones principales. Primero, dificulta la certificación, exigen poder demostrar que el sistema funciona correctamente en todos los escenarios posibles, algo que con un modelo opaco es prácticamente imposible porque no puedes conocer su comportamiento interno. Segundo, complica la supervisión humana: si un controlador recibe una recomendación del sistema, pero no entiende por qué se ha generado, no puede valorar si es fiable o no, lo que le pone en una situación muy difícil a la hora de decidir si seguirla o ignorarla[32]. Tercero, dificulta la atribución de responsabilidad en caso de fallo: si no hay forma de saber a qué llevó al sistema a tomar una decisión incorrecta, determinar si el error fue del modelo, de los datos o del operador se convierte en algo enormemente complicado tanto a nivel técnico como jurídico[33].

Para intentar solucionar esto existe el campo de la Inteligencia Artificial Explicable (*Explainable AI*, XAI), que desarrolla técnicas orientadas a hacer los modelos más interpretables. Algunas de las más utilizadas son LIME y SHAP[34], que permiten identificar qué variables han tenido más peso en una decisión concreta. Sin embargo, estas técnicas tienen una limitación importante: ofrecen aproximaciones al razonamiento que ha hecho el sistema, no descripciones exactas de su funcionamiento interno. En aviación, esa diferencia importa mucho, porque una explicación aproximada no es suficiente para garantizar la seguridad ni para superar un proceso de certificación.

EASA reconoce este problema directamente en su *AI Roadmap 2.0*, donde incluye la explicabilidad como uno de los requisitos básicos de confiabilidad (*trustworthiness*) para cualquier sistema de IA en aviación, junto con la robustez y la trazabilidad. Esto confirma que la opacidad no es solo un reto técnico sino también un obstáculo regulatorio real, y que resolverlo es una condición necesaria para poder integrar la IA de forma segura en el ATM.

[32] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[33] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

[34] Adadi, A., y Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

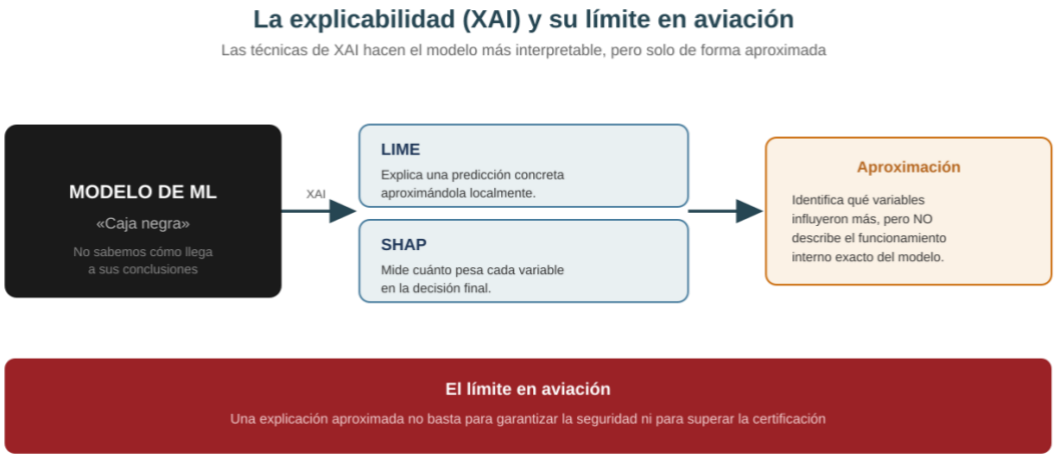


FIGURA 4. LAS TÉCNICAS DE XAI (LIME Y SHAP) Y SU LÍMITE EN AVIACIÓN. ELABORACIÓN PROPIA A PARTIR DE ADADI Y BERRADA (2018) Y EASA AI ROADMAP 2.0 (2023).

4.3. Limitaciones del Machine Learning [35]

El *machine learning* presenta limitaciones técnicas que condicionan su uso en entornos críticos, una de las más relevantes es su poca capacidad de generalización. Los modelos suelen funcionar correctamente en condiciones similares a las de su entrenamiento, pero pueden fallar cuando se enfrentan a situaciones nuevas o poco frecuentes, estas normalmente son situaciones de riesgo.

Otra limitación importante es la sensibilidad a los cambios en los datos, fenómeno conocido como *data drift*. Cuando el entorno operativo se modifica, por ejemplo, debido a cambios en el tráfico aéreo o en las condiciones meteorológicas, el rendimiento del modelo puede disminuir si no se actualiza como se debe.

Además, el ML requiere grandes volúmenes de datos de calidad para funcionar correctamente. En aviación, sin embargo, hay muy pocos datos de situaciones críticas o poco frecuentes por lo que el dificulta mucho entrenar el programa para estas situaciones. A este problema se le suma la dificultad para garantizar la seguridad del sistema ya que no siempre se puede demostrar que cumplirá con todos los requisitos en cada uno de los escenarios posibles.

Finalmente, existe el riesgo de sobreajuste[36], donde el modelo aprende en exceso los datos de entrenamiento y pierde capacidad de adaptación a situaciones reales. Todas estas limitaciones refuerzan la necesidad de mantener un enfoque en el que la IA actúe como sistema de apoyo y no como sustituto del operador humano. En este sentido, la Organización de Aviación Civil Internacional (OACI)[37] insiste en la importancia de garantizar la supervisión humana en todo momento.

---

[35] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[36] Adadi, A., y Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

[37] ICAO. (2025b). *Artificial intelligence (AI) contribution to aviation* (Nota de trabajo A42-WP/489 EX/220). International Civil Aviation

5. *Marco* Regulatorio y Certificación de Sistemas ATM

Cualquier sistema que tenga impacto sobre la seguridad operacional en aviación está sujeto a un conjunto de normas, estándares y procedimientos de certificación cuya finalidad es garantizar que su comportamiento es suficientemente seguro y predecible. El reto que plantea la IA es que sus características dificultan su adaptación en un marco regulatorio formado en base la lógica del software determinista.

5.1 Marco normativo europeo

La regulación de la aviación en Europa se construye sobre una base institucional muy sólida. La Agencia de la Unión Europea para la Seguridad Aérea (EASA) es la autoridad central encargada de garantizar la seguridad del sector, y actúa con el Reglamento (UE) 2018/1139, también conocido como el Reglamento de Base de EASA. Este reglamento le da la competencia para desarrollar normas técnicas y exigir que todos los sistemas con impacto en la seguridad aérea sean certificados antes de poder operar.[38]

Dentro del ATM, el marco normativo más específico es el Reglamento (UE) 2017/373, que regula los requisitos de los proveedores de servicios de navegación aérea, Air Navigation Service Provider (ANSP). Este reglamento obliga a estos proveedores a demostrar que sus sistemas y procedimientos cumplen los niveles de seguridad exigidos y que tienen implantados procesos necesarios de gestión del riesgo.[39]

A este marco normativo se ha sumado el Reglamento (UE) 2024/1689, el AI Act, que es la primera ley del mundo que regula la inteligencia artificial de forma vinculante. Lo relevante es que esta norma tiene carácter horizontal, es decir, se aplica a todos los sectores, incluida la aviación. Sus requisitos más exigentes entran en vigor en agosto de 2026, lo que significa que los sistemas de IA que ya se estén desarrollando ahora para ATM tendrán que cumplir tanto con la regulación aeronáutica como con el AI Act al mismo tiempo.[40]

Para ayudar a gestionar esta doble exigencia, EASA publicó en 2023 su AI Roadmap 2.0, donde anticipa el desarrollo de una nueva regulación específica para IA en aviación llamada Part-AI. Esta futura norma actuaría como puente entre el AI Act y la regulación aeronáutica existente, estableciendo requisitos concretos para organizaciones, autoridades y sistemas basados en IA.[41]

Organization. https://www.icao.int/sites/default/files/Meetings/a42/Documents/WP/wp_489_en.pdf

[38] Parlamento Europeo y Consejo de la UE. *Reglamento (UE) 2018/1139 — Reglamento de Base de EASA*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32018R1139

[39] Comisión Europea. *Reglamento (UE) 2017/373 — Proveedores ATM/ANS*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32017R0373

[40] Parlamento Europeo y Consejo de la UE. (2024). *Reglamento (UE) 2024/1689 [AI Act]*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[41] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

Lo que se puede concluir es que el marco existe y está en construcción, pero todavía no está completo. La Part-AI no está aprobada, el AI Act todavía no es plenamente aplicable y los estándares técnicos específicos para ML en aviación siguen desarrollándose. Esto crea una situación de cierta incertidumbre regulatoria para las organizaciones que quieran integrar IA en sistemas ATM hoy en día.

5.2 Estándares aplicables a software aeronáutico

Además de las normas con fuerza de ley que se han visto, la certificación en aviación se apoya en estándares técnicos desarrollados por organismos especializados como RTCA y EUROCAE[42]. Estos estándares no son obligatorios por sí mismos, pero en la práctica funcionan para demostrar que un sistema cumple los requisitos de seguridad exigidos por EASA, ya que son los medios de cumplimiento aceptados por la Agencia.

El estándar más importante para el software aeronáutico es el DO-178C (*Software Considerations in Airborne Systems and Equipment Certification*), conocido en Europa como ED-12C. Publicado en 2012, establece los objetivos que debe cumplir el proceso de desarrollo del software en aeronaves para poder ser certificado. Su elemento central es la clasificación del software en cinco niveles de criticidad, llamados DAL (*Design Assurance Level*): desde el nivel A, reservado para el software cuyo fallo podría causar un accidente catastrófico, hasta el nivel E, para el software sin impacto en la seguridad. Cuanto más alto es el nivel de criticidad, más estrictos son los requisitos de verificación que hay que cumplir.[43]

Sin embargo, el DO-178C está pensado para software embarcado en aeronaves. Los sistemas ATM son sistemas de tierra, y para ellos existe un estándar específico: el DO-278A. Este estándar es en gran medida equivalente al DO-178C pero adaptado a los sistemas CNS/ATM (*Communication, Navigation, Surveillance and Air Traffic Management*) y utiliza una escala de seis niveles de garantía (AL1 al AL6) en lugar de los cinco DAL del DO-178C. El AL1 es el más exigente y equivale aproximadamente al DAL A del DO-178C.[44]

[42] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[43] Rierson, L. (2013). *Developing safety-critical software: A practical guide for aviation software and DO-178C compliance*. CRC Press.

[44] RTCA/EUROCAE. (2011). *DO-278A/ED-109A: Software integrity assurance considerations for communication, navigation, surveillance and air traffic management (CNS/ATM) systems*. RTCA.

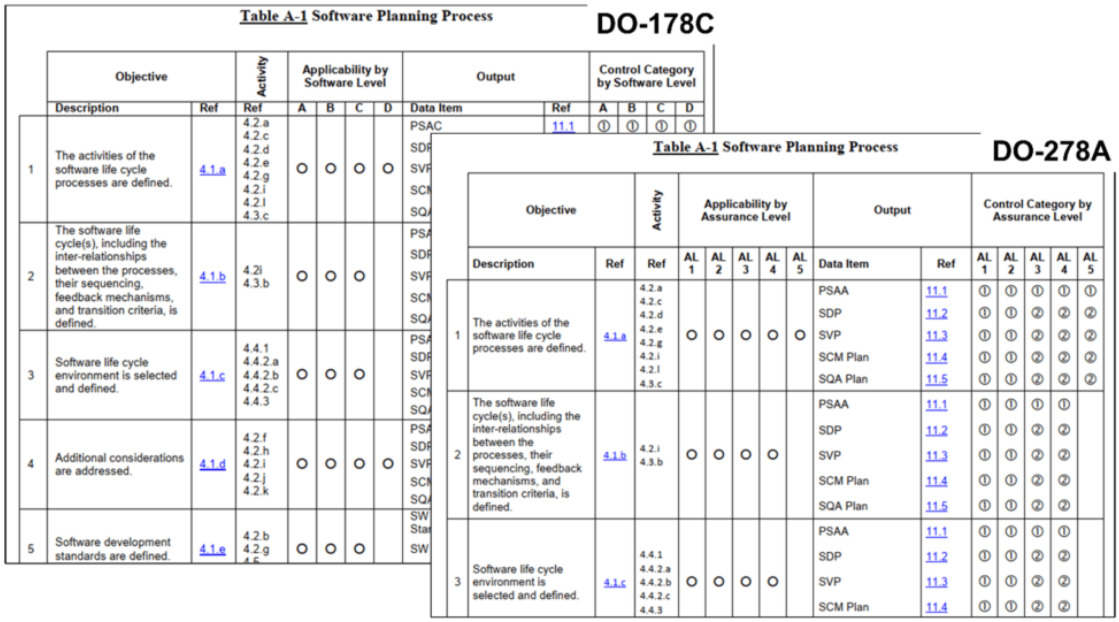


Table A-1 Software Planning Process — DO-178C

	Objective		Activity	Applicability by Software Level				Output		Control Category by Software Level			
	Description	Ref	Ref	A	B	C	D	Data Item	Ref	A	B	C	D
1	The activities of the software life cycle processes are defined.	4.1.a	4.2.a 4.2.c 4.2.d 4.2.e 4.2.g 4.2.i 4.2.l 4.3.c	O	O	O	O	PSAC	11.1	①	①	①	①
								SDP					
								SVP					
								SCM					
								SQA					
2	The software life cycle(s), including the inter-relationships between the processes, their sequencing, feedback mechanisms, and transition criteria, is defined.	4.1.b	4.2i 4.3.b	O	O	O		PSA					
								SDP					
								SVP					
								SCM					
								SQA					
3	Software life cycle environment is selected and defined.	4.1.c	4.4.1 4.4.2.a 4.4.2.b 4.4.2.c 4.4.3	O	O	O		PSA					
								SDP					
								SVP					
								SCM					
								SQA					
4	Additional considerations are addressed.	4.1.d	4.2.f 4.2.h 4.2.i 4.2.j 4.2.k	O	O	O	O	PSA					
								SDP					
								SVP					
								SCM					
								SQA					
5	Software development standards are defined.	4.1.e	4.2.b 4.2.g 4.6	O	O	O		SW					
								Star					
								SW					

Table A-1 Software Planning Process — DO-278A

	Objective		Activity	Applicability by Assurance Level					Output		Control Category by Assurance Level				
	Description	Ref	Ref	AL 1	AL 2	AL 3	AL 4	AL 5	Data Item	Ref	AL 1	AL 2	AL 3	AL 4	AL 5
1	The activities of the software life cycle processes are defined.	4.1.a	4.2.a 4.2.c 4.2.d 4.2.e 4.2.g 4.2.i 4.2.l 4.3.c	O	O	O	O	O	PSAA	11.1	①	①	①	①	①
									SDP	11.2	①	①	②	②	②
									SVP	11.3	①	①	②	②	②
									SCM Plan	11.4	①	①	②	②	②
									SQA Plan	11.5	①	①	②	②	②
2	The software life cycle(s), including the inter-relationships between the processes, their sequencing, feedback mechanisms, and transition criteria, is defined.	4.1.b	4.2.i 4.3.b	O	O	O	O		PSAA	11.1	①	①	①	①	
									SDP	11.2	①	①	②	②	
									SVP	11.3	①	①	②	②	
									SCM Plan	11.4	①	①	②	②	
									SQA Plan	11.5	①	①	②	②	
3	Software life cycle environment is selected and defined.	4.1.c	4.4.1 4.4.2.a 4.4.2.b 4.4.2.c 4.4.3	O	O	O	O		PSAA	11.1	①	①	①	①	
									SDP	11.2	①	①	②	②	
									SVP	11.3	①	①	②	②	
									SCM Plan	11.4	①	①	②	②	

FIGURA 5. COMPARATIVA DE LA TABLE A-1 (SOFTWARE PLANNING PROCESS) DEL DO-178C Y DEL DO-278A. FUENTE: RTCA/EUROCAE, DO-178C (2012) Y DO-278A (2011)

A nivel de sistema, por encima del software, el estándar de referencia es el ARP4754A, que define cómo debe gestionarse el desarrollo de un sistema complejo desde el punto de vista de la seguridad, incluyendo la asignación de niveles de criticidad y la gestión de requisitos desde las primeras fases del diseño.[45]

Lo importante de entender es que todos estos estándares están diseñados para un tipo de software que funciona de forma determinista: dado un conjunto de requisitos, se programa la lógica, se verifica que el código la cumple y se certifica. El problema es que esta lógica no aplica directamente a los sistemas basados en machine learning, ya que su comportamiento no viene del código sino de los datos con los que se han entrenado.

5.3 Concepto de certificación en aviación y principio de seguridad operacional

La certificación en aviación es el proceso mediante el cual EASA verifica que un sistema cumple los requisitos de seguridad necesarios para poder operar. No es una simple inspección puntual, sino un proceso que acompaña al sistema durante todo su ciclo: desde el diseño hasta las modificaciones posteriores, pasando por el desarrollo, las pruebas y la puesta en servicio.[46]

El principio que está detrás de todo esto es el de seguridad operacional o *safety*, que la OACI define como la reducción del riesgo de daños a personas o bienes hasta un nivel que se considera aceptable[47]. En aviación, ese nivel aceptable no es cero, porque la seguridad

[45] SAE International. (2010). *ARP4754A: Guidelines for development of civil aircraft and systems*. SAE International. https://www.sae.org/standards/content/arp4754a/

[46] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[47] ICAO. (2018). *Safety management manual* (Doc 9859, 4.ª ed.). International Civil Aviation Organization. https://www.icao.int/safety-management/SMI/SMM

absoluta no existe, pero sí se exige que la probabilidad de fallo catastrófico sea extremadamente baja. Para los sistemas más críticos, esa probabilidad no puede superar 10^{-9} por hora de vuelo, lo que en la práctica significa que un fallo de ese tipo no debería ocurrir más de una vez en mil millones de horas de operación.

Para alcanzar ese nivel de seguridad, el proceso de certificación se estructura en torno a varios elementos que deben estar documentados y auditados.[48] El primero es la evaluación de seguridad del sistema, donde se identifican todos los modos de fallo posibles, se analiza con qué probabilidad podrían ocurrir y qué consecuencias tendrían, y a partir de ahí se asigna el nivel de criticidad (DAL/AL) al que queda sujeto el software. El segundo elemento es la verificación y validación, que comprueba que el sistema hace lo que se supone que tiene que hacer y que no hace nada que no deba. En el software tradicional esto implica revisar el código, ejecutar pruebas y demostrar que todos los requisitos están cubiertos. El tercero es la trazabilidad, es decir, la capacidad de demostrar que cada parte del sistema responde a un requisito concreto y que no hay ningún requisito sin cubrir. El cuarto es la gestión de la configuración, que garantiza que el sistema que se ha certificado es exactamente el mismo que el que entra en operación, sin ninguna modificación no controlada.



FIGURA 6. FASES DEL PROCESO DE CERTIFICACIÓN DE SOFTWARE Y STAGES OF INVOLVEMENT (SOI). FUENTE: RTCA/EUROCAE, DO-178C (2012).

Todo este proceso tiene un denominador común: presupone que se puede describir con precisión qué tiene que hacer el sistema, comprobar que lo hace, y garantizar que seguirá haciéndolo en operación. Y eso es exactamente lo que los sistemas de machine learning ponen en cuestión, porque su comportamiento no se puede describir ni verificar de la misma manera. No hay unas reglas programadas que revisar, sino un modelo entrenado del que su razonamiento interno es opaco incluso para sus propios creadores.[49]

Esto no significa que la certificación de la IA sea imposible, pero sí que los métodos actuales no son directamente aplicables y que se necesitan enfoques nuevos.

---

[48] SAE International. (2010). *ARP4754A: Guidelines for development of civil aircraft and systems*. SAE International. https://www.sae.org/standards/content/arp4754a/

[49] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

5.4 La IA como sistema de alto riesgos

El AI Act clasifica los sistemas de IA según el riesgo que representan. Por lo que lo importante es saber qué aspectos implican a esta clasificación en sistemas de ATM.

El Anexo III del AI Act lista expresamente los ámbitos en los que los sistemas de IA se consideran de alto riesgo. Entre ellos se incluyen los sistemas utilizados en la gestión del tráfico y el transporte, lo que abarca directamente el ATM (AI Act, 2024, Anexo III).[50] Esto significa que cualquier sistema de IA que se quiera implantar en este entorno queda automáticamente sujeto a la norma más exigente, sin necesidad de ningún análisis adicional sobre si es o no suficientemente crítico.

Los requisitos que el AI Act impone a estos sistemas son bastante amplios y condicionan directamente cómo se puede desarrollar y desplegar la IA en ATM.[51] El artículo 9 exige que exista un sistema de gestión del riesgo documentado y actualizado durante todo el ciclo de vida del sistema, no solo en el momento de la certificación inicial. El artículo 10 establece requisitos específicos sobre la calidad de los datos utilizados para entrenar, validar y probar el modelo, exigiendo que sean representativos, que no tengan errores y no introduzcan sesgos que puedan afectar al rendimiento. Los artículos 11 a 13 obligan a mantener una documentación técnica completa que permita a las autoridades evaluar si el sistema cumple la normativa. El artículo 14 exige que el sistema esté diseñado para permitir una supervisión humana efectiva. Finalmente, el artículo 15 exige que el sistema mantenga niveles adecuados de precisión, robustez y ciberseguridad incluso ante errores o intentos de manipulación.

Un aspecto que conviene destacar y que no siempre se tiene en cuenta es que estos requisitos del AI Act no sustituyen a los que ya exige la regulación aeronáutica específica, sino que se suman a ellos. Un sistema de IA en ATM tendrá que cumplir simultáneamente con el DO-278A, con el Reglamento 2017/373, con el AI Act y con la futura Part-AI de EASA. Esto convierte el proceso de aprobación de estos sistemas en uno de los más complejos del panorama regulatorio europeo actual.[52]

Por su parte, la clasificación que hace EASA en su AI Roadmap 2.0 añade otra capa de análisis. Como ya se vio en el capítulo 3, EASA distingue entre sistemas de nivel 1 (asistencia al humano), nivel 2 (colaboración humano-máquina) y nivel 3 (automatización avanzada). Cuanto mayor es el nivel de autonomía del sistema, más profundos y exigentes son los requisitos de aseguramiento que aplican. La mayoría de los sistemas de IA que se están desarrollando actualmente para ATM se sitúan en el nivel 1 o en la transición hacia el nivel 2, precisamente porque los requisitos del nivel 3 todavía no tienen unos medios de cumplimiento suficientemente maduros para poder ser certificados.

Aquí aparece uno de los puntos de fricción más claros entre los dos marcos. El AI Act clasifica los sistemas de forma binaria: un sistema es de alto riesgo o no lo es, y si lo es, se le aplican todos los requisitos por igual. EASA, en cambio, gradúa los sistemas en niveles según cuánta

[50] Parlamento Europeo y Consejo de la UE. (2024). *AI Act*, Anexo III. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[51] Parlamento Europeo y Consejo de la UE. (2024). *AI Act*, arts. 9-15. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[52] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

autoridad tiene la IA, desde el nivel 1 de simple asistencia hasta el nivel 3 de automatización avanzada, y ajusta las exigencias a cada nivel. El resultado es que un sistema de nivel 1, que solo da recomendaciones, y uno de nivel 3, que actúa de forma casi autónoma, reciben del AI Act el mismo trato aunque su impacto sobre la seguridad sea muy distinto. Son dos lógicas de proporcionalidad que no terminan de encajar.

La mejor prueba de que este encaje no es automático es que EASA ha tenido que ponerse a construirlo. En noviembre de 2025 publicó la NPA 2025-07[53], su primera propuesta regulatoria para trasladar los requisitos de alto riesgo del AI Act al marco aeronáutico, que es el primer paso hacia la futura Part-AI. Si los dos marcos encajaran sin más, esta norma puente no sería necesaria. El hecho de que EASA esté desarrollándola en este momento confirma que entre el AI Act y la regulación aeronáutica hay huecos reales que todavía no están resueltos.

5.5 Retos de certificación específicos del Machine Learning[54]

Se ve con más claridad por qué integrar la IA en sistemas ATM es tan complicado desde el punto de vista regulatorio. No es solo una cuestión de trámites o burocracia, sino que hay un problema técnico de fondo: los métodos de certificación actuales no están diseñados para funcionar con sistemas que aprenden de datos.

Los estándares existentes no son suficientes

Los estándares de certificación como el DO-178C o el DO-278A presuponen que el comportamiento del software puede describirse de forma precisa mediante requisitos, verificarse a través del código y garantizarse para todos los casos posibles. En un sistema de ML esto no funciona así, porque el comportamiento del modelo no viene del código sino de los datos con los que se ha entrenado, y no existe una lista de reglas que revisar ni una lógica interna que se pueda auditar paso a paso.

EASA reconoce explícitamente este problema en su AI Roadmap 2.0, donde señala que los marcos de aseguramiento del desarrollo existentes no están adaptados para cubrir las particularidades de los sistemas de aprendizaje automático, y que hacen falta metodologías nuevas específicas para el ML.

El proceso en W y el concepto de Learning Assurance

Para dar respuesta a este problema, EASA ha desarrollado en su AI Concept Paper un nuevo marco de aseguramiento específico para sistemas de ML, conocido como proceso en W o *Learning Assurance*. La idea principal es que, a diferencia del proceso en V tradicional que va de los requisitos al código y del código a la verificación, el proceso en W añade un eje paralelo dedicado al aseguramiento de los datos, reconociendo que en el ML los datos son tan críticos como el propio código.[55]

---

[53] EASA. (2025). *Notice of Proposed Amendment (NPA) 2025-07: AI trustworthiness framework*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/newsroom-and-events/news/easas-first-regulatory-proposal-artificial-intelligence-aviation-now-open

[54] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[55] EASA. (2024a). *Concept paper: Guidance for level 1 & 2 machine learning applications — Issue 02*. European Union Aviation Safety

Dentro de este proceso, uno de los conceptos más importantes es el Dominio de Diseño Operacional o ODD (*Operational Design Domain*). El ODD define las condiciones y escenarios concretos para los que el sistema ha sido entrenado y dentro de los cuales su comportamiento puede considerarse suficientemente predecible. Fuera de ese dominio, el sistema puede comportarse de formas inesperadas y no hay garantías sobre su rendimiento. Esto tiene una implicación muy directa en ATM: si el sistema se enfrenta a una situación que no estaba dentro de su ODD, por ejemplo un escenario meteorológico poco frecuente o un patrón de tráfico inusual, no se puede garantizar que vaya a responder correctamente.

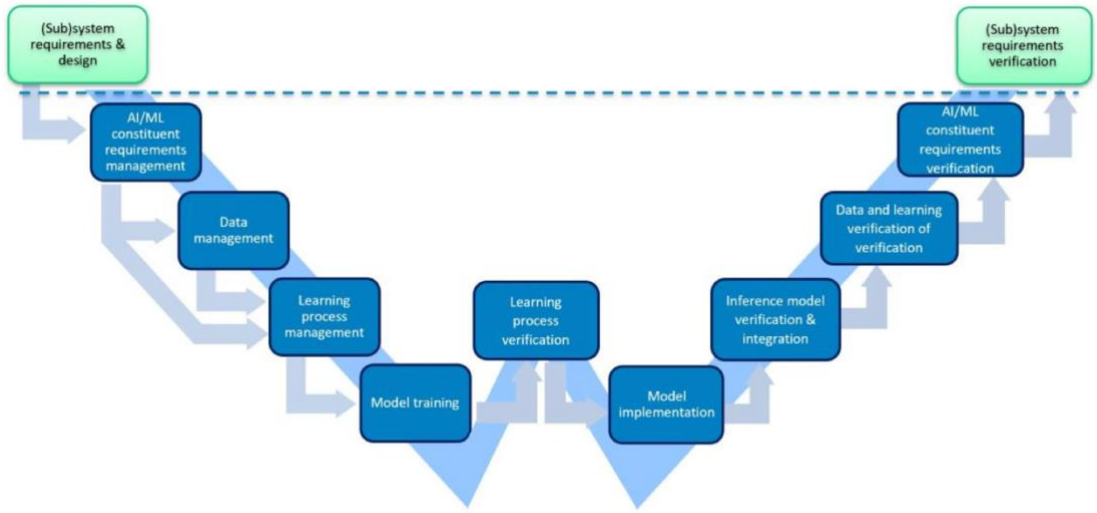


FIGURA 7. PROCESO DE LEARNING ASSURANCE EN FORMA DE W. FUENTE: EASA (2024)

El proyecto MLEAP y sus conclusiones

Para identificar métodos concretos que permitan cumplir los objetivos del proceso en W, EASA impulsó entre 2022 y 2024 el proyecto de investigación MLEAP (*Machine Learning Application Approval*), financiado por el programa Horizonte Europa. El proyecto fue desarrollado por un consorcio formado por Airbus Protect, LNE y Numalis, e investigó tres ejes técnicos principales: la representatividad de los datos de entrenamiento, las garantías del modelo y la robustez del algoritmo de aprendizaje.[56]

Las conclusiones del informe final son importantes porque muestran hasta dónde ha llegado la investigación y dónde siguen estando los límites. El proyecto identificó métodos útiles en los tres ejes, pero también reconoció que en varios aspectos los resultados no son todavía suficientemente sólidos para convertirse en medios de cumplimiento formalmente aceptados. En concreto, demostrar que el sistema funcionará correctamente fuera de sus datos de entrenamiento pero dentro del ODD, siguen siendo un área abierta que necesita más investigación.[57] Esto no es un detalle menor: significa que hoy por hoy no existe una forma completamente establecida de certificar que un modelo de ML va a funcionar bien en todas las situaciones para las que está previsto.

Agency. https://horizoneuropencpportal.eu/sites/default/files/2024-06/easa-concept-paper-guidance-for-level-1-and-2-machine-learning-applications-2024.pdf

[56] EASA. (2024b). *MLEAP D4 public report: Executive summary* (Issue 01). European Union Aviation Safety Agency. https://www.easa.europa.eu/sites/default/files/dfu/mleap-d4-public-report-executive_summary_expanded-issue01.pdf

[57] ERRIN. (2025). *Research objectives and MLEAP final report findings*. https://errin.eu/calls/research-help-shape-future-regulatory-framework-des

6. Human Oversight en Entornos ATC Asistidos por IA

6.1. Concepto de Human Oversight [58]

El *Human Oversight* o supervisión humana es definida en el Artículo 14 del Reglamento de Inteligencia Artificial de la Unión Europea que entra en vigor el 2 de agosto del 2026. El articulo establece que todos aquellos sistemas de inteligencia artificial que tengan un riesgo alto se tienen que diseñar para que su supervisión por humanos sea eficaz siempre que esté en uso.

Su finalidad es prevenir o reducir al mínimo todos los riesgos, como los que podrían ser la salud, la seguridad y los derechos fundamentales. En este sentido, la supervisión humana es un aspecto fundamental de control, es un complemento al resto de requisitos y regulaciones que ya tiene el propio sistema.

Esté artículo también determina que las medidas de supervisión humana deben ser proporcionales al nivel de riesgo de las actuaciones, la autonomía del sistema y su propio contexto de uso. Estas medidas están pensadas para garantizar que los operadores humanos tengan las capacidades suficientes y necesarias para poder tener un control sobre el sistema.

El reglamento también indica que los operadores que supervisan el sistema deben comprender las capacidades y limitaciones que tiene este para poder detectar anomalías o comportamientos inadecuados, además tienen que comprender correctamente los resultados generados por el sistema. Una parte fundamental es que deben ser conscientes del riesgo de sesgo de automatización, esta es una tendencia que se basa en confiar demasiado en los sistemas automatizados por lo que puede llegar a tomar malas decisiones.

La ley también exige que los operadores humanos tengan la capacidad de tener un control efectivo por tal de que sean capaces de no utilizar el sistema si es necesario, ignorar o modificar los resultados, intervenir en su funcionamiento o detenerlo. En algunos sistemas también es necesario una verificación adicional de al menos dos personas cualificadas antes de poder tomar decisiones basadas en el sistema.

En resumen, *Human Oversight* se define como un principio fundamental de la aplicación de la inteligencia artificial, para que en entornos muy automatizados, el control esté en todo momento en manos humanas.

6.2. Interacción humano-máquina[59]

La interacción humano-máquina no es la que era antes ya que en los softwares tradicionales los operadores trabajan con unos comandos predeterminados con unas respuestas

[58] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

[59] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

predefinidas. En cambio, los sistemas basados en la IA generan recomendaciones, predicciones, alertas… de forma autónoma por lo que la relación con el sistema es completamente distinta ya que el operador ya no usa una herramienta, sino que interactúa con el sistema. La diferencia principal es como se distribuye la responsabilidad en la toma de decisiones.

La interacción en ATM y el sistema se constituye básicamente en tres tipos de funciones: sistemas de apoyo a la decisión, sistemas de gestión de alertas y sistemas de optimización. Para todas estas situaciones, la IA actúa como intermediario entre los datos y el controlador, su función es filtrar, interpretar y presentar toda la información a tiempo real. Según como esté diseñado el sistema y el tipo de modelo humano-máquina tenga el sistema, mejores o peores van a ser los resultados.

Hay tres modelos que determinan los factores humanos de supervisión en los que se estructura esta relación. El modelo *Human-in-the-loop* requiere que el operador valide todas las acciones antes de que estas se ejecuten, de este modo se garantiza un control sobre cada una de las decisiones críticas. Por otro lado, el *Human-on-the-loop* permite que el sistema opere con más autonomía, pero siempre bajo supervisión de un operador que sea capaz de intervenir en todo momento. Finalmente, el *Human-in-command* determina que el operador siempre tiene la responsabilidad última y toda la autoridad en las decisiones sin tener en cuenta el grado de automatización del sistema. Este último es transversal y no puede ser delegado al sistema. [60]

El diseño de la interfaz es el elemento que hace posible esta relación. Una inteligencia artificial que ha sido correctamente programada puede ser ineficaz o peligroso si el controlador no puede comprender lo que está haciendo el sistema, el por qué lo hace y qué limitaciones tiene. Por esa razón, EASA y EUROCONTROL defienden que la mejor aplicación es la *human-centred design* para todos los sistemas de IA relacionados con la aviación. Según el EASA AI Roadmap 2.0 (2023) la inteligencia artificial se debe diseñar de manera que complemente las capacidades humanas y que no las sustituya. Por esta razón se deben crear interfaces comprensibles e interpretables fácilmente por el operador. Así mismo, el White Paper de EUROCONTROL[61] refuerza la idea, añaden que si el operador actúa como vigilante va a conducir a una pérdida de competencias y una reducción de la seguridad operacional.

La ley también respalda esta idea en el artículo 14 del Reglamento Europeo de Inteligencia Artificial (AI Act, 2024)[62] y lo transforma en obligaciones jurídicas: los sistemas de IA de alto riesgo se deben diseñar para que los operadores comprendan sus capacidades y limitaciones, para que puedan detectar anomalías e interpretar correctamente los resultados. La norma

[60] Comisión Europea, Grupo de Expertos de Alto Nivel en Inteligencia Artificial. (2019). *Directrices éticas para una IA fiable*. Comisión Europea. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[61] EUROCONTROL. (s.f.c). *Human factors integration in ATM system design: A white paper*. EUROCONTROL. https://www.eurocontrol.int/publication/human-factors-integration-atm-system-design

[62] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

también dicta que el operador debe tener siempre la capacidad de ignorar, revertir o detener el sistema.

La exigencia de esté diseño no es solo para reducir los propios riesgos que tiene la automatización sino también de posibilidad una dependencia del operador a las recomendaciones, también pueden provocar una pérdida de los conocimientos y habilidades de los controladores.

6.3.El rol del controlador aéreo en sistemas altamente automatizados

La automatización de las funciones ATM está transformando el perfil del controlador aéreo, y va más allá de la simple reducción de la carga de trabajo. Históricamente, el controlador ha tenido un control activo y directo del tráfico, gestionando la separación de aeronaves y la toma de decisiones tácticas en tiempo real. Los sistemas ATM actuales han modificado esto: muchas funciones han pasado a la máquina, por lo que el controlador ya no ejerce las mismas tareas ni necesita movilizar las mismas capacidades que antes.

Los tres modelos vistos en el apartado 6.2 describen cómo puede repartirse el control entre el humano y el sistema. En la práctica, el modelo que se está imponiendo en ATM combina varios de ellos bajo el concepto de human-machine teaming: el sistema opera con un grado de autonomía (on-the-loop), pero la autoridad última nunca se delega y permanece en el operador (in-command).

El modelo operativo actual se conoce como *human-machine teaming*, se basa en una cooperación con el sistema y el operador, en la que la IA analiza, predice y genera recomendaciones, por otro lado, el controlador supervisa, valida y gestiona decisiones o modificaciones. En este sistema, el controlador tiene toda la responsabilidad, pero sin tener tanta actividad sobre el tráfico, pasa a un plano de supervisión. EUROCONTROL (2024)[63] describe este cambio como una redefinición del rol del controlador hacia funciones de estrategia, gestión de la automatización y el sistema e intervención en situaciones imprevistas a tiempo real.

Este cambio no es simplemente de roles, sino que va más allá ya que el que supervisa no ejerce las mismas capacidades cognitivas que el que gestiona directamente el tráfico. En la supervisión hay una atención más difusa centrada en detectar anomalías en el sistema, por lo contrario, tener un control directo implica una atención más focalizada en su trabajo. Por esta razón, la supervisión pasiva es más vulnerable a cometer más errores[64].

[63] EUROCONTROL. (2024). *EUROCONTROL forecast update 2024–2030* (Autumn 2024). EUROCONTROL. https://www.eurocontrol.int/sites/default/files/2024-10/eurocontrol-seven-year-forecast-2024-2030-october-2024.pdf

[64] Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Cummings, M. L. (2004). Automation bias in intelligent time critical decision support systems. *AIAA 1st Intelligent Systems Technical Conference*. https://doi.org/10.2514/6.2004-6313

El principio *Human-in-command* determina que por mucho que haya cambiado el rol, no lo debe implicar la responsabilidad al sistema. El artículo 14 del AI Act (2024)[65] exige que los operadores puedan intervenir, revertir o detener el sistema en todo momento, es por esta razón que en todo momento tienen que mantener sus competencias para poder actuar de manera efectiva.

En un futuro la figura del controlador no será sustituida por la máquina, sino que el perfil de un controlador ya no será el mismo. Organizaciones como EUROCONTROL y IFATCA[66] coinciden en que las habilidades y procedimientos se están transformando hacia un perfil que pueda comprender y supervisar el sistema, pero además asumir el control en caso de que sea necesario y que en ese caso, de garantías de seguridad.

6.4. Riesgos del sesgo de automatización[67]

Definición y origen del concepto

El sesgo de automatización o también conocido como *automation bias* es una tendencia de los operadores a aceptar las decisiones de los sistemas como si todas fueran correctas y no tener la intención de buscar otra información o usar su propio criterio. Esto supone un grave problema en situaciones donde las decisiones son criticas como las que se encuentra un controlador aéreo.

Este fenómeno está muy relacionado con el *complacency* o complacencia el cual hace referencia a una actitud de poca vigilancia a los sistemas automatizados. En este caso no es que de como buenas todas las decisiones, sino que el operador comprueba de forma pasiva si el sistema funciona ya que da por sentado de que así es. Estas dos actitudes se pueden encontrar juntas y reforzarse mutuamente, esto genera un gran riesgo muy difícil de detectar.

Estos fenómenos no solo se encuentran en operadores con menos experiencia, se ha demostrado que incluso los profesionales muy cualificados y con gran experiencia pueden llegar a desarrollar estos efectos cuando trabajan durante largas jornadas. Esto supone un problema cuando el sistema falla ya que como el sistema tiene una alta fiabilidad, la confianza se vuelve excesiva y los controladores no se dan cuenta del error.

Manifestaciones concretas en entornos ATC

En el contexto del control del tráfico aéreo, el sesgo de automatización puede materializarse de dos formas: errores de omisión y errores de comisión.

Los errores de omisión son aquellos que se dan cuando un sistema no detecta o no alerta de un problema y el controlador tampoco por falta de vigilancia, es decir, al no tener la alerta del sistema el controlador interpreta que todo va bien. Estudios con simulaciones de alta fidelidad

[65] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

[66] IFATCA. (2024). *Technical and professional manual* (TPM, ed. 2024). International Federation of Air Traffic Controllers' Associations. https://ifatca.org/guidance-documents/

[67] Cummings, M. L. (2004). Automation bias in intelligent time critical decision support systems. *AIAA 1st Intelligent Systems Technical Conference*. https://doi.org/10.2514/6.2004-6313

han mostrado que cuando el sistema automatizado de detección de conflictos dejó de funcionar correctamente, un número significativamente menor de controladores detectó los conflictos en comparación con los que operaban sin automatización. Esto demuestra que los sistemas automatizados pueden reducir la capacidad de vigilancia de los operadores.

Los errores de comisión son los que ocurren cuando el controlador sigue las recomendaciones incorrectas del sistema, ignora otras fuentes que podrían dar el resultado correcto. Este es un claro caso de sesgo de automatización ya que el operador deja de tener un propio juicio. En un entorno ATC podría ocurrir que un controlador autorice una maniobra que el sistema ha recomendado sin asegurarse de si los datos son los correctos o si hay otros factores relevantes que el sistema no ha tenido en cuenta.

Estos errores son muy relevantes y sobre todo en ATC ya que las consecuencias de una mala acción son irreversibles y pueden afectar a la seguridad de más de una aeronave y a sus propios pasajeros.

Factores que agravan el riesgo en sistemas basados en ML

Los sistemas de IA basados en el *machine learning* añaden un riesgo relacionado con automatización de los sistemas. Tal y como se ha analizado anteriormente, los sistemas de ML dan recomendaciones basándose en la probabilidad y es muy difícil saber cuál es el razonamiento detrás de esta decisión. Esto tiene una grave consecuencia en caso de que el operador acepte la recomendación sin cuestionarla, esta confianza puede darse cuando el operador se encuentra fuera de su dominio de entrenamiento. [68]

Esta es la situación más peligrosa en entornos de ATC ya que un sistema ML entrenado bajo situaciones normales puede tener recomendaciones que pueden ser coherentes en situaciones excepcionales ya que el sistema no sabe reconocer cada una de las situaciones. El controlador no conoce los límites del sistema por lo que puede confiar en la recomendación sin saber que el modelo está trabajando más allá de su capacidad. Esta es una situación es más grave que en sistemas más deterministas ya que en caso de fallo es menos previsible y verificable que el otro.

Reconocimiento regulatorio del sesgo de automatización

La legislación europea ha incorporado explícitamente el concepto de sesgo de automatización en el marco normativo aplicable a los sistemas de IA de alto riesgo. El artículo 14(4)(b) del Reglamento (UE) 2024/1689, conocido como AI Act, establece que los operadores encargados de la supervisión de estos sistemas deben ser conscientes de la posible tendencia a depender o a depender excesivamente de los resultados producidos por el sistema, incluso cuando este proporciona información o recomendaciones para la toma de decisiones (AI Act, 2024, art. 14).[69]

---

[68] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency.* *https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

[69] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

Este reconocimiento normativo tiene un alcance importante: el sesgo de automatización ha dejado de ser un problema exclusivamente académico o técnico para convertirse en un riesgo que el propio legislador considera suficientemente relevante como para regularlo de forma expresa. En consecuencia, los sistemas de IA desplegados en entornos ATC no solo deben estar diseñados para minimizar técnicamente este riesgo, sino que los operadores humanos deben recibir formación específica para reconocerlo y contrarrestarlo activamente durante el ejercicio de sus funciones.

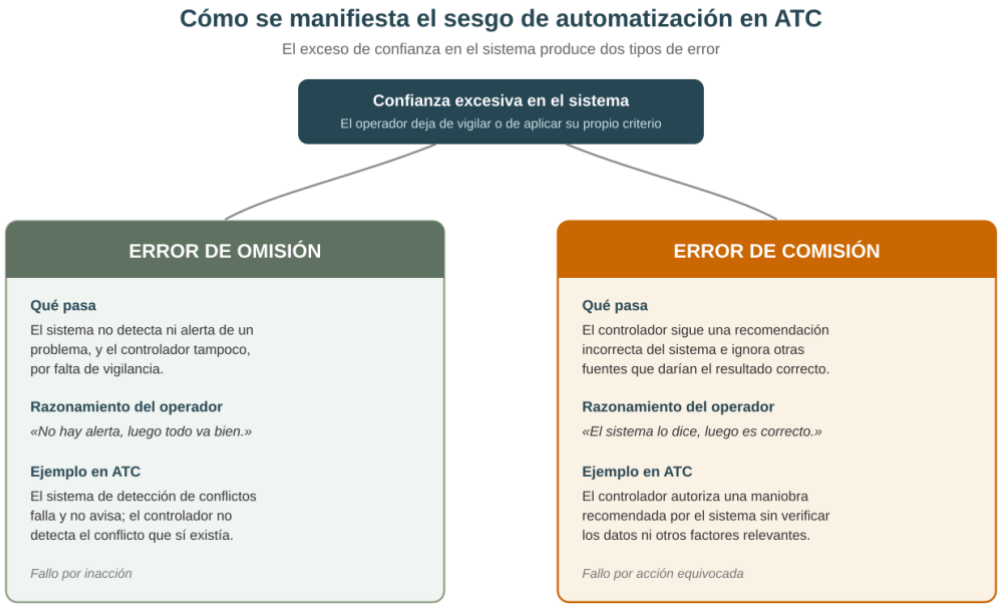


FIGURA 8. MANIFESTACIONES DEL SESGO DE AUTOMATIZACIÓN EN ATC: ERRORES DE OMISIÓN Y DE COMISIÓN. ELABORACIÓN PROPIA A PARTIR DE CUMMINGS (2004) Y EASA (2023)

6.5.Degradación de habilidades y pérdida de conciencia situacional

La transición del controlador hacia roles de supervisión pasiva genera dos fenómenos distintos que afectan directamente a la seguridad operacional: la pérdida conciencia situacional durante la monitorización y la erosión de las habilidades operacionales por falta de práctica activa. Aunque están relacionados, tienen causas, dinámicas y consecuencias diferentes y deben analizarse por separado.

Pérdida de conciencia situacional

El marco de referencia es el modelo de Endsley (1995)[70], que define la *situational awareness* (SA) como un proceso cognitivo de tres niveles: percepción de los elementos del entorno, comprensión de su significado, y proyección del estado del sistema en el futuro inmediato. Las pruebas muestran que la automatización afecta de manera diferente para cada uno de estos niveles. El *nivel 1* es la percepción: el operador detecta los elementos relevantes del entorno —posición de las aeronaves, altitudes, velocidades, alertas del sistema—. El *nivel*

[70] Endsley, M. R., y Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555

2 es la comprensión: el operador integra esos datos y entiende qué significan en relación con su objetivo, por ejemplo, si dos aeronaves están en ruta de colisión o si una desviación de rumbo supone un riesgo real. El *nivel 3* es la proyección: el operador anticipa cómo va a evolucionar la situación en los próximos minutos y actúa de forma preventiva antes de que el problema se materialice. Este tercer nivel es el que distingue a un operador experto de uno novato: no reacciona a lo que ya ha ocurrido, sino que se adelanta a lo que va a ocurrir (Endsley, 2001).

Este deterioro se documenta bajo el concepto de *out-of-the-loop* (OOTL) performance problem, definido por Endsley y Kiris (1995)[71] como la incapacidad del operador para retomar el control de forma eficaz cuando el sistema falla, consecuencia directa de haber pasado de un procesamiento activo a uno pasivo de la información. La evidencia en controladores aéreos es directa: Endsley y Rodgers (1998)[72] demostraron que los ATCOs mostraban un rendimiento inferior en la detección de conflictos cuando monitorizaban el sistema de forma pasiva, y Willems y Truitt (1999)[73] observaron respuestas más lentas y menor retención de información sobre el tráfico en las mismas condiciones. Investigaciones más recientes confirman además que la capacidad de atención sostenida se deteriora de forma medible en tareas de monitorización prolongada, con correlatos neurofisiológicos identificables que aumentan el riesgo de no detección de eventos críticos (Borghini et al., 2017).[74]

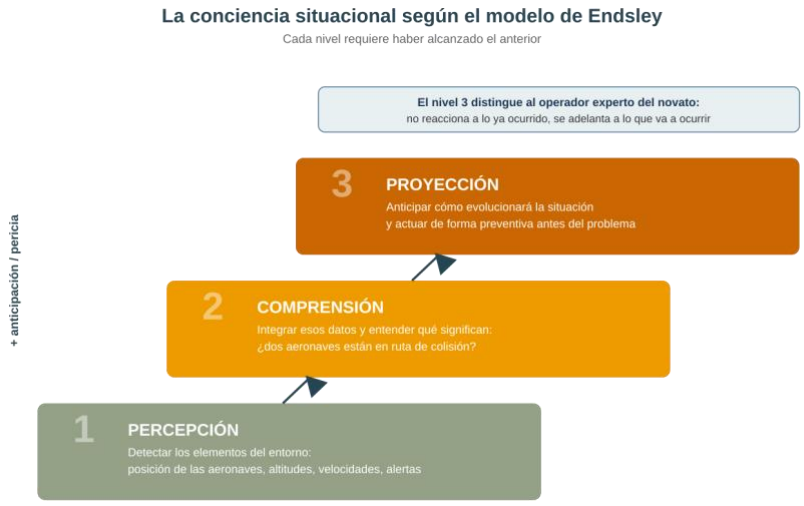


FIGURA 9. LOS TRES NIVELES DE LA CONCIENCIA SITUACIONAL APLICADOS AL CONTROL DE TRÁNSITO AÉREO. ELABORACIÓN PROPIA A PARTIR DE ENDSLEY (1995).

[71] Endsley, M. R., y Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555

[72] Endsley, M. R., y Rodgers, M. D. (1998). Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errors and automation. *Air Traffic Control Quarterly, 6*(1), 21–44. https://doi.org/10.2514/atcq.6.1.21

[73] Willems, B., y Truitt, T. R. (1999). *Implications of reduced involvement in en route air traffic control* (Informe técnico DOT/FAA/AM-99/19). Federal Aviation Administration Civil Aeromedical Institute.

[74] Borghini, G., Aricò, P., Di Flumeri, G., Cartocci, G., Colosimo, A., Bonelli, S., Golfetti, A., Imbert, J. P., Granger, G., y Benhacene, R. (2017). EEG-based cognitive control behaviour assessment: An ecological study with professional air traffic controllers. *Scientific Reports, 7*, 547. https://doi.org/10.1038/s41598-017-006

Degradación de habilidades

El *deskilling* es un fenómeno estructural de largo plazo: cuando un sistema automatizado asume el control de las tareas que antes ejecutaba el controlador, este deja de practicar estas competencias, por lo que tienden a deteriorarse con el tiempo. El White Paper de EUROCONTROL sobre Human Factors Integration in ATM System Design (2015)[75] señala que este proceso no solo implica pérdida de competencias técnicas, sino también una reducción del sentido de responsabilidad del operador sobre el resultado del sistema que supervisa. Este problema es especialmente crítico en ATM porque las situaciones que exigen intervención manual son precisamente las más excepcionales y complejas, y requieren exactamente las habilidades que la automatización ha tenido menos ocasión de mantener activas.

El círculo vicioso

La unión de ambos fenómenos configura un problema operacional: la automatización reduce simultáneamente la SA del controlador durante la monitorización y debilita sus habilidades por falta de práctica, de modo que el momento en que el sistema falla es el momento en que el operador se encuentra menos preparado para intervenir. Este círculo vicioso no es inevitable, pero sí previsible cuando el diseño de los sistemas no contempla mecanismos activos para mantener la implicación cognitiva del operador. El FLY AI Report (EUROCONTROL & SESAR JU, 2020)[76] identifica la necesidad de que la integración de la IA en ATM contemple el uso de las competencias humanas y que este sea considerado un requisito operacional imprescindible. Cuando esta condición no se cumple y el sistema falla, la determinación de responsabilidades adquiere una complejidad jurídica específica que merece un análisis propio.



FIGURA 10. EL CÍRCULO VICIOSO DE LA AUTOMATIZACIÓN EN ATC. ELABORACIÓN PROPIA A PARTIR DE ENDSLEY Y KIRIS (1995), EUROCONTROL (2015) Y EUROCONTROL & SESAR JU (2020).

[75] EUROCONTROL. (2015). *Human factors integration in ATM system design: A white paper*. EUROCONTROL. https://www.eurocontrol.int/publication/human-factors-integration-atm-system-design

[76] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

6.6. Requisitos para una supervisión humana efectiva

La supervisión humana en entornos ATC asistidos por IA no puede reducirse a una exigencia normativa formal. Como se ha expuesto al analizar los riesgos del sesgo de automatización, el deterioro de la conciencia situacional y la degradación de habilidades, garantizar que el controlador mantenga un control real sobre el sistema requiere condiciones operativas, técnicas y formativas concretas. Cumplir con la letra del artículo 14 del AI Act no equivale automáticamente a que la supervisión sea efectiva en la práctica. [77]

Transparencia y explicabilidad del sistema

Una supervisión efectiva exige que el operador comprenda qué está haciendo el sistema, por qué lo hace y cuáles son sus limitaciones. Sin esta comprensión, el controlador no puede evaluar críticamente las recomendaciones que recibe ni detectar errores con la antelación suficiente para intervenir. El EASA AI Roadmap 2.0 (2023)[78] identifica la transparencia y la explicabilidad como condiciones fundamentales para los sistemas de IA en aviación, añadiendo que estos deben ser predecibles, trazables y comprensibles para sus operadores. Esta exigencia es especialmente difícil de cumplir en los sistemas basados en machine learning, ya que la naturaleza estadística de los modelos hace que sus salidas no siempre puedan justificarse de forma intuitiva. Esto ha obligado a desarrollar mecanismos específicos de explicabilidad adaptados al contexto operacional del controlador, no solo documentación técnica.[79]

No basta con que el sistema sea comprensible en condiciones normales. El operador necesita información en tiempo real sobre el nivel de confianza del sistema, los márgenes de incertidumbre de sus predicciones y los límites que tiene. Sin esta información, el controlador no puede saber su grado de dependencia del sistema ni anticipar cuándo debe incrementar su nivel de atención o prepararse para asumir el control. [80]

Mantenimiento de la implicación cognitiva

Dado que la monitorización pasiva deteriora la conciencia situacional, una supervisión efectiva requiere que el diseño del sistema mantenga activo el pensamiento cognitivo del controlador. Esto puede lograrse mediante modelos de automatización adaptativa que ajustan el nivel de autonomía del sistema en función del estado del operador y de la situación del tráfico, evitando periodos de poca actividad que favorecen el fenómeno OOTL. El SESAR Joint Undertaking

---

[77] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial [Ley de Inteligencia Artificial]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

[78] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

[79] *EASA. (2023). Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20*

[80] Endsley, M. R., y Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555

(2023)[81] establece que los niveles de automatización en ATM deben diseñarse de forma que, cuando una situación supere el umbral de complejidad del sistema, este le de el control al operador con suficiente antelación y contexto para que la transición sea segura.

El White Paper de EUROCONTROL sobre Human Factors Integration in ATM System Design (2015)[82] afirma que mantener al operador cognitivamente activo no es solo una cuestión de diseño del sistema, sino de del modelo de trabajo: el controlador debe seguir teniendo tareas activas y significativas que requieran juicio, no simplemente confirmar lo que el sistema ya ha decidido. Reducir al operador a la función de validador pasivo de las decisiones no constituye una supervisión efectiva, sino una delegación de la responsabilidad.

Formación y mantenimiento de competencias

La degradación de habilidades requiere programas de formación específicamente orientados a mantener las competencias operacionales que el entorno automatizado no hace de forma natural. El FLY AI Report (EUROCONTROL & SESAR JU, 2020)[83] identifica la formación como uno de los aspectos clave para una integración segura de la IA en ATM, señala que los controladores deben estar preparados no solo para operar con el sistema en condiciones normales, sino para asumir el control en condiciones complicadas o ante fallos inesperados.

Esto implica, en la práctica, que los programas deben incorporar escenarios específicos de recuperación manual, ejercicios de detección de fallos del sistema y situaciones en las que la IA proporciona recomendaciones incorrectas que el controlador debe identificar y corregir. La formación orientada exclusivamente al uso del sistema no prepara al operador para el momento en que esta falla, que es precisamente el momento en que la supervisión humana debe demostrar su valor.

7. Responsabilidad Jurídica en Caso de Fallo de un Sistema IA en ATC

7.1 Principios generales de responsabilidad en aviación

En aviación hay múltiples actores, múltiples países y una normativa internacional muy específica que lleva décadas construyéndose.

El punto de partida de todo esto es el Convenio de Chicago de 1944, que es básicamente la "constitución" de la aviación civil internacional y que sigue vigente hoy en día. Entre otras cosas, establece que cada Estado tiene soberanía total sobre su espacio aéreo y que, por tanto, es su responsabilidad garantizar que los servicios de navegación aérea funcionen

[81] SESAR Joint Undertaking. (2023). *European ATM master plan: Automation roadmap* (ed. 2023). SESAR JU. https://www.sesarju.eu/node/4820

[82] EUROCONTROL. (2015). *Human factors integration in ATM system design: A white paper*. EUROCONTROL. https://www.eurocontrol.int/publication/human-factors-integration-atm-system-design

[83] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

correctamente. El artículo 28 del Convenio establece que cada Estado contratante se compromete a proveer en su territorio aeropuertos, servicios de radio, servicios meteorológicos y demás instalaciones para la navegación aérea para poder facilitar la navegación aérea internacional, de acuerdo con las normas y métodos establecidos en el propio Convenio. Por lo que, si hay control de tráfico aéreo en un país, ese país es el último responsable de que funcione bien. [84]

En Europa, esto se traduce en el Reglamento (UE) 2018/1139, que establece las normas comunes en el ámbito de la aviación civil y crea la EASA como organismo responsable de supervisar que todo se cumple. Cuando ocurre un accidente en aviación la investigación que se abre tiene como único objetivo aprender qué pasó para que no vuelva a ocurrir, no sirve para señalar culpables**.** Eso es algo completamente separado que se lleva por otra vía judicial o administrativa. [85]

Cuando hablamos de quién paga los daños si algo sale mal, la situación es bastante más complicada. Para las aerolíneas hay un convenio internacional, el Convenio de Montreal de 1999[86], que lo tiene todo bien definido: si un pasajero sufre un daño, la aerolínea responde, y hay límites de indemnización establecidos. Por otro lado, para los proveedores de servicios ATC, como ENAIRE en España, no existe nada parecido a nivel internacional. Cada país lo regula a su manera, y eso significa que, si un incidente ocurre en espacio aéreo que afecta a varios países, determinar quién asume qué responsabilidad se convierte en algo muy difícil de resolver.

Y aquí es donde entra la IA y donde el problema se complica todavía más. En el momento en el que entró la IA esté problema se complicó aún más. El marco que existe hoy en aviación está pensado para situaciones donde se puede seguir una cadena clara sobre una persona física. Pero cuando el fallo viene de un sistema de *machine learning*, esa cadena se rompe. Por ejemplo, si un sistema genera una recomendación incorrecta no se sabe de quien es la culpa ya que puede haber muchas variantes como el desarrollador, el proveedor, el que lo aprobó, el controlador… No se puede señalar a nadie directamente como único responsable y eso supone un grave problema.

[84] Convenio sobre Aviación Civil Internacional (Convenio de Chicago), 7 de diciembre de 1944, art.28. https://portales.segob.gob.mx/work/models/PoliticaMigratoria/CEM/UPM/MJ/II_30.pdf

[85] Reglamento (UE) 2018/1139 del Parlamento Europeo y del Consejo, de 4 de julio de 2018.https://www.boe.es/buscar/doc.php?id=DOUE-L-2018-81360

[86] Convenio de Montreal (1999). Ratificación en España: https://www.boe.es/buscar/doc.php?id=BOE-A-2004-9347

7.2 Responsabilidad del proveedor de servicios de navegación aérea (ANSP)

Actor	En qué se basa su responsabilidad	Normas clave
Proveedor de navegación aérea (ANSP)	Cómo despliega el sistema: formación de los controladores, protocolos de intervención manual y supervisión continua del rendimiento. Las negligencias organizativas son más fáciles de atribuir que el fallo del algoritmo.	Reglamento (UE) 2017/373 Ley 21/2003 de Seguridad Aérea AI Act, art. 26 (uso de sistemas de alto riesgo)
Fabricante / desarrollador	Defectos de origen y falta de actualización del modelo. Responsabilidad objetiva: basta probar el defecto y el daño, sin demostrar negligencia. El deterioro posterior del modelo le es imputable.	AI Act, arts. 16-20 Directiva (UE) 2024/2853 (productos defectuosos)
Controlador aéreo	Deber de supervisión activa e intervención ante riesgos, con independencia de lo que diga el sistema. Pero su responsabilidad queda condicionada si el ANSP no le dio formación, protocolos o autoridad suficientes.	Reglamento (UE) 2015/340 Ley 48/1960 de Navegación Aérea AI Act, art. 14
Estado	Doble posición: garante del servicio público de navegación aérea y autoridad certificadora. Responde por certificar lo que no debía o por no supervisar; su pasividad puede generar responsabilidad por omisión.	Convenio de Chicago (1944), art. 28 Ley 40/2015 (resp. patrimonial) AI Act, art. 74

TABLA 2 *Distribución de la responsabilidad jurídica entre los actores implicados ante el fallo de un sistema de IA en ATC. Elaboración propia.*

El principal actor en cuanto a proveedores de servicios de navegación aérea son los ANSP. En España sería ENAIRE, en el Reino Unido NATS, en Alemania DFS... Son las organizaciones que operan el sistema de control del tráfico aéreo en la práctica: deciden qué herramientas se usan, cómo se forman los controladores para usarlas y cómo se supervisa que todo funcione bien. Si un sistema de IA falla en ese entorno, el ANSP normalmente se encuentra dentro de está cadena de responsabilidad.

Desde el punto de vista de la normativa aeronáutica, los ANSPs europeos están sujetos al Reglamento de Ejecución (UE) 2017/373[87], de 1 de marzo de 2017, por el que se establecen requisitos comunes para los proveedores de servicios de gestión del tránsito aéreo y navegación aérea y otras funciones de la red de gestión del tránsito aéreo y su supervisión. Este

[87] Reglamento de Ejecución (UE) 2017/373 de la Comisión, de 1 de marzo de 2017:https://www.boe.es/buscar/doc.php?id=DOUE-L-2017-80442

reglamento les obliga a tener sistemas de gestión de la seguridad, a definir claramente quién es responsable de qué en cada función operacional, y a supervisar de forma continuada el rendimiento de los sistemas que tienen bajo su gestión. En España, la Ley 21/2003 de Seguridad Aérea[88] refuerza esto, atribuyendo al Ministerio la responsabilidad sobre el control de la circulación aérea y sobre los sistemas que se utilizan para prestarlo. Si se demuestra que el ANSP incumplió estos estándares y eso contribuyó al fallo, hay base para exigirle responsabilidad.

El AI Act.[89] El artículo 26 del Reglamento (UE) 2024/1689 regula las obligaciones de los ANSP cuando implantan un sistema de IA, no lo ha desarrollado, pero sí lo utiliza. Si una empresa utiliza un sistema de IA de alto riesgo, aunque no lo haya desarrollado, está obligada a utilizarlo conforme a las instrucciones de uso del proveedor y a asignar la supervisión humana a personas con la competencia, formación y autoridad necesarias. Además, deben vigilar el funcionamiento del sistema de IA de alto riesgo basándose en las instrucciones de uso, y esta obligación puede implicar informar a los proveedores, importadores, distribuidores y autoridades competentes e incluso suspender el uso del sistema. Todo esto entra en vigor de forma plena en agosto de 2026.

La responsabilidad del ANSP no depende solo de si el sistema falló técnicamente. Depende también de cómo lo desplegó. Si un incidente revela que el ANSP no formó bien a sus controladores para usar el sistema, que no había protocolos claros para intervenir manualmente o que nadie supervisaba de verdad cómo funcionaba en operación real, la responsabilidad es a causa de esas decisiones organizativas, no del fallo en sí. Eso es jurídicamente muy relevante porque esas negligencias son mucho más fáciles de atribuir y de demostrar que un error interno de un algoritmo.

Muchos ANSPs europeos son entidades públicas o semipúblicas. Eso significa que en algunos casos podría activarse el régimen de responsabilidad patrimonial del Estado por funcionamiento anormal de un servicio público, que tiene reglas procesales distintas a las de la responsabilidad civil.

7.3 Responsabilidad del fabricante o desarrollador del sistema

Si el ANSP es quien despliega el sistema, el desarrollador es quien lo construye. Y ahí es donde surge otra capa de responsabilidad igual de importante, si el sistema falla porque tiene un defecto de origen, la empresa que lo diseñó, lo entrenó y lo comercializó no puede quedarse al margen.

En el marco del AI Act[90], los desarrolladores tienen la figura de *proveedor* y están sujetos a un conjunto de obligaciones bastante exigentes. Los artículos 16 a 20 del Reglamento (UE) 2024/1689 les obligan a implantar un sistema de gestión de riesgos durante todo el ciclo de vida del sistema, mantener documentación técnica completa y actualizada y cooperar con las autoridades competentes. Esto no significa que incumplir estas reglas genere responsabilidad automáticamente, pero si ocurre un incidente y se demuestra que el desarrollador no cumplió

[88] Ley 21/2003, de 7 de julio, de Seguridad Aérea: https://www.boe.es/buscar/act.php?id=BOE-A-2003-13616

[89] AI Act, art. 26 (Reglamento (UE) 2024/1689): https://artificialintelligenceact.eu/es/article/26/

[90] AI Act, art. 26 (Reglamento (UE) 2024/1689): https://artificialintelligenceact.eu/es/article/26/

con estas obligaciones, eso tiene mucho peso como evidencia de que no actuó con la diligencia que se le exigía.

La pieza clave en términos de responsabilidad civil directa es la Directiva (UE) 2024/2853[91], aprobada en octubre de 2024. Esta Directiva moderniza el régimen anterior de 1985 y establece que la responsabilidad objetiva por daños causados por productos defectuosos se aplica a todos los bienes muebles, incluidos los programas informáticos. Los sistemas de IA quedan explícitamente incluidos, y sus desarrolladores quedan equiparados a fabricantes en cuanto a responsabilidad. La responsabilidad objetiva se basa en que no hace falta demostrar que el desarrollador actuó mal o que fue negligente. Basta con demostrar que el producto tenía un defecto y que ese defecto causó el daño.

Para los sistemas de *machine learning*: el desarrollador no se libra de responsabilidad porque el sistema funcionara bien al principio. Si el modelo se deteriora con el tiempo porque el entorno operativo y el desarrollador no hizo las actualizaciones necesarias para mantener la seguridad del sistema, ese deterioro le es imputable a él. Por lo que, en ATC, si un sistema de apoyo a la decisión empieza a dar recomendaciones incorrectas porque el tráfico aéreo ha cambiado y el modelo ya no está actualizado, el desarrollador no puede decir que cuando lo entregó funcionaba bien.

El mayor obstáculo para quien quiera reclamar al desarrollador sigue siendo la opacidad del sistema. Es muy difícil demostrar que lo que hizo el modelo en un momento concreto es un defecto y no simplemente una limitación de la que ya se tenía conocimiento. La Directiva (UE) 2024/2853 intenta abordar este obstáculo en el plano de la *prueba del defecto*: da a los tribunales la potestad de ordenar al fabricante la divulgación de evidencia técnica (datos de entrenamiento, documentación de validación, informes de rendimiento…) y establece que, si el desarrollador no cumple esa orden, el tribunal puede presumir directamente que el producto era defectuoso. Conviene retener esta distinción, porque esta presunción opera sobre el *defecto* del producto y no sobre la *relación causal* entre el defecto y el daño, que es un problema distinto y que, como se verá más adelante, el marco vigente resuelve peor.

7.4 Posible responsabilidad del controlador aéreo

El controlador aéreo es el último eslabón humano antes de que una decisión del sistema se convierta en una instrucción real en el espacio aéreo. Eso lo hace el actor más visible cuando algo falla y, históricamente, el primero sobre el que recaen las imputaciones. Pero en un entorno donde la IA lleva el peso de muchas decisiones, aplicar esa lógica sin matices puede ser profundamente injusto y jurídicamente problemático.

Para entender por qué, hay que tener claro qué se le exige a un controlador desde el punto de vista normativo. El Reglamento (UE) 2015/340[92] establece los requisitos técnicos y procedimientos administrativos relativos a las licencias y los certificados de los controladores de tránsito aéreo, con el objetivo de garantizar y mantener un nivel elevado y uniforme de seguridad. El controlador está obligado a supervisar activamente su espacio aéreo e intervenir cuando detecte riesgos, con independencia de lo que diga el sistema automatizado. La Ley

[91] Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024:https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81701
[92] Reglamento (UE) 2015/340 de la Comisión, de 20 de febrero de 2015: https://www.boe.es/buscar/doc.php?id=DOUE-L-2015-80440

48/1960[93] de Navegación Aérea española, establece que quien ejerce sobre el control del espacio aéreo no puede abandonarlo sin ser debidamente relevado. La responsabilidad operacional es personal.

Por lo que si el controlador aceptó sin verificar una recomendación incorrecta del sistema o no reaccionó a señales de riesgo que podía detectar, hay base para argumentar que incumplió con su deber.

En entornos de alta presión como la gestión del tráfico aéreo, los humanos tienden a seguir el resultado automatizado. A esto se debe tener en cuenta un deterioro en la capacidad de detectar anomalías durante tareas de supervisión pasiva prolongada. Cuando el sistema funciona bien durante mucho tiempo, el controlador se desconecta cognitivamente. El cerebro humano no está diseñado para mantenerse alerta de forma indefinida sin hacer nada. Y la ley no puede exigir que el controlador intervenga exactamente en el momento en que el sistema falla, que es precisamente cuando menos capacidad de reacción tiene.[94]

Y precisamente aquí conecta el AI Act[95]. El artículo 26 del Reglamento (UE) 2024/1689 obliga al ANSP a asignar la supervisión humana a personas con la competencia, formación y autoridad necesarias, así como el apoyo suficiente. Si el ANSP no cumplió con eso, la responsabilidad del controlador queda condicionada por ese incumplimiento. Esto no significa que el controlador quede automáticamente exonerado, su deber de diligencia profesional sigue existiendo, pero sí que cambia el peso de su responsabilidad en el conjunto de la cadena.

7.5 Responsabilidad del Estado como garante del sistema

El Estado no es un actor pasivo en la cadena de responsabilidad que se genera cuando un sistema de IA falla en un entorno ATC. Su implicación tiene dos dimensiones distintas que conviene analizar por separado: la de garante del servicio público de navegación aérea y la de autoridad certificadora de los sistemas que operan en ese entorno.

Desde la primera dimensión, el punto de partida es el artículo 28 del Convenio de Chicago de 1944,[96] que establece que cada Estado contratante asume el compromiso de proveer en su territorio los servicios de navegación aérea necesarios para facilitar la navegación aérea internacional. Esta obligación no desaparece cuando el Estado delega la prestación del servicio en un ANSP. El Estado sigue siendo el último responsable de que el sistema funcione correctamente, y esa posición de garante tiene consecuencias jurídicas directas cuando algo falla.

---

[93] Ley 48/1960, de 21 de julio, sobre Navegación Aérea: https://www.boe.es/buscar/act.php?id=BOE-A-1960-10905

[94] Htin, N. M. (2026). *Redefining the Standard of Human Oversight for AI Negligence*. Harvard Journal of Law & Technology Digest. https://jolt.law.harvard.edu/digest/redefining-the-standard-of-human-oversight-for-ai-negligence

[95] AI Act, art. 26 (Reglamento (UE) 2024/1689): https://artificialintelligenceact.eu/es/article/26/

[96] Organización de Aviación Civil Internacional. (1944). *Convenio sobre Aviación Civil Internacional* (Convenio de Chicago), art. 28. https://www.icao.int/publications/pages/doc7300.aspx

En el ordenamiento jurídico español esto se traduce en el régimen de responsabilidad patrimonial del Estado por funcionamiento de un servicio público, recogido en la Ley 40/2015 de Régimen Jurídico del Sector Público.[97] Este régimen establece que las administraciones públicas responden de los daños causados a los particulares como consecuencia del funcionamiento de los servicios públicos, con independencia de si ese funcionamiento fue normal o anormal. Lo relevante para este trabajo es que, si ENAIRE, que es una entidad pública empresarial, despliega un sistema de IA que falla y ese fallo causa un daño, el Estado no puede desvincularse completamente de la responsabilidad simplemente porque la decisión de implantar el sistema fuera del ANSP. La supervisión de que ese despliegue se hizo correctamente es también una obligación estatal.

La segunda dimensión es la del Estado como autoridad certificadora. En el marco europeo, EASA actúa como la autoridad central de certificación aeronáutica, pero son los Estados miembros quienes, a través de sus autoridades nacionales de aviación civil, supervisan que los sistemas operativos en su espacio aéreo cumplen los requisitos exigidos.[98] Esto significa que si un sistema de IA en ATM fue aprobado para operar sin cumplir realmente los requisitos del Reglamento (UE) 2017/373 o del AI Act, la autoridad que otorgó esa aprobación comparte responsabilidad en el resultado. No se trata de una responsabilidad por el fallo técnico del sistema sino por haber certificado algo que no debía haberse certificado, o por no haber detectado durante la supervisión continua que el sistema había dejado de cumplir los estándares exigidos.

El AI Act refuerza esta idea al establecer en su artículo 74[99] que los Estados miembros deben designar autoridades nacionales competentes para supervisar la aplicación del reglamento y darles los poderes necesarios para investigar incidentes, ordenar la retirada de sistemas y sancionar incumplimientos. Si una autoridad nacional no ejerce esa función de supervisión con la implicación exigida, su pasividad puede generar responsabilidad del Estado por omisión.

En definitiva, el Estado tiene responsabilidad estructural que deriva de su doble posición como garante del servicio público de navegación aérea y como autoridad supervisora de los sistemas que operan en ese entorno. Cuando un sistema de IA falla en ATC, la pregunta no es solo si el ANSP o el fabricante actuaron mal, sino también si el Estado cumplió con su obligación de garantizar que ese sistema nunca debería haber estado operando en condiciones inseguras.

7.6 Problemas de atribución de culpa en sistemas basados en ML

Uno de los problemas más complejos que plantea la IA en entornos ATC es determinar quién es responsable cuando algo falla. En los sistemas tradicionales esto es relativamente manejable: si un software falla se puede auditar el código, identificar el error y ver si fue un

[97] España. (2015). Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público. *Boletín Oficial del Estado*, 236. https://www.boe.es/buscar/act.php?id=BOE-A-2015-10566

[98] Parlamento Europeo y Consejo de la Unión Europea. (2018). Reglamento (UE) 2018/1139 del Parlamento Europeo y del Consejo, de 4 de julio de 2018, sobre normas comunes en el ámbito de la aviación civil [Reglamento de Base de EASA]. *Diario Oficial de la Unión Europea*, L 212, 1–122. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32018R1139

[99] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 74 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

problema de diseño, de desarrollo o de operación. Con los sistemas de ML esto no funciona así, y esa diferencia tiene consecuencias muy relevantes en lo jurídico.

El primer problema es la opacidad del modelo, desde el punto de vista de la responsabilidad, esta opacidad genera una dificultad muy práctica: cuando no es posible reconstruir por qué el sistema tomó una decisión concreta, demostrar que esa decisión fue el resultado de un defecto y no simplemente de una limitación conocida del sistema se convierte en algo muy difícil. Quien sufre el daño se encuentra en una posición muy complicada porque el desarrollador tiene acceso al modelo, a los datos de entrenamiento y a toda la documentación técnica, pero la víctima no tiene acceso a nada de eso. Es una asimetría de información muy pronunciada que el marco jurídico actual no resuelve bien.[100]

La Directiva (UE) 2024/2853 sobre responsabilidad por productos defectuosos intenta abordar parcialmente esta asimetría de información. Por un lado, permite que los tribunales ordenen al desarrollador que entregue la documentación técnica relevante (datos de entrenamiento, registros de rendimiento, informes de validación) cuando la víctima aporta indicios de que el producto podía tener un defecto. Por otro lado, y esto es lo más relevante, si el desarrollador se niega a entregar esa documentación, el tribunal puede presumir directamente que el producto era defectuoso. Ahora bien, esta herramienta resuelve la asimetría de información sobre el *defecto*, pero no elimina la dificultad de probar que ese defecto fue la causa concreta del daño, que en un sistema opaco sigue siendo el punto más difícil de acreditar.

El segundo problema es que en un sistema de ML en ATC hay muchos actores implicados al mismo tiempo. Como se ha visto en los apartados anteriores, cuando ocurre un incidente pueden estar en el centro del problema el desarrollador del modelo, el ANSP que lo desplegó, la autoridad que lo certificó, el controlador que aceptó la recomendación y el Estado como garante del servicio. Determinar qué parte de la responsabilidad corresponde a cada uno es enormemente complicado, sobre todo cuando el fallo no viene de un único error claro sino de la combinación de varios factores que por separado quizás no habrían generado ningún problema.

El tercer problema es uno de los más específicos del ML y que más dificultades genera: el deterioro del modelo con el tiempo. Un sistema que estaba bien certificado y funcionaba correctamente cuando se desplegó puede empezar a dar resultados incorrectos meses o años después, simplemente porque el entorno ha cambiado y el modelo ya no refleja bien la realidad. Por esta razón, el artículo 72 del AI Act[101] obliga a los desarrolladores a monitorizar el sistema después del despliegue y a notificar a las autoridades si detectan un riesgo grave, pero no concreta a partir de qué punto el deterioro del rendimiento obliga a actuar.

Frente a esta segunda dificultad, la de probar la causa, el instrumento que se había pensado no era la Directiva de productos defectuosos sino otra distinta: la propuesta de Directiva sobre responsabilidad en materia de IA de 2022. La diferencia entre las dos es importante. La Directiva (UE) 2024/2853 ayuda a demostrar que el producto tenía un defecto, mientras que

[100] Parlamento Europeo y Consejo de la Unión Europea. (2024). *Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos.* Diario Oficial de la Unión Europea. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024L2853

[101] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 72 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

esta propuesta ayudaba a demostrar que ese defecto fue el que causó el daño. Su idea era que, si se podía probar que el sistema incumplía los requisitos del AI Act y que ese incumplimiento tenía relación con el daño, el tribunal podía dar por hecho ese vínculo sin obligar a la víctima a demostrarlo de forma exhaustiva.

El problema es que esta propuesta fue retirada por la Comisión. Ya en su Programa de Trabajo para 2025 la incluyó en la lista de retiradas, alegando que no era previsible alcanzar un acuerdo entre los Estados miembros, y la retirada se formalizó mediante notificación publicada en el Diario Oficial de la Unión Europea el 6 de octubre de 2025. No es solo que la propuesta desapareciera, sino el motivo: la presión del sector tecnológico hacia la simplificación regulatoria y la falta de consenso sobre cómo abordar la responsabilidad por IA. Esto deja el obstáculo de la causalidad sin una herramienta pensada para resolverlo: la Directiva 2024/2853 cubre la parte del defecto, pero probar que ese defecto causó el daño en un sistema opaco vuelve a depender de las reglas generales de responsabilidad de cada país, que no se diseñaron para una tecnología como esta. Lejos de debilitar el argumento de este trabajo, la retirada lo refuerza: el propio legislador europeo identificó que las reglas generales no bastaban, diseñó un instrumento para resolverlo y acabó renunciando a él sin sustituirlo, dejando el vacío sin cubrir.[102]

En definitiva, los problemas de atribución en sistemas ML no son solo dificultades técnicas sino obstáculos reales que el marco normativo actual no resuelve completamente. La opacidad del modelo, la cantidad de actores implicados y el deterioro del rendimiento a lo largo del tiempo hacen que determinar responsabilidades en estos casos sea mucho más complicado que en cualquier otro tipo de sistema, y que las herramientas jurídicas necesarias para resolverlo todavía estén en construcción.

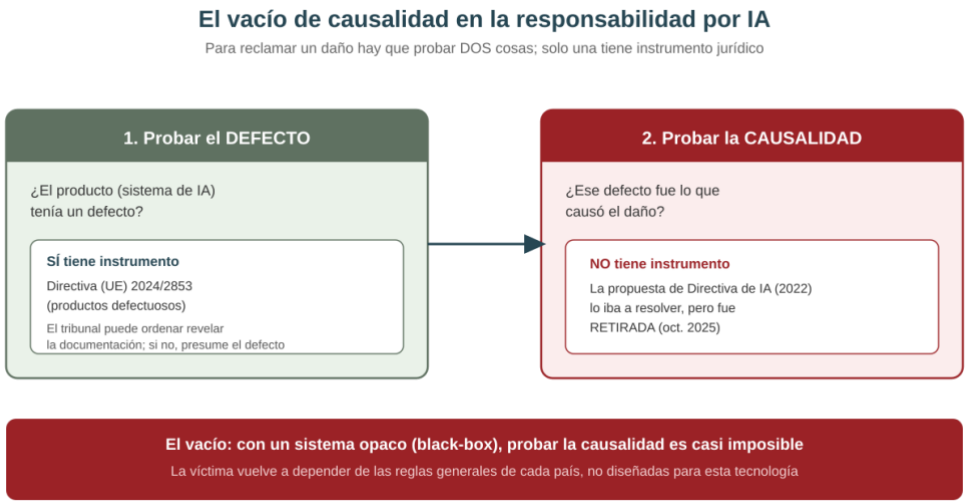


FIGURA 11. *EL VACÍO DE CAUSALIDAD EN LA RESPONSABILIDAD POR IA: LA DIRECTIVA (UE) 2024/2853 PERMITE PROBAR EL DEFECTO, PERO LA PROPUESTA DE DIRECTIVA SOBRE RESPONSABILIDAD EN MATERIA DE IA (2022), QUE CUBRÍA LA CAUSALIDAD, FUE RETIRADA EN OCTUBRE DE 2025. ELABORACIÓN PROPIA.*

[102] Comisión Europea. (2022). *Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial* [COM/2022/496 final]. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496

7.7 Escenarios hipotéticos de fallo y análisis jurídico

El análisis por actores desarrollado en los apartados anteriores permite identificar las obligaciones y responsabilidades de cada uno de forma individual. Sin embargo, en la práctica los incidentes no suelen tener una causa única ni un único responsable, sino que son el resultado de la combinación de varios factores que afectan a distintos actores al mismo tiempo. Para ilustrar cómo funcionaría el marco jurídico analizado en un caso concreto, se plantean a continuación tres escenarios hipotéticos que cubren las tres tipologías principales de fallo que pueden darse en un sistema de IA en ATC.

Escenario 1 — Fallo por data drift no detectado

Un sistema de apoyo a la decisión para la gestión de conflictos se instala en un centro de control en 2027 y funciona correctamente durante los primeros dieciocho meses de operación. A lo largo de 2028, el patrón de tráfico en el sector cambia de forma progresiva: la integración de nuevos operadores de drones de gran altitud y el aumento general de la densidad modifican las condiciones operativas para las que el modelo fue entrenado. El sistema empieza a generar recomendaciones subóptimas, pero no visiblemente erróneas, por lo que ni el ANSP ni los controladores detectan el deterioro. En un momento determinado, un controlador acepta una recomendación del sistema que resulta incorrecta y se produce una pérdida de separación entre dos aeronaves.

Desde el punto de vista jurídico, este escenario afecta a varios actores de forma simultánea. El fabricante del sistema tiene una responsabilidad clara por no haber implantado un sistema de monitorización post-despliegue adecuado. El artículo 72 del AI Act[103] obliga expresamente a los proveedores de sistemas de alto riesgo a recopilar y analizar datos sobre el rendimiento del sistema una vez en operación y a notificar a las autoridades competentes cuando detecten un riesgo grave. Si el fabricante no tenía ningún mecanismo para detectar que el modelo estaba deteriorándose, ese incumplimiento es difícilmente justificable. Además, la Directiva (UE) 2024/2853[104] establece que la responsabilidad objetiva del fabricante no desaparece porque el sistema funcionara bien en el momento de su entrega: si el deterioro posterior es consecuencia de no haber actualizado el modelo cuando era necesario, ese deterioro le es imputable.

El ANSP tiene también una responsabilidad importante en este escenario. El artículo 26 del AI Act[105] obliga a los operadores de sistemas de alto riesgo a supervisar el funcionamiento del sistema basándose en las instrucciones de uso del proveedor y a informar al fabricante si detectan anomalías. Si el ANSP no tenía procedimientos para evaluar de forma periódica si el rendimiento del sistema seguía siendo adecuado, o si tenía esos procedimientos, pero no los aplicó, hay base para exigirle responsabilidad por esa omisión organizativa. Esta es precisamente la situación más fácil de atribuir desde el punto de vista jurídico: no es el fallo

[103] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 72 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[104] Parlamento Europeo y Consejo de la Unión Europea. (2024). *Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por productos defectuosos*. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024L2853

[105] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 26 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

técnico del algoritmo lo que genera la responsabilidad del ANSP sino sus propias decisiones organizativas.

La responsabilidad del controlador en este escenario es más matizada. Si el sistema no daba señales visibles de estar funcionando mal, exigirle que detectara el deterioro de forma autónoma puede ser desproporcionado. Sin embargo, si existían indicios que un controlador diligente podría haber detectado como recomendaciones inusuales, patrones de alerta distintos a los habituales... su deber de supervisión activa sigue siendo exigible. En cualquier caso, el peso de la responsabilidad en este escenario recae principalmente sobre el fabricante y el ANSP, no sobre el controlador.

Escenario 2 — Fallo en zona fuera del ODD

En la primavera de 2029, una erupción volcánica obliga a cerrar parte del espacio aéreo europeo durante varios días. Los flujos de tráfico se reorganizan completamente: rutas habituales quedan inutilizadas, sectores que normalmente manejan poco tráfico se saturan y aparecen patrones de operación que no tienen precedente en los datos de entrenamiento del sistema. El modelo, entrenado exclusivamente para condiciones normales de tráfico, empieza a generar recomendaciones que los controladores de algunos sectores identifican como absurdas y deciden desactivar el sistema manualmente. Sin embargo, en otros sectores no hay un protocolo claro sobre qué hacer en esa situación y los controladores siguen utilizando el sistema porque nadie les ha indicado que deben pararlo. En uno de esos sectores se produce un incidente grave.

Este escenario ilustra de forma muy clara el concepto de Dominio de Diseño Operacional (ODD). El sistema estaba operando fuera de las condiciones para las que fue diseñado y entrenado, y en esas condiciones no hay ninguna garantía sobre su comportamiento.[106] La pregunta jurídica no es tanto si el sistema falló ya que es esperable que lo hiciera fuera de su ODD sino por qué seguía en operación en esas condiciones.

La responsabilidad en este escenario recae principalmente sobre el ANSP por una razón organizativa concreta: no tenía procedimientos claros de degradación del sistema para situaciones excepcionales. El artículo 26 del AI Act[107] obliga al operador a asignar la supervisión humana a personas con la competencia y autoridad necesarias para intervenir, lo que incluye la capacidad de detener el sistema cuando las condiciones operativas lo requieran. Si no existía un protocolo que indicara en qué circunstancias debía desactivarse el sistema y quién tenía autoridad para hacerlo, esa laguna organizativa es imputable al ANSP con independencia de lo que hiciera o dejara de hacer el controlador individual.

El fabricante también tiene una responsabilidad en este escenario, aunque más indirecta. La documentación técnica que el AI Act exige a los proveedores de sistemas de alto riesgo debe incluir una descripción clara del ODD del sistema, es decir, de las condiciones dentro de las

[106] EASA. (2024a). *Concept paper: Guidance for level 1 & 2 machine learning applications — Issue 02*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-concept-paper-guidance-level-1-2-machine-learning

[107] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 26 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

cuales su comportamiento puede considerarse fiable.[108] Si esa documentación existía y el ANSP no la tuvo en cuenta para diseñar sus procedimientos de degradación, la responsabilidad es del ANSP. Pero si la documentación no era suficientemente clara sobre los límites del sistema, el fabricante no puede desvincularse completamente de las consecuencias.

La responsabilidad del controlador individual que siguió usando el sistema en este escenario es la más difícil de evaluar. Si no había ningún protocolo que le indicara que debía pararlo, exigirle que tomara esa decisión de forma autónoma en medio de una situación de alta presión operativa puede ser desproporcionado. El deber de diligencia del controlador existe, pero está condicionado por el marco organizativo en el que opera: si ese marco falló, la responsabilidad individual queda significativamente mitigada.

Escenario 3 — Sesgo de automatización demostrable

Un controlador con más de diez años de experiencia está en su última hora de un turno de seis horas en un sector de alta densidad. Durante los meses anteriores, el sistema de apoyo a la decisión ha funcionado de forma impecable: sus recomendaciones han sido correctas en prácticamente todos los casos y el controlador ha desarrollado un alto nivel de confianza en el sistema. En ese contexto, el sistema genera una recomendación de separación que es claramente incorrecta ya que hay datos disponibles en pantalla que lo indican, pero el controlador, en un estado de fatiga cognitiva y con la confianza acumulada en el sistema, la acepta sin verificarla. Se produce una pérdida de separación.

Este escenario es el que más directamente conecta con el fenómeno del sesgo de automatización. La pregunta jurídica aquí no es si el sistema falló técnicamente, en este caso la recomendación era incorrecta pero el error era detectable, sino cómo se distribuye la responsabilidad entre el controlador que aceptó la recomendación, el ANSP que diseñó las condiciones de trabajo y el fabricante que diseñó el sistema.

La responsabilidad del controlador existe en este escenario. El Reglamento (UE) 2015/340[109] establece que el controlador tiene la obligación de supervisar activamente su espacio aéreo e intervenir cuando detecte riesgos, con independencia de lo que indique el sistema automatizado. Si los datos disponibles en pantalla permitían detectar el error, el deber de diligencia profesional del controlador exige que los hubiera revisado antes de aceptar la recomendación.

El ANSP tiene una responsabilidad organizativa clara en este escenario. Si el controlador llevaba seis horas de turno y estaba en un estado de fatiga cognitiva que reducía su capacidad de supervisión, la pregunta es si el ANSP había diseñado los turnos y las condiciones de trabajo de forma que ese riesgo estuviera contemplado y mitigado. El artículo 26 del AI Act[110] obliga al operador a asignar la supervisión humana a personas con la competencia y el apoyo suficientes

[108] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, arts. 11–13 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[109] Comisión Europea. (2015). Reglamento (UE) 2015/340 de la Comisión, de 20 de febrero de 2015, por el que se establecen requisitos técnicos y procedimientos administrativos relativos a las licencias y los certificados de los controladores de tránsito aéreo. *Diario Oficial de la Unión Europea*, L 63, 1–158. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32015R0340

[110] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 26 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

para ejercerla de forma efectiva. Un controlador fatigado en la última hora de un turno largo no está en las mejores condiciones para ejercer esa supervisión, y si el ANSP no tenía mecanismos para gestionar ese riesgo, esa omisión le es imputable.

El fabricante también tiene responsabilidad en este escenario, aunque por una razón distinta. El artículo 14(4)(b) del AI Act[111] establece que los sistemas de IA de alto riesgo deben diseñarse de forma que los operadores sean conscientes del riesgo de sesgo de automatización y que el sistema contemple mecanismos para contrarrestarlo. Si el sistema no tenía ningún mecanismo que forzara al controlador a verificar activamente las recomendaciones antes de aceptarlas, como por ejemplo, alertas de confirmación en situaciones de riesgo elevado o indicadores del nivel de confianza del modelo, ese diseño es cuestionable desde el punto de vista de los requisitos del AI Act. Un sistema que genera recomendaciones incorrectas sin ningún mecanismo que incentive la verificación activa del operador no cumple plenamente con la obligación de diseño para una supervisión humana efectiva.

Este tercer escenario es quizás el más relevante para el debate sobre responsabilidad en sistemas de IA, porque pone de manifiesto algo que los dos anteriores no muestran con tanta claridad: que la responsabilidad no siempre está en el fallo técnico del sistema sino en cómo ese sistema ha sido diseñado para interactuar con el ser humano, y en cómo el entorno operativo ha sido organizado para que esa interacción sea segura. Cuando ninguno de esos dos elementos funciona bien, determinar quién es el principal responsable es una cuestión que el marco normativo actual todavía no resuelve de forma completamente satisfactoria.



FIGURA 12 *TRES ESCENARIOS HIPOTÉTICOS DE FALLO DE UN SISTEMA DE IA EN ATC Y REPARTO DEL PESO DE LA RESPONSABILIDAD ENTRE LOS ACTORES IMPLICADOS. ELABORACIÓN PROPIA.*

[111] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 14(4)(b) [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

8. Hacia un Marco Integrado: Certificación, Supervisión y Responsabilidad

8.1 Necesidad de adaptación del modelo regulatorio actual

A lo largo de los capítulos anteriores ha ido quedando claro que el problema no es solo técnico. La IA plantea retos reales para los que el marco regulatorio actual no estaba preparado, y la respuesta no puede ser simplemente aplicar las mismas reglas de siempre a una tecnología que funciona de forma completamente diferente.

El modelo regulatorio que existe hoy en aviación se construyó pensando en sistemas deterministas: sistemas que hacen exactamente lo que se les programa, que se pueden verificar de forma exhaustiva antes de entrar en operación y cuyo comportamiento no cambia una vez certificados. Los estándares como el DO-278A o el DO-178C[112] presuponen que se puede describir con precisión qué tiene que hacer el sistema, comprobar que lo hace y garantizar que seguirá haciéndolo. Como se ha visto, esa lógica no funciona con el ML, y ese desajuste es el origen de la mayoría de los problemas que se han analizado.

El AI Act[113] ha supuesto un paso importante porque por primera vez introduce una regulación horizontal que reconoce que la IA es diferente y que necesita reglas propias. Sin embargo, su aplicación en aviación genera una doble exigencia que todavía no tiene una solución clara: los sistemas de IA en ATM tienen que cumplir al mismo tiempo con la regulación aeronáutica específica y con el AI Act, y entre ambos marcos hay lagunas y contradicciones que ninguno de los dos resuelve por sí solo. Como se vio, esto se manifiesta por ejemplo en que el AI Act trata el riesgo de forma binaria mientras que EASA lo gradúa por niveles, de modo que sistemas con un impacto muy distinto sobre la seguridad acaban sujetos a las mismas exigencias. EASA es consciente de este problema y por eso está desarrollando la futura Part-AI,[114] que actuaría como puente entre los dos marcos, pero mientras esa norma no esté aprobada, las organizaciones que quieran integrar IA en sistemas ATM se mueven en un entorno de incertidumbre regulatoria considerable.

A esto se suma que el modelo tradicional de certificación es esencialmente estático: se certifica el sistema en un momento concreto y esa certificación vale para toda su vida útil, salvo modificaciones significativas. Para el ML esto no tiene sentido, porque como se ha visto con el fenómeno del data drift, un sistema que estaba correctamente certificado puede dejar de serlo

[112] RTCA/EUROCAE. (2012). *DO-178C/ED-12C: Software considerations in airborne systems and equipment certification*. RTCA.
RTCA/EUROCAE. (2011). *DO-278A/ED-109A: Software integrity assurance considerations for communication, navigation, surveillance and air traffic management (CNS/ATM) systems*. RTCA.
[113] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689
[114] *EASA's first regulatory proposal on Artificial Intelligence for Aviation is now open for consultation | EASA*. (2025, 10 noviembre). EASA. https://www.easa.europa.eu/en/newsroom-and-events/news/easas-first-regulatory-proposal-artificial-intelligence-aviation-now-open

meses después simplemente porque el entorno ha cambiado. El marco regulatorio necesita incorporar una lógica de supervisión continua que hoy prácticamente no existe.[115]

La Directiva (UE) 2024/2853[116] sobre responsabilidad por productos defectuosos, ya aprobada y que equipara los sistemas de IA a cualquier otro producto a efectos de responsabilidad objetiva, supuso un paso en la dirección correcta. La propuesta de Directiva específica sobre responsabilidad en materia de IA, que habría introducido una presunción de causalidad pensada precisamente para los casos de opacidad, fue en cambio retirada por la Comisión en 2025[117] al no alcanzarse un acuerdo entre los Estados miembros. El resultado es un marco que avanza en lo relativo al producto defectuoso pero que deja sin resolver el problema probatorio específico de la IA, de modo que sigue siendo un marco incompleto para un problema que todavía no tiene una solución normativa definitiva.

En definitiva, lo que se necesita no es simplemente actualizar alguna norma concreta sino repensar de forma más profunda cómo se regula, certifica y supervisa un tipo de sistema que no se parece a nada para lo que el marco actual fue diseñado.

8.2 Certificación continua y monitorización post-despliegue

Uno de los cambios más importantes que necesita el modelo regulatorio actual es pasar de una lógica de certificación puntual a una lógica de supervisión continua. Como se ha visto a lo largo de este trabajo, el problema con los sistemas de ML no es solo conseguir que funcionen correctamente en el momento en que se certifican, sino garantizar que siguen funcionando correctamente una vez que están en operación real.

El modelo tradicional de certificación aeronáutica funciona básicamente así: se desarrolla el sistema, se verifica que cumple los requisitos, se certifica y se pone en operación. A partir de ese momento, salvo que se introduzcan modificaciones significativas, la certificación sigue siendo válida. Para el software determinista esto tiene sentido porque un sistema que funciona correctamente hoy va a seguir funcionando igual mañana, siempre que no se cambie nada. Para el ML no funciona así, porque el entorno operativo cambia y el modelo puede deteriorarse, aunque nadie haya tocado el código.[118]

Lo que se necesita es un enfoque de certificación continua que acompañe al sistema durante toda su vida útil, no solo en el momento inicial. Esto implica varias cosas concretas. En primer lugar, que el fabricante tenga obligación de monitorizar el rendimiento del sistema de forma sistemática una vez desplegado, comparando continuamente cómo se comporta en operación

---

[115] EASA. (2024a). *Concept paper: Guidance for level 1 & 2 machine learning applications — Issue 02*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-concept-paper-guidance-level-1-2-machine-learning

[116] Parlamento Europeo y Consejo de la Unión Europea. (2024). *Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por productos defectuosos*. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024L2853

[117] Comisión Europea. (2022). *Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial* [COM/2022/496 final]. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496

[118] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

real con cómo se esperaba que se comportara según su ODD. El artículo 72 del AI Act[119] ya establece esta obligación de forma general, pero para que sea efectiva en aviación necesita traducirse en métricas concretas, umbrales de alerta definidos y procedimientos claros sobre qué hacer cuando el rendimiento empieza a deteriorarse.

En segundo lugar, los ANSPs necesitan tener sus propios mecanismos de evaluación del rendimiento de los sistemas que operan, independientes de los del fabricante. No es seguro que la única fuente de información sobre si el sistema sigue funcionando bien sea el propio desarrollador. El Reglamento (UE) 2017/373[120] obliga a los ANSPs a tener sistemas de gestión de la seguridad, y esos sistemas deberían incluir de forma explícita procedimientos de evaluación continua del rendimiento de los sistemas de IA desplegados.

En tercer lugar, hace falta definir con claridad qué pasa cuando se detecta que el rendimiento del sistema ha empeorado. Hoy no existe un procedimiento estándar para gestionar esa situación, la ausencia de soluciones claras es precisamente lo que generó el problema del escenario 2 del apartado anterior, donde el sistema seguía en uso en algunos sectores porque nadie sabía qué hacer.

El proyecto MLEAP[121] identificó esta necesidad como una de las áreas prioritarias de investigación, reconociendo que los métodos actuales para demostrar que un modelo de ML va a seguir funcionando correctamente dentro de su ODD no son todavía suficientemente sólidos. Mientras esos métodos no estén disponibles, la monitorización continua post-despliegue es la única forma razonable de gestionar el riesgo residual que implica operar con sistemas cuyo comportamiento no puede garantizarse de forma exhaustiva en el momento de la certificación inicial.

En definitiva, la certificación continua no es una complicación adicional sino una consecuencia lógica de las características propias del ML. Un sistema que aprende de datos y cuyo rendimiento depende del entorno operativo necesita ser supervisado de forma permanente, y el marco regulatorio tiene que reflejar esa realidad de forma mucho más concreta de lo que lo hace hoy.

8.3 Redefinición del rol del Human Oversight

Si hay un concepto que atraviesa todo este trabajo de principio a fin es el de la supervisión humana. Aparece en la regulación, en el análisis de riesgos, en el capítulo de responsabilidad y en los tres escenarios hipotéticos.

---

[119] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 72 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[120] Comisión Europea. (2017). Reglamento de Ejecución (UE) 2017/373 de la Comisión, de 1 de marzo de 2017. *Diario Oficial de la Unión Europea*, L 62, 1–126. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32017R0373

[121] EASA. (2024b). *MLEAP D4 public report: Executive summary (Issue 01)*. European Union Aviation Safety Agency. https://www.easa.europa.eu/sites/default/files/dfu/mleap-d4-public-report-executive_summary_expanded-issue01.pdf

El artículo 14 del AI Act[122] establece que los sistemas de alto riesgo deben diseñarse para permitir una supervisión humana efectiva, y que los operadores deben tener la capacidad de intervenir, revertir o detener el sistema en todo momento. Es un punto de partida necesario, pero como se ha argumentado, cumplir con la letra de ese artículo no equivale automáticamente a que la supervisión sea real y efectiva en la práctica. Un controlador que tiene técnicamente la capacidad de detener el sistema pero que lleva meses sin necesitar hacerlo, que ha desarrollado una confianza elevada en sus recomendaciones y que está en la sexta hora de un turno exigente, no está en condiciones de ejercer una supervisión genuinamente crítica, aunque formalmente tenga todos los medios para hacerlo.

Lo que esto pone de manifiesto es que el Human Oversight no puede entenderse únicamente como una característica técnica del sistema sino como una condición operativa que depende de factores humanos, organizativos y de diseño que van mucho más allá de lo que cualquier reglamento puede especificar por sí solo. Redefinir el rol del Human Oversight significa reconocer esa complejidad y traducirla en requisitos concretos en tres niveles distintos.

El primer nivel es el del diseño del sistema. Para que la supervisión humana sea real, el sistema tiene que estar diseñado para mantener activa la implicación cognitiva del controlador, no para reducirla al mínimo. Esto significa que las interfaces deben proporcionar información sobre el nivel de confianza del modelo y los márgenes de incertidumbre de sus predicciones, que el sistema debe incluir mecanismos que fuercen la verificación activa en situaciones de riesgo elevado, y que los niveles de automatización deben ajustarse de forma dinámica en función de la complejidad de la situación.[123] El EASA AI Roadmap 2.0[124] defiende explícitamente el enfoque de *human-centred design* como requisito para todos los sistemas de IA en aviación, precisamente porque un sistema bien diseñado desde el punto de vista técnico puede ser ineficaz o incluso peligroso si el controlador no puede comprender lo que está haciendo ni por qué.

El segundo nivel es el organizativo. Como se ha visto en el escenario 3 del apartado anterior, la capacidad de supervisión del controlador no depende solo de él sino también de cómo el ANSP ha organizado el trabajo. Los turnos, la carga de trabajo, los procedimientos de degradación del sistema y los protocolos de intervención manual son decisiones organizativas que determinan en la práctica si la supervisión humana es posible o no. El FLY AI Report[125] señala que mantener las competencias operacionales de los controladores debe tratarse como un requisito operacional, no como una cuestión secundaria de formación, y que los ANSPs tienen

[122] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 14 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[123] EUROCONTROL. (2015). *Human factors integration in ATM system design: A white paper*. EUROCONTROL. https://www.eurocontrol.int/publication/human-factors-integration-atm-system-design

[124] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

[125] EUROCONTROL & SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

la responsabilidad de garantizar que sus operadores están en condiciones de asumir el control en cualquier momento.

El tercer nivel es el de la formación. La degradación de habilidades analizada no es inevitable, pero sí es previsible cuando los programas de formación no están diseñados para contrarrestarla. Los controladores necesitan entrenamiento específico no solo para operar con el sistema en condiciones normales sino para detectar sus fallos, reconocer el sesgo de automatización en su propio comportamiento y asumir el control manual en situaciones complejas. Como señalan Endsley y Kiris (1995),[126] el problema del out-of-the-loop no aparece de repente cuando el sistema falla: es el resultado de un proceso gradual de desconexión cognitiva que puede prevenirse si el diseño del trabajo y la formación están orientados a mantener al operador genuinamente implicado.

En definitiva, redefinir el Human Oversight significa pasar de entenderlo como un requisito formal que se cumple poniendo un operador delante de una pantalla, a entenderlo como una condición operativa que requiere un diseño adecuado del sistema, una organización del trabajo que lo haga posible y una formación que lo mantenga activo. Solo cuando esas tres condiciones se cumplen al mismo tiempo puede hablarse de una supervisión humana que añade valor real a la seguridad del sistema.



FIGURA 13. LAS TRES CONDICIONES DE UN HUMAN OVERSIGHT EFECTIVO. ELABORACIÓN PROPIA A PARTIR DEL ART. 14 DEL AI ACT, EASA AI ROADMAP 2.0 (2023), EUROCONTROL & SESAR JU (2020) Y ENDSLEY Y KIRIS (1995).

[126] Endsley, M. R., y Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555

8.4 Reparto equilibrado de la responsabilidad

Cuando un sistema de IA falla en un entorno ATC, la responsabilidad raramente es de un único actor. La cadena que va desde el desarrollador del modelo hasta el controlador que acepta la recomendación pasa por el ANSP que desplegó el sistema, la autoridad que lo certificó y el Estado que garantiza el servicio. Cada uno de esos actores tiene obligaciones concretas, y cuando alguna de ellas no se cumple, contribuye al riesgo que finalmente se materializa en el incidente.

El problema es que el marco jurídico actual no está bien preparado para gestionar esa responsabilidad compartida. La opacidad de los modelos de ML dificulta enormemente demostrar qué falló exactamente y por qué, y las reglas generales de responsabilidad civil no están diseñadas para situaciones en las que el daño es el resultado de la combinación de decisiones técnicas, organizativas y operativas tomadas por actores distintos en momentos distintos.

Lo que se necesita es un modelo de reparto de responsabilidad que sea proporcional a la capacidad de cada actor para haber prevenido el fallo. Esto no significa que todos respondan por igual, sino que cada uno responde en la medida en que incumplió las obligaciones que le correspondían. El fabricante responde por los defectos del modelo y por no haber monitorizado su rendimiento post-despliegue.[127] El ANSP responde por cómo desplegó el sistema, cómo formó a sus controladores y cómo organizó la supervisión.[128] El controlador responde por su deber de diligencia profesional, pero condicionado por el marco organizativo en el que opera.[129] Y el Estado responde por su función de garante y certificador, cuando esa función no se aplicó con la exigencia suficiente.[130]

Un elemento clave para que este reparto funcione en la práctica es la trazabilidad. Si no existe documentación suficiente sobre cómo se desarrolló el modelo, cómo se entrenó, qué límites tiene y cómo se monitorizó su rendimiento, determinar qué actor incumplió qué obligación se convierte en algo prácticamente imposible. Por eso los requisitos de documentación técnica del AI Act[131] no son solo una exigencia burocrática sino una condición necesaria para que el sistema de responsabilidad pueda funcionar. Sin trazabilidad no hay responsabilidad efectiva, y sin responsabilidad efectiva no hay incentivo real para cumplir con los estándares de seguridad.

---

[127] Parlamento Europeo y Consejo de la Unión Europea. (2024). *Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por productos defectuosos*. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024L2853

[128] Comisión Europea. (2017). Reglamento de Ejecución (UE) 2017/373 de la Comisión, de 1 de marzo de 2017. *Diario Oficial de la Unión Europea*, L 62, 1–126. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32017R0373

[129] Comisión Europea. (2015). Reglamento (UE) 2015/340 de la Comisión, de 20 de febrero de 2015. *Diario Oficial de la Unión Europea*, L 63, 1–158. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32015R0340

[130] Organización de Aviación Civil Internacional. (1944). *Convenio sobre Aviación Civil Internacional* (Convenio de Chicago), art. 28. https://www.icao.int/publications/pages/doc7300.aspx

[131] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, arts. 11–13 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

La propuesta de Directiva sobre responsabilidad en materia de IA de 2022 apuntaba en la dirección correcta al introducir mecanismos que facilitaban la prueba de la relación causal y que obligaban a los desarrolladores a colaborar con la divulgación de evidencia técnica. Sin embargo, esta propuesta fue retirada por la Comisión en 2025[132], lo que significa que esos mecanismos no llegarán a formar parte del derecho de la Unión, al menos en su forma original. Mientras no exista un instrumento que los recupere y mientras la Part-AI de EASA no esté en vigor, el marco sigue siendo incompleto, y buena parte de la carga de cerrar ese vacío recae ahora sobre la Directiva (UE) 2024/2853 y sobre las reglas generales de responsabilidad de cada Estado miembro, con el riesgo de fragmentación que ello conlleva.. Las organizaciones que hoy están desarrollando o desplegando sistemas de IA en ATM operan en un entorno en el que las reglas del juego no están del todo definidas, y esa incertidumbre es en sí misma un riesgo que el sector no puede permitirse ignorar.

En definitiva, un reparto equilibrado de la responsabilidad no es solo una cuestión de justicia sino también de eficacia. Si el marco jurídico no es capaz de atribuir correctamente las consecuencias de un fallo a quien pudo haberlo prevenido, los incentivos para invertir en seguridad se debilitan. Y en un entorno tan crítico como el ATC, donde las consecuencias de un error pueden ser irreversibles, esos incentivos son demasiado importantes como para dejarlos al azar.

9. Conclusiones

La integración de sistemas de inteligencia artificial basados en machine learning en el control del tráfico aéreo ya no es un escenario de futuro: es un proceso que está en marcha y para el que el marco normativo actual no estaba preparado. A lo largo del trabajo se ha analizado desde tres ángulos que, como se anunció en la introducción, no se pueden separar: la certificación de los sistemas, la supervisión humana y la responsabilidad jurídica cuando algo falla. Las conclusiones siguen el orden de los cuatro objetivos específicos y terminan con una valoración de conjunto.

Sobre los estándares de certificación

El primer objetivo era ver si los estándares de certificación aeronáutica actuales sirven para sistemas de ML. La respuesta es que no, y no por una cuestión de papeleo sino por un problema de fondo. Estándares como el DO-178C o el DO-278A parten de algo que el ML no cumple: que el comportamiento del software se puede describir con requisitos, verificar a través del código y garantizar para todos los casos posibles.[133] En un sistema que aprende de datos no hay una

[132] Comisión Europea. (2022). *Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial* [COM/2022/496 final]. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496

[133] EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

lógica interna que auditar paso a paso, y por eso la propia EASA admite que sus marcos de aseguramiento no están pensados para el aprendizaje automático.[134]

La respuesta a ese desajuste es el proceso en W o *Learning Assurance*, que añade un eje dedicado a los datos al clásico proceso en V. La idea es buena, pero no conviene darla por madura. El proyecto MLEAP, que buscaba métodos concretos para cumplir esos objetivos, concluyó que en varios puntos los resultados todavía no son lo bastante sólidos para servir como medios de cumplimiento formales.[135] En concreto, demostrar que un modelo va a funcionar bien fuera de sus datos de entrenamiento pero dentro de su ODD sigue siendo un problema sin resolver del todo. La consecuencia es incómoda: hoy no existe una forma plenamente establecida de certificar que un sistema de ML se va a comportar de forma fiable en todas las situaciones previstas. La certificación inicial, por sí sola, se queda corta para esta tecnología.

Sobre el encaje entre el AI Act y la normativa aeronáutica

El segundo objetivo era analizar cómo clasifica el AI Act los sistemas de IA en ATM y cómo convive con la normativa aeronáutica que ya existía. El AI Act es un avance necesario, porque por primera vez reconoce con una regulación horizontal que la IA necesita reglas propias. El problema es que, al superponerse a la normativa aeronáutica, hoy genera más fricciones que sinergias. La más visible es que cada marco mide el riesgo de forma distinta: el AI Act lo trata de manera prácticamente binaria, mientras que EASA lo gradúa por niveles. Así, sistemas con un impacto muy distinto sobre la seguridad pueden acabar sujetos a las mismas exigencias.[136]

La mejor prueba de que el encaje no es automático es que EASA ha tenido que ponerse a construirlo. La NPA 2025-07, publicada en noviembre de 2025 como primer paso hacia la futura Part-AI, es una norma puente que no haría falta si los dos marcos encajaran sin más.[137] Dicho de otro modo, el marco coherente que se buscaba todavía no existe, se está redactando ahora mismo. Mientras la Part-AI no entre en vigor, quien quiera integrar IA en sistemas ATM se mueve en bastante incertidumbre, obligado a cumplir a la vez dos marcos cuyas lagunas ninguno de los dos resuelve por su cuenta.

Sobre los riesgos para la supervisión humana

El tercer objetivo era examinar qué riesgos plantea la automatización creciente para que la supervisión humana sea efectiva. La conclusión principal es que cumplir el artículo 14 del AI

[134] EASA. (2024a). *Concept paper: Guidance for level 1 & 2 machine learning applications — Issue 02*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-concept-paper-guidance-level-1-2-machine-learning

[135] EASA. (2024b). *MLEAP D4 public report: Executive summary (Issue 01)*. European Union Aviation Safety Agency. https://www.easa.europa.eu/sites/default/files/dfu/mleap-d4-public-report-executive_summary_expanded-issue01.pdf

[136] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[137] EASA. (2025). *Notice of Proposed Amendment (NPA) 2025-07: AI trustworthiness framework*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/newsroom-and-events/news/easas-first-regulatory-proposal-artificial-intelligence-aviation-now-open

Act sobre el papel no garantiza que la supervisión sea real.[138] Un controlador que técnicamente puede detener el sistema, pero que lleva meses sin necesitar hacerlo, que confía mucho en sus recomendaciones y que está cansado, no está en condiciones de supervisar de verdad aunque formalmente tenga todos los medios.

El sesgo de automatización y la degradación de habilidades no son riesgos secundarios: son consecuencias previsibles del propio diseño cuando este no se piensa para contrarrestarlos. Por eso el *Human Oversight* no se puede entender solo como una característica técnica del sistema, sino como una condición operativa que depende de tres niveles que tienen que darse a la vez: un diseño que mantenga al controlador cognitivamente implicado, una organización del trabajo que haga posible la supervisión y una formación que la mantenga viva.[139] Si falla cualquiera de los tres, la supervisión humana se queda en algo formal justo cuando más debería servir: en el momento del fallo.

Sobre el reparto de la responsabilidad jurídica

El cuarto objetivo era determinar cómo se reparte la responsabilidad entre los distintos actores cuando un sistema de IA falla en ATC. La conclusión es que casi nunca hay un único responsable, y que el marco actual está mal preparado para gestionar esa responsabilidad compartida. El análisis por actores y los tres escenarios lo dejan claro: el fallo suele venir de la suma de decisiones técnicas, organizativas y operativas tomadas por personas distintas en momentos distintos, algo para lo que las reglas generales de responsabilidad civil no se diseñaron.

El balance es desigual. En la parte del defecto del producto se ha avanzado: la Directiva (UE) 2024/2853 equipara los sistemas de IA a cualquier otro producto a efectos de responsabilidad objetiva, e incluso permite presumir el defecto si el desarrollador se niega a entregar la documentación técnica.[140] Pero en la parte de la causalidad se ha retrocedido. La propuesta de Directiva sobre responsabilidad en materia de IA de 2022, que introducía una presunción de causalidad pensada justo para los casos de opacidad, fue retirada por la Comisión, y esa retirada se formalizó en el Diario Oficial de la Unión Europea el 6 de octubre de 2025.[141] Queda así un marco asimétrico, que ayuda a probar que el producto tenía un defecto pero deja sin herramienta específica lo más difícil en un sistema opaco: demostrar que ese defecto causó el daño. Lejos de debilitar el argumento del trabajo, la retirada lo refuerza: el propio legislador europeo vio que las reglas generales no bastaban, diseñó un instrumento para resolverlo y acabó renunciando a él sin sustituirlo, dejando el vacío sin cubrir y devolviendo el problema a las reglas nacionales, con el riesgo de fragmentación que eso supone.

---

[138] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, art. 14 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

[139] Endsley, M. R., y Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors*, 37(2), 381–394. https://doi.org/10.1518/001872095779064555

[140] Parlamento Europeo y Consejo de la Unión Europea. (2024). Directiva (UE) 2024/2853, de 23 de octubre de 2024, sobre responsabilidad por productos defectuosos. *Diario Oficial de la Unión Europea*. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024L2853

[141] Comisión Europea. (2022). Propuesta de Directiva relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial [COM/2022/496 final]. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496 — Retirada formalizada en el DOUE, C/2025/5423, de 6 de octubre de 2025.

Valoración de conjunto: un marco a medio hacer

Si los cuatro objetivos tienen algo en común es esto: en las tres dimensiones, certificación, supervisión y responsabilidad, el marco europeo está a medio camino. La certificación aún no tiene métodos sólidos para garantizar el rendimiento dentro del ODD, la articulación regulatoria depende de una Part-AI que todavía no existe y la responsabilidad ha perdido su instrumento más específico sobre la causalidad sin que nada lo reemplace. Tres piezas que deberían encajar y que hoy no lo hacen.

La aportación principal de este trabajo es mostrar que esas tres dimensiones no son problemas separados, sino un mismo problema. La opacidad del modelo dificulta a la vez la certificación, la supervisión y la prueba de la responsabilidad. La supervisión humana es a la vez un requisito de seguridad y el factor que más condiciona si se le puede atribuir culpa al controlador. Y la trazabilidad, lejos de ser papeleo, es lo que hace posible tanto la certificación continua como un reparto de responsabilidad proporcional a lo que cada actor pudo haber hecho para evitar el fallo.[142] Sin trazabilidad no hay responsabilidad efectiva, y sin responsabilidad efectiva se debilitan los incentivos para invertir en seguridad, algo que en un entorno como el ATC, donde un error puede ser irreversible, no se puede dejar al azar.

El camino, entonces, no es actualizar normas sueltas, sino cambiar la lógica de fondo: pasar de un modelo estático, pensado para sistemas deterministas que se certifican una vez y no cambian, a un modelo de supervisión continua que acompañe al sistema durante toda su vida útil. Mientras ese cambio no se complete, lo prudente es que la IA entre en ATC de forma gradual, manteniendo a la persona de verdad en el centro del control y no como una garantía formal delante de una pantalla. La tecnología avanza más deprisa que el marco que debe gobernarla, y reconocer ese desfase con honestidad es el primer paso para cerrarlo.

Limitaciones y futuras líneas de investigación

Conviene terminar señalando los límites del trabajo. El análisis se basa en revisión documental y análisis normativo, así que sus conclusiones se apoyan en el marco regulatorio existente y en la literatura disponible, no en datos sobre incidentes reales, que por lo nueva que es la tecnología son todavía escasos. Los tres escenarios son hipotéticos: sirven para ver cómo funcionaría el marco jurídico, no para sacar conclusiones estadísticas. Además, buena parte de la normativa está en plena evolución: el AI Act no despliega todos sus efectos hasta agosto de 2026 y la Part-AI sigue en consulta, de modo que algunas conclusiones podrían cambiar a medida que estos textos se concreten. Por eso resultarían especialmente útiles tres líneas futuras: analizar el texto definitivo de la Part-AI una vez aprobado, estudiar los primeros casos reales de despliegue de IA en ATM en Europa y elaborar métricas concretas de monitorización post-despliegue que traduzcan en umbrales operativos lo que el artículo 72 del AI Act hoy plantea de forma todavía genérica.

[142] Parlamento Europeo y Consejo de la Unión Europea. (2024). Reglamento (UE) 2024/1689, arts. 11–13 [AI Act]. *Diario Oficial de la Unión Europea*, L 206, 1–144. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32024R1689

Bibliografía:

Adadi, A., y Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access, 6*, 52138–52160. https://doi.org/10.1109/ACCESS.2018.2870052

Borghini, G., Aricò, P., Di Flumeri, G., Cartocci, G., Colosimo, A., Bonelli, S., Golfetti, A., Imbert, J. P., Granger, G., y Benhacene, R. (2017). EEG-based cognitive control behaviour assessment: An ecological study with professional air traffic controllers. *Scientific Reports, 7*, 547. https://doi.org/10.1038/s41598-017-00633-7

Comisión Europea. (2022). *Propuesta de Directiva del Parlamento Europeo y del Consejo relativa a la adaptación de las normas de responsabilidad civil extracontractual a la inteligencia artificial* [COM/2022/496 final]. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:52022PC0496 (Propuesta retirada por la Comisión Europea; retirada notificada en el DOUE C/2025/5423, de 6 de octubre de 2025).

Comisión Europea, Grupo de Expertos de Alto Nivel en Inteligencia Artificial. (2019). *Directrices éticas para una IA fiable*. Comisión Europea. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Cummings, M. L. (2004). Automation bias in intelligent time critical decision support systems. *AIAA 1st Intelligent Systems Technical Conference*. https://doi.org/10.2514/6.2004-6313

EASA. (2023). *Artificial intelligence roadmap 2.0: Human-centric approach to AI in aviation*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-artificial-intelligence-roadmap-20

EASA. (2024a). *Concept paper: Guidance for level 1 & 2 machine learning applications — Issue 02*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/document-library/general-publications/easa-concept-paper-guidance-level-1-2-machine-learning

EASA. (2024b). *MLEAP D4 public report: Executive summary* (Issue 01). European Union Aviation Safety Agency. https://www.easa.europa.eu/sites/default/files/dfu/mleap-d4-public-report-executive_summary_expanded-issue01.pdf

EASA. (2025). *Notice of Proposed Amendment (NPA) 2025-07: AI trustworthiness framework*. European Union Aviation Safety Agency. https://www.easa.europa.eu/en/newsroom-and-events/news/easas-first-regulatory-proposal-artificial-intelligence-aviation-now-open

ENAIRE. (2023, 20 de junio). *ENAIRE premia ideas de negocio y retos tecnológicos en gestión del tráfico aéreo*. ENAIRE. https://www.enaire.es/es_ES/2023_06_20/ndp_premio_ideas_negocio_crida

ENAIRE. (2026). *Plan anual de actuación 2026*. ENAIRE. https://www.enaire.es/docs/inline/es_ES/plan_anual_de_actuación/Plan_anual_actuacion_2026.pdf

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M. R., y Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(2), 381–394. https://doi.org/10.1518/001872095779064555

Endsley, M. R., y Rodgers, M. D. (1998). Distribution of attention, situation awareness, and workload in a passive air traffic control task: Implications for operational errors and automation. *Air Traffic Control Quarterly, 6*(1), 21–44. https://doi.org/10.2514/atcq.6.1.21

ERRIN. (2025). *Research objectives and MLEAP final report findings*. https://errin.eu/calls/research-help-shape-future-regulatory-framework-des

España. (1960). *Ley 48/1960, de 21 de julio, sobre Navegación Aérea*. Boletín Oficial del Estado. https://www.boe.es/buscar/act.php?id=BOE-A-1960-10905

España. (2003). *Ley 21/2003, de 7 de julio, de Seguridad Aérea*. Boletín Oficial del Estado. https://www.boe.es/buscar/act.php?id=BOE-A-2003-13616

España. (2015). *Ley 40/2015, de 1 de octubre, de Régimen Jurídico del Sector Público*. Boletín Oficial del Estado, 236. https://www.boe.es/buscar/act.php?id=BOE-A-2015-10566

EUROCONTROL. (2015). *Human factors integration in ATM system design: A white paper*. EUROCONTROL. https://www.eurocontrol.int/publication/human-factors-integration-atm-system-design

EUROCONTROL. (2024). *EUROCONTROL forecast update 2024–2030 (Autumn 2024)*. EUROCONTROL. https://www.eurocontrol.int/sites/default/files/2024-10/eurocontrol-seven-year-forecast-2024-2030-october-2024.pdf

EUROCONTROL. (s.f.a). *ATC-TBO project*. https://www.eurocontrol.int/project/atc-tbo

EUROCONTROL. (s.f.b). *Integrated network management*. https://www.eurocontrol.int/project/integrated-network-management

EUROCONTROL y SESAR JU. (2020). *FLY AI report: Demystifying and accelerating AI in aviation/ATM*. European Aviation High Level Group on Artificial Intelligence. https://www.eurocontrol.int/sites/default/files/2020-03/eurocontrol-fly-ai-report-032020.pdf

Htin, N. M. (2026). Redefining the standard of human oversight for AI negligence. *Harvard Journal of Law & Technology Digest*. https://jolt.law.harvard.edu/digest/redefining-the-standard-of-human-oversight-for-ai-negligence

ICAO. (2018). *Safety management manual* (Doc 9859, 4.ª ed.). International Civil Aviation Organization. https://www.icao.int/safety/SafetyManagement/Pages/GuidanceandTools.aspx

ICAO. (2025a). *Prácticas actuales y mejoras de inteligencia artificial para la aviación civil internacional* (Nota de trabajo A42-WP/401 EX/176). Organización de Aviación Civil Internacional. https://www.icao.int/sites/default/files/Meetings/a42/Documents/WP/wp_401_es.pdf

ICAO. (2025b). *Artificial intelligence (AI) contribution to aviation* (Nota de trabajo A42-WP/489 EX/220). International Civil Aviation

Organization. https://www.icao.int/sites/default/files/Meetings/a42/Documents/WP/wp_489_en.pdf

OCDE. (2024a). *Recommendation of the Council on Artificial Intelligence* (actualizada mayo 2024). OECD. https://oecd.ai/en/ai-principles

OCDE. (2024b). *Explanatory memorandum on the updated OECD definition of an AI system* (OECD Artificial Intelligence Papers, n.º 8). OECD. https://doi.org/10.1787/623da898-en

Organización de Aviación Civil Internacional. (1944). *Convenio sobre Aviación Civil Internacional (Convenio de Chicago)*, art. 28. https://www.icao.int/publications/pages/doc7300.aspx

Organización de Aviación Civil Internacional. (1999). *Convenio para la unificación de ciertas reglas para el transporte aéreo internacional (Convenio de Montreal)*. Boletín Oficial del Estado (ratificación en España, 2004). https://www.boe.es/buscar/doc.php?id=BOE-A-2004-9347

Páez, A. (2025). Transparencia, explicabilidad y confianza en los sistemas de aprendizaje automático. En J. D. Gutiérrez y R. Manrique (Eds.), *IA en acción: Desarrollo, impacto e implicaciones éticas en la sociedad*. Ediciones Uniandes. https://philarchive.org/archive/PEZTEY

Parlamento Europeo y Consejo de la Unión Europea. (2015). *Reglamento (UE) 2015/340 de la Comisión, de 20 de febrero de 2015*. Diario Oficial de la Unión Europea. https://www.boe.es/buscar/doc.php?id=DOUE-L-2015-80440

Parlamento Europeo y Consejo de la Unión Europea. (2017). *Reglamento de Ejecución (UE) 2017/373 de la Comisión, de 1 de marzo de 2017, por el que se establecen requisitos comunes para los proveedores de servicios de gestión del tránsito aéreo/navegación aérea*. Diario Oficial de la Unión Europea. https://www.boe.es/buscar/doc.php?id=DOUE-L-2017-80442

Parlamento Europeo y Consejo de la Unión Europea. (2018). *Reglamento (UE) 2018/1139 del Parlamento Europeo y del Consejo, de 4 de julio de 2018, sobre normas comunes en el ámbito de la aviación civil* [Reglamento de Base de EASA]. Diario Oficial de la Unión Europea, L 212, 1–122. https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX:32018R1139

Parlamento Europeo y Consejo de la Unión Europea. (2024a). *Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial* [Ley de Inteligencia Artificial]. Diario Oficial de la Unión Europea, L 206, 1–144. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81079

Parlamento Europeo y Consejo de la Unión Europea. (2024b). *Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, de 23 de octubre de 2024, sobre responsabilidad por los daños causados por productos defectuosos*. Diario Oficial de la Unión Europea. https://www.boe.es/buscar/doc.php?id=DOUE-L-2024-81701

Rierson, L. (2013). *Developing safety-critical software: A practical guide for aviation software and DO-178C compliance*. CRC Press.

SESAR Joint Undertaking. (2023). *European ATM master plan: Automation roadmap* (ed. 2023). SESAR JU. https://www.sesarju.eu/node/4820

Willems, B., y Truitt, T. R. (1999). *Implications of reduced involvement in en route air traffic control* (Informe técnico DOT/FAA/AM-99/19). Federal Aviation Administration Civil Aeromedical Institute.