INTABIOTECH

INTELIGENCIA ARTIFICIAL, APPCC Y SEGURIDAD ALIMENTARIA

Cómo integrar algoritmos sin crear nuevos puntos ciegos

Relación entre IA, análisis de peligros, puntos de control crítico, programas de prerrequisitos, alertas predictivas y decisiones automatizadas

TESIS EDITORIAL

La IA puede ampliar la capacidad preventiva de un sistema APPCC, pero también puede ocultar fallos bajo una apariencia de precisión. La integración correcta exige tratar al algoritmo como una fuente adicional de riesgo, evidencia y decisión: debe tener finalidad delimitada, datos gobernados, validación contextual, límites operativos, supervisión humana real, registros íntegros y un modo seguro de degradación.

José Manuel López

Director Técnico · INTABIOTECH

Nota editorial, alcance y cautelas de actualización

Este trabajo se concibe como un artículo técnico-jurídico de aplicación empresarial. Integra el Derecho alimentario de la Unión Europea, el marco horizontal de inteligencia artificial, protección de datos, ciberseguridad, responsabilidad por producto, seguridad de maquinaria y sistemas voluntarios de gestión. No sustituye la evaluación jurídica del caso concreto, la validación científica del proceso ni la responsabilidad indelegable del operador de empresa alimentaria. Las referencias normativas se han comprobado hasta el 18 de julio de 2026.

Existe una circunstancia temporal especialmente relevante. El Reglamento (UE) 2024/1689 —Reglamento de Inteligencia Artificial, conocido como *AI Act*— mantiene en su texto originario un calendario escalonado. No obstante, el Consejo de la Unión Europea anunció el 29 de junio de 2026 su aprobación final de la reforma conocida como *Digital Omnibus on AI*, que fija nuevas fechas para las obligaciones relativas a sistemas de alto riesgo: 2 de diciembre de 2027 para los sistemas autónomos y 2 de agosto de 2028 para los integrados en productos. Al cierre de esta edición, la consolidación y publicación final de esa reforma debía verificarse antes de invocar fechas en un expediente, contrato o dictamen. Por ello, el artículo separa cuidadosamente: a) las obligaciones ya aplicables; b) el contenido material del régimen de alto riesgo; y c) el calendario transitorio sujeto a la versión oficial publicada.

ADVERTENCIA DE USO

Una fecha anunciada institucionalmente no debe confundirse con una norma consolidada. Antes de una auditoría, compra tecnológica, lanzamiento o litigio, debe consultarse el DOUE y la versión consolidada de EUR-Lex vigente ese día. Esta cautela no modifica la recomendación práctica: las empresas alimentarias deben aplicar desde ahora una gobernanza equivalente a los requisitos de trazabilidad, supervisión, robustez y control de cambios, aunque su sistema concreto no sea jurídicamente “de alto riesgo”.

Resumen

La incorporación de inteligencia artificial a los sistemas de autocontrol alimentario promete *detectar desviaciones antes de que se materialicen, integrar señales dispersas, priorizar muestreos, anticipar contaminación ambiental y mejorar la respuesta ante incidentes*. Sin embargo, la automatización también puede *crear puntos ciegos inéditos*: falsos negativos estadísticamente invisibles, deriva del modelo, sensores descalibrados, correlaciones espurias, interfaces que inducen a la confianza excesiva, proveedores que actualizan el algoritmo sin control, ataques sobre datos y decisiones automáticas que alteran el destino de un lote sin una base probatoria suficiente. El problema no consiste en “añadir IA al APPCC”, sino en rediseñar el sistema de gestión para que la IA quede sometida a los mismos principios de prevención, validación, monitorización, verificación, documentación y mejora continua que rigen la seguridad alimentaria.

El artículo desarrolla una arquitectura de doble control. En la primera capa se mantiene el análisis clásico de peligros biológicos, químicos, físicos y alergénicos, con sus programas de prerrequisitos, prerrequisitos operativos y puntos de control crítico. En la segunda se incorpora un análisis de peligros algorítmicos y digitales: calidad y representatividad de datos, incertidumbre, deriva, integridad de sensores, robustez, ciberseguridad, explicabilidad operativa, supervisión humana, gestión de versiones y conservación de evidencias. Se propone el método HAZARD-AI, una matriz de decisión para determinar cuándo el algoritmo es mera ayuda, cuándo forma parte de una medida de control y cuándo su fallo debe activar un modo seguro o detener el proceso. Finalmente, se ofrece un modelo de gobernanza, validación y contratación aplicable a industrias certificadas bajo ISO 22000, FSSC 22000, BRCGS o IFS, y compatible con el Reglamento (CE) 852/2004, la Comunicación 2022/C 355/01, el AI Act, el RGPD, NIS2, el *Cyber Resilience Act* y el nuevo régimen europeo de responsabilidad por productos defectuosos.

Palabras clave: *inteligencia artificial; APPCC; HACCP; seguridad alimentaria; PCC; prerrequisitos; alertas predictivas; decisiones automatizadas; AI Act; deriva del modelo; trazabilidad probatoria; ciberseguridad industrial.*

Abstract

Artificial intelligence can strengthen food safety management by detecting weak signals, predicting loss of process control, prioritising environmental sampling and accelerating incident response. It can also create novel blind spots: silent false negatives, model and sensor drift, spurious correlations, automation bias, uncontrolled supplier updates, data poisoning and automated lot-disposition decisions unsupported by sufficient evidence. The relevant question is therefore not how to "add AI to HACCP", but how to subject AI to the preventive, validated, monitored, verified and documented logic of food safety management.

This paper proposes a two-layer assurance architecture. The first layer preserves the conventional analysis of biological, chemical, physical and allergen hazards through prerequisite programmes, operational prerequisite programmes and critical control points. The second layer addresses algorithmic and digital hazards: data governance, uncertainty, drift, sensor integrity, robustness, cybersecurity, operational explainability, meaningful human oversight, version control and evidentiary retention. A practical HAZARD-AI method is presented to determine whether an algorithm is a decision-support tool, part of a control measure or a safety-relevant component requiring fail-safe behaviour. The article also maps the interaction between EU food law, the AI Act, data protection, NIS2, the Cyber Resilience Act, machinery safety and product liability, and provides implementation, validation, audit and procurement templates for food businesses.

Keywords: artificial intelligence; HACCP; food safety; critical control points; prerequisite programmes; predictive alerts; automated decisions; AI Act; model drift; evidentiary traceability; industrial cybersecurity.

Índice

Síntesis ejecutiva: diez proposiciones para el consejo de dirección

Nº	Proposición	Consecuencia directiva
1	La IA no sustituye al operador	La responsabilidad primaria por la seguridad alimentaria permanece en el operador. Una recomendación algorítmica no desplaza el deber de organizar, validar y probar el autocontrol.
2	Un algoritmo no es un PCC por sí mismo	El PCC es una etapa en la que el control es esencial. El algoritmo puede monitorizarla, verificarla o actuar sobre ella; solo entonces forma parte de la medida de control y debe validarse como componente del sistema.
3	El dato también es materia prima crítica	Si la decisión depende de datos incompletos, sesgados, retrasados o manipulados, el proceso de control queda degradado, aunque el modelo sea técnicamente sofisticado.
4	La precisión media es insuficiente	En seguridad alimentaria importa la tasa de falsos negativos en condiciones críticas, por producto, línea, turno, estación, proveedor y categoría de peligro.
5	La deriva es una modificación del proceso	Cambios en formulación, envase, equipo, proveedor, microbiota, limpieza o estación pueden invalidar el modelo y obligan a revisar el APPCC conforme al artículo 5 del Reglamento 852/2004.
6	La alerta debe tener dueño y reloj	Toda alerta necesita severidad, responsable, plazo de respuesta, criterio de escalado, acción sobre producto y cierre documentado.
7	La supervisión humana debe poder contradecir	Un supervisor que solo pulsa “aceptar” no controla nada. Debe entender límites, disponer de tiempo, autoridad y acceso a evidencia independiente.
8	El modo degradado es obligatorio en la práctica	Si falla la IA, la empresa debe saber operar de forma segura mediante controles convencionales, frecuencia reforzada, retención de producto o parada.
9	La ciberseguridad es seguridad alimentaria	Un ataque que altera temperaturas, etiquetas, recetas, parámetros CIP o resultados de laboratorio puede materializar un peligro alimentario.
10	El valor comercial depende de la prueba	La IA genera retorno cuando reduce desviaciones, liberaciones tardías, desperdicio, muestreo indiscriminado y tiempo de investigación, sin rebajar el nivel de evidencia.

1. El problema: automatizar sin desplazar la responsabilidad

La digitalización de la seguridad alimentaria ha pasado de registrar datos a inferir riesgos. Las plantas ya no se limitan a capturar temperaturas, presiones, conductividades, imágenes, resultados microbiológicos o incidencias; comienzan a combinarlos para estimar la probabilidad de una desviación, clasificar el riesgo de un lote o recomendar una acción. Ese salto —de medir a inferir— altera la naturaleza del autocontrol. Cuando una persona observa un termómetro y aplica un límite crítico, la relación entre hecho, regla y decisión es relativamente visible. Cuando un modelo integra cientos de variables, produce un score y el sistema bloquea o libera producto, la cadena causal puede quedar oculta incluso para quienes operan la línea.

El atractivo empresarial es evidente. Los algoritmos pueden reconocer patrones que preceden a una pérdida de control, jerarquizar zonas de muestreo ambiental, anticipar un fallo de pasteurización, detectar anomalías en etiquetas o envases, estimar vida útil, identificar proveedores con combinación atípica de señales y reducir la fatiga de revisión. El error estratégico consiste en confundir capacidad predictiva con garantía de seguridad. Un modelo puede ser muy útil y, al mismo tiempo, estar mal gobernado; puede superar a una regla simple en promedio y fallar precisamente en la clase minoritaria que concentra el daño; puede funcionar durante seis meses y degradarse después de un cambio de formulación que nadie comunicó al equipo de datos.

La legislación alimentaria europea no permite externalizar la responsabilidad al software. El Reglamento (CE) 178/2002 atribuye al operador la responsabilidad de garantizar que los alimentos cumplen la legislación en las etapas bajo su control. El Reglamento (CE) 852/2004 exige implantar, aplicar y mantener procedimientos permanentes basados en los principios APPCC, revisar el sistema cuando se modifique el

producto o el proceso y conservar documentación que demuestre su aplicación efectiva. La IA puede ser una herramienta del sistema, pero no un sujeto responsable ni una presunción automática de diligencia.

NUESTRO PUNTO DE PARTIDA JURÍDICO

La empresa no cumple porque "tiene IA". Cumple si puede demostrar que ha identificado los peligros relevantes, seleccionado medidas de control adecuadas, validado sus límites y su tecnología, monitorizado el proceso, actuado ante desviaciones, verificado la eficacia y conservado la/s evidencia/s fiable/s. La IA solo añade valor probatorio cuando cada uno de esos verbos sigue siendo demostrable.

El riesgo de nuevos puntos ciegos aparece cuando la organización desplaza controles explícitos hacia una infraestructura que no comprende ni gobierna. Hay punto ciego cuando una señal relevante no se capta; también cuando se capta, pero el modelo no la pondera correctamente; y cuando el modelo alerta, pero nadie responde; o cuando la respuesta se ejecuta, pero no se conserva el razonamiento; e igualmente cuando una actualización modifica el comportamiento sin revisión del equipo APPCC. Por ello, la integración debe analizar el ciclo completo: sensor, dato, tratamiento, inferencia, interfaz, decisión, acción física, destino del producto, registro y aprendizaje posterior.

La hipótesis de este artículo es que la IA debe incorporarse como una capa de control sometida a una disciplina dual. La **primera disciplina es la alimentaria**: peligro, medida de control, nivel aceptable, monitorización, corrección, verificación y registro. La **segunda es la algorítmica:** finalidad prevista, calidad de datos, desempeño por escenarios, incertidumbre, robustez, deriva, supervisión, ciberseguridad, versionado y trazabilidad. Ninguna de las dos sustituye a la otra.

2. APPCC como arquitectura jurídica y probatoria, y no como formulario

El APPCC suele degradarse en la práctica a un conjunto de tablas heredadas. Esa visión documentalista resulta especialmente peligrosa al introducir IA, porque facilita incorporar una nueva columna de "alerta predictiva" sin revisar la lógica del sistema. El APPCC es, en realidad, una arquitectura de decisiones preventivas. Vincula conocimientos científicos sobre peligros con condiciones concretas de producto, proceso, instalaciones, proveedores, consumidores y uso previsto. Su función no es describir que la planta controla, sino explicar por qué determinadas medidas son necesarias, cómo se sabe que funcionan y qué ocurre cuando dejan de funcionar.

La Comunicación de la Comisión 2022/C 355/01 sitúa las buenas prácticas de higiene, los programas de prerrequisitos, los prerrequisitos operativos y los procedimientos basados en APPCC dentro de un sistema de gestión de la seguridad alimentaria. Esa arquitectura evita convertir todos los controles en PCC y exige asignar cada peligro a la capa adecuada. La IA no altera esta jerarquía; debe respetarla. Un modelo de visión que identifica acumulaciones de residuo puede reforzar un programa de limpieza. Un modelo que anticipa el riesgo de contaminación ambiental puede priorizar muestreos de verificación. Un algoritmo que ajusta automáticamente una etapa térmica puede integrarse en una medida de control crítica. La clasificación depende de la función y de las consecuencias de fallo, no de la etiqueta comercial del software.

Los **siete principios** del artículo 5 del Reglamento 852/2004 continúan siendo el núcleo: **identificar peligros; determinar PCC; fijar límites críticos; monitorizar; establecer acciones correctivas; verificar; y documentar.** La inteligencia artificial puede intervenir en todos, pero con distinta intensidad probatoria. En el análisis de peligros puede ayudar a explorar datos históricos, nunca sustituir la evaluación multidisciplinar. En la determinación de PCC puede aportar evidencia, pero no aplicar mecánicamente un árbol de decisión sin contexto. En la monitorización puede detectar patrones multivariantes, pero debe coexistir con límites observables y reglas de actuación. En la verificación puede revelar tendencias que una revisión mensual no detecta. En la documentación puede estructurar evidencias, pero un modelo generativo no debe inventar registros ni reconstruir retrospectivamente hechos inexistentes.

2.1. Los cuatro objetos que no deben confundirse

Para integrar IA con rigor conviene distinguir cuatro objetos: el peligro alimentario; la medida de control; el punto o etapa donde se aplica; y el mecanismo digital que observa o actúa. El peligro puede ser *Listeria monocytogenes*, un alérgeno no declarado, un fragmento metálico o una toxina. La medida de control puede ser tratamiento térmico, segregación, detector de metales, limpieza validada o control de proveedor. El PCC es la etapa donde el control resulta esencial. El mecanismo digital puede ser un sensor, un modelo predictivo, una cámara o un actuador. Confundir estos niveles lleva a afirmar que “la IA es el PCC”, una formulación técnicamente imprecisa.

2.2. La obligación de revisión como regla de control de deriva

El último inciso del artículo 5.2 del Reglamento 852/2004 obliga a revisar el procedimiento cuando se modifica el producto, el proceso o una etapa. En un entorno algorítmico, esta regla debe interpretarse de manera material. Hay modificación relevante no solo cuando cambia la receta o el equipo, sino cuando cambia la distribución de datos de entrada, la versión del modelo, la configuración de umbrales, la cámara, la iluminación, el firmware del sensor, la población de proveedores o el sistema de etiquetado. La deriva no es un incidente puramente informático; puede ser una modificación del proceso de control que activa el deber de revisión APPCC.

2.3. Evidencia proporcional, pero no ornamental

La documentación debe ser proporcional a naturaleza y tamaño, pero suficiente para demostrar aplicación efectiva. En IA, la proporcionalidad no significa prescindir de evidencia esencial. Incluso una pyme necesita saber qué versión utiliza, para qué sirve, con qué datos se validó, qué límites tiene, quién responde a sus alertas y qué hacer si falla. La sofisticación documental puede variar; la capacidad de reconstruir una decisión que afectó a la seguridad del producto no debería desaparecer.

3. Dónde encaja la IA: observar, inferir, recomendar, decidir y actuar

La expresión “IA en seguridad alimentaria” agrupa funciones jurídicamente diferentes. La evaluación debe comenzar por el grado de influencia del sistema sobre la realidad física y sobre el destino del producto. Cuanto mayor sea la autonomía y menor la reversibilidad, más intensa debe ser la garantía.

Función	Qué hace	Riesgo dominante	Ejemplo
Observación	Convierte señales en datos: visión, espectros, audio, sensores, lectura documental.	La IA no decide; puede, no obstante, omitir o distorsionar una señal crítica.	Cámara que detecta envases sin sello.
Inferencia	Estima una variable no medida o una probabilidad.	Riesgo de correlación espuria, incertidumbre y cambio de dominio.	Probabilidad de crecimiento microbiano o de fallo CIP.
Recomendación	Propone priorización o acción al usuario.	Automatización sesgada: la persona puede seguirla sin revisión real.	Priorizar zonas de hisopado ambiental.
Decisión	Clasifica, bloquea, libera o asigna destino.	La decisión requiere reglas de autoridad, evidencia y apelación interna.	Retención automática de un lote por score.
Actuación	Modifica parámetros o ejecuta una acción física.	El algoritmo forma parte de una función de control y puede generar un peligro por acción errónea.	Ajuste automático de tiempo-temperatura o dosificación.
Generación	Produce texto, instrucciones o documentación.	Riesgo de alucinación, fuente no verificable y contaminación del registro.	Borrador de análisis de causa o actualización de plan APPCC.

Esta escala permite fijar un principio de gobernanza: la intensidad del control no depende del nombre del modelo, sino de la influencia real de su salida. Un sistema de recomendación puede ser de facto decisorio si la organización sanciona al operario que lo contradice, si la interfaz oculta alternativas o si no hay tiempo para revisar. Del mismo modo, un sistema que actúa sobre un equipo puede ser menos arriesgado si sus acciones

están acotadas por límites duros independientes, dispone de interbloqueos y revierte automáticamente a un estado seguro.

La IA generativa merece un tratamiento separado. Es útil para estructurar información, comparar versiones, redactar borradores, resumir registros y preparar preguntas de auditoría. No es apropiada como fuente primaria de límites críticos, como sustituto de la validación científica ni como mecanismo autónomo para cerrar desviaciones. Su salida debe considerarse propuesta no verificada. Debe prohibirse que cree registros de monitorización, complete datos ausentes o redacte retrospectivamente una justificación que no existía cuando se tomó la decisión.

REGLA DE ORO

A mayor proximidad entre la salida algorítmica y la liberación de producto, la modificación de un parámetro crítico o la interrupción de una barrera de seguridad, mayor exigencia de validación, redundancia, registro y supervisión independiente.

4. Mapa normativo aplicable en la Unión Europea y España

La integración de IA en APPCC se encuentra en la intersección de varios regímenes. El error jurídico más frecuente consiste en preguntar únicamente si el sistema está sujeto al *AI Act*. Aunque no sea “de alto riesgo” conforme a ese Reglamento, su uso puede incumplir la legislación alimentaria, el RGPD, obligaciones de ciberseguridad, reglas de seguridad de maquinaria, contratos o deberes de diligencia. La evaluación correcta es acumulativa.

Instrumento	Núcleo	Implicación para IA-APPCC
Reglamento (CE) 178/2002	Responsabilidad del operador; alimentos no seguros; trazabilidad; retirada y notificación.	Impide delegar la decisión final al proveedor del algoritmo y exige que la trazabilidad digital sea operativa.
Reglamento (CE) 852/2004	Higiene y procedimientos permanentes basados en APPCC.	Obliga a revisar el plan cuando cambian proceso o producto; la versión del modelo y los datos pueden constituir cambios relevantes.
Reglamento (UE) 2021/382	Gestión de alérgenos, redistribución y cultura de seguridad alimentaria.	La automatización debe ser coherente con compromiso directivo, comunicación, recursos y formación.
Comunicación 2022/C 355/01	Guía sobre GHP, PRP, PPRO/OPRP, APPCC y auditoría.	Proporciona el marco funcional en el que ubicar cada caso de IA.
Reglamento (CE) 2073/2005	Criterios microbiológicos y reglas de muestreo/acción.	Un score no sustituye un criterio legal ni un método de referencia salvo base jurídica y validación científica adecuada.
Reglamento (UE) 2017/625	Controles oficiales, auditoría y acceso a información.	La empresa debe poder presentar evidencias, registros y explicación de funcionamiento a la autoridad.
Reglamento (UE) 2024/1689	Régimen horizontal de IA, alfabetización, prohibiciones, transparencia y alto riesgo.	Clasificación por finalidad; posible alto riesgo si es componente de seguridad de producto sujeto a evaluación de terceros.
RGPD y LO 3/2018	Datos personales, videovigilancia, perfiles, decisiones automatizadas y evaluación de impacto.	Relevante para cámaras, desempeño de trabajadores, biometría, voz y decisiones con efectos significativos.
Directiva NIS2 y normativa nacional	Gestión de riesgos y notificación de incidentes para entidades incluidas.	El sector de producción, procesado y distribución alimentaria industrial/Mayorista puede quedar comprendido según tamaño y transposición.
Cyber Resilience Act	Ciberseguridad de productos con elementos digitales comercializados.	Relevante para proveedores de sensores, *gateways*, software o equipos conectados, con aplicación escalonada.
Reglamento de Máquinas 2023/1230	Seguridad de maquinaria, software y comportamiento auto-evolutivo.	Puede activar el *AI Act* de alto riesgo cuando la IA es componente de seguridad y existe evaluación de

Instrumento	Núcleo	Implicación para IA-APPCC
		conformidad de terceros.
Directiva (UE) 2024/2853	Responsabilidad objetiva por productos defectuosos, incluido software.	Refuerza la importancia de actualizaciones, ciberseguridad, documentación y reparto contractual de responsabilidades.
Ley 17/2011	Marco español de seguridad alimentaria y nutrición, riesgos emergentes y sistema de alerta.	Conecta la gobernanza digital con deberes nacionales de prevención y cooperación.
ISO 22000 / FSSC / BRCGS / IFS	Sistemas certificables y exigencias contractuales de clientes.	Pueden imponer controles más detallados que el mínimo legal y convertir la gobernanza algorítmica en requisito de auditoría.

4.1. Primacía del Derecho alimentario sectorial

El *AI Act* no sustituye las obligaciones de seguridad alimentaria. La clasificación de un sistema como no alto riesgo no equivale a una declaración de seguridad para uso alimentario. La decisión de liberar un lote, aceptar una desviación, modificar una medida de control o reducir muestreo debe seguir apoyada en el Derecho sectorial, en evidencia científica y en el sistema documentado del operador.

4.2. Protección de datos: no todo dato industrial es anónimo

Las cámaras de línea pueden captar trabajadores; los registros de acceso y tiempos de respuesta pueden perfilar desempeño; los sistemas de voz pueden tratar datos personales; y un modelo de riesgo puede atribuir incidencias a turnos o personas. Cuando exista dato personal, se requiere base jurídica, minimización, información, seguridad, plazos de conservación y, cuando proceda, evaluación de impacto. El artículo 22 RGPD entra en juego cuando existe una decisión basada únicamente en tratamiento automatizado que produce efectos jurídicos o afecta significativamente a una persona. La decisión sobre un lote no activa por sí sola ese artículo, pero la evaluación automatizada de un empleado sí puede hacerlo.

4.3. Ciberseguridad como condición de validez

NIS2 incluye, en su anexo sectorial, empresas alimentarias dedicadas a distribución mayorista y producción o transformación industrial, con el alcance por tamaño y las particularidades de transposición. Incluso cuando una empresa no quede formalmente incluida, las medidas de gestión de riesgo, continuidad, control de acceso, gestión de proveedores y respuesta a incidentes representan un estándar de diligencia razonable para sistemas cuya manipulación puede afectar la seguridad del alimento.

4.4. Responsabilidad por producto y por decisión empresarial

La Directiva 2024/2853 considera el software un producto a efectos del régimen de responsabilidad por productos defectuosos y trata al desarrollador o proveedor de software como fabricante en los términos previstos. Esto no elimina la responsabilidad del operador que integra, configura o utiliza el sistema de forma inadecuada. En la práctica pueden coexistir acciones contra el proveedor tecnológico, el fabricante del equipo y el operador alimentario. De ahí la necesidad de definir finalidad prevista, límites, mantenimiento, actualizaciones, logs, cooperación en incidentes y cobertura aseguradora.

5. ¿Es la IA alimentaria un sistema de alto riesgo bajo el *AI Act*?

La respuesta general es: no necesariamente. El *AI Act* no clasifica como alto riesgo toda IA que pueda influir en la salud. El artículo 6 establece dos rutas principales. La primera se refiere a sistemas que son productos o componentes de seguridad de productos cubiertos por legislación de armonización del anexo I y sujetos a evaluación de conformidad por terceros. La segunda incluye los casos enumerados en el anexo III, centrados en biometría, infraestructuras críticas, educación, empleo, servicios esenciales, autoridades, migración, justicia y procesos democráticos. El uso interno de un modelo para predecir desviaciones de higiene o priorizar muestreos no aparece, por ese solo hecho, en el anexo III.

La consecuencia es importante: no debe afirmarse de forma automática que un algoritmo APPCC es “alto riesgo”. Esa sobreclasificación genera costes y confusión. Pero tampoco debe inferirse que carece de riesgo jurídico. El operador sigue sometido a la legislación alimentaria y, además, a las obligaciones generales del AI Act que resulten aplicables, como la alfabetización en IA del artículo 4, las prohibiciones del artículo 5 y determinadas reglas de transparencia del artículo 50.

5.1. ¿Cuándo puede convertirse en alto riesgo?

La hipótesis más relevante en industria alimentaria es la IA integrada como componente de seguridad de maquinaria o de otro producto armonizado. Para que opere el artículo 6.1 deben concurrir ambas condiciones: que el sistema sea componente de seguridad —o producto— cubierto por legislación del anexo I; y que el producto requiera evaluación de conformidad por un tercero. Un algoritmo que controla una función de seguridad de una máquina previene una condición peligrosa o interviene en un sistema sujeto a esa evaluación puede entrar en el régimen de alto riesgo. No basta con que sea “importante” para la empresa; se requiere el encaje jurídico concreto.

Una segunda vía puede aparecer cuando la misma herramienta se utiliza para evaluar a trabajadores, asignar tareas, monitorizar rendimiento o adoptar decisiones laborales. Esa finalidad puede quedar en el Anexo III aunque el módulo de seguridad alimentaria no lo esté. La clasificación se realiza por finalidad prevista y contexto de uso. Un mismo producto comercial puede contener módulos con regímenes distintos.

5.2. Proveedor, responsable del despliegue y fabricante del producto

La empresa alimentaria suele ser “deployer” o responsable del despliegue cuando utiliza un sistema bajo su autoridad. Sin embargo, puede convertirse en proveedor si desarrolla el sistema y lo comercializa bajo su nombre, o si modifica de manera sustancial un sistema de alto riesgo o cambia su finalidad para convertirlo en alto riesgo. La marca blanca, la personalización profunda y el reentrenamiento autónomo deben analizarse contractualmente. No es prudente asumir que toda responsabilidad permanece en el proveedor SaaS.

5.3. Buen gobierno voluntario para sistemas no alto riesgo

El hecho de que un sistema no sea alto riesgo permite una aplicación proporcionada, no una ausencia de control. En un sistema ligado a seguridad alimentaria, conviene adoptar voluntariamente varios elementos materiales del capítulo de alto riesgo: gestión de riesgos, documentación técnica suficiente, logs, transparencia para el usuario, supervisión humana, exactitud, robustez, ciberseguridad y monitorización posterior. Estas prácticas se alinean con ISO/IEC 42001, ISO/IEC 23894 y NIST AI RMF y reducen la exposición probatoria.

TEST PRÁCTICO DE CLASIFICACIÓN

1) ¿El sistema influye en personas, producto o maquinaria? 2) ¿Es componente de seguridad de un producto armonizado? 3) ¿Ese producto exige evaluación de terceros? 4) ¿Se usa para empleo, biometría u otro caso del anexo III? 5) ¿La empresa modifica finalidad o realiza cambios sustanciales? 6) Aunque no sea alto riesgo, ¿qué consecuencias tendría un falso negativo? La última pregunta gobierna el nivel interno de control, incluso cuando las anteriores sean negativas.

6. La IA frente a PRP, PPRO/OPRP y PCC: clasificación funcional

La relación entre IA y APPCC debe analizarse por función de control. La siguiente matriz evita dos extremos: llamar PCC a todo sistema digital o tratar la IA como un accesorio ajeno al plan.

Capa	Función	Uso de IA	Condición de integración
PRP/GHP	Condiciones básicas del entorno y operación.	Analítica de cumplimiento, visión de limpieza, detección de puertas	No suele fijar por sí sola un límite crítico. Su fallo puede degradar la base higiénica y debe

Capa	Función	Uso de IA	Condición de integración
		abiertas, tendencias de plagas.	gestionarse en el PRP.
PPRO/OPRP	Medida de control para peligros significativos que no se gestiona como PCC.	Predicción de riesgo de contaminación cruzada, priorización de saneamiento, control de concentración o secuencia.	Necesita criterio de acción, monitorización y corrección; el score debe traducirse a regla operativa.
PCC	Etapa donde el control es esencial para prevenir, eliminar o reducir un peligro a nivel aceptable.	Modelo que supervisa una combinación de tiempo-temperatura o interpreta un detector con actuación crítica.	La medida de control y sus límites deben validarse; el algoritmo no puede ocultar el criterio de aceptabilidad.
Verificación	Confirma que el sistema funciona eficazmente.	Análisis de tendencias, detección de patrones, revisión de registros, selección inteligente de auditoría.	No debe convertirse silenciosamente en monitorización primaria sin revalidación.
Validación	Obtiene evidencia previa de que la medida puede controlar el peligro.	Modelos predictivos, simulación, gemelo digital, análisis de sensibilidad.	La simulación complementa, no sustituye, ensayos, literatura, *challenge tests* o evidencia científica pertinente.
Trazabilidad/retirada	Reconstruye origen, proceso y destino.	Resolución de entidades, detección de lotes vinculados, priorización de retirada.	La IA puede sugerir alcance; la decisión debe basarse en datos maestros íntegros y reglas de precaución.
Gestión documental	Organiza y busca información.	NLP, extracción de certificados, comparación de especificaciones, generación de borradores.	No debe crear datos faltantes ni convertir documentos no verificados en hechos.

El criterio decisivo es la dependencia. Si la empresa puede mantener el control mediante un límite convencional independiente y la IA solo añade una señal anticipada, estamos ante apoyo. Si la acción de control depende del resultado del modelo, el algoritmo forma parte de la medida. Si su fallo puede permitir que producto no seguro avance sin otra barrera, la función debe tratarse como crítica, aunque la clasificación formal del software no use esa palabra.

En un PCC, el límite crítico debe separar aceptabilidad de inaceptabilidad mediante un criterio observable o medible. Un score de 0,73 no es un límite crítico jurídicamente útil si nadie puede explicar qué representa, cómo se relaciona con el nivel aceptable del peligro y qué incertidumbre incorpora. Puede emplearse como señal complementaria o convertirse en criterio de acción si ha sido validado, documentado y vinculado a una regla conservadora. En muchos casos es más robusto mantener un límite físico duro y utilizar la IA para anticipar aproximación, detectar incoherencias o activar muestreo reforzado.

La IA puede ser especialmente valiosa en PPRO/OPRP, donde la realidad multivariable y la frecuencia de control no siempre encajan en un único límite crítico. Pero el resultado debe traducirse a una instrucción: qué nivel de riesgo activa qué medida, durante cuánto tiempo, sobre qué producto y con qué evidencia de cierre. Una alerta sin procedimiento es información; no es control.

7. Taxonomía de nuevos puntos ciegos algorítmicos

Tipo	Descripción	Manifestación
Ceguera de cobertura	La variable relevante no se mide o la muestra no representa la realidad.	Sensor en zona no crítica; imagen sin ángulo de sellado; laboratorio sin turno nocturno.
Ceguera de calidad	El dato existe, pero está descalibrado, retrasado, mal etiquetado o mezclado.	Temperatura con offset; lote mal vinculado; ausencia codificada como cero.
Ceguera estadística	La métrica global oculta el fallo en escenarios raros o críticos.	99,5 % de exactitud con baja sensibilidad para la clase de contaminación.
Ceguera de dominio	El modelo se aplica fuera de condiciones de entrenamiento.	Nueva formulación, envase, iluminación, proveedor, estación o velocidad de línea.
Ceguera de interacción	Varios componentes funcionan por separado, pero fallan juntos.	Sensor válido + reloj desincronizado + regla de agregación incorrecta.

Tipo	Descripción	Manifestación
Ceguera de interfaz	La presentación induce confianza excesiva o es ambigua.	Semáforo verde sin intervalo de confianza ni evidencia de datos incompletos.
Ceguera organizativa	No hay propietario, turno de respuesta o autoridad para actuar.	Alerta emitida fuera de horario; calidad cree que mantenimiento la gestiona.
Ceguera de actualización	El proveedor cambia modelo, umbral o infraestructura sin revisión.	Actualización SaaS que altera sensibilidad o tratamiento de valores faltantes.
Ceguera adversarial	Un tercero manipula datos, modelos o comunicaciones.	*Data poisoning,* credenciales robadas, inyección de imagen, r*ansomware* en OT.
Ceguera probatoria	La decisión no puede reconstruirse posteriormente.	No se conserva versión, entradas, reglas, usuario, hora ni motivo de *over-ride*.
Ceguera de automatización	El humano acepta por defecto o no puede contradecir.	Revisión nominal de cientos de alertas sin tiempo ni datos alternativos.
Ceguera de retroalimentación	La propia acción cambia los datos y aparenta mejorar el modelo.	Solo se muestrean zonas de alto score; se pierde información sobre otras áreas.

Los falsos negativos merecen atención especial. En un entorno comercial, la presión suele favorecer métricas equilibradas o reducción de falsas alarmas. En seguridad alimentaria, el coste de no detectar una desviación puede ser muy superior al de retener temporalmente un lote. La función de pérdida del modelo y el umbral operativo deben reflejar esa asimetría. No significa maximizar sensibilidad sin límite —lo que haría el sistema inutilizable—, sino definir conscientemente el riesgo residual y las barreras compensatorias.

La incertidumbre debe hacerse visible. Un modelo puede estar bien calibrado en datos conocidos y ser inseguro cuando recibe entradas fuera de distribución. En vez de forzar una clasificación, debería poder declarar “no sé”, derivar a revisión humana o activar un modo conservador. La abstención controlada es una característica de seguridad, no un defecto comercial.

La ceguera de retroalimentación es frecuente en sistemas predictivos. Si el modelo dirige el muestreo hacia zonas de mayor score, se recopilan más positivos allí y menos información en zonas de bajo score. El modelo puede confirmar sus propias hipótesis y dejar de detectar nuevos patrones. Debe reservarse una fracción de muestreo aleatorio o basado en diseño higiénico, independiente de la recomendación algorítmica.

8. Método HAZARD-AI: análisis integrado de peligros y fallos digitales

Se propone un método de diez pasos para integrar el sistema algorítmico en el análisis de peligros. No pretende crear un segundo APPCC burocrático, sino añadir preguntas que el APPCC clásico no formulaba porque asumía controles deterministas y observables.

1. Definir la finalidad prevista y las decisiones que el sistema puede influir, incluyendo usos prohibidos y usos previsiblemente incorrectos.
2. Mapear el flujo de datos desde la fuente física hasta la acción: sensores, comunicaciones, transformación, almacenamiento, modelo, interfaz, decisión y actuador.
3. Identificar el peligro alimentario que se pretende controlar y la capa del FSMS en la que se integra: PRP, PPRO/OPRP, PCC, validación, verificación o trazabilidad.
4. Identificar modos de fallo algorítmico y digital: omisión, error, retraso, deriva, indisponibilidad, manipulación, incertidumbre, error humano inducido y pérdida de evidencia.
5. Valorar severidad, probabilidad y detectabilidad del fallo, pero también exposición: volumen de producto, población sensible, reversibilidad y existencia de barreras posteriores.
6. Definir controles preventivos del sistema: calidad de datos, calibración, límites duros, redundancia, abstención, segregación de funciones, control de acceso y pruebas adversariales.
7. Establecer criterios operativos: umbral, intervalo de confianza, tiempo máximo de respuesta, reglas de escalado, destino provisional del producto y autoridad para over-ride.

8. Validar en condiciones reales y desfavorables, por subgrupos y escenarios, comparando con método de referencia y con la práctica existente.
9. Diseñar monitorización y verificación continuas: desempeño, deriva, calidad de datos, alarmas ignoradas, cambios de versión, incidentes y eficacia de acciones.
10. Documentar el caso de garantía y revisar tras cambios, incidentes, tendencias o nueva evidencia científica y normativa.

8.1. Matriz de criticidad algorítmica

Clase	Influencia	Garantía mínima
A - Asistencia	No cambia el destino del producto ni parámetros; salida reversible.	Revisión normal; validación de utilidad; prohibición de presentar como hecho no verificado.
B - Priorización	Asigna orden de muestreo, revisión o mantenimiento.	Muestreo independiente mínimo; revisión de sesgo de cobertura; monitorización de omisiones.
C - Recomendación operativa	Propone acción sobre proceso o producto, pero requiere aprobación.	Supervisión competente; acceso a evidencia; tiempo real; registro de aceptación/(over-ride).
D - Decisión automatizada	Bloquea, clasifica o libera con intervención humana limitada.	Validación reforzada; doble barrera; reglas de apelación; registros íntegros; modo seguro.
E - Control físico autónomo	Modifica parámetros o acciona equipo relevante para seguridad.	Ingeniería de seguridad; límites independientes; análisis de maquinaria/AI Act; pruebas de fallo y ciberseguridad.

8.2. Registro de riesgos AI-FSMS

ID	Modo de fallo	Consecuencia	Criticidad	Control	Propietario	Verificación
R-01	Deriva de sensor de temperatura	Subestimación de exposición térmica	Alta	Calibración, plausibilidad cruzada, alarma de offset	Mantenimiento/Calidad	Mensual + tras intervención
R-02	Cambio de formulación no comunicado	Modelo fuera de dominio	Alta	Gestión de cambios vinculada a *MLOps*; bloqueo de versión	I+D/Calidad/IT	Antes de lanzamiento
R-03	Falso negativo de visión	Envase defectuoso no retenido	Alta	Muestra desafío, doble inspección, límites de iluminación	Producción/Calidad	Por turno y tendencia
R-04	Proveedor actualiza modelo	Cambio no validado	Alta	Cláusula de preaviso, entorno de prueba, *rollback*	Compras/IT/Calidad	Cada versión
R-05	*Ransomware* o pérdida de nube	Pérdida de monitorización y registros	Alta	Modo degradado, *backup* inmutable, continuidad	CISO/Operaciones	Simulacro anual
R-06	Automatización sesgada	Aceptación acrítica	Media-Alta	Formación, autoridad de over-ride, KPI de discrepancias	Dirección/Calidad	Trimestral

9. Integración de IA en los siete principios APPCC

9.1. Principio 1: análisis de peligros

La IA puede explorar grandes volúmenes de incidencias, resultados de laboratorio, reclamaciones, alertas regulatorias, variables de proceso y datos de proveedores. Puede detectar asociaciones y priorizar hipótesis. Sin embargo, la significación estadística no equivale a causalidad ni a relevancia para seguridad alimentaria. El equipo APPCC debe mantener competencia en microbiología, química, tecnología de proceso, alérgenos, ingeniería, mantenimiento, datos y Derecho.

El uso más sólido consiste en una secuencia humana-máquina-humana: el equipo define producto, uso previsto, población, proceso y peligros plausibles; el sistema analiza datos y propone patrones; el equipo contrasta con evidencia científica, diseño higiénico y experiencia; y la decisión queda documentada con

fuentes. La IA generativa puede ayudar a localizar vacíos, pero toda referencia debe verificarse en la fuente original.

CONTROL CRÍTICO DE CALIDAD

Nunca debe aceptarse un peligro, límite o medida de control porque “lo dice el modelo”. Debe existir una cadena de fundamentación: fuente científica o legal, aplicabilidad al producto/proceso, hipótesis, evidencia y decisión del equipo competente.

9.2. Principio 2: determinación de PCC

Un algoritmo puede asistir en la aplicación de árboles de decisión, pero la determinación de PCC exige juicio sobre esencialidad, barreras previas y posteriores, consecuencias de fallo y capacidad real de monitorización. Un sistema que aprende de planes históricos puede reproducir errores y variabilidad sectorial. La salida debe ser una propuesta trazable, no una clasificación automática. Si el algoritmo controla una etapa ya determinada como PCC, debe analizarse como componente de la medida de control.

9.3. Principio 3: límites críticos

Los límites críticos deben basarse en parámetros observables o medibles que separen aceptación e inaceptación. La IA puede combinar variables y estimar un estado latente, pero la organización debe definir el vínculo entre el score y el peligro. Siempre que sea viable, conviene conservar límites físicos o microbiológicos independientes. Cuando el límite sea algorítmico, la validación debe incluir calibración, sensibilidad, especificidad, incertidumbre, condiciones de abstención y trazabilidad de versión.

9.4. Principio 4: monitorización

La monitorización algorítmica ofrece continuidad y detección multivariante. Su debilidad es la opacidad: puede emitir un resultado sin que el operario perciba datos faltantes o degradados. El panel debe mostrar estado de sensores, frescura del dato, versión, confianza, razón de alerta y acción requerida. La ausencia de alerta no debe interpretarse como evidencia de control si el sistema está desconectado o fuera de dominio.

9.5. Principio 5: acciones correctivas

La acción correctiva debe separar dos problemas: recuperar el control del proceso y decidir el destino del producto potencialmente afectado. Una alerta predictiva puede emitirse antes de cruzar un límite crítico; en ese caso la acción es preventiva y no debería confundirse con desviación. Si el límite se ha superado, el score no puede “compensar” la no conformidad salvo que exista una evaluación científica y jurídica expresamente prevista. Debe registrarse quién actuó, cuándo, sobre qué lotes, con qué evidencia y qué medidas evitarán recurrencia.

9.6. Principio 6: verificación

La verificación debe evaluar tanto el APPCC como el sistema algorítmico. Incluye revisión de resultados, auditorías, calibraciones, muestreo independiente, análisis de alarmas, comparación con laboratorio, pruebas de recuperación, control de cambios y revisión de desempeño por subgrupos. Una métrica estable no demuestra por sí sola eficacia; puede existir degradación compensada por cambios en prevalencia o selección de datos.

9.7. Principio 7: documentación y registros

La trazabilidad mínima de una decisión algorítmica debería conservar: fecha y hora; producto, lote y línea; datos de entrada relevantes y su calidad; versión del modelo y configuración; salida y nivel de confianza; regla operativa aplicada; usuario o proceso que decidió; acción; *over-ride* y motivo; evidencia de cierre; y vínculos a registros originales. Los logs deben ser protegidos frente a alteración, sincronizados temporalmente y conservados según el riesgo, vida útil, plazos legales, contratos y necesidades probatorias.

10. Alertas predictivas: del score a la acción controlada

La alerta predictiva es uno de los usos más rentables y, a la vez, más mal diseñados. Un sistema puede producir cientos de avisos sin mejorar la seguridad si no existe una arquitectura de respuesta. La alerta eficaz no es una notificación; es un evento gobernado que modifica temporalmente el nivel de control.

Cada alerta debe tener cinco atributos: significado, severidad, horizonte temporal, evidencia y acción. “Riesgo alto” carece de utilidad si no se sabe qué peligro anticipa, cuánto tiempo existe antes de la posible pérdida de control, qué datos la sustentan y qué debe hacerse. El algoritmo debe evitar precisión falsa. Es preferible una categoría bien validada con intervalo de incertidumbre que un porcentaje con apariencia científica pero sin calibración.

10.1. Arquitectura de cuatro niveles

Nivel	Significado	Respuesta	Producto
Nivel 0 - Información	Tendencia sin evidencia suficiente de pérdida de control.	Revisión en rutina; no afecta producto.	Registrar para análisis.
Nivel 1 - Atención	Desviación incipiente o calidad de dato degradada.	Comprobar sensor/proceso dentro del turno.	No requiere retención salvo otros indicios.
Nivel 2 - Intervención preventiva	Probabilidad significativa de aproximación a límite o pérdida de PRP/PPRO.	Ajustar, limpiar, muestrear, aumentar frecuencia, informar a calidad.	Identificar producto expuesto y mantener trazabilidad.
Nivel 3 - Seguridad	Posible pérdida de control, falso negativo crítico o fallo del sistema sin barrera.	Retener/segregar, detener o pasar a modo seguro; activar investigación.	Decisión de producto por persona autorizada y evidencia independiente.
Nivel 4 - Incidente	Confirmación de producto potencialmente no seguro o pérdida grave de integridad.	Procedimiento de crisis, trazabilidad, notificación/retirada cuando proceda.	Preservar pruebas y cooperación regulatoria.

La gestión de alertas debe medir no solo cuántas se generan, sino su utilidad y respuesta: tiempo hasta reconocimiento, tiempo hasta acción, tasa de escalado, tasa de alertas ignoradas, repetición, confirmación por evidencia, falsos negativos descubiertos por otros canales y efecto sobre producto. Reducir alertas puede ser positivo si se eliminan duplicados; puede ser peligroso si se eleva el umbral para mejorar la comodidad operativa.

Las alertas de calidad de dato tienen prioridad. Un modelo que no recibe información fiable no debe continuar mostrando un semáforo verde. La interfaz debe diferenciar “riesgo bajo” de “incapacidad para evaluar”. En sistemas críticos, la pérdida de datos debe activar una regla conservadora: control convencional reforzado, retención o parada según el caso.

11. Decisiones automatizadas y supervisión humana significativa

La supervisión humana no se satisface añadiendo una casilla de aprobación. Debe ser material: la persona necesita competencia, tiempo, autoridad, información y posibilidad real de apartarse de la salida. La literatura sobre automatización muestra el riesgo de *automation bias*: aceptar la recomendación porque el sistema parece objetivo o porque contradecirlo exige más esfuerzo y responsabilidad.

En seguridad alimentaria deben distinguirse tres modelos de decisión. En el primero, la IA informa y la persona decide. En el segundo, la IA decide provisionalmente —por ejemplo, retiene— y una persona revisa antes de liberar. En el tercero, la IA actúa de forma autónoma dentro de límites físicos y la persona supervisa excepciones. El modelo adecuado depende de reversibilidad. Retener automáticamente suele ser más aceptable que liberar automáticamente, porque la retención es conservadora y reversible; la liberación expone el producto y puede resultar irreversible.

La autoridad de over-ride debe diseñarse en ambos sentidos. Un usuario puede anular una falsa alarma para evitar desperdicio, pero también debe poder imponer una retención, aunque el modelo indique bajo

riesgo. Todo *over-ride* debe registrar identidad, motivo, evidencia y resultado posterior. Una alta tasa de *over-ride* no es necesariamente mala; puede revelar cambio de dominio, mala interfaz o conocimiento experto que el modelo no incorpora. Ocultarla para proteger el KPI destruye aprendizaje y evidencia.

11.1. Test de supervisión significativa

Dimensión	Pregunta	Evidencia esperada
Comprensión	¿El supervisor entiende finalidad, límites, datos faltantes y métricas relevantes?	Formación específica y evaluación de competencia.
Información	¿Ve las variables y evidencias necesarias o solo un color?	Panel con trazabilidad, confianza, calidad de datos y acceso a registros.
Tiempo	¿Dispone de tiempo real para revisar antes de que la acción sea irreversible?	SLA y bloqueo temporal del proceso/producto.
Autoridad	¿Puede detener, retener, escalar o contradecir sin penalización indebida?	Matriz RACI, delegaciones y cultura de seguridad.
Independencia	¿Existe evidencia independiente o una segunda barrera?	Muestreo, regla física, verificación o doble aprobación.
Registro	¿La decisión humana y su fundamento quedan documentados?	Log de decisión, motivo y resultado.
Aprendizaje	¿Los desacuerdos alimentan revisión del modelo y del APPCC?	Revisión periódica de over-rides y falsos negativos.

Cuando el sistema afecta a trabajadores —asignación, productividad, disciplina, acceso o evaluación— deben añadirse las garantías laborales y de protección de datos. No debe reutilizarse silenciosamente un sistema instalado para seguridad alimentaria como mecanismo de vigilancia de personal. La ampliación de finalidad exige evaluación jurídica, información y, según el caso, consulta a representantes.

Las decisiones generadas por IA sobre producto deben integrarse en la matriz de autoridad existente. La liberación final de producto no conforme, la aceptación de una desviación o la modificación de un límite no deben quedar en manos de una persona sin competencia ni de un algoritmo. Deben existir criterios de escalado a dirección de calidad, responsable técnico y, cuando corresponda, asesoría jurídica o dirección general.

12. Validación, verificación, recalibración y gestión de deriva

Validar un modelo no es demostrar que funciona en un conjunto histórico. En seguridad alimentaria, la validación debe demostrar que el sistema es adecuado para la finalidad prevista y las condiciones reales de uso, incluida la variabilidad esperable y los escenarios adversos. Debe compararse con un método de referencia, con expertos o con un control convencional, según la función.

La validación previa debe definir población de datos, prevalencia, clases, criterios de inclusión, tratamiento de ausencias, partición temporal, prevención de fuga de datos, métricas y umbrales. En eventos raros, una partición aleatoria puede inflar resultados porque lotes o condiciones similares aparecen en entrenamiento y prueba. Conviene separar por tiempo, instalación, proveedor o campaña y realizar validación externa cuando el modelo se desplegará en contextos distintos.

Las métricas deben alinearse con el riesgo: sensibilidad para detectar la condición peligrosa, valor predictivo negativo, tasa de falsos negativos, tiempo de anticipación, calibración, tasa de abstención y desempeño bajo degradación. La exactitud global puede ser irrelevante. También deben medirse consecuencias operativas: retenciones innecesarias, carga de alertas y tiempo de respuesta.

12.1. Plan mínimo de validación

Bloque	Contenido	Evidencia
Finalidad y criterio de éxito	Decisión concreta que apoya; peligro; etapa; usuario; acción.	Protocolo aprobado por Calidad y propietario del proceso.
Datos	Origen, representatividad, calidad, etiquetas, exclusiones, sesgos.	Data sheet y trazabilidad de conjuntos.
Referencia	Método de laboratorio, regla física, inspección experta o resultado confirmado.	Justificación de método y competencia.
Diseño	Separación temporal/externa, escenarios normales y extremos, muestras desafío.	Plan estadístico previo.
Métricas	Sensibilidad, especificidad, FNR, calibración, latencia, abstención.	Criterios de aceptación y límites de confianza.
Robustez	Datos faltantes, ruido, cambio de sensor, ciberataque, interrupción.	Pruebas de estrés y respuesta segura.
Human factors	Comprensión, carga, errores de interpretación, over-ride.	Pruebas de usabilidad con usuarios reales.
Piloto	Operación paralela sin sustituir control vigente.	Informe de concordancia y desviaciones.
Aprobación	Riesgo residual, barreras y autorización.	Acta de equipo APPCC y gestión de cambios.
Monitorización	Indicadores, frecuencia, *triggers* de revalidación y *rollback.*	Plan post-despliegue.

12.2. Deriva de datos, concepto y desempeño

La deriva de datos aparece cuando cambia la distribución de entradas; la deriva de concepto, cuando cambia la relación entre entradas y resultado; y la deriva de desempeño, cuando caen las métricas observadas. En alimentos, puede deberse a estacionalidad, proveedores, formulaciones, equipos, limpieza, microbiota, clima, materias primas o comportamiento humano. No toda deriva exige reentrenar; puede requerir corregir un sensor, revisar un PRP o modificar la finalidad.

Los *triggers* de revalidación deben incluir: cambio de producto o proceso; nueva línea o planta; nueva versión; caída de sensibilidad; aumento de abstenciones; cambio de prevalencia; incidente; reclamación; cambio legal; modificación de método de referencia; ciber-incidente; y patrón sostenido de over-rides. El reentrenamiento automático en producción es especialmente delicado. Un modelo que aprende de decisiones no verificadas puede consolidar errores. En funciones relevantes para seguridad, el nuevo modelo debe pasar por un entorno controlado, comparación y aprobación antes de sustituir la versión vigente.

12.3. *Champion-challenger* y *rollback*

Una práctica robusta es mantener un modelo "*champion*" validado y evaluar versiones "*challenger*" en sombra. El *challenger* no decide hasta demostrar mejora sin deteriorar escenarios críticos. Debe existir *rollback* rápido a la versión previa y conservación de la configuración exacta. La frase "modelo continuamente mejorado" no es una garantía; sin gobierno puede significar comportamiento continuamente cambiante e imposible de auditar.

13. Datos, sensores, trazabilidad y cadena de custodia digital

La calidad de la inferencia no puede superar la calidad útil de los datos. En seguridad alimentaria, además, el dato debe ser íntegro, contextualizado y atribuible. Una temperatura sin identificación de sensor, calibración, posición, hora y lote puede ser numéricamente correcta pero probatoriamente inútil. La gobernanza debe abarcar datos maestros, eventos, metadatos, sincronización temporal y relaciones entre sistemas MES, LIMS, ERP, SCADA, QMS y plataforma de IA.

Los principios ALCOA+ —atribuible, legible, contemporáneo, original, exacto, completo, consistente, duradero y disponible— ofrecen una referencia útil, aunque procedan de entornos regulados distintos. En IA

se añade la necesidad de reproducibilidad: poder reconstruir qué versión y parámetros transformaron qué entradas en qué salida. No siempre será posible reproducir bit a bit un servicio externo; el contrato debe exigir al menos logs suficientes, versión identificable y cooperación en investigación.

La cadena de custodia digital debe proteger la evidencia desde captación hasta archivo. Hashes, firmas, controles de acceso, registros inmutables, *backups*, segregación y sellado temporal pueden ser apropiados según criticidad. *Blockchain* no es requisito ni solución universal; puede garantizar inmutabilidad de una entrada falsa. La prioridad es asegurar identidad, integridad, contexto y gobernanza de la fuente.

13.1. Controles de datos maestros

Objeto	Control	Riesgo que reduce
Identidad de producto/lote	Código único, versión de receta, proveedor, destino.	Evita mezclar poblaciones y permite retirada.
Sensor	ID, ubicación, rango, calibración, firmware, fecha de intervención.	Permite valorar validez y deriva.
Tiempo	Fuente de reloj, zona horaria, sincronización y latencia.	Reconstruye secuencia causal.
Etiquetas de entrenamiento	Definición, método, responsable, incertidumbre y conflictos.	Evita aprender de clases ambiguas.
Valores faltantes	Causa, codificación y regla de tratamiento.	Impide confundir ausencia con normalidad.
Versiones	Modelo, reglas, umbrales, software, librerías y configuración.	Reproducibilidad y control de cambios.
Retención	Plazo, formato, accesibilidad y destrucción segura.	Cumplimiento y defensa probatoria.
Interoperabilidad	Diccionario, unidades y transformaciones.	Evita errores de escala y semántica.

La trazabilidad alimentaria del artículo 18 del Reglamento 178/2002 no exige necesariamente una tecnología concreta, pero sí capacidad de identificar proveedores y clientes y organizar la información. Cuando la IA se utiliza para resolver relaciones entre lotes o estimar alcance de incidente, debe mantener una salida conservadora: una baja confianza no justifica excluir lotes potencialmente afectados. La decisión de alcance debe poder ampliarse por precaución.

En sistemas de visión, deben conservarse muestras representativas de imágenes, condiciones de iluminación y resultados, respetando protección de datos. En modelos basados en laboratorio, debe preservarse el vínculo con el resultado original, método, límite de detección y estado de acreditación. En modelos de proveedor, debe definirse quién puede acceder a datos y si se usan para reentrenar servicios compartidos.

14. Ciberseguridad OT/IT y continuidad del autocontrol

La convergencia entre tecnología operacional y sistemas de información convierte la ciberseguridad en una medida de seguridad alimentaria. Un atacante no necesita contaminar físicamente un producto si puede alterar una receta, desactivar una alarma, cambiar unidades, manipular un certificado, borrar resultados o cifrar los sistemas que permiten retener lotes. La evaluación de peligros debe contemplar amenazas intencionadas que puedan materializar peligros alimentarios, en coordinación con food defense y fraude.

NIS2 exige a las entidades incluidas medidas de gestión de riesgos de ciberseguridad y notificación. El sector alimentario industrial y mayorista figura en su ámbito sectorial, sujeto a tamaño y transposición. El *Cyber Resilience Act* establece requisitos para productos con elementos digitales comercializados y será relevante para proveedores de equipos y software. El Reglamento de Máquinas incorpora riesgos derivados de comportamiento auto-evolutivo y conectividad. Estas normas convergen en una exigencia empresarial: seguridad por diseño, gestión de vulnerabilidades y continuidad.

14.1. Amenazas específicas de IA

Amenaza	Mecanismo	Control
Data poisoning	Introducir datos falsos en entrenamiento o retroalimentación.	Separación de datos, aprobación de etiquetas, detección de anomalías, trazabilidad.
Adversarial examples	Modificar imágenes/señales para inducir clasificación errónea.	Pruebas adversariales, redundancia física, control de entorno.
Model tampering	Sustituir pesos, reglas o umbrales.	Firma de artefactos, control de acceso, verificación de integridad.
API/SaaS *compromise*	Manipular servicio o robar credenciales.	MFA, mínimo privilegio, *allowlists*, monitorización y plan de contingencia.
Ransomware	Indisponibilidad de monitorización y registros.	Segmentación, *backups* inmutables, restauración probada, modo degradado.
Prompt injection	Manipular IA generativa mediante documentos o entradas.	Aislamiento, listas de fuentes, revisión humana, no otorgar acciones críticas.
Supply-chain attack	Biblioteca, firmware o actualización comprometida.	SBOM cuando proceda, gestión de proveedores, pruebas y firma.
Clock/log manipulation	Alterar secuencia y evidencia.	NTP seguro, sellado temporal, logs centralizados e inmutables.

14.2. Modo degradado y continuidad

El plan de continuidad debe definir cómo se mantiene la seguridad cuando la IA, la nube, la red o los sensores fallan. Las opciones incluyen volver a límites y registros convencionales, aumentar frecuencia de medición manual, reducir velocidad, retener producto, suspender liberación o parar. El modo degradado debe ensayarse. Un procedimiento que existe solo en papel y requiere recursos no disponibles durante un ciber-incidente no es una barrera real.

La recuperación debe preservar el estado de lotes y decisiones. Tras restaurar, no debe asumirse que todo periodo sin visibilidad estuvo bajo control. Se requiere reconstrucción, evaluación de producto y documentación de incertidumbre. La presión por reanudar producción no puede convertir datos ausentes en evidencia favorable.

15. Auditoría, inspección oficial y prueba de cumplimiento

La autoridad competente puede exigir evidencia de cumplimiento del APPCC y examinar procedimientos, registros, instalaciones y prácticas. La opacidad contractual del proveedor no es una defensa suficiente. La empresa debe seleccionar sistemas cuya información permita demostrar que el control existe y funciona. La explicación necesaria no siempre implica revelar código fuente; puede consistir en finalidad, variables, lógica operativa, métricas, límites, validación, versión, registros y controles.

La auditoría debe diferenciar tres preguntas: conformidad del diseño, conformidad de la operación y eficacia. El diseño puede estar bien documentado y no aplicarse; la operación puede seguir el procedimiento y el modelo haber perdido desempeño; y el modelo puede ser preciso, pero no reducir el riesgo porque las alertas llegan tarde. La auditoría debe seguir la cadena desde un caso real hasta la evidencia original.

15.1. Paquete de evidencia exigible

Bloque	Contenido	Finalidad probatoria
Gobernanza	Inventario, clasificación, propietario, RACI, política y aprobación.	Demuestra responsabilidad y alcance.
Finalidad	Uso previsto, exclusiones, usuarios, acciones y límites.	Evita ampliación silenciosa de función.

Bloque	Contenido	Finalidad probatoria
Datos	Fuentes, diccionario, calidad, representatividad y protección.	Sustenta validez.
Validación	Protocolo, resultados, escenarios, métricas y aceptación.	Prueba aptitud previa.
Operación	SOP, formación, panel, alertas, *overrides*, mantenimiento.	Prueba aplicación efectiva.
Versiones	Historial, cambios, pruebas, aprobación y *rollback*.	Controla modificaciones.
Monitorización	Drift, desempeño, incidentes, falsos negativos y CAPA.	Prueba mantenimiento de eficacia.
Ciberseguridad	Arquitectura, accesos, vulnerabilidades, *backups* y simulacros.	Prueba resiliencia.
Proveedor	Contrato, SLA, sub-encargados, actualizaciones, auditoría y seguro.	Prueba control de terceros.
Decisión concreta	Entradas, salida, usuario, acción, lote y evidencia de cierre.	Permite reconstrucción probatoria.

En una inspección, la empresa debe evitar dos respuestas insatisfactorias: “el proveedor dice que funciona” y “es un algoritmo propietario”. La primera externaliza la responsabilidad; la segunda impide demostrar control. Debe existir un expediente interno de aceptación con evidencia suficiente y un mecanismo para obtener apoyo del proveedor.

La prueba de cumplimiento requiere conservar también evidencia desfavorable: alarmas ignoradas, fallos, over-rides, reentrenamientos y discrepancias. Borrar o no registrar excepciones produce una imagen artificialmente limpia y debilita la defensa. La diligencia se demuestra mejor mostrando detección, investigación y corrección que pretendiendo ausencia absoluta de fallos.

16. Cultura de seguridad alimentaria y alfabetización en IA

El Reglamento 2021/382 incorporó requisitos de cultura de seguridad alimentaria: compromiso de dirección y empleados, liderazgo, conciencia, comunicación y recursos. La IA puede fortalecer esa cultura si facilita información y aprendizaje; también puede erosionarla si centraliza decisiones en una “caja negra”, desincentiva preguntas o convierte el cumplimiento en un semáforo.

La alfabetización en IA exigida por el artículo 4 del AI Act debe adaptarse al rol. El consejo de dirección necesita comprender exposición, responsabilidad, inversión y límites. El equipo APPCC necesita entender datos, métricas, deriva y validación. Los operarios necesitan interpretar alertas, calidad de dato y modo degradado. IT/OT necesita comprender peligros alimentarios y consecuencias de cambios técnicos. Compras necesita evaluar contratos y dependencia. Jurídico necesita traducir clasificación normativa a controles reales.

16.1. Currículo mínimo por función

Rol	Contenido	Objetivo
Dirección	Responsabilidad, riesgo residual, apetito de riesgo, inversión, crisis y reporte.	Decidir con conocimiento y asignar recursos.
Equipo APPCC/Calidad	Finalidad, datos, métricas, validación, *drift*, falsos negativos, evidencia.	Integrar el algoritmo en el FSMS.
Producción	Interpretación de alertas, comprobaciones, acciones, *override,* escalado.	Evitar confianza ciega y demoras.
Mantenimiento/OT	Sensores, calibración, integridad, firmware, interbloqueos y continuidad.	Preservar capacidad de control.
IT/Ciberseguridad	Arquitectura, acceso, logs, vulnerabilidades, *backup* y respuesta.	Proteger disponibilidad e integridad.

Rol	Contenido	Objetivo
I+D	Gestión de cambios, dominio de aplicación, datos y revalidación.	Evitar deriva por innovación no comunicada.
Compras/Jurídico	Clasificación, SLA, datos, IP, auditoría, actualización, responsabilidad.	Controlar dependencia del proveedor.
Auditor interno	Muestreo de decisiones, evidencia, eficacia y pruebas de fallo.	Verificar diseño y operación.

Los indicadores de cultura deben incorporar el comportamiento frente a la IA: porcentaje de alertas revisadas en plazo, calidad de razones de over-ride, incidentes comunicados, participación en pruebas, dudas planteadas y capacidad de operar sin el sistema. Penalizar toda discrepancia con el algoritmo destruye la supervisión humana. La dirección debe premiar la detección temprana y la transparencia.

La formación no puede ser genérica. Un curso sobre conceptos de IA no habilita para liberar producto. La competencia debe evaluarse mediante casos, simulaciones, interpretación de métricas y respuesta a fallos. La autoridad para tomar decisiones críticas debe vincularse a esa competencia.

17. Contratación de proveedores y reparto de responsabilidades

La mayor parte de las empresas alimentarias comprará soluciones de terceros. El contrato es una medida de control, pero no sustituye la diligencia técnica. Debe impedir que el proveedor cambie silenciosamente el comportamiento, utilice datos de planta para fines incompatibles o limite tanto la información que la empresa no pueda demostrar cumplimiento.

La *due diligence* debe evaluar madurez del proveedor, historial, seguridad, metodología de validación, soporte, continuidad, subcontratistas y capacidad de exportar datos. Las demostraciones comerciales suelen realizarse en condiciones favorables y con métricas agregadas. Debe exigirse una prueba con datos y escenarios propios, incluyendo casos raros y condiciones degradadas.

17.1. Cláusulas esenciales

Materia	Contenido mínimo
Finalidad prevista	Descripción exacta de función, usuarios, producto, línea, decisiones y exclusiones.
Desempeño	Métricas por escenarios, método de medición, niveles mínimos y tratamiento de incertidumbre.
Cambios	Preaviso, documentación, pruebas, aprobación, compatibilidad y *rollback*.
Datos	Titularidad, roles RGPD, localización, retención, uso para entrenamiento, exportación y borrado.
Logs y auditoría	Acceso a registros, versiones, incidentes, evidencias y derecho de auditoría.
Ciberseguridad	Medidas, vulnerabilidades, parches, notificación, SBOM cuando proceda y cooperación.
Continuidad	SLA, recuperación, portabilidad, *escrow* si crítico, soporte y fin de servicio.
Responsabilidad	Garantías, indemnidad, límites razonables, seguros y cooperación ante reclamaciones.
Regulación	Obligaciones según rol *AI Act*/CRA/maquinaria y soporte documental.
Subcontratación	Lista, cambios, flujo de obligaciones y responsabilidad.
Propiedad intelectual	Derechos sobre modelo, configuraciones, datos derivados y desarrollos específicos.
Salida	Migración, entrega de datos, conservación de evidencia y eliminación segura.

Los límites de responsabilidad deben revisarse críticamente. Una cláusula que limita toda responsabilidad al importe mensual puede ser incompatible con la exposición real si el sistema interviene en liberación o control. La negociación debe distinguir daños directos, retirada, destrucción, interrupción, sanciones,

reclamaciones de clientes y daños personales. No todas las partidas serán asegurables o trasladables, pero deben conocerse.

El contrato debe definir la frontera entre configuración y modificación sustancial. Si la empresa reentrena o altera la finalidad, puede asumir obligaciones de proveedor. El suministrador debe advertir qué cambios invalidan la validación y cómo documentarlos. La ausencia de esta frontera genera un espacio de irresponsabilidad compartida.

18. Casos de uso sectoriales y criterios de aceptación

18.1. Cadena de frío y vida útil

Un modelo puede integrar perfiles tiempo-temperatura, aperturas, carga, rutas y características de producto para anticipar riesgo o vida útil remanente. Debe diferenciar estimación de calidad de evaluación de seguridad. La vida útil microbiológica requiere modelos y validación aplicables a matriz, pH, aw, atmósfera, flora, envase y condiciones. No debe utilizarse una predicción para ampliar fecha de caducidad sin un estudio formal. El criterio de aceptación incluye calibración por producto, detección de sensores anómalos, límites conservadores y reglas ante datos incompletos.

18.2. Tratamientos térmicos

La IA puede anticipar desviación, optimizar proceso o estimar letalidad. Si ajusta automáticamente parámetros, forma parte de la medida de control. Deben mantenerse límites duros, interbloqueos y cálculo validado independiente. La optimización no puede reducir margen de seguridad sin revalidación. Deben probarse fallos de sensor, latencia, receta incorrecta y cambio de carga.

18.3. Visión para envases, etiquetas y alérgenos

La visión es adecuada para verificar presencia, posición, código, sellado y correspondencia producto-etiqueta. El principal riesgo es el falso negativo por iluminación, reflejo, velocidad, variación de diseño o suciedad de lente. Para alérgenos, la consecuencia puede ser grave. Se requiere conjunto desafío, pruebas por variante, limpieza de lente, verificación de rechazo físico y confirmación de que el sistema falla de forma segura. Las imágenes con trabajadores deben gestionarse bajo RGPD.

18.4. Monitorización ambiental de Listeria

Los modelos pueden priorizar zonas y momentos de muestreo a partir de historial, flujo, humedad, temperatura, limpieza, mantenimiento y tráfico. No deben reemplazar el diseño higiénico ni eliminar muestreo aleatorio. La validación debe demostrar que aumenta la probabilidad de detección sin crear zonas no observadas. El modelo debe revisarse tras obras, cambios de equipo, positivos, nuevas rutas o estacionalidad.

18.5. Detectores de cuerpos extraños y clasificación

La IA puede reducir falsas alarmas en visión, rayos X o señales espectrales. Cuando decide el rechazo de producto, debe validarse por tamaño, material, posición, velocidad, densidad y producto. Debe conservarse un programa de prueba, *pieces* o muestras patrón. El algoritmo no debe filtrar señales que antes se consideraban críticas sin evidencia de equivalencia o mejora.

18.6. Fraude, proveedores y documentación

Modelos de anomalía pueden identificar precios, certificados, composición, rutas o resultados atípicos. Son herramientas de priorización, no prueba de fraude. Deben evitar sesgos contra proveedores pequeños o nuevos y permitir revisión. La IA generativa puede extraer datos de certificados, pero la autenticidad del documento y la validez del laboratorio requieren comprobación independiente.

18.7. Mantenimiento predictivo

El mantenimiento predictivo se convierte en seguridad alimentaria cuando el fallo del equipo afecta higiene, temperatura, dosificación, sellado o detección. El modelo debe integrarse con criticidad de activos. Una predicción de vida útil no justifica aplazar una intervención obligatoria. Las órdenes de mantenimiento y su impacto sobre limpieza, calibración y liberación deben quedar vinculadas.

Caso	Función	Punto ciego	Criterio de aceptación
Cadena de frío	Predicción de riesgo/vida útil	Confundir calidad con seguridad; datos incompletos	Validación por matriz y perfil; reglas conservadoras
Tratamiento térmico	Anticipación/optimización/control	Reducción de margen; sensor o receta errónea	Límite duro independiente; *challenge* y *fail-safe*
Etiquetado/alérgenos	Visión y correspondencia	Falso negativo por iluminación o variante	Conjunto desafío; rechazo verificado; prueba por SKU
Listeria ambiental	Priorización de muestreo	Sesgo de cobertura y retroalimentación	Cuota aleatoria; revisión por cambios y positivos
Cuerpos extraños	Clasificación/rechazo	Filtrado de señal crítica	Muestras patrón y validación multidimensional
Proveedores/fraude	Anomalías y documentación	Falsas acusaciones; documentos no auténticos	Revisión humana y evidencia independiente
Mantenimiento	Fallo anticipado	Aplazar intervención o perder barrera	Criticidad de activo y vínculo a APPCC

19. Hoja de ruta de implantación y modelo económico

La implantación debe comenzar por un problema concreto, no por la compra de una plataforma. Los mejores casos presentan peligro o pérdida relevantes, datos accesibles, una acción posible, un responsable claro y una métrica económica. La IA no corrige un proceso sin disciplina, sensores mal mantenidos o registros incoherentes; puede amplificar sus defectos.

19.1. Programa de 100 días

Periodo	Fase	Actividades	Salida
Días 1-15	Inventario y selección	Inventariar sistemas y casos; definir problema, peligro, decisión, propietario y *baseline*.	Ficha de caso y clasificación preliminar.
Días 16-30	Datos y arquitectura	Mapear flujo, calidad, RGPD, ciberseguridad, integración y modo degradado.	Data *map*, riesgo y requisitos.
Días 31-50	Prototipo controlado	Entrenar/evaluar sin sustituir control; métricas por escenarios; pruebas adversas.	Informe técnico y decisión *go/no-go*.
Días 51-70	Piloto en sombra	Operar en paralelo; medir alertas, tiempos, *overrides,* concordancia y carga.	Informe de piloto y CAPA.
Días 71-85	Validación y APPCC	Aprobar finalidad, umbrales, SOP, formación, gestión de cambios y contrato.	Acta APPCC y expediente de validación.
Días 86-100	Despliegue limitado	Activar con límites, monitorización, *rollback* y revisión intensiva.	Autorización de producción y plan post-mercado interno.

19.2. Modelo económico de valor ajustado a riesgo

El retorno debe calcularse como valor neto ajustado al riesgo, no como ahorro bruto de horas. Los beneficios pueden incluir reducción de desperdicio, reclamaciones, liberación tardía, muestreo indiscriminado, mantenimiento no planificado, paradas, tiempo de investigación y coste de retirada. Deben restarse

integración, sensores, datos, validación, formación, licencias, ciberseguridad, mantenimiento, auditoría y coste de falsas alarmas.

La fórmula directiva puede expresarse así: Valor anual esperado = ahorro operativo + pérdidas evitadas ponderadas por probabilidad + valor de capacidad/rapidez - coste total de propiedad - coste residual de fallos y dependencia. Las “pérdidas evitadas” deben estimarse con prudencia y escenarios, evitando vender como ahorro cierto un incidente hipotético.

Dimensión	Indicador	Cautela
Seguridad	Reducción de desviaciones detectadas tarde; tiempo de anticipación; falsos negativos.	Principal criterio de permanencia.
Calidad/merma	Producto retenido, reproceso, destrucción, sobre-procesado.	Debe separar mejora real de traslado de coste.
Productividad	Horas de revisión, muestreo, investigación y liberación.	No reducir supervisión por debajo de lo seguro.
Continuidad	Paradas evitadas, recuperación y mantenimiento.	Incluir coste de dependencia tecnológica.
Cumplimiento	Tiempo de auditoría, integridad de registros, respuesta a autoridad.	Valor probatorio, no solo eficiencia.
Comercial	Confianza de clientes, diferenciación, acceso a cuentas.	*Claims* prudentes; no afirmar “cero riesgo”.

La decisión de escalar debe depender de criterios técnicos y operativos previamente fijados. Un piloto puede ser estadísticamente prometedor y no escalable por latencia, mantenimiento, resistencia del usuario o falta de datos. También puede ser comercialmente atractivo, pero jurídicamente inmaduro. El gate debe exigir validación, continuidad, contrato y propietario.

La empresa debe evitar *lock-in* que impida exportar datos o reproducir decisiones. El coste de salida forma parte del TCO. Para funciones críticas, puede justificarse un modelo de arquitectura híbrida con límites locales, registros propios y servicios avanzados externos.

20. Conclusiones y posición estratégica

La inteligencia artificial puede transformar el APPCC desde un sistema predominantemente reactivo —basado en comprobar que un límite no se ha superado— hacia una capacidad anticipatoria que detecta tendencias, interacciones y señales débiles. Esa transformación es valiosa, pero no automática. La IA no convierte datos en seguridad por el mero hecho de procesarlos; convierte datos en inferencias. La seguridad aparece cuando esas inferencias se integran en medidas de control validadas, procedimientos claros, decisiones competentes y una organización capaz de reconocer incertidumbre y detenerse.

El principio jurídico central permanece intacto: la responsabilidad primaria corresponde al operador de empresa alimentaria. La compra de un algoritmo no compra una exoneración. El proveedor puede responder por defectos, incumplimientos o falta de seguridad, pero la empresa decide cómo integra, configura, supervisa y utiliza la herramienta. La diligencia exigible crece a medida que el sistema se acerca a la liberación de producto, la modificación de parámetros críticos o la sustitución de una barrera humana.

El *AI Act* aporta una gramática útil —riesgo, datos, registros, transparencia, supervisión, robustez y ciberseguridad—, pero no debe eclipsar el Derecho alimentario. Muchos sistemas APPCC no serán jurídicamente de alto riesgo; algunos sí podrán serlo cuando actúen como componentes de seguridad de maquinaria o se utilicen en finalidades laborales incluidas. En todos los casos, la clasificación formal es solo una parte. La pregunta empresarial relevante es qué daño puede producir el fallo, qué barreras lo detectan y qué evidencia permite reconstruir la decisión.

La integración madura exige un doble análisis: peligros del alimento y peligros del algoritmo. Deben evaluarse datos, sensores, métricas, deriva, interfaz, personas, actualizaciones, ciberataques y registros. Debe existir un modo degradado. Deben conservarse límites duros cuando aporten independencia. Las alertas deben tener dueño, reloj y acción. La supervisión humana debe ser capaz de contradecir. El reentrenamiento

debe someterse a gestión de cambios. La prueba de cumplimiento debe diseñarse antes del incidente, no después.

Desde una perspectiva comercial, la oportunidad no está en vender "IA para cumplir", una promesa jurídicamente simplista, sino en ofrecer sistemas que mejoren prevención y evidencia sin desestructurar el autocontrol. El producto diferencial combina conocimiento del peligro, tecnología de proceso, datos, validación, ciberseguridad, documentación y servicio técnico. La empresa que pueda demostrar esa integración reducirá pérdidas y, además, aumentará la confianza de clientes, auditores, aseguradoras y autoridades.

POSICIÓN FINAL

La IA debe ser una barrera visible, gobernada y verificable; nunca una presunción opaca de control. La empresa alimentaria del futuro no será la que más algoritmos despliegue, sino la que sepa exactamente qué decide cada uno, con qué evidencia, bajo qué límites, quién puede detenerlo y cómo se protege al consumidor cuando falla.

ANEXOS OPERATIVOS

Anexo I. Árbol de decisión para clasificar el uso de IA

1. ¿La herramienta realiza inferencias —predicciones, recomendaciones, decisiones o contenido— o solo ejecuta reglas totalmente definidas por personas? Si solo ejecuta reglas, puede quedar fuera de la definición de sistema de IA, aunque siga siendo software crítico.
2. ¿La salida influye en seguridad alimentaria, destino del producto, parámetros de proceso, personas o maquinaria? Documentar todas las influencias, incluidas indirectas.
3. ¿El sistema es mera asistencia, priorización, recomendación, decisión automatizada o control físico? Asignar clase A-E de criticidad.
4. ¿Está integrado como componente de seguridad de un producto cubierto por legislación de armonización y sujeto a evaluación de terceros? Si sí, analizar artículo 6.1 AI Act.
5. ¿Se utiliza para empleo, biometría u otra finalidad del anexo III? Si sí, analizar alto riesgo por finalidad.
6. ¿Se tratan datos personales o se adoptan decisiones sobre personas? Aplicar RGPD/LO 3/2018 y garantías laborales.
7. ¿La empresa desarrolla, reentrena, marca o modifica finalidad? Determinar si es proveedor, deployer, fabricante o varios roles.
8. ¿El sistema es producto con elementos digitales comercializado o componente de maquinaria? Analizar CRA y Reglamento de Máquinas.
9. ¿La empresa entra en NIS2 por sector, tamaño o designación? Integrar gestión de riesgo y notificación.
10. Aunque no se active alto riesgo legal, fijar control interno según consecuencia de fallo, exposición y reversibilidad.

Anexo II. Ficha mínima de modelo — *Model Card* alimentaria

Campo	Contenido
Identificación	Nombre, versión, proveedor, propietario interno, fecha de aprobación.
Finalidad	Problema, peligro, etapa, usuarios, salida y acción permitida.
Usos excluidos	Productos, líneas, condiciones, decisiones o poblaciones no cubiertas.
Datos	Fuentes, periodo, plantas, productos, etiquetado, calidad y limitaciones.
Método	Tipo de modelo, variables, transformaciones, umbral y abstención.
Desempeño	Métricas globales y por escenarios críticos, intervalos y prevalencia.
Limitaciones	Falsos negativos conocidos, dominio, incertidumbre y riesgos.
Human *oversight*	Rol, información, tiempo, autoridad, over-ride y escalado.

Campo	Contenido
Integración FSMS	PRP/PPRO/PCC/verificación; SOP; acciones sobre producto.
Ciberseguridad	Arquitectura, accesos, logs, actualizaciones y continuidad.
Monitorización	Indicadores, frecuencia, drift, incidentes y revalidación.
Cambios	Historial, aprobación, rollback y versiones retiradas.
Evidencia	Protocolo e informe de validación, referencias y acta APPCC.

Anexo III. SOP resumido para gestión de alertas predictivas

1. Recepción: el sistema registra alerta, nivel, hora, producto/lote, datos y versión.
2. Reconocimiento: el responsable confirma recepción dentro del SLA; si no, escalado automático.
3. Comprobación de integridad: verificar disponibilidad, frescura, calibración y calidad de datos.
4. Clasificación: distinguir aviso preventivo, desviación confirmada, fallo del sistema o incidente.
5. Control del producto: identificar exposición; retener o segregar cuando el riesgo no pueda excluirse.
6. Acción sobre proceso: ajustar, limpiar, reparar, muestrear, reducir velocidad o parar según matriz.
7. Evidencia independiente: medición manual, laboratorio, inspección o segundo sistema cuando proceda.
8. Decisión: persona autorizada determina disposición; toda anulación se justifica.
9. Cierre: registrar causa, acción, producto, verificación y CAPA.
10. Aprendizaje: revisar falsos positivos/negativos, umbral, proceso y necesidad de revalidación.

Anexo IV. Checklist de auditoría de IA integrada en APPCC

Área	Pregunta de auditoría
Gobierno	¿Existe inventario y propietario? ¿Se ha clasificado rol y criticidad?
Finalidad	¿Está definida y limitada? ¿Se controlan usos secundarios?
APPCC	¿El equipo ha evaluado cómo el fallo afecta peligros y medidas?
Datos	¿Son representativos, íntegros, contextualizados y trazables?
Validación	¿Existe protocolo previo, referencia y criterios de aceptación?
Métricas	¿Incluyen falsos negativos, calibración y escenarios críticos?
Operación	¿Hay SOP, responsables, SLA, producto y escalado?
Supervisión	¿El humano puede entender, contradecir y registrar?
Modo degradado	¿Existe, es seguro y se ha probado?
Cambio	¿Versiones, umbrales y actualizaciones requieren aprobación?
Drift	¿Se monitoriza y hay triggers de revalidación?
Ciberseguridad	¿Se gestionan accesos, vulnerabilidades, backups y respuesta?
Proveedor	¿Contrato permite evidencia, auditoría, cambios y salida?
Registros	¿Puede reconstruirse una decisión concreta?
Cultura	¿Formación por rol y reporte de incidentes sin represalia?
Eficacia	¿El sistema reduce riesgo y no solo produce métricas?

Anexo V. Cláusulas modelo — estructura orientativa

Las siguientes fórmulas son una estructura de negociación y deben adaptarse al contrato, jurisdicción, rol regulatorio y criticidad. No constituyen cláusulas finales para firma.

Control de cambios

El proveedor no modificará la finalidad prevista, lógica operativa, umbrales, tratamiento de datos faltantes, modelo, infraestructura crítica o métricas de desempeño sin preaviso documentado y sin permitir al cliente validar la modificación en un entorno de prueba. El cliente podrá mantener la versión anterior o ejecutar rollback cuando el cambio pueda afectar seguridad, cumplimiento o continuidad.

Evidencia y logs

El proveedor conservará y pondrá a disposición los registros necesarios para reconstruir la versión, configuración, entradas relevantes, salida, eventos de seguridad e incidencias vinculadas a una decisión. La limitación por secreto empresarial no impedirá proporcionar información suficiente para auditoría, investigación y autoridad competente, bajo confidencialidad.

Notificación de incidentes

El proveedor notificará sin dilación indebida cualquier vulnerabilidad, acceso no autorizado, alteración, pérdida de integridad, degradación significativa o incidente que pueda afectar la seguridad alimentaria, datos, disponibilidad o cumplimiento, y cooperará en contención, análisis de causa, preservación de evidencia y corrección.

Datos y entrenamiento

Los datos del cliente no se utilizarán para entrenar modelos compartidos ni para finalidades distintas sin autorización expresa, base jurídica y salvaguardas. Se definirán titularidad, ubicación, subencargados, retención, devolución, portabilidad y borrado verificable.

Continuidad y salida

El proveedor garantizará exportación en formato utilizable, asistencia de transición, conservación de evidencia y eliminación segura. Para funciones críticas se establecerán medidas de continuidad, recuperación y, cuando proceda, depósito de componentes o documentación esencial.

Anexo VI. Glosario técnico-jurídico

Término	Definición
Abstención	Capacidad del modelo de no emitir una clasificación cuando la entrada está fuera de dominio o la incertidumbre es excesiva.
APPCC/HACCP	Sistema preventivo basado en análisis de peligros y puntos de control crítico.
Calibración del modelo	Correspondencia entre probabilidad estimada y frecuencia observada del evento.
Concept drift	Cambio en la relación entre variables de entrada y resultado.
Data drift	Cambio en la distribución de datos de entrada.
Deployer	Persona física o jurídica que utiliza un sistema de IA bajo su autoridad, salvo uso personal no profesional.

Término	Definición
Falso negativo	Caso peligroso o no conforme que el sistema clasifica como seguro o normal.
Human-in-the-loop	Intervención humana dentro del proceso decisorio; solo es garantía si es material.
MLOps	Prácticas para desarrollar, desplegar, monitorizar y versionar modelos de aprendizaje automático.
OPRP/PPRO	Programa de prerrequisitos operativo para controlar peligros significativos sin tratarse como PCC, en la terminología de ISO 22000 y guías.
PCC/CCP	Etapa en la que el control es esencial para prevenir, eliminar o reducir un peligro a un nivel aceptable.
Proveedor de IA	Operador que desarrolla o hace desarrollar un sistema y lo comercializa o pone en servicio bajo su nombre o marca.
Rollback	Retorno controlado a una versión anterior validada.
Score	Valor numérico de salida; no equivale por sí solo a límite crítico ni a probabilidad calibrada.
Trazabilidad probatoria	Capacidad de reconstruir hechos, versiones, decisiones, actores y evidencias de forma íntegra y comprensible.
Valor predictivo negativo	Probabilidad de que un resultado negativo sea realmente negativo, dependiente de prevalencia.
Versionado	Identificación y control de cambios en modelo, datos, reglas, software y configuración.

Bibliografía y fuentes normativas

[1] Reglamento (CE) n.º 178/2002 del Parlamento Europeo y del Consejo, de 28 de enero de 2002, por el que se establecen los principios y requisitos generales de la legislación alimentaria.

[2] Reglamento (CE) n.º 852/2004 del Parlamento Europeo y del Consejo, de 29 de abril de 2004, relativo a la higiene de los productos alimenticios, versión consolidada.

[3] Reglamento (UE) 2021/382 de la Comisión, de 3 de marzo de 2021, sobre gestión de alérgenos, redistribución y cultura de seguridad alimentaria.

[4] Comisión Europea. Comunicación 2022/C 355/01 sobre sistemas de gestión de la seguridad alimentaria, buenas prácticas de higiene y procedimientos basados en APPCC.

[5] Codex Alimentarius Commission. General Principles of Food Hygiene, CXC 1-1969, revisión 2022.

[6] Reglamento (CE) n.º 2073/2005 de la Comisión, de 15 de noviembre de 2005, relativo a los criterios microbiológicos aplicables a los productos alimenticios, versión consolidada.

[7] Reglamento (UE) 2017/625 del Parlamento Europeo y del Consejo, de 15 de marzo de 2017, sobre controles oficiales.

[8] Reglamento (CE) n.º 853/2004, normas específicas de higiene de los alimentos de origen animal.

[9] Reglamento de Ejecución (UE) n.º 931/2011, requisitos de trazabilidad para alimentos de origen animal.

[10] 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.

[11] Consejo de la Unión Europea. Artificial Intelligence: Council gives final green light to simplify and streamline rules, 29 junio 2026.

[12] Comisión Europea. AI Act — Regulatory framework and implementation timeline.

[13] Reglamento (UE) 2016/679 — Reglamento General de Protección de Datos.

[14] Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

[15] Directiva (UE) 2022/2555 — NIS2, relativa a medidas para un elevado nivel común de ciberseguridad.

[16] Real Decreto-ley 12/2018, de 7 de septiembre, de seguridad de las redes y sistemas de información.

[17] Real Decreto 43/2021, de 26 de enero, de desarrollo del Real Decreto-ley 12/2018.

[18] Reglamento (UE) 2024/2847 — Cyber Resilience Act.

[19] Reglamento (UE) 2023/1230 del Parlamento Europeo y del Consejo, relativo a las máquinas.

[20] Directiva (UE) 2024/2853 del Parlamento Europeo y del Consejo, sobre responsabilidad por los daños causados por productos defectuosos.

[21] Reglamento (UE) 2023/2854 — Data Act.

[22] Reglamento (UE) 2022/868 — Data Governance Act.

[23] Directiva (UE) 2016/943, protección de secretos empresariales.

[24] Ley 17/2011, de 5 de julio, de seguridad alimentaria y nutrición.

[25] ISO 22000:2018. Food safety management systems — Requirements for any organization in the food chain.

[26] ISO/IEC 42001:2023. Information technology — Artificial intelligence — Management system.

[27] ISO/IEC 23894:2023. Information technology — Artificial intelligence — Guidance on risk management.

[28] ISO/IEC 22989:2022. Artificial intelligence concepts and terminology.

[29] ISO/IEC 25059:2023. Quality model for AI systems.

[30] ISO/IEC 27001:2022. Information security management systems — Requirements.

[31] NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023.

[32] NIST. AI RMF Core: Govern, Map, Measure and Manage.

[33] FDA. Relationship Between Food Safety Culture and Food Safety Management Systems, 2024.

[34] FDA. Human Foods Program 2026 Priority Deliverables — AI predictive models and risk management.

[35] FDA. FSMA Final Rule on Requirements for Additional Traceability Records for Certain Foods.

[36] Chen, Y. et al. Review of visual analytics methods for food safety risks. npj Science of Food, 2023.

[37] Taiwo, O.R. et al. Advancements in Predictive Microbiology: Integrating New Technologies for Efficient Food Safety Models. International Journal of Microbiology, 2024.

[38] Hassoun, A. et al. Enhancing Food Integrity through Artificial Intelligence and Machine Learning: A Comprehensive Review. Applied Sciences, 2024.

[39] Zoellner, C. et al. Design Elements of Listeria Environmental Monitoring Programs in Food Processing Facilities: A Scoping Review. Comprehensive Reviews in Food Science and Food Safety, 2018.

[40] Ivanek, R. et al. EnABLe: An agent-based model to understand Listeria dynamics in food processing facilities. Scientific Reports, 2019.

[41] GFSI. Benchmarking Requirements and recognised certification programmes.

[42] BRCGS. Global Standard Food Safety, Issue 9.

[43] IFS. IFS Food Standard, version 8.

[44] FSSC 22000. Scheme version 6.

Declaración de cierre editorial

El presente documento se ha preparado como obra de análisis técnico-jurídico y propuesta metodológica. Las referencias a estándares privados deben verificarse frente a la edición licenciada y los requisitos contractuales del centro. Las referencias normativas deben comprobarse en la versión consolidada vigente y, en particular, debe revisarse la publicación final de la reforma *Digital Omnibus on AI* y su efecto sobre fechas de aplicación antes de cualquier uso jurídico formal.