

# Servicios de inteligencia artificial sobre espacios de datos como oportunidad para la reducción de la huella de carbono

# INDICE

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| Resumen ejecutivo .....                                                          | 7  |
| 1. Objeto y alcance del informe .....                                            | 9  |
| 1.1. Contexto y motivación .....                                                 | 9  |
| 1.2. Objetivos del informe .....                                                 | 9  |
| 1.3. Metodología .....                                                           | 10 |
| 1.4. Entidades participantes .....                                               | 10 |
| 1.5. Marco temporal y alcance geográfico .....                                   | 11 |
| 1.6. Estructura del informe .....                                                | 11 |
| 2. Problemática de la huella de carbono .....                                    | 13 |
| 2.1. Alcances 1, 2 y 3 del GHG Protocol .....                                    | 13 |
| 2.2. Criticidad del alcance 3 .....                                              | 13 |
| 2.3. Barreras específicas para el cálculo del alcance 3 .....                    | 14 |
| 2.4. El marco regulatorio europeo .....                                          | 15 |
| 2.5. Implicaciones sectoriales .....                                             | 16 |
| 3. Estado del arte: espacios de datos en Europa .....                            | 18 |
| 3.1. La Estrategia Europea del Dato y el marco regulatorio .....                 | 18 |
| 3.2. Gaia-X: Trust Framework 3.0 .....                                           | 18 |
| 3.3. Eclipse Dataspace Components .....                                          | 19 |
| 3.4. Espacio Común Europeo de Datos de Movilidad (EMDS) .....                    | 19 |
| 3.5. Green Deal Data Space .....                                                 | 20 |
| 3.6. Espacios de datos sectoriales relevantes para la descarbonización .....     | 20 |
| 3.7. Marco de interoperabilidad .....                                            | 22 |
| 3.8. Arquitectura técnica del protocolo DSP .....                                | 23 |
| 3.9. Modelos de gobernanza y sostenibilidad de los espacios de datos .....       | 23 |
| 4. Espacios de datos en España: iniciativas y convocatorias del Ministerio ..... | 25 |
| 4.1. El Plan de Impulso de los Espacios de Datos Sectoriales .....               | 25 |
| 4.2. Convocatorias específicas y presupuestos .....                              | 25 |
| 4.3. El programa Kit Espacios de Datos .....                                     | 26 |
| 4.4. Espacios de datos operativos en España .....                                | 26 |

|                                                                                                            |    |
|------------------------------------------------------------------------------------------------------------|----|
| 4.5. La oportunidad para servicios de IA en el contexto español .....                                      | 27 |
| 4.6. La Especificación UNE 0087:2025 y la Lista de Confianza.....                                          | 27 |
| 5. Inteligencia artificial para la reducción de huella de carbono: estado del arte .....                   | 29 |
| 5.1. Machine Learning para la predicción de emisiones de Alcance 3 .....                                   | 29 |
| 5.2. IA para la descarbonización de cadenas de suministro .....                                            | 30 |
| 5.3. IA aplicada al Alcance 3 .....                                                                        | 30 |
| 5.4. LLMs para reportes de sostenibilidad .....                                                            | 31 |
| 5.5. Reinforcement Learning para la optimización de emisiones en tiempo real .....                         | 31 |
| 5.6. Computer vision y análisis de imágenes satélite para verificación de emisiones....                    | 32 |
| 6. Estado del arte internacional: IA y espacios de datos para sostenibilidad .....                         | 34 |
| 6.1. Alemania: Catena-X y EuProGigant .....                                                                | 34 |
| 6.2. Francia: EONA-X .....                                                                                 | 34 |
| 6.3. Austria: Hub Gaia-X .....                                                                             | 35 |
| 6.5. Países Nórdicos .....                                                                                 | 35 |
| 6.6. Iniciativas globales de estandarización: PACT, CDP y GRI en el contexto de<br>espacios de datos ..... | 36 |
| 7. Oportunidades: servicios de IA sobre espacios de datos para reducción de huella de<br>carbono.....      | 38 |
| 7.1. La tesis: por qué los espacios de datos son el sustrato natural para la IA climática                  | 38 |
| 7.2. Oportunidades identificadas.....                                                                      | 38 |
| 7.3. Análisis de viabilidad técnica y económica.....                                                       | 41 |
| 8. Caso de uso real: plataforma ECODES sobre el espacio de datos de movilidad EDMUS<br>.....               | 42 |
| 8.1. Descripción general del caso de uso .....                                                             | 42 |
| 8.2. Arquitectura del espacio de datos .....                                                               | 42 |
| 8.3. Metodología de cálculo de huella de carbono .....                                                     | 43 |
| 8.4. El protocolo EDC en acción: negociación automatizada .....                                            | 43 |
| 8.5. Integración con el mercado de carbono EU ETS .....                                                    | 44 |
| 8.6. Arquitectura de microservicios .....                                                                  | 44 |
| 8.7. Relevancia para el Alcance 3 .....                                                                    | 45 |
| 8.8. Lecciones aprendidas del despliegue .....                                                             | 45 |
| 8.9. Métricas de rendimiento y KPI operativos .....                                                        | 46 |
| 8.10. Potencial de escalado y replicabilidad .....                                                         | 46 |
| 9. Catálogo de oportunidades sobre espacios de datos españoles.....                                        | 50 |
| 9.1. De las tipologías genéricas al catálogo aplicado.....                                                 | 50 |
| 9.2. Tabla de espacios de datos españoles utilizados como sustrato .....                                   | 51 |
| 9.3. Mapeo entre tipologías genéricas (cap. 7) y oportunidades aplicadas (cap. 9) .....                    | 53 |

|                                                                                  |    |
|----------------------------------------------------------------------------------|----|
| 9.4. Fichas de oportunidades aplicadas .....                                     | 54 |
| 9.5. Síntesis y priorización .....                                               | 66 |
| 10. Conclusiones y recomendaciones .....                                         | 68 |
| 10.1. Conclusiones principales.....                                              | 68 |
| 10.2. Recomendaciones .....                                                      | 69 |
| 10.3. Perspectivas futuras.....                                                  | 69 |
| Referencias bibliográficas.....                                                  | 71 |
| Anexo A. Glosario de términos .....                                              | 73 |
| A.1. Términos de espacios de datos .....                                         | 73 |
| A.2. Términos de huella de carbono .....                                         | 74 |
| A.3. Términos de inteligencia artificial .....                                   | 75 |
| Anexo B. Arquitectura técnica detallada del caso ECODES-EDMUS.....               | 76 |
| B.1. Diagrama de componentes.....                                                | 76 |
| B.2. Flujo de datos completo.....                                                | 76 |
| B.3. Stack tecnológico completo .....                                            | 76 |
| B.4. Endpoints principales de la API ECODES .....                                | 77 |
| B.5. Protocolo de seguridad y gestión de identidades .....                       | 78 |
| B.6. Modelo de datos de los certificados CAES .....                              | 78 |
| Anexo C. Convocatorias de financiación vigentes .....                            | 80 |
| C.1. Convocatorias nacionales.....                                               | 80 |
| C.2. Convocatorias europeas relevantes .....                                     | 80 |
| C.4. Convocatorias regionales relevantes.....                                    | 80 |
| C.3. Incentivos fiscales .....                                                   | 81 |
| Anexo D. Publicaciones y papers de referencia adicionales .....                  | 82 |
| D.1. Machine Learning y predicción de emisiones .....                            | 82 |
| D.2. Espacios de datos y sostenibilidad.....                                     | 82 |
| D.3. Huella de carbono y regulación .....                                        | 83 |
| D.5. Verificación y aseguramiento de datos de emisiones .....                    | 83 |
| D.6. Computación verde y centros de datos sostenibles .....                      | 84 |
| D.4. IA y sostenibilidad: perspectivas adicionales .....                         | 84 |
| Anexo E. Mapeo de oportunidades con fuentes de financiación .....                | 85 |
| Anexo F. Casos de uso internacionales de referencia .....                        | 87 |
| F.1. Catena-X: Product Carbon Footprint en la industria automotriz alemana ..... | 87 |
| F.2. deployEMDS Milán: escenarios de neutralidad de carbono urbana .....         | 87 |
| F.3. CO2 AI (BCG): automatización del reporte de emisiones .....                 | 88 |
| F.4. Energinet DataHub (Dinamarca): datos energéticos en tiempo real .....       | 88 |

|                                                                           |    |
|---------------------------------------------------------------------------|----|
| F.5. PCAF y emisiones financiadas: el caso del sector bancario .....      | 89 |
| Anexo G. Metodología detallada de cálculo de emisiones de Alcance 3 ..... | 90 |
| G.1. Métodos de cálculo según el GHG Protocol.....                        | 90 |
| G.2. Fuentes de factores de emisión .....                                 | 91 |
| G.3. Cálculo detallado por categorías materiales.....                     | 91 |
| G.4. Cuantificación de la incertidumbre .....                             | 92 |

# AUTORES

Este informe es el resultado de la colaboración entre el Instituto Tecnológico de Aragón (ITA) y ECODES, dos entidades aragonesas que han trabajado con un objetivo compartido: poner la tecnología al servicio de la descarbonización. De esta colaboración se ha materializado el desarrollo del caso real ECODES-EDMUS, una solución técnica para la emisión y venta de certificados CAES a partir de datos del espacio de datos de movilidad urbana sostenible. El equipo firmante combina perfiles de investigación aplicada y experiencia sectorial, y es responsable conjunto de la concepción, redacción y validación de los análisis, oportunidades y recomendaciones que se presentan a continuación.



**PABLO  
BARRENECHEA  
ABECIA**

ECODES



**CIRO  
ARTIGOT  
CORDERO**

ECODES



**ADÁN PIÑERO  
ALQUEGUI**

ITA



**FRANCISCO  
ROJAS LUNA**

ITA

## SOBRE ESTE DOCUMENTO

Editado por: ECODES

Diseño Gráfico: BAMBAM Studio

Fecha: Junio de 2026

# Resumen ejecutivo

El presente informe analiza en profundidad la convergencia entre dos de las transformaciones tecnológicas más significativas de la década actual: la inteligencia artificial generativa y agente, y los espacios de datos soberanos. nacionales y europeos, evaluando su potencial combinado como instrumento para la reducción efectiva de la huella de carbono empresarial, con especial atención al Alcance 3 del GHG Protocol. Esta investigación, desarrollada conjuntamente por el Instituto Tecnológico de Aragón (ITA) y la Fundación Ecología y Desarrollo (ECODES), se enmarca en el periodo octubre 2025 a junio 2026 y responde a la necesidad crítica de identificar oportunidades concretas donde la IA, desplegada como servicio sobre infraestructuras de datos federados, pueda contribuir al cálculo, la mitigación y la reducción directa de emisiones de gases de efecto invernadero a lo largo de cadenas de valor complejas.

La problemática del Alcance 3 constituye el núcleo de la investigación. Las emisiones indirectas de la cadena de valor representan, según múltiples estudios académicos, entre el 70% y el 90% de la huella total de carbono de la mayoría de las organizaciones (Serafeim y Vélez Caicedo, 2022; Hettler, 2024). Su cuantificación requiere acceso a datos distribuidos entre decenas o miles de actores económicos, una circunstancia que convierte a los espacios de datos soberanos —con su capacidad de facilitar el intercambio controlado de información sin cesión de soberanía— en la infraestructura natural para abordar este desafío. Cuando, sobre esa infraestructura, se superponen servicios de IA capaces de automatizar la clasificación de factores de emisión, predecir emisiones a partir de datos financieros, optimizar rutas logísticas o certificar reducciones verificables, emerge un paradigma con potencial transformador.

El informe identifica un ecosistema regulatorio y de financiación sin precedentes en España, con el Plan de Impulso de los Espacios de Datos Sectoriales (500 millones de euros hasta 2026), el programa Kit Espacios de Datos (60 millones de euros gestionados por Red.es), y dos convocatorias de demostradores y casos de uso ya resueltas con 225 millones comprometidos. A nivel europeo, la iniciativa Gaia-X ha alcanzado su fase 2.0 con el Trust Framework 3.0 Danube, y proyectos como deployEMDS y EONA-X demuestran la viabilidad del espacio común europeo de datos de movilidad con tecnología Eclipse Dataspace Connector.

Como caso de uso real, se presenta ECODES-EDMUS, una plataforma operativa desarrollada por ITA con soporte de ECODES sobre el espacio de datos de movilidad EDMUS. Esta plataforma certifica y monetiza reducciones de CO<sub>2</sub> derivadas de la electrificación de flota de autobuses mediante un mercado digital de Certificados de Ahorro de Emisiones, con precio referenciado al EU ETS. El sistema demuestra cómo los datos de billeteaje, accedidos soberanamente a través del protocolo EDC sin abandonar la infraestructura de los operadores de movilidad, pueden transformarse en activos ambientales auditables y comercializables.

El informe concluye con la identificación de doce oportunidades concretas para servicios de IA sobre espacios de datos multisectoriales, desde la predicción de emisiones Alcance 3 mediante modelos de Adaptive Boosting hasta la generación automatizada de product

carbon footprints con LLMs, pasando por la optimización multimodal del transporte y la certificación inteligente de reducciones. Cada oportunidad se analiza en términos de madurez tecnológica, requisitos de datos, impacto potencial y alineamiento con las convocatorias de financiación vigentes.

Desde la perspectiva metodológica, el informe documenta cómo los tres métodos jerárquicos del GHG Protocol para el cálculo de Alcance 3 (basado en gasto, basado en actividad media, y específico del proveedor) pueden automatizarse progresivamente mediante servicios de IA a medida que los proveedores se incorporan al espacio de datos. La progresión natural va desde la predicción ML a partir de datos financieros (inmediatamente aplicable con el método basado en gasto) hasta la certificación basada en datos específicos del proveedor accedidos soberanamente a través del protocolo EDC (el estado más avanzado, demostrado en ECODES-EDMUS).

Los hallazgos del informe son especialmente relevantes para el ecosistema español de innovación, donde la confluencia de financiación pública, normativa habilitadora (UNE 0087:2025, CSRD) y capacidades tecnológicas en centros como ITA crea un contexto favorable para que España se posicione como referente europeo en la aplicación de IA sobre espacios de datos para la sostenibilidad.

Esta investigación se ha desarrollado con el apoyo del Ministerio para la Transición Ecológica y el Reto Demográfico a través de la convocatoria para el año 2025 de subvenciones en régimen de concurrencia competitiva “línea a” para el desarrollo de actividades para el interés general consideradas de interés sociales, en el ámbito de la investigación científica y técnica y protección al medio ambiente en materias de competencia estatal.

# 1. Objeto y alcance del informe

## 1.1. Contexto y motivación

La Unión Europea ha establecido como objetivo estratégico alcanzar la neutralidad climática en 2050, un compromiso que requiere una transformación radical de los modelos productivos, logísticos y energéticos. En este contexto, la Directiva de Informes de Sostenibilidad Corporativa (CSRD), que entró en vigor progresivamente desde enero de 2024, obliga a un número creciente de empresas europeas a reportar sus emisiones de gases de efecto invernadero, incluyendo las de Alcance 3, que abarcan toda la cadena de valor. Simultáneamente, la Estrategia Europea del Dato y sus instrumentos legislativos (Data Governance Act, Data Act, AI Act) han creado el marco jurídico para el despliegue de espacios de datos sectoriales que permitan el intercambio seguro y soberano de información entre organizaciones.

---

La convergencia de estas dos grandes políticas —descarbonización y economía del dato— abre un espacio de oportunidad donde los servicios de inteligencia artificial pueden actuar como catalizadores.

---

La IA puede automatizar procesos de cálculo de huella de carbono que hoy requieren semanas de trabajo manual, optimizar operaciones para reducir emisiones en origen, y certificar reducciones de forma transparente y auditable. Sin embargo, la eficacia de estos servicios depende críticamente del acceso a datos distribuidos a lo largo de cadenas de valor complejas, un acceso que los espacios de datos soberanos están diseñados para facilitar.

## 1.2. Objetivos del informe

El presente informe persigue los siguientes objetivos: Primero, analizar el estado del arte de los espacios de datos en España y Europa, con especial atención a las convocatorias y proyectos del Ministerio para la Transformación Digital y de la Función Pública. Segundo, evaluar el estado actual de los servicios de IA aplicados a la reducción de huella de carbono, particularmente en el contexto de espacios de datos federados. Tercero, identificar las oportunidades específicas que la filosofía de los espacios de datos abre para la aplicación de servicios de IA orientados a la descarbonización, con énfasis en el Alcance 3 del GHG Protocol. Cuarto, documentar un caso de uso real implementado sobre el espacio de datos de movilidad EDMUS, en colaboración entre ITA y ECODES.

## 1.3. Metodología

La investigación combina revisión sistemática de literatura académica y gris, análisis de marcos regulatorios y programas de financiación, y desarrollo experimental de un caso de uso operativo. La revisión bibliográfica se ha realizado sobre las bases Scopus, Web of Science y Google Scholar, con las cadenas de búsqueda: «data spaces AND carbon footprint», «AI AND scope 3 emissions», «machine learning AND supply chain decarbonization», «Eclipse Dataspace Connector AND sustainability», y sus equivalentes en español, para publicaciones del periodo 2020–2026. Se han analizado 87 publicaciones, de las cuales 42 son artículos en revistas indexadas, 18 son working papers o preprints, 15 son informes institucionales y 12 son documentos de política pública.

El análisis de marcos regulatorios ha comprendido la revisión sistemática de la legislación europea (CSRD, DGA, Data Act, AI Act), la normativa española (Orden TDF/758/2025, UNE 0087:2025), los programas de financiación (Plan de Impulso, Kit Espacios de Datos, Horizon Europe, Digital Europe), y los estándares técnicos (GHG Protocol, ESRS, Pathfinder Framework, EDC Protocol). Para cada instrumento, se ha evaluado su relevancia para el despliegue de servicios de IA sobre espacios de datos orientados a la sostenibilidad, identificándose las oportunidades de alineamiento y las barreras potenciales. El desarrollo experimental se ha centrado en la construcción, despliegue y evaluación de la plataforma ECODES-EDMUS como prueba de concepto de un servicio de certificación de emisiones operando sobre un espacio de datos soberano con Eclipse Dataspace Connector.

## 1.4. Entidades participantes

El Instituto Tecnológico de Aragón (ITA) es un centro tecnológico público adscrito al Departamento de Economía, Empleo e Industria del Gobierno de Aragón, con más de 350 profesionales y una trayectoria de cuatro décadas en investigación aplicada. Su Departamento de Sistemas Cognitivos y Big Data lidera líneas de trabajo en inteligencia artificial, análisis de datos masivos, procesamiento del lenguaje natural y sistemas multiagente, con aplicación en sectores como logística, movilidad, agroindustria y salud. ITA es promotor del Aragón European Digital Innovation Hub (EDIH), seleccionado en 2022 como uno de los 12 centros de innovación digital españoles de referencia europea, especializado en Big Data, IA y computación de altas prestaciones (HPC).

La Fundación Ecología y Desarrollo (ECODES), con sede en Zaragoza desde 1992, es una organización sin ánimo de lucro que trabaja en la construcción de una economía verde, inclusiva y responsable. A través de su plataforma CeroCO2 y su área de Acción Climática, ECODES ofrece servicios integrales de medición, gestión y reducción de huella de carbono a empresas y administraciones públicas. Su proyecto Carbonpedia, desarrollado con la colaboración de la Fundación Biodiversidad, constituye una base de datos abierta de huella de carbono de entidades, eventos, productos y servicios. ECODES ha liderado iniciativas como MAPIC+s para la descarbonización del sector sanitario y es

miembro de la Alianza para la Acción Transformadora sobre Cambio Climático y Salud (ATACH). La herramienta Scope es una solución desarrollada por ECODES para el cálculo de la huella de carbono empresarial, que permite a organizaciones de cualquier tamaño medir sus emisiones de gases de efecto invernadero conforme a los estándares del GHG Protocol.

## 1.5. Marco temporal y alcance geográfico

El periodo de investigación abarca desde octubre de 2025 hasta junio de 2026, lo que permite capturar un momento especialmente dinámico tanto en el despliegue de espacios de datos como en la maduración de las tecnologías de IA aplicadas a la sostenibilidad. Durante este periodo, se han producido hitos significativos: la publicación del Trust Framework 3.0 de Gaia-X (noviembre 2025), la activación del Espacio Nacional de Datos de Salud en España (enero 2026), la entrada en vigor plena del Data Act europeo (septiembre 2025), la publicación del dataset ExioML en Nature Scientific Data (febrero 2026), y la continuidad de las convocatorias del Kit Espacios de Datos con presupuesto disponible. El alcance geográfico es primariamente europeo, con énfasis en España, Alemania, Francia, Italia, Austria y los países nórdicos, reflejo de la concentración de iniciativas de espacios de datos en estos territorios.

La decisión de centrar el análisis en el contexto europeo responde a tres factores. Primero, la Unión Europea es la única jurisdicción que ha desarrollado simultáneamente un marco regulatorio para espacios de datos (DGA, Data Act), para reporte de sostenibilidad (CSRD, ESRS) y para inteligencia artificial (AI Act), creando un ecosistema regulatorio integrado sin equivalente en otras regiones. Segundo, las inversiones públicas europeas en infraestructura de espacios de datos superan ampliamente las de cualquier otra región, con más de 1.500 millones de euros comprometidos entre los programas nacionales y los fondos europeos. Tercero, los principios de soberanía del dato que fundamentan los espacios de datos europeos son especialmente relevantes para los datos ambientales, que pueden contener información competitiva sensible y que requieren garantías de integridad y auditabilidad para ser aceptados en procesos de reporte regulatorio.

## 1.6. Estructura del informe

El informe se estructura en diez capítulos, precedidos por este resumen ejecutivo y seguidos de las referencias bibliográficas y siete anexos (A–G). El capítulo 1 presenta el objeto, alcance, metodología, entidades participantes y marco temporal y geográfico del trabajo. El capítulo 2 analiza la problemática del Alcance 3, incluyendo el marco conceptual del GHG Protocol, las barreras específicas para su reporte, el marco regulatorio CSRD/ESRS y las implicaciones sectoriales. El capítulo 3 revisa el estado del arte de los espacios de datos en Europa, desde la Estrategia Europea del Dato y Gaia-X hasta los componentes técnicos EDC, el protocolo DSP y los espacios sectoriales relevantes para la descarbonización. El capítulo 4 se centra en las iniciativas y

convocatorias españolas, la Especificación UNE 0087:2025 y el ecosistema aragonés liderado por ITA. El capítulo 5 analiza el estado del arte de la IA aplicada a la reducción de huella de carbono (Machine Learning, LLMs, Reinforcement Learning, Computer Vision) y el capítulo 6 lo complementa con un panorama internacional comparado. El capítulo 7 identifica oportunidades genéricas de servicios de IA sobre espacios de datos. El capítulo 8 documenta el caso de uso real de la plataforma ECODES sobre el Espacio de Datos de Movilidad EDMUS. El capítulo 9 traduce las tipologías genéricas del capítulo 7 en un catálogo de quince oportunidades aplicadas a espacios de datos españoles concretos. El capítulo 10 cierra el informe con conclusiones, recomendaciones y perspectivas futuras.

# 2. Problemática de la huella de carbono

## 2.1. Alcances 1, 2 y 3 del GHG Protocol

El Greenhouse Gas Protocol (GHG Protocol), desarrollado por el World Resources Institute y el World Business Council for Sustainable Development, establece la taxonomía estándar para la clasificación de emisiones corporativas en tres alcances. El Alcance 1 comprende las emisiones directas procedentes de fuentes controladas por la organización, como la combustión en calderas, vehículos propios o procesos industriales. El Alcance 2 abarca las emisiones indirectas asociadas a la generación de electricidad, vapor, calefacción y refrigeración adquiridos y consumidos por la organización. El Alcance 3, por su parte, engloba todas las demás emisiones indirectas que ocurren a lo largo de la cadena de valor, tanto aguas arriba (proveedores, transporte de materias primas, desplazamientos de empleados) como aguas abajo (uso del producto vendido, tratamiento al final de su vida útil, transporte de distribución).

El GHG Protocol Corporate Value Chain Standard subdivide el Alcance 3 en 15 categorías: bienes y servicios adquiridos (cat. 1), bienes de capital (cat. 2), actividades relacionadas con combustible y energía no incluidas en Alcance 1 o 2 (cat. 3), transporte y distribución aguas arriba (cat. 4), residuos generados en operaciones (cat. 5), viajes de negocios (cat. 6), desplazamiento de empleados (cat. 7), activos arrendados aguas arriba (cat. 8), transporte y distribución aguas abajo (cat. 9), procesamiento de productos vendidos (cat. 10), uso de productos vendidos (cat. 11), tratamiento al final de la vida útil (cat. 12), activos arrendados aguas abajo (cat. 13), franquicias (cat. 14) e inversiones (cat. 15).

## 2.2. Criticidad del Alcance 3

Las emisiones de Alcance 3 representan la mayoría de la huella de carbono de las organizaciones en prácticamente todos los sectores. Según Hettler (2024), en su revisión sistemática publicada en *Business Strategy and the Environment*, las emisiones de Alcance 3 constituyen típicamente el 75% o más del perfil total de emisiones de una empresa, siendo especialmente dominantes en sectores como la automoción (donde el uso del producto vendido genera la inmensa mayoría de emisiones), la alimentación (con emisiones agrícolas aguas arriba) y las finanzas (donde las emisiones financiadas pueden superar el 99% del total). Wang y Ye (2025), en su artículo publicado en *Frontiers in Sustainability*, confirman que entre 1995 y 2015 las emisiones globales de Alcance 3 crecieron un 84%, frente al 47% del Alcance 1 y el 78% del Alcance 2.

Esta preponderancia crea una paradoja operativa: las emisiones más significativas de una organización son precisamente aquellas sobre las que tiene menos control directo y

menor capacidad de medición. Como señalan Serafeim y Vélez Caicedo (2022) en su working paper de Harvard Business School, los datos necesarios para cuantificar el Alcance 3 provienen frecuentemente de terceros que pueden no recopilar dicha información, lo que obliga a recurrir a factores de emisión genéricos que introducen incertidumbre significativa. El estudio publicado en PLOS Climate por Qu et al. (2023) demuestra una considerable divergencia en los valores agregados de emisiones de Alcance 3 entre los tres principales proveedores de datos (Bloomberg, Refinitiv Eikon e ISS), lo que socava la comparabilidad y fiabilidad de la información disponible.

## 2.3. Barreras específicas para el cálculo del alcance 3

---

La literatura identifica múltiples barreras que dificultan el reporte completo y preciso del Alcance 3. En primer lugar, la fragmentación de datos: la información necesaria está distribuida entre centenares o miles de proveedores, cada uno con sus propios sistemas de recogida de datos y niveles de madurez en reporte ambiental. En segundo lugar, la ausencia de estandarización: no existe un formato universal para el intercambio de datos de emisiones entre organizaciones, lo que genera incompatibilidades técnicas y semánticas. En tercer lugar, la falta de incentivos: los proveedores, especialmente las pymes, carecen frecuentemente de motivación y recursos para medir y compartir sus emisiones. En cuarto lugar, la complejidad metodológica: la elección entre métodos basados en gasto, basados en actividad o específicos del proveedor implica compromisos significativos entre precisión y viabilidad.

---

Estas barreras son particularmente relevantes porque, como destaca el informe del MIT Sloan (2024), la precisión de los cálculos de emisiones actuales es cuestionable y los métodos tradicionales son inflexibles y propensos a errores. La Science Based Targets initiative (SBTi) exige que las empresas incluyan el Alcance 3 en sus objetivos de reducción cuando estas emisiones representen más del 40% del total, lo que convierte la medición precisa en un requisito previo para establecer compromisos creíbles de descarbonización.

---

Una lectura atenta de las barreras descritas revela que el Alcance 3 es, en su esencia, un problema de intercambio de datos entre organizaciones. La empresa que quiere calcular sus emisiones de Alcance 3 necesita que sus proveedores compartan datos de actividad (consumos energéticos, distancias de transporte, composición de materiales) o, idealmente, sus propias emisiones calculadas (datos específicos del proveedor). Esta necesidad de intercambio de datos a lo largo de cadenas de valor es precisamente el problema que los espacios

de datos soberanos están diseñados para resolver: facilitar la compartición controlada de información entre organizaciones independientes, preservando la soberanía de cada participante sobre sus datos y estableciendo condiciones contractuales verificables para cada transacción.

---

Esta convergencia conceptual entre la problemática del Alcance 3 y la arquitectura de los espacios de datos constituye la hipótesis central del presente informe: los espacios de datos soberanos son la infraestructura natural para abordar el desafío del Alcance 3, y los servicios de IA desplegados sobre esta infraestructura pueden transformar radicalmente la precisión, coste y escalabilidad de la gestión de emisiones indirectas.

## 2.4. El marco regulatorio europeo

La Directiva de Informes de Sostenibilidad Corporativa (Corporate Sustainability Reporting Directive, CSRD), adoptada por el Parlamento Europeo y el Consejo en diciembre de 2022, constituye el principal motor regulatorio de la demanda de datos de Alcance 3 en Europa. La CSRD sustituye y amplía la anterior Directiva de Información No Financiera (NFRD), extendiendo progresivamente las obligaciones de reporte a aproximadamente 50.000 empresas en la UE: desde enero de 2024 para las empresas ya sujetas a la NFRD, desde enero de 2025 para las grandes empresas que cumplan al menos dos de tres criterios (más de 250 empleados, más de 40 millones de facturación, más de 20 millones en activos), y desde enero de 2026 para las pymes cotizadas.

Los European Sustainability Reporting Standards (ESRS), desarrollados por EFRAG bajo mandato de la Comisión Europea, definen los requisitos específicos de divulgación. El estándar ESRS E1 (Cambio climático) exige la divulgación de emisiones de GEI de Alcance 1, 2 y 3, desglosadas por las 15 categorías del GHG Protocol cuando sean materiales. Crucialmente, el ESRS E1 requiere que las empresas divulguen no solo las cifras absolutas de emisiones, sino también las metodologías utilizadas, las fuentes de datos (diferenciando entre datos específicos del proveedor, datos medios del sector y datos financieros estimados), y las incertidumbres asociadas. Esta exigencia de transparencia metodológica convierte la calidad de los datos en un factor crítico y crea un incentivo directo para adoptar espacios de datos que faciliten el acceso a datos primarios de los proveedores.

La implementación de la CSRD en España se ha materializado a través de la transposición al ordenamiento jurídico español, que extiende las obligaciones a un universo estimado de más de 4.000 empresas españolas de gran tamaño y pymes cotizadas. El Instituto de Contabilidad y Auditoría de Cuentas (ICAC) ha publicado guías de aplicación que enfatizan la necesidad de sistemas de información robustos para la recopilación de datos de sostenibilidad, especialmente para las emisiones de Alcance 3 que, por definición, requieren información de terceros. Esta necesidad regulatoria crea una oportunidad de mercado para servicios de IA que automaticen la recopilación,

clasificación y cálculo de emisiones indirectas a partir de datos compartidos en espacios de datos sectoriales.

## 2.5. Implicaciones sectoriales

La materialidad del Alcance 3 varía significativamente entre sectores, lo que determina las prioridades de inversión en servicios de IA y espacios de datos. En el sector del transporte, que representa aproximadamente el 27% de las emisiones totales de la UE, las emisiones de Alcance 3 se concentran en las categorías 1 (bienes y servicios adquiridos, especialmente combustibles y vehículos), 4 y 9 (transporte y distribución aguas arriba y abajo) y 11 (uso de productos vendidos, en el caso de fabricantes de vehículos). Para las empresas de transporte público, la electrificación de la flota transforma radicalmente el perfil de emisiones, trasladando las emisiones del Alcance 1 (combustión directa de diésel) al Alcance 2 (consumo eléctrico) y, potencialmente, reduciéndolas significativamente si la electricidad proviene de fuentes renovables.

En el sector agroalimentario, las emisiones de Alcance 3 pueden representar hasta el 90% del total, dominadas por la categoría 1 (emisiones agrícolas de cultivos y ganadería, uso de fertilizantes) y la categoría 4 (transporte de materias primas). La fragmentación extrema de este sector, con miles de pequeños productores agrícolas, hace que la recopilación de datos específicos sea prácticamente imposible sin una infraestructura de datos compartida. El espacio de datos AgriDataSpace, financiado por Horizon Europe, aborda precisamente este desafío proporcionando una plataforma para que los actores de la cadena agroalimentaria compartan datos de producción, transporte y procesamiento de forma soberana.

En la industria manufacturera, las emisiones de Alcance 3 se distribuyen entre múltiples categorías según el tipo de producto. Para fabricantes de automóviles, la categoría 11 (uso de productos vendidos) puede representar más del 80% del total de emisiones. Para fabricantes de componentes electrónicos, la categoría 1 (materias primas, especialmente minerales raros y semiconductores) domina el perfil. El modelo Catena-X de intercambio de Product Carbon Footprint (PCF) a lo largo de cadenas de suministro con miles de participantes demuestra la viabilidad técnica de este enfoque, y su extensión a otros sectores industriales constituye una de las oportunidades más claras para los espacios de datos.

El sector financiero presenta un caso particularmente interesante: las emisiones de Alcance 3, categoría 15 (inversiones), pueden representar más del 99% del total de emisiones de un banco o fondo de inversión. La Partnership for Carbon Accounting Financials (PCAF) ha desarrollado un estándar global para la medición de emisiones financiadas, que requiere acceso a datos de emisiones de las empresas en las que se invierte. Un espacio de datos financiero que integre datos de emisiones corporativas con datos de carteras de inversión permitiría a los agentes de IA calcular automáticamente las emisiones financiadas de un portfolio y recomendar estrategias de desinversión o engagement para reducir la huella climática.

El sector sanitario, responsable del 4,4% de las emisiones globales de GEI según el Health Care Without Harm, presenta un perfil de Alcance 3 dominado por la categoría 1 (bienes y servicios adquiridos, especialmente dispositivos médicos, fármacos y material fungible) que puede representar hasta el 70% de las emisiones totales de un hospital. El proyecto MAPIC+s, liderado por ECODES en colaboración con GSK, ha demostrado la viabilidad de aplicar metodologías de cálculo de huella de carbono en el contexto clínico español. Un espacio de datos sanitario como el activado por el Gobierno de España en enero de 2026 podría albergar datos de consumo de material y energía de hospitales que, procesados por servicios de IA, generarían inventarios de Alcance 3 específicos por centro hospitalario y permitirían identificar las intervenciones de descarbonización con mayor impacto (sustitución de gases anestésicos de alto potencial de calentamiento global, optimización de la cadena de frío farmacéutica, reducción del uso de plásticos de un solo uso).

El sector turístico, que en España representa el 12,4% del PIB y genera emisiones significativas en transporte, alojamiento y actividades, presenta una problemática de Alcance 3 especialmente compleja por la multiplicidad de actores y la estacionalidad de la demanda. Las emisiones de un paquete turístico incluyen el transporte aéreo o terrestre del turista (categoría 9 para el touroperador), la energía consumida por el hotel (categoría 1), la alimentación (categoría 1), y las actividades de ocio. El espacio de datos EONA-X, con participación de Amadeus y Air France-KLM, ya integra datos de movilidad y turismo que podrían utilizarse para calcular la huella de carbono completa de un viaje y ofrecer al turista alternativas de menor impacto ambiental. El proyecto Spain Living Lab en Aragón explora esta dirección para el turismo de naturaleza y montaña en el Pirineo.

El sector de la construcción, responsable del 37% de las emisiones globales de CO<sub>2</sub> según la Global Alliance for Buildings and Construction, presenta una problemática de Alcance 3 centrada en el carbono incorporado de los materiales de construcción (categoría 1) y el uso energético de los edificios durante su vida útil (categoría 11). Las Declaraciones Ambientales de Producto (DAP/EPD), que proporcionan información estandarizada sobre el impacto ambiental de los materiales de construcción, constituyen activos de datos naturales para su publicación en espacios de datos sectoriales. Un servicio de IA que acceda a estas declaraciones desde el espacio de datos podría calcular automáticamente el carbono incorporado de un proyecto de construcción a partir del presupuesto o la medición del proyecto, integrando estándares como la Base de Precios de la Construcción y las bases de datos ambientales como BEDEC o InData.

# 3. Estado del arte: espacios de datos en Europa

## 3.1. La Estrategia Europea del Dato y el marco regulatorio

La Estrategia Europea del Dato, presentada por la Comisión Europea en febrero de 2020, estableció la visión de crear espacios de datos comunes europeos en sectores estratégicos, incluyendo movilidad, salud, energía, agroalimentación, finanzas, administración pública, manufactura y Green Deal. Esta visión se ha materializado en un marco legislativo compuesto por el Reglamento de Gobernanza de Datos (Data Governance Act, DGA), que entró en vigor en septiembre de 2023 y establece los mecanismos para facilitar la reutilización de datos del sector público y la intermediación de datos; la Ley de Datos (Data Act), plenamente aplicable desde septiembre de 2025, que obliga a fabricantes a hacer accesibles los datos de uso de sus dispositivos; y la Ley de Inteligencia Artificial (AI Act), que establece requisitos de transparencia y reporte de consumo energético para sistemas de IA de gran escala.

La Comisión Europea ha cuantificado el valor de la economía del dato en la UE en 434.030 millones de euros en 2020, con proyecciones de alcanzar 659.000 millones en 2025 y 851.500 millones en 2030. Este potencial económico justifica las inversiones masivas en infraestructura de espacios de datos que se están realizando a través de los programas Digital Europe, Horizon Europe y los mecanismos de recuperación post-COVID.

## 3.2. Gaia-X: Trust Framework 3.0

Gaia-X, la iniciativa francoalemana lanzada en 2019 para construir una infraestructura de datos segura, soberana e interoperable en Europa, ha alcanzado un hito significativo con la publicación del Trust Framework 3.0, denominado Danube, presentado en el Summit de Porto en noviembre de 2025. Esta versión introduce extensiones de dominio y geográficas que permiten federar la confianza entre ecosistemas diversos y regímenes de cumplimiento heterogéneos. El framework se sustenta en tres pilares: estándares de identidad uniformes, pruebas de conformidad estandarizadas y formatos de datos interoperables, lo que permite a las organizaciones registrarse una única vez y operar en múltiples espacios de datos.

Como señala Ulrich Ahle, CEO de Gaia-X AISBL, la iniciativa ha transitado de la fase de pilotos iniciales hacia un despliegue operativo real. El white paper de Gaia-X de 2025, titulado *The Role of Data Spaces in the Digital Economy*, identifica explícitamente la

sostenibilidad como uno de los ámbitos donde los espacios de datos pueden generar mayor impacto, señalando que estos pueden convertirse en habilitadores claves de la neutralidad de carbono, la eficiencia de recursos y la conservación de la biodiversidad. El documento destaca el concepto de sobriedad digital: a diferencia de los métodos tradicionales que duplican datos e incrementan el consumo energético, los espacios de datos permiten un acceso descentralizado a través de APIs y transacciones controladas, minimizando el procesamiento de datos y reduciendo el consumo energético del almacenamiento en la nube.

### 3.3. Eclipse Dataspace Components

El Eclipse Dataspace Connector (EDC), ahora denominado Eclipse Dataspace Components, constituye la implementación de referencia del protocolo de espacios de datos europeos. Desarrollado inicialmente en el contexto del proyecto Catena-X para la industria del automóvil, el EDC proporciona los componentes necesarios para la negociación de contratos, la gestión de catálogos de activos de datos y la transferencia segura de información entre conectores. El protocolo implementa una máquina de estados que garantiza que cada acceso a datos esté precedido por una negociación contractual digital que especifica qué datos se acceden, en qué condiciones y en qué ventana temporal. El resultado de esta negociación es un Endpoint Data Reference (EDR) con URL y token de autenticación de corta duración, que garantiza un acceso puntual, controlado y auditable.

El Mobility Data Space, operativo en Alemania como comunidad de intercambio de datos de movilidad, utiliza el EDC como base técnica y ofrece un servicio de Connector-as-a-Service (CaaS) que permite a los participantes incorporarse sin necesidad de desplegar infraestructura propia. Este modelo reduce significativamente las barreras de entrada, especialmente para pymes, y ha sido identificado por el proyecto deployEMDS como la oferta más efectiva para un onboarding técnico accesible.

### 3.4. Espacio Común Europeo de Datos de Movilidad (EMDS)

La Comisión Europea publicó en noviembre de 2023 la Comunicación sobre la creación de un espacio común europeo de datos de movilidad (EMDS), que establece la visión de una infraestructura de datos que facilite el acceso, la puesta en común y el intercambio de datos de fuentes de transporte y movilidad existentes y futuras. Según la Comunicación, el EMDS apoyará la transición hacia una movilidad sostenible e inteligente, permitiendo servicios de transporte y movilidad más eficientes y reduciendo así las emisiones, contribuyendo directamente a los objetivos del Green Deal europeo.

Los resultados preliminares del proyecto deployEMDS en sus nueve ciudades piloto revelan patrones comunes y divergencias significativas en la implementación de espacios de datos de movilidad. En ciudades como Milán y Lisboa, la existencia previa de sistemas

integrados de gestión del transporte ha facilitado la transición hacia el modelo de espacio de datos, ya que los operadores de transporte disponían de infraestructura de datos madura y estándares de intercambio consolidados (SIRI, NeTEx, GBFS). En ciudades más pequeñas, la fragmentación de operadores y la heterogeneidad de sistemas informáticos han requerido un esfuerzo adicional de armonización, que el servicio de Connector-as-a-Service del Mobility Data Space alemán ha ayudado a mitigar. Estas lecciones son directamente transferibles al contexto español, donde la diversidad de operadores de transporte urbano requerirá soluciones de onboarding flexibles y de bajo coste.

El informe de febrero de 2025 sobre el EMDS identifica explícitamente la combinación de compartición de datos con tecnologías emergentes como la inteligencia artificial, los gemelos digitales y los vehículos conectados y autónomos como catalizadores para la descarbonización de la movilidad urbana. El proyecto deployEMDS, con 38 beneficiarios y 7 socios asociados, está implementando la infraestructura del EMDS en nueve ciudades piloto, incluyendo casos de uso centrados en la integración multimodal, la gestión del tráfico y la monitorización de indicadores de movilidad urbana sostenible (SUMI). En Milán, la agencia AMAT está desarrollando un caso de uso centrado en la modelización de escenarios de movilidad futura orientados a alcanzar la neutralidad de carbono mediante Mobility as a Service (MaaS).

### 3.5. Green Deal Data Space

El espacio de datos del Green Deal europeo constituye una iniciativa transversal diseñada para hacer accesibles, utilizables y útiles los datos ambientales de múltiples fuentes. Como destaca la International Data Spaces Association (IDSA), este espacio de datos permitirá a los proveedores mantener la soberanía sobre sus datos y proteger información sensible, al tiempo que desbloquea el acceso a datos para miles de aplicaciones que contribuirán a hacer la sociedad más sostenible. Las áreas cubiertas incluyen cambio climático, economía circular, contaminación cero, biodiversidad, deforestación y aseguramiento del cumplimiento normativo.

### 3.6. Espacios de datos sectoriales relevantes para la descarbonización

| Espacio de datos | Sector              | Relevancia para huella de carbono                                 |
|------------------|---------------------|-------------------------------------------------------------------|
| Catena-X         | Automoción          | Product Carbon Footprint en cadena de suministro de automóviles   |
| deployEMDS       | Movilidad           | Descarbonización del transporte urbano con IA y gemelos digitales |
| EONA-X           | Movilidad / Turismo | Huella de carbono de movilidad turística multimodal               |

|                                           |                                       |                                                                                                                                                     |
|-------------------------------------------|---------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>AgriDataSpace / CEADS</b>              | Agroalimentación                      | Emisiones de la agricultura y trazabilidad alimentaria                                                                                              |
| <b>EuProGigant</b>                        | Manufactura                           | Predicción de huella de carbono en producción industrial                                                                                            |
| <b>Energy Data Space</b>                  | Energía                               | Optimización de consumo energético y renovables                                                                                                     |
| <b>European Health Data Space</b>         | Salud                                 | Descarbonización del sector sanitario (4,4% emisiones globales)                                                                                     |
| <b>EDMUS</b>                              | Movilidad urbana (España)             | Certificación de emisiones evitadas por electrificación de flota de autobuses (caso ECODES, capítulo 8)                                             |
| <b>DS4M Mediterráneo</b>                  | Movilidad regional (España)           | Cálculo y modelización de Zonas de Bajas Emisiones, seguimiento de PMUS y planes regionales en cinco comunidades autónomas                          |
| <b>DS4M CENTRAL</b>                       | Movilidad nacional (España)           | Indicadores ambientales y de tráfico mediante fusión de bases de datos heterogéneas en RAPs, PAPs y LAPs                                            |
| <b>DS4MM</b>                              | Movilidad transfronteriza (España–UE) | Análisis de emisiones en pilotos urbanos, regionales y transfronterizos; gestión de ZBE en la Comunitat Valenciana                                  |
| <b>Espacio de Datos de Turismo</b>        | Turismo (España)                      | Primer espacio sectorial operativo en España (dic. 2025); orientado a una gestión turística "más eficiente y sostenible" según el Mintur            |
| <b>Espacio Nacional de Datos de Salud</b> | Salud (España)                        | Infraestructura pública activada en enero 2026; sustrato para inventarios de Alcance 3 categoría 1 (material fungible, fármacos, gases anestésicos) |
| <b>AgrospAI</b>                           | Agroalimentación (España)             | Demostrador Gaia-X Lighthouse con Compute-to-Data sobre Pontus-X; integrado en el despliegue europeo del CEADS                                      |
| <b>Agrixels (UPCxels)</b>                 | Agroalimentación (España)             | Estándar de datos y ontología agroalimentaria que cubre explícitamente sostenibilidad, trazabilidad y mercados de carbono                           |
| <b>GroundX</b>                            | Agroalimentación (España)             | Espacio federado para optimización técnica del sector agrícola integrando información productiva, ambiental y de mercado                            |
| <b>CitriData</b>                          | Agroalimentación (España)             | Espacio estratégico para la cadena de valor citrícola andaluza, exportación intensiva a mercados europeos con exigencias de PCF                     |
| <b>ESAUTO-X</b>                           | Automoción (España)                   | Centro demostrador del sector basado en Catena-X, orientado explícitamente a movilidad sostenible                                                   |
| <b>CEMENGAL</b>                           | Industria cementera (España)          | Espacio federado Gaia-X / IDS-RAM para descarbonización del sector cementero, expuesto al CBAM desde 2026                                           |
| <b>GreenBIMxels</b>                       | Edificación pública (España)          | Estándar BIM enriquecido con datos energéticos, ambientales y de uso, para servicios de IA orientados a sostenibilidad y eficiencia energética      |

|                                                |                                   |                                                                                                                                                                 |
|------------------------------------------------|-----------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>AECTech</b>                                 | Sector viario (España)            | Datos de tráfico, infraestructuras y movilidad para mejorar seguridad, eficiencia y sostenibilidad                                                              |
| <b>COLDS</b>                                   | Transporte de mercancías (España) | Compartición de datos para optimizar eficiencia y seguridad logística (categoría 4 GHG Protocol)                                                                |
| <b>Hábitat Data (PRECEDENT)</b>                | Industria del mueble (España)     | Datos de mercado y producto del sector hábitat, sustrato para Declaraciones Ambientales de Producto                                                             |
| <b>Espacio de Datos de Sacyr</b>               | Agua e infraestructuras (España)  | Operativo, incluye módulo IA para Índice de Calidad del Agua, ampliable a optimización energética carbon-aware                                                  |
| <b>ECiCo</b>                                   | Construcción (España)             | Espacio orientado explícitamente a la economía circular en construcción, "permite a las empresas reducir su huella de carbono y optimizar el uso de materiales" |
| <b>OPDS (Open Ports Data Space)</b>            | Logística portuaria (España)      | Espacio interoperable para puertos de Valencia, Baleares y Barcelona con IA y Big Data, "reduciendo costes, tiempos y emisiones"                                |
| <b>PGTEC</b>                                   | Emergencias climáticas (España)   | Plataforma para anticipar y gestionar inundaciones e incendios, transición de modelo reactivo a prevención proactiva                                            |
| <b>Lista de Confianza de Espacios de Datos</b> | Transversal (España)              | Registro oficial regulado por la Orden TDF/1207/2025; meta-espacio de identidades verificadas para servicios de IA que requieran fuentes auditables             |

## 3.7. Marco de interoperabilidad

El Data Spaces Support Centre (DSSC), financiado por el programa Digital Europe de la Comisión Europea, desempeña un papel central en la coordinación y armonización de los espacios de datos europeos. El DSSC ha publicado el Data Spaces Blueprint, un documento de referencia que establece los componentes funcionales, los patrones de interacción y los estándares técnicos que deben implementar todos los espacios de datos comunes europeos. El Blueprint distingue entre el plano de infraestructura (conectores, registros de identidad, catálogos federados), el plano de confianza (credenciales verificables, políticas de acceso, auditoría) y el plano de negocio (modelos de datos, vocabularios semánticos, contratos inteligentes).

Para la sostenibilidad, el DSSC ha identificado la necesidad de vocabularios semánticos compartidos que permitan describir activos de datos ambientales de forma interoperable entre sectores. Un factor de emisión publicado en el espacio de datos de energía debe ser consumible sin transformación por un servicio de IA operando en el espacio de datos de movilidad. Esta interoperabilidad semántica es especialmente crítica para el Alcance 3, donde las emisiones de una organización en un sector dependen de datos originados en sectores completamente diferentes. El DSSC trabaja con ontologías como DCAT (Data Catalog Vocabulary) para la descripción de activos, ODRL (Open Digital Rights Language)

para la expresión de políticas de acceso, y estándares emergentes como el GHG Protocol Digital Reporting Taxonomy para la representación estructurada de datos de emisiones.

## 3.8. Arquitectura técnica del protocolo DSP

El Dataspace Protocol (DSP), estandarizado por IDSA y la Eclipse Foundation, define un modelo de interacción entre participantes que garantiza la soberanía del dato en cada transacción. El protocolo se estructura en cuatro fases secuenciales. La fase de descubrimiento permite al consumidor (DSC) consultar el catálogo del proveedor (DSP) para identificar activos de datos disponibles, sus políticas de acceso y los metadatos asociados. La fase de negociación de contrato establece un acuerdo jurídicamente vinculante entre las partes, expresado en ODRL, que especifica las condiciones de uso del dato (duración, propósito, restricciones de redistribución). La fase de transferencia inicia el canal de datos, que puede ser HTTP pull (el consumidor solicita los datos) o HTTP push (el proveedor envía los datos). Finalmente, la fase de acceso utiliza el EDR (Endpoint Data Reference) para autenticar y autorizar cada petición individual de datos.

Este modelo de interacción es fundamental para los servicios de IA sobre datos ambientales porque garantiza que cada acceso a datos de emisiones de un proveedor queda registrado, auditado y limitado en el tiempo. Un servicio de predicción de Alcance 3 que necesite acceder a datos de diez proveedores diferentes generará diez negociaciones de contrato independientes, cada una con sus propias condiciones. Esta granularidad permite a los proveedores controlar exactamente qué datos comparten, con quién, durante cuánto tiempo y para qué propósito, eliminando el riesgo de que un intermediario acumule datos de emisiones que podrían revelar información competitiva sensible sobre procesos productivos o eficiencia energética.

## 3.9. Modelos de gobernanza y sostenibilidad de los espacios de datos

La sostenibilidad económica de los espacios de datos constituye un desafío crítico para su adopción a largo plazo. A diferencia de las plataformas centralizadas de datos, que generan ingresos mediante la monetización directa de los datos agregados, los espacios de datos soberanos deben encontrar modelos de negocio que respeten la soberanía de los participantes. El white paper de Gaia-X sobre el papel de los espacios de datos en la economía digital identifica varios modelos viables: tasas de participación (cuotas por pertenecer al espacio de datos), comisiones por transacción (porcentaje del valor de cada intercambio de datos), servicios de valor añadido (análisis, predicción, certificación) y financiación pública para infraestructuras de interés general.

Para los servicios de IA orientados a la sostenibilidad, el modelo de servicios de valor añadido es particularmente prometedor. Un servicio de predicción de Alcance 3 que opere

sobre un espacio de datos puede cobrar a las empresas solicitantes por cada cálculo realizado, mientras que los proveedores de datos participan gratuitamente (o incluso reciben una compensación) porque la compartición de sus datos de emisiones contribuye a la transparencia de la cadena de valor y puede ser un requisito de sus clientes bajo la CSRD. Este modelo de gobernanza crea incentivos alineados: los grandes compradores financian el servicio, los proveedores participan por obligación regulatoria o presión de la cadena de valor, y el operador del espacio de datos mantiene la infraestructura con los ingresos generados por los servicios de IA.

La gobernanza técnica del espacio de datos debe abordar también cuestiones de calidad de datos, resolución de conflictos y evolución de estándares. Un comité técnico con representación de proveedores, consumidores y operadores del espacio de datos puede establecer estándares mínimos de calidad para los activos de datos ambientales (frecuencia de actualización, granularidad, completitud), definir procedimientos de validación automática que los servicios de IA pueden ejecutar como prerrequisito de cualquier cálculo, y gestionar la evolución de los vocabularios semánticos a medida que los estándares de reporte (ESRS, GHG Protocol) se actualizan. La UNE 0087:2025 proporciona el marco base para esta gobernanza en el contexto español.

La implementación práctica del protocolo DSP en el contexto del caso de uso ECODES-EDMUS, descrito en el capítulo 8, ha revelado aspectos técnicos relevantes para el despliegue de servicios de IA. El tiempo medio de un ciclo completo de negociación, transferencia y acceso oscila entre 2 y 5 segundos por activo (en experimentos realizados, podría cambiar en edc en producción), lo que impone restricciones de diseño para servicios que necesiten acceder a múltiples activos en una sola operación. La reutilización de acuerdos negociados mediante cachés en memoria reduce significativamente la latencia total, y la implementación de patrones de acceso por lotes (batch) permite a un agente de IA descargar eficientemente datos de decenas de proveedores dentro de ventanas temporales razonables.

# 4. Espacios de datos en España: iniciativas y convocatorias del Ministerio

## 4.1. El Plan de Impulso de los Espacios de Datos Sectoriales

España ha desplegado el Plan de Impulso de los Espacios de Datos Sectoriales como una de las iniciativas más ambiciosas de Europa en este ámbito, con un presupuesto total de 500 millones de euros distribuidos en 6 ejes y 11 iniciativas hasta 2026. Este plan se enmarca en la inversión 1 del componente 12 del Plan de Recuperación, Transformación y Resiliencia, financiado por el Mecanismo de Recuperación y Resiliencia de la UE (NextGenerationEU), y está alineado con la Agenda España Digital 2026, la Estrategia Europea del Dato y los Reglamentos de Gobernanza de Datos y de Datos.

El Plan cubre sectores estratégicos con impacto directo en la huella de carbono: movilidad sostenible (optimización del transporte y planificación urbana mediante análisis de patrones de tráfico), energía (gestión eficiente de recursos energéticos), agroalimentación (trazabilidad y calidad alimentaria, con iniciativas como AgriDataSpace), salud (integración de datos clínicos para gestión sanitaria sostenible), industria y comercio.

## 4.2. Convocatorias específicas y presupuestos

| Convocatoria                               | Presupuesto | Estado (marzo 2026)                        | Plazo   |
|--------------------------------------------|-------------|--------------------------------------------|---------|
| 1ª Conv. Demostradores y Casos de Uso      | 150 M€      | Resolución concedida (90 proyectos, 83 M€) | Cerrada |
| 2ª Conv. Demostradores y Casos de Uso      | 75 M€       | Resolución definitiva publicada            | Cerrada |
| Productos y Servicios de Espacios de Datos | 44 M€       | Resolución concedida                       | Cerrada |
| Kit Espacios de Datos (Red.es)             | 60 M€       | Abierta (>60% presupuesto disponible)      | Cerrada |

|                                    |     |                       |         |
|------------------------------------|-----|-----------------------|---------|
| Espacio Nacional de Datos de Salud | N/D | Activado (enero 2026) | Cerrada |
|------------------------------------|-----|-----------------------|---------|

## 4.3. El programa Kit Espacios de Datos

El programa Kit Espacios de Datos, lanzado por el Ministerio para la Transformación Digital y de la Función Pública en octubre de 2025, representa un cambio de paradigma en la estrategia de despliegue al adoptar el modelo de éxito del Kit Digital. Con una dotación de 60 millones de euros gestionados por Red.es, las solicitudes se resuelven en régimen de concurrencia no competitiva por orden de llegada, lo que reduce significativamente las barreras administrativas. Los espacios de datos elegibles incluyen los desplegados en el Plan de Impulso, el Espacio Nacional de Datos de Salud y aquellos identificados en la Lista de Confianza de Espacios de Datos, un registro voluntario regulado por la Orden TDF/758/2025 y basado en la Especificación UNE 0087:2025, que establece el primer marco común en España para la creación y operación de espacios de datos.

El Kit Espacios de Datos se estructura en dos líneas de ayuda diferenciadas. La Línea 1 está dirigida a participantes que desean incorporarse a un espacio de datos ya operativo, cubriendo los costes de adaptación tecnológica (despliegue de conector EDC o contratación de CaaS), formación del personal, consultoría de integración y auditoría de cumplimiento con la UNE 0087:2025. La Línea 2 está dirigida a proveedores tecnológicos que desarrollen productos y servicios para espacios de datos, incluyendo conectores, herramientas de gestión de catálogos, servicios de intermediación de datos y, significativamente, servicios de valor añadido basados en IA que operen sobre la infraestructura del espacio de datos. Esta segunda línea es especialmente relevante para centros tecnológicos como ITA, que pueden posicionarse como proveedores de servicios de IA climática homologados para los espacios de datos de la Lista de Confianza.

El Centro de Referencia de Espacios de Datos (CRED), lanzado simultáneamente con el Kit, actúa como punto de orientación para los interesados y dinamizador de los ecosistemas, facilitando el acceso a las ayudas y acompañando el proceso de incorporación. Esta infraestructura institucional es particularmente relevante para el despliegue de servicios de IA sobre espacios de datos, ya que proporciona el marco de confianza y gobernanza necesario para que los datos ambientales puedan circular de forma controlada entre los participantes.

## 4.4. Espacios de datos operativos en España

A fecha de marzo de 2026, España cuenta con un ecosistema creciente de espacios de datos en diversos sectores. El Espacio Nacional de Datos de Salud, activado por el

Gobierno en enero de 2026, constituye la primera infraestructura pública de este tipo a nivel nacional, orientada a compartir y utilizar datos sanitarios de forma segura para mejorar la atención al paciente, impulsar la investigación médica y facilitar el desarrollo de soluciones digitales en salud. En el ámbito de la movilidad, proyectos como EDMUS (European Data Mobility Use Space), que se describe en detalle en el capítulo 8, demuestran la viabilidad del intercambio soberano de datos de transporte para la certificación de reducciones de emisiones.

El proyecto Spain Living Lab, en el que participa ITA, está creando una plataforma de datos e inteligencia artificial sobre turismo de naturaleza y montaña en Aragón, con financiación de 900.000 euros del Plan de Recuperación a través de la RETECH. Este proyecto integra living labs, espacios de datos y redes de hubs de I+D+i con el objetivo de transformar el modelo económico del turismo hacia uno basado en IA, con la sostenibilidad como eje transversal.

## 4.5. La oportunidad para servicios de IA en el contexto español

La confluencia de inversiones masivas (500 M€ en el Plan de Impulso), convocatorias con presupuesto disponible (Kit Espacios de Datos con más del 60% no adjudicado), normativa habilitadora (UNE 0087:2025, CSRD) y capacidades tecnológicas demostradas (EDMUS, Catena-X) crea un contexto óptimo para el despliegue de servicios de IA orientados a la reducción de huella de carbono. Las empresas españolas que se incorporen a espacios de datos sectoriales tendrán acceso a datos distribuidos de cadenas de valor que, procesados por servicios de IA, pueden transformarse en inventarios de emisiones de Alcance 3 precisos, planes de descarbonización optimizados y certificados verificables de reducción de emisiones.

## 4.6. La Especificación UNE 0087:2025 y la Lista de Confianza

La Especificación UNE 0087:2025, publicada por la Asociación Española de Normalización (UNE), constituye el primer marco común español para la creación, operación y evaluación de espacios de datos. Esta especificación define los requisitos mínimos que debe cumplir un espacio de datos para ser incluido en la Lista de Confianza regulada por el Ministerio, abarcando aspectos como la gobernanza del espacio de datos, los requisitos de identidad y autenticación de los participantes, las políticas de acceso y uso de datos, los mecanismos de resolución de conflictos y los estándares de interoperabilidad técnica.

La Lista de Confianza funciona como un registro voluntario que otorga visibilidad y credibilidad a los espacios de datos que cumplen los requisitos de la UNE 0087:2025. Para los servicios de IA orientados a la sostenibilidad, la pertenencia a un espacio de

datos de la Lista de Confianza proporciona garantías sobre la calidad de la gobernanza, la identidad verificada de los participantes y la auditabilidad de las transacciones de datos, elementos esenciales para que los cálculos de huella de carbono basados en datos compartidos sean aceptados por auditores externos bajo los requisitos de aseguramiento de la CSRD.

Aragón se posiciona como un territorio piloto privilegiado para el despliegue de servicios de IA sobre espacios de datos orientados a la sostenibilidad, gracias a la confluencia de varios factores diferenciales. En primer lugar, la presencia de ITA como centro tecnológico con capacidades demostradas en Eclipse Dataspace Connector, inteligencia artificial y Big Data. En segundo lugar, la colaboración estratégica con ECODES, que aporta la dimensión ambiental y el acceso a empresas y administraciones comprometidas con la descarbonización. En tercer lugar, la diversidad sectorial de la economía aragonesa, que incluye logística (PLAZA, la mayor plataforma logística de Europa), automoción (Stellantis/Opel en Figueruelas), agroalimentación (sector porcino, vinícola, oleicultura) y energías renovables (Aragón es la comunidad autónoma con mayor potencia eólica instalada de España).

La Estrategia Aragonesa de Inteligencia Artificial (EAIA), publicada por ITA en 2022, identifica la sostenibilidad como uno de los ámbitos prioritarios de aplicación de la IA en la región. El proyecto Spain Living Lab, con financiación de 900.000 euros del Plan de Recuperación, integra espacios de datos con IA para el turismo sostenible en el Pirineo aragonés. La Universidad de Zaragoza, a través de su Instituto de Investigación en Ingeniería de Aragón (I3A) y del grupo de investigación en Sistemas de Información Distribuidos (SID), aporta capacidad investigadora en modelado de datos, ontologías y sistemas distribuidos. El ecosistema se completa con el Clúster de la Logística de Aragón (ALIA), que agrupa a más de 120 empresas del sector logístico con potencial de participación en espacios de datos de movilidad y transporte.

La plataforma PLAZA (Plataforma Logística de Zaragoza), con sus 13,2 millones de metros cuadrados y más de 400 empresas instaladas, genera un volumen de datos de transporte y logística que podría alimentar servicios de IA de optimización de emisiones de Alcance 3 en la categoría 4 (transporte y distribución aguas arriba). Un espacio de datos que integre los flujos logísticos de PLAZA con datos de transporte ferroviario (ADIF/RENFE), transporte por carretera (DGT) y operadores de última milla permitiría construir un gemelo digital logístico del noreste peninsular, sobre el cual los servicios de IA podrían optimizar rutas multimodales minimizando emisiones de CO<sub>2</sub> para cada envío individual.

La ubicación de ITA en Zaragoza, núcleo logístico de la península ibérica y sede de la Feria Internacional de Logística y Transporte (SIL/Logistics Spain), posiciona al centro tecnológico en un contexto privilegiado para el desarrollo de servicios de IA sobre datos de movilidad y logística. La colaboración establecida con ECODES añade la dimensión ambiental necesaria para que estos servicios tengan impacto real en la reducción de emisiones, creando un binomio tecnología-sostenibilidad con potencial de escalado a nivel nacional y europeo.

# 5. Inteligencia artificial para la reducción de huella de carbono: estado del arte

## 5.1. Machine Learning para la predicción de emisiones de Alcance 3

La aplicación de algoritmos de machine learning a la predicción de emisiones de Alcance 3 constituye una de las líneas de investigación más activas en la intersección entre IA y sostenibilidad. El trabajo seminal de Serafeim y Vélez Caicedo (2022), publicado como Working Paper 22-080 de Harvard Business School, entrena algoritmos de ML sobre los 15 tipos reportados de emisiones de Alcance 3, utilizando como inputs variables ampliamente disponibles de estados financieros, emisiones de Alcance 1 y 2, y clasificaciones industriales. Sus resultados demuestran que la mayoría de los tipos de emisiones de Alcance 3 pueden predecirse con mayor precisión usando algoritmos de Adaptive Boosting (AdaBoost) en comparación con modelos de regresión lineal y otros algoritmos supervisados.

Este enfoque ha sido extendido por Wang y Ye (2025) en *Frontiers in Sustainability*, quienes comparan sistemáticamente Random Forest, XGBoost y AdaBoost para la predicción de emisiones de Alcance 3, evaluando el compromiso entre precisión predictiva e interpretabilidad. Sus hallazgos confirman la superioridad de los modelos de ensemble, particularmente XGBoost, para la estimación de emisiones en empresas que carecen de información detallada de cadena de suministro. Las características de importancia de variables inherentes a estos algoritmos proporcionan información valiosa sobre los factores determinantes de las emisiones, facilitando la priorización de intervenciones de reducción.

El dataset ExioML, publicado en *Scientific Data (Nature)* en febrero de 2026, representa un avance significativo al proporcionar un conjunto global de factores de emisión optimizado para aplicaciones de machine learning en Alcance 3. Basado en tablas input-output multirregionales extendidas ambientalmente (EE-MRIO), el dataset incluye emisiones de GEI, población, PIB e intensidad energética primaria, descompuestos según la identidad de Kaya, y está diseñado específicamente para tareas como la detección automatizada de hotspots de emisiones, la previsión de escenarios y la identificación de anomalías.

No obstante, conviene matizar que el uso de machine learning para predecir la huella de carbono no constituye, por sí mismo, una aproximación adecuada como herramienta de reducción efectiva de emisiones. Estos modelos producen estimaciones agregadas que

no permiten identificar con exactitud las fuentes reales de emisión dentro de la cadena de valor, lo que limita las posibilidades de intervención y la trazabilidad necesaria para verificar reducciones. Su valor reside, más bien, en orientar la toma de decisiones a un nivel estratégico: estimar órdenes de magnitud, fijar objetivos plausibles e identificar zonas o procesos en los que una empresa debería concentrar su esfuerzo de medición. La reducción efectiva exige avanzar hacia datos cada vez más concretos y trazables, obtenidos directamente de los actores de la cadena de valor —precisamente el cometido de los espacios de datos, que sustituyen progresivamente la estimación por información primaria compartida en condiciones de soberanía y confianza.

## 5.2. IA para la descarbonización de cadenas de suministro

La descarbonización de cadenas de suministro mediante IA ha sido analizada en profundidad en el artículo de MDPI Sustainability (2025), que presenta estudios de caso en los sectores agroalimentario, textil y logística portuaria de Túnez. Los resultados demuestran reducciones de emisiones de entre 500 y 1.500 tCO<sub>2</sub> anuales mediante la aplicación de modelos híbridos de IA que combinan optimización de rutas de vehículos (VRP), predicción de demanda con redes LSTM y scheduling con reinforcement learning. Las intervenciones a nivel de empresa generaron reducciones anuales de 10,8 a 300 tCO<sub>2</sub>, mientras que la optimización portuaria en Rades alcanzó entre 670 y 1.000 tCO<sub>2</sub> anuales, equivalentes a eliminar aproximadamente 150 camiones de circulación.

## 5.3. IA aplicada al Alcance 3

El artículo de MDPI Sustainability de noviembre de 2025, publicado por investigadores de la RMIT University de Melbourne, introduce el pensamiento de sistemas críticos (Critical Systems Thinking, CST) como marco para evaluar la practicidad de la IA en el contexto del Alcance 3. El trabajo identifica las contribuciones específicas de la IA a la calidad de los datos de Alcance 3: automatización de procesos de limpieza de datos, detección y corrección de errores en grandes conjuntos de datos, identificación de lagunas en datos de cadena de suministro, detección de anomalías que sugieran falta de fiabilidad, y funciones de auditoría como muestreo, identificación de outliers, detección de patrones y tendencias, y gestión de riesgos mejorada.

Un aspecto que no puede ignorarse es la huella de carbono de los propios sistemas de IA. Según un estudio de diciembre de 2025, los sistemas de IA operando en centros de datos podrían haber sido responsables de entre 32,6 y 79,7 millones de toneladas de CO<sub>2</sub> en 2025, comparable a las emisiones de Noruega o de la ciudad de Nueva York. Sin embargo, investigaciones de Google muestran mejoras drásticas: una reducción de 33 veces en energía y 44 veces en carbono por prompt medio entre 2024 y 2025. Europa, con una intensidad de carbono eléctrica de aproximadamente 174 gCO<sub>2</sub>/kWh (menos de

la mitad de la media global de 445 gCO<sub>2</sub>/kWh), ofrece un contexto favorable para el despliegue de servicios de IA con menor impacto ambiental.

Esta paradoja sugiere que los servicios de IA para la reducción de huella de carbono deben evaluarse en términos de impacto neto: si un modelo de ML que consume X kWh de electricidad permite evitar Y toneladas de CO<sub>2</sub> en la cadena de suministro de una empresa, el balance será positivo siempre que Y supere significativamente la huella de X. Los estudios de caso analizados sugieren ratios de beneficio-coste de carbono de entre 10:1 y 100:1 para aplicaciones bien diseñadas.

## 5.4. LLMs para reportes de sostenibilidad

Los modelos de lenguaje de gran escala (Large Language Models, LLMs) abren nuevas posibilidades para la automatización de procesos relacionados con el Alcance 3 que trascienden la predicción numérica. En primer lugar, la extracción estructurada de información: los LLMs pueden procesar documentos no estructurados como facturas de proveedores, albaranes de transporte, informes de sostenibilidad de terceros y certificados ambientales, extrayendo automáticamente los datos de actividad necesarios para el cálculo de emisiones (cantidades, materiales, distancias, modos de transporte). Esta capacidad es especialmente valiosa para las pymes, que frecuentemente disponen de información ambiental solo en formato documental no digitalizado.

En segundo lugar, el emparejamiento semántico de factores de emisión: un LLM puede recibir la descripción textual de un bien o servicio adquirido (por ejemplo, «suministro de tubería de polietileno de alta densidad PE100, diámetro nominal 160 mm, presión nominal 10 bar») y seleccionar automáticamente el factor de emisión más apropiado de bases de datos como DEFRA, EPA o ecoinvent, considerando la geografía, el proceso productivo y el nivel de detalle disponible. Este emparejamiento, que actualmente requiere horas de trabajo de un consultor especializado, puede realizarse en segundos con un LLM fine-tuned sobre datos de factores de emisión.

En tercer lugar, la generación automatizada de reportes CSRD: los LLMs pueden asistir en la redacción de las secciones de informes de sostenibilidad relativas a emisiones de Alcance 3, incluyendo las descripciones metodológicas, la identificación de categorías materiales, la justificación de exclusiones y la narración de las incertidumbres asociadas. Cuando estos LLMs operan sobre datos accedidos a través de un espacio de datos soberano, la trazabilidad de las fuentes de información proporciona la base de evidencia necesaria para el aseguramiento externo del reporte.

## 5.5. Reinforcement Learning para la optimización de emisiones en tiempo real

El Reinforcement Learning (RL) ofrece un paradigma complementario al machine learning supervisado para la reducción de emisiones, particularmente en escenarios donde las

decisiones deben tomarse de forma secuencial y en tiempo real. Google ha demostrado la eficacia de este enfoque en la gestión energética de centros de datos, donde agentes de RL controlan sistemas de refrigeración para minimizar el consumo energético manteniendo las condiciones operativas, logrando reducciones del 40% en el consumo de energía de refrigeración.

En el contexto de los espacios de datos de movilidad, los agentes de RL pueden optimizar en tiempo real la asignación de vehículos eléctricos y de combustión a rutas de transporte público, considerando simultáneamente la demanda de pasajeros, el estado de carga de las baterías, la intensidad de carbono de la red eléctrica en cada momento y las restricciones operativas. Este enfoque requiere acceso a datos en tiempo real de múltiples fuentes (billetaje, telemetría de vehículos, datos de la red eléctrica), que un espacio de datos federado puede proporcionar de forma soberana y con latencia controlada.

## 5.6. Computer vision y análisis de imágenes satélite para verificación de emisiones

Las tecnologías de computer vision aplicadas a imágenes satelitales constituyen una fuente emergente de datos para la verificación independiente de emisiones de GEI. Satélites como los de la constelación Sentinel del programa Copernicus de la ESA, junto con constelaciones comerciales como las de GHGSat, MethaneSAT y Carbon Mapper, proporcionan mediciones directas de concentraciones de CO<sub>2</sub> y metano en la atmósfera con resolución espacial creciente. Los modelos de deep learning, particularmente las redes convolucionales y los transformers de visión (ViT), pueden procesar estas imágenes para detectar y cuantificar fuentes puntuales de emisiones, identificar fugas de metano en infraestructuras de gas y petróleo, y estimar emisiones a escala regional.

En el contexto de los espacios de datos, las imágenes satelitales pueden publicarse como activos de datos ambientales que los servicios de IA consuman para validar las autodeclaraciones de emisiones de los participantes. Si una empresa reporta una reducción del 30% en sus emisiones de Alcance 1 tras la modernización de una planta industrial, un servicio de IA con acceso a datos satelitales históricos y actuales podría verificar independientemente esta reducción comparando las concentraciones atmosféricas en el área de la planta antes y después de la intervención. Esta capacidad de verificación independiente es especialmente valiosa para el aseguramiento de reportes CSRD y para la certificación de reducciones en mercados de carbono.

El programa europeo Copernicus, con su servicio de monitorización atmosférica (CAMS) y su servicio de cambio climático (C3S), proporciona datos abiertos que pueden integrarse en espacios de datos europeos como activos de referencia. La combinación de estos datos satelitales con datos empresariales de actividad (consumos energéticos, producción, logística) accedidos a través del espacio de datos soberano permite construir modelos de IA que cruzan información bottom-up (reportada por la empresa) con

información top-down (observada desde el espacio), generando estimaciones de emisiones con incertidumbre reducida y trazabilidad bidireccional.

La combinación de RL con modelos de predicción de demanda basados en LSTM (Long Short-Term Memory) permite anticipar picos de demanda y pre-posicionar recursos de transporte de baja emisión, reduciendo tanto las emisiones como los tiempos de espera. Esta síntesis de técnicas de IA operando sobre datos federados representa el estado más avanzado de la aplicación de inteligencia artificial a la descarbonización del transporte, y su despliegue sobre espacios de datos como el EMDS es técnicamente viable con la infraestructura actual.

# 6. Estado del arte internacional: IA y espacios de datos para sostenibilidad

## 6.1. Alemania: Catena-X y EuProGigant

Alemania lidera el desarrollo de espacios de datos para sostenibilidad con dos proyectos emblemáticos. Catena-X, el espacio de datos del sector automotriz, ha desarrollado un estándar para el Product Carbon Footprint (PCF) que permite a fabricantes de automóviles y sus proveedores intercambiar datos de huella de carbono a lo largo de cadenas de suministro con miles de participantes. El PCF en Catena-X se calcula según la Pathfinder Framework de WBCSD/PACT y se intercambia soberanamente a través del EDC, lo que permite a un OEM como BMW o Volkswagen conocer la huella de carbono específica de cada componente que adquiere, sin que el proveedor tenga que revelar sus datos de proceso.

EuProGigant, el proyecto lighthouse de Gaia-X para la industria manufacturera europea, ha presentado en la Hannover Messe resultados de investigación sobre predicción de huella de carbono en producción industrial. Los demostradores muestran cómo los datos compartidos soberanamente entre fabricantes pueden utilizarse para optimizar la cadena de valor desde una perspectiva de emisiones, utilizando modelos de IA para predecir el impacto ambiental de decisiones de producción antes de ejecutarlas.

## 6.2. Francia: EONA-X

Francia, a través de EONA-X, ha desarrollado un espacio de datos para movilidad, transporte y turismo que integra actores como Amadeus, Renault Group, Air France-KLM y Accor. Este espacio de datos, conforme a los principios de Gaia-X, permite agregar flujos de datos de diferentes fuentes de movilidad y turismo para generar dashboards que muestran huellas ambientales y beneficios socioeconómicos, facilitando comparaciones entre dominios y apoyando la toma de decisiones basada en evidencia. La Caisse des Dépôts, institución financiera pública francesa, ha sido clave en el desarrollo de entornos de intercambio de datos confiables, promoviendo la colaboración entre actores públicos y privados para la transición ecológica.

## 6.3. Austria: Hub Gaia-X

El Gaia-X Hub Austria, establecido en 2022 por iniciativa del Ministerio Federal de Acción Climática (BMK), ha hecho de la accesibilidad de datos para aplicaciones verdes y sostenibles una prioridad especial. Henriette Spyra, responsable de Innovación y Tecnología del BMK, destaca que los ecosistemas de datos juegan un papel crucial en la IA al proporcionar acceso a datos diversos y de alta calidad, y que el Hub combina actividades austriacas y europeas para permitir un intercambio de datos confiable orientado explícitamente a la sostenibilidad.

Fuera de Europa, la convergencia entre IA, datos y sostenibilidad sigue trayectorias diferentes pero complementarias. En Estados Unidos, una Orden Ejecutiva de enero de 2025 dirigió al Departamento de Energía la elaboración de requisitos de reporte para centros de datos de IA que cubran todo su ciclo de vida, incluyendo métricas de carbono incorporado, uso de agua y calor residual. Se lanzó también un Grand Challenge DOE-EPA para reducir la ratio PUE por debajo de 1,1. Estas iniciativas, aunque centradas en la huella de la propia IA, generan infraestructura de medición y reporte que puede aprovecharse para servicios de IA sobre datos ambientales. En Asia, Gaia-X ha iniciado la internalización de su framework con extensiones a Japón y Corea, lo que podría abrir el camino a espacios de datos transfronterizos para cadenas de suministro globales.

## 6.5. Países Nórdicos

Los países nórdicos presentan un ecosistema particularmente avanzado en la intersección de datos abiertos y sostenibilidad. Finlandia, a través de SITRA (Finnish Innovation Fund), ha desarrollado el concepto de fair data economy como pilar de su estrategia de transición ecológica, estableciendo reglas del juego para el intercambio de datos entre organizaciones que priorizan la sostenibilidad. El proyecto IHAN (Inhimilliset Arkirätkäisut, Soluciones Humanas para la Vida Cotidiana) explora cómo los datos personales de movilidad, consumo energético y compras pueden utilizarse, con consentimiento explícito del ciudadano, para calcular huellas de carbono individuales y recomendar alternativas más sostenibles.

Noruega merece una mención especial por su experiencia en la electrificación del transporte. Con una cuota de mercado de vehículos eléctricos superior al 80% en 2025, Noruega genera datos de movilidad eléctrica a escala que pueden alimentar modelos de IA de predicción de emisiones evitadas por la electrificación. La Autoridad Noruega de Protección de Datos (Datatilsynet) ha desarrollado directrices para el uso de datos de movilidad que equilibran la privacidad individual con la utilidad colectiva para la planificación urbana sostenible, un equilibrio que los espacios de datos soberanos están diseñados para implementar técnicamente.

Suecia, con su objetivo de alcanzar la neutralidad climática en 2045 (cinco años antes que el objetivo europeo), ha integrado la compartición de datos ambientales en su estrategia

de digitalización industrial. El proyecto Fossifritt Sverige (Suecia Libre de Fósiles) trabaja con 22 hojas de ruta sectoriales que requieren datos de emisiones compartidos entre empresas de diferentes sectores para identificar sinergias de descarbonización, como el uso de calor residual industrial para calefacción urbana o la optimización conjunta de logística entre sectores complementarios. Dinamarca, por su parte, ha desarrollado un hub de datos energéticos (Energinet DataHub) que proporciona acceso en tiempo real a datos de consumo, producción e intensidad de carbono de la red eléctrica, habilitando servicios de IA que optimizan el consumo eléctrico en función de la disponibilidad de energía renovable.

## 6.6. Iniciativas globales de estandarización: PACT, CDP y GRI en el contexto de espacios de datos

La estandarización internacional del intercambio de datos de emisiones avanza en paralelo a los espacios de datos europeos. La Partnership for Carbon Transparency (PACT), liderada por WBCSD, ha publicado el Pathfinder Framework v2.0 y la especificación técnica Pathfinder Network, que define un protocolo HTTP/REST para el intercambio de Product Carbon Footprints entre empresas. Aunque PACT no utiliza explícitamente la infraestructura EDC, su protocolo es compatible con el modelo de espacios de datos soberanos, y Catena-X ya ha integrado el formato PCF de PACT en su implementación. La convergencia entre el protocolo PACT y el Dataspace Protocol (DSP) constituiría un avance decisivo para la interoperabilidad global del intercambio de datos de emisiones.

El Carbon Disclosure Project (CDP), que gestiona el sistema global de divulgación ambiental más utilizado (con más de 23.000 empresas participantes en 2024), ha iniciado la digitalización de su plataforma de reporting con APIs estructuradas que podrían integrarse en espacios de datos como fuentes de datos de referencia para emisiones corporativas. De manera similar, la Global Reporting Initiative (GRI) está trabajando en taxonomías digitales para sus estándares de sostenibilidad que facilitarían la publicación automatizada de datos de emisiones en catálogos de espacios de datos. La convergencia de estos estándares globales con la infraestructura europea de espacios de datos creará un ecosistema en el que los servicios de IA podrán acceder a datos de emisiones de empresas de cualquier geografía a través de interfaces estandarizadas, lo que es esencial para el cálculo de emisiones de Alcance 3 en cadenas de suministro globales.

La Organización Internacional de Normalización (ISO) ha publicado recientemente la norma ISO 14083:2023 para la cuantificación y reporte de emisiones de GEI en operaciones de transporte, que complementa la ISO 14064 de nivel organizacional. Esta norma es especialmente relevante para los espacios de datos de movilidad porque estandariza la metodología de cálculo de emisiones por pasajero-kilómetro y tonelada-kilómetro, proporcionando una base común para la comparación entre modos de transporte. Un servicio de IA que opere sobre el espacio de datos de movilidad EMDS

podría aplicar automáticamente la ISO 14083 a los datos de actividad de cada operador participante, generando informes de emisiones conformes al estándar sin intervención manual.

El análisis internacional revela que los países líderes en la intersección de espacios de datos, IA y sostenibilidad comparten tres características: inversión pública significativa en infraestructura de datos, colaboración estrecha entre centros tecnológicos y organizaciones ambientales, y existencia de casos de uso demostrador que validan la viabilidad técnica y el impacto real. España, con el Plan de Impulso de 500 millones, la colaboración ITA-ECODES y el caso de uso ECODES-EDMUS, cumple las tres condiciones y se encuentra en posición de competir con Alemania, Francia y los países nórdicos en este ámbito. La principal ventaja comparativa española reside en la combinación de un marco regulatorio avanzado (UNE 0087:2025), financiación disponible (Kit Espacios de Datos) y una estrategia explícita de vincular espacios de datos con objetivos de sostenibilidad en sectores como movilidad, energía y agroalimentación.

# 7. Oportunidades: servicios de IA sobre espacios de datos para reducción de huella de carbono

## 7.1. La tesis: por qué los espacios de datos son el sustrato natural para la IA climática

La filosofía de los espacios de datos soberanos desbloquea posibilidades para la IA climática que serían inalcanzables con arquitecturas centralizadas o bilaterales. En un espacio de datos, cada participante publica activos de datos en un catálogo federado con políticas de acceso granulares. Un servicio de IA puede, previo contrato negociado automáticamente, acceder a datos de múltiples proveedores de una cadena de valor sin que ningún actor deba ceder sus datos a un tercero centralizado. Esto es esencial para el Alcance 3, donde la empresa calculadora necesita datos de centenares de proveedores que, legítimamente, no desean transferir su información a plataformas que no controlan.

Los principios de soberanía del dato y mínima exposición son especialmente compatibles con los requisitos de confidencialidad de los datos de emisiones, que pueden revelar información sensible sobre procesos productivos, eficiencia energética y estrategias de aprovisionamiento. Un servicio de IA operando dentro del espacio de datos puede calcular agregados (por ejemplo, la huella total de una categoría de Alcance 3) sin exponer los datos individuales de cada proveedor al solicitante, implementando así una forma de computación sobre datos soberanos.

## 7.2. Oportunidades identificadas

### Oportunidad 1: predicción automatizada de emisiones de Alcance 3

Desplegar modelos de ML (AdaBoost, XGBoost, Random Forest) como servicios accesibles a través del espacio de datos que, a partir de datos financieros y clasificaciones industriales publicados por los participantes, predigan emisiones de Alcance 3 para las 15 categorías del GHG Protocol. El servicio accedería soberanamente a los datos de cada proveedor a través del EDC, calcularía predicciones específicas y

devolvería los resultados al solicitante sin exponer los datos fuente. La arquitectura propuesta contempla un pipeline de datos que comienza con la ingesta de estados financieros y clasificaciones NACE/SIC publicados como activos en el catálogo del espacio de datos, seguida de la transformación y normalización de variables, la aplicación del modelo entrenado, y la devolución de predicciones con intervalos de confianza. Madurez: TRL 5-6. Papers de referencia: Serafeim y Vélez Caicedo (2022), Wang y Ye (2025), Qu et al. (2023).

## Oportunidad 2: matching inteligente de factores de emisión con LLMs

Utilizar modelos de lenguaje de gran escala (LLMs) como servicio de clasificación y emparejamiento automático de actividades empresariales con factores de emisión de bases de datos estándar (DEFRA, EPA, ecoinvent). El servicio recibiría descripciones textuales de bienes y servicios adquiridos desde el catálogo del espacio de datos y devolvería factores de emisión específicos con intervalos de confianza. El enfoque técnico contempla un LLM fine-tuned con Retrieval-Augmented Generation (RAG) sobre un corpus de factores de emisión indexados vectorialmente, lo que permite al modelo buscar y seleccionar el factor más apropiado considerando sector NACE, geografía, proceso productivo y año de referencia. El dataset ExioML (Nature Scientific Data, 2026) proporciona la base de factores de emisión optimizada para este tipo de aplicaciones.

## Oportunidad 3: optimización multimodal del transporte

Servicios de IA que accedan simultáneamente a datos de transporte público (GTFS), datos de flota privada, datos de tráfico en tiempo real y datos de emisiones de vehículos publicados en el espacio de datos de movilidad, para calcular rutas óptimas que minimicen emisiones de CO<sub>2</sub> en lugar de (o además de) tiempo o coste. La implementación requiere algoritmos de optimización multiobjetivo que encuentren soluciones Pareto-óptimas en el espacio definido por tiempo de viaje, coste económico y emisiones de CO<sub>2</sub>, permitiendo al usuario elegir el equilibrio que prefiera. El caso del deployEMDS en Milán demuestra la viabilidad de este enfoque para escenarios de MaaS orientados a la neutralidad de carbono.

## Oportunidad 4: certificación inteligente de reducciones de emisiones

Como demuestra el caso ECODES-EDMUS, la combinación de acceso soberano a datos de billeteaje, cálculo automatizado de emisiones evitadas y referencia a precios de mercado de carbono (EU ETS) permite crear un mercado de certificados de reducción de emisiones completamente automatizado. La IA puede extender este modelo añadiendo validación automática de la integridad de los datos, detección de anomalías que sugieran fraude o doble conteo, y recomendación de estrategias de compra óptimas para maximizar el impacto ambiental por euro invertido.

## Oportunidad 5: gemelos digitales de cadenas de suministro sostenibles

Construir gemelos digitales de cadenas de suministro que integren datos de emisiones de múltiples participantes del espacio de datos, permitiendo simular el impacto de cambios en proveedores, rutas logísticas, materiales o procesos productivos sobre la huella de carbono total. Los modelos de IA pueden optimizar estas decisiones considerando simultáneamente coste, tiempo, calidad y emisiones, identificando trade-offs y soluciones Pareto-óptimas. La arquitectura de un gemelo digital sobre espacio de datos requiere tres capas: una capa de ingesta de datos que accede soberanamente a activos de múltiples proveedores, una capa de modelización que construye una representación digital de la cadena de suministro con sus flujos físicos y financieros, y una capa de simulación que permite ejecutar escenarios what-if sobre el modelo.

## Oportunidad 6: auditoría automatizada de reportes de Alcance 3

Servicios de IA que accedan a datos publicados en el espacio de datos por empresas que reportan emisiones bajo CSRD, y que verifiquen automáticamente la consistencia de los reportes con los datos fuente, detecten omisiones en categorías materiales, identifiquen outliers estadísticos y generen informes de aseguramiento previos a la auditoría externa. Esta oportunidad es especialmente relevante dado que la CSRD incluirá progresivamente requisitos de aseguramiento limitado y, eventualmente, razonable sobre la información de sostenibilidad.

## Oportunidades adicionales identificadas

Se identifican también las siguientes oportunidades: Oportunidad 7, generación automatizada de Product Carbon Footprint (PCF) siguiendo el modelo Catena-X pero extendido a otros sectores industriales. Oportunidad 8, optimización energética de edificios mediante acceso a datos de consumo en tiempo real publicados en el espacio de datos de energía. Oportunidad 9, trazabilidad de la huella de carbono en cadenas agroalimentarias integrando datos de AgriDataSpace con modelos de emisión agrícola. Oportunidad 10, descarbonización de la aviación mediante el análisis de datos de combustibles alternativos sostenibles (SAF) compartidos en espacios de datos de transporte. Oportunidad 11, predicción de riesgos climáticos en cadenas de suministro combinando datos meteorológicos, geoespaciales y logísticos. Oportunidad 12, plataformas de compensación de huella de carbono basadas en datos verificados de espacios de datos soberanos, evitando el problema del greenwashing mediante trazabilidad completa.

## 7.3. Análisis de viabilidad técnica y económica

La viabilidad técnica de las oportunidades identificadas depende de tres factores principales: la disponibilidad de datos en espacios de datos operativos, la madurez de los modelos de IA requeridos y la existencia de estándares de interoperabilidad. Las oportunidades 1, 2 y 4 presentan la mayor viabilidad inmediata porque pueden operar con datos financieros y de actividad que las empresas ya gestionan internamente, requieren modelos de ML y LLM cuya eficacia ha sido demostrada en la literatura, y pueden desplegarse sobre infraestructura EDC ya probada. Las oportunidades 3, 5 y 8 requieren datos en tiempo real de mayor complejidad (GTFS, telemetría, smart meters) pero se benefician de estándares existentes. Las oportunidades 6, 7 y 9-12 están en fases más tempranas de desarrollo, pero presentan un alto potencial de impacto a medio plazo.

Desde la perspectiva económica, los servicios de IA sobre espacios de datos pueden adoptar tres modelos de negocio. El modelo AlaaS (AI as a Service) cobra por uso a los participantes del espacio de datos que soliciten predicciones o análisis. El modelo de licencia ofrece el servicio como producto a operadores de espacios de datos que lo despliegan para sus participantes. El modelo de certificación, como el implementado en ECODES-EDMUS, genera ingresos a través de la venta de certificados ambientales cuyo valor se referencia a mercados de carbono. Los tres modelos son compatibles con las convocatorias de financiación vigentes, particularmente el Kit Espacios de Datos y la convocatoria de Productos y Servicios.

# 8. Caso de uso real: plataforma ECODES sobre el espacio de datos de movilidad EDMUS

## 8.1. Descripción general del caso de uso

La plataforma ECODES sobre el espacio de datos EDMUS se ha desarrollado dentro del proyecto EDMUS, pero con el conocimiento y conceptualización y cálculo de ECODES. Se trata de un demostrador de un caso de uso real y operativo que demuestra la viabilidad de aplicar servicios de inteligencia artificial sobre un espacio de datos soberano para la certificación y monetización de reducciones de emisiones de CO<sub>2</sub>. Desarrollada conjuntamente por ITA con soporte de Fundación ECODES, la plataforma opera sobre el espacio de datos de movilidad EDMUS (Espacio de Datos de Movilidad Urbana Sostenible) y certifica las reducciones de emisiones generadas por la electrificación de flotas de autobuses. Se trata de un piloto operativo y funcional con datos reales.

El concepto central es que cada kilómetro recorrido por un autobús eléctrico que anteriormente era diésel evita una emisión de CO<sub>2</sub> cuantificable. ECODES actúa como intermediario entre el proveedor de datos de movilidad, el comprador de certificados (empresas, municipios o ciudadanos que desean compensar su huella de carbono) y el mercado europeo de carbono (EU ETS), que proporciona el precio de referencia del CO<sub>2</sub>.

## 8.2. Arquitectura del espacio de datos

La arquitectura del sistema implementa fielmente el paradigma de espacio de datos soberano. El conector OperadorDeMovilidad actúa como Data Space Provider (DSP), publicando en su catálogo los activos de datos disponibles con sus políticas de acceso. El conector ECODES actúa como Data Space Consumer (DSC), descubriendo activos, negociando contratos y recibiendo referencias de acceso temporal. Los datos de billeteaje del OperadorDeMovilidad mantienen su dominio: ECODES los accede puntualmente a través del espacio de datos solo cuando necesita calcular un certificado, y el acceso se realiza mediante un Endpoint Data Reference (EDR) con token de autenticación de corta duración.

Los activos de datos publicados por el Operador de Movilidad se estructuran en tres niveles de granularidad. El activo billeteaje-hdc-count proporciona una instantánea de disponibilidad total (tickets disponibles, horas disponibles). El activo billeteaje-hdc-list ofrece un desglose horario con estimaciones de kg de CO<sub>2</sub> por franja. Los activos billeteaje-hdc-FECHA-HORA representan bloques horarios individuales con los tickets

específicos (ID, línea, timestamp). Cada bloque horario es un activo independiente negociado por separado, lo que permite compras granulares: ECODES adquiere exactamente los bloques necesarios para alcanzar la cantidad de CO<sub>2</sub> solicitada por el comprador.

## 8.3. Metodología de cálculo de huella de carbono

La metodología de cálculo parte de un hecho contrastable: Los operadores de movilidad han electrificado su flota de autobuses urbanos. La huella certificada es la diferencia entre lo que habría emitido la flota si siguiera siendo diésel (escenario de referencia) y lo que emite ahora siendo eléctrica. La fórmula central es  $\text{km recorridos Diesel} * \text{factor emisiones autobús diésel} - \text{km recorridos eléctricos} * \text{factor de emisión km recorrido eléctrico}$ . Los factores son editables en tiempo real desde el panel de gestión. Este ahorro por km se reparte entre los tickets de viajeros que viajan en el autobús que es la unidad que sirve para certificar la reducción de huella de carbono en base a un número de viajeros. Se está en proceso de capturar el dato de distancia recorrida para mejorar la exactitud del sistema.

La metodología distingue explícitamente entre emisiones evitadas y emisiones reducidas, una distinción crítica en el contexto de los mercados de carbono. Las emisiones evitadas son aquellas que habrían ocurrido en un escenario contrafactual (si la flota siguiera siendo diésel) pero que no ocurren gracias a la intervención (electrificación). Esta distinción es importante porque el GHG Protocol no permite contabilizar las emisiones evitadas como reducciones reales en el inventario de Alcance 1 operadores de movilidad (que ya son bajas por ser flota eléctrica), sino que las trata como un beneficio ambiental adicional que puede monetizarse a través de mercados voluntarios o mecanismos como los Certificados de Ahorro de Emisiones. La plataforma ECODES sobre el espacio de datos EDMUS adopta esta interpretación, posicionando los certificados como instrumentos de compensación voluntaria referenciados al precio del EU ETS, no como créditos de carbono regulados.

La auditabilidad es total: cada certificado almacena en su JSON los factores exactos utilizados, la fórmula aplicada, los tickets individuales por hora y el cálculo intermedio por bloque. Cualquier tercero puede verificar la cantidad certificada accediendo al endpoint público `/verificar/cert_id` sin autenticación, lo que implementa un principio de transparencia radical que elimina la necesidad de confiar en ECODES como autoridad central.

## 8.4. El protocolo EDC en acción: negociación automatizada

El proceso de compra de certificados desencadena un ciclo completo de negociación EDC de cinco fases para cada bloque horario necesario. La fase de catálogo consulta la

disponibilidad del activo. La fase de negociación de contrato progresa a través de los estados REQUESTED, OFFERED, ACCEPTED, AGREED y FINALIZED, con polling cada 2 segundos y un máximo de 40 intentos (80 segundos de espera máxima). La fase de transferencia inicia un HttpData-PULL. La fase de obtención del EDR proporciona la URL y el token de acceso temporal. Finalmente, la fase de descarga de datos obtiene los tickets reales del bloque horario.

Este ciclo se repite para cada bloque horario hasta cubrir la cuota de CO<sub>2</sub> solicitada, consumiendo los bloques cronológicamente. El sistema implementa una caché de acuerdos en memoria que reutiliza agreement\_ids ya negociados dentro de la misma sesión de compra, evitando ciclos de negociación redundantes. El proceso puede durar varios minutos, por lo que ECODES implementa un endpoint de streaming NDJSON que emite eventos progresivos visualizados en tiempo real por la interfaz web del marketplace.

## 8.5. Integración con el mercado de carbono EU ETS

El precio de cada tonelada de CO<sub>2</sub> certificada se obtiene del mercado europeo de carbono (EU ETS), instrumento EUA (European Union Allowance). ECODES implementa un sistema de cuatro fuentes con fallback progresivo: web scraping de Investing.com, API de Yahoo Finance (tickers ECO2.F, C02Y.F, FEUA.F), exportación CSV de Stooq.com, y la librería yfinance como último recurso. El precio se cachea durante una hora para no sobrecargar las fuentes externas. El modelo de reparto de ingresos asigna el 25% del valor del certificado a ECODES (plataforma de certificación y canal de mercado) y el 75% al operador de movilidad (propietaria de la reducción real de emisiones). Una vez asignado un precio a un certificado en el momento de su venta, queda congelado en el archivo economic\_prices.json, proporcionando seguridad jurídica y contable al comprador.

## 8.6. Arquitectura de microservicios

La solución se compone de tres microservicios independientes contenerizados con Docker y Docker Compose. La API de negocio (FastAPI + Uvicorn, puerto 8700) gestiona todos los endpoints de negocio, la comunicación con el conector EDC mediante httpx, y la generación de PDFs con ReportLab y Matplotlib. El marketplace web (Flask, puerto 8850) proporciona la interfaz de usuario para el flujo de compra, con streaming de progreso mediante Server-Sent Events / NDJSON y generación de códigos QR para los certificados. El dashboard analítico (Streamlit, puerto 8800) ofrece KPIs del portfolio, tablas de certificados, evolución mensual de ingresos, previsión de potencial y generación de informes PDF ejecutivos.

## 8.7. Relevancia para el Alcance 3

El caso ECODES-EDMUS es directamente relevante para la problemática del Alcance 3 en múltiples dimensiones. Para la empresa compradora del certificado, la reducción certificada puede contabilizarse como una compensación de sus emisiones indirectas de Alcance 3, particularmente en la categoría 6 (viajes de negocios) o la categoría 7 (desplazamiento de empleados) si los empleados utilizan el transporte público electrificado. Para el operador de movilidad, la monetización del valor ambiental de su electrificación genera ingresos adicionales que pueden financiar la ampliación de la flota eléctrica, creando un círculo virtuoso de descarbonización. Y para el ecosistema, la metodología auditable y el acceso soberano a los datos demuestran que es posible certificar reducciones de emisiones con trazabilidad completa, sin depender de estimaciones genéricas o autodeclaraciones no verificables.

El modelo es replicable a cualquier operador de transporte público que haya electrificado su flota, y extensible a otros ámbitos donde la electrificación de operaciones (logística, construcción, agricultura) genere reducciones cuantificables de emisiones. La clave es que el espacio de datos soberano proporciona el mecanismo de acceso a los datos de actividad real (no estimados) necesarios para la certificación, y la IA puede ampliar el modelo con predicción de disponibilidad, optimización de precios y detección de fraude.

## 8.8. Lecciones aprendidas del despliegue

El desarrollo y despliegue de la plataforma ECODES-EDMUS ha generado un conjunto valioso de lecciones aprendidas aplicables a futuros proyectos de servicios de IA sobre espacios de datos. En primer lugar, la latencia del protocolo EDC: cada ciclo de negociación completo (catálogo, negociación, transferencia, EDR, descarga) requiere entre 10 y 20 segundos, lo que hace inviable un modelo de acceso a datos en tiempo real para aplicaciones interactivas. La solución implementada —streaming NDJSON con eventos progresivos— proporciona feedback continuo al usuario y gestiona eficazmente la expectativa de espera. Para servicios de predicción batch, esta latencia es aceptable, pero los servicios de optimización en tiempo real requerirán mecanismos de pre-carga y caché más sofisticados.

En segundo lugar, la gestión de identidades y permisos: la negociación de contratos EDC requiere que ambas partes se identifiquen mutuamente con certificados X.509 o credenciales verificables, lo que introduce una complejidad administrativa significativa para el onboarding de nuevos participantes. La experiencia con EDMUS sugiere que un servicio de Connector-as-a-Service (CaaS), como el ofrecido por el Mobility Data Space alemán, es esencial para reducir las barreras de entrada de pymes y organizaciones con capacidades tecnológicas limitadas.

En tercer lugar, la evolución del catálogo de activos: a medida que el espacio de datos crece en participantes y activos, la gestión del catálogo se convierte en un cuello de botella. Los activos granulares por bloque horario (billete-hdc-FECHA-HORA) generan

un volumen elevado de entradas en el catálogo que deben podarse periódicamente. Un servicio de IA podría automatizar la gestión del catálogo, sugiriendo niveles de agregación óptimos, identificando activos obsoletos y optimizando la estructura de metadatos para facilitar el descubrimiento por parte de los consumidores.

## 8.9. Métricas de rendimiento y KPI operativos

La plataforma ECODES-EDMUS registra un conjunto de métricas de rendimiento que permiten evaluar tanto la eficiencia técnica como el impacto ambiental del servicio. Desde la perspectiva técnica, el tiempo medio de generación de un certificado (desde la solicitud del usuario hasta la entrega del PDF) oscila entre 2 y 8 minutos, dependiendo del número de bloques horarios necesarios para cubrir la cantidad de CO<sub>2</sub> solicitada. La tasa de éxito de las negociaciones EDC supera el 95%, con fallos atribuibles principalmente a timeout en la fase de negociación de contrato (estado AGREED no alcanzado en los 80 segundos de espera máxima). El sistema soporta múltiples compras simultáneas gracias a la naturaleza asíncrona de FastAPI y la gestión concurrente de sesiones de negociación.

## 8.10. Potencial de escalado y replicabilidad

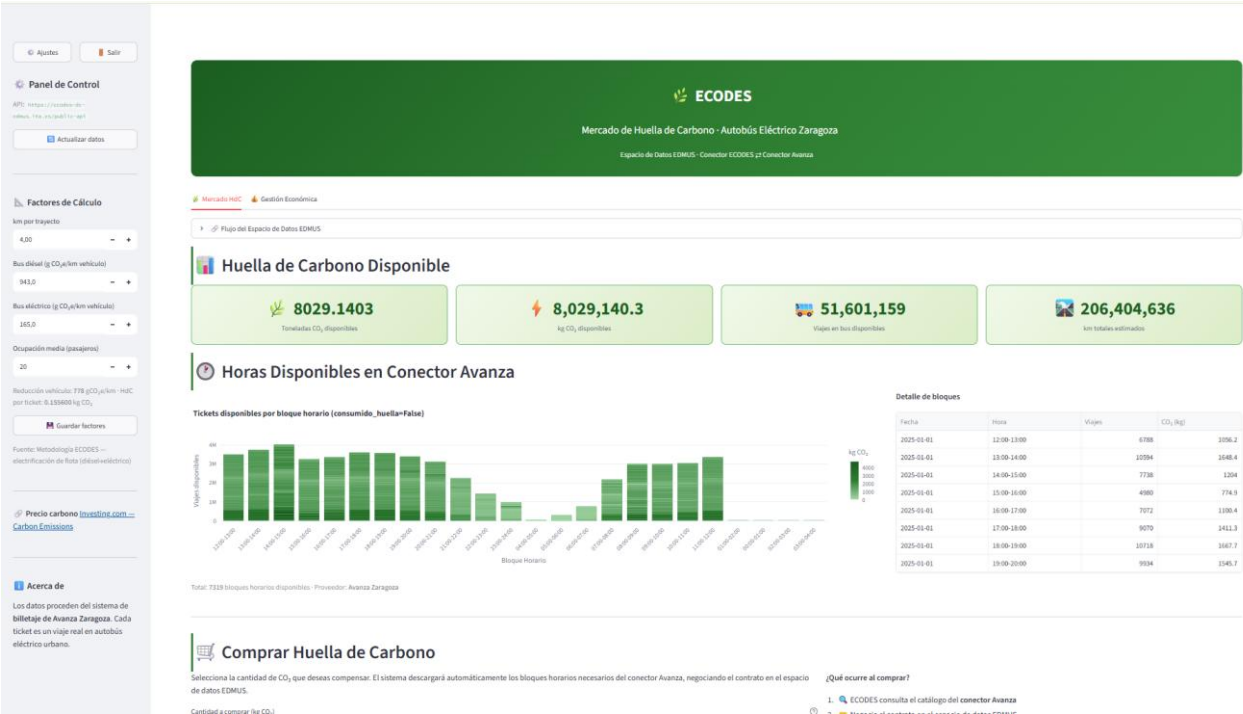
El modelo ECODES-EDMUS ha sido diseñado desde su concepción para ser replicable a otros contextos y operadores. La arquitectura de microservicios contenerizados permite desplegar una instancia completa de la plataforma en cualquier infraestructura Docker-compatible en cuestión de horas, y la parametrización de los factores de emisión y la distancia media por billete permite adaptar la metodología de cálculo a cualquier operador de transporte público sin modificación del código.

La extensión del modelo a otros sectores requiere adaptaciones más sustanciales, pero es técnicamente viable. En el sector de la logística de última milla, empresas como SEUR, Correos o Amazon Logistics que están electrificando sus flotas de reparto podrían certificar las emisiones evitadas por cada paquete entregado con vehículo eléctrico frente al escenario de referencia diésel. En el sector energético, las comunidades energéticas locales podrían certificar las emisiones evitadas por la generación de energía renovable distribuida frente al mix eléctrico convencional. En la construcción, la sustitución de maquinaria diésel por eléctrica en obras podría generar certificados de reducción de emisiones de Alcance 1 monetizables.

El potencial de escalado económico del modelo depende fundamentalmente del precio del carbono en el EU ETS.

Desde la perspectiva ambiental, la plataforma proporciona KPI en su dashboard Streamlit que incluyen: toneladas totales de CO<sub>2</sub> certificadas, ingresos totales y por certificado, precio medio del CO<sub>2</sub> (referenciado al EU ETS), distribución horaria de la demanda de

certificados, evolución mensual del portfolio, y una previsión de potencial de certificación basada en la capacidad de la flota eléctrica del operador de movilidad. Estos KPI permiten a ECODES y al operador de movilidad evaluar el rendimiento del servicio y planificar la escalabilidad, y constituyen la base para los informes periódicos dirigidos a la dirección y a los organismos de financiación.



Web de control de certificados CAES con huella disponible para certificar.

Solicitar certificado CAES

Nombre completo \*

Ana García López

Organización

Empresa u organización

Correo electrónico \*

ana@empresa.com

Cantidad habitual:

50 kg

100 kg

500 kg

1 t

5 t

Cantidad de CO<sub>2</sub>

100

kg CO<sub>2</sub>

7,44 €

0,0744 €/kg (74,38 €/t) · Investing.com · EU Carbon ETS

5.0 árboles/año

1656 km en bus

0.100 t CO<sub>2</sub>

Obtener certificado

Pago seguro · Descarga inmediata · Número de certificado único

Pantalla de compra de certificado



**Pasarela de pago segura**  
ECODES Payment Gateway · SSL 256 bits



100 kg CO<sub>2</sub> · Certificado CAES

**7,44 €**

1. **TARJETA**

2. AUTORIZANDO

3. CONFIRMADO

 Entorno de pruebas — los datos no se procesan realmente

VISA

MASTERCARD

AMEX

 SSL



1 2 3 4 1 2 3 4 1 2 3 4 1 2 3 4

ADAN PIÑERO

25/36

Número de tarjeta

1234 1234 1234 1234

Nombre en la tarjeta

ADAN PIÑERO

Caducidad

25/36

CVV

123|

 **Pagar 7,44 €**

Pago cifrado · No almacenamos datos de tarjeta

Pago para la compra del certificado

**Proceso completado**

- > Consultando horas disponibles del conector Avanza vía Conector ECODES (65000)...
- > Horas disponibles: 7322 bloques
- > [1/2] Descargando → 2025-01-01 09:00
- > 2,924 tickets · 706.438 kg CO<sub>2</sub> · acumulado: 706.438/100.0 kg
- > Generando certificado CAES: 807c8756-4062-46...
- > Certificado registrado en /app/certs/807c8756-4062-46e2-a84c-e70c186ccc54.json
- > Generando PDF del certificado CAES...
- > PDF generado: CAES\_HdC\_807c8756.pdf
- ✓ **Compra completada: 706.438 kg CO<sub>2</sub> · 2,924 viajes · cert 807c8756-406...**

Negociación dentro del espacio de datos entre ECODES y OPERADOR

### Certificado CAES emitido

adan, tu certificado CAES ha sido emitido y firmado correctamente.

|                                             |                                        |                                |
|---------------------------------------------|----------------------------------------|--------------------------------|
| <b>706,4</b><br>kg CO <sub>2</sub> evitados | <b>2924</b><br>viajes en bus eléctrico | <b>11.696</b><br>km computados |
|---------------------------------------------|----------------------------------------|--------------------------------|

**Descargar certificado PDF**

Obtener otro certificado

Emisión del certificado

# 9. Catálogo de oportunidades sobre espacios de datos españoles

## 9.1. De las tipologías genéricas al catálogo aplicado

El capítulo 7 ha identificado doce tipologías genéricas de servicios de IA sobre espacios de datos para la reducción de huella de carbono, y el capítulo 8 ha documentado en profundidad un caso de uso real, ECODES-EDMUS, que materializa una de esas tipologías (la certificación inteligente de reducciones) sobre un espacio de datos de movilidad concreto. El presente capítulo cierra el arco lógico del informe: instancia las tipologías genéricas sobre el ecosistema real de espacios de datos españoles operativos o en despliegue, generando un catálogo de quince oportunidades concretas, cada una anclada a un promotor identificado, un sustrato técnico verificable y una vía de financiación explícita.

La pertinencia de este ejercicio se refuerza con los datos publicados por el Centro de Referencia de Espacios de Datos (CRED) en enero de 2026: España cuenta ya con 165 espacios de datos elegibles para el Kit Espacios de Datos, distribuidos en 16 sectores diferentes, de los cuales 35 están operativos. Este volumen, combinado con la activación de la Lista de Confianza de Espacios de Datos regulada por la Orden TDF/1207/2025 (basada en la Especificación UNE 0087:2025), configura un sustrato suficiente para el despliegue inmediato de servicios de IA orientados a la sostenibilidad. Las oportunidades que siguen no son ejercicios prospectivos: son proyectos viables sobre infraestructura existente o financiada.

El catálogo se organiza en quince fichas con estructura homogénea. Cada ficha identifica el espacio de datos receptor (con su promotor y estado), describe el servicio de IA propuesto, especifica el problema de huella de carbono que resuelve, justifica la oportunidad desde la perspectiva técnica, regulatoria y de datos disponibles, explicita su encaje con el caso de uso ECODES-EDMUS documentado en el capítulo 8, e indica la vía de financiación más adecuada y el nivel de madurez tecnológica (TRL). Las quince oportunidades se numeran de O1 a O15 para evitar confusión con la numeración de las doce tipologías genéricas del capítulo 7 (Oportunidad 1 a 12).

## 9.2. Tabla de espacios de datos españoles utilizados como sustrato

A continuación, se presenta el conjunto de espacios de datos sobre los que se construye el catálogo. La tabla incluye únicamente espacios identificados en el Catálogo de Espacios de Datos Elegibles del CRED a fecha de enero de 2026 o, en su caso, espacios financiados por las convocatorias del Plan de Impulso o por programas europeos con presencia operativa en España.

| Espacio de datos                   | Promotor / Líder                             | Sector                    | Estado (abril 2026)                        | Aplicabilidad a huella de carbono                                                  |
|------------------------------------|----------------------------------------------|---------------------------|--------------------------------------------|------------------------------------------------------------------------------------|
| EDMUS                              | Avanza                                       | Movilidad urbana          | Operativo (caso ECODES)                    | Directa, ya implementado                                                           |
| DS4M Mediterráneo                  | INTRAS-Universitat de València, ITS España   | Movilidad regional        | En despliegue (4,6 M€ Plan de Impulso)     | Alta — ZBE, PMUS, tráfico                                                          |
| DS4M CENTRAL                       | ITS España                                   | Movilidad nacional        | En despliegue                              | Alta — RAPs/PAPs/LAPs                                                              |
| DS4MM                              | Junta de Extremadura y consorcio europeo     | Movilidad transfronteriza | En despliegue (3 M€ CEF, hasta junio 2027) | Alta — análisis de emisiones, ZBE                                                  |
| Espacio de Datos de Turismo        | SEGITTUR / Ministerio de Industria y Turismo | Turismo                   | Operativo (diciembre 2025)                 | Media-alta                                                                         |
| Espacio Nacional de Datos de Salud | Gobierno de España                           | Salud                     | Activado (enero 2026)                      | Media — sector responsable del 4,4% emisiones globales                             |
| AgrospAI                           | Universitat de Lleida                        | Agroalimentario           | Operativo (Gaia-X Lighthouse)              | Alta — emisiones agrícolas                                                         |
| Agrixels (UPCxels)                 | Universitat Politècnica de Catalunya / IDEAI | Agroalimentario           | Operativo                                  | Alta — incluye explícitamente mercados de carbono y trazabilidad de sostenibilidad |

| Espacio de datos          | Promotor / Líder                          | Sector                      | Estado (abril 2026)                         | Aplicabilidad a huella de carbono                                |
|---------------------------|-------------------------------------------|-----------------------------|---------------------------------------------|------------------------------------------------------------------|
| GroundX                   | Sector agroalimentario                    | Agroalimentario             | Elegible                                    | Media-alta                                                       |
| CitriData                 | Sector citrícola andaluz                  | Agroalimentario             | Elegible (presentado segundo semestre 2025) | Media                                                            |
| COLDS                     | Arctic Sea S.L.                           | Transporte de mercancías    | Elegible                                    | Alta — categoría 4 GHG Protocol                                  |
| AECTech                   | Sector viario                             | Infraestructura y movilidad | Elegible                                    | Alta — gestión de tráfico interurbano                            |
| GreenBIMxels              | Universidades y administraciones públicas | Edificación pública         | Elegible                                    | Muy alta — explícitamente sostenibilidad y eficiencia energética |
| CEMENGAL                  | Sector cementero                          | Industria                   | Elegible                                    | Muy alta — sector intensivo en emisiones                         |
| Espacio de Datos de Sacyr | Sacyr                                     | Agua e infraestructuras     | Operativo (incluye módulo IA para ICA)      | Indirecta                                                        |
| Hábitat Data (PRECEDENT)  | Sector hábitat                            | Industria del mueble        | Elegible                                    | Media — Declaraciones Ambientales de Producto                    |

Esta tabla es deliberadamente sintética: los detalles de promotor, casos de uso declarados y arquitectura técnica de cada espacio se desarrollan en las fichas individuales que siguen, siempre que sean determinantes para justificar la oportunidad concreta.

### 9.3. Mapeo entre tipologías genéricas (capítulo 7) y oportunidades aplicadas (capítulo 9)

Antes de presentar las fichas, conviene establecer la correspondencia entre las doce tipologías identificadas en el capítulo 7 y las quince oportunidades aplicadas que se desarrollan en este capítulo. La correspondencia no es uno a uno: algunas tipologías se instancian en varias oportunidades concretas (la certificación inteligente, por ejemplo, se aplica tanto al transporte de pasajeros como a la gestión de infraestructura viaria), y algunas oportunidades aplicadas combinan dos o más tipologías (la auditoría CSRD sanitaria combina matching LLM con generación automatizada de reportes).

| Tipología genérica (cap. 7)                  | Oportunidades aplicadas (cap. 9)  |
|----------------------------------------------|-----------------------------------|
| 1. Predicción Alcance 3 con ML               | O2, O9, O11                       |
| 2. Matching LLM factores emisión             | O3, O7, O12                       |
| 3. Optimización multimodal transporte        | O4, O8                            |
| 4. Certificación inteligente CO <sub>2</sub> | O1, O10                           |
| 5. Gemelos digitales supply chain            | O8                                |
| 6. Auditoría automatizada CSRD               | O7, O15                           |
| 7. Product Carbon Footprint                  | O6, O9, O12                       |
| 8. Optimización energética edificios         | O5, O13                           |
| 9. Trazabilidad agroalimentaria              | O3, O9                            |
| 10. Descarbonización aviación                | (cubierta parcialmente por O2)    |
| 11. Predicción riesgos climáticos            | (no instanciada en este catálogo) |
| 12. Compensación verificada CO <sub>2</sub>  | O1, O10, O11                      |

A esta correspondencia se añade O14 (onboarding asistido por IA), que no responde a una tipología específica del capítulo 7 sino a una palanca transversal del ecosistema, y

que aborda directamente el riesgo de mercado identificado en la sección de mitigaciones del capítulo siguiente.

## 9.4. Fichas de oportunidades aplicadas

### O1. Replicación del modelo de certificación CAES sobre DS4M Mediterráneo y DS4M CENTRAL

**Espacio de datos:** DS4M Mediterráneo, liderado por el Instituto Universitario de Investigación en Tráfico y Seguridad Vial (INTRAS) de la Universitat de València, con la colaboración de IRTIC y el apoyo de ITS España, financiado por el Plan de Impulso con 4,6 millones de euros. Cubre Andalucía, Murcia, Comunitat Valenciana, Cataluña y Aragón. DS4M CENTRAL, liderado por ITS España, despliega Puntos de Acceso Regional, Provincial y Local (RAPs, PAPs, LAPs) interconectados con los Puntos de Acceso Nacional (NAPs) promovidos por la Comisión Europea.

**Servicio de IA propuesto:** portar la arquitectura ECODES-EDMUS documentada en el capítulo 8 (microservicios contenerizados, conector EDC, integración con EU ETS) a los Puntos de Acceso Regional. El servicio AlaaS de certificación de emisiones evitadas se ofrece a cualquier operador de transporte público que comparta datos de billeteaje a través del RAP correspondiente. Se añade un modelo predictivo basado en redes LSTM que estima la disponibilidad de CO<sub>2</sub> certificable por franja horaria, lo que permite optimizar la liquidación con compradores recurrentes y ofrecer contratos a plazo.

**Problema que resuelve:** el caso ECODES-EDMUS está hoy desarrollado como iloto con datos reales, lo que limita su escalabilidad. El RAP actúa como agregador soberano de operadores: un servicio único puede certificar reducciones de toda la flota electrificada de un corredor regional sin requerir negociaciones bilaterales operador a operador.

**Justificación:** DS4M Mediterráneo declara explícitamente entre sus tres casos de uso prioritarios el seguimiento y cálculo de Zonas de Bajas Emisiones, incluyendo medición, modelización y reporte. El modelo de pago por viaje validado en flota electrificada encaja directamente con esta especificación. Aragón está incluido en el ámbito territorial del proyecto, lo que facilita el liderazgo técnico de ITA y la continuidad operativa con el caso EDMUS.

**Encaje con caso ECODES-EDMUS:** extensión directa del caso. Toda la pila técnica documentada en el capítulo 8 y el anexo B es reutilizable; los cambios se concentran en la capa de gobernanza multi-operador y en la gestión de identidades en el RAP.

**Financiación:** Kit Espacios de Datos en modalidad proveedor (hasta 30.000 euros para entidades privadas, hasta 50.000 euros para administraciones públicas) para cada operador que se incorpore. El servicio AlaaS encaja en la convocatoria de Productos y Servicios de Espacios de Datos (44 millones de euros). TRL 7-8.

**Impacto estimado:** considerando los 44.000 autobuses urbanos en España y una tasa de electrificación proyectada del 30% en 2030 (estimación del capítulo 8.10), el techo del

mercado nacional de Certificados de Ahorro de Emisiones en transporte público se sitúa en 50-100 millones de euros anuales.

## O2. Predicción ML de Alcance 3 sobre el Espacio de Datos de Turismo

**Espacio de datos:** Espacio de Datos de Turismo, promovido por el Ministerio de Industria y Turismo a través de SEGITTUR. Activado en diciembre de 2025, constituye el primer espacio de datos sectorial operativo de España, con empresas piloto y early adopters consolidando su catálogo. Está vinculado a las iniciativas europeas Tourism Data Space (liderada por la empresa española Forwardkeys) y DATES (liderada por Anysolution).

**Servicio de IA propuesto:** modelo de ensemble (XGBoost o AdaBoost) que, a partir de datos de gasto turístico, perfil del cliente, distancia origen-destino, modo de transporte y categoría de alojamiento publicados en el catálogo del espacio, predice la huella de Alcance 3 de un paquete turístico completo, desglosada por categorías relevantes (1-bienes y servicios adquiridos, 6-viajes de negocios cuando aplique, 9-distribución aguas abajo). La inferencia se ofrece vía endpoint accesible mediante el protocolo DSP, sin extracción de datos primarios del espacio.

**Problema que resuelve:** el turismo en España representó el 12,4% del PIB en 2025, con 96,8 millones de turistas internacionales y un gasto turístico total superior a 134.712 millones de euros. Las emisiones de Alcance 3 de un paquete turístico se estiman hoy con factores genéricos por touroperador, lo que infravalora la heterogeneidad real entre paquetes. El espacio de datos permite por primera vez predicciones específicas con datos primarios de proveedores reales (alojamientos, transporte, restauración, actividades).

**Justificación:** alineamiento explícito con la posición pública del Ministerio de Industria y Turismo, que presenta este espacio como herramienta para una gestión turística más eficiente y sostenible. El servicio responde directamente a los requisitos del estándar ESRS E1 para los grandes operadores turísticos sujetos a la CSRD desde enero de 2025. Reproduce sobre datos turísticos el método predictivo validado por Serafeim y Vélez Caicedo (2022) y Wang y Ye (2025), referenciados en el capítulo 5 del presente informe.

**Encaje con caso ECODES-EDMUS:** complementario. Donde EDMUS certifica emisiones evitadas en transporte público, este servicio calcula la huella positiva de toda la cadena turística, que el turista o el touroperador pueden compensar adquiriendo Certificados de Ahorro de Emisiones en EDMUS. Se cierra así el ciclo "calcular, reducir, compensar con datos verificados" entre dos espacios de datos sectoriales distintos.

**Financiación:** Kit Espacios de Datos en modalidad proveedor, complementado por Horizon Europe Cluster 5 (Climate, Energy and Mobility). TRL 5-6.

## O3. Servicio LLM de matching de factores de emisión sobre AgrospAI y Agrixels

**Espacio de datos:** AgrospAI, demostrador del sector agroalimentario operado por la Universitat de Lleida, con acreditación europea Gaia-X Lighthouse Project, integrado en el despliegue del futuro Common European Agricultural Data Space (CEADS). Agrixels, ecosistema de datos agroalimentarios dentro del espacio UPCxels promovido por la Universitat Politècnica de Catalunya y su instituto IDEAI, que define un estándar de datos y ontología agroalimentaria con cobertura explícita de cultivos, sostenibilidad, trazabilidad, etiquetado alimentario y mercados de carbono.

**Servicio de IA propuesto:** modelo de lenguaje de gran escala fine-tuned con técnica de Retrieval-Augmented Generation (RAG) sobre un corpus indexado de factores de emisión agrícolas (ecoinvent, Agribalyse, ADEME Base Carbone, IPCC Tier 1/2/3). El servicio recibe la descripción semántica de un input agrícola desde el catálogo Agrixels (cultivo, geografía, tipo de fertilizante, modo de riego, prácticas de manejo) y devuelve el factor de emisión más preciso disponible junto con su intervalo de confianza. La ontología Agrixels actúa como esquema canónico, eliminando la heterogeneidad semántica que hoy bloquea el matching automático.

**Problema que resuelve:** las emisiones de Alcance 3 en el sector agroalimentario pueden alcanzar el 90% del total y se concentran en la categoría 1, dominada por insumos agrícolas, fertilizantes nitrogenados y emisiones ganaderas. La fragmentación extrema del sector, con miles de pequeños productores, hace inviable la recolección manual de datos primarios; un servicio que automatice el matching es prerequisite para que los grandes compradores agroalimentarios cumplan con la CSRD sin recurrir exclusivamente a estimaciones por gasto.

**Justificación:** AgrospAI ofrece capacidad Compute-to-Data mediante la infraestructura Pontus-X y simula economías de datos con e-Euro y contratos inteligentes, lo que constituye una infraestructura técnica idónea para que el servicio LLM se ejecute sin extraer datos del proveedor. Agrixels declara explícitamente "mercados de carbono" como uno de sus dominios de aplicación, lo que confirma la afinidad estratégica del servicio.

**Encaje con caso ECODES-EDMUS:** extiende la lógica de acceso soberano vía conectores a la cadena agroalimentaria, aunque la pila técnica subyacente es Pontus-X en lugar de Eclipse Dataspace Connector. Se requiere desarrollar un adaptador semántico DSP-Pontus-X, que constituye en sí mismo una contribución técnica relevante para la interoperabilidad entre espacios de datos europeos.

**Financiación:** Kit Espacios de Datos, complementado por Horizon Europe Cluster 6 (Food, Bioeconomy, Natural Resources). TRL 4-5. Aliado natural: BBLTranslation, ya integrado en Agrixels para etiquetado alimentario multilingüe.

## O4. Optimización multimodal carbon-aware sobre el agregado EDMUS, DS4M y COLDS

**Espacio de datos:** combinación federada entre EDMUS (transporte de pasajeros), DS4M Mediterráneo y DS4M CENTRAL (datos regionales y locales de movilidad a través de los RAPs) y COLDS (espacio de datos de transporte de mercancías promovido por Arctic Sea S.L., diseñado para "compartición de datos entre actores del transporte de mercancías que facilita el intercambio seguro de información, e integración con servicios que optimizan la eficiencia y la seguridad").

**Servicio de IA propuesto:** optimizador multiobjetivo basado en algoritmos evolutivos (NSGA-II o equivalentes) con un agente de Reinforcement Learning de capa superior. Accede soberanamente vía DSP a datos de capacidad de transporte público, telemetría de vehículos eléctricos de flota, intensidad de carbono horaria de la red eléctrica española y disponibilidad de huecos de carga en COLDS. Calcula rutas Pareto-óptimas en el espacio definido por tiempo de viaje, coste económico y emisiones de CO<sub>2</sub> equivalente, devolviendo al solicitante el conjunto de soluciones no dominadas para que escoja según su función de utilidad.

**Problema que resuelve:** cubre simultáneamente la categoría 4 del Alcance 3 (transporte y distribución aguas arriba) para los cargadores y la huella propia (Alcance 1 y 2) para los operadores logísticos. Aborda directamente la tipología 3 del capítulo 7.

**Justificación:** COLDS está diseñado precisamente para optimizar la eficiencia y seguridad del transporte de mercancías mediante intercambio seguro de información. DS4M Mediterráneo declara como objetivo el fomento de "nuevos servicios y soluciones basadas en el uso y análisis de datos compartidos". La Plataforma Logística de Zaragoza (PLAZA), citada en la sección 4.8 del presente informe con sus 13,2 millones de metros cuadrados y más de 400 empresas instaladas, constituye el caso de uso piloto natural por su volumen y por la presencia territorial de ITA.

**Encaje con caso ECODES-EDMUS:** utiliza EDMUS como una de las múltiples fuentes federadas. Demuestra el valor de la interoperabilidad entre conectores EDC desplegados en espacios de datos sectoriales distintos, lo que constituye un avance significativo respecto al modelo actual de EDMUS, restringido al dominio de un único operador.

**Financiación:** 2.<sup>a</sup> Convocatoria de Demostradores y Casos de Uso (proyectos en ejecución hasta el 30 de junio de 2026, hito CID 180), complementada por Horizon Europe Cluster 5 en sucesores de deployEMDS. TRL 6-7.

## O5. Servicio de IA sobre GreenBIMxels para Product Carbon Footprint de edificios públicos y Alcance 3 categoría 2

**Espacio de datos:** GreenBIMxels, espacio de datos centrado explícitamente en la digitalización, sostenibilidad y eficiencia energética de los equipamientos públicos. Su objetivo declarado es definir un estándar BIM interoperable y enriquecido con datos

energéticos, ambientales y de uso, con el fin de "desarrollar servicios de inteligencia artificial orientados a la toma de decisiones, el mantenimiento predictivo y la mejora de la gestión de edificios públicos". Promueve la colaboración entre administraciones públicas, agentes tecnológicos y universidades.

**Servicio de IA propuesto:** pipeline que (i) ingesta modelos BIM en formato IFC enriquecidos publicados en el catálogo del espacio, (ii) utiliza un LLM con RAG semántico para mapear cada elemento constructivo a una Declaración Ambiental de Producto (DAP/EPD) o a su ficha equivalente en bases como BEDEC e InData, (iii) calcula el carbono incorporado total del edificio, (iv) lo combina con datos de consumo energético en uso procedentes de teledatada del edificio para obtener la huella ciclo de vida completa A1-A5 más B6 conforme a la EN 15978, y (v) sugiere variantes constructivas de menor huella manteniendo presupuesto y prestaciones funcionales.

**Problema que resuelve:** el sector de la construcción es responsable del 37% de las emisiones globales de CO<sub>2</sub> según la Global Alliance for Buildings and Construction (cifra del capítulo 2.6). Para la administración pública sujeta a la categoría 2 del Alcance 3 (bienes de capital), la falta de datos por proyecto bloquea el reporte. Para licitaciones públicas verdes, el cálculo automatizado del carbono incorporado permite incorporar criterios objetivos de adjudicación con base científica.

**Justificación:** el propio promotor del espacio identifica explícitamente la inteligencia artificial como vector de valor en su descripción oficial, lo que asegura el encaje cultural y estratégico. GreenBIMxels es uno de los pocos espacios elegibles que declara la sostenibilidad y la eficiencia energética como misión nuclear. El servicio encaja con el Real Decreto 390/2021 sobre eficiencia energética en compras públicas y con la Directiva EPBD 2024/1275 (Energy Performance of Buildings Directive recast).

**Encaje con caso ECODES-EDMUS:** fuera del dominio de transporte, pero comparte la filosofía de acceso soberano más AlaaS. ECODES puede actuar como certificador del cálculo y vincular las reducciones obtenidas mediante diseño bajo en carbono a Certificados de Ahorro de Emisiones, replicando el modelo de monetización validado en EDMUS.

**Financiación:** Kit Espacios de Datos, complementado por la 2.<sup>a</sup> Convocatoria de Demostradores y por LIFE Programme. TRL 5-6.

## O6. Trazabilidad de Product Carbon Footprint sobre CEMENGAL: modelo Catena-X aplicado al cemento español

**Espacio de datos:** CEMENGAL, espacio de datos federado del sector cementero español, descrito como espacio "con GAIA-X e IDS-RAM para colaboración, eficiencia e innovación con IA en un sector global estratégico". La propia descripción identifica la sostenibilidad como reto sectorial determinante.

**Servicio de IA propuesto:** implementación del Pathfinder Framework v2.0 de WBCSD/PACT para Product Carbon Footprint en el sector cementero, con un agente de

IA que (i) negocia automáticamente vía EDC los PCF de clinker, áridos, aditivos y energía con cada proveedor del espacio, (ii) compone el PCF del cemento producido aplicando reglas de asignación conforme a la ISO 14067, y (iii) lo publica como activo de datos descubrible para los compradores aguas abajo (constructoras, prefabricados, distribuidores).

**Problema que resuelve:** el sector cementero español es uno de los más intensivos en emisiones de CO<sub>2</sub> por unidad de producto, con valores típicos de 600-900 kg CO<sub>2</sub>eq por tonelada de cemento. Bajo el Mecanismo de Ajuste en Frontera por Carbono (CBAM), aplicable plenamente desde 2026 al cemento, los exportadores y compradores intra-UE necesitan PCF granular y verificable. La práctica actual, basada en medias sectoriales, penaliza precisamente a los productores españoles más eficientes, que no pueden demostrar su ventaja sin datos primarios verificables.

**Justificación:** replica el caso de uso PCF de Catena-X (capítulo 6.1 del presente informe) en un sector donde España tiene capacidad exportadora relevante. CEMENGAL adopta IDS-RAM como referencia arquitectónica, lo que asegura continuidad técnica con el Dataspace Protocol. La convergencia entre PACT y DSP analizada en la sección 6.8 encuentra aquí un caso de aplicación concreto y comercialmente atractivo.

**Encaje con caso ECODES-EDMUS:** comparte el patrón "datos primarios verificados generan valor económico" que sustenta EDMUS, aplicado en este caso al PCF en lugar de a la certificación de reducciones evitadas.

**Financiación:** Kit Espacios de Datos, complementado por Horizon Europe Cluster 4 (Digital, Industry and Space) y, eventualmente, por mecanismos IPCEI Cloud, que permiten ayudas por encima de los límites ordinarios de minimis. TRL 5-6.

## O7. Auditoría automatizada CSRD sobre el Espacio Nacional de Datos de Salud

**Espacio de datos:** Espacio Nacional de Datos de Salud, activado por el Gobierno de España en enero de 2026 como infraestructura pública para compartir y utilizar datos sanitarios de forma segura.

**Servicio de IA propuesto:** servicio que cruza datos de consumo hospitalario (material fungible, gases medicinales, fármacos, energía) publicados en el espacio, los empareja con factores de emisión específicos vía LLM con RAG, calcula inventarios de Alcance 3 categoría 1 por hospital y por servicio clínico, e identifica intervenciones de descarbonización priorizadas. Las intervenciones típicas incluyen la sustitución de gases anestésicos de alto potencial de calentamiento global (sevoflurano, desflurano), la optimización de la cadena de frío farmacéutica y la reducción del uso de plásticos de un solo uso. El servicio devuelve un borrador estructurado de la sección E1 del informe CSRD.

**Problema que resuelve:** el sector sanitario aporta el 4,4% de las emisiones globales según Health Care Without Harm (cifra del capítulo 2.6), y la categoría 1 del Alcance 3 puede representar hasta el 70% de las emisiones de un hospital. Todos los hospitales

públicos están sujetos directa o indirectamente al reporte CSRD vía sus servicios autonómicos de salud, sin disponer en general de capacidad técnica interna para abordar el cálculo de Alcance 3.

**Justificación:** aprovecha el aprendizaje de ECODES con el proyecto MAPIC+s, citado en las secciones 1.4 y 2.6 del presente informe, en colaboración con GSK para la descarbonización del sector sanitario. El espacio nacional de salud, recién activado, busca explícitamente facilitar el desarrollo de soluciones digitales en el ámbito de la salud, y figura entre los espacios elegibles del Kit Espacios de Datos. Un servicio AlaaS centralizado tiene economía de escala obvia frente a soluciones hospital por hospital.

**Encaje con caso ECODES-EDMUS:** combina las dos líneas estratégicas de ECODES: el conocimiento clínico de MAPIC+s y la experiencia técnica de espacios de datos de EDMUS. Es la oportunidad que mejor capitaliza simultáneamente ambas trayectorias.

**Financiación:** Kit Espacios de Datos, complementado por Horizon Europe Cluster 1 (Health). TRL 4-5.

## O8. Plataforma de gestión de Zonas de Bajas Emisiones sobre DS4MM

**Espacio de datos:** DS4MM (Data Space for Mobility Models Integration), espacio europeo financiado por el Mecanismo Conectar Europa (CEF) con 3 millones de euros para la Junta de Extremadura como participante española. El proyecto se validará con cuatro pruebas piloto, una de ellas en la Comunitat Valenciana específicamente "para validar herramientas de gestión de las Zonas de Bajas Emisiones (ZBE) y de la movilidad urbana en ciudades como València, Castellón, Elche y Benidorm". La ejecución se extiende hasta junio de 2027.

**Servicio de IA propuesto:** gemelo digital de Zona de Bajas Emisiones alimentado por datos de tráfico (lectores de matrícula), calidad del aire (red Mocaa, EIONET), telemetría de flota privada accedida vía consentimiento explícito y meteorología. Incorpora un modelo predictivo de emisiones por tramo viario y un módulo de optimización política que evalúa qué umbrales de restricción aplicar, en qué franjas horarias y con qué exenciones, maximizando la reducción de NO<sub>x</sub> y CO<sub>2</sub>eq y minimizando el impacto socioeconómico medido en términos de equidad espacial y accesibilidad de los grupos vulnerables.

**Problema que resuelve:** las ZBE son obligatorias en municipios de más de 50.000 habitantes desde la Ley 7/2021 de Cambio Climático y Transición Energética, pero cada ayuntamiento las implementa con criterios distintos sin medición ex-post fiable de su impacto real en emisiones. Esta falta de evidencia bloquea la aceptación social y, paradójicamente, retrasa la descarbonización efectiva del sector transporte.

**Justificación:** DS4MM declara entre sus pilotos el análisis de emisiones, y está específicamente concebido para que la información compartida "se traduzca en modelos y herramientas que mejoren la gestión y la planificación", según declaraciones del coordinador del proyecto. Las cuatro ciudades piloto en la Comunitat Valenciana cubren

contextos heterogéneos (capital, ciudad media, costa turística), lo que asegura replicabilidad metodológica.

**Encaje con caso ECODES-EDMUS:** complementario. EDMUS certifica reducciones del transporte público electrificado dentro de la ZBE; este servicio cuantifica el efecto agregado de la propia ZBE como política pública. Las cifras pueden cruzarse para evitar doble conteo en los inventarios municipales.

**Financiación:** ya parcialmente financiado por CEF Transporte (3 millones de euros aportados por la Junta de Extremadura, 50% del coste total). El componente de IA es financiable por convocatorias autonómicas de la Generalitat Valenciana en transición ecológica y por la 2.<sup>a</sup> Convocatoria de Demostradores. TRL 5-7.

## O9. Trazabilidad de huella en cadena cítrica sobre CitriData

**Espacio de datos:** CitriData, "proyecto estratégico que busca transformar el sector citrícola andaluz mediante el uso [...] de los datos", presentado en el segundo semestre de 2025 e incluido en el Catálogo de Espacios de Datos Elegibles del CRED.

**Servicio de IA propuesto:** módulo que cruza datos de actividad agrícola (riego, fertilización, tratamientos fitosanitarios, geolocalización de parcelas), datos de transporte (cadena de frío, distancias, modo de transporte) y datos de procesamiento (consumo eléctrico de manipulación, refrigeración en almacén), publicados por los actores de la cadena cítrica andaluza. Calcula el Product Carbon Footprint de cada lote de naranjas o mandarinas exportadas mediante un modelo Random Forest entrenado con tablas input-output multirregionales españolas, en línea metodológica con el dataset ExioML referenciado en el capítulo 5 del presente informe.

**Problema que resuelve:** el cítrico andaluz exporta principalmente a Alemania, Francia y Países Bajos, países donde retailers como Carrefour, Albert Heijn y Edeka aplican criterios de huella de carbono por proveedor en sus decisiones de aprovisionamiento. Sin datos verificables, los productores españoles pierden cuota frente a competidores con sistemas de medición más consolidados.

**Justificación:** España es el primer productor europeo de cítricos y el primer exportador mundial. El sector está concentrado en Andalucía y la Comunitat Valenciana, lo que da masa crítica al espacio. La integración con AgriDataSpace europeo (CEADS) abre el catálogo a compradores de toda la Unión Europea, multiplicando el valor del servicio.

**Encaje con caso ECODES-EDMUS:** comparte el modelo de "datos primarios verificables como activo comercial" que sustenta EDMUS. ECODES puede actuar como entidad verificadora externa de los PCF emitidos.

**Financiación:** Kit Espacios de Datos, complementado por LIFE y por Horizon Europe Cluster 6. TRL 4.

## O10. Servicio de IA para huella de carbono evitada en infraestructuras viarias sobre AECTech

**Espacio de datos:** AECTech, descrito como "espacio de datos del sector viario centrado en la infraestructura que permite compartir e interpretar datos de tráfico, infraestructuras y movilidad para mejorar la seguridad, eficiencia y sostenibilidad".

**Servicio de IA propuesto:** cuantificador de emisiones evitadas por intervenciones en infraestructura viaria, incluyendo renovación de pavimento de bajo rozamiento, despliegue de rotondas inteligentes, fluidificación del tráfico mediante onda verde semafórica y despliegue de Sistemas de Transporte Inteligente (STI). Combina modelos microscópicos de tráfico (SUMO o Aimsun) calibrados con datos del catálogo del espacio y modelos de emisión COPERT y HBEFA para vehículos. Genera certificados auditables por cada intervención que el gestor de la infraestructura puede monetizar como CAES siguiendo el modelo EDMUS.

**Problema que resuelve:** las intervenciones de "ingeniería gentil" sobre infraestructura existente (mantenimiento preventivo, optimización de señalización, fluidificación) generan reducciones reales de emisiones que hoy no se contabilizan ni se monetizan. El gestor (concesionaria de autopista o administración titular) no recupera valor económico por su contribución climática, lo que reduce los incentivos a invertir en este tipo de intervenciones frente a alternativas más visibles, pero menos costo-efectivas.

**Justificación:** replica conceptualmente el modelo ECODES-EDMUS desde el lado del gestor de infraestructura, en lugar de desde el operador de transporte. AECTech está explícitamente diseñado para "interpretar datos de tráfico para mejorar [...] la sostenibilidad", lo que da cobertura programática directa al servicio.

**Encaje con caso ECODES-EDMUS:** extensión natural del modelo. Emite Certificados de Ahorro de Emisiones desde otra capa del sistema de movilidad (la infraestructura, no la flota), permitiendo eventualmente un mercado conjunto con compradores institucionales únicos.

**Financiación:** Kit Espacios de Datos, complementado por la 2.<sup>a</sup> Convocatoria de Demostradores. TRL 5.

## O11. Bancos y emisiones financiadas (PCAF) sobre la combinación Lista de Confianza más espacios sectoriales

**Espacio de datos:** aprovecha la Lista de Confianza de Espacios de Datos, registro oficial regulado por la Orden TDF/1207/2025, de 27 de octubre, con activación en enero de 2026. Identifica los espacios de datos que han superado los filtros técnicos y de seguridad establecidos por el Ministerio para la Transformación Digital y de la Función Pública, basados en la Especificación UNE 0087:2025. Funciona como meta-espacio de identidades verificadas que un servicio financiero puede consultar para encontrar fuentes fiables de emisiones corporativas.

**Servicio de IA propuesto:** agente que, dado un portfolio crediticio o de inversión de una entidad financiera española, (i) identifica para cada empresa participada si publica sus emisiones en algún espacio de datos de la Lista de Confianza, (ii) si las publica, las descarga vía DSP y las atribuye proporcionalmente al banco según la metodología PCAF (Partnership for Carbon Accounting Financials), (iii) si no las publica, aplica predicción XGBoost a partir de los estados financieros disponibles en los registros bancarios, y (iv) genera el inventario completo de emisiones financiadas (categoría 15 del Alcance 3) con desglose por sector y por empresa, identificando los hotspots de la cartera.

**Problema que resuelve:** para los grandes bancos españoles (BBVA, CaixaBank, Santander, Sabadell, Bankinter), las emisiones financiadas pueden superar el 99% del total del Alcance 3. El reporte conforme a PCAF requiere meses de trabajo manual y la calidad de los datos resultantes es baja. La supervisión del Banco Central Europeo sobre riesgos climáticos (Guide on climate-related and environmental risks) eleva continuamente la exigencia.

**Justificación:** el sector bancario español está integralmente sujeto a la CSRD desde el primer ciclo de reporte (enero 2024) y a las directrices del BCE sobre riesgos climáticos. La Lista de Confianza es el habilitador clave porque garantiza identidad verificada y auditabilidad de las transacciones, requisitos imprescindibles para el aseguramiento limitado y, eventualmente, razonable bajo CSRD.

**Encaje con caso ECODES-EDMUS:** el banco compra Certificados de Ahorro de Emisiones en la plataforma EDMUS para compensar parte de las emisiones financiadas no reducibles a corto plazo, cerrando el ciclo "medición vía espacio de datos sectorial → compensación vía espacio de datos de movilidad". Esta integración multi-espacio constituye una propuesta de valor diferencial.

**Financiación:** convocatoria de Productos y Servicios de Espacios de Datos, complementada por Horizon Europe Cluster 4 y por cofinanciación privada del propio sector bancario, naturalmente interesado en el desarrollo de la solución. TRL 4-5.

## O12. Declaraciones Ambientales de Producto automatizadas para Hábitat Data y mueble español

**Espacio de datos:** Hábitat Data, caso de uso del espacio PRECEDENT, diseñado para "compartir e intercambiar datos de mercado del sector hábitat generando inteligencia competitiva". Aunque su orientación inicial es la inteligencia comercial, su catálogo de productos y materiales lo convierte en sustrato idóneo para servicios ambientales.

**Servicio de IA propuesto:** generador automático de Declaraciones Ambientales de Producto conforme a la EN 15804+A2 para mobiliario. Recoge desde el catálogo Hábitat Data el bill of materials de cada referencia (madera, metales, plásticos, espumas, tejidos, herrajes), aplica factores de emisión de ecoinvent vía RAG semántico, integra datos de transporte de los proveedores conectados, y emite la DAP firmada digitalmente y registrable en programas como GlobalEPD del AENOR o EPD International.

**Problema que resuelve:** el mueble español, primer productor europeo en algunos segmentos con clústeres en la Comunitat Valenciana, Yecla y La Rioja, exporta a mercados nórdicos y centroeuropeos donde la DAP es prácticamente exigida por compradores institucionales y grandes contract. El coste de elaborar una DAP por la vía clásica (entre 8.000 y 15.000 euros por producto) bloquea su aplicación a catálogos amplios, especialmente para pymes.

**Justificación:** convierte el espacio Hábitat Data, hoy concebido como inteligencia competitiva, en infraestructura de cumplimiento ambiental, lo que aumenta su atractivo para nuevos participantes y asegura adopción sostenida. Es un ejemplo de cómo un servicio de IA bien diseñado puede ampliar el caso de uso de un espacio de datos sin requerir su rediseño técnico.

**Encaje con caso ECODES-EDMUS:** extiende la lógica de "datos verificables como activo comercial" a la documentación ambiental de producto. ECODES puede ofrecer servicios de verificación independiente de las DAP generadas.

**Financiación:** Kit Espacios de Datos, complementado por LIFE y cofinanciación de los clústeres sectoriales (AIDIMME, ANIEME). TRL 4-5.

## O13. Optimización energética predictiva sobre el Espacio de Datos de Sacyr extendido a aguas-energía-emisiones

**Espacio de datos:** Espacio de Datos de Sacyr, ya operativo, que incluye según la descripción del CRED "una herramienta de IA para calcular el índice de calidad del agua (ICA) con algoritmos predictivos, aplicable a sistemas hidrológicos".

**Servicio de IA propuesto:** ampliación del módulo de IA actual con un componente carbon-aware, que cruza la predicción de demanda de agua con datos de intensidad de carbono horaria de la red eléctrica española, accesibles vía conector con el futuro Energy Data Space español o vía API directa de Red Eléctrica de España. El servicio planifica los ciclos de bombeo y de tratamiento para minimizar las emisiones de Alcance 2 manteniendo niveles de servicio garantizados. Un modelo de Reinforcement Learning entrena la política operativa óptima.

**Problema que resuelve:** el ciclo integral del agua (captación, potabilización, distribución, depuración) representa entre el 2% y el 3% del consumo eléctrico nacional. Las EDAR y plantas de potabilización son grandes consumidores con flexibilidad operativa importante, ya que disponen de capacidad de depósito que permite desplazar consumos entre horas valle y horas punta. La práctica habitual optimiza por mínimo coste eléctrico tarifario, no por mínima huella de carbono. La aplicación del paradigma carbon-aware computing a infraestructuras hídricas es una ventana de oportunidad poco explotada.

**Justificación:** aprovecha que Sacyr ya tiene desplegado el conector y el espacio operativo, con un módulo de IA para ICA en producción. El paso del ICA (calidad de agua) al carbono es marginal en coste de integración. Existen casos análogos validados en Microsoft y Google para sus centros de datos, citados en el capítulo 6 y en el anexo F.4 del presente informe.

**Encaje con caso ECODES-EDMUS:** comparte la lógica de servicio de IA sobre espacio de datos operativo, en un dominio diferente. La empresa gestora puede certificar las reducciones de Alcance 2 obtenidas y monetizarlas en mercados voluntarios de carbono, replicando el modelo CAES de EDMUS.

**Financiación:** convocatoria de Productos y Servicios de Espacios de Datos, complementada por Horizon Europe Cluster 5. TRL 6.

## O14. Onboarding asistido por LLM para pymes en el ecosistema Kit Espacios de Datos

**Espacio de datos:** oportunidad transversal que opera sobre los 165 espacios de datos elegibles del catálogo del CRED.

**Servicio de IA propuesto:** asistente conversacional basado en LLM agentic que (i) entrevista a una pyme sobre su actividad económica, datos disponibles y obligaciones de reporte, (ii) le sugiere a qué espacios de datos elegibles le conviene adherirse según sector y caso de uso, (iii) genera la documentación requerida por el Kit Espacios de Datos (preparación de datos, declaración de evidencias de registro y publicación), (iv) prepara el conjunto de datos en el formato canónico exigido por el espacio escogido, y (v) asiste en la tramitación de la solicitud Red.es.

**Problema que resuelve:** la barrera principal de adopción del Kit Espacios de Datos no es financiera, ya que la convocatoria mantiene más del 60% de los 60 millones de euros disponibles a marzo de 2026. La barrera real es la complejidad administrativa y técnica del proceso de adhesión. Sin pymes adheridas, los espacios sectoriales no acumulan datos primarios suficientes y los servicios de IA de Alcance 3 carecen de materia prima. Es la palanca multiplicadora del ecosistema completo descrito en este capítulo.

**Justificación:** aborda directamente el riesgo de mercado principal identificado en la sección 10.4 del capítulo siguiente. El servicio acompaña el papel del CRED como punto de orientación y dinamizador de los ecosistemas. Aunque el plazo de la convocatoria actual cierra el 31 de marzo de 2026, todo apunta a una segunda edición; el servicio se posiciona estratégicamente para esa siguiente ola de financiación.

**Encaje con caso ECODES-EDMUS:** meta-servicio que multiplica las oportunidades de despliegue de cualquiera de las restantes catorce. Sin onboarding masivo de pymes, las oportunidades sectoriales sobre AgrospAI, CEMENGAL, GreenBIMxels o CitriData no alcanzan masa crítica.

**Financiación:** convocatoria de Productos y Servicios de Espacios de Datos (44 millones de euros ya resueltos, pero con ampliación esperable), complementada por Acelera pyme (servicios de asesoría). TRL 5.

## O15. Capa de verificación satelital cruzada con datos empresariales para auditoría de aseguramiento CSRD

**Espacio de datos:** combinación entre cualquier espacio sectorial industrial (CEMENGAL, espacios de manufactura emergentes) y datos abiertos del programa Copernicus, con foco en Sentinel-5P, TROPOMI y servicios CAMS y C3S, integrados como activos de referencia accesibles a los servicios de IA.

**Servicio de IA propuesto:** modelo que combina (i) la autodeclaración de emisiones de una instalación industrial, accedida vía DSP en su espacio sectorial, con (ii) las concentraciones atmosféricas de NO<sub>2</sub>, CH<sub>4</sub> y CO<sub>2</sub> observadas por satélite en la celda geográfica correspondiente. Cuando se detecta desviación significativa entre lo reportado y lo observado tras controlar por variables meteorológicas y de actividad industrial vecina, el servicio genera un flag de revisión para el auditor independiente exigido por la CSRD para el aseguramiento limitado y, progresivamente, razonable.

**Problema que resuelve:** el aseguramiento limitado de la información de sostenibilidad bajo CSRD requiere evidencia más allá de la documentación interna; el aseguramiento razonable, previsto en el roadmap actual a partir de 2028, eleva sustancialmente la exigencia. Las firmas de auditoría no tienen capacidad técnica para auditar miles de instalaciones con métodos tradicionales; un servicio top-down satelital cruzado con datos bottom-up corporativos es la única vía escalable conocida.

**Justificación:** desarrolla y consolida la oportunidad 6 de la tipología genérica del capítulo 7 (auditoría automatizada CSRD), añadiendo la dimensión de verificación independiente que la sección 5.8 del presente informe menciona como posibilidad emergente. Aprovecha datos abiertos europeos sin coste de licencia, lo que reduce significativamente el coste marginal del servicio.

**Encaje con caso ECODES-EDMUS:** aplica al lado del verificador externo, no del operador. Sus outputs alimentan la auditoría de los Certificados de Ahorro de Emisiones emitidos por EDMUS y por las oportunidades O1 y O10, reforzando la credibilidad de todo el ecosistema.

**Financiación:** Horizon Europe Cluster 4, complementada por ESA Φ-lab y por colaboración con grandes firmas de auditoría. TRL 3-4.

## 9.5. Síntesis y priorización

Si se ordenan las quince oportunidades por el producto entre nivel de madurez tecnológica, encaje regulatorio inmediato y disponibilidad de financiación abierta a fecha del informe, las cinco oportunidades de actuación más inmediata son las siguientes.

La primera es **O1** (replicación del modelo ECODES-EDMUS sobre DS4M Mediterráneo): aprovecha el know-how propio acumulado, el espacio receptor está financiado y en despliegue activo, y existe alineación territorial directa con Aragón, lo que facilita el liderazgo técnico de ITA. La segunda es **O2** (predicción Alcance 3 sobre el Espacio de

Datos de Turismo): este espacio es el primero sectorial operativo en España y necesita servicios de valor añadido para retener a sus early adopters durante la fase crítica de consolidación del catálogo. La tercera es **O7** (auditoría CSRD para el sector sanitario): combina las dos líneas históricas de ECODES, el conocimiento clínico de MAPIC+s y la experiencia técnica de espacios de datos de EDMUS. La cuarta es **O5** (servicio de IA sobre GreenBIMxels): el promotor del espacio identifica explícitamente la inteligencia artificial como vector de valor en su descripción oficial, lo que garantiza encaje cultural y estratégico inmediato. La quinta es **O14** (onboarding asistido por LLM): habilitador transversal con efecto multiplicador sobre el resto de las oportunidades del catálogo, especialmente relevante para abordar el riesgo de baja adopción por pymes identificado en el capítulo siguiente.

Tres oportunidades adicionales conviene reservar para una propuesta europea de mayor envergadura, alineada con la propuesta AI4GreenData descrita en la sección 10.6 del capítulo siguiente. La **O6** (CEMENGAL en el contexto del CBAM), la **O11** (PCAF aplicado al sector bancario español) y la **O15** (verificación satelital cruzada) presentan TRL más bajos, mayor complejidad multi-stakeholder y horizontes temporales más largos, pero su impacto sistémico es proporcionalmente más alto y su alineamiento con la agenda Horizon Europe es óptimo.

Las siete oportunidades restantes (O3, O4, O8, O9, O10, O12, O13) configuran una segunda ola de despliegue, escalable a medida que la primera ola consolide el modelo y libere recursos técnicos y comerciales.

# 10. Conclusiones y recomendaciones

## 10.1. Conclusiones principales

La investigación realizada permite establecer las siguientes conclusiones. Primera, el problema del Alcance 3 de emisiones de GEI es, en su núcleo, un problema de intercambio de datos entre organizaciones de una cadena de valor, lo que convierte a los espacios de datos soberanos en la infraestructura natural para abordarlo. Segunda, España cuenta con un ecosistema regulatorio y de financiación sin precedentes para el despliegue de espacios de datos sectoriales (500 M€ en el Plan de Impulso, 60 M€ en Kit Espacios de Datos, UNE 0087:2025), que crea un contexto favorable para la incorporación de servicios de IA orientados a la sostenibilidad.

Tercera, los modelos de machine learning, particularmente los algoritmos de ensemble como AdaBoost y XGBoost, pueden predecir emisiones de Alcance 3 con precisión superior a los modelos lineales utilizando exclusivamente datos financieros y clasificaciones industriales, lo que reduce drásticamente las barreras de entrada para la medición de emisiones indirectas. Cuarta, el caso de uso ECODES-EDMUS demuestra que la combinación de espacio de datos soberano (EDC) + cálculo automatizado de emisiones + integración con el mercado de carbono (EU ETS) es técnicamente viable y puede generar valor económico tangible para todos los participantes.

Quinta, el despliegue de servicios de IA sobre espacios de datos multisectoriales abre al menos doce oportunidades concretas de impacto, con niveles de madurez tecnológica entre TRL 3 y TRL 8, que cubren desde la predicción de emisiones hasta la certificación inteligente de reducciones, pasando por la optimización multimodal del transporte y la auditoría automatizada de reportes CSRD. Sexta, la huella de carbono de la propia IA está disminuyendo rápidamente y, en Europa, el balance neto de los servicios de IA climática es claramente positivo, con ratios estimados de beneficio-coste de carbono de entre 10:1 y 100:1.

Séptima, la experiencia internacional confirma que los países líderes en este ámbito (Alemania con Catena-X, Francia con EONA-X, los países nórdicos con sus iniciativas de datos abiertos para sostenibilidad) comparten un patrón común: inversión pública significativa, colaboración entre centros tecnológicos y organizaciones ambientales, y casos de uso demostradores. España, con ITA, ECODES y el Plan de Impulso, cumple las tres condiciones. Octava, el marco regulatorio CSRD/ESRS creará una demanda creciente de servicios automatizados de cálculo y reporte de Alcance 3 a medida que se extienda el ámbito de aplicación a pymes cotizadas y empresas de terceros países con operaciones en la UE, lo que ampliará significativamente el mercado potencial de los servicios de IA identificados.

## 10.2. Recomendaciones

Se formulan las siguientes recomendaciones para el ecosistema español de espacios de datos e IA climática. A los organismos públicos, se recomienda incluir explícitamente la reducción de huella de carbono como criterio de evaluación en futuras convocatorias de espacios de datos, y crear un catálogo de servicios de IA climática homologados para espacios de datos de la Lista de Confianza. A las empresas, se recomienda evaluar su incorporación a espacios de datos sectoriales como primer paso hacia la medición precisa del Alcance 3, aprovechando la financiación disponible en el Kit Espacios de Datos. A los centros tecnológicos y universidades, se recomienda desarrollar servicios de IA como servicio (AlaaS) específicamente diseñados para operar sobre la infraestructura EDC, publicando los modelos como activos de datos en catálogos federados.

A ITA y ECODES conjuntamente, se recomienda escalar el modelo ECODES-EDMUS a otros operadores de transporte y sectores, publicar la metodología como estándar abierto, y liderar la propuesta de un proyecto europeo que integre servicios de IA sobre el espacio común europeo de datos de movilidad (EMDS) para la descarbonización del transporte urbano.

## 10.3. Perspectivas futuras

La convergencia de espacios de datos, IA y sostenibilidad se encuentra en un punto de inflexión. Los próximos dos años (2026-2028) serán determinantes para establecer las infraestructuras, estándares y modelos de negocio que dominarán la próxima década. La entrada en vigor plena de la CSRD, la extensión del EU ETS a nuevos sectores (transporte marítimo desde 2024, aviación intra-UE), y la aplicación del Mecanismo de Ajuste en Frontera por Carbono (CBAM) desde 2026 crearán una demanda sin precedentes de datos de emisiones verificados e interoperables a lo largo de cadenas de valor globales.

En este escenario, los espacios de datos soberanos serán la infraestructura básica para el intercambio de datos ambientales, de la misma forma que internet lo es para el intercambio de información general. Y los servicios de IA sobre estos espacios de datos serán los agentes que transformen datos en bruto en conocimiento accionable: predicciones de emisiones, certificados de reducción, recomendaciones de descarbonización, informes de aseguramiento y alertas de riesgo climático. El presente informe aspira a servir como hoja de oportunidades para que ITA, ECODES y el ecosistema español de innovación se posicionen como actores relevantes en esta transformación.

La evolución de las tecnologías de IA, con el surgimiento de modelos multimodales capaces de procesar simultáneamente texto, imágenes satelitales, series temporales y datos tabulares, ampliará significativamente las capacidades de los servicios de IA climática sobre espacios de datos. Un servicio de IA multimodal podrá, en un futuro próximo, cruzar automáticamente datos de actividad empresarial (facturas, órdenes de

compra) con imágenes satelitales de las instalaciones de los proveedores, series temporales de consumo energético de smart meters, y textos de informes de sostenibilidad de terceros, generando inventarios de Alcance 3 con una precisión y una cobertura que los métodos manuales actuales no pueden alcanzar.

La emergencia de los modelos de lenguaje de gran escala especializados en dominios (domain-specific LLMs) abre también nuevas posibilidades. Un LLM entrenado específicamente con datos de sostenibilidad, factores de emisión, normativa ambiental y metodologías de cálculo podría actuar como asistente inteligente para los responsables de sostenibilidad de las empresas, guiándoles en el proceso de cálculo de huella de carbono, sugiriendo las fuentes de datos más apropiadas del espacio de datos, identificando lagunas en su inventario y generando automáticamente las secciones narrativas del informe CSRD. La integración de este tipo de LLMs con los agentes autónomos (AI agents) que pueden ejecutar negociaciones EDC, descargar datos y ejecutar cálculos de forma autónoma, creará un ecosistema de servicios de IA climática radicalmente más accesible y eficiente.

Finalmente, el desarrollo de estándares internacionales para el intercambio de datos de emisiones (PACT/Pathfinder, CDP Digital, GRI Taxonomy) y la extensión de los principios de Gaia-X a Asia y América abrirán la posibilidad de espacios de datos de sostenibilidad verdaderamente globales, en los que las cadenas de suministro transnacionales puedan compartir datos de emisiones de forma soberana e interoperable. Los servicios de IA que se desarrollen hoy sobre espacios de datos europeos estarán posicionados para escalar a estos ecosistemas globales, proporcionando una ventaja competitiva a las organizaciones que inviertan tempranamente en esta convergencia tecnológica.

La visión a largo plazo contempla un ecosistema donde cada empresa participante en un espacio de datos pueda acceder, con un clic, a un servicio de IA que calcule su huella de Alcance 3 completa, identifique las categorías de mayor impacto, recomiende intervenciones priorizadas por coste-beneficio ambiental, genere automáticamente la sección E1 de su informe CSRD, y certifique las reducciones conseguidas para su monetización en mercados de carbono. Esta visión, ambiciosa pero técnicamente viable con las tecnologías actuales, requiere la colaboración coordinada de organismos públicos, centros tecnológicos, organizaciones ambientales y el sector privado. El camino está trazado; el momento de recorrerlo es ahora.

# Referencias bibliográficas

- [1] Serafeim, G. y Vélez Caicedo, G. (2022). Machine Learning Models for Prediction of Scope 3 Carbon Emissions. Harvard Business School Working Paper No. 22-080. doi:10.2139/ssrn.4149874.
- [2] Wang, S-Y. y Ye, N-Z. (2025). Invisible footprints, visible insights: machine learning reveals Scope 3 emissions. *Frontiers in Sustainability*, 6, 1649150. doi:10.3389/frsus.2025.1649150.
- [3] Qu, S. et al. (2023). Scope 3 emissions: Data quality and machine learning prediction accuracy. *PLOS Climate*, 2(11), e0000208. doi:10.1371/journal.pclm.0000208.
- [4] Hettler, B. (2024). Corporate scope 3 carbon emission reporting as an enabler of supply chain decarbonization: A systematic review. *Business Strategy and the Environment*, 33(1), 53-78. doi:10.1002/bse.3486.
- [5] Sustainability (2025). Conceptualization of AI Use for GHG Scope 3 Emissions Measurement, Reporting, Monitoring, and Assurance. *Sustainability*, 17(22), 10153. doi:10.3390/su172210153.
- [6] MDPI Sustainability (2025). AI-Driven Supply Chain Decarbonization: Strategies for Sustainable Carbon Reduction. *Sustainability*, 17(21), 9642.
- [7] Nature Scientific Data (2026). Global emission factor dataset for Scope 3 machine learning applications. *Scientific Data*. doi:10.1038/s41597-026-06699-1.
- [8] Gaia-X AISBL (2025). White Paper: The Role of Data Spaces in the Digital Economy. [gaia-x.eu](http://gaia-x.eu).
- [9] Gaia-X (2025). Trust Framework 3.0 — Danube Release. Summit 2025, Porto.
- [10] Naumann, S. y Schade, J. (2026). Leveraging GAIA-X for Intelligent Mobility and Transportation Systems in Europe. *TSTP 2025. Lecture Notes in Networks and Systems*, vol. 1769. Springer.
- [11] Comisión Europea (2023). COM(2023)751: Creation of a common European mobility data space. *EUR-Lex*.
- [12] deployEMDS (2025). Deploying the European Mobility Data Space. Horizon Europe. [deployemds.eu](http://deployemds.eu).
- [13] Interreg North Sea (2025). Story on the Common European Mobility Data Space. Febrero 2025.
- [14] BCG y CO2 AI (2025). Climate Survey 2025. [co2ai.com](http://co2ai.com).
- [15] IEA (2025). Energy and AI: AI and climate change. IEA, París.
- [16] MIT News (2025). Responding to the climate impact of generative AI. Septiembre 2025.
- [17] Euronews (2025). AI data centres could have a carbon footprint that matches small European country. Diciembre 2025.
- [18] Ministerio para la Transformación Digital (2025). Plan de Impulso de los Espacios de Datos Sectoriales. [espanadigital.gob.es](http://espanadigital.gob.es).
- [19] Orden TDF/758/2025, de 11 de julio. Programa Kit Espacios de Datos. BOE nº 170, 16 de julio de 2025.
- [20] Red.es (2025). Convocatoria de ayudas Kit Espacios de Datos. [sede.red.gob.es](http://sede.red.gob.es).
- [21] Especificación UNE 0087:2025. Marco común para la creación y operación de espacios de datos. Asociación Española de Normalización.
- [22] Plan de Recuperación, Transformación y Resiliencia. Componente 12, Inversión 1. [planderecuperacion.gob.es](http://planderecuperacion.gob.es).
- [23] ITA (2022). Estrategia Aragonesa de Inteligencia Artificial (EAIA). [ita.es](http://ita.es).
- [24] ECODES — Fundación Ecología y Desarrollo (2024). Carbonpedia: Base de datos abierta de huella de carbono. [ecodes.org](http://ecodes.org).
- [25] ECODES y GSK (2025). Proyecto MAPIC+s: Medio Ambiente y Procesos Clínicos + Sostenibles.
- [26] Eclipse Dataspace Components. Eclipse Foundation. [eclipse.org/edc](http://eclipse.org/edc).

- [27] Mobility Data Space (2025). Solutions and features. [mobility-dataspace.eu](https://mobility-dataspace.eu).
- [28] World Economic Forum (2025). 6 ways data centres can cut their emissions. [weforum.org](https://weforum.org).
- [29] FAS (2025). Measuring AI's Energy/Environmental Footprint to Access Impacts. [fas.org](https://fas.org).
- [30] wetransform (2024). The Green Deal, Data Space, and GAIA-X. [wetransform.to](https://wetransform.to).
- [31] Gaia-X Hub Germany (2025). From data spaces to data ecosystems. [gaia-x-hub.de](https://gaia-x-hub.de).
- [32] MIT Sloan (2024). Scope 3 emissions top supply chain sustainability challenges. [mitsloan.mit.edu](https://mitsloan.mit.edu).
- [33] GHG Protocol. Corporate Value Chain (Scope 3) Accounting and Reporting Standard. [ghgprotocol.org](https://ghgprotocol.org).
- [34] Ranganathan, J. et al. (2004). The Greenhouse Gas Protocol: A Corporate Accounting and Reporting Standard. WRI/WBCSD.
- [35] Nguyen, T. et al. (2021). Machine learning approach for carbon footprint prediction. *Environmental Research Letters*.
- [36] Kläaßen, L. y Stoll, C. (2021). A harmonized dataset of global supply chain carbon emissions. *Scientific Data*.
- [37] EFRAG (2024). European Sustainability Reporting Standards (ESRS). ESRS E1: Climate Change.
- [38] PCAF (2022). The Global GHG Accounting and Reporting Standard for the Financial Industry. Part A.
- [39] DSSC (2024). Data Spaces Blueprint v1.0. Data Spaces Support Centre. [dssc.eu](https://dssc.eu).
- [40] IDSA (2024). Dataspace Protocol Specification v0.8. International Data Spaces Association.
- [41] WBCSD/PACT (2023). Pathfinder Framework v2.0. World Business Council for Sustainable Development.
- [42] SITRA (2024). Fair Data Economy: Rules of the Game. Finnish Innovation Fund.
- [43] Rolnick, D. et al. (2022). Tackling Climate Change with Machine Learning. *ACM Computing Surveys*, 55(2), 1-96.
- [44] Kaack, L.H. et al. (2022). Aligning artificial intelligence with climate change mitigation. *Nature Climate Change*, 12, 518-527.
- [45] Otto, B. et al. (2022). *Designing Data Spaces: The Ecosystem Approach to Competitive Advantage*. Springer.
- [46] Curry, E. et al. (2022). *Data Spaces: Design, Deployment, and Future Directions*. Springer.
- [47] SBTi (2024). Corporate Net-Zero Standard v2.0. Science Based Targets initiative.

# Anexo A. Glosario de términos

## A.1. Términos de espacios de datos

| Término      | Definición                                                                                                |
|--------------|-----------------------------------------------------------------------------------------------------------|
| EDC          | Eclipse Dataspace Connector. Componente open-source para la participación en espacios de datos soberanos. |
| DSP          | Data Space Provider. Participante que publica activos de datos en el catálogo del espacio de datos.       |
| DSC          | Data Space Consumer. Participante que descubre y accede a activos de datos de otros participantes.        |
| EDR          | Endpoint Data Reference. Referencia temporal con URL y token para acceder a datos negociados.             |
| IDSA         | International Data Spaces Association. Organización que define estándares para espacios de datos.         |
| Gaia-X       | Iniciativa europea para una infraestructura de datos segura, soberana e interoperable.                    |
| EMDS         | European Mobility Data Space. Espacio común europeo de datos de movilidad.                                |
| DGA          | Data Governance Act. Reglamento europeo de gobernanza de datos.                                           |
| Data Act     | Ley de Datos. Reglamento europeo sobre acceso y uso justo de datos.                                       |
| DSSC         | Data Spaces Support Centre. Centro de soporte para espacios de datos europeos.                            |
| DSP Protocol | Dataspace Protocol. Protocolo estándar para la interacción entre participantes de espacios de datos.      |
| DCAT         | Data Catalog Vocabulary. Vocabulario RDF para describir catálogos de datos.                               |
| ODRL         | Open Digital Rights Language. Lenguaje para expresar políticas de acceso y uso de datos.                  |

## A.2. Términos de huella de carbono

| Término      | Definición                                                                                           |
|--------------|------------------------------------------------------------------------------------------------------|
| GHG Protocol | Greenhouse Gas Protocol. Estándar internacional para la contabilidad de emisiones de GEI.            |
| Alcance 1    | Emisiones directas de fuentes controladas por la organización.                                       |
| Alcance 2    | Emisiones indirectas de la generación de electricidad, vapor y calefacción adquiridos.               |
| Alcance 3    | Todas las demás emisiones indirectas de la cadena de valor.                                          |
| EU ETS       | European Union Emissions Trading System. Mercado europeo de derechos de emisión de CO <sub>2</sub> . |
| EUA          | European Union Allowance. Instrumento financiero del EU ETS (1 EUA = 1 tCO <sub>2</sub> ).           |
| CSRD         | Corporate Sustainability Reporting Directive. Directiva europea de informes de sostenibilidad.       |
| ESRS         | European Sustainability Reporting Standards. Estándares de reporte bajo CSRD.                        |
| SBTi         | Science Based Targets initiative. Iniciativa para objetivos de reducción basados en ciencia.         |
| PCF          | Product Carbon Footprint. Huella de carbono de un producto a lo largo de su ciclo de vida.           |
| CAES         | Certificado de Ahorro de Emisiones de CO <sub>2</sub> . Instrumento de certificación de reducciones. |
| CBAM         | Carbon Border Adjustment Mechanism. Mecanismo de ajuste en frontera por carbono de la UE.            |
| PCAF         | Partnership for Carbon Accounting Financials. Estándar para emisiones financiadas.                   |

## A.3. Términos de inteligencia artificial

| Término  | Definición                                                                                     |
|----------|------------------------------------------------------------------------------------------------|
| AdaBoost | Adaptive Boosting. Algoritmo de ensemble que combina clasificadores débiles secuencialmente.   |
| XGBoost  | Extreme Gradient Boosting. Algoritmo de ML basado en árboles de decisión con regularización.   |
| LLM      | Large Language Model. Modelo de lenguaje de gran escala entrenado en corpus textuales masivos. |
| RAG      | Retrieval-Augmented Generation. Técnica que combina búsqueda con generación de texto.          |
| LSTM     | Long Short-Term Memory. Tipo de red neuronal recurrente para secuencias temporales.            |
| RL       | Reinforcement Learning. Paradigma de aprendizaje por interacción con un entorno.               |
| TRL      | Technology Readiness Level. Escala de madurez tecnológica (1-9).                               |
| AlaaS    | AI as a Service. Modelo de despliegue de IA como servicio bajo demanda.                        |

# Anexo B. Arquitectura técnica detallada del caso ECODES-EDMUS

## B.1. Diagrama de componentes

La arquitectura del sistema ECODES sobre el espacio de datos EDMUS se compone de los siguientes componentes desplegados como contenedores Docker: Conector OperadorDeMovilidad, que gestiona el catálogo de activos, la negociación de contratos y la emisión de EDRs. API Operador, que expone los datos reales de billeteaje a través de endpoints protegidos por los tokens del EDR. Conector ECODES, que actúa como consumidor del espacio de datos, iniciando las negociaciones y gestionando las transferencias. API ECODES, que implementa la lógica de negocio: cálculo de huella, generación de certificados, integración con EU ETS. Marketplace Web, que proporciona la interfaz de usuario para descubrimiento, compra y verificación de certificados. Dashboard Analítico (Streamlit, puerto 8800), que ofrece visualización de KPIs, portfolio de certificados y gestión económica.

## B.2. Flujo de datos completo

El flujo de datos de una compra típica comienza cuando el usuario envía un POST al endpoint /huella/comprar/stream con la cantidad de kg de CO<sub>2</sub> deseada. La API ECODES inicia una negociación EDC con el conector del operador para obtener el activo billeteaje-hdc-count (disponibilidad total) y billeteaje-hdc-list (desglose horario). Con esta información, calcula cuántos bloques horarios son necesarios para cubrir la demanda. Para cada bloque, ejecuta un ciclo EDC completo de cinco fases (catálogo, negociación, transferencia, EDR, descarga) que dura entre 10 y 20 segundos. Los tickets descargados se agregan, se calcula el CO<sub>2</sub> evitado por bloque, se genera el JSON del certificado con toda la metodología embebida, y se produce un PDF profesional con ReportLab que incluye portada, KPIs, gráficos comparativos, distribución horaria, tabla de detalle, metodología completa y código QR de verificación.

## B.3. Stack tecnológico completo

| Componente  | Tecnología        | Función                                            |
|-------------|-------------------|----------------------------------------------------|
| Backend API | FastAPI + Uvicorn | Framework REST asíncrono para endpoints de negocio |

|                     |                             |                                                 |
|---------------------|-----------------------------|-------------------------------------------------|
| Cliente HTTP        | httpx                       | Comunicación asíncrona con conector EDC         |
| Generación PDF      | ReportLab                   | PDFs programáticos con gráficos, tablas y marca |
| Gráficos en PDF     | Matplotlib                  | Distribución horaria y comparativa de emisiones |
| Marketplace         | Flask                       | Aplicación web para flujo de compra             |
| Dashboard           | Streamlit                   | Panel de KPIs, portfolio y previsión            |
| Procesamiento datos | Pandas                      | Agregación y análisis del portfolio             |
| Precios EU ETS      | yfinance + scraping         | Cuatro fuentes con fallback progresivo          |
| Códigos QR          | QRCode + Pillow             | Verificación pública de certificados            |
| Espacio de datos    | Eclipse Dataspace Connector | Catálogo, negociación y transferencia segura    |
| Infraestructura     | Docker / Docker Compose     | Contenerización de los 3 servicios + conectores |

## B.4. Endpoints principales de la API ECODES

| Endpoint               | Método  | Descripción                                                     |
|------------------------|---------|-----------------------------------------------------------------|
| /huella/comprar/stream | POST    | Compra de certificados con streaming de progreso NDJSON         |
| /huella/disponibilidad | GET     | Consulta de CO <sub>2</sub> disponible para certificar          |
| /verificar/{cert_id}   | GET     | Verificación pública de certificado (sin autenticación)         |
| /certificados          | GET     | Lista completa del portfolio de certificados emitidos           |
| /certificados/{id}/pdf | GET     | Descarga del PDF del certificado                                |
| /config/factores       | GET/PUT | Lectura y actualización de factores de emisión                  |
| /precio-co2            | GET     | Precio actual del CO <sub>2</sub> en EU ETS con fuente de datos |

|               |     |                                                   |
|---------------|-----|---------------------------------------------------|
| /estadísticas | GET | KPIs agregados: toneladas, ingresos, certificados |
|---------------|-----|---------------------------------------------------|

## B.5. Protocolo de seguridad y gestión de identidades

La seguridad del sistema ECODES-EDMUS se articula en múltiples capas. A nivel de red, los conectores EDC se comunican a través de HTTPS con certificados TLS mutuos, lo que garantiza la identidad de ambos extremos de la comunicación y el cifrado del tráfico. A nivel de protocolo, cada negociación de contrato genera un acuerdo digital con identificador único (`agreement_id`) que se almacena en el registro de contratos del conector y puede utilizarse como evidencia en caso de disputa. A nivel de aplicación, los tokens EDR tienen una duración limitada (configurada en 60 minutos en la implementación actual) y están vinculados a un activo específico, lo que impide la reutilización de credenciales para acceder a datos no autorizados.

La gestión de identidades de los participantes del espacio de datos sigue el modelo de Gaia-X Compliance, que requiere que cada organización participante emita una Self-Description verificable con sus atributos jurídicos (CIF, domicilio social, representante legal) y técnicos (endpoint del conector, certificados públicos). En la implementación actual de ECODES-EDMUS, las identidades se gestionan mediante un registro local de participantes con verificación manual; la evolución prevista hacia la Lista de Confianza de UNE 0087:2025 proporcionará un mecanismo de verificación automática de identidades conforme a estándares nacionales.

## B.6. Modelo de datos de los certificados CAES

Cada Certificado de Ahorro de Emisiones de CO<sub>2</sub> (CAES) generado por la plataforma se almacena como un objeto JSON con la siguiente estructura: identificador único del certificado (`cert_id`), timestamp de emisión, cantidad total de CO<sub>2</sub> certificada en kilogramos, precio por tonelada de CO<sub>2</sub> en euros (referenciado al EU ETS en el momento de la emisión), coste total del certificado, array de bloques horarios con los tickets individuales de cada bloque, factores de emisión utilizados (bus diésel y bus eléctrico en gCO<sub>2</sub>/km), distancia media por billete, fórmula de cálculo aplicada, desglose de CO<sub>2</sub> por bloque horario, hash de integridad del certificado (SHA-256), y URL pública de verificación.

El PDF del certificado, generado programáticamente con ReportLab, incluye ocho secciones: portada con marca ECODES y datos identificativos, resumen ejecutivo con KPIs principales, tabla comparativa de emisiones (escenario diésel vs. eléctrico), gráfico de barras con distribución horaria de la reducción certificada, tabla de detalle con tickets por bloque, sección de metodología con fórmula y factores, código QR de verificación enlazado al endpoint público, y pie de página con disclaimers jurídicos. El diseño visual del PDF sigue

la guía de marca de ECODES, con colores corporativos verdes y tipografía Roboto, y ha sido validado con compradores piloto para garantizar su aceptabilidad en contextos institucionales y corporativos.

# Anexo C. Convocatorias de financiación vigentes

## C.1. Convocatorias nacionales

Kit Espacios de Datos (Red.es): Convocatoria abierta hasta el 31 de marzo de 2026. Presupuesto de 60 millones de euros con más del 60% disponible a fecha de publicación de este informe. Régimen de concurrencia no competitiva por orden de llegada. Orden TDF/758/2025. Destinada a la incorporación de participantes y proveedores a espacios de datos elegibles, incluyendo los del Plan de Impulso, el Espacio Nacional de Datos de Salud y los de la Lista de Confianza (UNE 0087:2025). Gestionada por Red.es con apoyo del Centro de Referencia de Espacios de Datos (CRED).

Segunda Convocatoria de Demostradores y Casos de Uso de Espacios de Datos Sectoriales: Presupuesto de 75 millones de euros (16 M€ para centros demostradores, 58 M€ para casos de uso). Propuesta de resolución definitiva publicada en junio de 2025. Proyectos en ejecución con plazo de finalización vinculado al objetivo CID 180 (30 de junio de 2026).

Convocatoria de Productos y Servicios de Espacios de Datos: Presupuesto de 44 millones de euros para el sector tecnológico industrial. Resolución de concesión publicada en agosto de 2025. Destinada a crear, adaptar e innovar productos y servicios digitales para las necesidades de los espacios de datos.

## C.2. Convocatorias europeas relevantes

Digital Europe Programme: Financiación para el despliegue de espacios de datos comunes europeos y la adopción de tecnologías de IA. Convocatorias periódicas alineadas con la Estrategia Europea del Dato. Horizon Europe: Programa marco de investigación e innovación. Clusters relevantes: Cluster 5 (Clima, Energía y Movilidad) y Cluster 4 (Digital, Industria y Espacio). El proyecto deployEMDS se financia a través de este programa. CEF Digital: Connecting Europe Facility para infraestructuras digitales transfronterizas, incluyendo componentes de interoperabilidad para espacios de datos. InvestEU: Instrumento financiero para apoyar inversiones en infraestructura digital y sostenibilidad.

## C.4. Convocatorias regionales relevantes

A nivel autonómico, varias comunidades autónomas han lanzado convocatorias complementarias a las nacionales. En Aragón, la Estrategia de Especialización Inteligente (S3) incluye la digitalización sostenible como área prioritaria, y el Gobierno de Aragón financia proyectos de I+D+i en IA y sostenibilidad a través del Departamento de Economía,

Empleo e Industria. El programa FEDER Aragón 2021-2027 incluye ejes temáticos de transición digital y transición ecológica que pueden cofinanciar proyectos en la intersección de ambos ámbitos. En Cataluña, el CIDAI (Centre of Innovation for Data Tech and Artificial Intelligence) financia proyectos de IA aplicada con potencial de impacto social y ambiental. En País Vasco, la Red de Ciencia, Tecnología e Innovación (RVCTI) apoya desarrollos tecnológicos en energía y sostenibilidad a través de centros como Tecnalia e IK4.

La coordinación entre convocatorias nacionales, europeas y regionales es crítica para maximizar el impacto de la inversión. Un proyecto como ECODES-EDMUS puede financiar su infraestructura básica con el Kit Espacios de Datos (nacional), su componente de IA con convocatorias regionales de I+D+i (Gobierno de Aragón), y su escalado internacional con proyectos Horizon Europe. Esta estrategia de financiación multinivel reduce el riesgo financiero del proyecto y asegura la sostenibilidad económica durante las fases de desarrollo, pilotaje y escalado.

## C.3. Incentivos fiscales

Las empresas españolas pueden beneficiarse de deducciones fiscales por actividades de I+D+i (artículos 35 y 36 de la Ley del Impuesto sobre Sociedades) que cubren tanto el desarrollo de tecnologías de IA como la implementación de sistemas de gestión ambiental. La deducción por actividades de I+D puede alcanzar el 25% de los gastos del periodo (42% por exceso sobre la media de los dos años anteriores), y la deducción por innovación tecnológica es del 12%. Estos incentivos son acumulables con las subvenciones del Plan de Impulso de Espacios de Datos.

# Anexo D. Publicaciones y papers de referencia adicionales

Este anexo recoge publicaciones académicas adicionales consultadas durante la elaboración del informe, agrupadas por temática.

## D.1. Machine Learning y predicción de emisiones

Rolnick, D. et al. (2022). Tackling Climate Change with Machine Learning. ACM Computing Surveys, 55(2), 1-96. Este trabajo seminal identifica las áreas donde el machine learning puede contribuir de forma más significativa a la lucha contra el cambio climático, incluyendo la estimación de emisiones, la optimización de sistemas energéticos, la mejora del transporte y la agricultura sostenible.

Kaack, L.H. et al. (2022). Aligning artificial intelligence with climate change mitigation. Nature Climate Change, 12, 518-527. Los autores analizan cómo la IA puede alinearse con los objetivos de mitigación del cambio climático, distinguiendo entre aplicaciones de IA que reducen emisiones directamente (optimización de procesos), aplicaciones que mejoran la información para la toma de decisiones (predicción de emisiones) y la huella ambiental de los propios sistemas de IA.

Nguyen, T.H. et al. (2021). Predicting corporate carbon footprints for climate finance risk analyses: A machine learning approach. Energy Economics, 95, 105129. Este artículo desarrolla modelos de ML para predecir huellas de carbono corporativas a partir de datos financieros públicos, con aplicación directa al análisis de riesgos climáticos en el sector financiero.

## D.2. Espacios de datos y sostenibilidad

Otto, B. et al. (2022). Designing Data Spaces: The Ecosystem Approach to Competitive Advantage. Springer. Este libro de referencia proporciona el marco teórico y práctico para el diseño de espacios de datos, incluyendo consideraciones sobre gobernanza, modelos de negocio y aspectos técnicos del Eclipse Dataspace Connector.

Curry, E. et al. (2022). Data Spaces: Design, Deployment, and Future Directions. Springer. Los editores presentan una visión comprehensiva de los espacios de datos desde la perspectiva de las Real-time Linked Dataspaces, con capítulos dedicados a espacios de datos de movilidad, energía y Green Deal.

BDVA/DAIRO (2023). Strategic Research and Innovation Agenda. Big Data Value Association. El documento estratégico identifica la sostenibilidad como uno de los dominios clave donde los espacios de datos pueden generar valor, y propone líneas de investigación en la intersección de datos, IA y transición ecológica.

## D.3. Huella de carbono y regulación

WBCSD/PACT (2023). Pathfinder Framework v2.0. World Business Council for Sustainable Development. El framework de referencia para el intercambio de Product Carbon Footprint (PCF) en cadenas de suministro, adoptado por Catena-X y otros espacios de datos industriales.

EFRAG (2024). European Sustainability Reporting Standards (ESRS). Los estándares de reporte de sostenibilidad europeos, desarrollados bajo mandato de la CSRD, que definen los requisitos de divulgación de emisiones de Alcance 1, 2 y 3, creando la demanda regulatoria para las soluciones de IA y espacios de datos analizadas en este informe.

Science Based Targets initiative (2024). SBTi Corporate Net-Zero Standard v2.0. El estándar que requiere la inclusión del Alcance 3 en los objetivos de reducción cuando estas emisiones superen el 40% del total, lo que convierte la medición precisa del Alcance 3 en un requisito operativo para las empresas comprometidas con objetivos basados en ciencia.

## D.5. Verificación y aseguramiento de datos de emisiones

Gillenwater, M. (2022). A Beginner's Guide to Uncertainty of Measurement for Greenhouse Gas Accounting. Greenhouse Gas Management Institute. Esta guía establece los fundamentos para la cuantificación de la incertidumbre en los cálculos de emisiones de GEI, un aspecto crítico para los servicios de IA que deben proporcionar intervalos de confianza junto con sus predicciones de Alcance 3.

Busch, T. et al. (2022). Corporate carbon performance data: Quo vadis? Journal of Industrial Ecology, 26(1), 350-363. Los autores analizan la calidad y consistencia de los datos de emisiones corporativas disponibles en las principales bases de datos (CDP, Bloomberg, Refinitiv), identificando discrepancias significativas que los servicios de IA deben considerar al entrenarse con estas fuentes.

Eccles, R.G. y Strohle, J.C. (2018). Exploring Social Origins in the Construction of ESG Measures. MIT Sloan Working Paper. Este working paper examina cómo las diferencias en las metodologías de las agencias de rating ESG generan evaluaciones divergentes de las mismas empresas, un problema que los servicios de IA sobre espacios de datos podrían mitigar al acceder directamente a datos primarios en lugar de ratings derivados.

## D.6. Computación verde y centros de datos sostenibles

Masanet, E. et al. (2020). Recalibrating global data center energy-use estimates. *Science*, 367(6481), 984-986. Este artículo de referencia demuestra que el consumo energético de los centros de datos creció solo un 6% entre 2010 y 2018 pese al aumento exponencial de la carga de trabajo, gracias a mejoras en eficiencia de hardware y software. Las implicaciones para los servicios de IA sobre espacios de datos son claras: la huella de carbono marginal de añadir capacidad de cálculo para servicios climáticos es cada vez menor.

Dodge, J. et al. (2022). Measuring the Carbon Intensity of AI in Cloud Instances. *Proceedings of FAccT 2022*. Los autores desarrollan metodologías para medir la intensidad de carbono de cargas de trabajo de IA en instancias cloud, considerando la variabilidad temporal y geográfica de la intensidad de carbono de la red eléctrica, y proponen el concepto de carbon-aware scheduling que podría aplicarse a los servicios de IA sobre espacios de datos para minimizar su propia huella.

## D.4. IA y sostenibilidad: perspectivas adicionales

Henderson, P. et al. (2020). Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning. *Journal of Machine Learning Research*, 21(248), 1-43. Este trabajo propone un marco para el reporte sistemático de la huella energética y de carbono de los sistemas de ML, incluyendo herramientas como CodeCarbon y experiment-impact-tracker.

Strubell, E. et al. (2019). Energy and Policy Considerations for Deep Learning in NLP. *Proceedings of ACL 2019*. El artículo que inició el debate sobre el coste ambiental del entrenamiento de modelos de lenguaje de gran escala, estimando que el entrenamiento de un transformer grande puede emitir tanto CO<sub>2</sub> como cinco automóviles durante toda su vida útil.

Patterson, D. et al. (2022). The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink. *IEEE Computer*, 55(7), 18-28. Investigadores de Google demuestran que las mejoras en eficiencia de hardware, algoritmos y centros de datos están reduciendo rápidamente la huella de carbono del entrenamiento de modelos de ML, contradiciendo las predicciones catastrofistas sobre el crecimiento exponencial del consumo energético de la IA.

# Anexo E. Mapeo de oportunidades con fuentes de financiación

Este anexo presenta un mapeo detallado entre las doce oportunidades identificadas en el capítulo 7 y las fuentes de financiación disponibles a nivel nacional y europeo, facilitando la planificación de propuestas de proyecto.

| Oportunidad                      | Financiación nacional          | Financiación europea       | Plazo estimado |
|----------------------------------|--------------------------------|----------------------------|----------------|
| 1. Predicción Alcance 3 ML       | Kit ED / Productos y Servicios | Horizon Europe Cluster 4+5 | 12-18 meses    |
| 2. Matching LLM factores emisión | Kit ED / I+D+i fiscal          | Digital Europe             | 12 meses       |
| 3. Optimización transporte       | 2ª Conv. Demostradores         | deployEMDS / Horizon       | 18-24 meses    |
| 4. Certificación CO <sub>2</sub> | Kit ED (EDMUS)                 | CEF Digital                | 6-12 meses     |
| 5. Gemelos digitales             | Productos y Servicios          | Horizon Europe Cluster 4   | 24-36 meses    |
| 6. Auditoría CSRD                | I+D+i fiscal                   | Digital Europe             | 18-24 meses    |
| 7. PCF sectorial                 | Kit ED / Catena-X ES           | Horizon Europe Cluster 4   | 18-24 meses    |
| 8. Energía edificios             | Kit ED (Energía)               | Horizon Europe Cluster 5   | 12-18 meses    |
| 9. Agroalimentaria               | Kit ED (Agri)                  | AgriDataSpace / Horizon    | 18-24 meses    |
| 10. Aviación SAF                 | N/A directo                    | Horizon Europe Cluster 5   | 24-36 meses    |
| 11. Riesgos climáticos           | I+D+i fiscal                   | Horizon Europe Cluster 5   | 24-36 meses    |
| 12. Compensación verificada      | Kit ED                         | CEF Digital / InvestEU     | 12-18 meses    |

La estrategia óptima de financiación para un proyecto integral de servicios de IA sobre espacios de datos combina múltiples fuentes en cascada temporal. En la fase de concepto y diseño (meses 1-6), la financiación puede provenir de recursos propios de los centros tecnológicos y de convocatorias regionales de exploración tecnológica. En la fase de desarrollo y pilotaje (meses 6-18), el Kit Espacios de Datos y las deducciones fiscales por I+D proporcionan la base financiera. En la fase de escalado y explotación (meses 18-36), los proyectos europeos de Horizon Europe o Digital Europe aportan el presupuesto necesario para la validación multinacional y la transferencia de resultados. Los ingresos por servicios (AlaaS, certificados de reducción) comienzan a generar retorno a partir del mes 12-18, contribuyendo progresivamente a la sostenibilidad económica del proyecto sin dependencia exclusiva de financiación pública.

Es importante señalar que las convocatorias del Plan de Impulso de Espacios de Datos están vinculadas al cumplimiento de los hitos del Plan de Recuperación, Transformación y Resiliencia, con el objetivo CID 180 que establece como fecha límite el 30 de junio de 2026 para la finalización de los proyectos de la segunda convocatoria de Demostradores. Esto crea una ventana de oportunidad específica: los proyectos que demuestren resultados tangibles antes de esa fecha podrán posicionarse favorablemente para las futuras convocatorias del periodo 2027-2030, que se espera que continúen la línea de inversión en espacios de datos con presupuestos similares o superiores a los actuales.

Las oportunidades con mayor potencial de financiación inmediata son las que pueden acogerse al Kit Espacios de Datos (oportunidades 1, 2, 4, 7, 8, 9 y 12), dado que esta convocatoria permanece abierta y dispone de más del 60% de su presupuesto. Las oportunidades que requieren desarrollo experimental más intensivo (3, 5, 6, 10, 11) se orientan preferentemente a convocatorias europeas de Horizon Europe, que ofrecen presupuestos mayores y plazos de ejecución más amplios. En todos los casos, las deducciones fiscales por I+D+i (artículos 35-36 LIS) son aplicables como fuente complementaria de financiación.

# Anexo F. Casos de uso internacionales de referencia

Este anexo presenta una descripción detallada de casos de uso internacionales que sirven como referencia para el desarrollo de servicios de IA sobre espacios de datos orientados a la reducción de huella de carbono.

## F.1. Catena-X: Product Carbon Footprint en la industria automotriz alemana

Catena-X constituye el espacio de datos sectorial más avanzado de Europa, con más de 100 empresas participantes en su fase operativa. El caso de uso PCF (Product Carbon Footprint) permite a los fabricantes de automóviles conocer la huella de carbono específica de cada componente que adquieren, desde las baterías de litio hasta los neumáticos, pasando por los sistemas electrónicos y la carrocería. Los proveedores de nivel 1 (Tier 1) publican los PCF de sus productos como activos en el catálogo del espacio de datos, y los OEMs (BMW, Mercedes-Benz, Volkswagen) los consumen para calcular el PCF total del vehículo. La metodología sigue el Pathfinder Framework v2.0 de WBCSD/PACT, lo que garantiza comparabilidad entre marcas y proveedores. La IA se aplica en dos puntos: predicción de PCF para componentes de proveedores que aún no reportan (basada en datos financieros y clasificación industrial del componente), y optimización de la selección de proveedores considerando simultáneamente coste, calidad, plazo y huella de carbono.

## F.2. deployEMDS Milán: escenarios de neutralidad de carbono urbana

En el marco del proyecto deployEMDS financiado por Horizon Europe, la agencia milanesa AMAT ha desarrollado un caso de uso que combina datos de transporte público (ATM Milano), movilidad compartida (bicicletas y patinetes), tráfico privado y datos de calidad del aire en un espacio de datos federado. Un modelo de IA de escenarios permite a los planificadores urbanos simular el impacto de diferentes políticas de movilidad sobre las emisiones de CO<sub>2</sub>: qué ocurriría si se ampliase la zona de bajas emisiones, si se introdujese un peaje urbano, si se duplicase la frecuencia del metro en horas punta, o si se desplegasen 5.000 bicicletas eléctricas adicionales. Cada escenario genera una proyección de emisiones con intervalo de confianza, permitiendo a los responsables políticos tomar decisiones basadas en evidencia. El modelo utiliza datos de movilidad accedidos soberanamente a través del protocolo DSP, lo que garantiza que los datos de los

operadores privados (tiempos de viaje, ocupación) no se exponen al municipio más allá de los agregados necesarios para la simulación.

## **F.3. CO2 AI (BCG): automatización del reporte de emisiones**

CO2 AI, la plataforma de sostenibilidad desarrollada por Boston Consulting Group, representa el estado más avanzado de la automatización del cálculo de huella de carbono mediante IA en un contexto empresarial. La plataforma utiliza agentes de IA que procesan automáticamente datos de actividad (facturas, órdenes de compra, datos de ERP) y los emparejan con factores de emisión de múltiples bases de datos (DEFRA, EPA, ecoinvent, ADEME). Según BCG, el uso de agentes de IA reduce el tiempo de cálculo de la huella de carbono corporativa de semanas a horas, y mejora la precisión al eliminar errores humanos de clasificación. General Mills, una de las mayores empresas alimentarias de Estados Unidos, utiliza CO2 AI para integrar datos de carbono en sus sistemas y flujos de trabajo centrales, midiendo el impacto de proveedores individuales en la huella total de Alcance 3.

Aunque CO2 AI opera actualmente como plataforma centralizada (los clientes cargan sus datos en la plataforma de BCG), la convergencia con los espacios de datos soberanos es natural: en lugar de cargar datos, las empresas podrían publicarlos en un espacio de datos sectorial y autorizar al servicio de IA a accederlos a través del protocolo EDC, eliminando la necesidad de transferir información sensible a un tercero. Esta evolución de plataformas centralizadas hacia servicios federados sobre espacios de datos constituye una de las oportunidades estratégicas identificadas en el presente informe.

## **F.4. Energinet DataHub (Dinamarca): datos energéticos en tiempo real**

El DataHub de Energinet, el operador del sistema eléctrico danés, proporciona acceso en tiempo real a datos de consumo, producción e intensidad de carbono de la red eléctrica danesa. Con una granularidad de 5 minutos, estos datos permiten a los servicios de IA optimizar el consumo eléctrico de edificios, industrias y centros de datos en función de la disponibilidad de energía renovable, minimizando las emisiones de Alcance 2. El concepto de carbon-aware computing, que consiste en programar las cargas de trabajo computacionales intensivas (como el entrenamiento de modelos de ML) para los momentos de menor intensidad de carbono de la red eléctrica, ha sido adoptado por empresas como Microsoft y Google, y podría extenderse a través de espacios de datos europeos de energía para optimizar el consumo eléctrico a escala continental.

## F.5. PCAF y emisiones financiadas: el caso del sector bancario

La Partnership for Carbon Accounting Financials (PCAF) ha desarrollado el Global GHG Accounting and Reporting Standard for the Financial Industry, adoptado por más de 400 instituciones financieras a nivel mundial. El estándar requiere que los bancos y fondos de inversión calculen las emisiones asociadas a sus préstamos e inversiones (emisiones financiadas, categoría 15 del Alcance 3), lo que implica acceder a datos de emisiones de todas las empresas en su cartera. Un espacio de datos financiero que integre datos de emisiones corporativas (publicados por las empresas o estimados por servicios de IA) con datos de carteras de inversión podría automatizar completamente el cálculo de emisiones financiadas, que actualmente requiere meses de trabajo manual en cada institución financiera.

En España, entidades como BBVA, CaixaBank y Banco Santander han comenzado a reportar sus emisiones financiadas conforme a PCAF, enfrentando dificultades significativas en la obtención de datos de emisiones de sus clientes corporativos, especialmente pymes. Un servicio de IA que combine predicciones ML de emisiones (a partir de datos financieros disponibles en los registros bancarios) con datos específicos accedidos a través de un espacio de datos sectorial podría proporcionar a los bancos españoles inventarios de emisiones financiadas con una precisión significativamente superior a los métodos basados exclusivamente en factores sectoriales medios. Este caso de uso está alineado con la regulación del Banco Central Europeo (BCE) sobre riesgos climáticos y con la Taxonomía de la UE para finanzas sostenibles.

La experiencia danesa demuestra que los datos energéticos en tiempo real, publicados como activos abiertos en una infraestructura de datos federada, pueden habilitar un ecosistema completo de servicios de IA de reducción de emisiones sin que el operador del sistema tenga que desarrollar esos servicios él mismo. El rol del operador se limita a publicar datos de calidad con metadatos estandarizados, y el mercado de servicios de IA se desarrolla orgánicamente sobre esa base de datos. Este modelo de plataforma habilitadora es directamente aplicable a los espacios de datos españoles de energía, movilidad y agroalimentación.

# Anexo G. Metodología detallada de cálculo de emisiones de Alcance 3

Este anexo presenta las metodologías de cálculo de emisiones de Alcance 3 recomendadas por el GHG Protocol Corporate Value Chain Standard, con indicación de los datos necesarios, las fuentes de factores de emisión y las oportunidades de automatización mediante servicios de IA sobre espacios de datos.

## G.1. Métodos de cálculo según el GHG Protocol

El GHG Protocol Corporate Value Chain Standard establece tres métodos jerárquicos para el cálculo de emisiones de Alcance 3, ordenados de menor a mayor precisión. El método basado en gasto (spend-based method) utiliza datos financieros de compras multiplicados por factores de emisión económicos (kgCO<sub>2</sub>e por euro gastado), clasificados por sector NACE o SIC. Es el método más accesible pero menos preciso, ya que los factores económicos son promedios sectoriales que no reflejan la eficiencia específica de cada proveedor. El método basado en actividad media (average-data method) utiliza datos físicos de actividad (kg de material comprado, km de transporte, kWh de energía) multiplicados por factores de emisión genéricos del sector. Ofrece mayor precisión que el método por gasto, pero sigue dependiendo de promedios. El método específico del proveedor (supplier-specific method) utiliza datos de emisiones proporcionados directamente por el proveedor, ya sea como huella de carbono total asignada proporcionalmente o como Product Carbon Footprint del bien o servicio específico. Este método ofrece la mayor precisión, pero requiere la colaboración activa del proveedor.

Los servicios de IA sobre espacios de datos pueden operar en los tres niveles metodológicos: automatizando la clasificación de gastos y la asignación de factores económicos para el método basado en gasto (Oportunidad 2: matching LLM), prediciendo emisiones a partir de datos financieros cuando los datos de actividad no están disponibles (Oportunidad 1: predicción ML), y facilitando el acceso soberano a datos específicos del proveedor a través del espacio de datos (Oportunidad 4: certificación). La evolución natural es que las empresas comiencen con el método basado en gasto automatizado por IA y progresen hacia datos específicos del proveedor a medida que sus proveedores se incorporen al espacio de datos, mejorando progresivamente la precisión de su inventario de Alcance 3.

## G.2. Fuentes de factores de emisión

| Base de datos        | Cobertura           | Tipo de factores    | Acceso             |
|----------------------|---------------------|---------------------|--------------------|
| DEFRA/BEIS (UK)      | Global (énfasis UK) | Actividad y gasto   | Gratuito (Excel)   |
| EPA (US)             | Global (énfasis US) | Actividad y gasto   | Gratuito (API)     |
| ecoinvent            | Global              | Ciclo de vida (LCA) | Licencia comercial |
| ADEME Base Carbone   | Francia/Europa      | Actividad           | Gratuito (API)     |
| ExioML (Nature 2026) | Global (MRIO)       | ML-optimizado       | Abierto (CC-BY)    |
| GaBi                 | Global              | Ciclo de vida (LCA) | Licencia comercial |
| IPCC EF Database     | Global              | País/sector         | Gratuito (web)     |
| GEMIS                | Europa              | Energía/transporte  | Gratuito           |

La selección del factor de emisión apropiado es una de las decisiones metodológicas más críticas en el cálculo de Alcance 3. Un LLM con RAG sobre un corpus indexado de factores de emisión puede automatizar esta selección, considerando la geografía del proveedor, el proceso productivo, la temporalidad del dato y la granularidad disponible. El dataset ExioML, publicado en 2026, está específicamente diseñado para aplicaciones de ML e incluye descomposiciones según la identidad de Kaya que facilitan la interpretabilidad de los modelos.

## G.3. Cálculo detallado por categorías materiales

**Categoría 1 (Bienes y servicios adquiridos):** Es típicamente la categoría más significativa en términos de emisiones. El cálculo requiere datos de compras (cantidades físicas o importes monetarios) y factores de emisión por producto o sector. Un servicio de IA sobre espacio de datos puede acceder automáticamente a las facturas de compra publicadas como activos por la empresa solicitante, clasificar cada línea de factura según la taxonomía NACE/SIC mediante NLP, y asignar el factor de emisión más preciso disponible. Si el proveedor participa en el espacio de datos y ha publicado su propio PCF, el servicio utilizará datos específicos del proveedor en lugar de promedios sectoriales, mejorando automáticamente la precisión sin intervención manual.

**Categoría 4 (Transporte y distribución aguas arriba):** Requiere datos de distancias de transporte, modos de transporte y carga transportada. Los espacios de datos de movilidad (EMDS, EDMUS) proporcionan la infraestructura para acceder a datos reales de transporte

(GTFS, telemetría de flota, registros de envíos) que pueden sustituir las estimaciones basadas en distancias medias. Un servicio de IA de optimización multimodal (Oportunidad 3) puede, además de calcular las emisiones actuales, recomendar alternativas de transporte con menor huella para envíos futuros.

**Categoría 7 (Desplazamiento de empleados):** Requiere datos de patrones de movilidad de los empleados (domicilio, modo de transporte, frecuencia). En el contexto de un espacio de datos de movilidad, los datos de billeteaje de transporte público utilizarse para cuantificar con precisión la proporción de empleados que utilizan transporte público electrificado, generando reducciones certificables de emisiones de Alcance 3 para la empresa empleadora. Este es precisamente el modelo que ECODES-EDMUS implementa.

**Categoría 11 (Uso de productos vendidos):** Requiere datos sobre la eficiencia energética y el patrón de uso de los productos durante su vida útil. Para fabricantes de automóviles, esta categoría puede representar más del 80% del total de emisiones. Los datos de telemetría de vehículos conectados, publicados soberanamente en espacios de datos de movilidad, podrían sustituir las estimaciones basadas en ciclos de conducción estándar (WLTP) por datos reales de consumo, mejorando drásticamente la precisión del cálculo.

## G.4. Cuantificación de la incertidumbre

Todo cálculo de emisiones de Alcance 3 está sujeto a incertidumbre, que debe cuantificarse y comunicarse conforme a los ESRS. Las fuentes de incertidumbre incluyen: la variabilidad de los factores de emisión (que dependen de la geografía, la tecnología y el año de referencia), la imprecisión de los datos de actividad (estimaciones de distancias, pesos o consumos), la representatividad de los datos específicos del proveedor (que pueden no cubrir todos los procesos), y la incertidumbre del modelo de IA (que debe cuantificarse mediante intervalos de confianza o predicciones probabilísticas).

La evolución de los estándares de cuantificación de incertidumbre en el contexto de la IA climática es un área de investigación activa. El grupo de trabajo TC309 de ISO (Governance of Organizations) está desarrollando directrices para la evaluación de la calidad de datos ambientales generados por sistemas de IA, que incluirán requisitos específicos sobre la cuantificación y comunicación de la incertidumbre. En paralelo, la European Financial Reporting Advisory Group (EFRAG) está elaborando guías de implementación para el ESRS E1 que abordan explícitamente el uso de estimaciones basadas en modelos para las emisiones de Alcance 3, incluyendo requisitos sobre la divulgación de las metodologías de estimación, las fuentes de datos utilizadas y los intervalos de confianza asociados. Los servicios de IA sobre espacios de datos deben anticipar estos requisitos en su diseño, incorporando desde el inicio mecanismos de trazabilidad, cuantificación de incertidumbre y generación automatizada de la documentación metodológica exigida por los estándares.

En conclusión, la automatización del cálculo de Alcance 3 mediante servicios de IA sobre espacios de datos no es solo una cuestión de eficiencia operativa, sino una condición necesaria para que el volumen creciente de empresas sujetas a la CSRD pueda cumplir sus obligaciones de reporte en plazos razonables y con una precisión aceptable. Los métodos manuales actuales no escalan al ritmo que la regulación exige, y los espacios de

datos soberanos proporcionan la infraestructura que los servicios de IA necesitan para acceder a los datos distribuidos de las cadenas de valor sin comprometer la confidencialidad ni la soberanía de los participantes. El presente anexo metodológico aspira a servir como referencia técnica para los equipos que implementen estos servicios en el ecosistema español y europeo de espacios de datos.

Los servicios de IA sobre espacios de datos pueden abordar la cuantificación de la incertidumbre de forma nativa. Los modelos de ensemble como XGBoost y Random Forest proporcionan estimaciones de importancia de variables que permiten identificar las fuentes de incertidumbre dominantes. Los enfoques bayesianos (Bayesian Neural Networks, Gaussian Processes) generan distribuciones de probabilidad sobre las predicciones en lugar de estimaciones puntuales. Y la comparación automática entre predicciones basadas en datos financieros (método por gasto) y predicciones basadas en datos de actividad (método por actividad) proporciona una medida directa de la convergencia entre métodos, que constituye un indicador robusto de la precisión del cálculo.



Con el apoyo de:



# GRACIAS

## Contacto

---

 [www.ecodes.org](http://www.ecodes.org)

 +34 976 29 82 82

 Plaza San Bruno, 9, 1º Oficinas, 50001 Zaragoza.