

Diseños para el uso de la IA en la reducción de la huella de carbono

Una propuesta de especificaciones de
diseño en diferentes sistemas IA para el
cálculo y reducción de la huella de
carbono

ABRIL 2025

AGRADECIMIENTOS



Al Instituto Tecnológico de Aragón por su perspectiva y capacidad para imaginar el futuro de la Inteligencia Artificial aplicada a la reducción de la huella de carbono. Han sido una contribución esencial para la elaboración del presente informe.

EDITA: ECODES

DISEÑO GRÁFICO: ECODES

FECHA: Abril 2025

Índice

Resumen ejecutivo	5
1. Introducción	7
1.1. Análisis de tipos de Inteligencia Artificial para la descarbonización	7
1.1.1. Aprendizaje Supervisado (Supervised Learning)	7
1.1.2. Aprendizaje No Supervisado (Unsupervised Learning)	8
1.1.3. Aprendizaje por Refuerzo (Reinforcement Learning)	9
1.1.4. Machine Learning (ML)	9
1.1.5. Modelos de Lenguaje Extensos (LLMs)	10
1.1.6. Modelos de Lenguaje Multimodales	10
1.1.7. Generación Aumentada por Recuperación (RAG)	11
1.1.8. Sistemas Multiagente (MAS)	11
2. Propuestas de diseño de herramientas específicas	13
2.1. Asistente para la Interpretación y Cumplimiento de Legislación Ambiental	13
2.2. Asistente para Recomendaciones de Acciones de Mitigación Personalizadas	14
2.3. Sistema de Generación de Informes a partir de Consultorías Grabadas	14
3. Diseño detallado: Asistente General para la Gestión Integral de la Descarbonización ..	16
3.1. Descripción funcional detallada	16
3.2. Arquitectura general de la solución orientada a IA	17
3.3. Módulos principales y su interacción con la IA	19
3.4. Especificaciones técnicas detalladas	21
3.4.1. Especificaciones Generales de la Plataforma	21
3.4.2. Especificaciones del Módulo de Ingesta de Datos (ID)	22
3.4.3. Especificaciones del Módulo de Cálculo de Huella de Carbono	25
3.4.4. Especificaciones del Módulo de Generación de Recomendaciones (IA Centric)	28
3.4.5. Especificaciones del Módulo de Elaboración de Memorias y Planes	31
4. Diseño avanzado: Agente IA Colaborativo para la optimización de la descarbonización en Ecosistemas Multiagente (AIC-MAS)	34

4.1.	Visión y objetivos estratégicos del Agente IA Colaborativo (AIC-MAS).....	34
4.2.	Arquitectura conceptual del Agente IA Colaborativo (AIC-MAS) y su ecosistema.	35
4.3.	Capacidades fundamentales detalladas del Agente IA Colaborativo (AIC-MAS) ..	37
4.4.	Tecnologías IA y habilitadoras clave para el AIC-MAS	38
4.5.	Desafíos significativos y líneas de investigación futuras.....	39
5.	Conclusiones	41
5.1.	El Asistente General para la Gestión Integral de la Descarbonización (AGID): la democratización de la acción climática empresarial	41
5.2.	El Asistente para la Interpretación y Cumplimiento de Legislación Ambiental (ALA): navegando la complejidad regulatoria	42
5.3.	El Agente IA Colaborativo en Entornos Multiagente (AIC-MAS): hacia la descarbonización sistémica y cooperativa.....	42
5.4.	Conclusiones generales e implicaciones transversales	43
6.	Anexo. Casos de uso sectoriales.....	45
6.1.	Anexo A. Casos de uso sectoriales para herramientas de IA en descarbonización	45
A.1.	Sector Manufacturero (Industria General).....	45
A.2.	Sector Energético (Generación, Transmisión y Distribución)	46
A.3.	Sector Transporte y Logística	46
A.4.	Sector Agrícola y Alimentario	47
A.5.	Sector Construcción y Promoción Inmobiliaria.....	48
A.6.	Sector Retail (Comercio Minorista) y Bienes de Consumo.....	49
A.7.	Sector TIC (Tecnologías de la Información y Comunicación)	50
A.8.	Sector Sanitario	50
A.9.	Sector Turismo y Hostelería	51
A.10.	Sector Público y Planificación Urbana	52

Resumen ejecutivo

El presente informe aborda el desafío crítico de la descarbonización empresarial mediante la propuesta y especificación detallada de soluciones innovadoras basadas en Inteligencia Artificial (IA). Ante la creciente urgencia climática y la complejidad de la transición hacia una economía baja en carbono, la IA se presenta como un catalizador fundamental, capaz de procesar vastas cantidades de datos, optimizar procesos, generar conocimiento accionable y facilitar la toma de decisiones estratégicas.

Se analizan diversos paradigmas de IA –desde el aprendizaje supervisado y no supervisado hasta los Modelos de Lenguaje Extensos (LLMs), Modelos Multimodales (VLM), Generación Aumentada por Recuperación (RAG) y Sistemas Multiagente (MAS)–, explorando su aplicabilidad directa en la reducción de la huella de carbono. A partir de este análisis, se conceptualizan tres herramientas progresivamente ambiciosas:

1. **Asistente para la Interpretación y Cumplimiento de Legislación Ambiental (ALA):** Conceptualizado como una herramienta especializada, el ALA emplearía LLMs y RAG para ofrecer respuestas claras y actualizadas sobre la compleja normativa ambiental y de reporte de emisiones. Su objetivo es democratizar el acceso al conocimiento legal, reduciendo riesgos de incumplimiento y facilitando la planificación estratégica.
2. **Asistente General para la Gestión Integral de la Descarbonización (AGID):** Sus especificaciones técnicas describen una plataforma SaaS integral diseñada para automatizar el ciclo completo de gestión de la huella de carbono. Entre sus capacidades clave, impulsadas por IA, destacan:
 - **Ingesta Inteligente de Datos:** Utilización de OCR avanzado, NLP y VLM para la extracción automática de información de facturas y otros documentos.
 - **Cálculo Preciso de Huella:** Adhesión a estándares (GHG Protocol, ISO 14064) para Alcances 1, 2 y un soporte para Alcance 3, con una base de factores de emisión actualizable.
 - **Recomendaciones de Mitigación Personalizadas:** Un motor híbrido de IA (ML/LLM con RAG) analiza la huella y propone medidas específicas, cuantificando su impacto, coste y ROI, y explicando su justificación.
 - **Generación Asistida de Informes y Planes:** LLMs fine-tuned para redactar memorias de sostenibilidad y planes de reducción, integrando datos y visualizaciones.

Las más de 100 especificaciones técnicas detalladas para el AGID cubren desde la arquitectura de microservicios en la nube, la gestión de datos y los módulos funcionales, hasta la interfaz de usuario, la seguridad y el ciclo de vida de los modelos de IA, asegurando una solución robusta, escalable y fiable.

3. **Agente IA Colaborativo para la Optimización de la Descarbonización en Ecosistemas Multiagente (AIC-MAS):** Representa una visión a medio y largo plazo de una IA avanzada capaz de operar en entornos empresariales complejos y conectados. Este agente se enlazaría continuamente con bases de datos diversas (legislación, espacios de datos de movilidad, industriales, mix eléctricos) y los sistemas de otras empresas (agentes IA). Su función sería colaborar para que las decisiones

colectivas (ej. en cadenas de suministro, simbiosis industriales, gestión de redes energéticas) minimicen la huella de carbono global (Alcances 1, 2 y 3), considerando el cálculo, la mitigación y la auditoría de forma integrada. Además, el AIC-MAS podría generar recomendaciones estratégicas para administraciones públicas y facilitar un seguimiento global del progreso en descarbonización.

El informe también incluye un anexo con **diez casos de uso sectoriales** (Anexo A), ilustrando cómo estas herramientas podrían aplicarse en la manufactura, energía, transporte, agricultura, construcción, retail, TIC, sanidad, turismo y el sector público, detallando beneficios específicos y desafíos de implementación.

Las conclusiones derivadas de las especificaciones técnicas resaltan la centralidad de los datos, el auge de los LLMs y RAG como tecnologías transversales, la necesidad de enfoques híbridos de IA y la evolución desde la optimización individual hacia una inteligencia colaborativa para la sostenibilidad. Se reconoce que la materialización de estas soluciones, especialmente las más avanzadas como el AIC-MAS, implica desafíos técnicos, éticos y de gobernanza significativos.

Las propuestas aquí detalladas constituyen una hoja de ruta posible para aprovechar el poder de la IA, impulsando una descarbonización más eficiente, precisa y sistémica, y reafirma su compromiso con la investigación y el desarrollo de tecnologías que contribuyan a un futuro sostenible.

1. Introducción

La transición hacia una economía baja en carbono es uno de los desafíos más apremiantes de nuestro tiempo. Las organizaciones, independientemente de su tamaño o sector, se enfrentan a la creciente necesidad de medir, reducir y reportar su huella de carbono, impulsadas tanto por la regulación como por la demanda de los consumidores y los inversores. La Inteligencia Artificial (IA) emerge como una tecnología catalizadora con el potencial de transformar radicalmente la forma en que abordamos la descarbonización, ofreciendo herramientas para analizar grandes volúmenes de datos, optimizar procesos, predecir escenarios y facilitar la toma de decisiones informadas.

Este informe, elaborado en colaboración con el Instituto Tecnológico de Aragón, tiene como objetivo proporcionar una visión técnica detallada de cómo la IA puede ser aplicada para crear soluciones efectivas para la reducción de la huella de carbono. Se explorarán diferentes enfoques de IA y se propondrán herramientas concretas, culminando con propuestas de diseño y con las especificaciones técnicas para un asistente general integral.

1.1. Análisis de tipos de Inteligencia Artificial para la descarbonización

La Inteligencia Artificial (IA) no es una tecnología monolítica, sino un campo diverso con múltiples enfoques y técnicas. Cada una de estas posee características únicas que pueden ser aprovechadas para abordar los complejos desafíos de la descarbonización. A continuación, se describen algunos tipos de IA y se profundiza en su potencial aplicación en herramientas para la reducción de la huella de carbono:

1.1.1. APRENDIZAJE SUPERVISADO (SUPERVISED LEARNING)

El aprendizaje supervisado es un paradigma de la IA donde los modelos se entrenan utilizando conjuntos de datos previamente etiquetados. Esto significa que cada dato de entrada en el conjunto de entrenamiento está asociado con una etiqueta o resultado correcto. El objetivo es que el modelo aprenda una función de mapeo que pueda predecir la etiqueta de salida para nuevos datos no vistos.

Posibilidades para la descarbonización

Predicción de consumo energético: Esta aplicación es crucial para la optimización energética. Mediante el entrenamiento de modelos (como regresiones, SVM, o redes neuronales) con datos históricos detallados de consumo energético (electricidad, gas, etc.) y variables exógenas influyentes (condiciones climáticas, niveles de producción industrial, ocupación de edificios, precios de la energía), es posible generar predicciones precisas de la demanda futura. Estas predicciones permiten a las empresas optimizar la compra de energía, programar el uso de fuentes de energía renovable intermitentes de manera más eficiente, y dimensionar adecuadamente la capacidad de generación propia, evitando tanto el derroche por sobreproducción como la escasez.

Clasificación de fuentes de emisión: La correcta identificación y categorización de las fuentes de emisión es fundamental para un cálculo preciso de la huella de carbono. Se pueden entrenar modelos de clasificación para analizar automáticamente datos provenientes de diversas fuentes, como facturas de proveedores (energía, combustible, agua, gestión de residuos), datos de sensores IoT (monitoreo de procesos, vehículos) o informes de actividad. Por ejemplo, un modelo podría aprender a distinguir si el consumo en una factura de combustible corresponde a una flota de vehículos (Alcance 1) o a un generador de emergencia, o si ciertos productos químicos adquiridos son precursores de emisiones fugitivas. Esto agiliza enormemente la recopilación de datos y reduce errores manuales.

Estimación de emisiones: En muchos casos, las emisiones de Gases de Efecto Invernadero (GEI) no se miden directamente, sino que se estiman a partir de datos de actividad y factores de emisión. El aprendizaje supervisado puede entrenar modelos para predecir estas emisiones basándose en una variedad de parámetros. Por ejemplo, se podría desarrollar un modelo que estime las emisiones de CO₂ de una flota de vehículos a partir de los kilómetros recorridos, el tipo de vehículo, el tipo de combustible y el estilo de conducción (si se dispone de datos telemáticos). De manera similar, se pueden estimar las emisiones de procesos industriales basadas en el consumo de kWh o el volumen de producción, aprendiendo la relación entre estos parámetros y las emisiones resultantes.

1.1.2. APRENDIZAJE NO SUPERVISADO (UNSUPERVISED LEARNING)

A diferencia del aprendizaje supervisado, el aprendizaje no supervisado trabaja con conjuntos de datos que no tienen etiquetas predefinidas. El objetivo principal de estas técnicas es explorar la estructura inherente de los datos para descubrir patrones, agrupaciones o anomalías por sí mismas.

Posibilidades para la descarbonización

Detección de anomalías: Los algoritmos de detección de anomalías (como Isolation Forest o One-Class SVM) pueden analizar series temporales de consumo energético o flujos de emisiones para identificar patrones inusuales o desviaciones significativas del comportamiento normal. Estos patrones anómalos pueden ser indicativos de fugas de energía, mal funcionamiento de equipos, ineficiencias operativas no detectadas previamente, o incluso errores en la medición de datos. La detección temprana de estas anomalías permite una intervención rápida, evitando el desperdicio de energía y la generación innecesaria de emisiones.

Segmentación de perfiles de consumo: Mediante técnicas de clustering (como K-Means o DBSCAN), es posible agrupar automáticamente usuarios, edificios, procesos industriales o incluso clientes con comportamientos de consumo energético o perfiles de emisión similares. Esta segmentación permite a las empresas comprender mejor la diversidad de sus fuentes de emisión y aplicar estrategias de reducción de huella de carbono más personalizadas y efectivas para cada segmento, en lugar de enfoques genéricos. Por ejemplo, se podrían identificar grupos de edificios con alto consumo nocturno para investigar posibles causas.

Identificación de patrones ocultos: El aprendizaje no supervisado puede descubrir correlaciones y relaciones complejas entre diferentes variables operativas y su impacto en la huella de carbono que no serían evidentes mediante un análisis manual. Por ejemplo, podría revelar cómo ciertas combinaciones de procesos productivos, turnos de trabajo o condiciones

ambientales específicas se correlacionan con picos de emisión, ofreciendo nuevas perspectivas para la optimización y la reducción de emisiones que no se habían considerado.

1.1.3. APRENDIZAJE POR REFUERZO (REINFORCEMENT LEARNING)

El aprendizaje por refuerzo es un área de la IA donde un agente aprende a tomar decisiones interactuando con un entorno. El agente recibe recompensas o penalizaciones en función de las acciones que realiza, y su objetivo es aprender una política (una estrategia de toma de decisiones) que maximice la recompensa acumulada a lo largo del tiempo.

Posibilidades para la descarbonización

Optimización de sistemas energéticos: Los agentes de aprendizaje por refuerzo pueden ser entrenados para controlar de forma inteligente y dinámica sistemas complejos como la climatización (HVAC) en grandes edificios, la operación de redes eléctricas inteligentes (smart grids) con generación distribuida y almacenamiento, o la optimización de procesos industriales intensivos en energía. El agente aprende a ajustar los parámetros del sistema (temperatura, flujos, velocidades) en tiempo real para minimizar el consumo energético y, por ende, las emisiones, mientras se mantienen los niveles de confort, la estabilidad de la red o la calidad del producto.

Gestión de flotas de vehículos: En la logística y el transporte, el aprendizaje por refuerzo puede optimizar la gestión de flotas de vehículos. Un agente podría aprender las rutas más eficientes energéticamente (considerando tráfico, topografía, tipo de vehículo), los modos de conducción óptimos para cada segmento de ruta, o las estrategias de carga para vehículos eléctricos (cuándo y dónde cargar para minimizar costes y emisiones, considerando la disponibilidad de energía renovable). Esto puede llevar a reducciones significativas en el consumo de combustible o electricidad y las emisiones asociadas.

1.1.4. MACHINE LEARNING (ML)

Machine Learning es el campo más amplio de la IA que engloba las técnicas anteriores (supervisado, no supervisado, refuerzo) y otras muchas. Su aplicación en la descarbonización es transversal y fundamental, ya que proporciona las herramientas para convertir datos brutos en conocimiento accionable.

Posibilidades para la descarbonización: El ML es el motor que impulsa muchas de las soluciones específicas. Sus capacidades se extienden desde el **procesamiento de datos inicial**, como la extracción automática de información relevante de facturas energéticas y documentos de transporte mediante técnicas de Procesamiento de Lenguaje Natural (NLP) y Reconocimiento Óptico de Caracteres (OCR), hasta la **construcción de modelos predictivos sofisticados** para la eficiencia energética en edificios e industrias. Además, el ML es clave para la **optimización de la cadena de suministro**, permitiendo modelar y reducir las emisiones indirectas (Alcance 3) al identificar los eslabones con mayor impacto. Finalmente, los modelos de ML pueden potenciar sistemas de **recomendación de medidas de mitigación**, sugiriendo las acciones más coste-efectivas y con mayor potencial de reducción de emisiones, personalizadas para el contexto de cada organización.

1.1.5. MODELOS DE LENGUAJE EXTENSOS (LLMS)

Los LLMS son modelos de IA, específicamente redes neuronales profundas, entrenados con cantidades masivas de datos textuales. Han demostrado una capacidad sobresaliente para comprender, generar y manipular el lenguaje humano de manera coherente y contextual.

Posibilidades para la descarbonización

Asistentes conversacionales (chatbots): Se pueden desarrollar chatbots impulsados por LLMS para actuar como asistentes virtuales. Estos asistentes pueden guiar a los usuarios, especialmente a aquellos con menos experiencia técnica, a través del complejo proceso de cálculo de la huella de carbono. Pueden responder preguntas frecuentes sobre metodologías, normativas ambientales aplicables (como el GHG Protocol o la ISO 14064), explicar conceptos técnicos (ej. Alcance 1, 2 y 3) de forma sencilla, y ayudar a localizar los factores de emisión correctos.

Generación de informes: Una de las tareas más laboriosas en la gestión de la sostenibilidad es la redacción de informes. Los LLMS pueden automatizar parcialmente la generación de secciones de memorias de sostenibilidad, informes de huella de carbono o planes de reducción. A partir de datos estructurados (resultados del cálculo de la huella, listas de medidas de mitigación) y directrices predefinidas, un LLM puede redactar borradores de texto coherentes, descriptivos y ajustados a los estándares de reporte, ahorrando un tiempo considerable.

Análisis de documentos: Las empresas deben navegar por una gran cantidad de documentación: textos legales y normativos sobre emisiones, artículos científicos sobre nuevas tecnologías de descarbonización, informes sectoriales, etc. Los LLMS pueden procesar y analizar estos documentos para extraer información clave, resumir contenidos extensos, identificar obligaciones regulatorias o detectar tendencias relevantes, facilitando la toma de decisiones informadas y el cumplimiento normativo.

1.1.6. MODELOS DE LENGUAJE MULTIMODALES

Estos modelos representan una evolución de los LLMS, capaces de procesar y relacionar información proveniente de múltiples modalidades, como texto, imágenes, audio y vídeo. Esta capacidad abre nuevas vías para la recopilación y análisis de datos en el contexto de la descarbonización.

Posibilidades para la descarbonización

Análisis de imágenes satelitales y aéreas: Los modelos multimodales pueden analizar imágenes de satélite o drones para evaluar la deforestación (una fuente importante de emisiones), monitorizar cambios en el uso del suelo (ej. expansión urbana sobre terrenos agrícolas), verificar la instalación y el rendimiento de grandes plantas de energía solar o parques eólicos, o incluso estimar la biomasa y el secuestro de carbono en áreas forestales.

Interpretación de datos de sensores visuales: En el ámbito industrial o de edificios, las imágenes térmicas (termografías) pueden ser analizadas por modelos multimodales para detectar fugas de calor en aislamientos, equipos o conducciones, identificando así oportunidades de ahorro energético. Del mismo modo, pueden analizar imágenes de cámaras en plantas de tratamiento de residuos para identificar y clasificar tipos de residuos, optimizando los procesos de reciclaje y valorización.

Procesamiento de información de vídeos/audio: Las grabaciones de vídeo o audio de consultorías de descarbonización, inspecciones visuales de instalaciones o entrevistas con personal operativo contienen información valiosa que a menudo es difícil de estructurar. Los modelos multimodales pueden transcribir el audio, identificar entidades relevantes (equipos mencionados, datos de consumo, problemas detectados), analizar expresiones visuales de problemas (ej. fugas visibles) y extraer datos clave para generar resúmenes o alimentar bases de conocimiento sobre oportunidades de mejora y reducción de emisiones.

1.1.7. GENERACIÓN AUMENTADA POR RECUPERACIÓN (RAG)

La técnica RAG combina el poder generativo de los LLMs con la capacidad de recuperar información específica de una base de conocimiento externa. En lugar de depender únicamente de la información internalizada durante su entrenamiento, el LLM consulta esta base de datos para obtener datos relevantes y actualizados antes de generar una respuesta.

Posibilidades para la descarbonización

Acceso a información actualizada y fiable: Para cuestiones críticas como la descarbonización, la precisión y actualidad de la información son vitales. RAG permite que un LLM acceda a bases de datos constantemente actualizadas de normativas ambientales (nacionales e internacionales), factores de emisión oficiales, las últimas investigaciones sobre tecnologías limpias, o las mejores prácticas sectoriales. Esto asegura que las respuestas y recomendaciones proporcionadas por el sistema de IA sean precisas, contextualizadas y basadas en la evidencia más reciente, minimizando el riesgo de "alucinaciones" o información desactualizada del LLM.

Personalización avanzada de recomendaciones: Al planificar acciones de mitigación, la especificidad es clave. Utilizando RAG, un LLM puede consultar una base de conocimiento detallada de medidas de mitigación (con sus costes, potenciales de reducción, requisitos técnicos, casos de éxito) y cruzar esta información con el perfil específico de una empresa (sector, tamaño, principales fuentes de emisión, recursos disponibles). El resultado son recomendaciones de descarbonización altamente personalizadas y priorizadas según la relevancia y viabilidad para esa organización en particular.

1.1.8. SISTEMAS MULTIAGENTE (MAS)

Un Sistema Multiagente está compuesto por múltiples agentes autónomos e inteligentes que interactúan entre sí y con su entorno. Cada agente tiene sus propios objetivos y capacidades, y la inteligencia global del sistema emerge de estas interacciones coordinadas o competitivas.

Posibilidades para la descarbonización

Simulación de mercados de carbono y políticas energéticas: Los MAS pueden utilizarse para crear modelos de simulación complejos que representen a los diferentes actores de un mercado de carbono (empresas, reguladores, inversores) o de un sistema energético. Permiten estudiar cómo interactúan estos actores bajo diferentes escenarios de precios del carbono, políticas de incentivos a las renovables o regulaciones de emisiones, ayudando a diseñar políticas más efectivas y a prever sus impactos.

Optimización descentralizada de la energía: En el contexto de las comunidades energéticas locales o las micro-redes inteligentes, donde coexisten múltiples pequeños

productores (ej. paneles solares en tejados) y consumidores, los MAS pueden coordinar de forma descentralizada la producción, el almacenamiento y el consumo de energía. Cada participante puede ser modelado como un agente que busca optimizar sus propios objetivos (ej. minimizar costes, maximizar autoconsumo) mientras contribuye a la estabilidad y eficiencia del sistema global, promoviendo un uso más eficiente de los recursos renovables locales.

Gestión de cadenas de suministro complejas y sostenibles: La huella de carbono de un producto a menudo se extiende a lo largo de cadenas de suministro globales y complejas (Alcance 3). Los MAS pueden modelar estas cadenas, donde cada empresa o nodo logístico es un agente. Estos sistemas pueden simular y optimizar los flujos de materiales, energía e información entre los agentes para identificar ineficiencias y oportunidades de reducción de emisiones a lo largo de toda la cadena de valor, fomentando la colaboración para alcanzar objetivos de descarbonización compartidos.

2. Propuestas de diseño de herramientas específicas

Basándonos en el potencial de las diversas técnicas de IA analizadas, proponemos el diseño conceptual de tres herramientas específicas orientadas a resolver desafíos clave en el proceso de descarbonización empresarial. Estas herramientas podrían funcionar de manera independiente o integrarse en una solución más amplia.

2.1. Asistente para la Interpretación y Cumplimiento de Legislación Ambiental

Concepto: Esta herramienta tiene como objetivo desmitificar el complejo y a menudo cambiante panorama de la legislación ambiental y las normativas de reporte de emisiones. Se concibe como un sistema inteligente que utiliza Modelos de Lenguaje Extensos (LLMs) y la técnica de Generación Aumentada por Recuperación (RAG) para proporcionar a las empresas respuestas claras, concisas y actualizadas sobre sus obligaciones legales en materia de sostenibilidad y emisiones (a nivel nacional, europeo o internacional, según se configure).

Funcionalidades clave

Chatbot conversacional para consultas en lenguaje natural: Permitiría a los usuarios (incluso sin formación legal) realizar preguntas como "¿Qué obligaciones de reporte de GEI tiene mi empresa si factura X en el sector Y?" o "¿Cuáles son los últimos cambios en la directiva europea de eficiencia energética?", obteniendo respuestas comprensibles.

Base de datos de legislación relevante, viva y actualizada: El sistema se conectaría a una base de conocimiento (gestionada mediante RAG) que contenga textos legales, reglamentos, directivas y guías oficiales, la cual sería objeto de actualización continua para reflejar nuevos desarrollos normativos.

Capacidad de resumir textos legales complejos: El LLM podría generar resúmenes ejecutivos de documentos legales extensos, destacando los puntos clave y las implicaciones para la empresa.

Alertas sobre cambios normativos o nuevas obligaciones: El sistema podría configurarse para enviar notificaciones proactivas a los usuarios cuando se publiquen nuevas leyes o se modifiquen las existentes que afecten a su perfil de empresa.

IA Principal: La combinación de **LLM** para la comprensión de las consultas y la generación de respuestas en lenguaje natural, **RAG** para asegurar que la información se basa en fuentes legislativas actualizadas y fidedignas, y técnicas de **NLP (Procesamiento de Lenguaje Natural)** para el análisis y la estructuración de los textos legales.

2.2. Asistente para Recomendaciones de Acciones de Mitigación Personalizadas

Concepto: Diseñado para ofrecer inteligencia procesable, este sistema va más allá del simple cálculo de la huella de carbono. A partir del análisis detallado de los datos de emisiones de una empresa, su perfil sectorial, su tamaño y otros factores contextuales, el asistente sugeriría un conjunto de acciones de mitigación (tecnológicas, operativas, de gestión) que sean las más efectivas en términos de reducción de emisiones y que ofrezcan el mejor retorno de inversión o viabilidad para esa organización específica.

Funcionalidades clave

Análisis profundo de la huella de carbono para identificar principales focos de emisión:

El sistema identificaría automáticamente las actividades o procesos que contribuyen en mayor medida a las emisiones totales de la empresa.

Base de datos extensa y curada de medidas de mitigación: Contendría información sobre una amplia gama de acciones (ej. mejora de eficiencia energética, adopción de energías renovables, optimización logística, economía circular, tecnologías de captura de carbono), incluyendo datos sobre su potencial de reducción, costes estimados, complejidad de implementación y casos de éxito.

Modelo de Machine Learning (ML) para predecir el impacto: Utilizaría modelos de ML supervisado para estimar con la mayor precisión posible la reducción de emisiones esperada, el coste de implementación y el posible Retorno de la Inversión (ROI) o periodo de amortización de cada medida sugerida en el contexto de la empresa. Los modelos ML para recomendación podrían modificarse por sistemas multiagente como analizaremos más adelante.

Priorización inteligente de recomendaciones: Las acciones sugeridas se presentarían priorizadas según criterios personalizables por el usuario (ej. mayor reducción de CO₂, menor coste inicial, menor tiempo de implementación, cumplimiento de objetivos específicos).

IA Principal: Principalmente **Machine Learning (supervisado)** para los modelos predictivos de impacto y coste-beneficio. Se apoyaría en **LLM** para explicar de forma clara y convincente cada recomendación y su justificación, y en **RAG** para acceder a información contextual adicional como benchmarks sectoriales, casos de estudio relevantes o detalles técnicos actualizados sobre las tecnologías de mitigación.

2.3. Sistema de Generación de Informes a partir de Consultorías Grabadas

Concepto: Esta innovadora herramienta busca transformar la información valiosa, pero a menudo no estructurada, contenida en las grabaciones de vídeo y/o audio de sesiones de consultoría sobre descarbonización, inspecciones de campo o reuniones de trabajo. El sistema procesaría estos archivos multimedia para extraer automáticamente datos clave, identificar discusiones relevantes y generar un borrador estructurado de informe de huella de carbono, hallazgos principales y medidas de descarbonización propuestas durante la sesión.

Funcionalidades clave

Transcripción de audio a texto (Speech-to-Text) de alta precisión: Convertiría el contenido hablado de las grabaciones en texto escrito, sirviendo como base para análisis posteriores.

Identificación y extracción de entidades relevantes: Mediante NLP, el sistema identificaría y extraería del texto transcrito elementos cruciales como fuentes de emisión mencionadas (ej. "caldera de gasoil", "flota de camiones"), datos cuantitativos (ej. "consumo de 10.000 kWh", "reducción del 15%"), problemas detectados y acciones de mitigación propuestas por los consultores o el personal.

Resumen y estructuración inteligente de la información: El sistema agruparía la información extraída, la resumiría y la organizaría de manera lógica y coherente, por ejemplo, por áreas de mejora, tipos de emisión o medidas propuestas.

Generación de un borrador de informe en formato editable: Con la información procesada, el sistema generaría automáticamente un borrador preliminar del informe de consultoría, incluyendo secciones sobre la huella de carbono estimada (si se discute), los principales hallazgos, y un listado de las recomendaciones de descarbonización, lo que reduciría significativamente el tiempo de redacción manual.

IA Principal: Se basaría fuertemente en **Modelos de Lenguaje Multimodales** para el procesamiento integrado de audio y, potencialmente, vídeo (si se analizan elementos visuales de las grabaciones). El **NLP** sería crucial para el análisis del texto transcrito y la extracción de entidades. Finalmente, un **LLM** podría ser empleado para la redacción y estructuración del borrador del informe.

3. Diseño detallado: Asistente General para la Gestión Integral de la Descarbonización

Este asistente se concibe como una solución **SaaS (Software as a Service) de vanguardia, tecnológicamente avanzada y especializada**, diseñada para guiar y empoderar a las organizaciones en cada fase del complejo ciclo de vida de la gestión de su huella de carbono. Su arquitectura y funcionalidades están intrínsecamente ligadas al uso intensivo de Inteligencia Artificial, incluyendo Modelos de Lenguaje Extensos (LLMs), Modelos de Lenguaje Visual (VLMs) y técnicas avanzadas de Machine Learning, para ofrecer una precisión, eficiencia y capacidad de análisis sin precedentes.

3.1. Descripción funcional detallada

El Asistente General para la Gestión Integral de la Descarbonización proporcionará un conjunto de capacidades altamente automatizadas y analíticamente potentes:

1. **Recopilación y Procesamiento Inteligente de Datos:** El sistema irá más allá de la simple entrada manual. Empleará **motores OCR (Optical Character Recognition) de alta precisión**, enriquecidos con **componentes de NLP (Natural Language Processing) basados en LLMs**, para la extracción automática y estructuración semántica de información crítica de una amplia gama de documentos no estructurados o semi-estructurados. Esto incluye facturas energéticas (electricidad, gas, otros combustibles), facturas de agua, informes de gestión de residuos, registros de viajes de negocios, datos de consumo de flotas de vehículos y parámetros de producción. Se explorará activamente el uso de **VLMs para la interpretación de layouts complejos en facturas o el análisis de documentos escaneados de baja calidad**, mejorando drásticamente la tasa de extracción y reduciendo la necesidad de intervención manual.
2. **Cálculo Riguroso y Transparente de la Huella de Carbono:** La plataforma facilitará la ejecución de cálculos precisos de emisiones de Gases de Efecto Invernadero (GEI) para los **Alcances 1 y 2, y ofrecerá capacidades robustas y configurables para el complejo Alcance 3**. El sistema se adherirá estrictamente a metodologías estándar reconocidas internacionalmente, como el GHG Protocol Corporate Standard y la ISO 14064. Integrará una base de datos de factores de emisión viva, actualizable dinámicamente (potencialmente asistida por LLMs con RAG para incorporar las últimas publicaciones de organismos oficiales), y permitirá la trazabilidad completa de cada cálculo, desde el dato de actividad original hasta la emisión final.
3. **Generación Avanzada de Memorias e Informes Asistida por LLMs:** El asistente producirá automáticamente memorias de huella de carbono, informes de sostenibilidad y otros documentos regulatorios o de comunicación. Esta funcionalidad se basará en **LLMs fine-tuned con corpus específicos del dominio de la sostenibilidad y el reporting ambiental**. Estos modelos recibirán como entrada los datos de huella calculados, las recomendaciones de mitigación, los planes de reducción y las directrices de la plantilla para generar borradores narrativos coherentes, precisos y alineados con los estándares exigidos (ej. MITECO, GRI, etc.).

4. **Obtención de Recomendaciones de Mitigación Hiper-Personalizadas y Accionables:** El sistema no se limitará a sugerencias genéricas. Un **motor de recomendación basado en IA (ensamblaje de modelos de ML y LLMs con RAG)** analizará la huella de carbono desglosada, el sector de actividad, el tamaño de la empresa y otros parámetros contextuales para proponer medidas de mitigación específicas. Estas recomendaciones incluirán una estimación cuantitativa del impacto (reducción de tCO₂eq, ahorro económico, ROI), el nivel de inversión requerido y la complejidad de implementación, todo ello justificado y explicado en lenguaje natural por el LLM.
5. **Desarrollo Asistido de Planes Estratégicos de Reducción:** La plataforma guiará a los usuarios en la creación de planes de reducción de emisiones robustos y realistas. Facilitará la definición de objetivos SMART (Específicos, Medibles, Alcanzables, Relevantes, Temporales), la asignación de acciones concretas (derivadas de las recomendaciones de IA), la designación de responsables, el establecimiento de cronogramas y la asignación de presupuestos. **Un LLM podría asistir en la formulación de estos objetivos y acciones**, asegurando su alineación con los resultados del análisis de la huella.
6. **Seguimiento Dinámico y Predictivo del Progreso:** El asistente permitirá una monitorización continua del progreso del plan de reducción. Los dashboards interactivos visualizarán KPIs clave, y el sistema alertará sobre desviaciones o la necesidad de acciones correctivas. Se implementarán **modelos de ML para análisis predictivos**, estimando la probabilidad de alcanzar los objetivos y el impacto acumulado de las medidas implementadas, permitiendo ajustes proactivos al plan.

3.2. Arquitectura general de la solución orientada a IA

La solución se erigirá sobre una arquitectura de microservicios nativa de la nube, diseñada para la máxima escalabilidad, resiliencia, agilidad en el despliegue y facilidad de mantenimiento, con un enfoque central en la orquestación eficiente de los componentes de IA.

- **Frontend:** Una aplicación web progresiva (PWA) o Single Page Application (SPA) desarrollada con frameworks modernos (ej. React, Angular, Vue.js), optimizada para ofrecer una experiencia de usuario fluida e intuitiva en navegadores de escritorio y tabletas. Incorporará visualizaciones de datos interactivas (ej. D3.js, Chart.js) y un **asistente conversacional (chatbot) integrado, impulsado por un LLM**, para soporte contextual y guía al usuario.
- **Backend (API Gateway y Microservicios):**
 - **API Gateway (ej. AWS API Gateway):** Punto de entrada único para todas las solicitudes del frontend y de sistemas externos, gestionando la autenticación, el enrutamiento, el balanceo de carga y la limitación de tasa hacia los microservicios subyacentes.
 - **Microservicios Especializados:**

- **Servicio de Autenticación y Autorización:** Basado en estándares como OAuth 2.0 y OpenID Connect, para una gestión segura de identidades y accesos granulares.
- **Servicio de Ingesta y Procesamiento de Documentos (IDPS):** Este será un microservicio crítico con alta carga de IA. Integrará:
 - **Motores OCR avanzados** (ej. Tesseract con fine-tuning, Abbyy FineReader, o soluciones cloud como AWS Textract, Google Document AI).
 - **Modelos NLP/LLM para Named Entity Recognition (NER), extracción de relaciones y clasificación de información** a partir del texto extraído por OCR.
 - **Capacidades VLM (potencialmente)** para el análisis de la estructura visual de documentos complejos (ej. facturas con múltiples tablas, diagramas en informes técnicos).
 - Colas de procesamiento para gestionar la carga asíncrona de documentos.
- **Servicio de Cálculo de Huella de Carbono:** Contendrá la lógica de cálculo conforme a los estándares, interactuando con la base de datos de factores de emisión y los datos de actividad procesados por el IDPS.
- **Servicio de Modelos de IA y Recomendaciones:** Orquestrará la ejecución de modelos de Machine Learning para predicciones (ej. consumo futuro, impacto de medidas) y el motor de recomendaciones. Este servicio interactuará intensamente con los LLMs y la Vector Database para las funcionalidades de RAG y la generación de explicaciones.
- **Servicio de Generación de Informes y Documentos:** Utilizará LLMs (con acceso a datos estructurados y plantillas) para componer los borradores de informes y memorias.
- **Servicio de Planificación y Seguimiento:** Gestionará la lógica de creación y monitorización de planes de reducción.
- **Bases de Datos Optimizadas:**
 - **Base de Datos Relacional (ej. PostgreSQL, MySQL con optimizaciones para lectura/escritura frecuente):** Para almacenar datos estructurados como perfiles de usuario, información de empresas, planes de reducción, factores de emisión estructurados y metadatos de los cálculos.
 - **Base de Datos NoSQL Documental (ej. MongoDB, AWS DocumentDB):** Ideal para almacenar los documentos originales (PDFs, imágenes), los resultados intermedios del procesamiento OCR/NLP (ej. JSON con texto y entidades), logs detallados del sistema y configuraciones flexibles.
 - **Vector Database (ej. Pinecone, Weaviate, FAISS integrada):** Componente esencial para la implementación eficiente de RAG. Almacenará embeddings vectoriales de la base de conocimiento (normativas, papers técnicos, mejores

prácticas, catálogos de medidas de mitigación) para permitir búsquedas semánticas rápidas y precisas por parte de los LLMs.

- **Infraestructura de IA Dedicada:**
 - **Plataforma de MLOps (ej. Kubeflow, AWS SageMaker, Azure ML):** Para gestionar el ciclo de vida completo de los modelos de ML y LLMs, incluyendo el versionado de datos y modelos, entrenamiento distribuido (con acceso a GPUs/TPUs), despliegue como endpoints de inferencia escalables, y monitorización continua del rendimiento y drift de los modelos.
 - **Servidores de Inferencia Optimizados:** Para el despliegue de LLMs y VLMs, que pueden requerir hardware específico (GPUs) y software de optimización (ej. TensorRT, ONNX Runtime) para inferencias de baja latencia.
- **Cola de Mensajes (ej. RabbitMQ, Apache Kafka):** Para la comunicación asíncrona entre microservicios, asegurando el desacoplamiento, la resiliencia y la capacidad de manejar picos de carga (ej. en la ingesta masiva de documentos o en el lanzamiento de cálculos complejos).

3.3. Módulos principales y su interacción con la IA

- **Módulo de Ingesta de Datos Inteligente:**
 - **Entrada manual de datos:** Formularios web validados y estructurados.
 - **Carga de archivos (PDF, CSV, Excel) con extracción automática avanzada:** Al cargar un PDF de una factura eléctrica, por ejemplo, el **motor OCR** digitalizará el texto. Seguidamente, **modelos NLP (posiblemente basados en arquitecturas Transformer o LLMs especializados en extracción)** identificarán y extraerán entidades clave como el periodo de consumo, los kWh consumidos, el coste total, el proveedor, etc. Para documentos con estructuras visuales complejas o imágenes de mala calidad, se considerará la aplicación de **Modelos de Lenguaje Visual (VLM)** que puedan interpretar el layout (tablas, secciones) y el contenido textual de forma conjunta para una extracción más robusta. Los datos extraídos serán validados y presentados al usuario para una posible supervisión y corrección, donde un **LLM podría incluso sugerir correcciones** basadas en datos históricos o patrones comunes.
 - **Conectores API con sistemas de terceros (ERP, sistemas de gestión de flotas, plataformas IoT):** Para la ingesta directa y automatizada de datos estructurados, minimizando la intervención manual.
- **Módulo de Cálculo de Huella de Carbono Riguroso:**
 - **Configuración guiada de límites organizacionales y operacionales:** Un **asistente basado en LLM** podría ayudar al usuario a interpretar y aplicar correctamente los criterios del GHG Protocol para definir los límites de su inventario.

- **Selección inteligente de metodologías y factores de emisión:** El sistema propondrá los factores de emisión más adecuados basados en la ubicación geográfica, el tipo de fuente y el año de reporte, utilizando una base de datos que se actualiza mediante procesos automatizados y, potencialmente, **RAG para consultar las últimas publicaciones de organismos oficiales (IEA, EPA, DEFRA, MITECO)** y ofrecer la justificación de dichos factores.
- **Cálculo transparente de emisiones de Alcance 1, 2 y Alcance 3 (configurable):** La plataforma ofrecerá herramientas específicas para las categorías más relevantes del Alcance 3, donde la **IA podrá asistir en la estimación de emisiones para categorías de difícil cuantificación**, como bienes y servicios adquiridos (analizando datos de compras y aplicando modelos de gasto-emisión) o viajes de negocios (procesando informes de gastos o integrándose con plataformas de gestión de viajes).
- **Visualización interactiva y analítica de resultados:** Dashboards dinámicos permitirán desglosar la huella por fuente, alcance, unidad de negocio, gas GEI, etc., facilitando la identificación de "hotspots" de emisión.
- **Módulo de Generación de Recomendaciones Proactivas y Personalizadas (Núcleo de IA):**
 - **Motor de IA polifacético para el análisis de la huella y la proposición de medidas:** Este motor integrará **modelos de Machine Learning (ej. árboles de decisión, gradient boosting, redes neuronales)** entrenados para predecir el impacto cuantitativo (reducción de tCO₂eq, ahorro económico, periodo de amortización) de un amplio catálogo de medidas de mitigación (eficiencia energética, adopción de renovables, optimización de procesos, electrificación, economía circular, etc.). Un **LLM, potenciado por RAG para acceder a una base de conocimiento actualizada y detallada de tecnologías, proveedores y casos de éxito**, generará explicaciones detalladas, comprensibles y personalizadas para cada recomendación, justificando su idoneidad para la empresa específica.
 - **Estimación precisa de potencial de reducción, coste de implementación y ROI:** Las predicciones de los modelos de ML se basarán en los datos específicos de la empresa (consumos, procesos, equipamiento) y en benchmarks sectoriales.
 - **Vinculación con bases de datos de tecnologías y proveedores:** El sistema (a través del LLM con RAG) podrá ofrecer información contextual sobre tecnologías específicas o incluso sugerir tipos de proveedores, basándose en la medida recomendada.
- **Módulo de Elaboración de Memorias y Planes Asistido por LLM:**
 - **Plantillas dinámicas y personalizables para informes (GHG Protocol, memorias de sostenibilidad MITECO, GRI, etc.).**
 - **Generación automática de texto de alta calidad basada en LLMs:** Los **LLMs** recibirán como entrada los datos calculados de la huella, las recomendaciones de IA seleccionadas, los objetivos del plan de reducción y las directrices de la plantilla del informe. A partir de esto, generarán borradores

de las secciones narrativas del informe (descripción de la metodología, análisis de resultados, plan de reducción, etc.), manteniendo un estilo profesional y coherente. El usuario podrá luego editar y refinar estos borradores.

- **Herramientas colaborativas para definir objetivos, acciones, responsables y cronogramas del plan de reducción:** La IA podrá sugerir KPIs relevantes o ayudar a estructurar el plan de forma lógica.
- **Módulo de Seguimiento y Monitorización Inteligente:**
 - **Dashboards con KPIs de progreso del plan y proyecciones basadas en IA:** Visualización del avance hacia los objetivos, incluyendo **predicciones de ML sobre la probabilidad de alcanzar las metas** en los plazos establecidos, y alertas tempranas si se detectan desviaciones significativas.
 - **Alertas proactivas sobre desviaciones o hitos:** El sistema notificará automáticamente a los responsables.
 - **Recálculo periódico y automatizado de la huella para evaluar el impacto real de las medidas implementadas:** Permitiendo un ciclo de mejora continua y la validación del rendimiento de las acciones de mitigación.

3.4. Especificaciones técnicas detalladas

A continuación, se detallan las especificaciones, con un párrafo explicativo para cada una, resaltando los aspectos técnicos y la implicación de las herramientas de IA.

3.4.1. ESPECIFICACIONES GENERALES DE LA PLATAFORMA

1. **GP001: La plataforma será una aplicación web moderna y accesible.**
Se desarrollará como una Single Page Application (SPA) o Progressive Web Application (PWA) utilizando frameworks como React, Angular o Vue.js, garantizando una experiencia de usuario fluida, interactiva y accesible desde las últimas versiones de navegadores estándar (Chrome, Firefox, Edge, Safari). Se asegurará la compatibilidad con resoluciones de escritorio y tabletas.
2. **GP002: Arquitectura robusta basada en microservicios.**
La arquitectura de backend se fundamentará en un patrón de microservicios, desplegados preferentemente en contenedores (Docker) y orquestados mediante Kubernetes. Esta elección promueve la escalabilidad independiente de los componentes (especialmente los servicios de IA que pueden tener demandas variables), la resiliencia ante fallos de servicios individuales, y facilita las actualizaciones y el mantenimiento evolutivo de la plataforma.
3. **GP003: Despliegue nativo en infraestructura cloud.**
La solución se desplegará en una de las principales plataformas cloud (AWS, Azure o GCP), aprovechando sus servicios gestionados para bases de datos, orquestación de contenedores, funciones serverless, colas de mensajes, y, crucialmente, su infraestructura especializada para el entrenamiento y la inferencia de modelos de IA (GPUs, TPUs, servicios de MLOps).

4. **GP004: Alta disponibilidad del servicio (99.9%).**
Se diseñará la infraestructura para una alta disponibilidad, implementando redundancia en todos los niveles críticos (balanceadores de carga, instancias de microservicios, réplicas de bases de datos entre zonas de disponibilidad) y estrategias de failover automático. El objetivo es un Service Level Objective (SLO) de disponibilidad del 99.9%.
5. **GP005: Soporte multi-idioma con capacidad de expansión.**
Inicialmente, la interfaz de usuario y los contenidos generados (informes, recomendaciones) estarán disponibles en español e inglés. La arquitectura de internacionalización (i18n) y localización (l10n) permitirá la fácil adición de nuevos idiomas en el futuro. Los LLMs utilizados para generación de texto deberán ser capaces de operar en los idiomas soportados.
6. **GP006: Diseño responsive y adaptativo.**
La interfaz de usuario se diseñará siguiendo principios de "mobile-first" aunque el enfoque principal sea desktop/tablet, asegurando que todos los elementos visuales y funcionales se adapten correctamente a diferentes tamaños de pantalla y orientaciones, manteniendo una usabilidad óptima.
7. **GP007: Sistema granular de gestión de usuarios con roles y permisos.**
Se implementará un sistema de Role-Based Access Control (RBAC) que permita definir roles (ej. Administrador de Organización, Calculador de Huella, Editor de Informes, Lector) con permisos específicos sobre las diferentes funcionalidades y conjuntos de datos, asegurando la segregación de funciones y la seguridad de la información.
8. **GP008: Arquitectura multi-tenant para aislamiento de datos y personalización.**
La plataforma estará diseñada para dar servicio a múltiples organizaciones (tenants) de forma completamente aislada, tanto a nivel de datos como de configuración. Cada tenant tendrá su propio esquema de datos o base de datos lógicamente separada, y podrá personalizar aspectos como factores de emisión específicos o plantillas de informes.
9. **GP009: Sistema centralizado de logs y monitorización de actividad.**
Se implementará un sistema de logging centralizado (ej. ELK Stack - Elasticsearch, Logstash, Kibana; o Grafana Loki) para registrar eventos de aplicación, errores, accesos de usuarios y actividad de los modelos de IA. Esto es crucial para la depuración, la auditoría de seguridad y la monitorización del rendimiento del sistema.
10. **GP010: Documentación técnica exhaustiva y actualizada.**
Se mantendrá una documentación técnica completa que cubra la arquitectura de la solución, el diseño de las APIs (siguiendo el estándar OpenAPI), los modelos de datos, los flujos de trabajo de los componentes de IA y los procedimientos de despliegue y mantenimiento.

3.4.2. ESPECIFICACIONES DEL MÓDULO DE INGESTA DE DATOS (ID)

1. **ID001: Soporte para carga manual de datos mediante formularios web dinámicos.**

Se proveerán interfaces de usuario intuitivas para la introducción manual de datos de actividad que no puedan ser importados automáticamente. Estos formularios serán dinámicos, adaptándose al tipo de fuente de emisión y a la metodología seleccionada, e incluirán validaciones en tiempo real para minimizar errores.

2. **ID002: Soporte para carga masiva de datos mediante archivos estructurados (CSV, Excel).**

Los usuarios podrán cargar grandes volúmenes de datos de actividad preexistentes mediante plantillas de archivos CSV o Excel. El sistema permitirá el mapeo flexible de columnas del archivo de origen a los campos requeridos por la plataforma y realizará validaciones exhaustivas durante el proceso de importación.

3. **ID003: Extracción de datos de documentos PDF y formatos de imagen mediante tecnología OCR avanzada y VLM.**

La plataforma incorporará un motor de Reconocimiento Óptico de Caracteres (OCR) robusto, capaz de procesar documentos PDF (nativos y escaneados), así como formatos de imagen comunes (JPEG, PNG, TIFF). Este motor será entrenado o afinado (fine-tuning) para reconocer una amplia variedad de plantillas de facturas energéticas, de combustible, agua y gestión de residuos. Se priorizará la capacidad de manejar documentos con layouts variables, diferentes calidades de escaneo y la presencia de artefactos visuales. Se evaluará activamente la integración de componentes de preprocesamiento de imagen (ej. mejora de contraste, binarización adaptativa, eliminación de ruido) y, para casos particularmente complejos o documentos multimodales, se investigará y aplicará el uso de **Modelos de Lenguaje Visual (VLM)** que puedan interpretar la estructura tabular, las secciones y el contenido textual de forma conjunta, superando las limitaciones de los OCR tradicionales en documentos complejos.

4. **ID004: Precisión del OCR y VLM superior al 95% en campos clave de documentos estructurados y semi-estructurados.**

El objetivo de precisión para la extracción de campos de datos críticos (ej. consumo, unidades, fechas, importes) de documentos comunes como facturas será superior al 95% tras el procesamiento por OCR/VLM y la posterior capa de NLP. Esta métrica se monitorizará continuamente y se utilizará para el reentrenamiento y mejora de los modelos.

5. **ID005: Aplicación de NLP y LLMs para la identificación y extracción semántica de entidades clave.**

Una vez que el OCR/VLM haya extraído el texto crudo, se aplicarán modelos de Procesamiento de Lenguaje Natural (NLP), incluyendo **LLMs especializados o fine-tuned para Named Entity Recognition (NER) y extracción de relaciones**, para identificar semánticamente y estructurar la información relevante. Por ejemplo, el sistema deberá ser capaz de entender que "Consumo Periodo Punta" se refiere a una cantidad de energía eléctrica, asociarle su unidad (kWh) y el valor numérico correspondiente, así como el periodo al que aplica, incluso si la redacción o el formato varían entre facturas. Esta capa de NLP manejará la normalización de unidades, fechas y la desambiguación de términos.

6. **ID006: Implementación de robustos mecanismos de validación de datos en la ingesta.**

Todos los datos, ya sean introducidos manualmente, cargados masivamente o extraídos por IA, pasarán por un proceso de validación configurable. Esto incluirá validaciones de tipo de dato, rango (ej. un consumo no puede ser negativo), formato, consistencia entre campos (ej. la fecha final debe ser posterior a la inicial) y reglas de negocio específicas (ej. comparación con consumos históricos para detectar outliers significativos, lo cual podría ser asistido por modelos de detección de anomalías).

7. **ID007: Interfaz de usuario para la supervisión, corrección y validación de datos extraídos por IA.**

Se proporcionará una interfaz ("human-in-the-loop") donde los usuarios puedan revisar los datos extraídos automáticamente por los modelos OCR/VLM/NLP. La interfaz resaltará los campos donde el modelo tenga baja confianza y permitirá una corrección fácil e intuitiva. Esta retroalimentación del usuario será crucial y se utilizará para el reentrenamiento y la mejora continua de los modelos de extracción (aprendizaje activo). Un **LLM podría asistir sugiriendo correcciones o completando información** basándose en el contexto.

8. **ID008: Almacenamiento seguro y versionado de los documentos originales cargados.**

Todos los documentos fuente cargados por los usuarios (PDFs, imágenes, CSVs) se almacenarán de forma segura en un sistema de almacenamiento de objetos (ej. AWS S3, Azure Blob Storage) con políticas de acceso estrictas y cifrado en reposo. Se implementará un sistema de versionado para estos documentos, permitiendo la trazabilidad y la auditoría.

9. **ID009: Desarrollo de conectores API predefinidos para fuentes de datos comunes.**

Se desarrollarán y mantendrán conectores API para interactuar directamente con sistemas de terceros que son fuentes comunes de datos de actividad, como plataformas de proveedores energéticos (si ofrecen APIs públicas o de partners), sistemas ERP populares, software de gestión de flotas o plataformas IoT que recopilen datos de sensores. Esto automatizará aún más la ingesta de datos.

10. **ID010: Framework extensible para el desarrollo de conectores API personalizados.**

Se proporcionará un framework o SDK que facilite a los usuarios (o a un equipo de integración) el desarrollo de conectores personalizados para sus sistemas legacy o fuentes de datos específicas que no estén cubiertas por los conectores predefinidos, promoviendo la flexibilidad y adaptabilidad de la plataforma.

11. **ID011: Implementación de un sistema de versionado de datos de actividad.**

Cada dato de actividad ingresado en el sistema, independientemente de su origen, será versionado. Esto permitirá rastrear cualquier modificación, corrección o eliminación de datos a lo largo del tiempo, asegurando una completa auditabilidad y la capacidad de recalcular la huella de carbono con diferentes instantáneas de los datos.

12. ID012: Capacidad avanzada de mapeo de campos de origen a campos del sistema.

Para la importación de datos desde archivos o APIs, el sistema ofrecerá una interfaz gráfica e inteligente (potencialmente asistida por IA para sugerir mapeos comunes) que permita a los usuarios mapear los campos de sus datos de origen a los campos internos requeridos por la plataforma para el cálculo de la huella, manejando transformaciones simples si fuera necesario.

13. ID013: Funcionalidad para la programación de ingestas automáticas recurrentes.

Los usuarios podrán configurar tareas de ingesta de datos para que se ejecuten automáticamente de forma recurrente (ej. mensual para facturas energéticas, diaria para datos de sensores), utilizando los conectores API o accediendo a ubicaciones de archivos predefinidas.

14. ID014: Sistema de notificaciones al usuario sobre el estado y resultados de las ingestas.

Se informará a los usuarios (mediante notificaciones en la plataforma y/o por correo electrónico) sobre el progreso y el resultado de los procesos de ingesta de datos, incluyendo el número de registros procesados, los errores encontrados, y los datos que requieren validación manual.

15. ID015: Soporte integral para múltiples unidades de medida y conversión automática inteligente.

El sistema manejará una amplia variedad de unidades de medida para energía, combustible, distancia, peso, volumen, etc. Realizará conversiones automáticas a las unidades estándar requeridas para los cálculos de emisiones, basándose en factores de conversión estándar. La capa de NLP/LLM en la ingesta ayudará a identificar correctamente las unidades originales, incluso si están expresadas de forma no estándar.

3.4.3. ESPECIFICACIONES DEL MÓDULO DE CÁLCULO DE HUELLA DE CARBONO

1. HC001: Cumplimiento estricto con metodologías estándar internacionales (GHG Protocol Corporate Standard, ISO 14064-1).

El motor de cálculo se adherirá rigurosamente a los principios y requisitos establecidos en el GHG Protocol Corporate Accounting and Reporting Standard y en la norma ISO 14064-1. Esto incluye la definición de límites organizacionales y operativos, la identificación de fuentes de emisión, la selección de factores de emisión y los métodos de cálculo específicos para cada tipo de emisión. Un **asistente conversacional basado en LLM** podrá guiar al usuario en la interpretación de estos estándares.

2. HC002: Soporte completo para el cálculo de emisiones de Alcance 1 (emisiones directas).

La plataforma permitirá el cálculo detallado de todas las emisiones directas de GEI provenientes de fuentes que son propiedad o están controladas por la organización. Esto incluye la combustión estacionaria (calderas, hornos), la combustión móvil (flota

de vehículos), las emisiones de procesos industriales y las emisiones fugitivas (ej. fugas de refrigerantes HFC).

3. **HC003: Soporte completo para el cálculo de emisiones de Alcance 2.**

Se calcularán las emisiones indirectas asociadas a la generación de electricidad, calor, vapor o refrigeración adquirida y consumida por la organización.

- **HC003.1: Aplicación de enfoques duales: basado en mercado (market-based) y basado en ubicación (location-based).** El sistema permitirá el cálculo y reporte de las emisiones de Alcance 2 utilizando ambos enfoques, tal como lo requiere el GHG Protocol Scope 2 Guidance, permitiendo a las empresas reflejar el impacto de sus contratos de energía renovable (certificados, PPA) a través del enfoque de mercado.

4. **HC004: Soporte configurable y metodológicamente sólido para categorías relevantes de Alcance 3.**

Dado que el Alcance 3 es a menudo el más complejo y extenso, la plataforma ofrecerá herramientas y guías metodológicas (potencialmente asistidas por un **LLM con RAG que acceda a la Guía Técnica de Alcance 3 del GHG Protocol**) para calcular las 15 categorías. Se priorizará el soporte para las más comunes:

- **HC004.1: Categoría 1: Bienes y servicios adquiridos.** Se podrán utilizar datos de gasto (spend-based) combinados con factores de emisión medioambientales de entrada-salida (EEIO), o datos más específicos del proveedor cuando estén disponibles. La IA podría ayudar a clasificar las compras y asignar factores adecuados.
- **HC004.2: Categoría 3: Actividades relacionadas con combustible y energía (no incluidas en Alcance 1 o 2).** Principalmente emisiones de la extracción, producción y transporte de combustibles y energía comprada por la organización.
- **HC004.3: Categoría 4: Transporte y distribución aguas arriba.** Emisiones del transporte de productos adquiridos por la organización.
- **HC004.4: Categoría 5: Residuos generados en las operaciones.** Emisiones del tratamiento y eliminación de residuos. La IA podría ayudar a estimar las emisiones a partir de tipos y cantidades de residuos.
- **HC004.5: Categoría 6: Viajes de negocios.** Se calcularán emisiones de vuelos, trenes, vehículos de alquiler, etc., a partir de datos de distancia, tipo de transporte o datos de plataformas de gestión de viajes.
- **HC004.6: Categoría 7: Desplazamiento de empleados (commuting).** Estimación de emisiones del transporte de los empleados desde sus hogares al lugar de trabajo, a menudo mediante encuestas o modelos. Un **LLM podría ayudar a diseñar y analizar los resultados de estas encuestas.**

5. **HC005: Base de datos de factores de emisión exhaustiva, actualizable y personalizable, gestionada con IA.**

La plataforma integrará una base de datos completa de factores de emisión para CO₂, CH₄, N₂O y otros GEI relevantes, cubriendo una amplia gama de fuentes, combustibles, redes eléctricas y procesos.

- **HC005.1: Inclusión de factores de emisión de fuentes reconocidas internacionalmente y nacionalmente.** Se incorporarán factores de organismos como el IPCC, EPA (EE.UU.), DEFRA (Reino Unido), IEA, Ecoinvent, y fuentes nacionales específicas (ej. MITECO en España). La base de datos será actualizada periódicamente. Un **sistema RAG podría alertar sobre nuevas publicaciones de factores** y un LLM asistir en su interpretación e integración. Los usuarios también podrán añadir sus propios factores de emisión personalizados con la debida justificación.

6. **HC006: Trazabilidad completa y auditable de todos los cálculos de emisiones.**

Cada cifra de emisión calculada será completamente trazable hasta los datos de actividad originales, los factores de emisión utilizados (incluyendo su fuente y versión) y las fórmulas de cálculo aplicadas. Esta transparencia es fundamental para la verificación externa y la credibilidad del inventario.

7. **HC007: Capacidad para definir un año base y realizar seguimiento comparativo y recalculos.**

Los usuarios podrán definir un año base para su inventario de GEI y la plataforma facilitará la comparación de las emisiones de años posteriores con este año base. Se soportarán políticas de recalculo del año base si se producen cambios estructurales significativos en la organización, conforme a las directrices del GHG Protocol.

8. **HC008: Agregación de emisiones por gas de efecto invernadero (GEI) y conversión a CO₂ equivalente (CO₂eq).**

Las emisiones se calcularán individualmente para cada GEI relevante (CO₂, CH₄, N₂O, HFCs, PFCs, SF₆, NF₃) y luego se convertirán a toneladas de CO₂ equivalente utilizando los Potenciales de Calentamiento Global (GWP) publicados por el IPCC en sus informes de evaluación (AR).

9. **HC009: Cálculo automático de indicadores de intensidad de emisiones.**

La plataforma calculará automáticamente indicadores de intensidad de emisiones (ej. tCO₂eq por unidad de producción, tCO₂eq por empleado, tCO₂eq por facturación), permitiendo a las empresas normalizar sus emisiones y realizar un seguimiento de su eficiencia de carbono a lo largo del tiempo, independientemente de los cambios en el nivel de actividad.

10. **HC010: Visualización avanzada de resultados mediante gráficos y tablas interactivas y personalizables.**

Se ofrecerán múltiples opciones de visualización (gráficos de barras, tartas, Sankey, mapas de calor, series temporales) para explorar los resultados del inventario de GEI. Estas visualizaciones serán interactivas, permitiendo filtrar, desglosar y profundizar en los datos desde diferentes perspectivas (fuente, alcance, unidad de negocio, geografía).

11. HC011: Exportación flexible de resultados de cálculo en formatos estándar (CSV, Excel, PDF).

Los usuarios podrán exportar los datos detallados del inventario de emisiones, los resúmenes y los gráficos en formatos comunes como CSV y Excel para análisis adicionales, y en PDF para su inclusión en informes.

12. HC012: Configuración detallada de límites organizacionales según diferentes enfoques de consolidación.

El sistema permitirá a las empresas definir sus límites organizacionales utilizando los enfoques de consolidación especificados en el GHG Protocol: control (financiero u operativo) o participación accionarial (equity share), asegurando que todas las emisiones relevantes se incluyan de acuerdo con la estructura de la empresa.

13. HC013: Recálculo automático y selectivo de la huella ante cambios en datos de actividad o factores de emisión.

Si se modifican datos de actividad históricos o se actualizan factores de emisión, la plataforma permitirá un recálculo eficiente de la huella de carbono para los periodos afectados, manteniendo la consistencia y la integridad de los datos a lo largo del tiempo.

14. HC014: Soporte para diferentes series de Potenciales de Calentamiento Global (GWP) según los informes del IPCC.

Los usuarios podrán seleccionar qué conjunto de valores GWP (ej. del Cuarto Informe de Evaluación - AR4, o del Quinto - AR5, o el más reciente AR6) utilizar para la conversión a CO₂eq, de acuerdo con los requisitos de los programas de reporte a los que se adhieran. La plataforma almacenará los GWP de diferentes informes del IPCC.

15. HC015: Posibilidad de documentar y justificar la exclusión de fuentes de emisión no relevantes.

Conforme al principio de relevancia del GHG Protocol, el sistema permitirá a los usuarios identificar y documentar formalmente cualquier fuente de emisión que se excluya del inventario por no ser significativa, proporcionando la justificación correspondiente para fines de auditoría.

3.4.4. ESPECIFICACIONES DEL MÓDULO DE GENERACIÓN DE RECOMENDACIONES (IA CENTRIC)

1. GR001: Motor de IA híbrido (ML/LLM con RAG) para el análisis de la huella y la proposición de medidas de mitigación.

El núcleo del sistema de recomendaciones residirá en un sofisticado motor de Inteligencia Artificial. Este motor será un ensamblaje de múltiples modelos:

- **Modelos de Machine Learning (ML) supervisado** (ej. regresores XGBoost para estimar ahorros energéticos/económicos, clasificadores RandomForest para categorizar la idoneidad de una medida según el sector/proceso) entrenados con datos históricos anonimizados, benchmarks sectoriales, catálogos de tecnologías y simulaciones de impacto.

- Un **Modelo de Lenguaje Extenso (LLM)**, fine-tuned para el dominio de la descarbonización y la eficiencia energética, que orquestrará la presentación de las recomendaciones. Este LLM utilizará **Generación Aumentada por Recuperación (RAG)** para consultar en tiempo real una **Vector Database** que contendrá una base de conocimiento extensa y actualizada sobre tecnologías de mitigación, estudios de caso, normativas de incentivos, y mejores prácticas. Esto asegura que las recomendaciones sean contextualmente relevantes y basadas en la información más reciente. El LLM será responsable de generar explicaciones detalladas y personalizadas en lenguaje natural para cada recomendación, adaptando el nivel técnico al perfil del usuario.

2. **GR002: Base de conocimiento dinámica de medidas de mitigación, categorizadas y enriquecidas continuamente.**

La Vector Database contendrá embeddings de una vasta colección de documentos y datos estructurados sobre medidas de mitigación, categorizadas por tipo (eficiencia energética, energías renovables, economía circular, optimización de procesos, cambio modal en transporte, etc.), sector de aplicación, tecnología implicada, coste estimado de implementación, potencial de reducción de emisiones, y co-beneficios (ej. mejora de la calidad del aire, reducción de costes operativos).

- **GR002.1: Actualización periódica y semiautomatizada de la base de conocimiento.** Se implementarán procesos para incorporar nuevas tecnologías, investigaciones, políticas de incentivos y mejores prácticas. Esto podría involucrar web scraping de fuentes relevantes, procesamiento de feeds de noticias científicas y regulatorias mediante NLP, y la curación asistida por expertos humanos. Un LLM podría ayudar a resumir y categorizar nueva información para su inclusión.

3. **GR003: Estimación cuantitativa del potencial de reducción de emisiones para cada medida propuesta.**

Para cada recomendación, los modelos de ML del motor de IA proporcionarán una estimación cuantitativa del potencial de reducción de emisiones de GEI (en tCO₂eq/año), basada en los datos específicos de la huella de carbono de la empresa (ej. consumo energético de un equipo específico que se propone reemplazar) y en los parámetros de rendimiento de la tecnología o práctica recomendada.

4. **GR004: Estimación del coste de implementación, ahorros potenciales y Retorno de la Inversión (ROI).**

Además del impacto en emisiones, el motor de IA estimará el coste de inversión inicial (CAPEX), los ahorros operativos anuales (OPEX, ej. por reducción de consumo energético) y calculará métricas financieras clave como el Retorno de la Inversión (ROI) y el periodo de amortización simple para cada medida. Estas estimaciones se basarán en modelos de costes y en datos de mercado accesibles a través del RAG.

5. **GR005: Priorización multicriterio de recomendaciones, configurable por el usuario.**

Las recomendaciones se presentarán al usuario en un formato priorizado. El algoritmo de priorización considerará múltiples criterios (ej. máximo potencial de reducción de

CO2, menor coste de implementación, mayor ROI, menor complejidad técnica, alineación con objetivos estratégicos de la empresa). Los usuarios podrán ajustar los pesos de estos criterios para personalizar la priorización según sus necesidades y capacidades.

6. **GR006: Presentación de recomendaciones en lenguaje claro, comprensible y accionable, con justificación detallada por el LLM.**

El LLM del motor de IA será crucial para traducir los resultados técnicos de los modelos de ML en explicaciones claras, concisas y persuasivas. Cada recomendación irá acompañada de una justificación que explique por qué es relevante para la empresa, cómo funciona la tecnología o práctica propuesta, cuáles son los beneficios esperados y los próximos pasos sugeridos para su implementación. Se evitará la jerga técnica excesiva a menos que el perfil del usuario lo demande.

7. **GR007: Vinculación explícita de las recomendaciones a las fuentes de emisión específicas identificadas en la huella.**

Cada medida de mitigación propuesta estará claramente vinculada a las fuentes de emisión específicas (ej. "Sustitución de la caldera X del edificio Y", "Optimización de rutas para la flota de vehículos de reparto Z") que aborda, permitiendo al usuario comprender directamente cómo cada acción contribuye a reducir su huella de carbono en áreas concretas.

8. **GR008: Interfaz interactiva para que el usuario acepte, rechace, posponga o personalice las recomendaciones.**

El usuario tendrá control total sobre las recomendaciones. Podrá aceptarlas para incluirlas en su plan de reducción, rechazarlas (idealmente proporcionando un motivo, lo que retroalimentaría al sistema), posponer su consideración, o incluso personalizar parámetros de la recomendación (ej. ajustar la escala de implementación).

9. **GR009: Funcionalidad de Benchmarking anónimo y contextualizado.**

Cuando sea posible y con datos agregados y anonimizados, el sistema podrá ofrecer un benchmarking comparando la intensidad de emisiones de la empresa o la adopción de ciertas medidas con otras organizaciones similares en sector y tamaño. Esta información, proporcionada con cautela y respetando la privacidad, puede ser un fuerte motivador. **Modelos de ML (clustering)** podrían identificar grupos de empresas comparables.

10. **GR010: Uso intensivo de RAG para enriquecer las recomendaciones con información externa, hiper-actualizada y específica.**

La técnica de Generación Aumentada por Recuperación (RAG) será fundamental. El LLM no solo se basará en su conocimiento pre-entrenado, sino que consultará activamente la Vector Database para obtener fragmentos de información específica (ej. detalles técnicos de un nuevo modelo de bomba de calor, información sobre una subvención local para eficiencia energética recién publicada, un caso de estudio de una empresa similar que implementó una medida específica) para enriquecer y validar sus recomendaciones, haciéndolas mucho más precisas, actuales y accionables.

3.4.5. ESPECIFICACIONES DEL MÓDULO DE ELABORACIÓN DE MEMORIAS Y PLANES

Este módulo es fundamental para traducir los datos de emisiones y las estrategias de mitigación en documentos formales y planes de acción estructurados. La Inteligencia Artificial, especialmente los LLMs, juega un papel transformador en la automatización y mejora de la calidad de estos entregables.

1. **MP001: Generador de informes con un repositorio de plantillas predefinidas y actualizables (ej. MITECO, GHG Protocol, GRI, CDP).**

La plataforma ofrecerá una biblioteca extensible de plantillas de informes que cumplan con los principales estándares y marcos de reporte de sostenibilidad y emisiones a nivel nacional e internacional. Estas plantillas no serán meros documentos estáticos, sino estructuras dinámicas que un **Modelo de Lenguaje Extenso (LLM), entrenado específicamente en la semántica y los requisitos de cada estándar (potencialmente utilizando RAG para consultar las guías oficiales del estándar durante la generación)**, podrá interpretar y poblar. La biblioteca de plantillas será actualizada periódicamente por el proveedor de la plataforma para reflejar cambios en los estándares o la aparición de nuevos marcos de reporte relevantes.

2. **MP002: Capacidad de personalización avanzada de plantillas de informes y creación de plantillas corporativas.**

Además de las plantillas estándar, las organizaciones podrán personalizar profundamente las existentes (modificando secciones, añadiendo su identidad corporativa como logos y estilos) o crear sus propias plantillas de informe desde cero utilizando un editor integrado. Un **LLM podría asistir en este proceso de personalización, sugiriendo la estructura de nuevas secciones basadas en el contenido que el usuario desea incluir, o ayudando a mapear campos de datos de la plataforma a las secciones personalizadas del informe**, asegurando la coherencia y la correcta inserción de información.

3. **MP003: Generación de texto asistida por LLMs (fine-tuned) para secciones descriptivas y analíticas del informe, basándose en datos cuantitativos y cualitativos.**

Un **LLM (o un ensemble de LLMs especializados y fine-tuned con corpus de informes de sostenibilidad y descarbonización de alta calidad)** será capaz de redactar borradores de las secciones narrativas de los informes. Este LLM tomará como entrada:

- Los datos cuantitativos del inventario de GEI (emisiones totales, desgloses por alcance, fuente, gas, indicadores de intensidad).
- Las medidas de mitigación seleccionadas y su impacto estimado.
- Los objetivos del plan de reducción.
- Información contextual de la empresa (sector, tamaño, estrategia de sostenibilidad).
- Los requisitos específicos de la plantilla de informe seleccionada. **Utilizando técnicas de RAG, el LLM podrá enriquecer el texto con**

explicaciones metodológicas (citando cláusulas del GHG Protocol, por ejemplo), comparaciones con benchmarks sectoriales (si están disponibles y el usuario lo consiente), o interpretaciones de tendencias observadas en los datos de emisiones. El resultado será un borrador de texto coherente, preciso, con el tono adecuado y que requerirá una revisión y edición significativamente menor por parte del usuario.

4. MP004: Exportación de informes en formatos estándar y editables (PDF, DOCX) con preservación de formato.

Los informes generados, incluyendo el texto redactado por IA y las visualizaciones de datos integradas, podrán ser exportados a formatos universalmente aceptados. El formato PDF se utilizará para la distribución final y la presentación oficial, mientras que el formato DOCX (Microsoft Word) permitirá a los usuarios realizar ediciones y ajustes finos posteriores, integrando el trabajo de la IA en sus flujos de revisión y aprobación documental existentes. Se prestará especial atención a la fidelidad del formato en la exportación.

5. MP005: Herramienta dedicada para la creación de planes de reducción de emisiones estructurados y cuantificables.

Se proporcionará un módulo específico para que las empresas puedan desarrollar sus planes estratégicos de descarbonización. Este módulo no será una simple lista de tareas, sino una herramienta de planificación completa.

- **MP005.1: Definición asistida por IA de objetivos SMART (Específicos, Medibles, Alcanzables, Relevantes, Temporales) para la reducción de emisiones.** Un asistente conversacional basado en LLM guiará al usuario en la formulación de objetivos de reducción de emisiones que cumplan con los criterios SMART. Por ejemplo, si un usuario indica "quiero reducir mucho mis emisiones", el LLM podría preguntar sobre plazos, porcentajes específicos, alcances a los que aplica, y analizar la viabilidad basándose en la huella actual y las recomendaciones de mitigación para proponer una redacción más precisa y medible del objetivo.
- **MP005.2: Asignación detallada de acciones, responsables internos/externos, plazos de ejecución y presupuestos estimados.** La herramienta permitirá desglosar los objetivos en acciones concretas, asignar responsables, establecer fechas de inicio y fin, y vincular presupuestos. **Modelos de ML, si se dispone de datos históricos suficientes (anonimizados entre clientes o de la propia empresa), podrían sugerir plazos realistas o rangos presupuestarios para tipos de acciones comunes.**

6. MP006: Integración inteligente y bidireccional de las recomendaciones de mitigación aceptadas en el plan de reducción.

Las medidas de mitigación que el usuario haya revisado y aceptado desde el módulo de recomendaciones (GR) se podrán incorporar automáticamente como acciones o proyectos dentro del plan de reducción. Esta integración será bidireccional; es decir, el progreso en la implementación de estas acciones en el plan se reflejará en el módulo de seguimiento, y a su vez, el impacto real de las medidas podrá retroalimentar al

motor de recomendaciones. Un **LLM podría ayudar a traducir una recomendación técnica en una serie de tareas más granulares para el plan de acción, identificando dependencias o prerequisites.**

7. **MP007: Sistema de versionado exhaustivo para planes e informes, con capacidad de comparación de versiones.**

Todos los planes de reducción e informes generados serán almacenados con un sistema de control de versiones completo. Esto permitirá a los usuarios acceder a iteraciones anteriores, rastrear la evolución de su estrategia de descarbonización y de sus reportes, y comparar diferentes versiones para identificar cambios. **Un LLM podría ser empleado para generar automáticamente un resumen de "cambios significativos" entre dos versiones de un plan o un informe extenso, facilitando la revisión de las modificaciones.**

8. **MP008: Implementación de flujos de trabajo (workflows) configurables para la revisión y aprobación interna de planes e informes.**

La plataforma soportará la definición de flujos de trabajo de aprobación personalizables. Las empresas podrán configurar secuencias de revisión y aprobación que involucren a diferentes roles o usuarios (ej. responsable de sostenibilidad, director financiero, dirección general) antes de que un informe se considere final o un plan de reducción se active. El sistema gestionará las notificaciones y el estado de cada documento dentro del flujo. **Un LLM podría realizar una verificación de completitud o cumplimiento de directrices internas antes de que un documento pase a la siguiente etapa del workflow.**

9. **MP009: Inclusión automática y contextual de gráficos y tablas de datos relevantes en los informes generados por IA.**

El motor de generación de informes no solo redactará texto, sino que también seleccionará e incrustará de forma inteligente las visualizaciones de datos (gráficos, tablas) más relevantes para cada sección del informe, basándose en el contenido que el LLM está generando. Por ejemplo, al describir las emisiones de Alcance 2, el sistema podría insertar automáticamente un gráfico de la evolución de estas emisiones y la contribución del enfoque de mercado. **El LLM también será responsable de generar pies de foto descriptivos y analíticos para estas visualizaciones, explicando qué muestran y cuáles son sus implicaciones.**

10. **MP010: Soporte para adjuntar documentación de respaldo y evidencia verificable, con posible análisis preliminar por IA.**

Los usuarios podrán cargar y adjuntar archivos (PDF, Excel, imágenes) como evidencia o documentación de respaldo directamente a secciones específicas de sus informes (ej. certificados de energía renovable, facturas de compra de equipos eficientes, informes de auditoría de terceros) o a acciones concretas de sus planes de reducción. Se explorará la capacidad de que modelos NLP/LLM realicen un análisis preliminar de los documentos adjuntados para extraer metadatos clave, verificar su relevancia para la sección a la que se adjuntan, o incluso resumir su contenido para facilitar la revisión por parte del usuario o auditores.

4. Diseño avanzado: Agente IA Colaborativo para la optimización de la descarbonización en Ecosistemas Multiagente (AIC-MAS)

Este apartado describe una visión a medio y largo plazo para una herramienta de Inteligencia Artificial de próxima generación: un **Agente IA Colaborativo (AIC)** diseñado para operar y prosperar dentro de **Sistemas Multiagente (MAS)** complejos y dinámicos. El objetivo fundamental del AIC-MAS es catalizar una descarbonización profunda y sistémica, no solo optimizando las operaciones de una única entidad, sino influyendo y coordinando las decisiones de múltiples actores interconectados para minimizar la huella de carbono colectiva, considerando integralmente los Alcances 1, 2 y 3.

4.1. Visión y objetivos estratégicos del Agente IA Colaborativo (AIC-MAS)

La transición hacia una economía neutra en carbono requiere un nivel de coordinación y optimización que excede las capacidades de los sistemas centralizados tradicionales. El AIC-MAS se concibe como una entidad inteligente autónoma que:

1. **Opera como un nodo de inteligencia para la descarbonización dentro de redes empresariales y sectoriales:** Se integra con los sistemas de planificación y operación de las empresas (y de otros agentes) para que las consideraciones de carbono sean un factor intrínseco en la toma de decisiones operativas y estratégicas.
2. **Aprovecha la conectividad continua con fuentes de datos heterogéneas:** Accede y procesa en tiempo real información de bases de datos internas de las empresas, legislación ambiental actualizada, factores de emisión dinámicos, datos de espacios de datos sectoriales (movilidad, industria, energía – incluyendo el mix eléctrico en tiempo real), condiciones de mercado y datos geoespaciales.
3. **Colabora y negocia con otros agentes IA (o sistemas legados representados por proxies de agentes):** Dentro de los ecosistemas empresariales (ej. agentes de producción, logística, compras, gestión energética) y entre empresas (cadena de suministro), el AIC-MAS buscará soluciones cooperativas que reduzcan la huella de carbono global, incluso si esto implica compromisos locales.
4. **Integra el ciclo completo de la gestión del carbono en sus deliberaciones:** El cálculo de la huella de carbono (Alcances 1, 2 y 3), la identificación de oportunidades de mitigación y la necesidad de auditoría y trazabilidad son elementos fundamentales en el proceso de razonamiento del agente al proponer o evaluar acciones.
5. **Extiende su influencia más allá de las operaciones directas:**
 - Genera **recomendaciones estratégicas y basadas en datos agregados para las administraciones públicas**, informando el diseño de políticas, incentivos y regulaciones más efectivas para la descarbonización.

- Facilita el **seguimiento agregado y la evaluación del progreso global o sectorial** hacia los objetivos de neutralidad climática, contribuyendo a una mayor transparencia y rendición de cuentas.

4.2. Arquitectura conceptual del Agente IA Colaborativo (AIC-MAS) y su ecosistema

El AIC-MAS no es una aplicación aislada, sino un componente sofisticado dentro de un ecosistema digital más amplio.

- **Componentes Centrales del Agente IA Colaborativo (AIC):**

1. **Módulo de Percepción y Adquisición de Conocimiento Distribuido (P-KAD):**

- **Subsistema de Conectores Dinámicos:** Interfaces robustas para enlazar con una multitud de fuentes de datos: APIs de bases de datos corporativas (ERP, SCM, MES), servicios web de legislación ambiental (ej. EUR-Lex), plataformas de espacios de datos (Gaia-X, IDS), APIs de operadores del sistema eléctrico (para mix de generación), servicios meteorológicos, y fuentes de datos de mercado.
- **Subsistema de Fusión de Sensores y Datos Heterogéneos:** Utiliza técnicas de IA (incluyendo LLMs y ontologías semánticas) para interpretar, validar, limpiar, normalizar y fusionar datos de diversas modalidades y niveles de granularidad, creando un modelo de conocimiento coherente y actualizado del entorno.
- **Repositorio de Conocimiento Semántico (Knowledge Graph):** Una base de conocimiento gráfica que representa entidades (empresas, procesos, tecnologías, normativas, emisiones) y sus relaciones, permitiendo un razonamiento complejo.

2. **Módulo de Razonamiento Estratégico y Planificación para la Descarbonización (SR-DP):**

- **Núcleo de Aprendizaje por Refuerzo Profundo Multiagente (MADRL):** El corazón del agente. Utiliza algoritmos MADRL (ej. MAPPO, MADDPG, QMIX) para aprender políticas de decisión óptimas en entornos donde las acciones de múltiples agentes interdependientes afectan el resultado global (la huella de carbono). El agente aprende a cooperar o competir estratégicamente para alcanzar objetivos de descarbonización.
- **Motor de Simulación y Predicción de Impacto de Carbono:** Un "Digital Twin" o modelo de simulación integrado que permite al agente predecir el impacto en emisiones (Alcances 1, 2 y 3) de diferentes secuencias de acciones o decisiones conjuntas antes de proponerlas o ejecutarlas. Este motor se alimenta de los datos del P-KAD y los modelos de huella de carbono.

- **Planificador Jerárquico de Objetivos:** Capaz de descomponer objetivos de descarbonización de alto nivel (ej. "reducir emisiones del Alcance 3 en un 20% en 5 años") en sub-objetivos y acciones coordinadas para los diferentes agentes involucrados.
3. **Módulo de Interacción, Comunicación y Negociación Multiagente (ICN-MA):**
- **Protocolos de Comunicación Multiagente:** Implementación de lenguajes de comunicación de agentes (ACL) estándar como FIPA-ACL, y protocolos de negociación (ej. Contract Net Protocol, subastas combinatorias) adaptados a los objetivos de descarbonización.
 - **Capacidades de Comunicación Basadas en LLM:** Utilización de LLMs para generar y comprender mensajes en lenguaje natural (o formal estructurado) con otros agentes (humanos o artificiales), explicar propuestas, justificar decisiones y facilitar la resolución de conflictos durante la negociación.
 - **Modelado de Otros Agentes (Opponent/Collaborator Modeling):** Capacidad de inferir los objetivos, creencias y posibles acciones de otros agentes en el sistema para anticipar sus respuestas y adaptar sus propias estrategias de colaboración o negociación.
4. **Módulo de Cálculo, Auditoría y Trazabilidad Continua de la Huella de Carbono (CAV-CFC):**
- **Motor de Cálculo de Huella de Carbono Dinámico e Integrado:** Realiza cálculos de emisiones (Alcances 1, 2 y 3) en tiempo real o bajo demanda, basándose en los datos de actividad que fluyen a través del P-KAD y las decisiones tomadas por los agentes.
 - **Subsistema de Trazabilidad Basado en Blockchain (potencial):** Exploración del uso de tecnologías de registro distribuido para crear un rastro inmutable y verificable de los datos de actividad, factores de emisión utilizados, cálculos realizados y acciones de mitigación implementadas, facilitando la auditoría externa y la creación de certificados de carbono fiables.
 - **Framework de Verificación y Alerta de Desviaciones:** Monitoriza continuamente la consistencia de los datos y los cálculos, alertando sobre posibles errores, fraudes o desviaciones respecto a los objetivos.
5. **Módulo de Generación de Recomendaciones Estratégicas y Políticas (SRP-Gen):**
- **Motor de Análisis Agregado y Síntesis de Conocimiento:** Utiliza técnicas de ML y LLMs para analizar los datos agregados de múltiples empresas y sectores (de forma anonimizada y respetando la privacidad) para identificar tendencias, barreras comunes y oportunidades de descarbonización a gran escala.
 - **Generador de Informes y Recomendaciones para Administraciones Públicas:** Basándose en el análisis anterior, el LLM

redacta informes y recomendaciones políticas (ej. sobre áreas prioritarias para inversión en I+D, diseño de incentivos fiscales, necesidad de nuevas infraestructuras) dirigidas a facilitar la transición ecológica.

- **Ecosistema de Interacción del AIC-MAS:**

1. **Agentes IA Empresariales y Sistemas Legados:** Otros agentes dentro de las empresas (gestión de la producción, optimización logística, trading de energía, gestión de la demanda) con los que el AIC-MAS debe interactuar y coordinarse.
2. **Plataformas de Espacios de Datos Soberanos:** Infraestructuras de compartición de datos (ej. Catena-X para automoción, futuros espacios de datos para energía o industria) que proporcionan acceso seguro y gobernado a datos sectoriales cruciales.
3. **Organismos Reguladores y Administraciones Públicas:** Entidades que reciben recomendaciones del AIC-MAS y que, a su vez, pueden establecer el marco normativo y los objetivos que el agente debe considerar.
4. **Plataformas de Auditoría y Certificación:** Sistemas externos que pueden interactuar con el módulo CAV-CFC para verificar la información de emisiones y la efectividad de las medidas de mitigación.

4.3. Capacidades fundamentales detalladas del Agente IA Colaborativo (AIC-MAS)

1. **CF-001: Conectividad y Fusión Semántica de Datos Heterogéneos en Tiempo Real:** El AIC-MAS se conectará de forma persistente a flujos de datos de alta velocidad y variedad (datos de sensores IoT industriales, transacciones de mercado eléctrico, previsiones meteorológicas, datos de movilidad urbana, actualizaciones legislativas). Utilizará ontologías compartidas y LLMs para interpretar semánticamente estos datos, resolver conflictos y construir un modelo de conocimiento unificado y dinámico del estado del ecosistema relevante para la descarbonización.
2. **CF-002: Modelado y Simulación Predictiva del Impacto de Carbono en Decisiones Multiagente Complejas:** Antes de comprometerse con una línea de acción o proponer una estrategia colaborativa, el AIC-MAS utilizará su motor de simulación integrado (potencialmente un "Digital Twin" del sistema socio-técnico) para ejecutar escenarios "what-if". Esto permitirá predecir con alta fidelidad las consecuencias en términos de emisiones de CO₂eq (desglosadas por Alcance 1, 2 y 3) de las decisiones coordinadas de múltiples agentes, considerando efectos de segundo y tercer orden.
3. **CF-003: Negociación y Colaboración Inteligente y Adaptativa con Otros Agentes (IA y Humanos):** El AIC-MAS empleará estrategias de negociación sofisticadas (derivadas de la teoría de juegos y el aprendizaje por refuerzo multiagente) para alcanzar acuerdos que beneficien los objetivos de descarbonización. Podrá adaptar su estilo de negociación (más competitivo o más cooperativo) en función del contexto y del comportamiento de los otros agentes. Los LLMs facilitarán la comunicación,

permitiendo al AIC-MAS explicar sus propuestas, entender las contrapropuestas y justificar sus razonamientos de forma comprensible.

4. **CF-004: Auditoría Continua y Trazabilidad Criptográficamente Segura de la Huella de Carbono en el Ecosistema MAS:** A través de su módulo CAV-CFC y la potencial integración con blockchain, el AIC-MAS mantendrá un registro auditable y a prueba de manipulaciones de todos los datos de actividad relevantes, los factores de emisión aplicados, los cálculos de huella de carbono realizados por los diferentes agentes (o para las actividades conjuntas) y la implementación de medidas de mitigación. Esto es crucial para la transparencia, la rendición de cuentas y la prevención del "greenwashing".
5. **CF-005: Generación Proactiva y Contextualizada de Estrategias de Mitigación Holísticas (Alcances 1, 2 y 3):** El AIC-MAS no se limitará a reaccionar, sino que buscará proactivamente oportunidades de reducción de emisiones. Esto puede implicar la reconfiguración de cadenas de suministro (para optimizar el Alcance 3 aguas arriba y aguas abajo), la promoción de simbiosis industriales (donde los residuos o subproductos de un agente se convierten en insumos para otro), la optimización del uso de energía renovable compartida, o la coordinación de la logística para minimizar las emisiones del transporte.
6. **CF-006: Elaboración de Recomendaciones Basadas en Evidencia y Análisis Predictivo para Administraciones Públicas:** Analizando los datos agregados (anonimizados) de las interacciones y los resultados obtenidos en el ecosistema MAS, el AIC-MAS podrá identificar cuellos de botella sistémicos, la efectividad de diferentes tipos de incentivos, o las necesidades de inversión en infraestructura. Utilizará LLMs para sintetizar estos hallazgos en informes claros y recomendaciones accionables para los responsables políticos, ayudando a diseñar intervenciones más efectivas.
7. **CF-007: Contribución al Seguimiento Agregado y la Visualización del Progreso de Descarbonización a Nivel Sectorial o Regional:** Al operar como un nodo en una red más amplia, el AIC-MAS puede contribuir con datos (debidamente agregados y anonimizados) a plataformas de seguimiento más amplias, permitiendo una visión en tiempo casi real del progreso de la descarbonización en un sector industrial, una cadena de valor o una región geográfica, facilitando la evaluación de políticas y la identificación de áreas rezagadas.

4.4. Tecnologías IA y habilitadoras clave para el AIC-MAS

El desarrollo y la operación exitosa del AIC-MAS dependerán de la madurez y la integración de varias tecnologías de vanguardia:

1. **TIA-001: Aprendizaje por Refuerzo Profundo Multiagente (Multi-Agent Deep Reinforcement Learning - MADRL):** Esencial para que el agente aprenda estrategias de colaboración y negociación efectivas en entornos complejos con múltiples actores con objetivos parcialmente alineados.
2. **TIA-002: Modelos de Lenguaje Extensos (LLMs) Avanzados:** Para la interpretación de información no estructurada (legislación, informes), la comunicación natural con otros agentes (humanos y artificiales), la generación de explicaciones, la redacción de informes y recomendaciones, y la asistencia en la planificación estratégica.

3. **TIA-003: Teoría de Juegos Algorítmica y Mecanismos de Diseño de Incentivos:** Para modelar las interacciones estratégicas entre agentes y diseñar mecanismos que promuevan la cooperación hacia la descarbonización.
4. **TIA-004: Digital Twins y Simulación Avanzada:** Para crear réplicas virtuales de sistemas empresariales, industriales y energéticos que permitan al AIC-MAS probar y evaluar el impacto de diferentes estrategias antes de su implementación real.
5. **TIA-005: Bases de Conocimiento Gráficas (Knowledge Graphs) y Ontologías Semánticas:** Para representar de forma estructurada y rica el conocimiento del dominio (empresas, procesos, tecnologías, regulaciones, emisiones) y permitir un razonamiento complejo y la inferencia de nuevas relaciones.
6. **TIA-006: Tecnologías de Registro Distribuido (Blockchain):** Para garantizar la transparencia, la inmutabilidad y la auditabilidad de los datos de emisiones, las acciones de mitigación y los certificados de carbono, aumentando la confianza en el sistema.
7. **TIA-007: Computación Confidencial y Técnicas de Preservación de la Privacidad (Privacy-Enhancing Technologies - PETs):** Crucial para permitir que los agentes colaboren y compartan información sensible (ej. datos de producción, costes) de forma segura, sin revelar información competitiva más allá de lo estrictamente necesario para los objetivos de descarbonización.

4.5. Desafíos significativos y líneas de investigación futuras

La realización plena del AIC-MAS enfrenta importantes desafíos que requieren investigación y desarrollo continuos:

1. **DEF-001: Escalabilidad y Eficiencia de los Algoritmos MADRL:** Entrenar y desplegar agentes MADRL en entornos con un gran número de actores y espacios de acción complejos sigue siendo un desafío computacional.
2. **DEF-002: Confianza, Explicabilidad y Robustez (Trustworthy AI):** Asegurar que las decisiones del AIC-MAS sean comprensibles, justificables, robustas ante perturbaciones o ataques, y alineadas con valores éticos es fundamental para su adopción.
3. **DEF-003: Seguridad en Entornos Multiagente Abiertos:** Proteger al AIC-MAS y al ecosistema de agentes maliciosos o comportamientos no deseados es un área crítica.
4. **DEF-004: Gobernanza de Datos y de las Interacciones entre Agentes:** Establecer reglas claras sobre la propiedad de los datos, el acceso, el uso y la resolución de conflictos en un ecosistema distribuido y soberano.
5. **DEF-005: Alineación de Incentivos y Manejo de Objetivos Divergentes:** Diseñar mecanismos para que agentes con objetivos primarios diferentes (ej. maximizar beneficio vs. minimizar emisiones) puedan encontrar soluciones colaborativas que avancen en la descarbonización.

6. **DEF-006: Interoperabilidad Semántica y Técnica:** Garantizar que agentes desarrollados por diferentes organizaciones y con diferentes tecnologías puedan comunicarse y colaborar de manera efectiva.

El AIC-MAS representa una frontera ambiciosa pero necesaria en la aplicación de la IA para la crisis climática. Su desarrollo progresivo permitirá pasar de optimizaciones aisladas a una gestión coordinada y sistémica de la descarbonización, acelerando la transición hacia un futuro sostenible.

5. Conclusiones

La necesidad de acometer una descarbonización profunda y acelerada de nuestras economías y sociedades exige la movilización de todas las herramientas a nuestro alcance. En este contexto, la Inteligencia Artificial (IA) emerge no como una solución única y mágica, sino como un conjunto de tecnologías habilitadoras con un potencial transformador sin precedentes. El presente informe ha delineado las especificaciones técnicas para tres sistemas basados en IA, cada uno con un alcance y una ambición crecientes, pero todos orientados a facilitar y optimizar los esfuerzos de reducción de la huella de carbono: el Asistente General para la Gestión Integral de la Descarbonización (AGID), el Asistente para la Interpretación y Cumplimiento de Legislación Ambiental (ALA), y el Agente IA Colaborativo para la Optimización de la Descarbonización en Ecosistemas Multiagente (AIC-MAS). El análisis de sus especificaciones nos permite extraer conclusiones significativas sobre el estado del arte y las futuras direcciones de la IA aplicada a la sostenibilidad.

5.1. El Asistente General para la Gestión Integral de la Descarbonización (AGID): la democratización de la acción climática empresarial

Las especificaciones detalladas para el AGID (Sección 4) revelan una arquitectura diseñada para la **automatización inteligente y la democratización del cálculo y gestión de la huella de carbono**. La clave reside en:

- **Ingesta de Datos Inteligente (especificaciones ID003-ID005, ID007):** La combinación de OCR avanzado, Modelos de Lenguaje Visual (VLM) para interpretar layouts complejos de facturas, y Procesamiento de Lenguaje Natural (NLP) basado en LLMs para la extracción semántica de entidades, transforma radicalmente la laboriosa tarea de recopilación de datos. Esto no solo reduce drásticamente el esfuerzo manual, sino que también mejora la precisión y la cobertura de los datos de actividad, que son la base de cualquier cálculo fiable.
- **Cálculo robusto y conforme a estándares (especificaciones HC001-HC005, HC013):** La adhesión estricta a metodologías como el GHG Protocol, el soporte para los tres Alcances (con especial énfasis en la complejidad del Alcance 3), y una base de datos de factores de emisión viva, potencialmente actualizada mediante RAG para consultar fuentes oficiales, garantizan la credibilidad y comparabilidad de los inventarios.
- **Recomendaciones de mitigación personalizadas y accionables (especificaciones GR001-GR010):** El motor de IA híbrido, que integra Machine Learning para la predicción del impacto y LLMs con RAG para la contextualización y explicación de medidas, representa un salto cualitativo. Pasa de ofrecer listas genéricas a proponer acciones específicas, cuantificadas en términos de reducción, coste y ROI, y justificadas de forma comprensible para la empresa.
- **Generación de informes asistida por LLM (especificaciones MP001-MP003, MP009):** La capacidad de los LLMs fine-tuned para redactar secciones narrativas de memorias de sostenibilidad e informes de huella de carbono, integrando datos y

gráficos, agiliza enormemente el proceso de reporting, permitiendo a las empresas centrarse en la estrategia y la acción en lugar de en la redacción.

En conjunto, las especificaciones del AGID apuntan a una herramienta que puede empoderar a un amplio espectro de organizaciones, especialmente PYMEs, para que gestionen activamente su impacto climático con un nivel de sofisticación que antes solo estaba al alcance de grandes corporaciones con equipos dedicados.

5.2. El Asistente para la Interpretación y Cumplimiento de Legislación Ambiental (ALA): navegando la complejidad regulatoria

Aunque sus especificaciones no se detallaron con el mismo nivel granular que el AGID en este informe, el concepto del ALA (Sección 3.1) se fundamenta en la capacidad de los **LLMs y la técnica RAG para desentrañar la complejidad del marco normativo ambiental**. Las especificaciones implícitas clave serían:

- **Acceso conversacional a información legal:** Un chatbot basado en LLM permitiría consultas en lenguaje natural sobre requisitos legislativos específicos.
- **Base de conocimiento actualizada vía RAG:** El LLM consultaría una base de datos viva de legislación, directivas y guías oficiales, asegurando respuestas basadas en la información más reciente.
- **Capacidad de resumen y explicación de textos legales:** El LLM simplificaría la jerga legal, facilitando la comprensión de las obligaciones.

Un ALA así especificado se convertiría en un aliado indispensable para las empresas, reduciendo el riesgo de incumplimiento, ahorrando costes de consultoría legal especializada y democratizando el acceso al conocimiento regulatorio, un factor crítico para la planificación de la descarbonización.

5.3. El Agente IA Colaborativo en Entornos Multiagente (AIC-MAS): hacia la descarbonización sistémica y cooperativa

Las especificaciones propuestas para el AIC-MAS (Sección 5) representan una visión de futuro donde la IA no solo optimiza entidades individuales, sino que **orquesta la colaboración para la descarbonización a nivel de ecosistemas complejos**. Sus especificaciones definitorias incluyen:

- **Conectividad masiva y fusión semántica de datos (especificación CF-001):** La capacidad de conectarse a espacios de datos industriales, de movilidad, legislativos y energéticos, y de fusionar esta información heterogénea mediante ontologías y LLMs en un modelo de conocimiento coherente, es fundamental para una visión holística.
- **Razonamiento estratégico y planificación multiagente (especificación CF-002, TIA-001):** El uso de Aprendizaje por Refuerzo Profundo Multiagente (MADRL) permitiría al AIC-MAS aprender políticas de decisión que optimicen la huella de

carbono del sistema global, no solo de sus partes. Esto implica la capacidad de simular el impacto de decisiones coordinadas.

- **Negociación y comunicación avanzada (especificación CF-003, TIA-002):** La interacción con otros agentes (IA o humanos) mediante protocolos de comunicación y la asistencia de LLMs para la explicación y justificación de propuestas es crucial para lograr acuerdos colaborativos en cadenas de valor o simbiosis industriales.
- **Auditoría continua y trazabilidad (especificación CF-004, TIA-006):** La integración de cálculos de huella dinámicos (Alcances 1, 2 y 3) y la potencial aplicación de blockchain para la trazabilidad de acciones y emisiones, aportarían un nivel de transparencia y fiabilidad sin precedentes, esencial para mercados de carbono y la verificación de compromisos.
- **Generación de recomendaciones a Administraciones (especificación CF-006):** La capacidad de analizar datos agregados y proponer políticas basadas en evidencia a los reguladores cierra el bucle entre la acción empresarial y el marco habilitador público.

El AIC-MAS, tal como se especifica, trasciende la optimización local para abordar la descarbonización como un problema de optimización de sistemas complejos interconectados. Representa un cambio de paradigma hacia la inteligencia colectiva aplicada a la sostenibilidad.

5.4. Conclusiones generales e implicaciones transversales

De la revisión de estas tres propuestas de sistemas y sus especificaciones, emergen varias conclusiones transversales:

- **La Centralidad de los Datos y su Procesamiento Inteligente:** Todas las soluciones dependen críticamente de la disponibilidad, calidad y procesamiento avanzado de datos. Las especificaciones de ingesta del AGID, la conexión a bases de conocimiento del ALA y la fusión de datos heterogéneos del AIC-MAS subrayan que la IA para la descarbonización es, en gran medida, una IA intensiva en datos.
- **El Auge de los LLMs y RAG como Tecnologías Fundamentales:** Los Modelos de Lenguaje Extensos, especialmente cuando se combinan con la Generación Aumentada por Recuperación, se perfilan como herramientas versátiles para la interacción con el usuario, la interpretación de documentos, la generación de texto, la explicación de recomendaciones y el acceso a conocimiento especializado y actualizado.
- **La Necesidad de Enfoques Híbridos de IA:** La combinación de diferentes técnicas de IA (ML para predicción, LLMs para lenguaje, MADRL para estrategia multiagente, VLM para visión) es esencial para abordar la multifacética naturaleza de los desafíos de descarbonización.
- **El Potencial de la IA para Reducir Barreras:** Tanto el AGID como el ALA están diseñados para simplificar procesos complejos y hacerlos accesibles a un público más amplio, reduciendo la dependencia de la experticia humana altamente especializada y costosa.

- **La Evolución hacia la Inteligencia Colaborativa:** El AIC-MAS ilustra una trayectoria desde la optimización individual hacia la optimización de sistemas, donde la IA facilita la cooperación y la coordinación entre múltiples actores para lograr objetivos comunes de sostenibilidad que serían inalcanzables de forma aislada.
- **Desafíos Éticos y de Gobernanza:** A medida que estos sistemas se vuelven más potentes y autónomos (especialmente el AIC-MAS), la necesidad de marcos éticos robustos, explicabilidad (XAI), seguridad, privacidad y gobernanza de datos se vuelve primordial. Las especificaciones de seguridad (SC) y de modelos de IA (AI) para el AGID ya apuntan en esta dirección.

En definitiva, las especificaciones técnicas presentadas no son meros ejercicios teóricos, sino un esbozo de las capacidades que la Inteligencia Artificial puede y debe aportar a la lucha contra el cambio climático. Desde la automatización de tareas fundamentales y la provisión de inteligencia accionable a nivel de empresa, hasta la facilitación de la comprensión regulatoria y la orquestación de la colaboración a gran escala, la IA se posiciona como un aliado estratégico indispensable. La materialización de estas soluciones requerirá inversión continua en investigación y desarrollo, colaboración entre el sector público y privado, y un compromiso firme con la aplicación ética y responsable de estas poderosas tecnologías para forjar un futuro más sostenible.

6. Anexo. Casos de uso sectoriales

6.1. Anexo A. Casos de uso sectoriales para herramientas de IA en descarbonización

Este anexo explora la aplicación potencial de las herramientas de Inteligencia Artificial descritas en el informe – el Asistente General para la Gestión Integral de la Descarbonización (AGID), el Asistente para la Interpretación y Cumplimiento de Legislación Ambiental (ALA), y el Agente IA Colaborativo para la Optimización de la Descarbonización en Ecosistemas Multiagente (AIC-MAS) – en diez sectores económicos clave. Para cada sector, se identifica un caso de uso específico, los beneficios potenciales y los desafíos de aplicación.

A.1. SECTOR MANUFACTURERO (INDUSTRIA GENERAL)

- **Herramienta(s) IA principal(es) aplicables:** AGID (primordial), ALA, AIC-MAS (para cadenas de suministro complejas).
- **Caso de uso específico:** Una empresa de fabricación de componentes metálicos utiliza el **AGID** para automatizar el cálculo de su huella de carbono. El sistema ingiere facturas de electricidad y gas natural (Alcance 2 y 1), datos de consumo de gasóleo de la flota de carretillas elevadoras (Alcance 1) y datos de compra de acero y otros metales (base para Alcance 3, Categoría 1). El motor de IA del AGID analiza los patrones de consumo energético y recomienda la optimización de los ciclos de los hornos y la inversión en sistemas de recuperación de calor residual, estimando ahorros y reducción de tCO₂eq. El **ALA** es consultado para entender las obligaciones de reporte de emisiones bajo el régimen ETS si la planta supera ciertos umbrales.
- **Beneficios potenciales**
 - Reducción significativa del tiempo y coste asociado al cálculo manual de la huella.
 - Identificación precisa de los principales focos de emisión y oportunidades de eficiencia energética.
 - Recomendaciones de mitigación personalizadas con análisis de viabilidad económica.
 - Mejora del cumplimiento normativo y preparación para auditorías.
- **Desafíos de aplicación**
 - Disponibilidad y calidad de los datos de actividad, especialmente para el Alcance 3 (datos de proveedores).
 - Integración del AGID con sistemas de planificación de recursos empresariales (ERP) o de ejecución de manufactura (MES) legados.

- Resistencia al cambio en la adopción de nuevas tecnologías o procesos productivos recomendados.
- Coste inicial de implementación de la solución AGID y de las medidas de mitigación.

A.2. SECTOR ENERGÉTICO (GENERACIÓN, TRANSMISIÓN Y DISTRIBUCIÓN)

- **Herramienta(s) IA principal(es) aplicables:** AGID (para operaciones propias), AIC-MAS (para optimización de red y mercado).
- **Caso de uso específico:** Múltiples generadores de energía (renovable y fósil), operadores de red y grandes consumidores industriales participan en un ecosistema gestionado por **AIC-MAS**. Los agentes IA de cada entidad colaboran para optimizar el despacho de energía, minimizando las emisiones del mix de generación en tiempo real. El AIC-MAS utiliza previsiones de demanda y generación renovable (eólica, solar) para coordinar el almacenamiento de energía y la respuesta de la demanda. Recomienda inversiones en la red para reducir pérdidas y facilitar la integración de más renovables. El **AGID** es usado por cada central para su huella operativa específica (ej. emisiones fugitivas de CH₄ en plantas de gas).
- **Beneficios potenciales**
 - Maximización de la penetración de energías renovables y minimización del uso de combustibles fósiles.
 - Reducción de las pérdidas en transmisión y distribución.
 - Mayor estabilidad y resiliencia de la red eléctrica.
 - Optimización de las inversiones en infraestructura energética a nivel sistémico.
- **Desafíos de aplicación**
 - Necesidad de marcos de compartición de datos seguros y soberanos entre entidades competidoras.
 - Complejidad algorítmica del aprendizaje por refuerzo multiagente (MADRL) para la optimización de sistemas a gran escala.
 - Estandarización de protocolos de comunicación y ontologías semánticas entre los diversos agentes.
 - Gobernanza del sistema AIC-MAS y alineación de incentivos entre los participantes.

A.3. SECTOR TRANSPORTE Y LOGÍSTICA

- **Herramienta(s) IA principal(es) aplicables:** AGID (para flotas individuales), AIC-MAS (para optimización de cadenas logísticas).

- **Caso de uso específico:** Una gran empresa de logística utiliza **AGID** para calcular las emisiones de su flota de camiones (Alcance 1) y de sus almacenes (Alcance 2). El AGID recomienda la renovación de la flota hacia vehículos eléctricos o de hidrógeno, y optimiza las rutas de entrega individuales. A un nivel superior, esta empresa participa en un ecosistema **AIC-MAS** con otras empresas de transporte, cargadores y receptores. Los agentes IA colaboran para consolidar cargas, optimizar rutas multimodales (carretera, ferrocarril, marítimo) entre diferentes proveedores, y reducir los kilómetros en vacío, minimizando la huella de carbono de toda la cadena logística.
- **Beneficios potenciales**
 - Reducción del consumo de combustible y de las emisiones de GEI de las flotas.
 - Optimización de la ocupación de vehículos y reducción de costes operativos.
 - Mejora de la eficiencia de las cadenas de suministro y reducción de emisiones de Alcance 3 para los clientes.
 - Facilitación de la transición hacia combustibles y tecnologías de transporte más limpias.
- **Desafíos de aplicación**
 - Integración de datos en tiempo real (tráfico, disponibilidad de puntos de recarga, etc.).
 - Compartición de datos sensibles (rutas, clientes) entre empresas competidoras en el AIC-MAS.
 - Heterogeneidad de las flotas y de la infraestructura de combustibles alternativos.
 - Complejidad de la optimización dinámica en redes logísticas extensas.

A.4. SECTOR AGRÍCOLA Y ALIMENTARIO

- **Herramienta(s) IA principal(es) aplicables:** AGID, ALA.
- **Caso de uso específico:** Una cooperativa agrícola implementa **AGID** para ayudar a sus agricultores miembros a calcular la huella de carbono de sus explotaciones. Considera emisiones de la fermentación entérica del ganado, gestión del estiércol, uso de fertilizantes nitrogenados (fuente de N₂O), y consumo de energía en maquinaria y regadío. El AGID sugiere prácticas de agricultura regenerativa, optimización del uso de fertilizantes mediante agricultura de precisión, y mejoras en la gestión del estiércol. El **ALA** informa a los agricultores sobre las normativas relativas a la gestión de nitratos y los posibles incentivos para prácticas de secuestro de carbono en suelo.
- **Beneficios potenciales**

- Identificación de las principales fuentes de emisión en las explotaciones agrarias.
- Adopción de prácticas agrícolas más sostenibles y con menor impacto climático.
- Potencial acceso a mercados de carbono agrícola o certificaciones de sostenibilidad.
- Mejora de la salud del suelo y reducción de la contaminación por nutrientes.
- **Desafíos de aplicación**
 - Dificultad en la medición precisa de emisiones difusas (ej. N₂O de suelos).
 - Variabilidad biológica y dependencia de factores climáticos.
 - Necesidad de capacitación y apoyo técnico para la adopción de nuevas prácticas por parte de los agricultores.
 - Fragmentación del sector y diversidad de tipos de explotación.

A.5. SECTOR CONSTRUCCIÓN Y PROMOCIÓN INMOBILIARIA

- **Herramienta(s) IA principal(es) aplicables:** AGID, ALA.
- **Caso de uso específico:** Una empresa promotora utiliza **AGID** para evaluar y reducir el carbono embebido (Alcance 3, Categoría 1) de sus nuevas construcciones. El sistema se integra con bases de datos de Declaraciones Ambientales de Producto (DAP) para materiales de construcción (hormigón, acero, aislamiento, etc.) y permite comparar el impacto de diferentes diseños y selecciones de materiales. También calcula las emisiones operativas (Alcance 1 y 2) de la fase de construcción (maquinaria, energía de obra). El **ALA** proporciona información actualizada sobre códigos de edificación sostenible, requisitos de eficiencia energética y certificaciones tipo BREEAM o LEED.
- **Beneficios potenciales**
 - Diseño y construcción de edificios con menor huella de carbono (tanto embebida como operativa).
 - Optimización del uso de materiales y reducción de residuos de construcción.
 - Cumplimiento de normativas de edificación sostenible y obtención de certificaciones.
 - Mayor valor de los activos inmobiliarios sostenibles.
- **Desafíos de aplicación**

- Disponibilidad y estandarización de datos de DAP para todos los materiales de construcción.
- Complejidad del análisis del ciclo de vida de los edificios.
- Coordinación entre múltiples actores en el proceso constructivo (arquitectos, ingenieros, contratistas, proveedores).
- Coste inicial potencialmente mayor de algunos materiales o soluciones constructivas sostenibles.

A.6. SECTOR RETAIL (COMERCIO MINORISTA) Y BIENES DE CONSUMO

- **Herramienta(s) IA principal(es) aplicables:** AGID (para operaciones y Alcance 3), AIC-MAS (para la cadena de suministro).
- **Caso de uso específico:** Una cadena de supermercados emplea **AGID** para calcular las emisiones de sus tiendas (electricidad para iluminación y refrigeración – Alcance 2; fugas de refrigerantes – Alcance 1) y de su logística de distribución. Para el complejo Alcance 3, el AGID ayuda a estimar las emisiones de los bienes comprados y vendidos mediante modelos de gasto-emisión o datos de proveedores. A través de una plataforma **AIC-MAS**, la cadena colabora con sus principales proveedores de bienes de consumo para obtener datos más precisos de la huella de carbono de los productos (PFC) y co-diseñar productos con menor impacto (ej. envases más ligeros, ingredientes de origen sostenible).
- **Beneficios potenciales**
 - Reducción de la huella de carbono operativa de las tiendas y la logística.
 - Mayor transparencia y capacidad de gestión de las emisiones de Alcance 3.
 - Oferta de productos más sostenibles a los consumidores.
 - Mejora de la reputación de marca y fidelización de clientes concienciados.
- **Desafíos de aplicación**
 - Enorme complejidad del Alcance 3 debido a la diversidad de productos y proveedores.
 - Dificultad para obtener datos fiables de PFC de toda la cadena de suministro.
 - Influencia limitada sobre las prácticas de fabricación de proveedores globales.
 - Necesidad de educar al consumidor para fomentar la demanda de productos sostenibles.

A.7. SECTOR TIC (TECNOLOGÍAS DE LA INFORMACIÓN Y COMUNICACIÓN)

- **Herramienta(s) IA principal(es) aplicables:** AGID.
- **Caso de uso específico:** Un proveedor de servicios en la nube utiliza **AGID** para monitorizar y calcular de forma continua la huella de carbono de sus centros de datos. Se enfoca en el consumo eléctrico (Alcance 2), la eficiencia energética (PUE – Power Usage Effectiveness), las emisiones de generadores diésel de respaldo (Alcance 1) y el carbono embebido en los servidores y equipos de red (Alcance 3). El AGID recomienda estrategias de optimización del enfriamiento, consolidación de servidores mediante virtualización, y la adquisición de energía renovable a través de PPAs.
- **Beneficios potenciales**
 - Reducción del consumo energético y de los costes operativos de los centros de datos.
 - Mejora de la eficiencia y reducción de la huella de carbono por unidad de cómputo.
 - Mayor atractivo para clientes que buscan soluciones cloud sostenibles.
 - Contribución a la descarbonización del sector digital.
- **Desafíos de aplicación**
 - Alta densidad energética y rápido crecimiento de la demanda de servicios digitales.
 - Ciclos de vida cortos del hardware y dificultad para cuantificar el carbono embebido.
 - Necesidad de asegurar la fiabilidad y disponibilidad del servicio 24/7.
 - Transparencia en la cadena de suministro de componentes electrónicos.

A.8. SECTOR SANITARIO

- **Herramienta(s) IA principal(es) aplicables:** AGID, ALA.
- **Caso de uso específico:** Un hospital utiliza **AGID** para analizar su huella de carbono, que incluye el consumo energético de edificios e equipos médicos (Alcance 2), el uso de gases anestésicos con alto PCA (Potencial de Calentamiento Atmosférico) como el desflurano (Alcance 1), y las emisiones asociadas a la gestión de residuos sanitarios y la cadena de suministro de productos farmacéuticos y material médico (Alcance 3). El AGID recomienda la sustitución de gases anestésicos por alternativas de menor impacto, la mejora de la eficiencia energética de los edificios y la optimización de las rutas de gestión de residuos. El **ALA** ayuda a cumplir con la compleja normativa sobre residuos peligrosos y emisiones.
- **Beneficios potenciales**

- Reducción de costes operativos a través de la eficiencia energética.
- Minimización del impacto ambiental de los gases anestésicos.
- Mejora de la gestión de residuos y reducción de su huella de carbono.
- Creación de un entorno hospitalario más saludable y sostenible.
- **Desafíos de aplicación**
 - Priorización de la atención al paciente y la seguridad clínica sobre algunas medidas ambientales.
 - Complejidad de la cadena de suministro farmacéutica y de material médico.
 - Gestión especializada y costosa de los residuos sanitarios.
 - Necesidad de formación y concienciación del personal sanitario.

A.9. SECTOR TURISMO Y HOSTELERÍA

- **Herramienta(s) IA principal(es) aplicables:** AGID.
- **Caso de uso específico:** Una cadena hotelera implementa **AGID** para que cada uno de sus establecimientos calcule y gestione su huella de carbono. Se miden consumos de energía (electricidad, gas para calefacción y agua caliente – Alcance 1 y 2), agua, generación de residuos, y se estima el impacto de los alimentos y bebidas adquiridos (Alcance 3). El AGID sugiere la instalación de sistemas de gestión energética en habitaciones, la optimización de lavanderías, la compra de productos locales y de temporada, y programas de reducción de desperdicio alimentario.
- **Beneficios potenciales**
 - Reducción de costes de suministros (energía, agua).
 - Mejora de la imagen de marca y atracción de turistas eco-conscientes.
 - Cumplimiento con criterios de certificaciones de turismo sostenible.
 - Contribución a la preservación de los destinos turísticos.
- **Desafíos de aplicación**
 - Alta variabilidad de la ocupación y estacionalidad.
 - Dificultad para influir en el comportamiento de los huéspedes.
 - Complejidad en el seguimiento de la huella de carbono de la cadena de suministro de alimentos y bebidas.
 - Necesidad de inversión en tecnologías eficientes o renovables.

A.10. SECTOR PÚBLICO Y PLANIFICACIÓN URBANA

- **Herramienta(s) IA principal(es) aplicables:** ALA, AIC-MAS.
- **Caso de uso específico:** Un ayuntamiento utiliza **AIC-MAS** como herramienta de apoyo a la decisión para su plan de acción climática. Agentes IA modelan diferentes subsistemas urbanos (transporte, edificación, energía, residuos). El AIC-MAS simula el impacto en emisiones de diferentes políticas (ej. peatonalización de zonas, fomento del transporte público eléctrico, incentivos para la rehabilitación energética de edificios, implantación de redes de calor urbanas). El **ALA** asegura que los planes municipales estén alineados con los objetivos y la legislación climática nacional y regional. El AIC-MAS puede también facilitar el seguimiento agregado del progreso de la ciudad hacia sus metas de descarbonización.
- **Beneficios potenciales**
 - Planificación urbana basada en datos y optimizada para la reducción de emisiones.
 - Identificación de las políticas más coste-efectivas para la descarbonización urbana.
 - Mejora de la calidad del aire y de la calidad de vida de los ciudadanos.
 - Mayor capacidad para cumplir con los compromisos climáticos de la ciudad.
- **Desafíos de aplicación**
 - Necesidad de integrar datos de fuentes muy diversas (tráfico, catastro, consumo energético, etc.).
 - Complejidad del modelado de sistemas urbanos interdependientes.
 - Consideraciones políticas y sociales en la toma de decisiones que pueden primar sobre la optimización puramente técnica.
- Largos horizontes temporales para la implementación y la observación del impacto de las políticas urbanas.



Con el apoyo de:



Elaborado por:

ecodes
tiempo de actuar