

Inteligencia artificial y big data

INFORME SITUACIÓN 2022

DISRUPTIV

Plataforma Tecnológica Española
de Tecnologías Disruptivas

Con financiación de:



Convocatoria 2020 Plataformas Tecnológicas
y de Innovación (PTR2020-001159)

Secretaría técnica a cargo de:



INDICE

- 03** Introducción
- 04** Nuevas tendencias
- 07** Ecosistema
- 09** Retos y oportunidades
- 15** Casos de uso
- 17** Enlaces de interés

INTRODUCCIÓN

El objetivo de este informe es actualizar los datos del informe previo de 2021 sobre la situación actual de la tecnología en España.

En este informe se ha recopilado información relativa a las novedades en tendencias, ecosistema, retos y oportunidades y casos de uso



Este informe pretende ser una foto del momento actual en el que se encuentra dicha tecnología en España y que sirva para poder comparar la situación en un futuro próximo.

NUEVAS TENDENCIAS



Entre las tendencias de IA-BD de 2022 destacan el desarrollo de modelos distribuidos basados en el **aprendizaje automático distribuido** y **aprendizaje federado**, así como propiciar una mejor comprensión de los usuarios de las variables involucradas en el diseño de los algoritmos, aplicando la **transparencia de los modelos de aprendizaje**, que permitan una mejor toma de decisión en nuevos modelos de negocio utilizado Big Data Analytics, así como una mayor concienciación sobre la necesidad de promover una **IA sostenible o "verde"**.

¿Qué es el aprendizaje automático distribuido?

La aparición de tecnologías como big data (volumen), smart data (calidad), IoT-Internet of Things (ubicuidad) y la masificación del uso de 5G (velocidad) ha dado lugar a varios conjuntos de datos distribuidos y estos grandes conjuntos de datos almacenados en un repositorio central imponen enormes requisitos de procesamiento e informática.

La forma tradicional de utilizar herramientas integradas para la extracción de datos y su posterior análisis ya no es práctica, ya que los datos son demasiado grandes y diversos para gestionarlos.

El aprendizaje automático distribuido permite a las empresas tomar decisiones informadas y sacar conclusiones significativas a partir de grandes cantidades de datos.

Si bien existen desventajas asociadas al ML distribuido son inferiores a todos los ahorros de costes en mantenimiento, personal y energía que administrar un gran servidor central, por lo que muchos investigadores afirman que este procesamiento distribuido de datos es la plataforma informática adecuada para enfrentar el continuo cambio de los datos del presente y futuro, para que sean escalables, flexibles y eficientes.



¿Qué es aprendizaje federado?

El aprendizaje federado aplica cuando el aprendizaje se realiza de modo cooperativo, un tema de gran importancia para las cadenas de valor y para ámbitos como el logístico, debido a la forma orquestada de aprovecharse el conocimiento. El aprendizaje federado hace que todas las iteraciones del entrenamiento se realicen en dispositivos locales, por lo que no pone en riesgo los datos originales ni los expone.



¿A qué llamamos transparencia de modelos?

La transparencia de los modelos está relacionada con las buenas prácticas que deben acometerse en IA y tiene que ver con la posibilidad de que los usuarios conozcan cuáles son los criterios de la toma de decisiones del algoritmo. Es decir, la transparencia de los modelos busca que no sea una preocupación el cómo se hace, sino que se tenga la certidumbre de la respuesta. Además, está muy conectado a todo el tema regulatorio y estandarización de la IA (Data Act) junto con a la ética en la IA-BD.

¿Qué es la IA verde?

La IA verde es la IA respetuosa con el medio ambiente y en ese sentido, se puede abordar desde 2 ámbitos (Green by design y Green by AI):

- 1) Apuesta por el diseño de algoritmos y hardware que, en sí mismos, sean más sostenibles y produzcan un menor gasto energético (Green by design).
- 2) Algoritmos que, en su uso, contribuyen a encontrar soluciones eficaces energéticamente en cualquier industria. (Green by AI).

En España se lanzó en 2021 el Programa Nacional de Algoritmos Verdes para fomentar la colaboración público-privada en el desarrollo de algoritmos Green by design. Este programa busca crear un ecosistema y hubs de datos que se regule con una perspectiva sostenible.

Tendencias del Big Data:

Según IEBS Digital School, estas son algunas de las tendencias más importantes del big data:

- **Data Strategy:** la estrategia de datos integrada en la estrategia de crecimiento de las empresas.
- **Datos como servicio:** utilizan la tecnología de la nube para dar a los usuarios o a las aplicaciones la información sin depender de dónde puedan estar estos usuarios o aplicaciones
- **Data Lakehouse:** Un Data Lakehouse es un Datawarehouse que tiene como almacenamiento de datos a un Data Lake.



ECOSISTEMA:



Hub Español de Gaia X:

Gaia-X es una iniciativa europea del sector privado cuyo objetivo es la creación de una infraestructura de datos abierta, federada y soberana, enfocada al negocio y constituida sobre los valores de soberanía digital y disponibilidad de los datos. Su principal objetivo es acelerar la respuesta europea en materia de compartición de datos y soberanía.

La iniciativa logra la capilaridad en el territorio de los diferentes estados miembro a través de 15 hubs nacionales de Gaia-X.

El hub español de Gaia-X busca que las industrias sectoriales de todos los tamaños, con especial atención para las PYMEs y microPYMEs, creen comunidad en torno al dato, desarrollen nuevas soluciones y fomenten la innovación de datos e IA. El objetivo de este hub nacional es acelerar la capacidad española en materia de compartición de datos y soberanía digital: desplegar la economía del dato y apostar por el liderazgo de espacios de datos en sectores estratégicos como turismo o salud; así como generar un ecosistema sólido en el ámbito de la compartición de datos industriales integrado por centros de investigación y desarrollo tecnológico, grandes empresas, PYMEs, microempresas y startups.

El 18 de marzo de 2022 tuvo lugar en Talavera de la Reina el acto de la Asamblea Constituyente de la asociación del hub español de Gaia-X. Los vocales designados para la creación de la Asociación fueron los siguientes:

La Junta de Comunidades de Castilla-La Mancha, la Red Española de Ciudades Inteligentes, AMETIC, INDESIA-Asociación Industrial para el Impulso de la Economía del Dato y de la Inteligencia Artificial, el Instituto Tecnológico de Informática, Tecnalía, Fundación CTIC Centro Tecnológico, Aire Networks, Telefónica, Amadeus IT Group, Inetum, Universidad Politécnica de Madrid, Fundación General CSIC, Cruz Roja, Clúster Big Data Madrid, EIT Urban Mobility, Gigas Hosting, AnySolution, Tinanmica, Eccocar Sharing, Cloudglobal y la Confederación Española de Organizaciones Empresariales.



ECOSISTEMA:

El Chief Data Officer de España asume la presidencia del Consejo Gubernamental de Gaia X:

El pasado 11 de julio de 2022 se informó de que Alberto Palomo Lozano, primer Chief Data Officer del Gobierno de España, asumía la presidencia del Consejo Gubernamental (Governmental Advisory Board) de Gaia-X a nivel europeo.

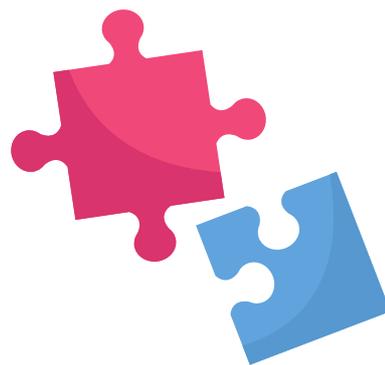
Entre sus funciones se incluyen las labores de coordinación de las actividades del Consejo, la facilitación de los debates y el proceso de toma de decisiones tanto dentro como fuera de las reuniones del organismo y su orientación en el cumplimiento de las funciones y tareas propias del Consejo.

Alianza por la Nueva Economía de la Lengua:

La colaboración público-privada del PERTE de Economía de la Lengua se articula a través de la Alianza para la Nueva Economía de la Lengua, cuya comisionada es Cristina Gallach. Esta Alianza tiene como objetivo sumar distintos agentes en el impulso de las medidas propuestas y de proyectos complementarios alineados con los objetivos del PERTE. Además, contará con un Consejo Consultivo formado por personalidades vinculadas a la promoción de la lengua y su digitalización en el ámbito de la tecnología, la ciencia y la cultura. También estarán representadas instituciones como el Instituto Cervantes, la Biblioteca Nacional de España, la Secretaría General de Estados Iberoamericanos y la Organización de Estados Iberoamericanos, entre otros. Y está previsto que se celebre anualmente un evento con sede itinerante para realizar actividades de encuentro y divulgación en el que se fomentará la participación de empresas, universidades y redes profesionales.

El primero de estos eventos se celebró el pasado 28 de septiembre en Alicante y reunió a más de 100 empresas punteras en Procesamiento del Lenguaje Natural con motivo de la puesta en marcha del Foro de la Alianza por la Nueva Economía de la Lengua.

RETOS Y OPORTUNIDADES



Según el estudio publicado por el **ONTSI** en enero de 2022, solo un 8% de empresas españolas utilizan IA. Madrid es la comunidad donde más empresas han adoptado tecnologías de IA: un 11% del total.

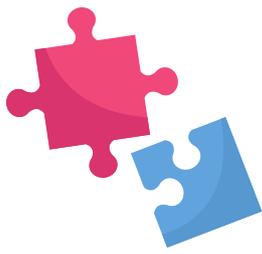
Sin embargo, en el **Hype Cycle de Gartner** de agosto de 2022, entre las 25 tecnologías emergentes que identifica, señala la oportunidad en la aceleración de la automatización de la IA. En relación con esta oportunidad concreta las siguientes tecnologías:

La inteligencia artificial causal: identifica y utiliza las relaciones de causa y efecto para ir más allá de los modelos predictivos basados en la correlación y hacia sistemas de IA que puedan prescribir acciones de forma más eficaz y actuar de forma más autónoma.

Modelos fundacionales: son modelos basados en arquitecturas transformadoras, como los grandes modelos lingüísticos, que encarnan un tipo de arquitectura de red neuronal profunda que computa una representación numérica del texto en el contexto de las palabras circundantes, haciendo hincapié en las secuencias de palabras.

Diseño generativo o el diseño aumentado por la IA: es el uso de tecnologías de IA, aprendizaje automático (ML) y procesamiento del lenguaje natural (NLP) para generar y desarrollar automáticamente flujos de usuario, diseños de pantalla, contenido y código de capa de presentación para productos digitales.

Herramientas de generación de código de aprendizaje automático: incluyen modelos de ML alojados en la nube que se conectan a los entornos de desarrollo integrados (IDE) de los desarrolladores profesionales, que son extensiones que ofrecen sugerencias de código basadas en descripciones en lenguaje natural o fragmentos de código parciales.



RETOS Y OPORTUNIDADES

Con respecto a Big Data, el estudio del **ONTSI** informa de que un 11% de las empresas analiza grandes cantidades de datos; una cifra que aumenta en el caso de las grandes empresas, donde lo hace el 29%. Su uso lo ha facilitado y hecho más asequible la existencia de herramientas analíticas más sencillas y el cloud computing.

En el **Hype Cycle de Gartner de agosto de 2022** se destacan un bloque de tecnologías emergentes relacionadas con la entrega de tecnología optimizada (optimized Technology delivery) y que se centran en los componentes clave de la creación de un negocio digital y entre las cuales se citan las siguientes:

Sostenibilidad de la nube: es el uso de los servicios en la nube para lograr beneficios de sostenibilidad dentro de los sistemas económicos, medioambientales y sociales.

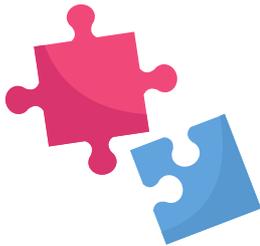
Almacenamiento computacional (CS): descarga el procesamiento del host de la memoria principal de la unidad central de procesamiento (CPU) al dispositivo de almacenamiento.

La arquitectura de malla de ciberseguridad (CSMA): es un enfoque emergente para la arquitectura de controles de seguridad distribuidos y comprobables que mejoran la eficacia general de la seguridad.

Observabilidad de los datos: es la capacidad de comprender la salud del panorama de los datos de una organización, los conductos de datos y la infraestructura de datos mediante la supervisión, el seguimiento, la alerta, el análisis y la resolución de incidentes de forma continua.

Respuesta al reto de disponer de más datos: se crea el Hub Español de Gaia - X:

El pasado 18 de marzo se creaba en la Asociación Gaia-X España en Talavera de la Reina (Toledo), con el que se crea un hub de la economía del dato que permitirá desarrollar el tejido empresarial nacional centrado en la creación de espacios de datos abiertos para promover su compartición de forma segura.



RETOS Y OPORTUNIDADES

Esta iniciativa se enmarca en la Agenda España Digital que impulsa el Gobierno, que tiene entre sus ejes avanzar hacia una economía del dato, aprovechando las oportunidades que ofrecen la inteligencia artificial y el Big Data para que al menos un 25% de las empresas usen estas tecnologías en 2025.

Convocatorias de ayudas y proyectos estratégicos relacionadas con la IA y BD:

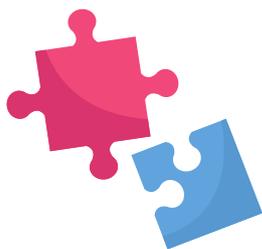
Sin duda los Fondos Next Generation de la Comisión Europea, la Agenda España Digital 2026 y los PERTE que está poniendo en marcha el Gobierno español pueden ser una gran oportunidad para potenciar el desarrollo y liderazgo de estas tecnologías.

PERTE Nueva Economía de la Lengua Española e Inteligencia Artificial:

Este PERTE se plantea como una oportunidad para aprovechar el potencial del español y de las lenguas cooficiales como factor de crecimiento económico y competitividad internacional en áreas como la inteligencia artificial, la traducción, el aprendizaje, la divulgación cultural, la producción audiovisual, la investigación y la ciencia. Para ello, cuenta con un presupuesto de 1.100 millones de millones de euros de inversión pública, con el objetivo de movilizar otros 1.000 millones de inversión privada.

Se impulsarán 14 proyectos tractores mediante la acción coordinada de las administraciones públicas, las universidades, los centros de investigación, las empresas e industrias, basados en cinco ejes estratégicos:

1. Conocimiento en español y lenguas cooficiales: Incluye el apoyo al **Proyecto Lengua Española e Inteligencia Artificial (LEIA)** para crear un banco de datos de la lengua, enriquecido con las aportaciones de las comunidades autónomas; y la creación del **Observatorio del español**, que se dedicará a estudiar la situación del español en el mundo, incluyendo la perspectiva de la nueva economía digital y evaluará el seguimiento de los proyectos tractores incluidos en este PERTE.



RETOS Y OPORTUNIDADES

2. *Inteligencia Artificial en español*: Se invertirán más de 330 millones de euros para poner en marcha proyectos como la creación de una **Red de Excelencia en Inteligencia Artificial** y los Programas de Lenguas accesibles, con el objetivo de proporcionar herramientas digitales para personas con discapacidad y diversidad funcional y actuaciones del Plan Nacional de Tecnología del Lenguaje y su industria.

3. Ciencia en español.

4. Aprendizaje del español.

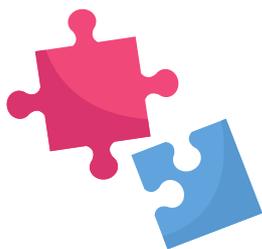
5. Industrias culturales.

RETEC: Redes Territoriales de Especialización Tecnológica

Es una de las nuevas iniciativas de la actualización de la Agenda Digital España 2025, es decir, Agenda Digital España 2026.

El objetivo de este programa mejorar la co-gobernanza del Plan de Recuperación a través de la articulación de proyectos regionales orientados a la transformación y especialización digital. El apoyo a estos proyectos fomentará el intercambio de conocimiento y las oportunidades de cada comunidad autónoma a través de redes de impacto nacional que permitan maximizar el equilibrio territorial y la cohesión social.

Las Comunidades Autónomas podrá presentar proyectos colaborativos en las diferentes áreas temáticas definidas dentro de la iniciativa RETECH (**Inteligencia Artificial** y otras tecnologías habilitadoras aplicadas a las industrias, Gemelos Digitales, GreenTech, Ciberseguridad, Emprendimiento Digital, Tecnología con impacto social, RuralTech, FashionTech y Salud Digital), contribuyendo con ello además al cumplimiento de hitos y objetivos del Plan de Recuperación, Transformación y Resiliencia.

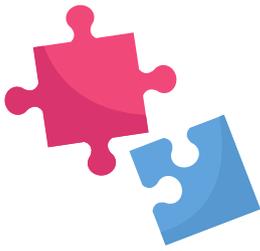


RETOS Y OPORTUNIDADES

EJE 4 de la Agenda España Digital 2026: Economía del Dato e inteligencia artificial:

Este eje cuenta con las medidas 13, 14, 15, 16 y 17 que entre otras acciones incluyen las siguientes

- Agencia Nacional de Supervisión de la Inteligencia Artificial.
- Investigación y sensibilización en IA: Plan de Protección para Colectivos Vulnerables en inteligencia artificial y Plan de sensibilización y confianza hacia la inteligencia artificial.
- Observatorio del impacto social y ético de los algoritmos (OBISAL).
- Sandboxes regulatorios para una aplicación ética de la IA.
- Sello de IA confiable.
- Convocatoria de subvenciones para integración de la inteligencia artificial en cadenas de valor: se financiarán proyectos de desarrollo experimental cuyo estado de madurez tecnológica sea avanzado.
- Programa Nacional de Algoritmos Verdes (PNAV): se creará una guía de buenas prácticas, un catálogo de algoritmos eficientes y otro de algoritmos para abordar problemas ambientales.



RETOS Y OPORTUNIDADES

España presenta el primer piloto del sandbox de regulación de Inteligencia Artificial en la UE:

El pasado 27 de junio el Gobierno de España, presentaba en Bruselas el proyecto piloto para poner en marcha el primer sandbox (banco de pruebas) regulatorio de la Unión Europea sobre Inteligencia Artificial.

El sandbox pretende acercar a las autoridades competentes a las empresas desarrolladoras de Inteligencia Artificial para definir de forma conjunta buenas prácticas que sirvan de base para la implementación de la futura regulación europea de Inteligencia Artificial (Artificial Intelligence Act), y garantizar que se puede aplicar en el tiempo previsto, en dos años.

Mientras todo el ecosistema se prepara para su aplicación, se espera que este piloto liderado por España genere directrices de buenas prácticas y guías que preparen y sensibilicen a las empresas, especialmente a pymes y startups, para facilitar su puesta en marcha. Este piloto propuesto por el Gobierno español estudiará la operatividad de los requisitos del futuro reglamento, así como las evaluaciones de conformidad o las actividades posteriores a la comercialización.

CASOS DE USO

A continuación os mostramos algunos de los ejemplos de casos de uso que DISRUPTIVE ha ido recopilando durante 2022 y que puedes consultar con más detalle pinchando [aquí](#)



HONEY.AI -Sonicat System

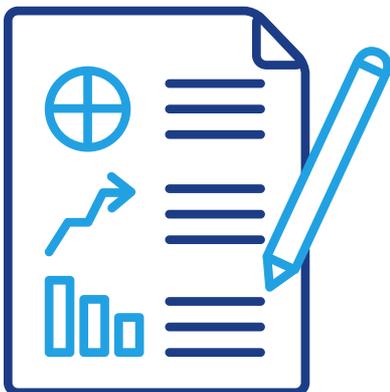
Honey.AI permite realizar distintos análisis de calidad en miel in-situ (análisis polínico, recuento de levaduras y grado de cristalización), en menos de 1 hora.

El sistema Honey.AI funciona con un microscopio óptico automatizado de diseño propio que escanea automáticamente toda la superficie del portaobjetos conteniendo el sedimento de miel.

El equipo Honey.AI alcanza una precisión de movimiento mecánico en sus ejes XYZ de hasta 1 micra.

Honey.AI funciona mediante algoritmos de procesamiento de imagen y varias redes neuronales de aprendizaje profundo.

La inteligencia artificial de Honey.AI ha sido ampliamente entrenada desde 2019 con cientos de muestras de miel de diferentes orígenes geográficos y botánicos para poder detectar, clasificar y contar los granos de polen contenidos en la muestra.



Previsión de stock y tendencias - Cloud Levante

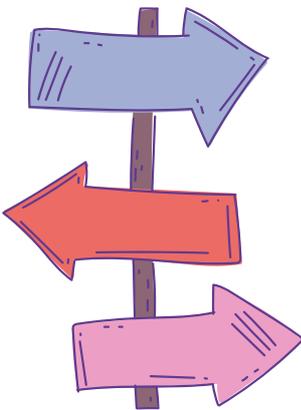
Extracción y análisis de datos en tiempo real o por batch para extraer información relevante de la empresa. Entrenamos modelos probabilísticos o estadísticos capaces de detectar patrones y definir previsiones de demanda, stock, tendencias, etc. Además, con el Machine Learning podemos automatizar procesos de negocio, permitiendo a las organizaciones reducir el tiempo dedicado a las tareas diarias y dedicarlo a las áreas importantes de su empresa.

CASOS DE USO



BD EN DEPORTE - Universidad de Valladolid

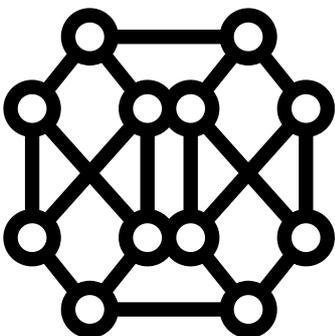
Uso de Big Data para analizar el rendimiento en el deporte. En el ámbito deportivo se generan muchos datos de los de los deportistas (ya sean profesionales, futbolistas, por ejemplo, o recreativos). Son datos referidos a su rendimiento físico táctico o referidos a sus hábitos de consumo deportivo, recorridos, datos ergonómicos y fisiológicos. Estos datos, están en manos de empresas que los comercializan. Algunos de esos datos pueden estar relacionados con la salud y la gestión de la salud de entidades públicas, por lo que tienen un valor estratégico.



GESTIÓN DE TRÁFICO - ACISA (Grupo Aldesa)

Integran algoritmos basados en IA y Big Data en sus plataformas de gestión de tráfico.

En todos sus casos de uso orientados a la administración y gestión de infraestructura de movilidad, la ciberseguridad es un tema clave. Sus soluciones se realizan bajo sistema de gestión SGSI certificado ISO 27001.



REDES NEURONALES - Neurogenesis IA Technologies S.L.

Creación de redes neuronales basadas en aprendizaje profundo con interpretación. Modelos de deep learning que aprende tanto del resultado como de los diferentes procesos que le han llevado a tomar dicho resultado.

CASOS DE USO

LOGMEAL - Algecko Technologies S.L.

¿Sabes lo que comes realmente? Con "LogMeal, el shazam de la comida" con sólo una foto puedes tener toda la información nutricional en tu móvil. LogMeal es un completo sistema de monitorización digital de ingestas para nutricionistas deportivos y endocrinos.



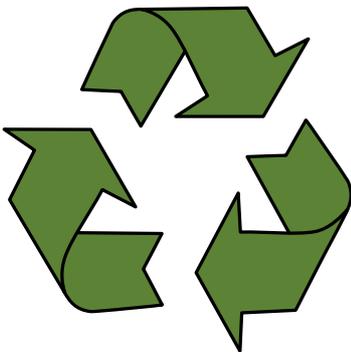
EVOCONSTRUCTOR® - Evolution Construction System SL (EVOCONS)

Se trata de un gantry robot para la construcción robotizada a gran escala e impresión 3D en hormigón que automatiza el proceso constructivo de edificios y otras obras de construcción de gran envergadura: ya que no solo ejecuta paredes y muros a través de la impresión 3D en hormigón, que es lo único que hace la competencia, sino que también automatiza el 60% del proceso constructivo como: tareas de alicatado, enfoscado, bases de pavimento y aplicación de aislante térmicos y acústicos.



SIGHO BUFFET- Visionary Hospitality

Este aplicativo usando visión e inteligencia artificial permite cuantificar los residuos de comida del buffet diferenciándolos por tipología, y vinculados al perfil del tipo de cliente alojado. Con esta iniciativa se espera, por un lado, reducir los desperdicios de alimentos y por otro, optimizar los menús y hacerlos más adecuados a los gustos y exigencias de cada tipo de huésped. "Si conozco lo que tiro sé lo que no tengo que comprar".



ENLACES DE INTERÉS

Aquí os remitimos a enlaces de noticias, webs, documentos normativos o informes de interés, así como a bibliografía sobre esta tecnología.

Tecnologías emergentes en 2022 según Gartner

Tendencias big data 2022

El gobierno apoya el nacimiento de la Asociación de Gaia-X

PERTE Nueva Economía de la Lengua

Nadia Calviño: "Con el PERTE Nueva Economía de la Lengua queremos que España sea el nodo central de la inteligencia artificial en español"

Algoritmos verdes

Uso de tecnologías digitales por empresas en España

Informe tendencias 2022 de Globant

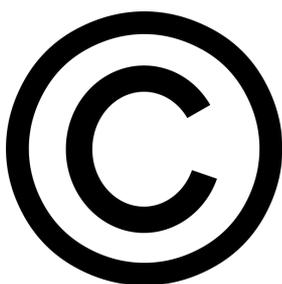
España Digital 2026

Eje 4. Economía del Dato e Inteligencia Artificial

El Gobierno de España presenta, en colaboración con la Comisión Europea, el primer piloto del sandbox de regulación de Inteligencia Artificial en la UE

El PERTE de la Nueva Economía de la Lengua reúne a más de 100 empresas punteras de Procesamiento del Lenguaje Natural como puesta en marcha del Foro de la Alianza

Nace la Fundación PyTorch, el mayor proyecto abierto para acelerar la investigación en inteligencia artificial



Informe realizado por la **Asociación de Parques Científicos y Tecnológicos de España (APTE)**, entidad que gestiona la secretaría técnica de la **Plataforma Tecnológica Española de Tecnologías Disruptivas (DISRUPTIVE)** con la colaboración de su **grupo de trabajo de inteligencia artificial y big data** durante el último trimestre de 2022



Con financiación de:



Convocatoria 2020 Plataformas Tecnológicas y de Innovación (PTR2020-001159)

Secretaría técnica a cargo de:

