Inteligencia artificial como servicio: Potenciando la innovación y eficiencia en la industria y las metodologías ágiles

Artificial intelligence as a service: Boosting innovation and efficiency in industry and agile methodologies

Diego Monsalves[1]* https://orcid.org/0000-0001-6780-3754
Pablo Olivares[1] https://orcid.org/0009-0003-5200-3289
Fabián Riquelme[1] https://orcid.org/0000-0003-4491-0148
Héctor Cornide-Reyes[2] https://orcid.org/0000-0002-9130-184X

Recibido 12 de octubre de 2024, aceptado 25 de noviembre de 2024
Received: October 12, 2024 Accepted: November 25, 2024

RESUMEN

La transformación digital ha revolucionado la manera en que operan las industrias, mejorando la eficiencia y calidad de productos y servicios mediante la integración de tecnologías avanzadas. La recolección y análisis de grandes volúmenes de datos, junto con la Inteligencia Artificial (IA), son esenciales para la toma de decisiones informadas y la optimización de operaciones. La Inteligencia Artificial como Servicio (AIaaS) permite el acceso a tecnologías avanzadas como cajas negras, facilitando a las empresas implementar soluciones de IA sin grandes inversiones en conocimientos técnicos o infraestructura. Este artículo explora el impacto de AIaaS en la industria, destacando su papel en la transición de la Industria 4.0 a la Industria 5.0, que se centra en la colaboración humano-máquina y la sostenibilidad. Además, se examina la evolución de las metodologías ágiles y su integración con tecnologías emergentes. Se presenta una revisión de las actuales y potenciales aplicaciones soportadas por AIaaS, junto con un análisis de las oportunidades, desafíos y limitaciones de estas tecnologías.

Palabras clave: Inteligencia artificial como servicio, industria 5.0, industria 4.0, metodologías ágiles, innovación tecnológica.

ABSTRACT

Digital transformation has revolutionized how industries operate, improving the efficiency and quality of products and services through the integration of advanced technologies. The collection and analysis of large data volumes and Artificial Intelligence (AI) are essential for informed decision-making and the operations optimization. Artificial Intelligence as a Service (AIaaS) enables access to advanced technologies as black boxes, making it easier for companies to implement AI solutions without significant technical expertise or infrastructure investments. This article explores AIaaS impact on the industry and highlighting its role in the transition from Industry 4.0 to Industry 5.0, which focuses on human-machine collaboration and sustainability. In addition, the evolution of agile methodologies and their integration with emerging technologies is examined. A review of current and potential applications supported by AIaaS is presented, along with an analysis of these technologies' opportunities, challenges, and limitations.

Keywords: *Artificial Intelligence as a service, industry 5.0, industry 4.0, agile methodologies, technological innovation.*

[1] Universidad de Valparaíso. Escuela de Ingeniería Informática. Valparaíso, Chile.
Email: diego.monsalves@uv.cl; pablo.olivareszu@uv.cl; fabian.riquelme@uv.cl
[2] Universidad de Atacama. Departamento de Ingeniería Informática y Ciencias de la Computación. Copiapó, Chile.
Email: hector.cornide@uda.cl *
Autor de correspondencia: diego.monsalves@uv.cl

INTRODUCCIÓN

La era digital ha transformado profundamente la manera en que operan las grandes industrias, impulsando una constante evolución en sus procesos y estrategias empresariales. La transformación digital y los avances tecnológicos han introducido nuevas herramientas y metodologías que optimizan la eficiencia, reducen costos y mejoran la calidad de los productos y servicios [1]. La recolección y análisis de grandes volúmenes de datos es fundamental en el contexto industrial. La capacidad de procesar y analizar estos datos permite a las empresas extraer información valiosa para mejorar la toma de decisiones, optimizar operaciones y predecir tendencias futuras [2]. Esto es crucial para identificar ineficiencias, mejorar la calidad de los productos y servicios, y personalizar la experiencia del cliente, promoviendo la innovación y adaptabilidad en un entorno de rápida evolución tecnológica [3].

En este contexto, la Inteligencia Artificial (IA) emerge como un pilar fundamental de la innovación, ofreciendo capacidades avanzadas de análisis y toma de decisiones que revolucionan la industria [4]. Se espera que el mercado global de la transformación digital crezca a una tasa compuesta anual (CAGR) del 27,6% de 2024 a 2030 [5], donde la IA está desempeñando un papel crucial en este crecimiento, con aplicaciones que abarcan desde la automatización de procesos en la manufactura hasta la mejora en la gestión de cadenas de suministro [6], la personalización de la experiencia del cliente en sectores como la salud [7] y las finanzas [8]. En este mismo ámbito, las grandes empresas que encabezan esta transformación tecnológica también ofrecen soluciones disponibles a los usuarios finales para integrar IA en sus servicios. Por ejemplo, Google ofrece soluciones avanzadas de IA que permiten a las pymes y microempresas integrar tecnologías en sus operaciones, proporcionando herramientas como la personalización de ofertas y recomendaciones de productos mediante IA, la optimización de la cadena de suministro y la gestión de inventarios en tiempo real.

La Inteligencia Artificial como Servicio (AIaaS) se refiere a la provisión de capacidades de inteligencia artificial a través de plataformas en la nube, lo que permite a las empresas acceder y utilizar tecnologías avanzadas de IA sin la necesidad de desarrollar y mantener una infraestructura interna costosa [9], [10]. AIaaS funciona mediante la oferta de herramientas y servicios, como algoritmos de aprendizaje automático, procesamiento del lenguaje natural, modelos de lenguaje grande (en inglés, Large Language Models o LLM) y análisis predictivo, que los usuarios pueden integrar en sus propias aplicaciones y procesos empresariales. Este modelo de servicio democratiza el acceso a la IA, permitiendo que tanto grandes empresas como pymes puedan aprovechar estas tecnologías.

Por otra parte, las metodologías ágiles, originadas en la informática, han revolucionado la gestión de proyectos en general, ofreciendo desde su teoría y práctica un enfoque flexible y adaptativo que contrasta con los métodos tradicionales [11]. Frameworks como Scrum, Kanban y Extreme Programming priorizan la colaboración, la adaptabilidad y la entrega continua de valor, mejorando la eficiencia y la satisfacción del cliente. Su adopción se ha extendido más allá del desarrollo de software, encontrando aplicaciones en sectores como marketing [12], casos clínicos [13] y desarrollo de leyes [14]. Los equipos de trabajo que adoptan metodologías ágiles se caracterizan por su autonomía, multifuncionalidad y comunicación fluida. Prácticas como las reuniones diarias y las revisiones periódicas promueven la transparencia y el alineamiento de objetivos. En el contexto actual, las metodologías ágiles están evolucionando para integrar tecnologías emergentes, no solo en el desarrollo de soluciones de IA, sino también en la optimización de los propios procesos ágiles [15].

La evolución industrial ha progresado desde la Industria 4.0 hacia la Industria 5.0, marcando un cambio significativo en los paradigmas de producción. La Industria 4.0 se caracterizó por la integración de tecnologías digitales avanzadas como sistemas ciberfísicos, Internet de las Cosas (IoT), big data e Inteligencia Artificial (IA) en los procesos de fabricación, mejorando la automatización y eficiencia [16]. Sin embargo, este modelo fue criticado por su falta de enfoque en la sostenibilidad y los aspectos humanos [17]. La Industria 5.0 ha surgido recientemente como respuesta, reintroduciendo el elemento humano y enfatizando la colaboración humano-robot [16], [18]. Este nuevo paradigma se basa en tres pilares: centralidad humana, resiliencia y sostenibilidad, utilizando tecnologías como robots colaborativos y gemelos digitales para crear

soluciones más ergonómicas y eficientes [19]. Aspira a ser un enfoque más equilibrado y ético, promoviendo la justicia social y la gestión ambiental responsable [17], [20]. En este contexto, la AIaaS y los LLM juegan un papel crucial, mejorando la colaboración humano-IA y potenciando sistemas multi-agente más robustos [21], [22].

Este artículo tiene como objetivo describir el potencial de AIaaS para mejorar las aplicaciones e investigación actuales en la industria, considerando aspectos de Industria 4.0 y 5.0, así como de metodologías ágiles. Se presenta una lista de aspectos de investigación y desarrollo en los que AIaaS puede contribuir dependiendo del área. Además, se presenta una tabla descriptiva con un listado de las empresas, tecnologías y herramientas más relevantes hasta la fecha y disponibles para usuarios finales.

ESTADO DEL ARTE

En esta sección se revisan conceptos, aplicaciones e investigaciones actuales presentes en la literatura que engloban a las metodologías ágiles, industria 4.0-5.0, y AIaaS.

Metodologías ágiles y la industria 4.0-5.0

La gestión de proyectos ha experimentado una significativa transformación en respuesta a un entorno empresarial caracterizado por su dinamismo y avance tecnológico [23]. Las metodologías tradicionales, conocidas por su estructura rígida y secuencial, han cedido terreno a enfoques ágiles que facilitan una adaptación más eficiente a las fluctuantes demandas del mercado. Los equipos de trabajo, ahora más autónomos y multifuncionales, se benefician de frameworks ágiles que promueven la flexibilidad, la comunicación constante y la entrega continua de valor [24]. Estos equipos ágiles pueden responder rápidamente a cambios en los requisitos del proyecto, ajustando planes y prioridades según sea necesario, mejorando la colaboración y facilitando que todos los miembros estén alineados con los objetivos del proyecto. Estudios recientes sugieren que la implementación de metodologías ágiles ha tenido un impacto positivo en la optimización de la gestión de proyectos, fomentando la innovación y la eficiencia en una amplia gama de sectores industriales [24], [25].

Es en estos contextos industriales que durante los años 2010 se desarrolló la denominada Industria 4.0, descrita en la Sección I, que se basa en la revolución tecnológica y los datos como fuente del desarrollo industrial. Sin embargo, la adopción de estas tecnologías ha presentado múltiples desafíos para las empresas, que han tenido que enfrentarse a la complejidad y costos asociados con la integración de sistemas avanzados [26], para reestructurar sus procesos y adaptarse a nuevas formas de operar que demandan habilidades y conocimientos especializados. La implementación efectiva de la Industria 4.0 requiere no solo una inversión en infraestructura tecnológica, sino también un cambio cultural dentro de las empresas, promoviendo la innovación y la flexibilidad [26], [27]. A pesar de estos desafíos, las empresas que han logrado integrar estas tecnologías han reportado mejoras en la eficiencia operativa, la toma de decisiones basada en datos y la capacidad de respuesta a las demandas del mercado [26].

En el último tiempo, se ha comenzado a hablar sobre la Industria 5.0, que, según algunos autores, constituye una nueva etapa en que las grandes empresas han completado el proceso de adaptación con la Industria 4.0 y ahora están desarrollando nuevos modelos de negocio que abarcan tecnologías relacionadas con enfoques innovadores [26], que dependen de la colaboración entre humanos y máquinas inteligentes. Su foco está en integrar más estrechamente la IA y la robótica avanzada con las capacidades humanas, priorizando la centralidad humana, la resiliencia y la sostenibilidad [17].

Aplicaciones en la industria con IA

Los avances recientes en IA constituyen una tecnología emergente con el potencial de transformar diversas áreas de la sociedad y la industria. Su capacidad para procesar grandes volúmenes de datos y aprender de ellos permite optimizar procesos, mejorar la toma de decisiones y abrir nuevas oportunidades para la innovación. En la era de la Industria 4.0 y el surgimiento de la Industria 5.0, la IA se ha convertido en un componente esencial para mantener la competitividad y eficiencia en un mercado globalizado [28].

En el sector financiero, la IA ha mostrado buenos rendimientos en la previsión financiera, la protección financiera y el análisis de riesgos, permitiendo una detección más precisa de fraudes y una gestión automatizada de carteras de inversión [29]. En

la manufactura, se ha utilizado para optimizar la cadena de suministro, llevar a cabo controles de calidad automatizados, realizar mantenimiento predictivo, logrando desarrollar tecnología para predecir fallos y planificar intervenciones de manera más efectiva [30], mejorando así, según algunos autores, los estándares de calidad y reduciendo costos para las empresas [31]. En el ámbito de la salud, la IA ha revolucionado el diagnóstico y tratamiento de enfermedades mediante técnicas de aprendizaje profundo aplicadas a la imagen médica y la personalización de tratamientos basados en el análisis de datos del paciente [32], [33]. En la agricultura, la IA se utiliza para el monitoreo de cultivos, la predicción de rendimientos y la gestión de plagas, ayudando a los agricultores a tomar decisiones informadas sobre el manejo de sus recursos [34]. En la educación superior, se ha aplicado para el análisis predictivo del rendimiento académico, la deserción y el comportamiento estudiantil [35], además de complementar la enseñanza con investigación en la educación superior [36].

Según algunos autores, la integración de IA en la Industria 5.0 potencia la colaboración entre humanos y máquinas, permitiendo la capacitación de *cobots* para adaptarse a la lógica humana [28], redefiniendo así los procesos industriales hacia una mayor eficiencia y adaptabilidad [28]. La adopción de estas tecnologías para su uso en aplicaciones es un aspecto a tomar en cuenta tanto en entornos industriales como de futuros profesionales. Esto subraya la importancia de desarrollar competencias específicas que aumenten la empleabilidad de los profesionales en un mercado laboral cada vez más digitalizado. Así, por ejemplo, un estudio con estudiantes de arquitectura de interiores reveló una limitada conciencia sobre las capacidades tecnológicas de las IA avanzadas, como ChatGPT, Stable Diffusion y Midjourney [37]. No obstante, aunque los estudiantes expresan inquietudes sobre el impacto de la IA en el empleo, también muestran disposición para integrar en sus carreras para mejorar la productividad y la creatividad.

Desarrollos actuales en IA y AIaaS

Los desarrollos actuales en IA y AIaaS están revolucionando múltiples sectores de la industria con innovaciones de gran impacto. Un ejemplo relevante es el de la plataforma NVIDIA Earth-2 [38], que ha aumentado la eficiencia en las predicciones climáticas y meteorológicas con simulación interactiva de alta resolución potenciada por IA. Esta plataforma abierta incluye simulaciones físicas con modelos numéricos como ICON e IFS, modelos de aprendizaje de máquina o machine learning cómo FourCastNet y GraphCast, así como un modelo de predicción meteorológica basado en aprendizaje profundo, llamado Deep Learning Weather Prediction (DLWP) desarrollado con su framework NVIDIA Modulus. Además, en el ámbito de los servicios de IA, Earth-2 ofrece acceso a modelos basados en IA generativa como los modelos de difusión CorrDiff, capaces de generar grandes conjuntos de predicciones útiles de alta resolución mediante.

Por otra parte, la ampliación de investigación y desarrollo (I+D) en productos y servicios con IA está siendo impulsada por plataformas como SILO AI, que integran la IA en tecnologías existentes para evolucionar desde experimentos fragmentados hacia desarrollos que generan un valor significativo. Estas aplicaciones mejoran la eficiencia en sectores como la fabricación, donde garantizan mejor calidad y confiabilidad mediante visión artificial y machine learning; en vehículos y maquinaria pesada, proporcionando información útil para el mantenimiento preventivo y eficiente; y en la medicina, la educación y la prensa, transformando cada campo con soluciones innovadoras y personalizadas. En el contexto de la Industria 5.0, las plataformas como SILO AI permiten a las empresas desarrollar sistemas que combinan la capacidad analítica de las máquinas con la creatividad y el pensamiento crítico en las personas.

La IA está desempeñando también un papel transformador en la medicina, en particular para buscar mejoras en el diagnóstico y el tratamiento de diversas enfermedades. Estudios recientes demuestran que la IA puede analizar grandes volúmenes de datos médicos con alta precisión, permitiendo la detección temprana de enfermedades y la personalización de tratamientos [39]. Además, la integración de IA en los sistemas de salud está optimizando la gestión hospitalaria y reduciendo costos operativos [40].

En educación, la IA está proporcionando soluciones innovadoras que transforman la experiencia de aprendizaje. Los sistemas de tutoría inteligentes, por ejemplo, utilizan algoritmos de aprendizaje

automático para personalizar el contenido educativo y adaptarlo a las necesidades individuales de los estudiantes [41]. Además, la IA está ayudando a los educadores a identificar áreas donde los estudiantes pueden estar teniendo dificultades, permitiendo intervenciones más rápidas y efectivas [41]. Una aplicación disponible para los estudiantes con IA es Celebras Voice, un agente inteligente capaz de establecer una conversación en inglés, para poder ejercitar sin la necesidad de un hablante nativo, esta aplicación diseñada con LiveKit Agents. En investigaciones, aplicaciones con IA están creando un entorno educativo más dinámico y accesible, beneficiando tanto a estudiantes como a profesores, según indican en [42]. Por otro lado, los LLM con capacidad para manejar contextos largos han sido especialmente estudiados en investigaciones recientes. Han surgido varios puntos de referencia, como LongIns, LongICLBench y Babilong, para evaluar su rendimiento [43]-[45], lo que demuestra la necesidad de mejorar las capacidades para comprender y procesar grandes cantidades de información [44], [45]. Además, la introducción de frameworks como LC-Boost enfatiza el potencial de los LLM para abordar tareas de contexto largo de manera eficiente a través del acceso y la utilización adaptativos de contexto [46]. Además, innovaciones como la incrustación extensible ofrecen métodos rentables y flexibles para ampliar el contexto de los LLM de manera efectiva, mejorando su performance en el manejo de aplicaciones críticas. OpenAI ha marcado una nueva hoja de ruta para los LLM como se conocía hasta este momento. El proyecto Strawberry salió a la luz como OpenAI o1, un serie de modelos capaces de "razonar" lógicamente antes de entregar una respuesta. Según sus desarrolladores, entrenaron modelos para que dediquen más tiempo a pensar en los problemas antes de responder, de forma similar a como lo haría una persona [47]. A través del entrenamiento, aprenden a refinar su proceso de pensamiento, a probar diferentes estrategias y a reconocer sus errores. Según los primeros datos de referencias [47], [48], se comporta de manera similar a estudiantes de doctorado en tareas de referencia desafiantes en física, química y biología. En cuanto a habilidades de codificación, los modelos se evaluaron en concursos humanos y alcanzaron el percentil 89 en las competencias de Codeforces. Según pruebas de rendimiento realizadas en septiembre de 2024, el modelo o1 preliminar entregado lidera en todos los aspectos, especialmente en matemáticas, ejercicios difíciles y programación. Estos datos marcan la tendencia de que la investigación y desarrollo de modelos LLM no se estarían estancado.

POTENCIALES APLICACIONES DE AIAAS EN LA INDUSTRIA

Esta sección explora diversas aplicaciones de AIaaS en la industria y en las metodologías ágiles, destacando cómo estas tecnologías pueden contribuir a la evolución hacia la Industria 5.0.

Los siguientes son ocho puntos donde AIaaS puede apoyar y mejorar soluciones actuales en la industria:

- **Mantenimiento autónomo de infraestructuras**: Sistemas de monitoreo continúo basados en aprendizaje automático para detección de anomalías y predicción de fallas en tiempo real. Estos sistemas podrían activar protocolos de reparación automática o alertar a operarios humanos con instrucciones precisas, minimizando el tiempo de inactividad y optimizando la eficiencia operativa.
- **Sistemas de energía inteligente:** Sistemas basados en aprendizaje profundo para optimización de la gestión energética en tiempo real, predicción de patrones de demanda a corto y largo plazo, y optimización de la distribución de carga entre diferentes fuentes de energía.
- **Software generado por representaciones informales de requisitos**: Sistemas capaces de analizar y comprender automáticamente representaciones informales de requisitos, tales como historias de usuario, para generar prototipos de código y sugerir arquitecturas de software optimizadas. Esto aceleraría el ciclo de desarrollo y mejoraría la alineación entre las necesidades del usuario y el producto final.
- **Asistentes virtuales en entorno industrial:** Sistemas de asistentes virtuales potenciados por AIaaS para proporcionar soporte en tiempo real a los trabajadores, ofreciendo información contextual, guiando en procedimientos complejos y facilitando la toma de decisiones.
- **Generación de API automática**: Sistemas de interfaces que simplifiquen la integración de servicios y sistemas mediante la generación automática de API RESTful o GraphQ.
- **Fabricación adaptativa en tiempo real**: Sistemas de fabricación que se adapten

instantáneamente a cambios en la demanda, calidad de los materiales o condiciones de producción. Esto permitiría una personalización masiva eficiente, optimizando el uso de recursos y minimizando los desperdicios.

- **Gemelos digitales colaborativos:** Sistemas de gemelos digitales con servicios AIaaS para crear representaciones virtuales más precisas y dinámicas de los sistemas físicos. Estos gemelos podrían colaborar entre sí y con sistemas humanos, facilitando la simulación de escenarios complejos, la optimización de procesos y la toma de decisiones predictivas a escala industrial.
- **Realidad aumentada para supervisión y mantenimiento**: Sistemas de realidad aumentada que proporcionen información en tiempo real sobre el estado de los equipos, guíen en procedimientos de mantenimiento complejos y ofrezcan asistencia remota experta. Esto mejoraría la eficiencia del mantenimiento, reduciría los errores y facilitaría la transferencia de conocimientos entre expertos y trabajadores en formación.

En la actualidad, la literatura ya muestra evidencias de que la AIaaS tiene el potencial de transformar la manera en que los equipos ágiles operan, mejorar su desempeño y la toma de decisiones [49]. AIaaS puede apoyar en automatizar tareas repetitivas y administrativas, permitiendo a los miembros del equipo centrarse en actividades de mayor valor para los clientes. También, ayudan a analizar en tiempo real los grandes volúmenes de datos que se generan durante el proceso de desarrollo del software. Lo anterior proporciona valiosa información sobre el desempeño del equipo lo que permite tomar decisiones de forma oportuna. En la misma línea, el uso de la IA generativa basada en LLM junto con la ingeniería de software ágil ayudan a interpretar y comprender mejor el lenguaje natural, lo que facilita la captura y análisis de requisitos escritos en lenguaje común, sin necesidad de traducción técnica. Esto permite una mejor comunicación entre los stakeholders y los ingenieros de requisitos [50]. Sin lugar a duda, todos estos avances van a generar desafíos importantes a la gestión de proyectos ágiles debido a que deberá facilitarse la planificación y priorización de tareas, optimizando así la eficiencia del equipo. Un aspecto muy importante por considerar es que AIaaS también ayuda a mejorar la comunicación y la colaboración entre los miembros de un equipo ágil. Herramientas como los chatbots y asistentes virtuales proporcionan respuestas rápidas, programan reuniones y facilitan la documentación compartida.

Finalmente, podemos señalar que la AIaaS puede contribuir significativamente a los equipos que trabajan con métodos ágiles de las siguientes formas:

- Automatización de tareas repetitivas.
- Gestión del Backlog.
- Análisis de los datos generados por el proceso de desarrollo en tiempo real.
- Optimización de la planificación del proyecto.
- Facilitación de la comunicación y la colaboración entre todos los stakeholders.
- Soporte en la toma de decisiones.
- Retrospectivas más eficientes que potencian la mejora continua.

SOLUCIONES AIAAS ACTUALES

Las AIaaS actuales se encuentran en constante actualización, posicionándose a la vanguardia del desarrollo tecnológico y económico mundial. Estas tecnologías, en rápida transformación, permiten a empresas de todos los tamaños integrar IA en sus operaciones. Las grandes corporaciones tecnológicas están liderando su uso, colaborando en proyectos de investigación y estableciendo estándares tecnológicos, mientras compiten por ofrecer soluciones más avanzadas y atractivas al mercado (Tabla 1). Esta dualidad de colaboración y competencia ha impulsado numerosos procesos de innovación en breves lapsos de tiempo.

La Tabla 2 muestra las empresas y servicios más relevantes en AIaaS, junto con las descripciones de cada tecnología y modalidades de análisis de los modelos IA. Estas tablas destacan las diferentes modalidades disponibles para satisfacer diversas necesidades.

Con respecto a la modalidad de texto, Alibaba ofrece modelos como Qwen para creación de textos y traducción (no es multi-lenguaje), mientras que Amazon proporciona servicios de procesamiento de lenguaje natural (Natural Language Processing o NLP) con Amazon Comprehend para propósito general, al igual que Mistral. En cambio, Nvidia ofrece soluciones enfocadas al descubrimiento de

Tabla 1. Novedades en investigación de IA en la industria.

Organización	Investigación reciente	Ejemplos / Detalles
Meta	Generación de imágenes y videos	Uso de IA generativa para crear imágenes y videos para publicidad y contenido digital. Disponible aquí
Google	Modelos de lenguaje y robótica	Fusión de modelos de lenguaje con robótica para desarrollar sistemas más flexibles como Gemini y RT-2, capaces de realizar tareas complejas y preguntar para mejorar la interacción con el mundo real. Disponible aquí
OpenAI	Transparencia y seguridad en IA	Desarrollo de técnicas para aumentar la transparencia de sus sistemas, como parte de sus esfuerzos para abordar los riesgos asociados con la IA generativa. Disponible aquí
IBM	Optimización de modelos y despliegue flexible	Desarrollo de técnicas como Low Rank Adaptation (LoRA) y la cuantización para optimizar el rendimiento de modelos compactos, haciendo la IA más accesible a startups y desarrollos locales. Disponible aquí
Synthesia	Deepfakes para marketing y entrenamiento	Herramientas para generar avatares de deepfake a partir de una actuación única, utilizadas por muchas empresas de la lista Fortune 100. Disponible aquí
McKinsey	Valor empresarial de la IA generativa	Uso de IA analítica y generativa para obtener beneficios significativos en la cadena de suministro, gestión de inventarios y operaciones de servicio. Disponible aquí
Stanford University	Evaluación y estándares de IA responsable	En su informe AI Index, destacan la falta de estandarización en la evaluación de la IA responsable, lo que complica la comparación de riesgos y limitaciones de los modelos principales. Disponible aquí
Runway	Festival de cine con IA	Organización de un festival de cine que premia películas experimentales creadas con IA, reflejando la creciente integración de la IA en la industria del entretenimiento. Disponible aquí

fármacos con su modelo BioNemo. Servicios de audio y texto, como Azure AI Speech de Microsoft y Watson Speech to Text de IBM, permiten la conversión y análisis preciso del habla. Nvidia, con su SDK Maxine, y Amazon Rekognition, abordan las necesidades de video y audio, ofreciendo herramientas avanzadas para la creación de contenido y reconocimiento de imágenes.

En el ámbito multimodal, varias empresas ofrecen modelos avanzados capaces de manejar texto, imágenes, videos y código. Anthropic, con su modelo Claude, y Google, con su serie de modelos Gemini, permiten a las empresas desarrollar soluciones que integran múltiples tipos de datos. OpenAI ofrece modelos como GPT, que combinan texto, imágenes y videos para crear contenido generativo; servicios que Microsoft ocupa y disponibiliza a través de Azure OpenAI Service en un entorno seguro y escalable. Meta, con LLaMA, su LLM avanzado, ofrece capacidades multimodales y de código abierto. Huawei, con EIHealth y HiLens, y Nvidia, con Clara, también destacan en este campo, proporcionando herramientas para la genómica, el monitoreo de pacientes, y la creación de contenido visual.

La Tabla 3 muestra suites de desarrollo en AIaaS ofrecidas por empresas tecnológicas líderes. Estas suites proporcionan plataformas integrales que permiten a las empresas y desarrolladores acceder a avanzadas capacidades de IA y herramientas para el aprendizaje automático y profundo, integración con múltiples frameworks y lenguajes de programación, y facilitar la creación y despliegue de modelos de IA generativa y predictiva. Los beneficios clave de estas suites incluyen la automatización del flujo de trabajo, la personalización y escalabilidad de aplicaciones, y la optimización de procesos mediante el análisis de grandes volúmenes de datos. Estas plataformas democratizan el acceso a tecnologías avanzadas de IA y apoyan la transición hacia la Industria 5.0, donde la colaboración humano-máquina y la sostenibilidad son esenciales. Por ejemplo, Google con Vertex AI, proporciona una plataforma unificada para el desarrollo de aplicaciones utilizando modelos multimodales y herramientas avanzadas de IA, propias, de otras empresas o de la comunidad. Al igual que Google, Microsoft a través de Azure Studio ofrece un amplio catálogo de herramientas y modelos preentrenados de reconocidas plataformas como OpenAI, Hugging Face y Meta, que pueden

Tabla 2. Mayores empresas que prestan servicios de AIaaS.

Empresa	Servicio	Modalidad	Descripción
Alibaba	Qwen model	Texto	Serie de modelos LLM que permiten la creación de texto, servicios de traducción, simulación de diálogos y visualización de datos. Los modelos incluyen Qwen-Turbo, Qwen-Plus y Qwen-Max, y solo funcionan en chino e inglés.
	AIRec	Texto	Servicios de recomendación personalizada para empresas y desarrolladores.
	Intelligent Robot	Texto	Plataforma de diálogo basada en NLP que permite diálogos inteligentes a través de diversos clientes, como sitios web, aplicaciones móviles y robots.
	Intelligent Speech Interaction	Texto/Audio	Reconocimiento de voz, síntesis de voz y comprensión del lenguaje natural, ofreciendo una experiencia de interacción humano-computadora inmersiva.
Amazon	Amazon Q	Texto	Asistente basado en la IA generativa para el desarrollo de software y aprovechar los datos internos de las empresas.
	Amazon Comprehend	Texto	Servicio NLP que se utiliza para descubrir información y conexiones valiosas en textos.
	Amazon Lookout	Multimodal	Servicios que incluye Metrics, Vision y Equipment. Metrics detecta anomalías en métricas e identifica automáticamente su causa raíz. Vision detecta defectos en productos, automatizando la inspección de calidad. Equipment previene tiempos de inactividad no planeados mediante la detección automática de comportamientos anormales en equipos.
	Amazon Rekognition	Image/Video	Reconocimiento de imágenes y el análisis de videos con el machine learning.
	Amazon Polly	Texto/Audio	Sintetizar habla que se asemeja a una voz humana, lo que le permite convertir el texto de artículos a voz.
	Amazon Transcribe	Audio/Texto	Convierte automáticamente audio a texto.
Anthropic	Claude	Multimodal	Modelos LLM multimodales de propósitos generales. Claude 3.5 Sonnet, ideal para casos de uso de alto volumen; Claude Haiku, el más rápido para acciones ligeras; y Claude Opus, el de mayor rendimiento para análisis complejos y tareas avanzadas.
Cohere	Command R	Texto	LLM altamente escalables que equilibran un alto rendimiento con una gran precisión.
	Rerank	Texto	Servicio semántico de búsqueda en cualquier sistema de búsqueda de palabras clave o vectores.
	Embed	Texto	NLP de representación de texto para resultados de búsqueda, RAG, clasificación y agrupación.
Google	Gemini	Multimodal	Serie de modelos LLM multimodales de Google para propósito general, incluyendo texto, imágenes, vídeos y código.
	Gemma	Multimodal	Conjunto de modelos LLM abiertos. Disponibles para ejecutarse en aplicaciones y hardware, dispositivos móviles o servicios alojados.
	Speech-to-Text	Audio/Texto	Servicio para convertir audio en transcripciones de texto y, además, integra el reconocimiento de voz en aplicaciones con APIs.
	Text-to-Speech	Texto/Audio	Servicio para convertir texto en voz natural mediante una API.
	Vision AI	Video	Modelos de visión avanzados para automatizar las tareas de visión, optimizar el análisis y obtener estadísticas prácticas. O bien, crea apps personalizadas con entrenamiento de modelos sin código.
	Natural Language	Texto	Modelo IA de procesamiento de lenguaje natural para relacionar texto no estructurado con el aprendizaje automático de Google.
	Document AI	Texto	Procesadores de alta precisión para extraer, clasificar y dividir información no estructurada.
Huawei	Image Recognition	Image	Servicio de reconocimiento de imágenes para identificar objetos, escenas y conceptos en imágenes.
	EIHealth	Multimodal	Plataforma profesional de I+D de IA para investigaciones y aplicaciones con IA en genómica, descubrimiento de fármacos e imágenes médicas.
	Video Intelligent Analysis Service	Video	VIAS es una solución integral para análisis de video impulsado por IA, detección de eventos y soporte para la toma de decisiones.
	Question Answering Bot	Texto	Crear, publicar y gestionar rápidamente robots inteligentes de preguntas y respuestas.
	Huawei HiLens	Multimodal	Plataforma de desarrollo de IA multimodal que permite la sinergia entre dispositivos y la nube.
	Content Moderation	Multimodal	Servicio de moderación de contenido, donde se adopta tecnologías de detección de imágenes, texto y video que detectan pornografía y anuncios.
IBM	watsonx Assistant	Texto/Audio	IA generativa que proporciona experiencias de autoservicio sin fricciones para los clientes. Ofrece modelos LLM, de voz, procesamiento y comprensión de lenguaje natural (NLP, NLU), y recopilación de contexto inteligente.
	Watson Discovery	Texto	Permite descubrir insights en los documentos empresariales. IA de lenguaje natural de grado empresarial, impulsada por modelos NLP y LLMs de IBM Research.
	Knowledge Catalog	Texto	Repositorio de metadatos empresarial en la nube para IA, ML y aprendizaje profundo. Permite encontrar, comprender, curar y acceder a datos y activos de conocimiento, independientemente de su ubicación.

Continúa en página siguente.

Continuación Tabla 2.

Empresa	Servicio	Modalidad	Descripción
IBM	Watson Natural Language Understanding	Texto	Servicio para extraer significado y metadatos de texto no estructurado. Mediante análisis de texto, permite extraer categorías, clasificaciones, entidades, palabras clave, sentimiento, emociones, relaciones y sintaxis.
	Natural Language Processing Library	Texto	Biblioteca de NLP contenedorizada, diseñada para ofrecer flexibilidad a los socios de IBM al integrar potente IA de lenguaje natural en sus soluciones.
	Watson Speech to Text	Audio/Texto	Tecnología que permite la transcripción rápida y precisa del habla en múltiples idiomas para diversos casos de uso.
	Watson Text to Speech	Texto/Audio	Servicio que convierte texto escrito en audio natural en varios idiomas y voces.
Meta	Meta AI	Multimodal	Chatbot basado en LLaMA. Meta AI se integra en las interacciones digitales diarias mediante la tecnología de chat de Meta.
	Llama	Multimodal	LLM más avanzado de Meta. Es multimodal, de código abierto y está diseñado con investigación abierta para la comunidad. Ofrece diferentes versiones de modelos según la cantidad de ventana de contexto.
Microsoft	Azure OpenAI Service	Multimodal	Ofrece acceso a los LLM de OpenAI (GPT-4/4o/4o mini, GPT-3.5, DALL-E 2, Whisper) en un entorno seguro y escalable.
	Azure AI Document Intelligence	Texto	Servicio avanzado para extraer texto, pares clave-valor, tablas y estructuras de documentos de forma automática y precisa.
	Azure AI Language	Texto	Servicio administrado para desarrollar aplicaciones de NLP. Identifica términos y frases clave, analiza opiniones, resume textos y crea interfaces conversacionales.
	Azure AI Vision	Video	Servicio que ofrece funcionalidades de computer vision. Proporciona la capacidad de analizar imágenes, leer texto y detectar caras con el etiquetado de imágenes predefinido, la extracción de texto con OCR y reconocimiento facial.
	Azure AI Speech	Audio/Texto	Servicios de procesamiento de voz, incluyendo la conversión de voz a texto, la síntesis de texto a voz, la traducción automática y la identificación de oradores.
	Azure AI Translator	Texto	Servicio de traducción que permite a las empresas traducir texto y documentos entre más de 100 idiomas.
	Phi-3	Multimodal	Familia de modelos de lenguaje pequeños y multimodales desarrollados por Microsoft AI.
	Copilot Azure	Multimodal	Asistente IA que simplifica la gestión de aplicaciones e infraestructuras en la nube, desde el diseño hasta la operación y optimización, tanto en la nube como en el edge.
	Azure AI Search	Texto	Servicio de respuestas de alta calidad con una base de datos de vectores creada para la generación RAG.
Mistral	Mistral Models	Texto	Conjunto de LLM que se distinguen por su eficiencia y portabilidad. Mistral se destaca por su compromiso con la IA responsable y transparente, ofreciendo sus modelos bajo licencias de código abierto.
Nvidia	Maxine	Video/Audio	SDK acelerado por GPU con funciones IA para que los desarrolladores creen aplicaciones de creación de contenido y colaboración virtual.
	Clara	Multimodal	Plataforma para la generación de imágenes, la genómica, el monitoreo de pacientes y el descubrimiento de fármacos.
	Picasso	Texto/Imagen	Servicios para crear e implementar modelos de IA generativa para contenido visual. NVIDIA Picasso ofrece entrenamiento y personaliza modelos de IA generativa visual.
	BioNeMo	Texto	IA para el descubrimiento de fármacos que simplifica y acelera el entrenamiento de modelos utilizando sus propios datos.
	Omniverse Cloud	Framework	Servicios para desarrolladores empresariales integren tecnologías de Universal Scene Description (OpenUSD) y RTX en sus aplicaciones de digitalización industrial 3D.
	AI Enterprise	Multimodal	Software nativa para acelerar los procesos de ciencia de datos y agilizar el desarrollo y la implementación de copilotos de nivel de producción y otras aplicaciones de IA generativa.
OpenAI	GPT	Multimodal	LLM multimodal entrenado con una gran cantidad de datos de texto, código e imágenes.
	Whisper	Audio/Texto	Modelo para transcribir y traducir voz en tiempo real con alta precisión.
	Sora	Texto/Video	Modelo de inteligencia artificial multimodal que genera videos realistas a partir de descripciones de texto.
	Dall-e	Texto/Imagen	Modelo de inteligencia artificial multimodal que genera videos realistas a partir de descripciones de texto.
	o1	Texto	Serie de modelos de IA diseñados para dedicar más tiempo a pensar antes de responder. Pueden razonar sobre tareas complejas y resolver problemas más difíciles que los modelos anteriores en ciencia, programación y matemáticas.

Tabla 3. Empresas con Suite de desarrollo de AIaaS.

Empresa	Suite	Descripción
Alibaba	Platform for AI	Plataforma de aprendizaje automático y profundo diseñada para empresas y desarrolladores. Ofrece plug-ins para diversos escenarios industriales. Con más de 140 algoritmos de optimización incorporados, proporciona capacidades completas de IA, incluyendo etiquetado de datos, construcción de modelos y entrenamiento de modelos.
Amazon	Amazon Bedrock	Servicio que ofrece una selección de modelos fundacionales de las principales empresas de IA como AI21 Labs, Anthropic, Cohere, Meta, Mistral AI, Stability AI y Amazon, junto con un amplio conjunto de funciones necesarias para crear aplicaciones de IA generativa con seguridad, privacidad e IA responsables.
DataRobot	DataRobot	Plataforma de ciencia de datos que permite a las empresas utilizar herramientas para construir y desplegar rápidamente modelos de aprendizaje automático. Soporta Python y R, y se integra con herramientas como TensorFlow, PyTorch y servicios en la nube. DataRobot automatiza la construcción y despliegue de modelos, facilitando la gestión del flujo de trabajo para científicos de datos y analistas.
Google	Vertex AI	Plataforma unificada y gestionada para desarrollar y usar IA generativa. Ofrece acceso a más de 150 modelos, incluyendo Gemini, donde los desarrolladores pueden utilizar texto, imágenes, vídeos y código para crear aplicaciones. Vertex AI integra herramientas y soporta modelos de terceros (como Anthropic's Claude 3 y OpenAI's GPT) y abiertos.
Huawei	ModelArts	Plataforma de inteligencia artificial totalmente gestionada por Huawei Cloud. Ofrece herramientas y servicios integrados para el desarrollo, entrenamiento y despliegue de modelos de IA. La plataforma soporta múltiples frameworks de aprendizaje automático y profundo, incluyendo TensorFlow y PyTorch, y proporciona un entorno colaborativo para equipos de ciencia de datos. ModelArts permite la implementación de modelos en la nube, en entornos locales o en dispositivos edge.
Microsoft	Azure Studio	Plataforma unificada para desarrollar e implementar aplicaciones de inteligencia artificial generativas. Incluye un catálogo sólido de herramientas y modelos de código abierto o empresas como OpenAI, Hugging Face, Meta y más.
Nvidia	NeMo	Plataforma de extremo a extremo para desarrollar IA generativa personalizada en cualquier lugar. Ofrezca modelos listos con curación de datos precisa, personalización de vanguardia, RAG y rendimiento acelerado.

integrarse y aplicarse a datos específicos, permitiendo también incorporar modelos propios soportados por frameworks como TensorFlow y PyTorch.

A nivel de infraestructuras de IA, la Tabla 4 muestra los servicios de las empresas para desplegar aplicativos IA a gran escala. Estas infraestructuras permiten a las empresas innovar, optimizar procesos y reducir costos, acelerando la adopción de soluciones de IA en diversos sectores industriales. Alibaba, con su PAI-Lingjun, soporta computación inteligente y opciones sin servidor en Alibaba Cloud. Amazon SageMaker ofrece un conjunto completo de herramientas de alto rendimiento y bajo costo, adecuado para múltiples casos de uso. Google AI Infrastructure proporciona GPUs, TPUs y CPUs para el entrenamiento y procesamiento de datos a gran escala, permitiendo una configuración eficiente de clústeres. Microsoft Azure Machine Learning gestiona todo el ciclo de vida de proyectos de ML en la nube, desde el desarrollo hasta la implementación y optimización. Por último, Nvidia DGX Cloud permite entrenar modelos de IA para desarrolladores empresariales mediante una experiencia sin servidor.

OPORTUNIDADES, DESAFÍOS Y LIMITACIONES

La AIaaS presenta un panorama complejo de oportunidades, desafíos y limitaciones para las organizaciones que buscan aprovechar estas tecnologías avanzadas.

Las oportunidades que ofrece AIaaS son significativas. Como se evidencia en la Figura 2, una amplia gama de servicios está disponible, permitiendo la democratización de la IA. Empresas de todos los tamaños pueden ahora acceder a tecnologías sofisticadas sin necesidad de grandes inversiones iniciales en infraestructura o personal

Tabla 4. Empresas con infraestructura para IA en cloud.

Empresa	Infraestructura	Descripción
Alibaba	PAI-Lingjun	PAI-Lingjun Intelligent Computing Service es un servicio PaaS para aprendizaje profundo a gran escala y computación inteligente integrada. Ofrece una edición Serverless en la nube pública de Alibaba Cloud y una Edición Exclusiva.
Amazon	Amazon Sagemaker	Servicio totalmente administrado que reúne un amplio conjunto de herramientas para permitir ML de alto rendimiento y bajo costo para cualquier caso de uso.
Google	AI Infrastructure	GPUs, TPUs o CPUs para cubrir distintos casos prácticos, como entrenamiento de alto rendimiento, inferencias de bajo coste y procesamiento de datos a gran escala. Permite configurar rápidamente entornos de aprendizaje automático, automatizar la orquestación, gestionar clústeres grandes y configurar aplicaciones de baja latencia.
Microsoft	Azure Machine Learning	Plataforma integral en la nube que permite a los usuarios acelerar y gestionar el ciclo de vida completo de sus proyectos ML. Desde el desarrollo y entrenamiento de modelos hasta su implementación, monitorización y optimización.
Nvidia	DGX Cloud	Plataforma de entrenamiento de IA como servicio que ofrece una experiencia sin servidor para desarrolladores empresariales optimizada para IA generativa.

especializado. Esto facilita la innovación acelerada, permitiendo a las organizaciones experimentar con soluciones de IA de forma rápida y con menor riesgo financiero. Además, la naturaleza escalable de estos servicios ofrece la flexibilidad de ajustar recursos según las necesidades cambiantes, optimizando costos operativos. Quizás lo más importante es que AIaaS proporciona herramientas analíticas avanzadas que pueden mejorar significativamente la toma de decisiones empresariales, ofreciendo insights valiosos basados en el análisis de grandes volúmenes de datos.

Sin embargo, la adopción de AIaaS también conlleva desafíos considerables. Uno de los más apremiantes es la eficiencia energética. El procesamiento intensivo requerido para ejecutar modelos de IA complejos consume una cantidad sustancial de energía, planteando preocupaciones de sostenibilidad. Este desafío se extiende a la necesidad de una infraestructura robusta capaz de manejar cargas de trabajo intensivas y crecientes sin comprometer el rendimiento o la rentabilidad. La gestión efectiva de datos presenta otro desafío significativo, ya que las empresas deben navegar la complejidad de escalar sus capacidades de procesamiento y almacenamiento. Además, la seguridad y privacidad de los datos se convierten en preocupaciones críticas al utilizar servicios en la nube, requiriendo estrategias sólidas para proteger información sensible y cumplir con regulaciones cada vez más estrictas. La integración de soluciones AIaaS con sistemas existentes también puede resultar compleja, necesitando adaptaciones significativas en la infraestructura tecnológica de las organizaciones.

En cuanto a las limitaciones, aunque AIaaS reduce las barreras de entrada, los costos operativos continuos pueden ser una limitante, especialmente para pequeñas y medianas empresas. Esto se agrava por la escasez crítica de talento especializado capaz de implementar y mantener estas soluciones de manera efectiva. Esta brecha de habilidades se ve exacerbada por una brecha educativa, donde muchas instituciones académicas aún no han incorporado tecnologías como AIaaS en sus planes de estudio. Además, la adopción de soluciones AIaaS puede crear una dependencia significativa de proveedores específicos, como se puede observar en la diversidad de servicios ofrecidos como lo detallan las Tablas presentes en el documento, lo que podría limitar la flexibilidad futura de las empresas.

CONCLUSIONES

La Inteligencia Artificial como Servicio está emergiendo como un componente crucial en la evolución de la industria hacia modelos más avanzados y eficientes, facilitando la transformación hacia la Industria 5.0. Los hallazgos de este estudio revelan la creciente importancia de AIaaS, con una amplia gama de servicios ofrecidos por empresas líderes que abarcan desde el procesamiento de lenguaje natural hasta el análisis predictivo y la visión por computadora. Estas soluciones están entregando el acceso a tecnologías de IA avanzadas, permitiendo a organizaciones y usuarios de todos

los tamaños implementar capacidades de IA sin grandes inversiones iniciales.

Se observa una tendencia hacia la integración de AIaaS con metodologías ágiles, mejorando la eficiencia en la gestión de proyectos y optimizando procesos. Sin embargo, la adopción generalizada de estas tecnologías presenta desafíos significativos, como la necesidad de gestionar grandes volúmenes de datos, abordar preocupaciones de seguridad y privacidad, y adaptar las estructuras organizativas. Es crucial reconocer que la implementación efectiva de AIaaS requiere no solo inversión en tecnología, sino también un cambio cultural y organizativo significativo.

El artículo presenta una valiosa recopilación de recursos a la fecha del artículo en AIaaS a través de cuatro tablas detalladas. Se detallan las últimas investigaciones en IA de grandes empresas tecnológicas, se ofrece un panorama completo de los servicios AIaaS disponibles, incluyendo modalidades, descripciones y suites de desarrollo y las infraestructuras de IA en la nube presente actualmente. Esta síntesis de los modelos y servicios más populares proporciona una visión integral del estado actual de AIaaS, permitiendo a los lectores comprender las opciones disponibles y las tendencias del mercado. Además, se identifican oportunidades prometedoras para la innovación en áreas como el mantenimiento predictivo, la optimización de la cadena de suministro y la personalización de productos y servicios mediante AIaaS. A medida que la industria continúa su transformación digital, será esencial mantener un equilibrio entre la adopción de tecnologías AIaaS y la consideración cuidadosa de sus implicaciones a largo plazo. El éxito en la integración de estas soluciones dependerá de la capacidad de las organizaciones para adaptarse, innovar y abordar de manera proactiva los desafíos que surjan en el camino hacia una industria más inteligente y conectada.

Finalmente, la integración de AIaaS en la Industria 5.0 promueve la sostenibilidad al optimizar el uso de recursos y reducir el impacto ambiental mediante procesos más eficientes y precisos. Además, fomenta la empleabilidad al crear nuevas oportunidades laborales enfocadas en la gestión y desarrollo de tecnologías de inteligencia artificial. Sin embargo, esta transformación requiere una capacitación continua de la fuerza laboral para adaptarse a roles más especializados y colaborativos entre humanos y máquinas. Así, se equilibra el avance tecnológico con la creación de empleos sostenibles, promoviendo un desarrollo industrial responsable y resiliente.

AGRADECIMIENTOS

Este trabajo fue parcialmente financiado por la Escuela de Ingeniería Informática, Universidad de Valparaíso, Chile, a través de la beca REXE N.° 4054/2022. Diego Monsalves parcialmente financiado por Beca Doctorado Nacional de la Agencia Nacional de Investigación y Desarrollo resolución exenta 4402/2023. Pablo Olivares financiado por Beca Doctorado Nacional de la Agencia Nacional de Investigación y Desarrollo resolución exenta 6147/2024.

REFERENCIAS

[1] K. Niemiec, K. Niemiec, and D. Krenczyk, "Intelligent tools for improving quality and efficiency in production environment," *Journal of Physics: Conference Series*, vol. 2198, no. 1, Art. no. 012009, doi: 10.1088/1742-6596/2198/1/012009.

[2] C. Li, Y. Chen, and Y. Shang, "A review of industrial big data for decision making in intelligent manufacturing," *Engineering Science and Technology, an International Journal*, vol. 29, Art. 101021, 2022, doi: 10.1016/j.jestch.2021.06.001.

[3] A. Aldoseri, K.N. Al-Khalifa, and A.M. Hamouda, "Re-thinking data strategy and integration for artificial intelligence: Concepts, opportunities, and challenges," *Applied Sciences*, vol. 13, no. 12, 2023, doi: 10.3390/app13127082 .

[4] A. Brem, F. Giones, and M. Werle, "The AI digital revolution in innovation: A conceptual framework of artificial intelligence technologies for the management of innovation," *IEEE Transactions on Engineering Management*, vol. 70, no. 2, pp. 770-776, 2023, doi: 10.1109/TEM.2021.3109983.

[5] Mordor Intelligence Research & Advisory, "Digital transformation market size & share analysis - growth trends & forecasts (2024 - 2029)," mordorintelligence.com

2024. [Online]. Available: https://www.mordorintelligence.com/industry-reports/digital-transformation-market

[6] C. Hendriksen, "Artificial intelligence for supply chain management: Disruptive innovation or innovative disruption?," *Journal of Supply Chain Management*, vol. 59, no. 3, pp. 65-76, 2023, doi: 10.1111/jscm.12304.

[7] J. Xue *et al.*, "Evaluation of the current state of chatbots for digital health: Scoping review," *Journal Medical Internet Research*, vol. 25, Art. no. e47217, Dec 2023, doi: 10.2196/47217.

[8] M. Ridha and K. Haura Maharani, "Implementation of artificial intelligence chatbot in optimizing customer service in financial technology company pt. finaccel finance indonesia," *Proceedings*, vol. 83, no. 1, 2022, doi: 10.3390/proceedings2022083021.

[9] D. Monsalves, F. Riquelme, and H. Cornide-Reyes, "Enhancing Education through Multimodal Learning Analytics and AI-as-a-Service," in *2023 IEEE CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON),* Valdivia, Chile, 2023, pp. 1-6, doi: 10.1109/CHILECON60335.2023.10418760.

[10] N. Guntupalli and V. Rudramalla, "Artificial Intelligence as a Service: Providing Integrity and Confidentiality," in *Multi-disciplinary Trends in Artificial Intelligence*, R. Moruspalli, T.S. Dandibhotla, V.V. Atluri, D. Windridge, P. Lingras, and V.R. Komati, Eds., Switzerland: Springer, 2023, pp. 309-315, isbn: 978-3-031-36402-0.

[11] P. Spagnoletti, N. Kazemargi, and A. Prencipe, "Agile practices and organizational agility in software ecosystems," *IEEE Transactions on Engineering Management*, vol. 69, no. 6, pp. 3604-3617, 2022, doi: 10.1109/TEM.2021.3110105.

[12] A.T. Barbosa, C.C. Da Silva, R.L. Caetano, D.P.S. Da Silva, J.V. Barbosa, and Z.T. Pinto, "Metodologias Ágeis: E su aplicabilidade na Área de marketing," *Revista Ibero-Americana de Humanidades, Ciências e Educação*, vol. 8, no. 3, pp. 1659-1669, 2022, doi: 10.51891/rease.v8i3.4751.

[13] P. Bylund, J. Mehta, N. Mathavan, K. Trowbridge, B. Taylor, H. El Sharu, and N. Campbell, "Using agile methodology and nudge strategies to improve enrollment in clinical trials," Innovation in Aging, vol. 6, pp. 652-652, 2022, doi: 10.1093/geroni/igac059.2409.

[14] T. Broda *et al.*, "The agile working method for legal departments," *Jusletter IT*, no. 20, 2022, 2021, doi: 10.38023/fc0c658d-1230-45ca-bf08-bf3113e2d678.

[15] M. Lourens, R. Raman, P. Vanitha, R. Singh, G. Manoharan, and M. Tiwari, "Agile Technology and Artificial Intelligent Systems in Business Development," 2022 5th International Conference on Contemporary Computing and Informatics (IC3I), Uttar Pradesh, India, 2022, pp. 1602-1607, doi: 10.1109/IC3I56241.2022.10073410.

[16] B. Alojaiman, "Technological Modernizations in the Industry 5.0 Era: A Descriptive Analysis and Future Research Directions," *Processes*, vol. 11, no. 5, 2023, doi: 10.3390/pr11051318.

[17] A. Agote Garrido, A. Martín Gómez y J. Lama, "Industry 5.0 values. A bibliometric analysis of the new industrial paradigm from the social approach", *Dyna*, vol. 98, 2023, doi: 10.6036/10834.

[18] J. Pizoń and M. Kulisz, "Industry 5.0: New approaches to innovation in the industry," in *Innovation in the Digital Economy. New Approaches to Management for Industry 5.0*, A. Rzepka Ed., New York, USA: Routledge, 2023, pp. 38-50, doi: 10.4324/9781003384311-4.

[19] M. Golovianko, V. Terziyan, V. Branytskyi, and D. Malyk, "Industry 4.0 vs. Industry 5.0: Co-existence, Transition, or a Hybrid," in *Procedia Computer Science*, *4th International Conference on Industry 4.0 and Smart Manufacturing*, vol. 217, pp. 102-113, 2023, doi: 10.1016/j.procs.2022.12.206.

[20] D. Slavic, "The main concepts of Industry 5.0: A Bibliometric Analysis Approach," in *2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH),* East Sarajevo, Bosnia and Herzegovina, 2023, pp. 1-5, doi: 10.1109/INFOTEH57020.2023.10094143.

[21] Y. Talebirad and A. Nadiri, "Multi-Agent Collaboration: Harnessing the Power of Intelligent LLM Agents," 2023, ArXiv: 2306.03314.

[22] I. de Zarzà, J. de Curtò, G. Roig, P. Manzoni, and C. T. Calafate, "Emergent Cooperation and Strategy Adaptation in Multi-Agent Systems: An Extended Coevolutionary Theory with LLMs," *Electronics*, vol. 12, no. 12, 2023, doi: 10.3390/electronics12122722.

[23] F.S. Karal and A. Soyer, "The impact of digital enterprise transformation strategies on project managers' competencies," in *Handbook of Research on Management and Strategies for Digital Enterprise Transformation*, Kamaljeet Sandhu, Eds., Hershey, Pa, USA: IGI Global Scientific Publishing, 2020, pp. 274-295, doi: 10.4018/978-1-7998-5015-1.CH014.

[24] T. Gustavsson, M. Berntzen, and V. Stray, "Changes to team autonomy in largescale software development: a multiple case study of Scaled Agile Framework (SAFe) implementations," *International Journal of Information Systems and Project Management*, vol. 10, no. 1, p. 29-46, Apr. 2022, doi: 10.12821/ijispm100102.

[25] M. Tenemaza and C. Sarzosa, "Agile methodologies in the fashion industry," *Intelligent Human Systems Integration*, T. Ahram, W. Karwowski, P. Di Bucchianico, R. Taiar, L. Casarotto, and P. Costa, Eds., vol. 22, 2022, doi: 10.54941/ahfe1001050.

[26] A. Erdil, "Industry 4.0 perception regarding to new developments and new trends of industries," *Avrupa Bilim ve Teknoloji Dergisi*, no. 28, pp. 228-240, 2021, doi: 10.31590/ejosat.996172.

[27] H. Yüksel, "Industry 4.0 transformation: factors affecting adoption and impacts on companies," *International Journal of Industrial Engineering and Operations Management*, vol. 4, no. 3, pp. 63-89, 2022, doi: 10.1108/ijieom-06-2022-0020.

[28] P.K.R. Maddikunta *et al.*, "Industry 5.0: A survey on enabling technologies and potential applications," *Journal of Industrial Information Integration*, vol. 26, Art. no. 100257, 2022, doi: 10.1016/j.jii.2021.100257.

[29] X. Li, A. Sigov, L. Ratkin, L.A. Ivanov, and L. Li, "Artificial intelligence applications in finance: a survey," *Journal of Management Analytics*, vol. 10, no. 4, pp. 676-692, 2023,

[30] S.J. Plathottam, A. Rzonca, R. Lakhnori, and C.O. Iloeje, "A review of artificial intelligence applications in manufacturing operations," *Journal of Advanced Manufacturing and Processing,* vol. 5, no. 3, Art. no. e10159, 2023, doi: 10.1002/amp2.10159.

[31] T.V. Andrianandrianina Johanesa, L. Equeter, and S.A. Mahmoudi, "Survey on AI Applications for Product Quality Control and Predictive Maintenance in Industry 4.0," *Electronics*, vol. 13, no. 5, Mar. 2024, doi: 10.3390/electronics13050976.

[32] K. Denecke, R. May, and O. Rivera-Romero, "Transformer Models in Healthcare: A Survey and Thematic Analysis of Potentials, Shortcomings and Risks," *Journal of Medica*l Systems, vol. 48, no. 1, 2024, doi: 10.1007/ s10916-024-02043-5.

[33] C.H. Lee, C. Wang, F. Li, Q. Deng, and D. Chang, "A Survey of Artificial-Intelligence Enabled Digital Transformation in Elderly Healthcare Field," in *Transdisciplinarity and the Future of Engineering. Advances in Transdisciplinary Engineering,* vol. 28, B. Moser, P. Koomsap, and J. Stjepandic, Eds., Clifton, VA, USA: Ios Press, 2022, pp.330-339, doi: 10.3233/ATDE220662.

[34] E. Elbasi *et al.*, "Artificial Intelligence Technology in the Agricultural Sector: A Systematic Literature Review," *IEEE Access*, vol. 11, pp. 171-202, 2023, doi: 10.1109/ACCESS.2022.3232485.

[35] A. Younas, K. Subramanian, M. Al-Haziazi, S. Hussainy, and A. Kindi, "A Review on Implementation of Artificial Intelligence in Education," *International Journal of Research and Innovation in Social Science*, vol. 7, no. 8, pp. 1092-1100, 2023, doi: 10.47772/ IJRISS.2023.7886.

[36] N.V. Salastekar, C. Maxfield, T.N. Hanna, E.A. Krupinski, D. Heitkamp, and L.J. Grimm, "Artificial Intelligence/Machine Learning Education in Radiology: Multiinstitutional Survey of Radiology Residents in the United States," *Academic Radiology*, vol. 30, no. 7, pp. 1481-1487, 2023, doi: 10.1016/j.acra.2023.01.005.

[37] Y. Cao, A.A. Aziz, and W.N.R.M. Arshard, "University students' perspectives on Artificial Intelligence: A survey of attitudes and awareness among Interior

Architecture students," *International Journal of Educational Research and Innovation*, no. 20, 2023, doi: 10.46661/ijeri.8429.

[38] M. Govett *et al.*, "Exascale Computing and Data Handling: Challenges and Opportunities for Weather and Climate Prediction," *Bulletin of the American Meteorological Society*, vol. 105, no. 12, 2024, doi: 10.1175/BAMS-D-23-0220.1.

[39] P. Gupta and M. Pandey, "Role of AI for Smart Health Diagnosis and Treatment," *Smart Medical Imaging for Diagnosis and Treatment Planning*, Boca Raton, FL, USA: Chapman and Hall/CRC, 2024, pp. 23-45, doi: 10.1201/9781003464884-4.

[40] Z. Xinrui, "Application of artificial intelligence technology in hospital management systems," *Applied and Computational Engineering*, vol. 50, pp. 229-233, 2024, doi: 10.54254/2755-2721/50/20241548.

[41] O. Onesi-Ozigagun, Y.J. Ololade, N.L. Eyo-Udo, and D. O. Ogundipe, "Revolutionizing education through ai: a comprehensive review of enhancing learning experiences," *International Journal of Applied Research in Social Sciences*, vol. 6, no. 4, 2024, doi: 10.51594/ijarss.v6i4.1011.

[42] J. Young, "The rise of artificial intelligence in education," *International Journal of Innovative Research and Development*, vol.13, no. 2, 2024, doi: 10.24940/ijird/2024/v13/i2/feb24019.

[43] S. Gavin *et al.*, "LongIns: A Challenging Long-context Instruction-based Exam for LLMs," 2024, arXiv: 2406.17588.

[44] T. Li, G. Zhang, Q.D. Do, X. Yue, and W. Chen, "Long-context LLMs Struggle with Long In-context Learning ," 2024, arXiv: 2404.02060.

[45] Y. Kuratov, A. Bulatov, P. Anokhin, I. Rodkin, and M. Burtsev, "BABI-Long: Testing the Limits of LLMs with Long Context Reasoning-in-a-Haystack," 2024, arXiv: 2406.10149.

[46] H. Qian, Z. Liu, P. Zhang, K. Mao, Y. Zhou, C. Xu, and Z. Dou, " Are Long-LLMs A Necessity For Long-Context Tasks? ," 2024, arXiv.2405.15318.

[47] OpenAI, "OpenAI o1 System Card," openai.com, 2024. [Online]. Available: https://cdn.openai.com/o1-system-card.pdf

[48] R. Marino, "Fast Analysis of the OpenAI O1-Preview Model in Solving Random K-SAT Problem: Does the LLM Solve the Problem Itself or Call an External SAT Solver?," 2024, arXiv.2409.11232.

[49] P. Tominc, D. Oreški, and M. Rožman, "Artificial intelligence and agility-based model for successful project implementation and company competitiveness," *Information*, vol. 14, no. 6, Art. no. 337, 2023, doi: 10.3390/info14060337.

[50] D.L. Young, P. Boyette, J. Moreland, and J. Teske, "Generative ai agile assistant," in *Disruptive Technologies in Information Sciences VIII*, B.T. Wysocki, M. Blowers, and R. Bharadwaj, Eds., 2024, doi: 10.1117/12.3011173.