

# ChatGPT: la IA está aquí y nos desafía

*“si actúa de manera inteligente... es inteligente”*

**Alan Turing**

## Génesis del ChatGPT

La inteligencia artificial (IA) tiene como objetivo lograr una inteligencia general similar a la humana, lo que es considerado uno de los objetivos más ambiciosos de la ciencia. Este objetivo se compara con otros grandes objetivos científicos, como explicar el origen de la vida o conocer la estructura de la materia. A lo largo del tiempo, se han utilizado diferentes modelos o metáforas del cerebro. En 1975, Allen Newell y Herbert Simon presentaron la hipótesis del Sistema de Símbolos Físicos (SSF), según la cual cualquier sistema de símbolos físicos tiene la capacidad para llevar a cabo acciones inteligentes. Los seres humanos también son considerados sistemas de símbolos físicos. La IA busca verificar esta hipótesis en computadoras convenientemente programadas.

La inteligencia de tipo general es el objetivo de la IA y difiere de la inteligencia específica, que se limita a un ámbito muy concreto. La IA fuerte se refiere a la capacidad de los ordenadores para tener una conducta inteligente de tipo general. En cuanto a los modelos de IA, se ha propuesto una aproximación bioinspirada, como la computación evolutiva, que busca mejorar automáticamente las soluciones a problemas mediante la imitación de la evolución. Aunque tiene limitaciones, ha producido buenos resultados en la resolución de problemas de optimización.

En los últimos 20 años, la IA ha experimentado un rápido avance gracias a la disponibilidad de grandes cantidades de datos y la capacidad de procesamiento de alta velocidad. Esto ha permitido el desarrollo de sistemas de IA especializados en tareas específicas, como el reconocimiento de voz, el procesamiento del lenguaje natural, la visión por computadora, los vehículos autónomos y los asistentes virtuales. Uno de los mayores avances en la IA en los últimos años ha sido el aprendizaje profundo, que utiliza redes neuronales artificiales para aprender patrones y características de los datos de entrada y realizar tareas como la clasificación, la detección de objetos y la generación de texto y audio. Esto ha permitido el desarrollo de sistemas de IA más precisos y avanzados, como AlphaGo de DeepMind, que derrotó al campeón mundial de Go, y GPT-3 de OpenAI, que puede generar texto de manera impresionante similar a los humanos.

GPT-3 (Generative Pre-trained Transformer 3) es un modelo de lenguaje natural de última generación desarrollado por OpenAI que utiliza aprendizaje profundo para generar texto de alta calidad. Basado en la arquitectura Transformer, el modelo es capaz de procesar grandes cantidades de datos y detectar patrones en ellos. Además, GPT-3 es un modelo preentrenado, lo que significa que se ha condicionado previamente en una gran cantidad de texto para entender la estructura del lenguaje natural y las relaciones semánticas entre las palabras.

Utilizando un modelo de lenguaje de aprendizaje profundo, GPT-3 es capaz de predecir la siguiente palabra en una oración o el texto completo de una respuesta a una pregunta. El modelo de lenguaje está entrenado en una tarea de "auto-supervisión", lo que significa que se entrena para predecir la siguiente palabra en una secuencia dada la información anterior, sin necesidad de una respuesta externa o una etiqueta de objetivo.

GPT-3 es un modelo de lenguaje de "muestreo autoregresivo", lo que significa que genera texto de forma secuencial, prediciendo la siguiente palabra en función de las palabras anteriores. Para generar texto, utiliza una técnica llamada "decodificación por haz", generando varias secuencias de palabras simultáneamente y seleccionando la secuencia más probable como resultado final.

### **Sentido de utilidad entre luces y sombras**

ChatGPT apareció a finales de noviembre de 2022 y a mediados de febrero de 2023 alcanzó su versión estable. Además está decir que la comunidad científica aún no se ha manifestado sobre este prototipo de IA específico, aunque se han marcado desventajas y desviaciones en modelos y productos anteriores. Desde el punto de vista del impacto y adopción de la tecnología, podemos estimar que en sólo dos meses se encuentra transitando el pico de expectativas infladas (Ciclo de Gartner) en base a su éxito relativo y a las audaces promesas que maravillan a los curiosos, y en este contexto se espera una pendiente de desilusión como paso previo a la mesta de la productividad que finalmente determine el nivel de adopción que alcanzará en cada uno de los ámbitos donde los humanos la incorporemos.

Los modelos de redes neuronales artificiales recurrentes o asociativas, trabajan por un sistema de aprendizaje supervisado seguido por un aprendizaje por refuerzo. Para el año 2010 ya se hablaba de machine learning aprovechando el crecimiento exponencial de almacenamiento y procesamiento de datos, mientras que para 2017 aparecen los algoritmos "transformer" de Google que fueron aprovechados por OpenAI para generar las primeras versiones de GPT en 2019. El problema principal que presentaron estos prototipos fue la existencia de sesgos marcados en la información de alimentación de la IA que estaba integrada por fuentes diversas de internet. En el GPT-3 aparece la versión de aprendizaje por refuerzo sobre una base de fuentes verificadas en conjunto con la información de realimentación, por tanto, tiene un modelo híbrido que ha demostrado efectividad para llegar a la difusión actual. Mientras tanto, los grandes operadores del mercado tecnológico (Google, Amazon, Meta, etc.) están anunciando fuertes inversiones en IA para los próximos años e incluso han presentado en algunos casos versiones prototipo que han arrojado varios errores. Recientemente Meta ha indicado que su línea de trabajo en IA será un producto de apoyo a investigadores, maximizando la fiabilidad de las fuentes para diferenciarse del resto de los desarrolladores.

En el ámbito empresarial, por el contrario, muchas empresas implementaron o piensan integrar la IA en sus operaciones. Puntualmente, de los 1.000 líderes empresariales del informe de Resume Builder (Resumebuilder.com), muchos aseguraron que ChatGPT reemplazó a trabajadores en sus empresas. El 66% declara usar ChatGPT para escribir código, el 58% para redacción y creación de contenido, el 57% para atención al cliente y el 52% para resúmenes de reuniones y otras tareas. Asimismo, más del 65% le han encontrado utilidades en los procesos de administración de personal (escribir descripciones de puestos, citar a entrevistas, responder solicitudes, etc.). Dado que esta nueva tecnología está aumentando en el lugar de trabajo, los trabajadores seguramente deben pensar en cómo puede afectar las responsabilidades de su trabajo actual. Los resultados de esta encuesta muestran que los empleadores buscan simplificar algunas responsabilidades laborales utilizando ChatGPT.

En la sociedad misma han aparecido estas últimas semanas un sinnúmero de estudios relativos a las potencialidades del ChatGPT, que se han difundido ampliamente en los medios de prensa digital e impresa. Desde la experiencia de Kevin Roose (New York Times) que se declara perturbado e incómodo por la dinámica de su charla, hasta la experiencia de Mariano Vidal (Clarín.com) que destaca la potencia de la herramienta.

Es por todo esto que desde AACINI-RIII quisimos experimentar una aproximación a la IA del ChatGPT (<https://openai.com/>) e intentar navegar entre los temores y las potencialidades que los docentes e investigadores asignamos a priori a este nuevo actor con el que deberemos convivir.

## **Nuestra experiencia con el ChatGPT**

En tanto que el mismo bot se define como un “modelo de lenguaje”, quisimos probar la capacidad de interpretación y análisis de consignas y la potencialidad para responder cuestiones relativas a la Ingeniería Industrial.

Comenzamos la prueba alimentando al sistema con una serie de consignas simples de parcial. En general encontramos que la capacidad de resolución de cuestiones simples es potente, incluso al pedido de elaboración de código y de análisis que soliciten relaciones de conceptos. Claramente, las fuentes y recursos del sistema no son necesariamente las mismas usadas por una cátedra en particular, por lo que la especificidad mejora cuando uno pregunta y solicita el punto de vista de autores específicos.

Para el análisis de casos y la provisión de ejemplos específicos, el sistema arroja respuestas genéricas y en general no converge a un resultado satisfactorio, aunque se intenten varias reformulaciones para ajustar los errores.

La herramienta se ha mostrado potente a la hora de proponer ejercicios, evaluaciones integrales, consignas y cuestionarios sobre temas específicos, aunque el nivel de desarrollo y relación requerido es escaso. Sin duda esta herramienta puede ayudar con ideas iniciales, pero difícilmente pueda elaborar un problema de elevada complejidad que plantee situaciones hipotéticas complejas.

Otra de las aplicaciones experimentadas tiene que ver con la capacidad de interpretación y procesamiento de texto. Para este ejercicio le copiamos los resúmenes de los artículos del presente número de la AACINI-RIII que corresponden al XV COINI y le pedimos que establezca las directrices comunes del conjunto. En este punto, el sistema es capaz de ofrecer un resumen integrado de contenido que extrae los 4 ejes principales de estudio.

Le preguntame en este punto su impresión sobre el rol que cumple la tecnología en la calidad de vida de las personas y nos indica que no cuenta con suficiente información, por lo que agregamos las conclusiones de los artículos y el resultado es asombroso. El sistema encuentra las 4 áreas de interés donde la tecnología ha sido estudiada en relación con la calidad de vida de las personas.

Luego de este experimento completamos el ejercicio escribiendo un resumen tomando las partes de interés, incorporando impresiones personales nuestras y agregando nueva información, y le pedimos nuevamente que evalúe si el texto propuesto por nosotros era representativo del contenido del presente número de la RIII. Si a estas alturas usted duda de la efectividad del proceso, debo decir que el ChatGPT ha quedado muy conforme con el trabajo y ha marcado los principales lineamientos del texto.

Si desea consultar el texto elaborado con este asistente puede leer nuestro artículo editorial... que es el producto final obtenido en la interacción. Asimismo, la síntesis de conceptos elaborada por el ChatGPT se ha colocado como resumen del artículo, y las palabras clave que figuran son las sugeridas también por el ChatGPT. En definitiva, luego de leer el presente número de la RIII le sugiero que relea el editorial y juzgue usted mismo la potencia de la herramienta, si uno puede convertirse en un interlocutor válido en el proceso.

## **Últimas impresiones**

En la interacción con el ChatGPT hemos descubierto una relativa potencia de resolución de ejercicios en tanto los mismos resulten de aplicación directa. En la resolución de casos de estudio y consignas complejas con ejemplos arroja resultados no convergentes con la respuesta esperada.

La potencia de interpretación y procesamiento de textos, tanto para la comunicación con el interlocutor como para el análisis de datos ingresados es elevada. Es verosímil, aunque no es válido. Las respuestas a la misma pregunta hecha por diferentes usuarios son diferentes, y está influida por el contenido de la charla previa. La validez la aporta el humano que chatea con el bot, y esto ha quedado de manifiesto en las diferentes iteraciones para lograr la respuesta buscada. Sin conocer el tema, posiblemente el usuario podría arribar a conclusiones inconsistentes.

La IA puede tener beneficios para la eficiencia y la productividad en el lugar de trabajo, también plantea dudas legítimas sobre el impacto en el empleo, la justicia y la equidad, y la falta de control humano. Por lo tanto, es importante que las empresas consideren cuidadosamente las implicaciones de la automatización antes de adoptarla de manera generalizada. La posibilidad actual de sustitución del criterio humano frente al criterio de la IA para la toma de decisiones es limitada, aunque para tareas rutinarias puede resultar un buen complemento y herramienta de optimización del trabajador. Aunque la inteligencia artificial ha avanzado rápidamente en los últimos años, todavía hay tareas para las que la tecnología no es adecuada. Algunos argumentan que los trabajos que requieren habilidades humanas únicas, como la creatividad, la empatía y la toma de decisiones complejas, no pueden ser completamente automatizados y limitar siempre la intervención humana.

Se hace necesario llegado este punto, mover a la acción a la comunidad científica para generar las evaluaciones y el corpus empírico que sirva de validación del modelo. Esta validación vendrá sin duda de la producción de trabajos científicos en diversas áreas de aplicación, y es por eso que instamos a la comunidad de lectores que AACINI-RIII a enviar trabajos que busquen general las evidencias para asegurar que el camino del desarrollo que van tomando las IA se corresponde con el sentido de utilidad que busque maximizar la calidad de vida de los humanos.