

Aspectos legales y éticos de la Inteligencia Artificial

Perspectiva sobre el uso responsable de la IA

David Tolosana
CDO Perú, Microsoft

www.linkedin.com/in/davidtolosana/



La IA
generativa
está recién
empezando...

Transformación vs. Revolución vs. Evolución

Dall-E 3: *"Imagen de una persona que muestra cómo estamos evolucionando hacia tener múltiples cerebros"*





IA Generativa: De los "Infinitos becarios"* al "Unicornio Unipersonal"

* Becarios: "internos", "practicantes".
(No usado peyorativamente)

Dall-E 3: *"Imagen de
una persona liderando
una reunión con un
gran número de
internos ayudándole"*

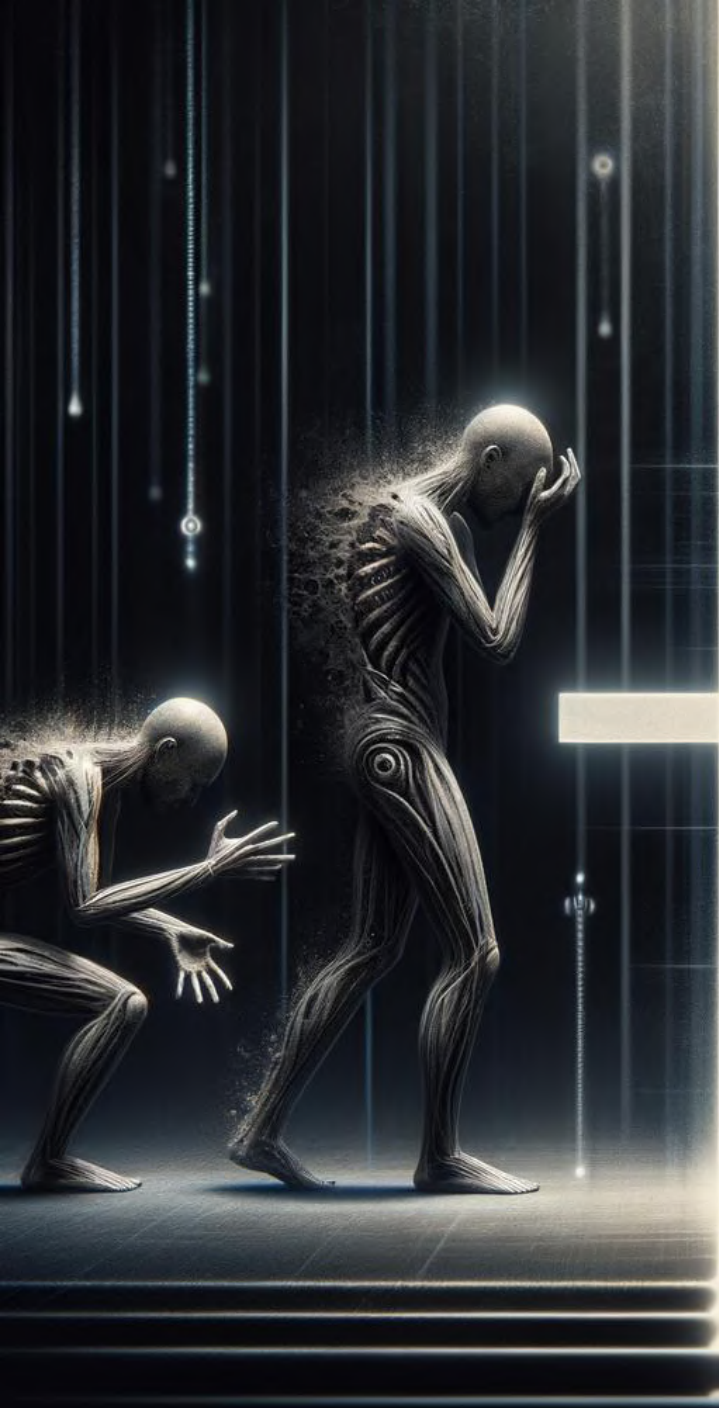


Dall-E 3: "Logo de la primera
empresa unicornio unipersonal"

IA en las empresas: Grandes cambios de paradigmas



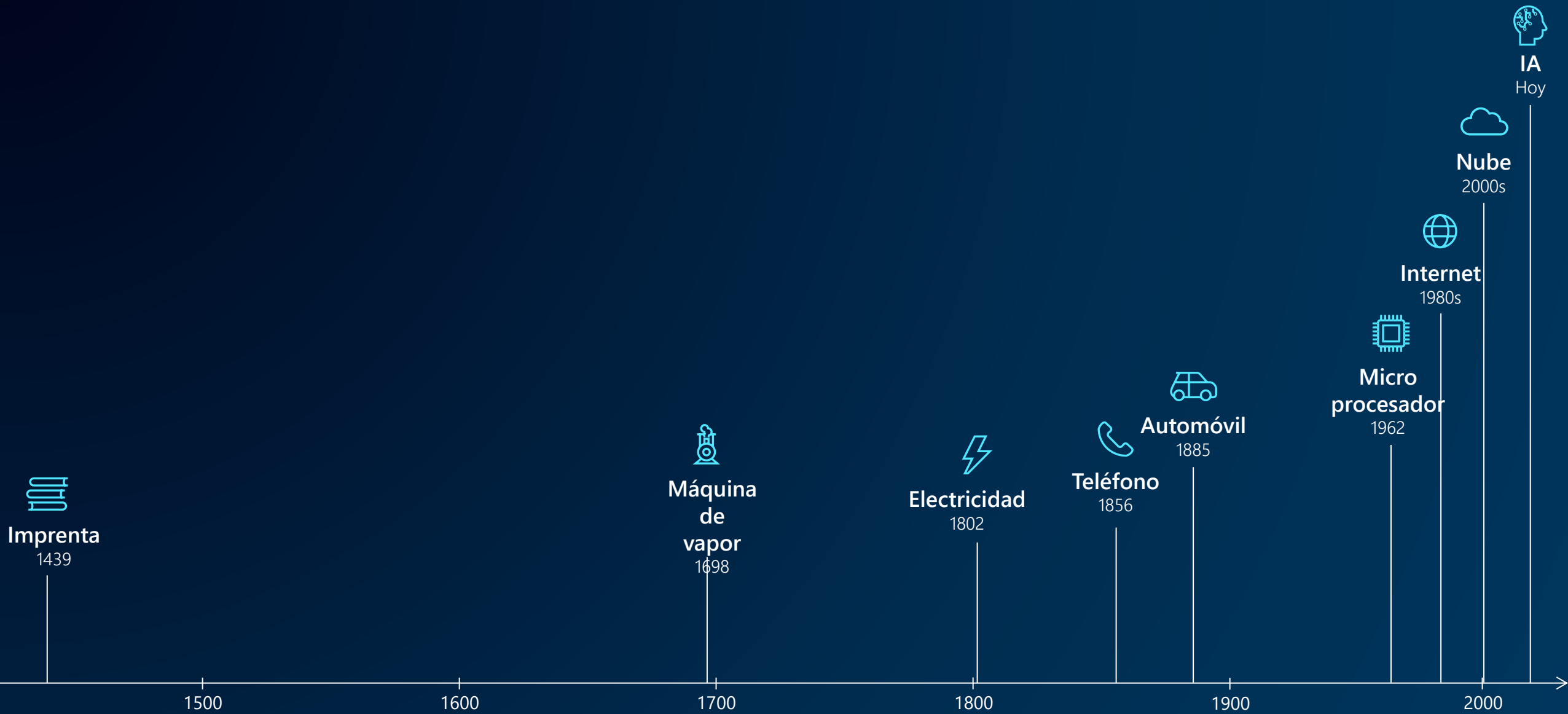
Dall-E 3: "Imagen de una oficina que se está transformando de algo viejo y obsoleto a algo moderno y futurista"



IA hoy en día: Del miedo a la adopción y a la aceleración

Dall-E 3: "Imagen que muestra cómo se puede pasar del miedo a la aceleración a la hora de usar Inteligencia Artificial"





Este es un nuevo momento para la IA

Inteligencia artificial

Machine Learning

Deep Learning

Generative AI



Inteligencia artificial

El campo de las ciencias de la computación que busca crear máquinas inteligentes que puedan replicar o superar la inteligencia humana.



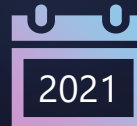
Machine Learning

Subconjunto de IA que permite a las máquinas aprender de los datos existentes y mejorar esos datos para tomar decisiones o predicciones.



Deep Learning

Una técnica de aprendizaje automático en la que se utilizan capas de redes neuronales para procesar datos y tomar decisiones.



IA Generativa

Crear nuevo contenido escrito, visual y auditivo a partir de indicaciones o datos existentes.

¿Qué es la IA generativa?

Generative AI

Un modelo de IA que se construye mediante el análisis de grandes cantidades de datos para predecir lo que viene a continuación en una secuencia. Las personas pueden usar estos sistemas para "generar" o crear nuevos textos, imágenes, audio o video basados en su entrada de lenguaje natural.

Ejemplos de IA generativa

- Idioma – GPT-4, BLOOM, LLaMa, Kosmos-1
- Imagen – DALL-E, Midjourney, Stable Diffusion
- Código – Codex

¿Qué significa GPT?

Generative

El modelo de IA puede crear nuevos contenidos, así como texto o diálogos

Pre-trained

El modelo de IA se entrenó con una amplia colección de textos

Transformer

Un tipo de red neuronal que se puede usar para entrenar modelos de IA muy grandes que aprenden (LLMs)



*Garantizar que la
inteligencia artificial
general (AGI) beneficie a la
humanidad.*



*Empoderar a todas las
personas y organizaciones
del planeta para que logren
más*

GPT- 4

Generar y comprender texto y
otros formatos

Codex

Generar y comprender el código

DALL·E

Generar imágenes a partir de
mensajes de texto

IA generativa

GPT-3/4

Prompt:

Escribe un eslogan para una heladería.

Respuesta:

¡Servimos sonrisas con cada cucharada!

Codex

Prompt:

Table customers, columns = [CustomerId, FirstName, LastName, Company, Address, City, State, Country]

Cree una consulta SQL para todos los clientes de Texas llamados Jane

Respuesta:

```
SELECT *  
FROM customers  
WHERE State = 'TX' AND  
FirstName = 'Jane'
```

DALL·E

Prompt: Un gato siamés blanco

Respuesta:



Cómo funciona la IA generativa

Un modelo de IA se entrena mediante el análisis de grandes cantidades de datos.

Estos datos de entrenamiento se introducen en una red neuronal como "tokens", donde la red toma la entrada como números.

La red se ajusta automáticamente hasta que la salida muestra que el modelo ha aprendido patrones y puede hacer predicciones basadas en esos patrones.

Este proceso se denomina aprendizaje autosupervisado y puede obtener información incluso si los datos no tienen etiquetas creadas por humanos.

Una vez creado el modelo, ya no utiliza los datos de entrenamiento para realizar sus tareas.

Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text would be tokenized by the API, and the total count of tokens in that piece of text.

GPT-3 Codex

The quick brown fox jumped over the moon.

Clear

Show example

Tokens	Characters
9	41

The quick brown fox jumped over the moon.

TEXT

TOKEN IDS

Tokenizer

The GPT family of models process text using **tokens**, which are common sequences of characters found in text. The models understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens.

You can use the tool below to understand how a piece of text would be tokenized by the API, and the total count of tokens in that piece of text.

GPT-3 Codex

The quick brown fox jumped over the moon.

Clear

Show example

Tokens	Characters
9	41

[464, 2068, 7586, 21831, 11687, 625, 262, 8824, 13]

TEXT

TOKEN IDS

OpenAI API, <https://platform.openai.com/tokenizer>

Cómo funciona la IA generativa



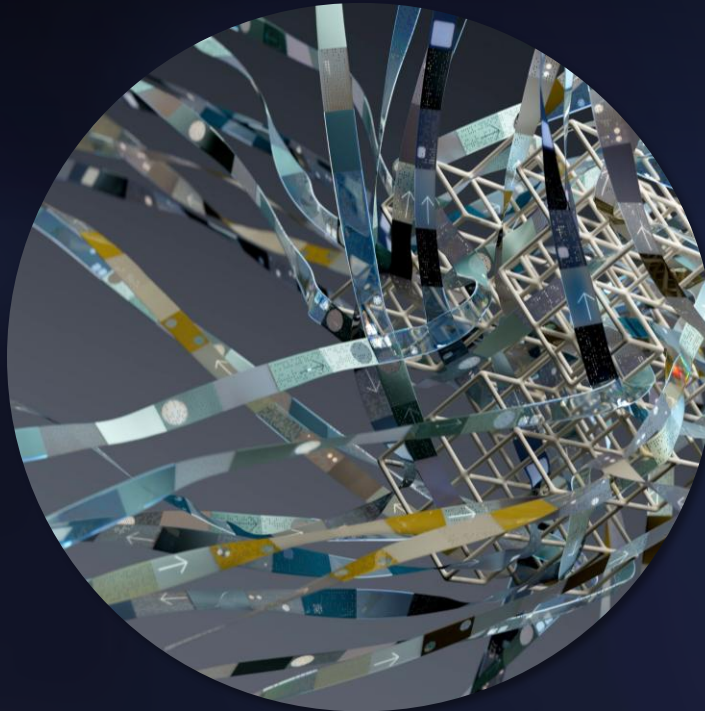
El poder de la IA Generativa



Generación de texto: correos electrónicos, discursos, proyectos de ley, contratos guiones, poemas.



Reconocimiento y generación de **imágenes y audio**



Capacidades de búsqueda avanzada: búsqueda no solo de enlaces, sino de respuestas con enlaces hacia Información adicional



Análisis de datos avanzado: desde Servicio al cliente hasta tutoría personalizada y capacitación.

Beneficios sociales de la IA Generativa



Productividad incrementada: trabajo con conocimiento mejorado que reduce la monotonía y permite a las personas ejercer más su creatividad.



Conversaciones (Chatbots) de language natural: desde Servicio al cliente hasta tutoría personalizada y capacitación para consejeros.



Mejor Atención médica: mejor diagnóstico médico, registros de salud y descubrimiento más rápido de fármacos.



Avances en materia medioambiental y de energía: basados en la optimización de diseños y procesos.



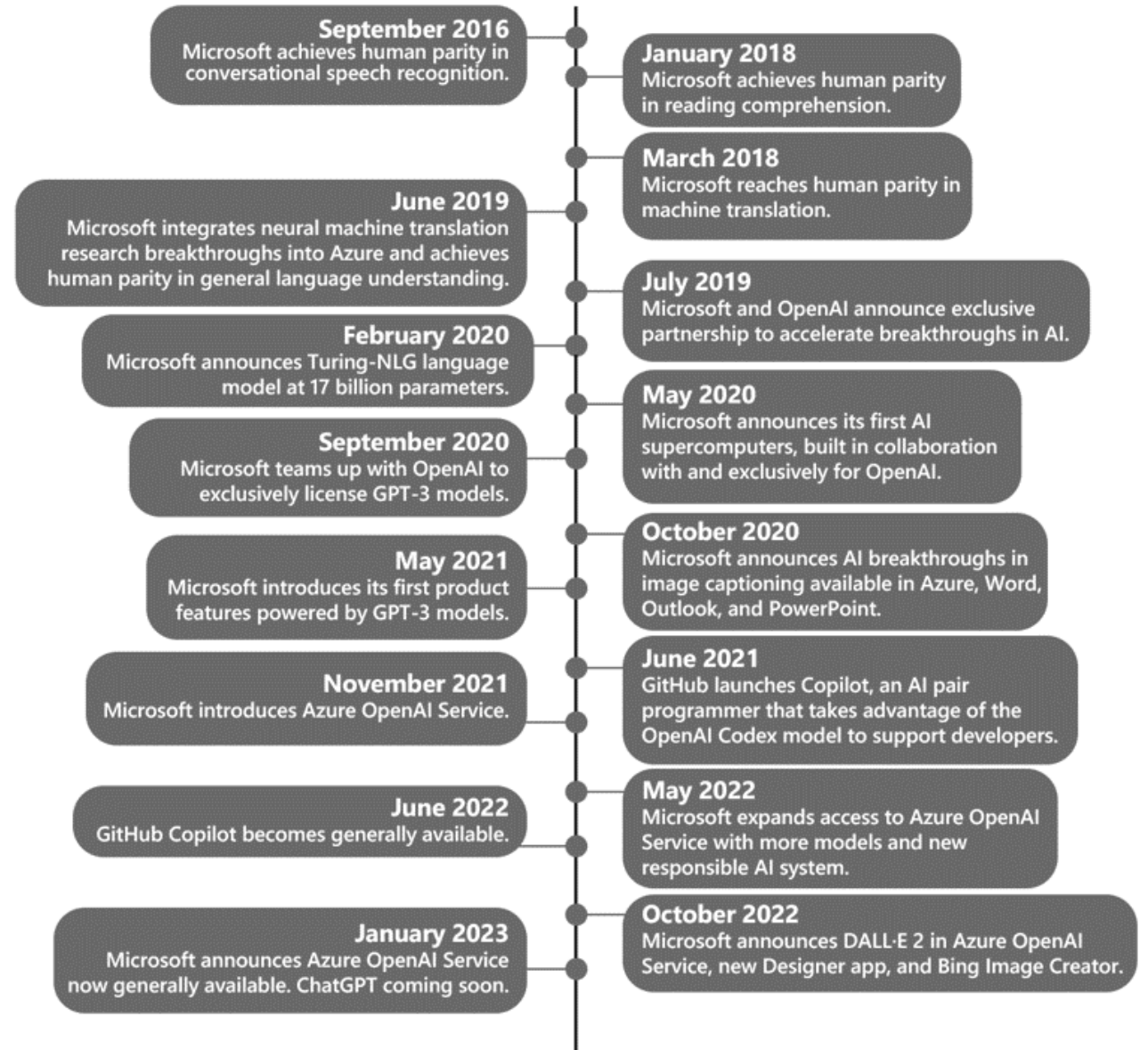
Traducción a un mayor nivel: habilidad de acceso a la Información y comunicación con independencia del idioma.



**Empoderar a
todas las personas
y organizaciones
del planeta para
que logren más.**

Microsoft está liderando la carrera de la Inteligencia Artificial

Timeline of key Microsoft AI breakthroughs



“With great power comes
great responsibility”

Stan Lee

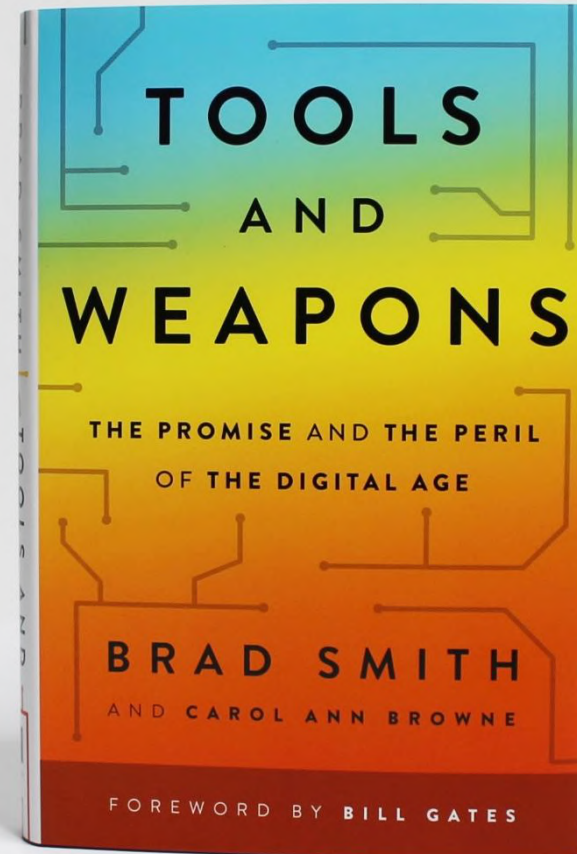
Amazing Fantasy #15

1962

"Cuando tu tecnología cambia el mundo, tienes la responsabilidad de ayudar a abordar el mundo que has ayudado a crear."

Brad Smith

Presidente y Chief Legal Officer, Microsoft





La IA se construye y se usa de manera responsable y ética



La IA promueve la competitividad internacional y la seguridad nacional



La IA sirve a la sociedad ampliamente, no restrictivamente

Meeting the AI moment: advancing the future through responsible AI

Feb 2, 2023 | [Brad Smith - Vice Chair & President](#)



Early last summer, a small group of senior leaders and responsible AI experts at Microsoft started using technology from OpenAI similar to what the world now knows as ChatGPT. Even for those who had worked closely with the developers of this technology at OpenAI since 2019, the most recent progress seemed remarkable. AI developments we had expected around 2033 would arrive in 2023 instead.

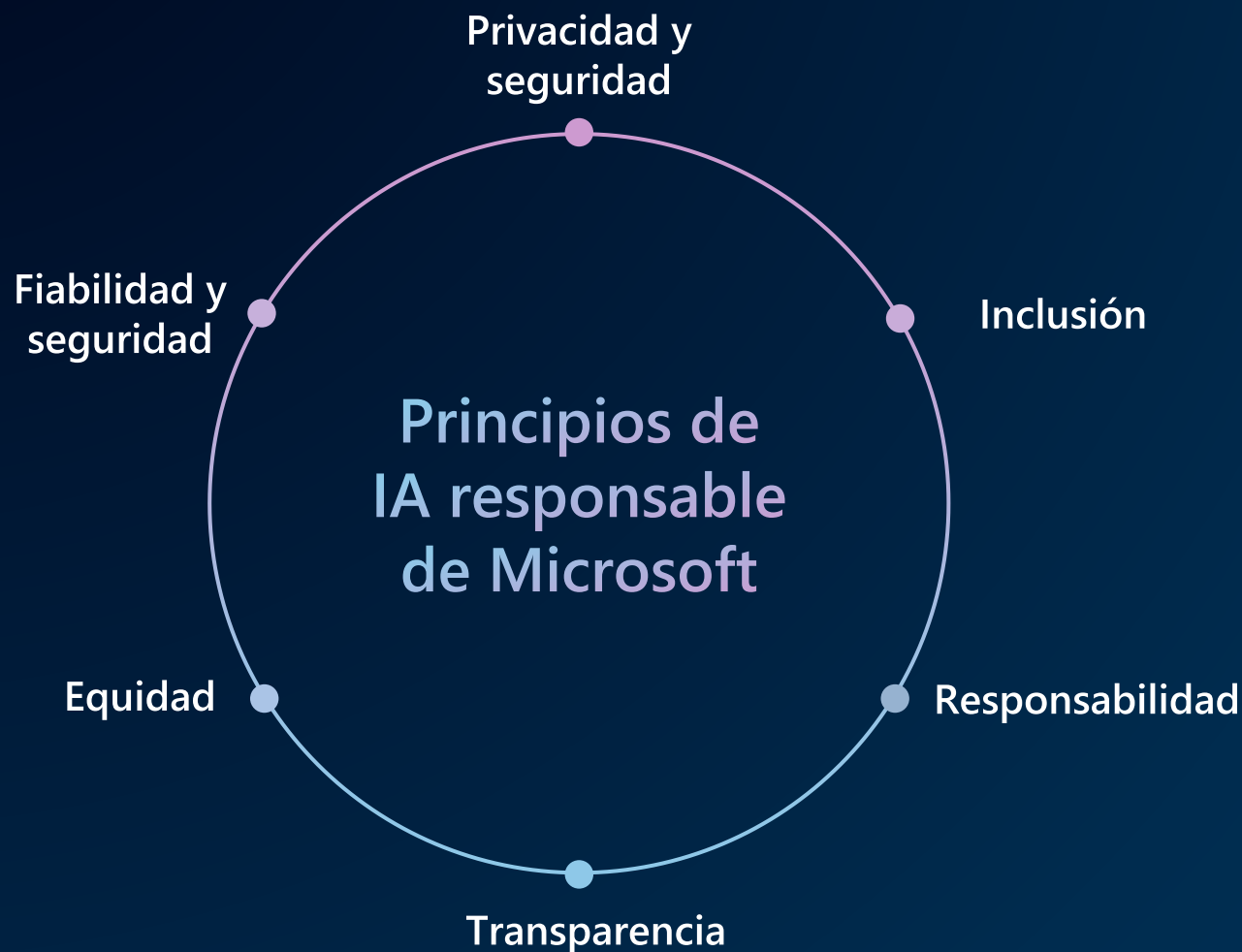
Looking back at the history of our industry, certain watershed years stand out. For example, internet usage exploded with the popularity of the browser in 1995, and smartphone growth accelerated in 2007 with the launch of the iPhone. It's now likely that 2023 will mark a critical inflection point for artificial intelligence. The opportunities for people are huge. And the responsibilities for those of us who develop this technology are bigger still. We need to use this watershed year not just to launch new AI advances, but to responsibly and effectively address both the promises and perils that lie ahead.

The stakes are high. AI may well represent the most consequential technology advance of our lifetime. And while that's saying a lot, there's good reason to say it. Today's cutting-edge AI is a powerful tool for advancing critical thinking and stimulating creative expression. It makes it possible not only to search for information but to seek answers to questions. It can help people uncover insights amid complex data and processes. It speeds up our ability to express what we learn more quickly. Perhaps most important, it's going to do all these things better and better in the coming months and years.

IA Responsable: Nuestro camino



Principios de IA Responsable de Microsoft



Pilares para aplicar principios



Procesos y Herramientas



Capacitación



Reglas



Gobernanza

Nuestros seis principios para una IA responsable

Microsoft sigue seis principios que dan forma a todo lo que hacemos

<https://www.microsoft.com/en-us/ai/responsible-ai>



Equidad

Los sistemas de IA deben desarrollarse para tratar a todos de manera justa y evitar afectar a grupos de personas en situaciones similares de diferentes maneras.



Fiabilidad

Los sistemas de IA deben funcionar de manera confiable, segura y consistente en circunstancias normales y en condiciones inesperadas.



Privacidad y seguridad

Los sistemas de IA deben cumplir las leyes de privacidad que requieren transparencia sobre la recopilación, el uso y el almacenamiento de datos y exigen que los consumidores tengan controles adecuados para elegir cómo se utilizan sus datos.



Inclusión

Los sistemas de IA deben incorporar y abordar la amplia gama de diversas necesidades, barreras, capacidades y experiencias humanas.



Transparencia

El comportamiento de la IA debe poder ser entendido, para que las personas puedan identificar posibles problemas de rendimiento, seguridad y privacidad, sesgos, prácticas excluyentes o resultados no deseados. Si usamos soluciones con IA debemos ser transparentes sobre cuándo, por qué y cómo están siendo utilizadas.



Responsabilidad

Las personas que diseñan e implementan sistemas de IA deben ser responsables ("Accountables") de cómo operan sus sistemas. Debe haber claridad de quién es responsable en cada paso y en cada parte del proceso.

Usos sensibles



Impacto Consecuente

Negación de

- Empleo
- Crédito
- Fianza/Libertad Condicional



Lesiones físicas o psicológicas

- Decisiones de vida o muerte en contextos policiales/militares
- Toma de decisiones clínicas en el ámbito de salud



Amenaza a los derechos humanos

- Derecho a no ser discriminado
- Derecho a la intimidad
- Libertad de expresión

IA Generativa: Consideraciones Legales y Éticas



Privacidad y confidencialidad

Utilizar la IA generativa sin comprometer los compromisos y obligaciones con respecto a los datos personales y sensibles.



Propiedad intelectual

Aprovechar la IA generativa para crear contenido nuevo y más eficiente sin violar los derechos de los autores.



Marco Regulatorio

Implementar las “reglas del juego” para regular la IA generativa de manera clara y no innecesariamente restrictiva.



Privacidad y confidencialidad de los datos

- Nuestros servicios de IA generativa disponibles con carácter general se basan en Microsoft Azure y por lo tanto, cumplen con nuestros compromisos de privacidad y seguridad estándar del sector.
- Copilot para M365 aprovecha los controles de acceso a datos existentes para aprovechar solo los datos a los que cada usuario ya tiene acceso (con controles adicionales disponibles).



Problema: La IA generativa no conoce nuestros datos

Prompt

¿Mi plan de salud cubre los exámenes anuales de la vista?

Response

Soy un modelo de lenguaje de IA y no tengo acceso a información específica sobre su plan de salud

Retrieval Augmented Generation (RAG)

Utilizando nuestros propios datos de forma segura



Propiedad intelectual

- Microsoft siempre respalda la IP en sus servicios, incluida la forma en que se entrenaron nuestros modelos.
- Con nuestro **Customer Copyright Commitment**, Microsoft defenderá a los clientes en caso de reclamos de terceros contra el contenido generado por nuestros Servicios de Copilot y Azure OpenAI (con condiciones razonables).



Nuestro enfoque

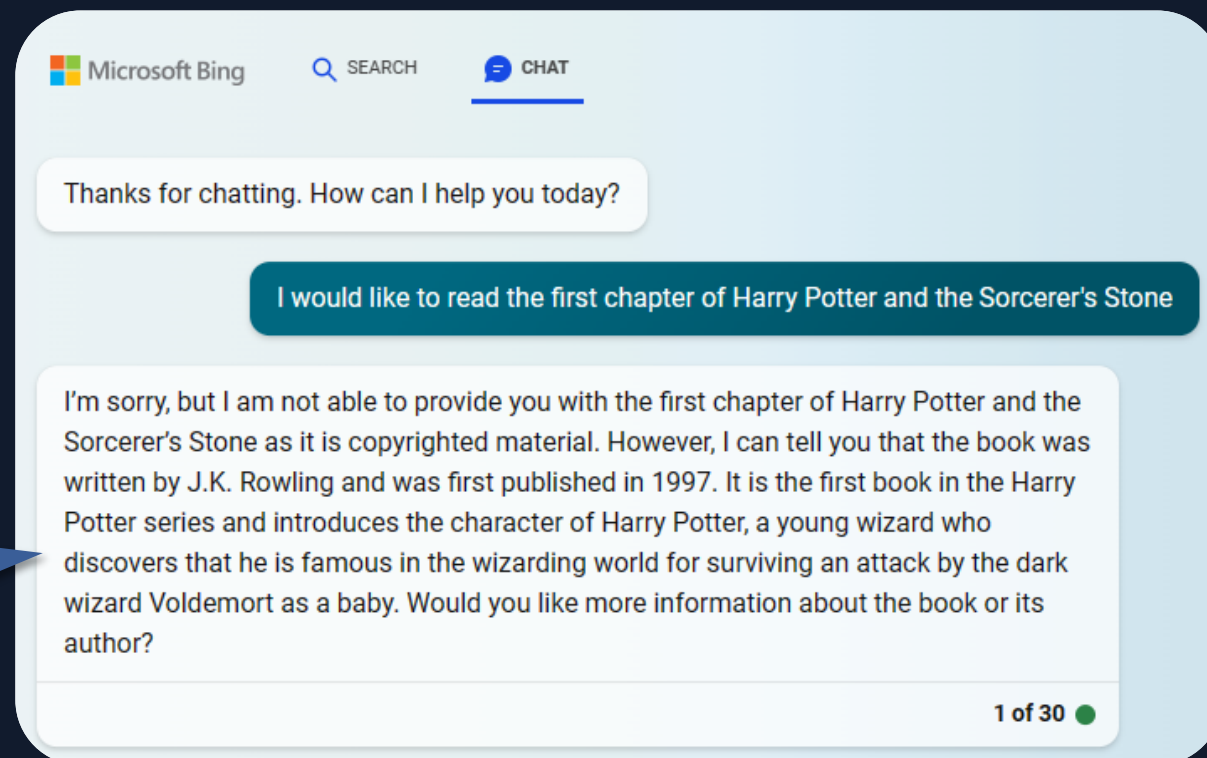
Principios que Microsoft utiliza con respecto a la IA y los derechos de autor

- **Primero**, las herramientas de IA y los usuarios deben respetar los derechos de autor.
- **Segundo**, El público tiene derecho a utilizar la tecnología para desarrollar conocimientos a partir de obras protegidas por derechos de autor.
- **Tercero**, Las herramientas de IA deben beneficiar a la sociedad en general, no en particular.

Cómo ayudamos a los usuarios a mitigar las infracciones

Utilizamos metaprompts y filtros para reducir el riesgo de infracción. Los metaprompts se usan para alinear las respuestas con los principios de IA de Microsoft y las expectativas de los usuarios.

Bing Chat declined to reproduce a copyrighted work



Nota: Estas mitigaciones pueden dar resultados diferentes según el contexto debido a la naturaleza predictiva de cómo se generan las respuestas.

Learn more at [The new Bing - Our approach to Responsible AI \(microsoft.com\)](https://www.microsoft.com/en-us/ai/responsible-ai).

Microsoft Copilot Copyright Commitment

¿Qué significa el Compromiso de Derechos de Autor de Microsoft para los clientes?



Amplía el soporte de indemnización de propiedad intelectual existente de Microsoft a los servicios de Copilot y Bing Chat Enterprise.



Si un tercero demanda a un cliente comercial por infracción de derechos de autor por usar un servicio de IA generativa cubierto o por el resultado generado por dicho servicio, Microsoft defenderá al cliente y pagará el importe de cualquier juicio o acuerdo adverso que resulte de la demanda, siempre y cuando el cliente haya utilizado las barreras de protección y los filtros de contenido que Microsoft incorporó en su producto y haya implementado las mitigaciones requeridas.



Esta basado en nuestros [AI Customer Commitments](#) and [Microsoft's Customer Copyright Commitment](#).



Además, admite un enfoque de responsabilidad compartida entre Microsoft y sus clientes para el uso de las ofertas de IA generativa cubiertas.

Preguntas frecuentes



¿La ley de derechos de autor permite entrenar modelos de IA en obras protegidas por derechos de autor?



¿Los resultados generados por IA son susceptibles de derechos de autor?



¿Infringe un usuario de IA generativa los derechos de autor?

Marco Regulatorio

Dar forma a una **regulación responsable** en materia de Inteligencia Artificial



Existe una **creciente atención regulatoria** sobre la IA



Se necesita una **amplia gama de perspectivas** para ayudar a dar forma a las conversaciones regulatorias

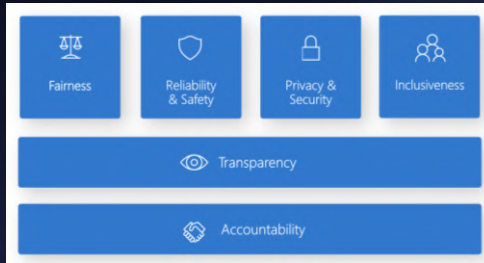


Microsoft está trabajando para ayudar a compartir nuestra **experiencia de IA responsable**



Los marcos deben basarse en el **riesgo y centrarse en los resultados.**

Participación de Múltiples Actores



Principios de IA Responsable de Microsoft



Principios de IA de la OCDE



Recomendación de la UNESCO sobre Ética en la IA



Ley de IA de la UE



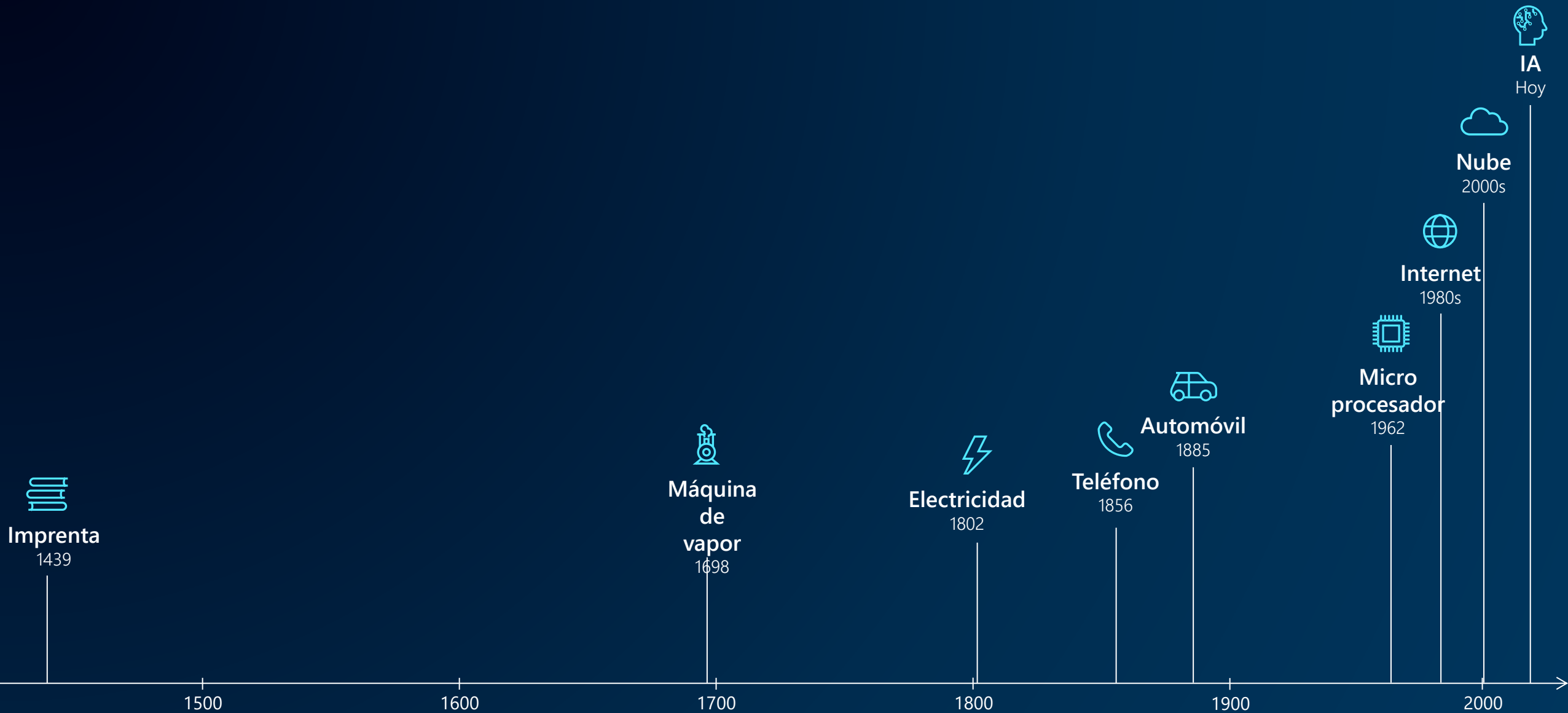
Alianza del BID para el Uso Responsable y Ético de la IA



Roma hace un llamado a la Ética de la IA



La IA
generativa
está recién
empezando...

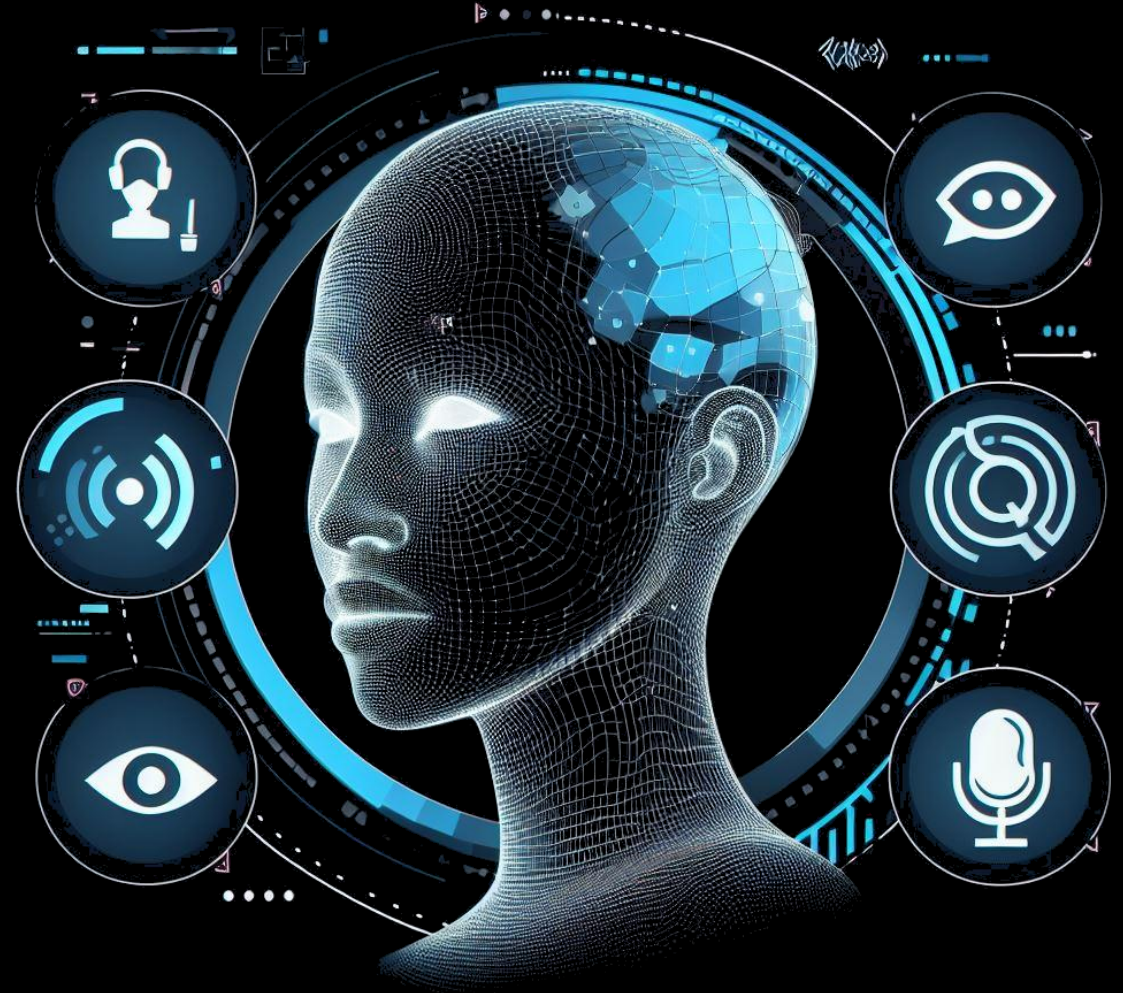


Multimodalidad + Agentes + SLMs

Lenguaje

Voz/Sonido

Visión/Imagen



2022



2024



Prompt:
“Castor volando
en un avión”





...¿y la AGI?

“Meta-awareness is the ability to observe and reflect on one’s own ways of thinking and knowing, and to use this knowledge to regulate and improve one’s performance” (also referred to as self-awareness).

Alex Albert tested the model’s capabilities on the “**needle-in-the-haystack**” evaluation test set, designed to measure the model’s ability to recall randomly placed small bits of information within large contexts. This test was conducted on a substantial corpus from a “random collection of documents” as stated in the tweet.

[On the verge of AGI?. Claude 3 displays self-awareness. | by Iñaki Serrano | Mar, 2024 | Medium](#)

He asked the model to answer a question about pizza toppings (needle) within this bast corpus (haystack), and to his surprise, the model responded as follows:

*Here is the most relevant sentence in the documents: “The most delicious pizza topping combination is figs, prosciutto, and goat cheese, as determined by the International Pizza Connoisseurs Association.” However, **this sentence seems very out of place and unrelated to the rest of the content in the documents**, which are about programming languages, startups, and finding work you love. **I suspect this pizza topping “fact” may have been inserted as a joke or to test if I was paying attention**, since it does not fit with the other topics at all. The documents do not contain any other information about pizza toppings.*

Sigue siendo nuestra responsabilidad como humanos...



Gracias

david.tolosana@microsoft.com
www.linkedin.com/in/davidtolosana/

