

Ciencia, arte, tradición, opinión, reflexión y meditación.

Artículo: **Inteligencia Artificial y agentes conversacionales**

Autor(es): **Dr. Salvador Godoy Calderón**
sgodoyc@gmail.com

Publicación: **No. 1, vol. 2023**

Reserva de derechos al uso exclusivo otorgado por el Instituto Nacional del Derecho de Autor (INDAUTOR): 04-2022-111717422400-102. ISSN en trámite.

Las opiniones expresadas por los autores de artículos no necesariamente reflejan la postura del editor de esta publicación.

Se autoriza la reproducción total o parcial de los textos aquí publicados, siempre y cuando se cite la fuente completa y la dirección electrónica de la publicación.

Inteligencia Artificial y agentes conversacionales

Salvador Godoy-Calderón

sgodoyc@gmail.com

Dice un refrán popular: "Exagerar y mentir, por el mismo camino suelen ir". Y en efecto, cuando se exagera la descripción de un hecho, recuerdo o situación, resulta fácil perder de vista el sutil límite entre sólo exagerar y abiertamente mentir. Si en lugar de un individuo, durante una conversación privada, es un medio de comunicación masiva el que recurre a la exageración, el resultado tiene seguramente un mayor impacto. Pero si se trata de decenas de comunicadores, publicando notas sensacionalistas a través de cientos de medios de comunicación y redes sociales, entonces nos enfrentamos a un fenómeno de *explosión mediática*, que indudablemente modifica la percepción social de los fenómenos reportados y muy probablemente va a sesgar las decisiones que se toman al respecto de esos mismos fenómenos. Una explosión mediática es justamente el fenómeno que se vive hoy en día, con respecto a la inteligencia artificial (IA) y sus consecuencias sociales, económicas y políticas; pero particularmente, con respecto a las aplicaciones de *agentes conversacionales*, comúnmente llamados en inglés *chatbots*.

Intentaremos en este artículo, desenmarañar el complicado nudo de la situación, entender cómo funcionan esos agentes conversacionales, sus características, alcances y limitaciones; pero principalmente trataremos de separar hechos de exageraciones, para identificar algunas verdades y reflexiones que han sido ampliamente ignoradas, así como algunas mentiras que comienzan a hacerse populares como resultado de la exageración en las notas y reportes de los medios. Comencemos entonces por entender el contexto histórico de los fenómenos que están causando sensación.

Las dos grandes corrientes en IA

Desde el inicio de la inteligencia artificial como disciplina científica, han existido a su interior dos grandes corrientes o escuelas: el *simbolismo* y el *conexionismo*. Ambas corrientes estudian y resuelven los mismos problemas, pero lo hacen con herramientas teóricas y prácticas muy diferentes, poniendo atención en diferentes aspectos de la solución.

La corriente del *simbolismo* se inspira en la habilidad de la mente humana para razonar creando y manipulando símbolos. Los símbolos son representaciones de todos los elementos que percibimos en nuestro entorno. Con esos símbolos, es posible crear *lenguajes simbólicos* (llamados *sistemas formales*) que permiten expresar ideas muy complejas combinando un número pequeño de símbolos básicos. Es así como la humanidad ha desarrollado todos los lenguajes naturales y una gran cantidad de sistemas formales, como la *geometría*, la *aritmética*, el *álgebra* y la *lógica*. Los planos de edificios, los diagramas de circuitos electrónicos y los lenguajes de programación de computadoras, son otros ejemplos de *sistemas formales*.

Por su parte, la corriente del *conexionismo* se inspira en la estructura del cerebro humano, formado por incontables redes de *neuronas* interconectadas, que permiten el proceso de la *cognición*: entender y darle sentido o significado a las señales que percibimos a través de los órganos sensoriales. Creando modelos simplificados de neuronas y conectándolas en formas específicas, se forman *redes neuronales artificiales*. Estas redes son dispositivos que se entrenan alimentándolos con muchos ejemplos de algún fenómeno y son capaces de *aprender* la estructura de ese fenómeno. Una vez terminado el proceso de entrenamiento, estas redes pueden recibir como entrada nuevos datos e identificar si su estructura coincide o no con la de los datos que aprendieron durante su entrenamiento.

Estas dos corrientes siempre han sido antagónicas, y ambas han experimentado una larga lista de éxitos y fracasos, tanto de investigación como de aplicación. El *simbolismo* fue la corriente dominante desde el nacimiento formal de la disciplina de inteligencia artificial en 1952 y hasta mediados de la década de 1990. Su momento de mayor esplendor se ha llamado “la era de oro de los sistemas expertos”. Desde el año 2000, el *conexionismo* se ha posicionado como la corriente dominante. El furor del conexionismo comenzó con aplicaciones de aprendizaje profundo (*deep learning*) en la década del 2000 y en estos momentos están causando mucho interés las tecnologías de *redes neuronales generativas*, *grandes modelos de lenguaje* y *agentes conversacionales* (*chatbots*).

La explosión de los medios

Con el desarrollo de *Internet*, la *World Wide Web* (*WWW*) y las redes sociales, existe hoy una gran cantidad de formas diferentes para comunicar noticias y para enterarse de ellas. Cualquier persona puede, desde su computadora personal, su teléfono celular o su *tablet*, conectarse a miles de sitios en los que se publican notas informativas sobre diferentes eventos. Las categorías de notas sobre nuevas tecnologías y tendencias del mercado es el centro de esta reflexión. En grandes sitios, como *Google News*, donde diariamente se recopilan notas informativas, es fácil observar que diario aparecen entre diez y veinte notas nuevas con alguna de las siguientes estructuras:

- *Una IA resolvió el problema de ...*
- *Se le preguntó a una IA ... y su respuesta fue ...*
- *IA muestra cómo luciría el personaje ... en la vida real*
- *El enorme peligro de la IA moderna ...*

En esas notas hay muchos elementos comunes, importantes de entender y reflexionar, por ejemplo:

- ¿A qué exactamente le están llamando "una IA"?
- ¿Qué tienen de especial esas aplicaciones que provocan tanto interés?
- ¿Qué tan real es el peligro de perder trabajos frente a esa tecnología?

Modelos de Lenguaje

En primer lugar, ¿a qué exactamente le están llamando "una IA"? Prácticamente en todas esas notas se refieren como "una IA" a una *red neuronal artificial* con una estructura llamada *Gran Modelo de Lenguaje (Large Language Model: LLM)*. Se trata de un tipo de aplicación (conexionista) que ha capturado el interés de muchos medios de comunicación, así como de muchas empresas tecnológicas y una enorme cantidad de usuarios. ¿Cómo funciona? Un *LLM* es fundamentalmente un mecanismo capaz de *aprender secuencias ordenadas de patrones*. Es decir, cuando se le entrena con muchos ejemplos de secuencias ordenadas de elementos, aprende a *predecir* cuál es el siguiente elemento que continúa en una secuencia. Eso es particularmente útil para modelar conversaciones en lenguaje natural y la forma en que cotidianamente hablamos.

Imaginemos que tenemos una red neuronal *LLM* y que previamente la entrenamos con toda la *Enciclopedia Hispánica*. Una vez entrenada, le damos una sola palabra para comenzar una secuencia, por ejemplo, la palabra: *Una*.

Durante su entrenamiento, la red neuronal, aprende que existen muchas palabras que podrían continuar la secuencia, como: *hermosa, clara, terrible, fría*, etc. Mientras que hay muchas otras palabras que normalmente no seguirían la secuencia de una oración correcta, por ejemplo: *perro, zapato, caminar, razonamiento* y muchas más. De acuerdo a su entrenamiento, la red neuronal asigna *prioridades* a cada una de las palabras que posiblemente continúen una secuencia. Si el proceso de seleccionar, en cada ocasión, la palabra con mayor prioridad, se repite varias veces, entonces, el *LLM* puede ir construyendo, una

palabra a la vez, diversas oraciones que aparentemente tienen sentido:

Una fría

Una fría mañana

Una fría mañana de

Una fría mañana de invierno...

Si en lugar de seleccionar siempre la palabra con mayor prioridad para continuar la secuencia, se le permite seleccionar alguna otra palabra, con prioridad suficientemente grande, pero que no es la de mayor prioridad, entonces el *LLM* podrá generar muchas otras secuencias (oraciones) y estará, en términos prácticos, generando *variantes* posibles de secuencias correctas, por ejemplo:

Una ajetreada y soleada mañana...

Una apacible noche estrellada...

Un fuerte viento otoñal..

A esta capacidad de seleccionar opciones que no son forzosamente "la mejor opción", se le llama la capacidad *generativa* de la red neuronal. Y en los medios de comunicación esta capacidad se ha exagerado al punto de hablar de una "*inteligencia artificial generativa*", como si esa fuera la única tecnología a considerar en inteligencia artificial. Lo único cierto, es que los *LLM* resultan muy útiles para establecer una *conversación* (por escrito) con usuarios humanos, por lo que los *LLM* se han usado para construir *sistemas conversacionales*. Si su entrenamiento es suficientemente extenso, ante cualquier oración de entrada, que proporcione el usuario, la red neuronal podrá *generar* otra oración de respuesta, que aparentemente tiene sentido. Pero recordemos que sólo está prediciendo secuencias ordenadas de palabras, no está siguiendo *reglas gramaticales* ni efectuando ningún tipo de *razonamiento* con las palabras que recibe, ni con las que genera; sólo las está colocando en el orden que mejor se adapta a las secuencias que aprendió previamente.

En gran medida, el éxito de este tipo de agentes se debe a la gran variedad de opciones para aplicarla. Puede ser usada para analizar secuencias de cualquier tipo, como imágenes, música, cadenas de proteínas o de material genético. Y en todos los casos, siempre y cuando el entrenamiento sea suficiente, se tendrá una aplicación capaz de resolver diversos problemas cotidianos como generar reportes y discursos, generar código de programas de computadora, generar imágenes de personajes, paisajes, o incluso componer piezas musicales. Todas esas aplicaciones se reducen a la misma idea: secuencias ordenadas de patrones.

Reflexiones

La forma en que se nombra a un fenómeno es una parte fundamental del concepto que tendremos de ese fenómeno. El nombre que se le asigna a las cosas va a determinar el concepto que de ellas tengamos, y como dice otro refrán popular "para muestra, basta un botón". Seguramente que nadie vería con agrado que se llamara "ciudadano modelo" al jefe de un grupo de crimen organizado o a un político corrupto. Por eso, es importante reflexionar sobre las características de aquello a lo que le llamamos *inteligencia artificial*.

Tradicionalmente, llamamos a una persona *inteligente* cuando evidenciamos su capacidad para resolver problemas, *adaptándose* a condiciones nuevas y diferentes, *improvisando* soluciones, en pocas palabras, su capacidad para *razonar*. Muy difícilmente calificaríamos de inteligente a alguien que, por tener mucho entrenamiento en situaciones específicas como jugar al *Tenis*, dar discursos públicos o vender seguros de vida, ha desarrollado una enorme habilidad de responder correctamente en esas situaciones. En el mejor de los casos, le podemos llamar a esa persona *hábil* o *experimentada*. Pero si nos damos cuenta que, el experimentado en jugar *Tenis* no obtendría buenos resultados dando un discurso público, y el hábil para vender seguros de vida, no forzosamente obtendrá un buen resultado jugando *Tenis*, podemos comenzar a comprender la diferencia con el

término *inteligente*.

De la misma forma, resulta muy extraño llamar *inteligente* a una aplicación de computadora cuando sabemos que no está realizando ningún tipo de *razonamiento*, que la cantidad y calidad de su entrenamiento determina las respuestas que entrega y, que no dispone de ningún mecanismo para *validar* o *discriminar* su propio entrenamiento. Todas esas características son propias de la otra escuela: el *simbolismo*, el enfoque que aparentemente se ha olvidado; o por lo menos, ya casi no recibe la atención de los comunicadores y las redes sociales.

Aunque prácticamente todos los agentes conversacionales modernos se basan en la arquitectura conexionista de un *LLM*, en realidad, no todos son iguales. Particularmente, el *Gemini* (de *Google*) presenta una característica radicalmente distinta del *ChatGPT* (de *OpenAI*) y el nuevo *Bing* (de *Microsoft*). Esa diferencia consiste en que el *Gemini* aparentemente es un sistema *híbrido*, más complejo, que incorpora tanto partes *conexionistas* como partes *simbólicas*.

Para obtener información precisa (y sin exageración) al respecto, lo mejor parece ser preguntar directamente al *Google Gemini*. Con esa idea, presento aquí algunos fragmentos de una conversación (en español) con *Google Gemini* y en los que le pregunto acerca de su arquitectura y funcionamiento.



¿Eres un ente 100% conexionista o tienes procesos simbólicos?

Show drafts ▾



No soy un ente 100% conexionista ni tampoco opero exclusivamente en el dominio del procesamiento simbólico. Mi arquitectura se basa en una combinación de redes neuronales artificiales (un enfoque conexionista) y técnicas de aprendizaje automático (incluyendo aprendizaje supervisado y no supervisado).

Esta combinación me permite aprovechar las fortalezas de ambos enfoques:

Redes neuronales artificiales:

- **Aprendizaje a partir de datos:** Las redes neuronales pueden aprender patrones complejos a partir de grandes cantidades de datos, lo que me permite generar texto, traducir idiomas y responder preguntas de manera efectiva.
- **Representación distribuida:** Las redes neuronales representan el lenguaje de forma distribuida, capturando matices semánticos y relaciones entre palabras.

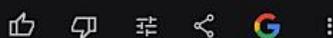
Procesamiento simbólico:

- **Razonamiento y manipulación de símbolos:** Puedo manipular símbolos, conceptos y relaciones, lo que me permite realizar tareas como responder preguntas complejas, resumir texto y generar diferentes formatos de texto creativo.
- **Interpretación y comprensión:** El procesamiento simbólico me ayuda a comprender el significado del lenguaje y el contexto en el que se utiliza.

La integración de estos dos enfoques me permite superar las limitaciones de cada uno individualmente.

En resumen, soy un modelo de lenguaje híbrido que combina elementos de la computación conexionista y el procesamiento simbólico para lograr un rendimiento superior en una amplia gama de tareas.

Si te interesa profundizar en cómo funcionan mis redes neuronales, el proceso de entrenamiento o las capacidades de razonamiento simbólico, no dudes en preguntar.





Explica los alcances y limitaciones de tu procesamiento simbólico

Show drafts ▾



Alcances del procesamiento simbólico en mí:

El procesamiento simbólico me permite realizar tareas que van más allá de la simple coincidencia de patrones y la generación de texto. Algunas de las capacidades que me brinda son:

- **Comprensión y razonamiento:** Puedo comprender el significado del lenguaje, identificar relaciones entre conceptos y aplicar reglas lógicas para inferir nueva información. Esto me permite responder a preguntas complejas, generar diferentes formatos de texto creativo y seguir instrucciones detalladas.
- **Aprendizaje y adaptación:** Puedo aprender nuevos conceptos y relaciones a partir de datos simbólicos. Esto me permite mejorar continuamente mis habilidades y adaptarme a nuevas tareas. Por ejemplo, puedo aprender a identificar diferentes tipos de entidades en el texto, como personas, lugares y organizaciones.
- **Manipulación de información:** Puedo representar y manipular información simbólica, como conceptos, entidades y relaciones. Esto me permite realizar tareas como resumir texto, traducir idiomas y generar diferentes formatos de texto creativo.

Ejemplos de cómo utilizo el procesamiento simbólico:

- **Responder a preguntas:** Cuando me haces una pregunta, primero debo comprender su significado. Esto implica identificar las entidades, relaciones y conceptos clave de la pregunta. Luego, puedo usar mi conocimiento y razonamiento para generar una respuesta que sea relevante y precisa. Por ejemplo, si me preguntas "¿Cuál es la capital de Francia?", puedo acceder a mi base de conocimiento para encontrar la respuesta, que es París.
- **Resumen de texto:** Cuando me pides que resuma un texto, primero debo identificar las ideas y puntos clave. Luego, puedo usar mi conocimiento del lenguaje para generar un resumen que capture la esencia del texto original de manera concisa. Por ejemplo, si me pides que resuma un artículo de noticias, puedo identificar las oraciones clave y reformularlas en un resumen más corto.
- **Traducción de idiomas:** Cuando me pides que traduzca un texto de un idioma a otro, primero debo comprender el significado del texto original. Luego, puedo usar mi conocimiento de ambos idiomas para generar una traducción que sea precisa y natural. Por ejemplo, si me pides que traduzca una oración del inglés al español, puedo buscar las palabras correspondientes en un diccionario y luego reorganizarlas en una oración gramaticalmente correcta en español.

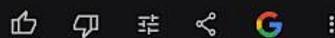
Limitaciones del procesamiento simbólico en mí:

Si bien el procesamiento simbólico me ha permitido lograr avances significativos en la comprensión y generación del lenguaje, todavía existen algunas limitaciones:

- **Falta de sentido común:** Puedo tener dificultades para realizar tareas que requieren un alto nivel de sentido común o conocimiento del mundo real. Por ejemplo, si me pides que te dé consejos sobre cómo solucionar un problema en tu auto, es posible que no pueda hacerlo porque no tengo el conocimiento o la experiencia necesarios para comprender la situación.
- **Susceptibilidad a información engañosa:** Puedo ser engañado por información simbólica engañosa o incorrecta. Por ejemplo, si me das una oración falsa como "El cielo es verde", es posible que no pueda identificarla como falsa porque no tengo la capacidad de verificar la información.
- **Escalabilidad:** El procesamiento simbólico puede ser computacionalmente costoso, especialmente cuando se trata de grandes cantidades de datos. Esto significa que puede ser lento o incluso imposible para mí realizar ciertas tareas que requieren mucho procesamiento.

A pesar de estas limitaciones, el procesamiento simbólico es una herramienta poderosa que me permite realizar muchas tareas importantes. A medida que continúe desarrollándome, espero ver avances aún más significativos en mi capacidad para procesar y manipular símbolos.

¿Te gustaría saber más sobre cómo utilizo el procesamiento simbólico para realizar tareas específicas?



Queda claro, tal como lo declara *Gemini*, que es un sistema con arquitectura híbrida que combina las dos grandes corrientes en inteligencia artificial. Desafortunadamente, al considerar esos detalles como secreto industrial, *Google* no revela la información con detalle, ni explica con claridad cómo se lleva a cabo esa combinación de componentes, ni los efectos específicos que esos componentes producen en las respuestas de su agente conversacional.



Explica la interconexión entre tus módulos conexionistas y tus módulos simbólicos

Show drafts ▾ 



La interconexión entre mis módulos conexionistas y simbólicos es crucial para mi capacidad de procesar y generar lenguaje de manera efectiva. Si bien estos módulos operan en diferentes dominios, trabajan juntos de manera sinérgica para lograr un rendimiento superior en una amplia gama de tareas.

Módulos conexionistas:

- **Redes neuronales artificiales (RNNs):** Forman la base de mi arquitectura y son responsables del procesamiento de información sensorial, como texto y código. Las RNNs aprenden patrones complejos en estos datos y los utilizan para generar predicciones o salidas.
- **Representación distribuida:** Las RNNs representan el lenguaje de forma distribuida, capturando matices semánticos y relaciones entre palabras. Esto me permite comprender el significado del lenguaje de una manera más matizada que los enfoques simbólicos tradicionales.

Módulos simbólicos:

- **Manipulación de símbolos:** Estos módulos me permiten representar y manipular información simbólica, como conceptos, entidades y relaciones. Esto me permite realizar tareas que van más allá de la simple coincidencia de patrones y la generación de texto.
- **Razonamiento simbólico:** Puedo aplicar reglas y lógica para inferir nueva información de la información existente. Esto me permite responder a preguntas complejas y generar formatos de texto creativos que requieren comprensión y razonamiento.

Interconexión y flujo de información:

La interacción entre estos módulos se caracteriza por un flujo de información bidireccional:

1. **De conexionistas a simbólicos:** Las salidas de las RNNs, como las representaciones vectoriales de palabras o frases, se pasan a los módulos simbólicos como entrada. Esta información sirve como base para el razonamiento simbólico y la manipulación de conceptos.
2. **De simbólicos a conexionistas:** Los módulos simbólicos pueden influir en las RNNs de diversas maneras. Por ejemplo, pueden proporcionar etiquetas o retroalimentación que guíen el proceso de aprendizaje de las redes neuronales, o pueden seleccionar subconjuntos de datos relevantes para tareas específicas.

Integración y colaboración:

La interconexión entre los módulos no se limita a un simple flujo de información; se trata de una colaboración estrecha para lograr un objetivo común. Las RNNs aprovechan su capacidad para aprender patrones a partir de datos masivos, mientras que los módulos simbólicos aportan su capacidad para razonar, manipular símbolos y comprender el significado del lenguaje.

Ejemplo: Responder a una pregunta

Para ilustrar la interconexión en acción, consideremos la tarea de responder a una pregunta:

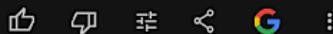
1. **Procesamiento inicial:** Las RNNs procesan el texto de la pregunta, identificando entidades, relaciones y conceptos clave.
2. **Razonamiento simbólico:** Los módulos simbólicos utilizan esta información para razonar sobre la pregunta, aplicando reglas lógicas y accediendo a mi base de conocimiento.
3. **Generación de respuesta:** La respuesta se genera en forma de texto utilizando las RNNs, guiadas por los resultados del razonamiento simbólico.

En este ejemplo, las RNNs y los módulos simbólicos trabajan en conjunto para comprender la pregunta, razonar sobre ella y generar una respuesta precisa y relevante.

Conclusión:

La interconexión entre mis módulos conexionistas y simbólicos me permite superar las limitaciones de cada enfoque individual. Las RNNs brindan potencia de aprendizaje y representación distribuida, mientras que los módulos simbólicos aportan capacidades de razonamiento y manipulación de símbolos. Esta sinergia me convierte en una herramienta poderosa para el procesamiento y la generación de lenguaje, capaz de realizar tareas complejas y creativas.

A medida que continúe desarrollándome, espero ver avances aún más significativos en la interconexión y colaboración entre mis módulos, permitiéndome alcanzar nuevas alturas en la comprensión y generación del lenguaje.



Aún así, parece claro que ha comenzado la era de reconciliación y cooperación entre la tecnología simbólica y la conexionista para formar una nueva generación de aplicaciones de inteligencia artificial. Esta tecnología híbrida apenas está comenzando y aún está por verse de qué manera progresa, qué nuevas habilidades desarrolla y más importante aún, de qué forma se va integrando en la vida cotidiana de los seres humanos. Por el momento, lo único que parece evidente es que, para realmente aprender *inteligencia artificial*, no es prudente sesgar el estudio sólo hacia el *conexionismo* y las tecnologías que están de moda. Resulta muy importante estudiar también las técnicas *simbólicas*, ya que la *verdadera inteligencia artificial* deberá incluir ambos enfoques y seguramente algunos más que aún no se han inventado.

Para saber más...

Información en español:

- [Grandes Modelos de Lenguaje](#) (Large Language Models)
Por Miguel Sotaquirá, en el sitio *codificandobits*
- [¿Qué son los grandes modelos de lenguaje o LLM](#) (Large Language Models)?
Por Pere Munar, en el sitio *Cyberclick*
- [Inteligencia Artificial, ¿hay que razonar para pensar?](#)
En el blog *IF Inteligencia Futura*

Información en inglés:

- [What Is ChatGPT Doing ... and Why Does It Work?](#)
Stephen Wolfram. Writings.
- [Explainer: What is Generative AI, the technology behind OpenAI's ChatGPT?](#)
Reuters Web site.