

# *Disrupción*

en los ecosistemas de comunicación:  
representaciones, sesgos y nuevas  
dinámicas de consumo ante la  
Inteligencia Artificial (IA)

Antony Flores Mérida

*"...todo lo que escribimos o decimos sobre la IA ya es viejo."*

*Carlos Alberto Scolari, en una entrada de blog en mayo de 2023*

Las herramientas de inteligencia artificial sugieren una nueva forma de relacionarnos con las tecnologías digitales de comunicación interactiva, sin embargo, requerimos identificar los sesgos en las representaciones (o en las interpretaciones de la realidad) que nos ofrecen. En este texto se presentan un par de ejercicios y reflexiones sobre los sesgos algorítmicos que las IA pueden estar reproduciendo, así como la importancia de impulsar procesos de apropiación que tomen esto en cuenta desde el contexto latinoamericano.

AI

Cuando en el mes de febrero de 2023, el servicio de OpenAI, el muy popular *ChatGPT* alcanzó sus primeros 100 millones de usuarios, llegamos a un hito no solo numérico sino cultural: nuestros procesos de apropiación de las tecnologías digitales de comunicación interactiva (como prefiero llamar a estas en lugar del acrónimo TIC) ahora están marcados por una nueva herramienta, los modelos de *inteligencia no humana*, como las llama Scolari (2023).

Previo a esa fecha, las cuestiones sobre apropiación técnica y cultural de tecnologías digitales consideraban poco a este tipo de herramientas que en cuestión de unos meses han empezado a irrumpir en distintas interfaces y plataformas: no solo es posible hacer uso de ChatGPT y Dall-e desde OpenAI sino que la alianza de esta compañía (alianza, ante todo, comercial) con Microsoft ahora han llevado el servicio al buscador, al navegador y en breve, al propio sistema operativo de Windows con *Copilot*. En tanto, Google y su *Bard* tratan de mantener al día al otrora líder de búsquedas en internet. Los generadores de imágenes como *Midjourney* y *Stable Difussion* marcan un nuevo ritmo en materia de visualidad mientras que otros desarrollos nos plantean la generación de música, video y un etcétera que aún está por verse.

Si bien los modelos de inteligencia no humana (seguiremos en esto a Scolari) han estado presentes desde hace ya bastante tiempo, estos no se habían puesto a disposición del usuario promedio del modo en que OpenAI y otros desarrollos lo han hecho. El genio ha sido liberado de la botella y ahora cualquier persona, desde cualquier lugar, puede hacer uso de esta nueva herramienta.

Sin embargo, antes de apresurarnos a alabar las posibilidades (el tiempo que podríamos ahorrar, las soluciones que podríamos encontrar) o a condenar sus riesgos (la seguridad de los datos, el desplazamiento y sustitución de puestos de trabajo, etcétera), la irrupción de estas herramientas también nos obliga a plantearnos algunas preguntas en términos de las intervenciones a las matrices culturales: si los modelos de IA, en tanto que "nuevas" herramientas puestas a disposición nuestra, han de modificar los ecosistemas de comunicación, ¿qué clase de lecturas -discursos, representaciones, subjetividades- ponderarán o privilegiarán y cuáles quedarán invisibilizados por la máquina?

### Nuevas herramientas, viejos defectos

Cuando Safiya Umoja Noble en *Algorithms of Oppression* (2018) documentó los sesgos algorítmicos del motor de búsqueda Google como una herramienta que refuerza el racismo, tuvo la caución de señalar al menos dos cosas: que el origen de estas discriminaciones son *humanos* (quienes diseñan, construyen y "nutren" a las máquinas) y, que la sub-representaciones de las minorías en el diseño de las tecnologías tiene consecuencias sociales como la invisibilización de esas identidades.

Es decir, muchas prácticas y discursos que reproducen y/o agudizan la desigualdad han sido *transfundidos* a la *arquitectura* de las tecnologías digitales y desde ahí, en ocasiones, permiten perpetuar dichas inequidades (Harmer & Lumsden, 2019). ¿Habrá de ser esto diferente con las IA?



Figura 1. Resultados de Dall-E en Bing al *prompt* "una persona sonriendo frente al espejo".

En la Figura 1 hay cuatro resultados de Dall-e en Bing a partir del *prompt* "una persona sonriendo frente al espejo". El ejercicio se repitió durante un tiempo tratando de identificar algún patrón y estas imágenes de alguna manera lo representan: la mayoría de los resultados corresponden a personas de piel blanca y rasgos anglosajones, en promedio el 25% corresponden a alguna minoría (rasgos afroamericanos o asiáticos).



Figura 2. Dall E 2 y algunos ejemplos de lo que puede llegar a generar

Otros ejercicios realizados por el autor, han identificado que herramientas de generación de imagen tienden a sobrerrepresentar en sus resultados los rasgos de personas del "norte global". Esto era incluso más acusado en las primeras versiones de la herramienta y parece irse matizando en las últimas versiones.

De momento, parece estar claro que hay un sesgo racial en los modelos generativos de imágenes. Estos sesgos no son inocentes, tienen su origen en prácticas y discursos dominantes que son reproducidos (no siempre con conciencia de ello) por quienes llevan a cabo estos desarrollos. A ello se puede añadir el hecho de que, siguiendo a Noble (2018), a menor presencia de minorías involucradas en el desarrollo de estas tecnologías, hay una menor probabilidad de que sus inquietudes, preocupaciones, subjetividades, etcétera, se vean representadas en estas tecnologías.

### ¿De qué color es la vida, según la IA?

La forma en que ciertas ideas, desde lo más sublime hasta lo más mundano, es interpretado por las inteligencias no humanas, tendrá consecuencias sociales. Si los modelos generativos han de convertirse en un nuevo actor de los ecosistemas de comunicación, es decir, si han de actuar como mediadores entre los usuarios y las formas culturales contemporáneas, eso implica que favorecerán ciertas interpretaciones del mundo en lugar de otras.

El uso y apropiación de las herramientas de IA habrá de modificar cómo consumimos otras tecnologías digitales de comunicación interactiva. Si, como Microsoft pretende, serán un *copiloto*, sus respuestas nos guiarán para encuadrar segmentos de la realidad. Las IA tienen el potencial de cambiar la forma de "viejos medios" para habilitar "nuevas" formas de consumo, de *remediarlos* pues, entretejer lógicas mediáticas en nuevos procesos de mediatización (Hepp, 2020; Hjarvard, 2014). Pero eso implica que ciertas lógicas (es decir, muy particulares gramáticas y formas de sentido) sean ponderadas en lugar de otras.

Aquí valdría la pena volver la mirada hacia el concepto de representación y pensar qué tipo de propuestas simbólicas nos presentan (o ponderan o resultan de nuestra interacción con) las inteligencias no humanas.

La *Figura 2* es un ejercicio orientado en este sentido. En ella se muestran cuatro resultados cada uno a partir de un *'prompt'* distinto. El primero (superior izquierda) usó solo la palabra "pobreza" como indicación a Dall-E; en tanto, el segundo (superior derecha) usó la palabra "riqueza". En la parte inferior se usaron las propuestas "tristeza" (izquierda) y "alegría" (derecha).



Figura 3. Resultados de Dall-E a los *'prompts'* pobreza (superior izquierda), riqueza (superior derecha), tristeza (inferior izquierda) y alegría (inferior derecha)

El ejercicio se repitió en distintas ocasiones con resultados muy similares. El patrón que parece surgir es una forma particular de representar los conceptos usados como *'prompts'*. La pobreza es gris, de avanzada edad, solitaria; la tristeza tiene rasgos muy similares en términos de composición y color, aunque no de edad. ¿Por qué las imágenes de lo "pobre" refieren a la senectud? En la otra mano lo que tenemos es que la riqueza es joven, ostentosa (los fajos de billetes son siempre visibles), poderosa, con una actitud retadora mientras que la alegría brilla con colores cercanos al dorado (los amarillos resaltan siempre) y tiene rostro de mujer, de una mujer rubia para ser precisos.

Algunas otras dicotomías ofrecen resultados que siempre dan para reflexionar: la obediencia es animal, la desobediencia tiene rostro de mujer, la responsabilidad es más masculina que femenina, la irresponsabilidad tiene rostro juvenil, la enseñanza (el acto de enseñar, educar, etcétera) casi siempre es femenino, como el de cuidar, mientras que vigilar es masculino. Las representaciones de emociones, conceptos, nociones, parecen reproducir un tipo de discurso y forma de interpretar la realidad que hay que poner en cuestión y no reproducir acríticamente.

## Datificación avasallante

El desarrollo de internet nos condujo de aquella web preponderantemente de texto a una red centrada en la interacción y contenidos generados por los usuarios. El camino por la web social y la integración de aplicaciones ha dejado una víctima en el camino: nuestros datos.

Las plataformas de redes sociales lograron lo que las puntocom no pudieron, la forma de convertir los datos producidos por los usuarios en una fuente de producción de valor (van Dijck, 2016; van Dijck *et al.*, 2018). Estos procesos de *datificación* han sido fundamentales para el propio desarrollo, tanto técnico como económico, de lo sociodigital. Sin embargo, esto ha sido bastante a costa de los usuarios.

Las consecuencias sociales de la *datificación* no han dejado de analizarse, desde el caso *Cambridge Analytica* hasta los procesos de polarización política contemporáneos. Las herramientas de IA han requerido sus propios procesos de datificación para emerger.



Figura 4. Cambridge Analytica y el caso de datificación

*Common Crawl* es el nombre del repositorio consistente en más de 300 Terabytes de información usados para entrenar a ChatGPT. Los *datos* de este repositorio tienen muy distintas fuentes, incluyendo informes de autores y empresas. Pero modelos de IA como este siguen en entrenamiento y requerirán más y más datos para alimentarse.

Sin embargo, el hecho de que prácticamente “todos los datos” puedan ser ingeridos por los modelos de IA no implica que todos los discursos, prácticas, identidades, terminen representadas en los resultados. Por el contrario, aquellos recursos y discursos propios de los actores dominantes bien pueden convertirse en los que terminen más representados en los resultados de estas herramientas (nótese el resultado de los ejercicios aquí expuestos).

¿Cómo habrán de subvertirse estos procesos para que la balanza de la relación de poder no vuelva a caer en el lado históricamente conocido?

Desde Latinoamérica, contamos con ejemplos en los que las lógicas de dominación propias de la datificación se han subvertido para ser aprovechadas por la ciudadanía en procesos que hemos llamado de *datificación crítica* (Ábrego Molina & Flores Mérida, 2021). Esto implica procesos que tomen distancia para analizar y luego iniciativa para apropiarse de las herramientas que surgen en lo sociodigital.

Las IA no deberían quedar exentas a estos procesos de apropiación crítica por parte de distintos actores, sobre todo en el contexto subcontinental del que somos parte.

Una de las ventajas de que el genio haya salido de la botella es que los modelos generativos son ahora más accesibles que los modelos de IA privados preexistentes. Estos pueden “entrenarse” con datos propios para poder llevar sus potencialidades hacia los problemas que nuestros contextos requieren.

Por supuesto que ello requiere de conocimientos y habilidades técnicas que deben adquirirse. Sin embargo, estas mismas herramientas pueden facilitar dicho proceso.

Los sesgos que las IA tienen no son inamovibles y los procesos de reproducción simbólica que sugieren no tienen que ser imparables. Si han de ser, como creo que serán, un actor más de los ecosistemas comunicacionales y su irrupción modificará nuestras pautas de consumo y producción, antes que celebrarlos o vetarlos, hay que encontrar formas de apropiarlos y fomentar procesos que los pongan en manos de las ciudadanías.

## Referencias

- Ábrego Molina, V. H., & Flores Mérida, A. (2021). *Datificación crítica: Práctica y producción de conocimiento a contracorriente de la gubernamentalidad algorítmica. Dos ejemplos en el caso mexicano*. Administración Pública y Sociedad (APyS), 11, 211-231.
- Harmer, E., & Lumsden, K. (2019). Online Othering: An Introduction. En K. Lumsden & E. Harmer (Eds.), *Online Othering* (pp. 1-33). Springer International Publishing. [https://doi.org/10.1007/978-3-030-12633-9\\_1](https://doi.org/10.1007/978-3-030-12633-9_1)
- Hepp, A. (2020). *Deep mediatization*. Routledge, Taylor & Francis Group.
- Hjarvard, S. (2014). Mediatization and cultural and social change: An institutional perspective. En K. Lundby (Ed.), *Mediatization of Communication* (pp. 199-226). De Gruyter. <https://doi.org/10.1515/9783110272215.199>
- Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. New York University Press.
- Scolari, C. A. (2023, abril 1). *ChatpGPT: Del Test de Turing al Test de Baricco. Hipermediaciones*. <https://hipermediaciones.com/2023/04/01/baricco-turing-chatgpt/>
- van Dijck, J. (2016). *La cultura de la conectividad: Una historia crítica de las redes sociales* (H. Salas, Trad.). Siglo Veintiuno Editores, Argentina.
- van Dijck, M. J., Poell, T., & Waal, M. de. (2018). *The platform society*. Oxford University Press.

Intelligence  
computer graphics/CSDevelopers

```
#include <stdio.h>
int main()
{
  int i;
  printf("Dec/Oct/Hex\n");
  printf("-----\n");
  for(i=0; i=15; i++)
  {
    printf(" %d\t%o\t%X\n", i, i, i);
  }
}
```

### Ficha de autor

**Dr. Antony Flores Mérida:** [nflores@colmex.mx](mailto:nflores@colmex.mx)

Doctor en Ciencia Social con especialidad en Sociología por El Colegio de México (COLMEX) y Maestro en Estudios Culturales y Licenciado en Ciencias de la Comunicación por la Universidad Autónoma de Chiapas (UNACH). Desarrollará el proyecto posdoctoral "Prácticas de comunicación y resistencia en ambientes sociodigitales" en el Instituto de Investigaciones Sociales de la UNAM en el periodo 2023-24.

Figura 5. Desarrollo de la IA en el mundo empresarial, artístico y político

**ci**

**CANAL IBEROAMERICANO**  
*Señal que nos une*

**PROYECTA LA ESENCIA LATINOAMERICANA**

**DESCARGA NUESTRA App**

**ILCE player**

DISPONIBLE EN Google Play

Disponible en el App Store