

A PROPÓSITO DE METÁFORAS: “INTELIGENCIA ARTIFICIAL” Y NECESIDAD DE UNA ÉTICA ANTICIPATORIA

Fernando Lolas Stepke¹

En sistemas sociotécnicos complejos, por ejemplo “medicina-ciencias de la salud-cuidado”, las preguntas fundamentales se asocian a utilidad, aplicación razonable, respeto a la dignidad y las diversidades humanas, y prevención o evitación de daños a personas y ambiente.

Todo pensamiento creativo se asocia a metáforas. Decir con otras palabras e imágenes lo inusual o necesitado de belleza y evocación es atributo de todo lenguaje. Las expresiones metafóricas hacen visible lo invisible, permiten aprehender imaginativamente lo que cuesta expresar. Hacen migrar de un campo experiencial a otro los referentes que permiten expresión e impresión.

Algunas metáforas son más exitosas que otras y se convierten en formas habituales de aludir a realidades. Se acepta que “los ojos son los espejos del alma”, que “la vida humana es un río o un camino”, que el “Ser Divino es bondad”.

Escoger palabras es arte antiguo. Basta leer *Agudeza y arte de ingenio* de Baltazar Gracián, suma y compendio de muchos tratados de gramática, retórica y elocuencia. La retórica clásica, con sus etapas de *Inventio*, *Dipositio*, *Elocutio*, construye realidades. Gran logro y gran peligro el de las *metáforas ocultas* que heredamos con el lenguaje de nuestra cultura. Las culturas moldean las mentes de sus miembros.

Esto es relevante cuando nuevas disciplinas científicas o tecnologías reciben denominaciones que sin iluminar su contenido y fines los entenebrecen. Lo que las metáforas reemplazan tiene vida propia y sugiere infinidad de comparaciones y significados.

“Inteligencia artificial” (IA) alude a operaciones tecnológicas y dispositivos que generan respuestas contextualmente acertadas. Por ejemplo, escribir textos con sentido y estilo, interpretar imágenes, anticipar comportamientos. Una aceptable definición se encuentra en un documento europeo:

“Software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structure or unstructured data, reasoning on the knowledge, or process the information, derived from this data and deciding the best actions(s) to take to achieve the given goal” (European Commission Joint Research Centre et al)(1).

La expresión “inteligencia artificial” es metafórica y fue acuñada por John McCarthy en 1956, pero tiene sus orígenes en los trabajos de George Boole del siglo XIX, quien propuso que todo razonamiento lógico puede reducirse a un sistema de ecuaciones. Clásicamente, la inteligencia humana alude a capacidades humanas relacionadas con atención, cognición, memoria, comprensión, adaptabilidad creativa y acción prudente. ¿A qué se refiere, exactamente, la expresión “inteligencia” en el sintagma inteligencia artificial (AI)? A un atributo humano imitado por artilugios y algoritmos parcialmente, aunque se espera que su desarrollo lleve a una imitación más completa, incluyendo creatividad y acción adecuada (IA generativa).

Las “herramientas” de IA son producto de actividad humana. Alimentar bases de datos y algoritmos depende del anónimo trabajo de miles de personas, ejemplo de colaboración social. Impresionan porque se desconstruyen o se ignora de donde proceden sus informaciones. El efecto *black box*, que impide comprensión por no expertos, fomenta el misterio. Algunas variantes indican fuentes, lo que revela que no hay un ser pensante sino un algoritmo que reconstituye y recombina lo realizado por otras personas. La mayor parte de las denuncias contra las compañías de IA se refiere a propiedad intelectual y a la explotación no autorizada de

¹ Profesor Titular, Universidad de Chile y Universidad Central de Chile, Chile. Director de *Acta Bioethica*, flolas@uchile.cl, <https://orcid.org/0000-0002-9684-2725>

fuentes existentes. Eso expone a sesgos (eurocentrismo, machismo, economicismo, entre otros).

Las expresiones metafóricas sugieren que tratamos con “seres” pensantes. “Regular” la inteligencia artificial es como “regular” cerebros. Tener neuroleyes y neuroderechos se confunde con tener regulaciones de “inteligencias” naturales y artificiales. Hay entusiasmo y temor.

Nuevas tecnologías designadas con palabras antiguas. Su carga semántica y emocional obliga a reaccionar como si se tratara de realidades concretas. Lo que imita al sistema nervioso humano no puede escapar a las fantasías asociadas, pues no hay un solo “Sistema Nervioso Central” (SNC) sino muchos “sistemas nerviosos conceptuales” (SNC) según los métodos para su investigación. La *neuromitología* se entroniza y el prefijo “neuro” parece explicar o dar fundamento.

Lamentablemente, para difundir inventos novedosos hay que usar expresiones no novedosas. Eso permite que se entiendan y difundan. Pero no estimulan la creatividad, alimentan temores y expectativas no realistas.

Existe el concepto de “*data dignity*”. Parece tratarse de una “justicia contributiva” que contextualiza datos del ciberespacio. La expresión “*data as labor*” sugiere que la contribución individual, aunque modesta (a veces solo pasiva entrega de informaciones por usuarios), merece ser reconocida. Las *black boxes* de GPT y otras “inteligencias”, al ser abiertas, demostrarían el origen contextualizado de las capacidades, todas provenientes de trabajo humano procesado por ingeniería creativa.

Es preciso “desmitologizar” y buscar designaciones que no evoquen antropomórficas capacidades. Norbert Wiener y su *cibernética* parecían novedades impresionantes y no hay cambios en la actitud hacia el “pensamiento artificial”, como decía Pierre de Latil(2). Las *mitologías metafóricas* persisten e incluso los científicos reproducen el “conocimiento convencional” expresado en novelas, filmes y productos de la imaginación.

La cultura es invencible porque es invisible.

Los artefactos y algoritmos ejecutan acciones. Cuando éstas se transforman en funciones (conjuntos de acciones con “fines macro” aparentes) parecen ser inteligentes en el sentido de tener “intenciones” o “finalidades”. Esta confusión antropomórfica es mítica, produce expectativas ilusorias y peligrosidades que exigen un uso correcto del idioma.

Las reflexiones éticas sobre transformaciones tecnológicas suelen ser “reacciones” a desarrollos científicos e industriales. Numerosas regulaciones promulgadas por organismos internacionales y sociedades científicas plantean limitaciones y restricciones para evitar males sociales. Muchas son deliberaciones jurídicas o sociales que subestiman aspectos económicos y políticos. Son generalidades. Tienen valor normativo sin considerar el mundo real y las motivaciones de quienes desarrollan tecnologías. Aluden a dilemas posteriores a la adopción de las innovaciones que se difunden y comercializan, a menudo con gran éxito popular. Proponen proteger privacidad, transparentar procedimientos, aplicación supervisada, acuerdos sociales para moderar su uso y permitir acceso equitativo a sus beneficios. El carácter general y abstracto de tales admoniciones exige una actitud anticipatoria, prospectiva, que incluya desde su mismo origen consideraciones éticas en el diseño, las aplicaciones y la utilidad en campos específicos(3). No es lo mismo el campo de la medicina que el de la justicia criminal o el trabajo académico. La especificidad de la IA generativa debe ser estipulada en cada nueva aplicación.

Referencias

1. Samoili S, López Cobo M, Gómez E, De Prato G, Martínez-Plumed F, Delipetrev B. AI watch: defining artificial intelligence: towards an operational definition and taxonomy of artificial intelligence. *European Commission, Joint Research Centre Publications Office*, 2020. <https://doi.org/10.2760/382730> [Cited 2023 Aug 22].
2. De Latil P. *El pensamiento artificial*. Buenos Aires: Losada; 1958.
3. Arbelaez Ossa et al. Integrating ethics in AI development: a qualitative study. *BMC Medical Ethics* 2024; 25: 10. <https://doi.org/10.1186/s12910-023-010>