**Máquinas intencionales: una defensa de la confianza en la inteligencia artificial médica**

Georg Starke, Rik van den Brule, Bernice Simone Elger, Pim Haselager

18 de junio de 2021

Abstract:

La confianza constituye una estrategia fundamental para afrontar los riesgos y la incertidumbre en sociedades complejas. De acuerdo con la vasta literatura que enfatiza la importancia de la confianza en las relaciones médico-paciente, la confianza se sugiere regularmente como una forma de lidiar con los riesgos de la inteligencia artificial médica (IA). Sin embargo, este enfoque se ha hecho cargo desde diferentes ángulos. Se pueden distinguir al menos dos líneas de pensamiento: (1) que confiar en la IA es conceptualmente confuso, es decir, que no *podemos* confiar en la IA; y (2) que también es peligroso, es decir, que no *debemos* confiar en la IA, especialmente si lo que está en juego es tan alto como lo es habitualmente en la medicina. En este artículo, nuestro objetivo es defender una noción de confianza en el contexto de la IA médica contra ambos cargos. Para hacerlo, destacamos las intenciones técnicamente mediadas que se manifiestan en los sistemas de inteligencia artificial, lo que hace que la confianza sea una postura conceptualmente plausible para lidiar con ellos. Con base en la literatura de las interacciones humano-robot, la psicología y la sociología, proponemos un modelo novedoso para analizar las nociones de confianza, distinguiendo entre tres aspectos: confiabilidad, competencia e intenciones. Discutimos cada aspecto y hacemos sugerencias sobre cómo la IA médica puede ser digna de nuestra confianza.

Referencia:

Starke, Georg, Rik van den Brule, Bernice Simone Elger, y Pim Haselager. 2021. «Intentional Machines: A Defence of Trust in Medical Artificial Intelligence». *Bioethics* 1-8. doi: [10.1111/bioe.12891](https://doi.org/10.1111/bioe.12891).