



Frente al flujo incesante de necesidades insatisfechas en todo el mundo, los actores humanitarios son muy susceptibles a las profecías del *solucionismo tecnológico*: cuando una nueva solución ofrece el *potencial* de mejorar la eficacia de nuestro trabajo, nos ilusionamos. Guiados por la esperanza de que esto mejore la situación de la gran cantidad de personas que vemos sufrir en desempeño de nuestra labor, a veces nos dejamos llevar por esa ilusión. Este sesgo natural en favor de la innovación se amplifica con el *determinismo tecnológico* que nos rodea. Convivimos con una fuerte sensación de que la IA ya no es una opción, sino un imperativo, un requisito innegociable para el humanitarismo contemporáneo, que no puede permitirse quedarse atrás ante los avances tecnológicos.

En otras palabras, no hay tiempo para pensarlo; no usar IA sería irresponsable.

Por otro lado, el *pánico moral* y los miedos distópicos en torno a los riesgos que entraña esta innovación pueden paralizar a los profesionales humanitarios. Puede resultar difícil comprender qué importancia podría tener la IA para comunidades que viven en lugares remotos donde escasean los alimentos y donde la conectividad sigue siendo una realidad lejana. Asimismo, la *terminología* alienante y la opacidad técnica que caracterizan las conversaciones sobre IA generan distancia con los actores humanitarios. Mientras crece la *lista* de innovaciones con IA para el sector humanitario (desde análisis predictivo para facilitar la planificación de programas y las evaluaciones de necesidades hasta el uso de *chatbots* para gestionar solicitudes de información e interacciones con poblaciones afectadas), también se multiplican los debates en el sector en torno a *los riesgos y las responsabilidades* asociados al uso de estas tecnologías. A decir verdad, los desafíos por delante pueden ser abrumadores.

En otras palabras, no hay tiempo para pensarlo; usar IA sería irresponsable.

Durante los últimos dos años, hemos observado cómo se manifiesta esta dualidad en el CICR. Cuando hace 18 meses empezamos a documentar nuestro uso de la IA, era muy evidente que había interés y entusiasmo entre los colegas de explorar su potencial para mejorar los efectos humanitarios. También había una cuota similar de miedo e incertidumbre frente a una tecnología que muchos de nosotros no entendemos en profundidad.

En nuestras conversaciones con miembros de toda la organización, nos dimos cuenta de que la IA ya estaba muy integrada a las actividades del CICR — desde los sistemas de gestión de información a los de apoyo logístico— y que era el motor detrás de muchos *proyectos* de innovación. También quedaron de manifiesto las disparidades en la alfabetización digital y la capacidad de nuestros equipos de entender cómo funciona la IA y qué efectos podría tener en la labor humanitaria. Mientras que algunos tenían amplia experiencia con esta tecnología, otros no lograban percibir los riesgos básicos que plantea. Mientras que algunos proponían acelerar la implementación con un enfoque más audaz y competitivo, otros tenían sus recaudos y querían aminorar la marcha o dar un paso al costado.

Este sondeo inicial dejó en evidencia la necesidad imperiosa de brindar orientaciones para que el personal del CICR pudiera aprender y explorar si la IA puede ayudar a la organización a llevar adelante su cometido y, en ese caso, cómo, cuándo y dónde utilizarla.

En ese sentido, crear una política para el uso de la IA con fines humanitarios fue todo un desafío. Dejamos que los principios humanitarios *guiaran nuestras decisiones éticas* y los usamos como un punto de partida común para abordar el dilema del uso de la IA con fines humanitarios.

De todas formas, nada de esto fue fácil.

## La necesidad de deconstruir nuestros sesgos en cuanto a la IA

Nuestro principal obstáculo era deconstruir los sesgos cognitivos que nos impedían analizar la IA de forma objetiva.

Nuestro primer sesgo surge del marco binario que suele aplicarse al debate en torno a la IA: o será nuestra salvación o destruirá el mundo. Estas visiones polarizadas y enfrentadas no sirven de mucho. No es la IA la que va a salvar o destruir el planeta; *sino los seres humanos*. Por lo tanto, nuestra primera medida fue tomar como eje de la política un enfoque *centrado en las personas*.

El segundo sesgo es una mezcla de fascinación y emociones preconcebidas que nos invita a ilusionarnos con cualquier tecnología nueva que pueda ayudarnos a cumplir con nuestros objetivos. En la práctica, este sesgo cognitivo suele manifestarse en un enfoque guiado por las soluciones, que toma la IA como punto de partida. Esto plantea la pregunta: ¿qué podemos hacer con ella? Sin embargo, la IA *no* es una tecnología nueva, y el objetivo de los actores humanitarios no es innovar por el mero hecho de *innovar*. Mediante nuestro enfoque centrado en las personas y *guiado por los problemas*, el punto de partida son las necesidades humanitarias, no la IA. De esta forma, la pregunta pasa a ser: ¿cuál el problema humanitario y es posible que la IA nos ayude a dar respuesta de forma ética y segura? De ser así, ¿cuándo, dónde y cómo?

El tercer sesgo es la idea de que las tecnologías, como la IA, son solo meras herramientas, es decir, que no acarrearán problemas o valores intrínsecos. Eso implica que todo depende de nosotros y que, si la usamos para el bien (por ejemplo, para la labor humanitaria), la IA será “buena”. Pero la tecnología *nunca* ha sido *neutral*, y la IA es un excelente ejemplo de la naturaleza *política* y el *doble propósito* de este tipo de innovaciones. Como es bien sabido que, si no nos interesa la política, la política se interesa por nosotros —también a través de la IA—, centramos nuestro enfoque en los principios humanitarios de neutralidad e independencia para gestionar esas tensiones.

De esta manera, la identificación y el tratamiento de esas distintas trampas cognitivas nos ayudó a crear un espacio de reflexión compartida menos sesgado para deconstruir la complejidad de la IA y crear nuestro propio enfoque *humanitario*.

## Hacia un enfoque humanitario para el uso de la IA

El principio de humanidad es el motor de la labor humanitaria, así como de la acción del CICR. Es un *recordatorio* muy útil cuando hablamos de IA. Para evitar caer en la óptica general centrada en la productividad, nuestra política adopta una perspectiva de derechos humanos y vulnerabilidades guiada por el valor. Tenemos la determinación de hacer todo lo que está a nuestro alcance —incluso usar IA— para aliviar el sufrimiento y proteger los derechos y la dignidad de las personas afectadas por conflictos armados y otras crisis humanitarias. Un uso responsable exige identificar salvaguardas eficaces para mitigar el riesgo de que las soluciones de IA contribuyan, por medio de algoritmos, al abuso o la discriminación de las personas que ya afrontan situaciones de sufrimiento; la clave es hacerlo *antes* de implementar dichas soluciones.

Este enfoque de precaución implica analizar cuidadosamente los riesgos de utilizar IA en cada situación y usarla únicamente en caso de que tenga un efecto positivo tangible en la vida de las personas afectadas o en la capacidad del CICR de protegerlas y prestarles asistencia. En consecuencia, se evitaría usar esta tecnología si erosiona nuestra capacidad de interactuar de forma directa y tener presencia física cuando se nos necesita. También implica no usarla si no están dadas las condiciones de salvaguarda necesarias para cumplir con el propósito de *no hacer daño*.

Las medidas de salvaguarda eficaces deben evaluar para qué se usan las soluciones de IA, qué consecuencias tienen sobre la población afectada y qué tipos de datos se utilizan de base, en especial si son datos personales, confidenciales o sensibles. La IA pone bajo la lupa la disponibilidad, calidad y fiabilidad de los datos; básicamente, la eficacia de los algoritmos depende de la calidad de los datos que usan. Para aplicarla correctamente al sector humanitario, primero tenemos que adoptar el enfoque correcto para la protección de datos. En ese sentido, a pesar del enorme progreso alcanzado en los últimos *diez* años, todavía queda *mucho por hacer*.

El aumento en la polarización y la *competencia* estratégica global en torno a estas tecnologías también hace que el uso humanitario de la IA se vuelva un asunto político complejo. Si se usa o no, para qué y con quién se han vuelto decisiones políticas que pueden influir en cómo nos perciben tanto las personas y las comunidades como los Gobiernos y las partes en conflicto. Para los actores humanitarios, es esencial gestionar esa percepción, junto con la confianza, el acceso y la seguridad que nos confiere. En consecuencia, es imperioso mejorar nuestras capacidades en materia de diligencia debida y procurar que los procesos de compra contemplen como corresponde los riesgos que representan las soluciones de IA que adoptamos y sus proveedores para la imagen del CICR.

## El mito del vacío normativo: no hace falta reinventar la rueda

A menudo, con este tema, circula un discurso que da la impresión de que la IA surge en medio de un vacío total. Como suele pasar con las innovaciones tecnológicas, pareciera que, como es una tecnología “nueva” y sus usos se multiplican con tanta rapidez, no tenemos los instrumentos o marcos normativos para encuadrarla.

Sin embargo, esto no es así. Tenemos un gran abanico de normas jurídicas, principios éticos y buenas prácticas de innovación que pueden aplicarse, por ejemplo, el *derecho internacional humanitario* y los *derechos humanos*, así como legislación nacional e internacional en materia de *protección de datos*, *responsabilidad sobre productos*, *propiedad intelectual* y *ética* en la innovación. Es fundamental utilizar esas normas para regular los sistemas de IA *antes* de implementarlos, en lugar de inventar reglas nuevas para justificar su incumplimiento de forma retroactiva.

La otra buena noticia es que los especialistas en IA ya han identificado los *principios esenciales* —así como las *trampas*— que definen cómo sería un enfoque de uso ético y responsable. En los últimos años, se han publicado *muchas* directrices éticas para adecuar el uso de esta tecnología a los valores humanos. En consecuencia, nuestra *tarea* como actores humanitarios no es reinventar la rueda en cuanto a la ética del uso de la IA, sino definir cómo podemos incorporar esos principios a nuestro ámbito. Para eso, es necesario traducir los conceptos de la IA al lenguaje de la acción humanitaria, de modo que nuestro sector pueda dialogar y reflexionar acerca de este tema en términos que le sean pertinentes y propios. Ese es justamente el propósito de la política del CICR relativa a la IA.

Al tener la terminología correcta, es más fácil dialogar con los especialistas en IA, plantear problemáticas y aprender a partir de su experiencia práctica con el fin de explorar soluciones. Solicitamos activamente el apoyo de expertos externos en ética relativa a la IA y aprendizaje automático para verificar nuestro enfoque y distinguir las incógnitas conocidas de las desconocidas. Eso marcó una gran diferencia. No hay nada más poderoso que usar la *inteligencia humana* colectiva para darle sentido a la *inteligencia artificial*.

## El contexto lo es todo

Algo que hemos aprendido de estas conversaciones es que siempre se debe pensar la IA en contexto: el contexto *externo* de los lugares donde se procura usar estas soluciones o causar un efecto puntual, y el contexto *interno* y el ecosistema donde se van a implementar, que contempla la infraestructura digital y el entorno de datos, los puntos fuertes y débiles de la organización y factores humanos y socioculturales existentes. No es suficiente pensar en qué puede hacer la IA en favor de la acción humanitaria; sino que también hay que considerar cómo puede *afectar* a la acción humanitaria.

En un contexto en que los presupuestos humanitarios están estancados, es normal precipitarse y buscar soluciones que podrían ahorrar costos. Mientras que el cálculo suele restringirse a los ahorros de corto plazo, las inversiones en IA deberían analizarse con miras a un futuro más lejano. Para estimar los ahorros de costos que podrían producirse al automatizar la labor humana, se debe considerar la inversión necesaria para contratar especialistas en IA y adquirir capacidad computacional adicional, el tiempo que le llevará al personal evaluar medidas de precaución y mitigación, la capacitación para que el personal pueda usar la IA en consonancia con los compromisos de la política y los riesgos de  *depender de un proveedor específico*. La IA tiene el potencial de generar una buena relación costo-beneficio, pero si no se analiza el costo integral, no es más que una promesa vacía.

Para un análisis de costos frente a beneficios, se requiere un enfoque multidisciplinario en el que las distintas áreas de la organización y toda la cadena de valor humanitaria puedan aportar su perspectiva crítica. Entender las limitaciones sistémicas y operacionales, qué necesitan los usuarios y su configuración operacional, o en qué casos los datos representan una oportunidad o una amenaza lleva tiempo y esfuerzo. A su vez, exige confiar tanto en aquellas personas que tienen la pericia técnica como en quienes pueden enmarcar adecuadamente las soluciones propuestas en la ética humanitaria y las realidades operacionales.

## Los pilares de un uso responsable de la IA

Para crear un enfoque responsable para el CICR, usamos los *principios éticos* que han surgido como normas mínimas a escala internacional. Se basan en los marcos normativos mencionados anteriormente y pueden resumirse en las siguientes orientaciones para el personal del CICR que debe tomar decisiones relativas a la IA.

### ***El principio de precaución***

Frente a la gran cantidad de preguntas sin respuesta acerca de los efectos y los riesgos de la IA en el ámbito humanitario, es evidente por qué el *principio de precaución* es tan importante. Junto con un enfoque basado en la proporcionalidad, nos ayuda a usar esta tecnología para resolver problemas cuando hay demostración empírica de que tendrá un efecto positivo en contexto.

### ***No causar daño***

Priorizar la protección y la seguridad de las personas es parte del requisito de *no causar daño* que impone la ética humanitaria. En la práctica, implica identificar explícitamente el tipo de datos que se incorporarán a los sistemas de IA, y la índole y los efectos potenciales de los riesgos que trae aparejado su uso. Se debería implementar sistemas *diferenciados* de mitigación de riesgos de ciberseguridad y de protección de datos en función del grado de sensibilidad de los datos procesados y las consecuencias de los riesgos para el usuario final. Mientras que usar soluciones de IA en procesos administrativos e internos puede no representar un riesgo muy grande, su uso en actividades externas con incidencia en la realización de programas o las interacciones con poblaciones afectadas puede tener fuertes repercusiones en la vida de las personas y debe gestionarse adecuadamente. Proteger los datos literalmente significa proteger a las personas.

### ***Transparencia***

La mayoría de los sistemas de IA a la venta en el mercado son “*cajas negras*”, de modo que es casi imposible analizar por completo su funcionamiento interno. Sin embargo, en el sector humanitario, la transparencia y la “*explicabilidad*” son de gran *importancia*. Es imperioso que las organizaciones humanitarias hagan un esfuerzo activo y continuo de ser transparentes y explicar —como mínimo— cuándo, por qué y cómo usan la IA en el desarrollo de sus actividades. La prueba de fuego es sencilla: si no pueden dar esas explicaciones, probablemente no deberían estar usándola. En ese sentido, las soluciones de IA de código abierto tienen la ventaja de que ofrecen mayor nivel de transparencia que los modelos patentados cerrados.

Por otro lado, puede ser difícil encontrar la forma correcta de explicar nuestro enfoque a las personas a las que prestamos servicios en términos que puedan entenderlo. Además, es sabido que las organizaciones humanitarias ya tienen un *déficit* en lo que respecta a la rendición de cuentas ante las poblaciones afectadas. Existe un riesgo considerable de que el uso de la IA pueda desvirtuar iniciativas en curso relacionadas y el progreso logrado si la *rendición de cuentas* se pierde en la *implementación* algorítmica. Es por eso que son tan importantes los principios de responsabilidad y rendición de cuentas a la hora de usar la IA con fines humanitarios.

### ***Gobernanza y rendición de cuentas***

Un enfoque centrado en las personas también exige una revisión periódica de las herramientas de IA durante todo su ciclo de vida a manos de profesionales calificados. Para eso, debemos mejorar nuestro conjunto de herramientas y estructurar nuestros mecanismos de gobierno y control para que el uso de la IA cumpla con las leyes y normas aplicables y con la *ética* humanitaria. En consecuencia, los expertos jurídicos deberán prestar apoyo a los equipos de ingeniería y a la toma de decisiones, y las estructuras de gobierno actuales deberán incorporar nuevas habilidades para gestionar las implicaciones del uso de IA en la acción humanitaria. Esto significa tener procesos de denuncia y reparación viables y poder brindar sistemas alternativos eficaces y accesibles a aquellos usuarios que no puedan —o no quieran— depender de soluciones de IA. También implica tener en cuenta los efectos negativos que pueden tener en nuestra capacidad de *reducir el impacto* ambiental de las actividades humanitarias y de mejorar la *localización* de la ayuda.

Es importante recordar que la responsabilidad y la rendición de cuentas van de la mano. Son independientes y solo pueden dar resultado mediante la transparencia genuina y la *participación* continua de usuarios y partes interesadas pertinentes. Es esencial tener un ciclo de retroalimentación humano para evitar los riesgos de *colapso* que plantea el ciclo de la IA.

## El comienzo de un camino de aprendizaje

Diseñar una política para el uso de la IA que guíe a nuestra organización y nos ayude a sortear las dificultades que propone esta tecnología no fue tarea fácil. Sin embargo, la adopción de una política solo es el primer paso en el camino del CICR hacia el uso de la IA con fines humanitarios.

Todavía nos queda mucho recorrer, pero ya contamos con un programa de formación virtual para ayudar a nuestro personal a adquirir los conocimientos digitales y de datos necesarios para explorar el uso de la IA con el fin de mejorar los efectos humanitarios. El consejo de Gobernanza Digital del CICR

utilizará la política relativa a la IA para supervisar su acatamiento e implementación, y nuestro equipo seguirá consultando con pares y especialistas externos para aprender y mejorar nuestro enfoque.

Es posible que algunas de nuestras iniciativas con IA fracasen o no brinden los resultados esperados. Se puede recurrir al método de prueba y error siempre y cuando no afecte negativamente a las personas con las que trabajamos y a las cuales prestamos servicio, y no suponga un costo desproporcionado.

La mayor dificultad será no confundir la necesidad urgente de *pensar* en el uso de la IA con la urgencia de *usarla*. En este sentido, la política del CICR relativa a la IA pretende ser aspiracional y sentar las bases necesarias para que abordemos de forma ética y con éxito el desafío de usar esta tecnología con fines humanitarios. Dado nuestro cometido humanitario, es esencial evitar la deshumanización, el abuso y la discriminación que un mal uso de la IA podría generar en las personas con las que trabajamos y a las cuales prestamos servicios.