

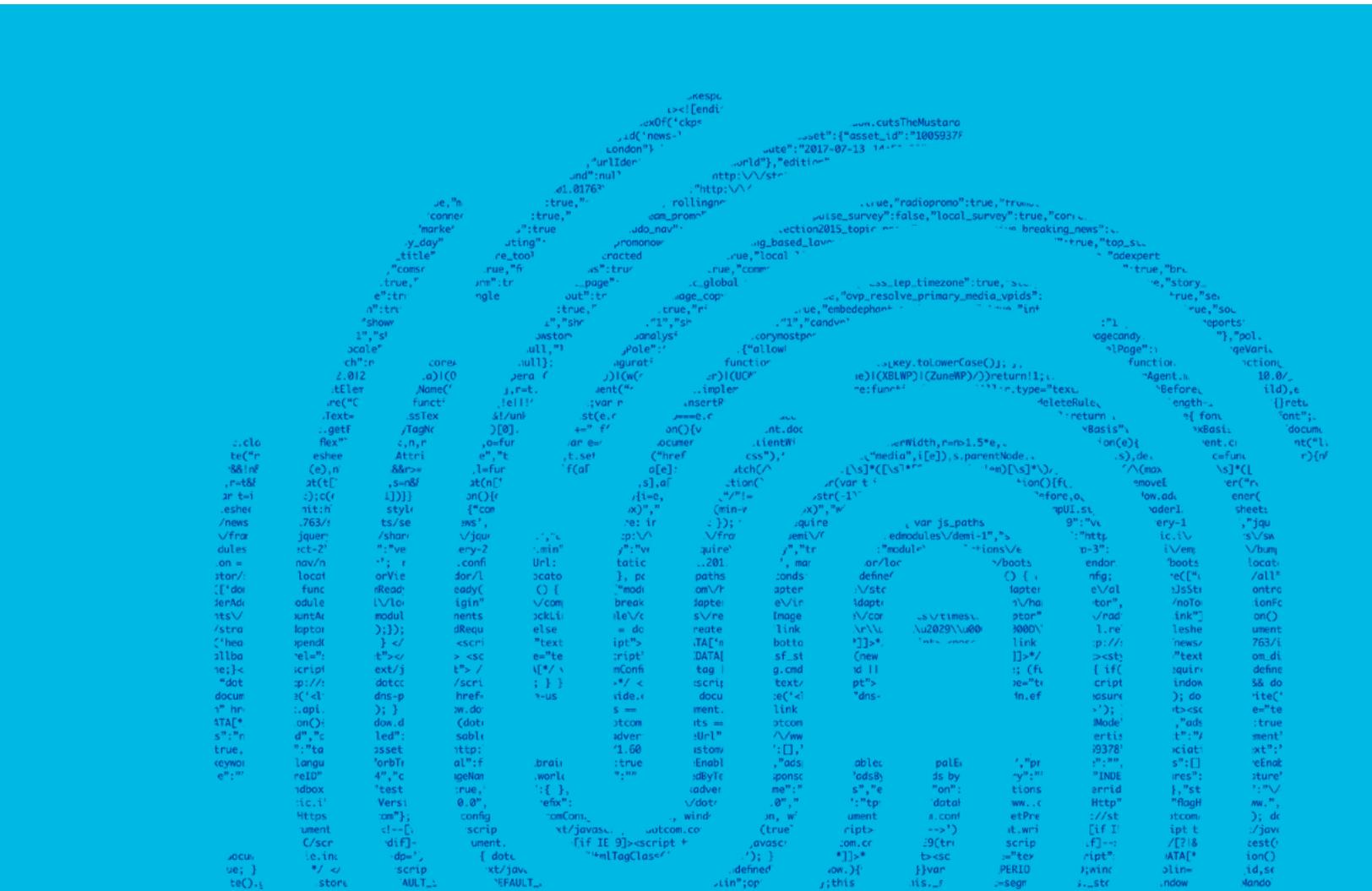


< >

Declaración de Montreal
IA responsable_

</ >

DECLARACIÓN DE MONTREAL PARA UN DESARROLLO RESPONSABLE DE LA INTELIGENCIA ARTIFICIAL 2018



Este documento forma parte de la **DECLARACIÓN DE MONTREAL PARA UN DESARROLLO RESPONSABLE DE LA INTELIGENCIA ARTIFICIAL** de 2018. Acceda al informe completo [AQUÍ](#).

ÍNDICE

LECTURA DE LA DECLARACIÓN	5
PREÁMBULO	7
PRINCIPIOS	
1. PRINCIPIO DE BIENESTAR	8
2. PRINCIPIO DE RESPETO A LA AUTONOMÍA	9
3. PRINCIPIO DE PROTECCIÓN DE LA PRIVACIDAD Y LA INTIMIDAD	10
4. PRINCIPIO DE SOLIDARIDAD	11
5. PRINCIPIO DE PARTICIPACIÓN DEMOCRÁTICA	12
6. PRINCIPIO DE EQUIDAD	13
7. PRINCIPIO DE INCLUSIÓN DE LA DIVERSIDAD	14
8. PRINCIPIO DE PRECAUCIÓN	15
9. PRINCIPIO DE RESPONSABILIDAD	16
10. PRINCIPIO DE DESARROLLO SOSTENIBLE	17
GLOSARIO	18
EQUIPO	I
COLABORADORES	II

LECTURA DE LA DECLARACIÓN

¿CUÁL ES EL PROPÓSITO DE LA DECLARACIÓN?

La Declaración de Montreal para un desarrollo responsable de la IA tiene tres objetivos principales:

1. crear un marco ético para el desarrollo y la implementación de la IA;
2. guiar la transición digital para que todos puedan beneficiarse de esta revolución tecnológica;
3. abrir un foro de discusión nacional e internacional a fin de lograr conjuntamente un desarrollo equitativo, inclusivo y ecológicamente sostenible de la IA.

¿DE QUÉ TRATA ESTA DECLARACIÓN?

PRINCIPIOS

El primer objetivo de la Declaración consiste en identificar los principios y los valores éticos que respaldan los intereses fundamentales de las personas y los grupos. Si aplicamos estos principios al ámbito de la inteligencia artificial y digital, siguen siendo generales y abstractos. Para hacer una lectura correcta, es importante tener en cuenta los siguientes puntos:

- > Si bien se los presenta en forma de lista, no hay un orden jerárquico. El último principio no es menos importante que el primero. No obstante, es posible, en función de las circunstancias, darle mayor peso a un principio que a otro o considerar que un principio es más relevante que otro.
- > Si bien son diversos, se los debe interpretar de manera coherente a fin de prevenir cualquier tipo de conflicto que pudiera impedir su aplicación. Como regla general, los límites de la aplicación de un principio se definen a partir del ámbito de aplicación de otro principio.
- > Si bien reflejan los valores morales y políticos de la sociedad en la que se formularon, sientan las bases para un diálogo intercultural e internacional.
- > Si bien es posible interpretarlos de diferentes maneras, no se los puede interpretar de cualquier manera. Es imprescindible que se los interprete de forma coherente.
- > Si bien se trata de principios éticos, es posible traducirlos al lenguaje político e interpretarlos en términos legales.

Sobre la base de estos principios, se elaboraron algunas recomendaciones, cuyo propósito es indicar pautas para llevar a cabo la transición digital dentro del marco ético de la Declaración. Su objetivo es abordar algunos temas transversales clave a fin de reflexionar sobre la transición hacia una sociedad en la que la IA ayude a promover el bien común: la gobernanza algorítmica, la alfabetización digital, la inclusión digital de la diversidad y la sostenibilidad ecológica.

¿PARA QUIÉN ES ESTA DECLARACIÓN?

La Declaración de Montreal se dirige a cualquier persona, organización y empresa que desee ser partícipe del desarrollo responsable de la inteligencia artificial, ya sea para contribuir en términos científicos o tecnológicos, preparar proyectos sociales, elaborar normas (reglas, códigos) que sean aplicables, refutar enfoques inadecuados o insensatos, o alertar a la opinión pública cuando sea necesario.

También se dirige a representantes políticos, electos o nombrados, ya que sus ciudadanos esperan que evalúen los cambios sociales que se están produciendo, establezcan rápidamente un marco que permita una transición digital en pos del bien común y prevean los riesgos graves que plantea el desarrollo de la IA.

¿EN QUÉ MÉTODO SE BASA LA DECLARACIÓN?

La Declaración fue el producto de un proceso de deliberación inclusivo que inició el diálogo entre los ciudadanos, los expertos, las autoridades públicas, las partes interesadas de la industria, las organizaciones civiles y las asociaciones profesionales. Este enfoque presenta tres ventajas:

1. **permite mediar conjuntamente en las controversias éticas y sociales de la IA;**
2. **mejora la calidad de las reflexiones sobre la IA responsable;**
3. **fortalece la legitimidad de las propuestas de IA responsable.**

La elaboración de principios y recomendaciones fue una construcción conjunta que contó con la participación de diferentes sectores en espacios públicos, en las salas de juntas de las organizaciones profesionales, en mesas redondas de expertos a nivel internacional, en oficinas de investigación, en aulas o en línea, siempre con el mismo rigor.

¿Y QUÉ SUCEDE LUEGO DE LA DECLARACIÓN?

Dado que la Declaración se refiere a una tecnología que no ha dejado de progresar desde la década de los cincuenta y cuyo ritmo en materia de innovaciones importantes es cada vez más vertiginoso, es fundamental considerar la Declaración como un documento de guía abierto, que se deberá revisar y adaptar conforme a la evolución del conocimiento y de las técnicas, así como la opinión de los usuarios sobre el uso de la IA en la sociedad. Al final del proceso de elaboración de la Declaración, llegamos a un punto de partida para propiciar un debate abierto e inclusivo sobre cómo se utilizarán las tecnologías de la inteligencia artificial en el futuro de la humanidad.

PREÁMBULO

Por primera vez en la historia de la humanidad, se pueden crear sistemas autónomos capaces de llevar a cabo tareas complejas que considerábamos que solo la inteligencia natural era capaz de hacer: procesar grandes volúmenes de información, calcular y predecir, aprender y adaptar las respuestas en función de situaciones cambiantes, y reconocer y clasificar objetos. Puesto que se trata de tareas intangibles, por analogía con la inteligencia humana, diseñamos esta gran variedad de sistemas y los llamamos "inteligencia artificial". La inteligencia artificial constituye una forma importante de progreso científico y tecnológico, que puede generar grandes beneficios sociales al mejorar las condiciones de vida y la salud, agilizar la justicia, crear riquezas, reforzar la seguridad pública, y atenuar el efecto de las actividades humanas en el medio ambiente y el clima. Las máquinas inteligentes no solo hacen mejores cálculos que los humanos; también pueden interactuar con los seres conscientes, acompañarlos y cuidarlos.

No obstante, el desarrollo de la inteligencia artificial plantea desafíos éticos y riesgos sociales importantes. Efectivamente, las máquinas inteligentes pueden restringir las opciones de individuos y grupos, reducir el nivel de vida, perturbar la organización del trabajo y del mercado laboral, influir en la política, entrar en conflicto con los derechos fundamentales, exacerbar las desigualdades sociales y económicas, y afectar a los ecosistemas, al clima y al medio ambiente. Si bien el progreso científico y la vida en sociedad siempre implican un riesgo, les corresponde a los ciudadanos determinar los fines morales y políticos que dan sentido a los riesgos que se corren en un mundo incierto.

Cuanto menores sean los riesgos al hacer uso de la inteligencia artificial, mayores serán sus beneficios. El primer peligro del desarrollo de la inteligencia artificial consiste en creer que podemos controlar

el futuro mediante cálculos. Reducir la sociedad a una serie de números y gobernarla mediante procedimientos algorítmicos es una antigua utopía que sigue impulsando las ambiciones humanas. Pero, en lo que se refiere a los asuntos humanos, el futuro rara vez se asemeja al presente, y los números no pueden determinar qué tiene valor moral, ni qué es lo más conveniente a nivel social.

Los principios de la presente declaración son como los puntos de una brújula moral que orientarán el desarrollo de la inteligencia artificial para alcanzar objetivos convenientes en términos morales y sociales. También constituyen un marco ético que promueve los derechos humanos reconocidos a nivel internacional en los ámbitos que se ven afectados por el lanzamiento de la inteligencia artificial. En su conjunto, los principios planteados sientan las bases para fomentar la confianza de la sociedad en los sistemas de inteligencia artificial.

Los principios de esta declaración se basan en la creencia común de que todos los humanos buscamos crecer como seres sociales con sensaciones, pensamientos y sentimientos, y nos esforzamos por desarrollar nuestro potencial al poner en práctica nuestras capacidades emocionales, morales e intelectuales libremente. Las diferentes partes implicadas en los ámbitos público y privado, y los políticos a nivel local, nacional e internacional son responsables de garantizar que, con el desarrollo y la implementación de la inteligencia artificial, se protejan las capacidades y los objetivos humanos fundamentales, y se contribuya a su plena realización. Con este objetivo en mente, es fundamental interpretar los principios propuestos de manera coherente y, a la vez, tener en cuenta los contextos sociales, culturales, políticos y legales específicos de su aplicación.

1

PRINCIPIO DE BIENESTAR

El desarrollo y el uso de los sistemas de inteligencia artificial (SIA) deben permitir que se cultive el bienestar de todos los seres conscientes.

1. Los SIA deben contribuir a que las personas mejoren sus condiciones de vida, su salud y sus condiciones laborales.
2. Los SIA deben permitir que las personas vayan tras sus deseos, siempre y cuando no causen daño a otros seres conscientes.
3. Los SIA deben permitir que las personas pongan en práctica sus capacidades mentales y físicas.
4. Los SIA no deben generar malestar, a menos que nos permitan alcanzar un bienestar superior que no se podría lograr de otra manera.
5. El uso de los SIA no debe contribuir a un aumento del estrés o la ansiedad, ni debe generar la sensación de que el entorno digital nos acosa.



2

PRINCIPIO DE RESPETO A LA AUTONOMÍA

Los SIA se deben desarrollar y utilizar respetando la autonomía de las personas, a fin de que tengan un mayor control sobre sus vidas y lo que las rodea.

1. Los SIA deben permitir que los individuos alcancen sus objetivos morales y tengan una vida que consideren que vale la pena vivir.
2. No se deben desarrollar ni utilizar los SIA para imponer un estilo de vida específico a las personas, ya sea de forma directa o indirecta, mediante la implementación de mecanismos opresivos de vigilancia y evaluación o de incentivos.
3. Las instituciones públicas no deben utilizar los SIA para promover o desmerecer una concepción particular de buena vida.
4. Es fundamental empoderar a los ciudadanos respecto de las tecnologías digitales al garantizarles el acceso a las formas de conocimiento relevantes, fomentar el aprendizaje de habilidades fundamentales (alfabetización digital y mediática) y promover el desarrollo del pensamiento crítico.
5. No se deben desarrollar los SIA para difundir información no confiable, mentiras o propaganda; se los debe diseñar con el objetivo de contener su difusión.
6. En el desarrollo de los SIA, se debe evitar la creación de dependencias mediante técnicas que capten la atención o imiten las características humanas (apariencia, voz, etc.) que puedan causar confusión entre los SIA y los humanos.



3

Se deben proteger la privacidad y la intimidad contra cualquier intrusión de los SIA y los sistemas de adquisición y archivo de datos (SAAD).

PRINCIPIO DE PROTECCIÓN DE LA PRIVACIDAD Y LA INTIMIDAD

1. Los espacios personales en los que las personas no se encuentran bajo vigilancia o evaluación digital se deben proteger de la intrusión de los SIA y los sistemas de adquisición y archivo de datos (SAAD).
2. Se debe proteger estrictamente la intimidad de los pensamientos y las emociones contra los usos de SIA y SAAD que puedan ocasionar daños, en especial aquellos usos que imponen juicios morales sobre las personas o sus elecciones de vida.
3. Las personas siempre deben tener derecho a desconectarse del ámbito digital en sus vidas privadas, y los SIA deben ofrecer explícitamente la opción para desconectarse periódicamente, y no deben alentar a las personas a que sigan conectadas.
4. Las personas deben tener un gran control sobre la información relativa a sus preferencias. Los SIA no deben crear perfiles de preferencias individuales para influir en el comportamiento de las personas sin su consentimiento libre e informado.
5. Los SAAD deben garantizar la confidencialidad de los datos y el anonimato de los perfiles personales.
6. Toda persona debe tener un gran control sobre sus datos personales, en especial en lo que respecta a su recopilación, uso y difusión. Las personas deben poder acceder a los SIA y los servicios digitales sin tener que perder el control o la propiedad de sus datos personales.
7. Las personas deben tener la posibilidad de elegir donar sus datos personales a organizaciones de investigación a fin de contribuir al avance del conocimiento.
8. Se debe garantizar la integridad de la identidad personal. No se deben utilizar los SIA para imitar o alterar la apariencia, la voz u otras características específicas de una persona a fin de dañar su reputación o manipular a otras personas.

4

El desarrollo de los SIA debe intentar preservar los lazos de solidaridad entre las personas y las generaciones.

PRINCIPIO DE SOLIDARIDAD

1. Los SIA no deben amenazar la preservación de las relaciones humanas afectivas y morales gratificantes, y deben desarrollarse con el objetivo de propiciar dichas relaciones y reducir la vulnerabilidad y el aislamiento de las personas.
2. Los SIA se deben desarrollar con el propósito de colaborar con los humanos en tareas complejas y deben fomentar el trabajo conjunto entre los humanos.
3. No se deben implementar los SIA para reemplazar a las personas en tareas que requieren relaciones humanas de calidad, sino que se los debe desarrollar para facilitar dichas relaciones.
4. En los sistemas de salud donde se utilizan los SIA, se debe tener en cuenta la importancia de las relaciones del paciente con su familia y el personal sanitario.
5. El desarrollo de los SIA no debe fomentar un comportamiento cruel hacia los robots que fueron diseñados para parecerse humanos o animales y actuar como ellos.
6. Los SIA deben permitir que se mejore la gestión de los riesgos y crear las condiciones necesarias para una sociedad con una distribución mutua y más equitativa de los riesgos individuales y colectivos.

5

PRINCIPIO DE PARTICIPACIÓN DEMOCRÁTICA

Los SIA deben cumplir con los criterios de inteligibilidad, capacidad de justificación y accesibilidad, y se los debe someter a un análisis, un debate y un control democráticos.

1. Los procesos de SIA en los que se toman decisiones que afectan la vida, la calidad de vida o la reputación de una persona deben ser inteligibles para sus creadores.
2. Las decisiones que toman los SIA y afectan la vida, la calidad de vida o la reputación de una persona deben poder justificarse en un idioma que entiendan las personas que los usan o se ven afectadas por su uso. La justificación consiste en hacer que los factores y los parámetros que respaldan la decisión sean transparentes y adopten la misma forma que la justificación que le solicitaríamos a una persona que tome una decisión del mismo tipo.
3. Las autoridades públicas y las partes interesadas relevantes siempre deben poder acceder al código de los algoritmos, ya sean públicos o privados, a efectos de verificación y control.
4. Al descubrir SIA con errores, efectos inesperados o no deseables, vulneraciones de seguridad y pérdidas de datos, se los debe informar imperiosamente a las autoridades públicas y las partes interesadas pertinentes, y a aquellos que se ven afectados por la situación.
5. De acuerdo con el requisito de transparencia para las decisiones públicas, todos deben poder acceder al código de los algoritmos que utilizan las autoridades públicas para la toma de decisiones, a excepción de los algoritmos que puedan plantear un gran riesgo de peligro si se los utiliza indebidamente.
6. Para los SIA públicos que afectan de forma significativa la vida de los ciudadanos, estos deben tener la oportunidad y las habilidades para debatir sobre los parámetros sociales de estos SIA, sus objetivos y los límites de su utilización.
7. En todo momento, debemos ser capaces de verificar que los SIA estén haciendo aquello para lo que se los programó y para lo que se los utiliza.
8. Toda persona que utiliza un servicio debe saber si un SIA tomó una decisión que le concierne o la afecta.
9. Todo usuario de un servicio que utiliza un bot de chat debe poder identificar con facilidad si está interactuando con un SIA o una persona real.
10. Las investigaciones sobre inteligencia artificial deben ser abiertas y todos deben poder acceder a ellas.

6

PRINCIPIO DE EQUIDAD

El desarrollo y el uso de los SIA deben contribuir a crear una sociedad justa y equitativa.

1. Se deben diseñar los SIA y entrenarlos de manera tal que no creen, refuercen ni reproduzcan patrones de discriminación basados en diferencias sociales, sexuales, étnicas, culturales o religiosas, entre otras.
2. El desarrollo de los SIA debe ayudar a eliminar las relaciones de dominación entre grupos y personas en función de diferencias de poder, riqueza o conocimiento.
3. El desarrollo de los SIA nos debe beneficiar a todos en el ámbito social y económico al reducir las desigualdades y las vulnerabilidades.
4. El desarrollo de los SIA en el ámbito industrial debe ser compatible con condiciones laborales aceptables en cada etapa de su ciclo de vida, desde la extracción de recursos naturales hasta el reciclaje, lo que incluye el procesamiento de datos.
5. Se debe reconocer la actividad digital de los usuarios de los SIA y los servicios digitales como una labor que contribuye al funcionamiento de los algoritmos y genera valor.
6. Se debe garantizar que todos puedan acceder a los recursos, los conocimientos y las herramientas digitales fundamentales.
7. Debemos respaldar el desarrollo de algoritmos comunes (y de los datos abiertos necesarios para entrenar los SIA) y ampliar su uso, como un objetivo para lograr la equidad social.

7

PRINCIPIO DE INCLUSIÓN DE LA DIVERSIDAD

El desarrollo y el uso de los SIA deben contribuir a preservar la diversidad social y cultural y no debe restringir las posibilidades en lo que respecta a las elecciones de vida o las experiencias personales.

1. El desarrollo y el uso de los SIA no deben conducir a una homogeneización de la sociedad mediante la normalización de los comportamientos y las opiniones.
2. A partir del momento en que se conciben los algoritmos, el desarrollo y la implementación de los SIA deben tener en cuenta la gran cantidad de expresiones de diversidad social y cultural que se encuentran presentes en la sociedad.
3. Los entornos de desarrollo de la IA, ya sea en el ámbito de la industria o la investigación, deben ser inclusivos y dar cuenta de la diversidad de las personas y los grupos de la sociedad.
4. Los SIA deben evitar usar los datos obtenidos para encerrar a las personas en un perfil de usuario, determinar su identidad personal o confinarlas a un filtro burbuja, lo que restringiría y limitaría sus posibilidades de desarrollo personal, en especial en ámbitos tales como la educación, la justicia o los negocios.
5. No se debe desarrollar ni utilizar los SIA para limitar la libre expresión de ideas o la oportunidad de escuchar diferentes opiniones, condiciones que resultan fundamentales para alcanzar una sociedad democrática.
6. Para cada categoría de servicios, se debe diversificar la oferta de los SIA para evitar que se formen monopolios de facto y se socaven las libertades individuales.

8

PRINCIPIO DE PRUDENCIA

Toda persona que esté involucrada en el desarrollo de la IA debe actuar con prudencia y anticipar, en la medida de lo posible, los efectos adversos del uso de los SIA y tomar las medidas adecuadas para evitarlos.

1. Es necesario desarrollar mecanismos que tengan en cuenta las posibilidades de uso doble (beneficioso o perjudicial) de la investigación en IA y el desarrollo de los SIA (ya sean públicos o privados) a fin de limitar los usos perjudiciales.
2. Cuando el uso indebido de un SIA puede poner en peligro la salud o la seguridad públicas y las probabilidades de que ocurra son altas, es prudente limitar el acceso abierto y la difusión pública de su algoritmo.
3. Antes de lanzarlos al mercado, sin importar si son gratuitos o pagos, los SIA deben cumplir requisitos rigurosos de fiabilidad, seguridad e integridad, y se los debe someter a pruebas que no pongan en peligro la vida de las personas, perjudiquen su calidad de vida o afecten de forma negativa su reputación o integridad psicológica. Estas pruebas deben estar abiertas a las autoridades públicas y las partes implicadas pertinentes.
4. El desarrollo de los SIA debe prevenir los riesgos del uso indebido de los datos de los usuarios y proteger la integridad y la confidencialidad de los datos personales.
5. Las instituciones públicas y empresas de sectores que representan un peligro significativo para la integridad de las personas y la organización social deben compartir públicamente y a escala mundial los errores y las fallas que se descubran en los SIA y los SAAD.

9

PRINCIPIO DE RESPONSABILIDAD

El desarrollo y el uso de los SIA no deben contribuir a reducir la responsabilidad de los seres humanos a la hora de tomar decisiones.

1. Solo se puede responsabilizar a los seres humanos por las decisiones que surjan de las recomendaciones de los SIA y las acciones que resulten de ello.
2. En los ámbitos en que se deben tomar decisiones que afectan la vida, la calidad de vida o la reputación de una persona, siempre que el tiempo y las circunstancias lo permitan, un ser humano debe tomar la decisión final y esta decisión debe ser libre e informada.
3. La decisión de matar siempre debe estar a cargo de seres humanos, y no se debe transferir la responsabilidad de esta decisión a un SIA.
4. Las personas que autorizan a un SIA a cometer un crimen o delito, o demuestran negligencia al permitir que un SIA los cometa, son responsables de este crimen o delito.
5. Cuando un SIA provoca daños o perjuicios y se demuestra que el SIA es responsable y se utilizó según lo previsto, no es razonable culpar a las personas involucradas en su desarrollo o uso.

10

El desarrollo y el uso de los SIA se deben llevar a cabo con el objetivo de garantizar una sostenibilidad fuerte del medio ambiente.

PRINCIPIO DE DESARROLLO SOSTENIBLE

1. El hardware de los SIA, su infraestructura digital y los objetos relevantes en los que estos se basan, como los centros de datos, deben intentar alcanzar la mayor eficiencia energética posible y mitigar las emisiones de gas invernadero a lo largo de todo su ciclo de vida.
2. El hardware de los SIA, su infraestructura digital y los objetos relevantes en los que estos se basan deben intentar generar la menor cantidad posible de residuos eléctricos y electrónicos, y contar con procedimientos de mantenimiento, reparación y reciclaje de acuerdo con los principios de la economía circular.
3. El hardware de los SIA, su infraestructura digital y los objetos relevantes en los que estos se basan deben minimizar nuestro impacto en los ecosistemas y la biodiversidad en todas las etapas de su ciclo de vida, en particular con respecto a la extracción de recursos y la disposición final de los equipos al final de su vida útil.
4. Los actores públicos y privados deben respaldar el desarrollo responsable con el medio ambiente de los SIA para hacer frente al despilfarro de recursos naturales y bienes producidos, crear cadenas de suministro y un comercio sostenibles, y reducir la contaminación mundial.

GLOSARIO

Actividad en línea

La actividad en línea hace referencia a todas las actividades que lleva a cabo una persona en un entorno digital, ya sea a través de una computadora, un teléfono o cualquier objeto conectado.

Alfabetización digital

La alfabetización digital de una persona hace referencia a su capacidad para acceder, administrar, comprender, integrar, comunicar, evaluar y crear información de manera segura y adecuada mediante las herramientas digitales y las tecnologías de red a fin de participar en la vida económica y social.

Algoritmo

Un algoritmo es un método para resolver un problema mediante una serie de operaciones finitas y precisas. Concretamente, en el contexto de la inteligencia artificial, se refiere a una serie de operaciones que se aplican para introducir datos a fin de lograr el resultado deseado.

Aprendizaje automático

El aprendizaje automático es la rama de la inteligencia artificial que consiste en programar un algoritmo para que pueda aprender por sí solo.

Es posible clasificar las diferentes técnicas en tres tipos principales de aprendizaje automático:

- > En el aprendizaje supervisado, el sistema de inteligencia artificial (SIA) aprende a predecir un valor a partir de los datos que se le introdujeron. Para esto, durante el entrenamiento, se requieren pares anotados que consisten en una entrada y un valor. Por ejemplo, un sistema puede aprender a reconocer un objeto que aparece en una imagen.

- > En el aprendizaje no supervisado, el SIA aprende a encontrar similitudes entre datos que no están anotados, por ejemplo, para dividirlos en varios grupos homogéneos. De este modo, un sistema puede reconocer comunidades de usuarios de redes sociales.
- > Mediante el aprendizaje por refuerzos, el SIA aprende a actuar en su propio entorno a fin de maximizar la recompensa que recibe durante el entrenamiento. Esta es la técnica mediante la cual los SIA pudieron vencer a los humanos en el juego Go o en el videojuego *Dota 2*.

Aprendizaje profundo

El aprendizaje profundo es la rama del aprendizaje automático que usa redes neuronales artificiales en muchos niveles. Es la tecnología que se encuentra detrás de los últimos avances de la IA.

Bienes comunes digitales

Los bienes comunes digitales son las aplicaciones o los datos que produce una comunidad. A diferencia de los bienes materiales, se los puede compartir fácilmente y no se deterioran con el uso. Por lo tanto, a diferencia del software propietario, se considera el software de código abierto (que suele ser el fruto de la colaboración de diferentes programadores) como bienes comunes digitales, ya que su código fuente es abierto y todos pueden acceder a él.

Bot de chat

Un bot de chat es un sistema de IA que puede conversar con un usuario en una lengua natural.

Capacidad de justificación de las decisiones

Se justifica la decisión de un SIA cuando hay razones no triviales que motivan esta decisión, y se debe poder comunicar dichas razones en lengua natural.

Datos abiertos

Los datos abiertos son datos digitales a los cuales los usuarios pueden acceder libremente. Por ejemplo, este es el caso de la mayoría de los resultados de las investigaciones publicadas sobre IA.

Datos personales

Los datos personales son aquellos que ayudan a identificar a una persona directa o indirectamente.

Dependencia del camino

Mecanismo social a través del cual las decisiones tecnológicas, organizativas o institucionales — que antes se consideraban racionales, pero ahora son obsoletas— siguen influyendo en la toma de decisiones. Un mecanismo que se mantiene debido al sesgo cognitivo o porque, para llevar a cabo un cambio, se requeriría demasiado dinero o esfuerzo. Tal es el caso de la infraestructura vial urbana: en lugar de implementar un cambio para reducir significativamente las emisiones de carbono, se llevan adelante programas de optimización del tráfico. Se debe estar al tanto de este mecanismo al utilizar la IA para proyectos especiales, ya que, en el aprendizaje supervisado, los datos de entrenamiento pueden reforzar antiguos paradigmas organizacionales que ahora son cuestionados.

Desarrollo sostenible

El desarrollo sostenible hace referencia a un desarrollo de la sociedad que sea compatible con la capacidad que tienen los sistemas naturales para ofrecer los recursos y los servicios necesarios a dicha sociedad. Se trata de un desarrollo económico y social que satisface las necesidades actuales sin poner en riesgo a las generaciones venideras.

Desconexión digital

La desconexión digital se refiere al cese temporal o permanente de la actividad en línea de una persona.

Efecto rebote

El efecto rebote es el mecanismo a través del cual una mayor eficiencia energética o un mejor comportamiento ambiental de los bienes, los equipos y los servicios da lugar a un aumento más que proporcional en su uso. Por ejemplo, el tamaño de las pantallas aumenta, la cantidad de dispositivos electrónicos en los hogares también, y se recorren mayores distancias en automóvil o avión.

El resultado global es una mayor presión en los recursos y el medioambiente.

Entrenamiento

El entrenamiento es el proceso de aprendizaje automático mediante el cual los SIA construyen un modelo a partir de datos. El desempeño de los SIA depende de la calidad del modelo, que depende de la cantidad y la calidad de los datos utilizados durante el entrenamiento.

Fiabilidad

Un SIA es fiable cuando lleva a cabo la tarea para la cual fue diseñado de la manera esperada. La fiabilidad es la posibilidad de éxito que oscila entre el 51 % y 100 %, valores estrictamente superiores al azar. Cuanto más fiable es un sistema, más previsible es su comportamiento.

Filtro burbuja

La expresión "filtro burbuja" hace referencia a la información "filtrada" que le llega a una persona en internet. Diferentes servicios, como las redes sociales o los motores de búsqueda, ofrecen resultados personalizados para los usuarios. Esto puede aislar a las personas (en "burbujas") porque ya no pueden acceder a la información común.

Inteligencia artificial (IA)

La inteligencia artificial (IA) hace referencia a una serie de técnicas que permiten que una máquina imite el aprendizaje humano, en concreto para aprender, hacer predicciones, tomar decisiones y percibir lo que la rodea. En el caso de un sistema informático, la inteligencia artificial se aplica a los datos digitales.

Inteligibilidad

Se considera que un SIA es inteligible cuando un ser humano que cuenta con los conocimientos necesarios puede comprender sus operaciones, es decir, su modelo matemático y los procesos que lo determinan.

RGAs

La sigla de "red generativa antagónica". En una RGA, dos redes antagónicas compiten para generar una imagen. Por ejemplo, se las puede utilizar para crear una imagen, una grabación o un video que es posible que cualquier persona acepte como reales.

Sistema de adquisición y archivo de datos (SAAD)

Un SAAD hace referencia a todo sistema informático que pueda recopilar y registrar datos. En última instancia, estos datos se utilizan para entrenar los sistemas de IA o como parámetros para la toma de decisiones.

Sistema de inteligencia artificial (SIA)

Un SIA es cualquier sistema informático que use algoritmos de inteligencia artificial, ya sea que se trate de un software, un objeto conectado o un robot.

Sostenibilidad fuerte del medio ambiente

La noción de sostenibilidad fuerte del medio ambiente se remonta a la idea de que, para ser sostenible, el ritmo de consumo de los recursos naturales y las emisiones contaminantes debe ser compatible con los límites ambientales del planeta, el ritmo de renovación de los recursos y del ecosistema, y la estabilidad del clima.

A diferencia de la sostenibilidad débil, que requiere un esfuerzo menor, la sostenibilidad fuerte no permite sustituir la pérdida de recursos naturales con capital artificial.

EQUIPO

La redacción de la Declaración de Montreal para un desarrollo responsable de la inteligencia artificial es el fruto del trabajo de un equipo científico multidisciplinario e interuniversitario que se basa en un proceso de consulta ciudadana y en el diálogo con expertos y partes interesadas en el desarrollo de la IA.

Christophe Abrassart: profesor adjunto de la Escuela de Diseño y codirector de Lab Ville Prospective de la Facultad de planeamiento de la Universidad de Montreal, miembro de Centre de recherche en éthique (CRÉ).

Yoshua Bengio: profesor titular del Departamento de Ciencias de la Computación e Investigación de Operaciones, Universidad de Montreal, director científico de MILA e IVADO.

Guillaume Chicoisne: director de programas científicos, IVADO.

Nathalie de Marcellis-Warin: profesora titular, Escuela Politécnica de Montreal; presidenta y directora ejecutiva, Centro Interuniversitario de Investigación y Análisis de las Organizaciones (CIRANO).

Marc-Antoine Dilhac: profesor adjunto, Departamento de Filosofía, Universidad de Montreal; presidente del Grupo de Ética y Política, Centre de recherche en éthique (CRÉ); titular de la Cátedra de Investigación de Canadá en Ética Pública y Teoría Política; director de Institut Philosophie Citoyenneté Jeunesse.

Sébastien Gambs: profesor de informática de la Universidad de Quebec en Montreal; titular de la Cátedra de Investigación de Canadá en Preservación de la Privacidad y Análisis Ético de Macrodatos.

Vincent Gauthrais: profesor titular, Facultad de Derecho, Universidad de Montreal; director de Centre de recherche en droit public (CRDP); titular de la cátedra L.R. Wilson en Tecnología de la Información y Derecho del Comercio Electrónico.

Martin Gibert: asesor ético en IVADO e investigador en Centre de recherche en éthique (CRÉ).

Lyse Langlois: profesora titular y vicedecana de la Facultad de Ciencias Sociales; directora de Institut d'éthique appliquée (IDÉA); investigadora del Centro Interuniversitario de Investigación sobre la Globalización y el Trabajo (CRIMT).

François Laviolette: profesor titular del Departamento de Ciencias de la Computación e Ingeniería de Software, Universidad Laval; director de Centre de recherche en données massives (CRDM).

Pascale Lehoux: profesora titular de la Escuela de Salud Pública de la Universidad de Montreal (ESPU); titular de la cátedra Innovación Responsable en Salud.

Jocelyn Maclure: profesor titular, Facultad de Filosofía, Universidad Laval; presidente de la Comisión de Ética en Ciencia y Tecnología de Quebec (CEST).

Marie Martel: profesora de la Escuela de Biblioteconomía y Ciencias de la Información, Universidad de Montreal.

Joëlle Pineau: profesora titular, Escuela de Ciencias de la Computación, Universidad McGill; directora del Laboratorio de IA de Facebook en Montreal; codirectora de Reasoning and Learning Lab.

Peter Railton: Gregory S. Kavka Distinguished University Professor; John Stephenson Perrin Professor; Arthur F. Thurnau Professor, Departamento de Filosofía, Universidad de Michigan; miembro de la Academia Estadounidense de las Artes y las Ciencias.

Catherine Régis: profesora adjunta, Facultad de Derecho, Universidad de Montreal; titular de la Cátedra de Investigación de Canadá en Cultura Colaborativa en Derecho y Políticas Sanitarias; investigadora, Centre de recherche en droit public (CRDP).

Christine Tappolet: profesora titular, Departamento de Filosofía, Universidad de Montreal; directora de Centre de recherche en éthique (CRÉ).

Nathalie Voarino, estudiante del Doctorado en Bioética, Universidad de Montreal.

COLABORADORES

Université 
de Montréal



CENTRE DE RECHERCHE EN ETHIQUE



ICRA
Programme
IA et
société



Québec 
Fonds de recherche – Nature et technologies
Fonds de recherche – Santé
Fonds de recherche – Société et culture



