



Retos sociales y éticos del metaverso

Se abre el debate

Richard Benjamins | Chief AI & Data Strategist Telefónica

Yaiza Rubio | Chief Metaverse Officer Telefónica

Chema Alonso | Chief Digital Officer Telefónica

1. Resumen.....	3
2. Introducción	3
3. Cómo definir el metaverso	4
3.1 Ingredientes clave del metaverso	5
4. Una visión pragmática del metaverso	6
4.1 Evoluciones clave.....	6
4.2 Los diferentes papeles que pueden desempeñar las empresas en el metaverso	8
5. Retos sociales y éticos del metaverso	8
5.1 Aprender de los retos asociados a la inteligencia artificial.....	8
5.1.1 Retos directos	9
5.1.2 Desafíos indirectos.....	12
5.2 Nuevos retos	13
5.2.1 Problemas de salud, incluidos los mentales	13
5.2.2 Desafíos a la privacidad.....	14
5.2.3 Desafíos a la libertad	14
5.2.4 Desafíos del mundo dual	15
5.2.5 Desafíos a la igualdad.....	15
5.3 Desafíos derivados de un uso malicioso deliberado del metaverso	16
5.4 Enfoques para afrontar los desafíos y mitigar los riesgos sociales y éticos...	16
6. Conclusión	17
6.1 Hacia los valores y principios éticos del metaverso	18
7. Reconocimientos.....	19

1. Resumen

El metaverso está atrayendo una enorme atención como la próxima versión de Internet, y las empresas están empezando a explorar las muchas y nuevas oportunidades de negocio que ofrece. Sin embargo, basándonos en la experiencia con la inteligencia artificial, sabemos que también existen posibles consecuencias éticas y sociales negativas del uso masivo de la tecnología que deben ser tratadas. En este libro blanco describimos brevemente qué es el metaverso, qué tecnologías forman parte de su ecosistema y de dónde procede. A continuación, nos centramos en los posibles riesgos sociales y éticos del metaverso. Afirmamos que las empresas que están aplicando el uso responsable de la IA están bien preparadas para los riesgos sociales y éticos del metaverso. No porque conozcan el futuro, sino porque cuentan con la gobernanza y la cultura adecuadas para afrontar esos riesgos.

2. Introducción

¿Quién no ha oído hablar del Metaverso? El Internet del futuro con una experiencia inmersiva en 3D donde vivimos en mundos virtuales a través de avatares. Un mundo digital completamente integrado con nuestro mundo físico. Hay mucho debate y opiniones sobre lo que es, o mejor, será el metaverso, o las aplicaciones que ofrecerá. Y al igual que Internet hace 30 años, nadie sabe lo importante que será, pero la mayoría coincide en que es algo que hay que observar con atención. Piense en una nueva forma de teletrabajo en la que pueda sentarse virtualmente con sus colegas en la misma sala y mantener conversaciones e interacciones como si fuera real. O sobre nuevas experiencias de marketing, como conducir un coche que quizás quiera comprar desde su casa, eligiendo rápidamente entre diferentes alternativas. O tratamientos psicológicos para ayudar a curar a las personas con fobias. Y nuevas formas de inmersión en la educación o en la formación de instrucciones complejas, donde los estudiantes pueden "vivir" en la Edad Media, y los ingenieros técnicos pueden practicar costosas operaciones de mantenimiento sin riesgos. El metaverso podría ser tan grande como Internet, con enormes oportunidades, aunque en su mayoría todavía desconocidas. Incluso ya existe una organización de estándares (Metaverse Standards Forum) como la de la web (Worldwide Web Consortium, W3C).

Las oportunidades del metaverso abundan, y aunque creemos que los beneficios potenciales superarán con creces los riesgos, también es importante reflexionar, con antelación, sobre esos posibles riesgos, con el objetivo de mitigarlos antes de que se produzcan. El objetivo de este libro blanco es revisar los riesgos éticos y sociales que pueden asociarse al metaverso, y orientar sobre cómo las organizaciones pueden prepararse para los mismos. Hay otros riesgos, como los jurídicos, fiscales, delitos y de propiedad intelectual, que quedan fuera del ámbito de este libro blanco, aunque podemos referirnos a ellos en caso de que tengan un aspecto social importante.

El metaverso no ha hecho más que empezar y esto nos da una oportunidad única de construir un metaverso en el que queramos vivir o con el que queramos vivir. Tendrá un profundo impacto en nuestras vidas y, por tanto, es muy importante —desde el principio— pensar en las posibles consecuencias y actuar en consecuencia. Esto ayudará a dejar de lado el enfoque actual de "romper, pedir disculpas y arreglar" para pasar a uno más proactivo.

3. Cómo definir el metaverso

Mucha gente asocia el metaverso con la experiencia inmersiva en 3D de los videojuegos. Y aunque esto está claramente incluido, no es únicamente eso. Será mucho más. Al igual que era difícil definir Internet hace 30 años, también es difícil definir de antemano lo que será el metaverso. Parisi (2021), en su publicación ampliamente referenciada¹, habla sobre siete reglas del metaverso, que ayudan a definir lo que es, pero también, lo que no es:

- Solo hay un metaverso.
- El metaverso es abierto.
- El metaverso es para todos.
- Nadie controla el metaverso.
- El metaverso es una red.
- El metaverso es independiente del hardware.
- El metaverso es Internet.

Esta definición del Grupo de Análisis² es un buen intento de definir el metaverso de forma descriptiva. *"El metaverso, tal como se concibe hoy, es considerado por muchos como un "sucesor" de Internet. Aunque no existe una definición consensuada del metaverso, una forma de concebirlo es como una red expansiva de espacios digitales, incluidas las experiencias inmersivas en 3D en realidad aumentada, virtual y mixta, que están interconectados y son interoperables para que puedas moverte fácilmente entre ellos, y en los que puedes crear y explorar con otras personas que no están en el mismo espacio físico que tú. Algunos se han referido al metaverso como un "Internet personificado" en el que los individuos se sentirán como si estuvieran realmente "presentes" en las experiencias y no simplemente observándolas a través de sus pantallas. Esto significa que la interacción con Internet (y los dispositivos que proporcionan acceso a Internet) tiene el potencial de ser mucho más naturales, incorporando modos de comunicación que incluyan el gesto y la voz, de manera que*

¹ <https://medium.com/meta-verses/the-seven-rules-of-the-metaverse-7d4e06fa864c>

² <https://www.analysisgroup.com/globalassets/insights/publishing/2022-the-potential-global-economic-impact-of-the-metaverse.pdf>

los individuos no se limiten a teclear o a pulsar. Además, se prevé que el metaverso pueda albergar casi todas las actividades en las que participamos actualmente (por ejemplo, socializar, trabajar, aprender, entretenerse, comprar, crear contenidos, etc.) y que también haga posible nuevos tipos de actividades".

Algunos ejemplos adicionales de aplicaciones en el metaverso descritos en otro lugar: un espacio en el que se podrá ir de compras, jugar, reunirse con amigos, asistir a conciertos, trabajar y, en general, construir una vida virtual; revolucionar la forma de trabajar, con avatares ultrarrealistas y salas de conferencias virtuales llenas de personal internacional; de forma similar a la propiedad y los bienes inmuebles tradicionales, los propietarios virtuales pueden cobrar cuotas de alquiler a través de condiciones predeterminadas que se negocian con los inquilinos y se aplican mediante contratos inteligentes; regalías que los creadores pueden obtener al vender o revender los NFT en los mercados secundarios; organizaciones autónomas descentralizadas (DAO), que representan un conjunto de personas que actúan juntas a través de reglas bien definidas y codificadas. Las DAO pueden actuar como prestamistas de la comunidad, creando la próxima generación de servicios financieros descentralizados, transformando sectores económicos como la educación, la sanidad, la producción, la formación laboral, las comunicaciones, el entretenimiento y el comercio minorista.

3.1 Ingredientes clave del metaverso

El metaverso es y será posible gracias a las grandes innovaciones en hardware, interfaz persona-ordenador, infraestructura de red, herramientas de creación, y economías digitales, incluyendo, entre otras, la realidad virtual y aumentada, la inteligencia artificial, los datos, las pantallas holográficas, el IdC (internet de las cosas), el despliegue de fibra, el 5G autónomo, la edge computing, la cadena de bloques, los contratos inteligentes, tokenomics/NFT, el IFPS, la ciberseguridad y la "softwarización" de las redes³.

El metaverso, como cualquier sociedad, necesita una economía funcional. En una sociedad virtual, la economía depende de la verificación de las propiedades digitales, como la casa, el coche, la granja, los libros, la ropa y los muebles del metaverso. Para prosperar, también necesita la capacidad de viajar y comerciar libremente entre reinos del metaverso que pueden tener leyes y reglas diferentes⁴. Los tokens no fungibles (NFT), como los registros de propiedad digital almacenados en una cadena de bloques, serán una parte esencial de la economía del metaverso al permitir la verificación de las posesiones, la propiedad e incluso la identidad. Dado que cada NFT

³ <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/telecoms-future-in-the-web3-era-jose-maria-alvarez-pallete-lopez>

⁴ <https://www.ft.com/partnercontent/crypto-com/nfts-the-metaverse-economy.html>

está asegurado por una clave criptográfica que no puede ser borrada, copiada o destruida, permite la verificación robusta y descentralizada —de la propia identidad virtual y de las posesiones digitales— necesaria para que la sociedad del metaverso tenga éxito e interactúe con otras sociedades del metaverso. Los NFT, por tanto, permiten las transacciones de propiedad que impulsan el metaverso.

Otro ingrediente clave del metaverso serán los datos, dado el aumento previsto de la actividad e interactividad virtuales. Los datos podrán flotar entre los diferentes reinos del metaverso en lugar de permanecer en silos de la empresa como en la Web 2.0. La propiedad de los datos de los usuarios se transferirá de las empresas (de vuelta) a los usuarios, que decidirán soberanamente lo que ocurrirá con sus datos (explotación, eliminación, portabilidad, acceso, etc.) de forma que se preserve la privacidad. La inteligencia artificial y el aprendizaje automático se utilizarán de forma masiva para operar y dirigir el metaverso y, al mismo tiempo, optimizar la experiencia de usuario; incluso más de lo que ocurre hoy en día.

4. Una visión pragmática del metaverso

Como nadie sabe cómo será el metaverso, para avanzar, ser pragmático puede ayudar: en lugar de esperar lo que será el futuro, es mejor ayudar a crearlo. Creemos que la experiencia inmersiva en 3D es solo la punta del iceberg del metaverso. Una forma de entender la génesis del metaverso es considerarlo como la culminación de varias evoluciones importantes que han tenido, y siguen teniendo, lugar en las últimas décadas en, respectivamente, *la interfaz de usuario*, *la programación*, *la conectividad* y *la economía de Internet*.

4.1 Evoluciones clave

Interfaz de usuario. En los últimos 20 o 30 años, el paradigma de la interfaz de usuario ha pasado gradualmente de la adaptación de las personas a la tecnología a la adaptación de la tecnología a las personas. Ha evolucionado desde el texto simple hasta el texto enriquecido, pasando por las interfaces gráficas mediante ventanas, los navegadores y las aplicaciones, y la voz mediante interfaces cognitivas. Las interfaces 3D no son más que la siguiente evolución del mismo concepto que explota nuevas tecnologías (RV, RA, RM) y capacidades (p. ej., HPC, nube y conectividad de alta velocidad). En el futuro, además de la vista y el oído, la interfaz de usuario podrá incluir elementos de los otros tres sentidos, como el tacto (en cierta medida ya disponible a través de las interfaces hápticas), el gusto y el olfato. ¿Se imagina "imprimir" el olor de un bosque o el sabor de un café con leche?

La **programación** (codificación) ha pasado de los lenguajes de procedimiento simples a las reglas complejas y los enfoques orientados a objetos, en los que se escriben instrucciones claras (algoritmos) para "decir" a las máquinas lo que deben hacer. Las capacidades de la programación visual (sin código/*low code*) también han aumentado considerablemente en los últimos años. Y hoy es la inteligencia artificial / el aprendizaje automático el que "escribe" esas instrucciones para la máquina, basadas en el procesamiento de enormes cantidades de datos, para llevar a cabo tareas complejas. Algunos grandes modelos de IA son incluso capaces de convertir las instrucciones de voz en código de programación formal, como Copilot o GPT-3.

La **conectividad** también ha experimentado importantes evoluciones, empezando por las conexiones de datos digitales que se conectaban a *hosts*, y pasando a la arquitectura cliente-servidor a través de las LAN. Después llegó Internet con su arquitectura de tres niveles de ISP. La era de los puntocom se vio favorecida por la creciente disponibilidad de DSL y fibra de banda ancha, y la Web 2.0 asistió a la evolución de la conectividad móvil de 2G a 3G y 4G. La Web3, que será la base del metaverso, vivirá otra evolución que incluirá Wifi6, 5GSA, FTTR, edge computing, *slicing* de la red, SDN y NaaS (la red como servicio).

El **modelo económico** de la primera web (web 1.0) se basaba en que los editores publicaban contenidos y los consumidores los consumían. El motor económico que lo movía todo era el contenido de pago o la publicidad. La Web 2.0 era (y es) la red de contenidos generados por los usuarios, siendo las redes sociales el fenómeno más popular. Pero también se han disparado otros servicios como el vídeo en línea y el comercio electrónico. La actividad se concentra en unos pocos actores (p. ej., las FANG⁵) que dominan el mercado de forma monopolística. Aunque los usuarios contribuyen significativamente al éxito de esas empresas, no participan plenamente de los beneficios. Además, los datos personales de los usuarios se explotan a gran escala para mejorar los servicios y los ingresos publicitarios. La Web3 promete cambiar este modelo dando a los usuarios más beneficios del ecosistema en el que participan. Además de ser consumidores de los servicios Web3, los usuarios pueden convertirse en propietarios y responsables. Nacen las nociones de tokenomics y organizaciones autónomas descentralizadas (DAO).

Juntas, estas cuatro evoluciones conducen naturalmente a lo que se denomina el metaverso: una nueva Internet con una experiencia mejorada y la posibilidad de difuminar las fronteras de poder y económicas entre empresas y usuarios.

⁵ FANG: Facebook, Amazon, Netflix, Google.

4.2 Los diferentes papeles que pueden desempeñar las empresas en el metaverso

Como ocurre con cualquier nueva tecnología, las empresas deben decidir qué papel(es) quieren desempeñar en el metaverso. Pueden intentar construir grandes mundos virtuales y atraer a cantidades masivas de usuarios (probablemente el objetivo de empresas como Meta y Microsoft). Pueden construir componentes tecnológicos del metaverso (por ejemplo, empresas emergentes de realidad virtual o aumentada, NFT, Inteligencia Artificial o cadena de bloques). Pueden crear facilitadores para que otras empresas aprovechen el metaverso (p. ej., conectividad de alta velocidad para que las familias y los pacientes hospitalizados se reúnan de forma realista). Pueden utilizar el metaverso como un nuevo canal para interactuar con sus clientes, como experiencias de marketing y ventas (ver cómo le queda la ropa, cómo se adaptan los muebles nuevos a su casa, conducir ese coche nuevo, etc., pero también la posventa del cliente, como las reclamaciones y las reparaciones). Aplicando la noción de tokenomics, las empresas pueden involucrar a sus clientes y grupos de interés en nuevas formas de su negocio proporcionando una distribución más equilibrada de los beneficios.

De este modo, las empresas pueden participar activamente en el desarrollo del metaverso —cada una a su manera— aprendiendo mucho durante el proceso, en lugar de adoptar una actitud de espera.

5. Retos sociales y éticos del metaverso

Los retos a los que se enfrenta el metaverso son de índole muy diversa: técnicos, jurídicos, empresariales, fiscales, de cadena de valor, de seguridad, de experiencia de usuario y económicos. Además, los riesgos pueden provenir de acciones malintencionadas deliberadas —normalmente prohibidas— o de consecuencias imprevistas de acciones supuestamente inocentes. Este libro blanco se centra sobre todo en los retos sociales y éticos del metaverso derivados de las consecuencias imprevistas de acciones inocuas, pero también introduce algunos desafíos derivados del uso malicioso deliberado del metaverso.

5.1 Aprender de los retos asociados a la inteligencia artificial

Hay pocas tecnologías TIC que hayan provocado más debates sobre los retos éticos y sociales que la inteligencia artificial. Dado que muchos de los riesgos asociados a la IA podrían incluso estar más presentes en el metaverso, es un buen lugar para empezar. Distinguimos entre retos sociales y éticos *directos e indirectos*. "Directas" se refiere a las consecuencias atribuibles e inmediatas de un determinado uso de la IA o del

metaverso, como la discriminación no deseada. "Indirectas" se refiere a las consecuencias que no son obvias ni están previstas al principio.

5.1.1 Retos directos

Hay dos tipos de retos directos de la IA, y por lo tanto también del metaverso: los que las empresas pueden manejar individualmente, y los que requieren la participación del gobierno.

Desafíos a los que deben hacer frente las organizaciones individuales

Entre los desafíos que deben afrontar las organizaciones individuales se encuentran los sesgos y la discriminación no deseada, la explicabilidad y los algoritmos de caja negra, la intervención humana y la adecuada autonomía del sistema, la privacidad, la seguridad y la protección, y la huella de carbono de los algoritmos de IA. Si una organización toma medidas en relación con esos desafíos, podría ser posible prevenirlos o mitigarlos. Pero si esos desafíos ni siquiera se tienen en cuenta, la probabilidad de que se produzca un impacto negativo será mayor. A continuación, explicaremos brevemente cada uno de esos desafíos y su relación con el metaverso.

Sesgos y discriminación no deseada. Se han escrito cientos de artículos sobre este tema, en el que los sistemas de inteligencia artificial, basados en los sesgos, podrían dar lugar a una discriminación no deseada de grupos vulnerables. Las personas de color son tratados de forma diferente a las personas blancas en algunos sistemas judiciales estadounidenses⁶, las mujeres reciben menos préstamos de los bancos⁷ y son menos contratadas por las empresas⁸ solo por su género, etc. El metaverso estará lleno de aplicaciones que utilizan la inteligencia artificial para la predicción y la clasificación, por lo que este desafío es también un reto para el metaverso.

Explicabilidad y algoritmos de caja negra. Los sistemas de IA pueden tomar decisiones sin intervención humana y se espera que el metaverso esté repleto de estas decisiones autónomas. Actualmente, los algoritmos de aprendizaje profundo proporcionan los mejores resultados, pero su funcionamiento interno es muy difícil de entender para las personas. Si las decisiones tienen un impacto significativo en la vida de las personas, entonces tenemos que entender cómo el algoritmo llega a una determinada conclusión o decisión; necesitamos algoritmos de caja blanca.

Intervención humana y autonomía adecuada del sistema. Cuando las decisiones tienen un impacto en la vida de las personas, es posible que no queramos dejar todo en manos de una máquina. Por ejemplo, no querríamos diagnosticar a un paciente con una enfermedad grave sin la participación de un profesional médico, especialmente si

⁶ <https://www.propublica.org/datastore/dataset/compas-recidivism-risk-score-data-and-analysis>

⁷ <https://www.nytimes.com/2019/11/10/business/Apple-credit-card-investigation.html>

⁸ <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>

esto implica tratamientos intrusivos. De nuevo, muchos algoritmos de IA se ejecutarán en el metaverso, y es importante tomar decisiones explícitas sobre el grado de autonomía que asignamos a un algoritmo.

Privacidad. La privacidad es un reto para cualquier sistema digital accesible a través de Internet. La inteligencia artificial ha aumentado este reto porque los datos, que incluyen los datos personales, son el combustible de la IA. Algunos modelos de negocio explotan la inteligencia artificial para analizar enormes cantidades de datos personales, a veces de forma poco transparente o incluso ilegal. En el metaverso, se espera que este desafío se agrave debido a la creciente oportunidad de registrar cualquier tipo de interacción digital y virtual que implique los datos personales.

Seguridad y protección. Los algoritmos de la inteligencia artificial nunca son correctos al 100%. Por lo tanto, es importante evaluar de antemano los posibles daños a la seguridad y la protección cuando los algoritmos cometen errores, y actuar en consecuencia para evitarlos o mitigarlos. Los sistemas de IA de aprendizaje profundo son vulnerables a los llamados ataques adversarios, y algunos hackers lograron engañar al sistema Autopilot de Tesla para que confundiera una señal de stop con una de límite de velocidad⁹. Además del reto de la IA, el metaverso introduce toda una serie de nuevos riesgos de seguridad y protección. Piense, por ejemplo, en las personas que mueven enérgicamente los brazos y las piernas mientras llevan puestas las gafas 3D.

Huella de carbono de los algoritmos. Recientemente se ha identificado un reto para la IA que está relacionado con el consumo de energía de los grandes modelos de lenguaje natural. Los modelos como el GPT-3 consumen millones de dólares en electricidad y, por tanto, tienen una importante huella de carbono. Esto ha dado lugar al término "IA verde", en el que la consideración del consumo de electricidad es una parte integral del diseño del algoritmo. Los NFT se basan en la cadena de bloques, que utiliza algoritmos de consenso para añadir nuevos bloques a la cadena. Todos sabemos por Bitcoin que algunos de esos algoritmos incentivan el consumo de electricidad, en particular el algoritmo de consenso *proof of work*: cuanto más trabajo de "minería" se realice, mayor será la probabilidad de ser compensado. Otro tipo de algoritmo de consenso llamado *proof of stake* es mucho menos intensivo en energía y, por lo tanto, es preferible para los NFT a gran escala que sustentarán la Web3 y el metaverso. Además, acercar la computación a la aplicación final a través del *edge computing* también puede reducir la huella de carbono informática necesaria para las experiencias inmersivas en el metaverso. Esperamos ver más el término "metaverso verde" en el futuro.

⁹ <https://medium.com/self-driving-cars/adversarial-traffic-signs-fd16b7171906>

Desafíos a los que deben hacer frente los gobiernos

Entre los desafíos que deben afrontar los gobiernos se encuentran las noticias falsas y los *deep fakes*, la relación entre las personas y las máquinas, los derechos de autor, la concentración de datos y riqueza, el aumento de la desigualdad, y la IA para la defensa y la guerra. Para prevenir o mitigar dichos desafíos, se requieren acciones gubernamentales. Las empresas no pueden afrontar satisfactoriamente esos retos a nivel individual.

Noticias falsas y deep fakes (videos falsos). La inteligencia artificial es relevante en dos aspectos para las noticias falsas y las *deep fakes*. En primer lugar, mediante el uso de grandes modelos de lenguaje natural y el aprendizaje profundo es posible, respectivamente, generar automáticamente noticias a partir de algunas palabras clave y vídeos falsos de personas que dicen cosas que no han dicho. Si hay una mala intención, se pueden producir grandes cantidades de noticias falsas y *deep fakes* con poco esfuerzo. En segundo lugar, los algoritmos de recomendación pueden difundir noticias falsas y *deep fakes* de forma viral, llegando así a millones de usuarios. En el metaverso, este riesgo se amplifica porque la experiencia inmersiva hace más difícil distinguir entre el contenido real y el falso.

La relación entre las personas y las máquinas. ¿Puede una máquina ser tu jefe? ¿Pueden las personas casarse con robots? ¿Hasta qué punto pueden los robots cuidar de las personas mayores? ¿Son necesarias unas normas que regulen la relación entre las personas y las máquinas? Todas esas preguntas, y otras similares, cobran aún más importancia en el metaverso debido a los avatares. ¿Pueden los avatares entablar una relación formal y legal? ¿Puede un avatar ser responsable de una acción en contra de otro avatar?

Derechos de autor. Si bien la noción de derechos de autor es un concepto jurídico, resuena de manera especial en la sociedad en general, como lo demuestran las quejas airadas de los artistas en un concurso de arte que se ganó con la importante ayuda de un sistema de IA¹⁰. ¿Quién tiene los derechos de autor de una obra de arte generada por un sistema de IA, como una canción, un cuadro generado por Dalle2¹¹, un texto de LaMDA¹² o incluso un código de programación de Copilot¹³? El metaverso constará de millones de líneas de código y albergará millones de contenidos. ¿Generados por personas, por IA?

La concentración de datos y riqueza y la creciente desigualdad. Actualmente, la inteligencia artificial no está ayudando a reducir la brecha entre ricos y pobres, tanto entre distintos países como dentro de los mismos. Algunas grandes empresas están

¹⁰ <https://edition.cnn.com/2022/09/03/tech/ai-art-fair-winner-controversy/index.html>

¹¹ <https://openai.com/dall-e-2/>

¹² <https://blog.google/technology/ai/lamda/>

¹³ <https://github.com/features/copilot>

concentrando una parte importante de todos los datos digitales del mundo y los utilizan para entrenar y construir los mejores algoritmos de IA. A nivel macro, la productividad y la riqueza aumentan, pero los beneficios no se distribuyen equitativamente. Existe el riesgo de que esta tendencia aumente con el metaverso; además, los ciudadanos mejor situados podrán aprovechar antes el metaverso que los peor situados.

IA para la defensa y la guerra. Los robots asesinos o los sistemas de armas autónomas letales son temas delicados en los debates sobre la IA para la defensa y la guerra. Los gobiernos y las organizaciones internacionales se enfrentan a este debate, ya que existe el riesgo de una carrera armamentística de IA. Pero, ¿cuál será el papel del metaverso en la guerra? Aunque (todavía) no se le llama guerra, en el ciberespacio ya están teniendo lugar muchas batallas entre países. ¿Cómo funcionará eso en el metaverso? ¿Podríamos o deberíamos imaginar una guerra librada en el metaverso? ¿Como complemento digital a una guerra tradicional? ¿Como una guerra totalmente independiente?

5.1.2 Desafíos indirectos

Los desafíos indirectos se refieren a los problemas que empiezan a surgir después de algún tiempo, como efecto secundario, en su mayoría imprevistos en el momento del lanzamiento al mercado. Entre los más conocidos, todos ellos debidos a las redes sociales, se encuentran el aumento de la polarización en las sociedades, el incremento de la anorexia y la vigorexia en los adolescentes por la presión social, y la adicción a la tecnología de los juegos o las aplicaciones de las redes sociales.

Aumento de la polarización en las sociedades. Los algoritmos de recomendación pueden crear burbujas de filtrado en las que las personas solo ven lo que les interesa, reforzando su propio pensamiento y cerrándoles el paso a perspectivas alternativas. Como hemos visto en algunas elecciones y otros acontecimientos democráticos importantes, los algoritmos de IA pueden marcar la diferencia cuando las elecciones son muy reñidas. Dada la experiencia inmersiva del metaverso, es probable que este riesgo aumente.

Adicción a la tecnología, a los juegos o a las aplicaciones de las redes sociales. Algunos algoritmos de recomendación son tan buenos que los usuarios no son capaces de desconectar de los contenidos que ofrecen. Ejemplos concretos son Instagram y TikTok, que presentan un vídeo tras otro perfectamente personalizado a los intereses de cada usuario. Algunas personas, especialmente los jóvenes, pueden convertirse en adictos a esas aplicaciones de medios sociales. Pueden pasar muchas horas al día en ellas y cuando finalmente se desconectan, sienten ansiedad y necesitan volver a conectarse cuanto antes. Los juegos plantean un reto similar, cuando la inteligencia artificial optimiza tanto la experiencia que la gente no puede dejar de jugar y, cuando lo hace, se siente infeliz. Aunque esto puede ocurrirle a cualquier persona, el riesgo es más grave para los adolescentes jóvenes. De nuevo,

debido a la experiencia inmersiva del metaverso, es probable que este riesgo aumente¹⁴.

El aumento de la anorexia y la vigorexia en los adolescentes por la presión social. En relación con el riesgo anterior sobre la adicción a las redes sociales, esas aplicaciones también pueden ejercer una profunda presión social y crear una experiencia distorsionada de la realidad. La combinación de horas de exposición y la presión social de ser físicamente perfecto, está provocando ya un aumento de la anorexia y la vigorexia en adultos jóvenes, que son vulnerables para que se produzcan esos riesgos. En el metaverso, es probable que este riesgo aumente porque la gente podría disociarse más fácilmente de su imagen física y empezar a preferir la imagen de su avatar.

5.2 Nuevos retos

El metaverso también introducirá nuevos retos^{15,16} aunque algunos de ellos son una extensión de los retos a los que ya nos enfrentamos con la IA, adquieren una nueva dimensión en el metaverso. Aunque no podemos prever de antemano todos esos retos, probablemente algunos serán:

5.2.1 Problemas de salud, incluidos los mentales

- Repetimos aquí la adicción a la tecnología porque no hay que subestimar el poder de los algoritmos de recomendación en un entorno inmersivo como el metaverso. Las recomendaciones pueden ser tan buenas que, a partir de cierto punto, a la gente le puede resultar (demasiado) difícil salir del metaverso cuando debe/quiere.
- El uso obsesivo del metaverso para escapar de la realidad: las personas que no son felices en el mundo real, pueden encontrar una alternativa atractiva en el mundo virtual donde pueden ser quienes quieren ser. En lugar de verse estimulados a mejorar su vida real, huyen de ella, percibiendo la vida real cada vez peor.
- Los niños son especialmente vulnerables a las tecnologías inmersivas, ya que es más probable que confundan la realidad con el mundo virtual.

¹⁴ <https://www.sciencedirect.com/science/article/pii/S0261517721001138>

¹⁵ <https://www.makeuseof.com/reasons-to-worry-about-facebook-metaverse/>

¹⁶ <https://www.makeuseof.com/metaverse-worsen-effects-social-media/>

- El ciberacoso probablemente aumentará¹⁷ y tendrá un mayor impacto negativo gracias a la experiencia digital mejorada e inmersiva, que se percibe casi como una realidad.
- Tristeza pos-RV¹⁸: el mundo real se vuelve decepcionante y las personas experimentan sentimientos de tristeza.
- La "resaca" de la realidad virtual o ciberenfermedad, que denota sensaciones de náuseas, fatiga, mareos y desorientación corporal¹⁹.

5.2.2 Desafíos a la privacidad

- La tecnología de realidad extendida tiene una capacidad aún mayor de captar datos personales como ubicaciones, movimientos, reflejos, movimientos oculares y patrones de voz. ¿Qué ocurre con esta información y cómo pueden los usuarios mantenerla bajo su control?
- Los anuncios dirigidos y la preocupación por la privacidad de los datos asociada se dispararán y pueden llegar a ser muy intrusivos.

5.2.3 Desafíos a la libertad

- Encontrar el equilibrio adecuado entre la libertad de expresión y evitar que se publiquen y difundan contenidos maliciosos en el metaverso.
- Como hemos visto, un elemento clave del metaverso es la realidad mixta (RM), una mezcla de los mundos digital y real mediante el uso de tecnologías de realidad virtual (RV) y realidad aumentada (RA). Con el tiempo, esta mezcla puede llegar a ser tan inmersiva y omnipresente que las vidas virtuales y reales de las personas se unan y sean indistinguibles. Si esto ocurre, quien controle (una parte importante del) el metaverso podría controlar una parte importante de la realidad.

¹⁷ El metaverso más allá de la propaganda:

<https://www.sciencedirect.com/science/article/pii/S0268401222000767>

¹⁸ <https://www.theatlantic.com/technology/archive/2016/12/post-vr-sadness/511232/>

¹⁹ La ética de la tecnología de la realidad virtual:

<https://link.springer.com/article/10.1007/s11948-017-9979-y>

5.2.4 Desafíos del mundo dual

- ¿Hay actos considerados ilegales en el mundo real que están permitidos en el metaverso? ¿Cuáles son esos actos y quién debe regularlos?
- Pasar horas y horas en mundos virtuales, puede hacer que las personas se vuelvan menos sensibles a las consecuencias de las acciones en el mundo real con todas las posibles consecuencias negativas (disociación).
- Como vimos, una organización autónoma descentralizada (DAO) es un programa informático que incorpora reglas de gobierno y toma de decisiones que funcionan sobre una red descentralizada (cadena de bloques). Una DAO proporciona un sistema que traslada esas reglas a un contrato inteligente, de manera que no es necesario confiar en personas e instituciones (no existe un órgano de gobierno en una DAO), evitando posibles manipulaciones o corrupción. Sin embargo, aunque en teoría parece un sistema eficiente sin necesidad de relegar una decisión a un comité o a personas concretas, existen riesgos potenciales como la compra de votos para saltarse las reglas programadas, o el posible sesgo que se codifica en los contratos inteligentes y que puede acabar perjudicando a la organización. Las DAO como Decentraland²⁰ ni siquiera tienen una empresa detrás, lo que crea incertidumbre sobre cómo interactuar en el mundo virtual y el real en caso de problemas.

5.2.5 Desafíos a la igualdad

- Un hardware caro puede excluir a las personas con menos poder adquisitivo, lo que aumenta la brecha digital.
- El acceso al metaverso y a sus mundos virtuales debe ser justo en el sentido de que las comisiones de los creadores no deben ser abusivas.
- Existe un riesgo de accesibilidad e inclusión si el metaverso solo se desarrolla para la mayoría, dejando fuera a grupos específicos como las personas mayores con menos conocimientos digitales (brecha digital) o a las personas con visión o audición limitadas.

²⁰ <https://decentraland.org/>

5.3 Desafíos derivados de un uso malicioso deliberado del metaverso

Hasta ahora, hemos hablado de los retos del metaverso derivados de las intenciones legítimas. Sin embargo, también existen desafíos debido al uso malicioso deliberado del metaverso. Aunque la mayoría de estos actos están prohibidos por la ley, siguen ocurriendo y es probable que aumenten con el metaverso.

- Suplantación de identidad. En el metaverso, puedes acceder a diferentes mundos virtuales con la misma identidad. Las consecuencias negativas de la suplantación de identidad son, por tanto, mucho más graves que en el caso actual de los servicios digitales, donde cada servicio tiene un nombre de usuario diferente.
- Si tu identidad es secuestrada, eres más vulnerable al programa de secuestro y a la extorsión. Con los servicios digitales, puedes restablecer tu contraseña. Sin embargo, en el metaverso no puedes cambiar simplemente tu avatar, ya que está directamente conectado a tu existencia virtual. ¿Y cómo van a desempeñar los *deep fakes* un papel en el robo de identidad?
- Robo de propiedades virtuales.
- ¿Deben considerarse un delito las agresiones virtuales en el metaverso? ¿Puede esto implicar incluso daños físicos en el mundo real, por ejemplo, si se utiliza un traje háptico?
- Manipulación del proceso de votación de las DAO con intención maliciosa a través de la compra de votos.

5.4 Enfoques para afrontar los desafíos y mitigar los riesgos sociales y éticos

Como se está experimentando actualmente con la IA, hay diferentes formas de tratar los riesgos éticos y sociales previstos en el metaverso: autorregulación, recomendaciones de instituciones internacionales o regulación. Tenga en cuenta que no son necesariamente excluyentes.

La **autorregulación** significa que las empresas consideran, detectan y mitigan los posibles riesgos sociales y éticos de su actividad en el metaverso de forma voluntaria. En el mundo de la IA, la noción de un uso responsable de la misma desde el diseño está siendo adoptada cada vez por más organizaciones. Los riesgos éticos y sociales se tienen en cuenta durante todo el ciclo de vida del sistema y, cuando se detectan, se mitigan mediante acciones específicas. Si es imposible mitigar los riesgos previstos, las empresas podrían y deberían abstenerse de desplegarlo hasta que se conozcan mejor los riesgos y se puedan evitar. Las acciones de autorregulación relacionadas

con los riesgos sociales y éticos pueden gestionarse a través de iniciativas ESG (ambientales, sociales y de gobernanza) o códigos de conducta. Basándonos en la experiencia con la IA y las ESG, creemos que las empresas empezarán a trabajar también en el uso responsable del metaverso desde el diseño. La autorregulación es muy importante con una nueva tecnología como el metaverso, ya que hay poca experiencia real de su impacto en la sociedad.

Recomendaciones de las instituciones internacionales y de las colaboraciones público-privadas. Las partes interesadas de varios países pueden reunirse para debatir cómo la tecnología puede constituir un riesgo social o ético y presentar recomendaciones a las empresas y los gobiernos para hacerles frente. Entre los ejemplos de este tipo de iniciativas para la IA se encuentran los principios de la IA de la OCDE²¹ y la recomendación sobre IA de la UNESCO²². Las organizaciones individuales pueden entonces adherirse a las directrices. Insistimos en la importancia de estos debates interdisciplinarios y globales sobre los riesgos sociales y éticos del metaverso, incluidas las colaboraciones público-privadas. Esos debates son una aportación esencial a una posible regulación.

La **regulación** significa que algunos usos de la tecnología están regulados por ley. El RGPD (Reglamento general de protección de datos²³) europeo y la Regulación de IA (próximo reglamento europeo de IA²⁴) son ejemplos de ello. Un enfoque de la regulación basado en el riesgo implica que cuanto mayor sea el riesgo, más normas se aplicarán. Dado el impacto potencial de largo alcance del metaverso, sería prudente empezar a pensar en los riesgos importantes que —como sociedades— queremos evitar. Existe el riesgo de que la regulación del metaverso llegue demasiado pronto, cuando todavía haya poca experiencia real con su impacto social y ético. Como se ha dicho, los debates interdisciplinarios y globales son esenciales en este momento.

6. Conclusión

Todo el mundo habla del metaverso, pero hoy en día sigue siendo más una promesa que una realidad. En este trabajo hemos demostrado que el metaverso no aparece de la nada, sino que es una consecuencia lógica de varias evoluciones importantes. Hemos explorado los posibles retos sociales y éticos del metaverso a partir de una experiencia similar con la inteligencia artificial.

²¹ <https://oecd.ai/assets/files/OECD-LEGAL-0449-en.pdf>

²² <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

²³ <https://gdpr-info.eu/>

²⁴ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206>

Pocos dudan de que el metaverso tendrá un impacto significativo en nuestras vidas, sociedad y economías. Teniendo en cuenta esta expectativa, es importante pensar con antelación (ahora) en los posibles retos sociales y éticos, y diseñar estrategias para mitigarlos. Solo así la humanidad podrá disfrutar plenamente de las oportunidades y, al mismo tiempo, reducir las posibles consecuencias negativas. Podemos y debemos aprender de la experiencia con la inteligencia artificial, una tecnología transformadora con también un enorme impacto en la sociedad, cuyo uso debe gestionarse con cuidado. Como hemos visto, actualmente hay muchas iniciativas en marcha para garantizar un buen uso de la inteligencia artificial, incluyendo la autorregulación, las recomendaciones internacionales y la regulación.

Nadie sabe cómo se desarrollarán los posibles riesgos sociales y éticos del metaverso. Pero las empresas que han empezado a aplicar el uso ético de la IA, estarán en una buena posición para enfrentarse a ellos. En Telefónica estamos utilizando una metodología denominada "uso responsable de la IA desde el diseño", que consiste en los principios de la IA²⁵, la concienciación y formación de los empleados, un cuestionario con las preguntas adecuadas, herramientas técnicas y un modelo de gobernanza²⁶ que define las funciones y responsabilidades. Hemos identificado un nuevo rol llamado *campeón responsable de la IA*²⁷ que es la persona a la que hay que acudir para las cuestiones relacionadas con el uso ético de la IA. Como parte de nuestro modelo de gobernanza, hemos creado un comité de ética de la IA formado por expertos multidisciplinares. También tenemos un acuerdo con la UNESCO en relación con su Recomendación sobre la ética de la inteligencia artificial²⁸. Nuestra intención es utilizar un enfoque similar para el metaverso, aprovechando también la amplia experiencia de Telefónica en ciberseguridad y privacidad desde el diseño.

6.1 Hacia los valores y principios éticos del metaverso

Dados los muchos retos que se presentan asociados al metaverso, se necesitan valores y principios éticos ahora —antes de que siga creciendo— para garantizar un metaverso justo en el futuro. Para abrir el debate público sobre cuáles deberían ser esos valores y principios éticos, sugerimos tomar como punto de partida los valores identificados en la recomendación de la UNESCO sobre la ética de la inteligencia artificial²⁹, entre los que se encuentran:

²⁵ <https://www.telefonica.com/wp-content/uploads/sites/7/2021/11/principios-ai-eng-2018.pdf>

²⁶ <https://www.telefonica.com/es/wp-content/uploads/sites/4/2021/06/ia-responsible-governance.pdf>

²⁷ <https://business.blogthinkbig.com/a-new-organizational-role-for-artificial-intelligence-the-responsible-ai-champion/>

²⁸ <https://www.telefonica.com/es/sala-comunicacion/la-unesco-y-telefonica-se-comprometen-a-promover-impulsar-e-implementar-la-recomendacion-sobre-la-etica-de-la-inteligencia-artificial/>

²⁹ <https://unesdoc.unesco.org/ark:/48223/pf0000380455>

- Respect, protection and promotion of human rights and fundamental freedoms and human dignity.
- Environment and ecosystem flourishing
- Ensuring diversity and inclusiveness
- Living in peaceful, just and interconnected societies

Los principios éticos de un metaverso justo incluirán probablemente los siguientes conceptos: equidad, igualdad, responsabilidad y rendición de cuentas, privacidad, seguridad, ecología y sostenibilidad, niños y otros grupos vulnerables, modelos de negocio que no dañen, transparencia, inclusión y libertad.

¡Bienvenidos al debate!

7. Reconocimientos

Nos gustaría agradecer a las siguientes personas por sus comentarios sobre las versiones anteriores: Francisco Montalvo, Joaquina Salado, Eusebio Felguera, Irene Bernal y Silvia Díaz.



www.telefonica.com