Universidad Internacional de La Rioja
Facultad de Derecho

Máster Universitario en el Ejercicio de la Abogacía y la Procura

Regulación global de la inteligencia artificial generativa: datos, privacidad y propiedad intelectual en Europa, Estados Unidos y China.

Trabajo fin de estudio presentado por:	Darío Florencio Sansano Larrea
Tipo de trabajo:	Investigación teórica
Área jurídica:	Derecho Digital
Director/a:	Ignacio Lopez Lopez
Fecha:	28/11/2024

Resumen

Este trabajo de investigación jurídica examina los marcos regulatorios de la inteligencia artificial (IA) en la Unión Europea, Estados Unidos y China, aplicándolos al proyecto empresarial **NeuroBlock**, del cual soy cofundador. **NeuroBlock** es una startup ganadora de la 13ª edición del programa de creación de startups del Parque Científico de la Universidad Miguel Hernández de Elche, donde actualmente se está desarrollando. Su enfoque principal es la recopilación de datos personales anonimizados a través de una extensión de navegador, recompensando a los usuarios con una criptomoneda propia. Estos datos, junto con información obtenida mediante web scraping, son procesados por sistemas de IA para generar estudios de mercado dirigidos a pequeñas y medianas empresas.

El estudio identifica los principales desafíos legales relacionados con la protección de datos, la propiedad intelectual, y propone estrategias para cumplir con normativas como la LOPDGDD, el GDPR, el Artificial Intelligence Act, el SB 1047 y el Plan de Desarrollo de IA de China, entre otras novedosas regulaciones. Además, con este trabajo, se pretende subrayar la importancia de marcos regulatorios armonizados que permitan promover la innovación tecnológica al tiempo que se protegen los derechos de los ciudadanos.

Palabras clave: Regulación de Inteligencia Artificial, Protección de Datos, Big Data, Criptomonedas, Web Scraping, Propiedad Intelectual

Abstract

This legal research examines the regulatory frameworks of artificial intelligence (AI) in the European Union, the United States, and China, applying them to the entrepreneurial project **NeuroBlock**, of which I am the co-founder. **NeuroBlock** is a startup that won the 13th edition of the startup creation program at the Scientific Park of the Universidad Miguel Hernández de Elche, where it is currently being developed. Its primary focus is the collection of anonymized personal data through a browser extension, rewarding users with a proprietary cryptocurrency. This data, along with information obtained through internet web scraping, is processed by AI systems to generate market studies aimed at small and medium-sized enterprises.

The study identifies the main legal challenges related to data protection, intellectual property, and proposes strategies to comply with regulations such as the LOPDGDD, GDPR, Artificial Intelligence Act, SB 1047, and China's AI Development Plan. Additionally, this work aims to underscore the importance of harmonized regulatory frameworks that promote technological innovation while protecting citizens' rights.

Keywords:

Artificial Intelligence Regulation, Data Protection, Big Data, Web Scraping, Intellectual property

índice

1. Introducción

La inteligencia artificial (IA) ha experimentado un rápido desarrollo en los últimos años, impactando diversos sectores sociales, económicos y jurídicos. Esta aceleración tecnológica ha generado una necesidad urgente de establecer marcos regulatorios que salvaguarden derechos fundamentales, particularmente la protección de datos y otros derechos vinculados a la privacidad y la intimidad, así como derechos de tipo industrial, como puedan ser los derechos de propiedad intelectual, todo ello sin frenar la innovación.

La justificación para este trabajo radica en la creciente presencia de sistemas de IA que utilizan técnicas como el Web Scrapping masivo de datos de internet para los entrenamientos de los nuevos sistemas de Inteligencia Artificial Generativa, basasdos principalmente en la arquitectura de *transformers* y en los conflictos legales derivados de la creación de contenidos por IA, que generan diversas controversias, especialmente en lo que respecta a la titularidad de los derechos de autor y la violación de la privacidad.

1.1. Justificación del tema elegido

La inteligencia artificial se ha convertido en un área estratégica para el desarrollo económico, lo que ha llevado a la adopción de normativas en varias jurisdicciones. Sin embargo, estas normativas presentan diferencias significativas en su enfoque y aplicación. Por ejemplo, en Europa se ha propuesto y aprobado el ***Artificial Intelligence Act*** (REGLAMENTO (UE) 2024/1689, 2024), mientras que Estados Unidos adopta un enfoque sectorial, destacándose la ***Safe and Secure Innovation for Frontier Artificial Intelligence Models Act*** (SB 1047, 2024) de California, y China prioriza la seguridad nacional a través de su ***Plan de Desarrollo de IA de Próxima Generación*** (PLAN DE DESARROLLO DE IA, 2017).

Este trabajo pretende analizar estos marcos regulatorios, con especial atención a sus intersecciones y vacíos legales, aplicándolos a la startup **NeuroBlock**, proyecto ganador del programa de incubación de startups de la que soy cofundador y que, tanto desde un punto de vista jurídico como técnico en el desarrollo de software, estoy desarrollando en el parque

científico de la Universidad Miguel Hernández de Elche.

1.2. Problema y finalidad del trabajo

El problema central que aborda este trabajo es la falta de uniformidad y claridad en la regulación de la IA, principalmente aunque no de forma exclusiva, en cuanto a la protección de datos y la propiedad intelectual de los contenidos generados por estos sistemas. La finalidad del trabajo es proponer una comparación crítica entre los marcos regulatorios en España, la Unión Europea, Estados Unidos y China como principales bloques geopolíticos, de cuya tensión se espera, surjan y de hecho ya está sucediendo, los principales instrumentos regulatorios de la nueva IA Generativa, identificando áreas de mejora y recomendando posibles soluciones regulatorias que equilibren la innovación con la protección de derechos fundamentales, aplicándolo directamente al proyecto NeuroBlock como caso de estudio particular.

1.3. Objetivos

Objetivo general: Realizar un análisis comparativo de los principales marcos regulatorios en el ámbito de la inteligencia artificial, con especial atención a las áreas de protección de datos y propiedad intelectual, aplicándolos a la startup de Inteligencia Artificial NeuroBlock.

Objetivos específicos:

1. Analizar la regulación actual en España y su relación con el Artificial Intelligence Act de la Unión Europea.
2. Examinar los enfoques regulatorios de la Unión Europea, Estados Unidos y China en el ámbito de la IA.
3. Identificar vacíos legales y áreas de controversia en la regulación de la IA, tales como el uso de Scrapping masivo y su impacto en la privacidad y los conflictos de propiedadintelectual que esta técnica de recopilación de datos en auge genera.
4. Proponer alternativas que permitan un desarrollo ético y regulado de la IA en el contexto de NeuroBlock.

2. Marco teórico y desarrollo

2.1. Regulación de la Inteligencia Artificial en la Unión Europea

La Unión Europea ha demostrado ser un referente global en la regulación de tecnologías emergentes, estableciendo marcos legales robustos que equilibran la protección de derechos fundamentales con la promoción de la innovación. En un contexto donde la inteligencia artificial (IA) y el Big Data están transformando profundamente la economía y la sociedad, la UE busca liderar con un enfoque normativo que priorice la privacidad, la transparencia y la seguridad.

La regulación de la IA en la UE no solo responde a los riesgos asociados con el uso de datos personales, sino que también fomenta un entorno de confianza para el desarrollo y adopción de tecnologías responsables. Normativas como el Reglamento General de Protección de Datos (GDPR), el Data Governance Act (DGA) y el recientemente adoptado Artificial Intelligence Act configuran un ecosistema jurídico coherente y detallado que plantea tanto oportunidades como desafíos para empresas tecnológicas como NeuroBlock. Estos marcos subrayan la importancia de principios como la soberanía de los datos, la explicabilidad de los sistemas y la interoperabilidad, aspectos que demandan soluciones técnicas avanzadas y una adaptación constante por parte de las organizaciones.

En este apartado, se analizarán las principales normativas de la UE que impactan el desarrollo y uso de la inteligencia artificial, evaluando su relevancia en la protección de datos y explorando cómo estas regulaciones pueden guiar a NeuroBlock en la implementación de prácticas alineadas con los valores europeos, sin comprometer su capacidad para innovar en un mercado altamente competitivo.

2.1.1. Reglamento General de Protección de Datos (GDPR)

El **GDPR** es la normativa más robusta en cuanto a la protección de datos personales dentro de la Unión Europea, y establece principios rectores para cualquier actividad que implique el tratamiento de datos. De particular relevancia es el **artículo 5**, que fija los principios de licitud, lealtad y transparencia en el uso de datos. Estos principios afectan directamente el uso de técnicas de **Web Scraping**, ya que para que este tipo de actividad sea legal, debe garantizarse el consentimiento explícito del usuario, tal y como señala el **artículo 7** del GDPR.

Otra disposición clave es el **derecho al olvido,** recogido en el **artículo 17 del RGPD**, que permite a los usuarios solicitar la eliminación de sus datos. Sin embargo, este derecho plantea desafíos técnicos significativos en el contexto de la IA generativa. Una vez que los datos se utilizan para entrenar modelos basados en transformadores, eliminarlos del sistema sin comprometer su integridad resulta prácticamente imposible. Durante el entrenamiento, cada palabra o subpalabra se convierte en un token, que se traduce en un vector numérico en un espacio de alta dimensionalidad mediante una capa de embedding. Estos modelos, mediante mecanismos de atención (*self-attention*), establecen relaciones contextuales entre tokens utilizando cálculos matriciales (VASWANI et al., 2017). Este proceso permite que, dado un conjunto inicial de palabras, el modelo sea capaz de predecir la siguiente secuencia de forma coherente y contextualizada. Los datos utilizados en el entrenamiento quedan integrados indirectamente en los pesos del modelo, influyendo tanto en su capacidad para generalizar conocimiento como en la calidad de sus respuestas. Este embebimiento dificulta la eliminación específica de datos sin afectar al rendimiento general del modelo.

Según ROSA (2023), el **artículo 17 del GDPR** plantea dificultades prácticas para las empresas tecnológicas, ya que "la eliminación de datos en un contexto donde los modelos de IA dependen de vastos conjuntos de datos históricos puede resultar técnicamente inviable". Esta apreciación resulta acertada, ya que estas nuevas tecnologías exigen la adopción de paradigmas tanto legales como arquitectónicos.

Empresas como Anthropic han logrado avances significativos en sus investigaciones sobre las

redes neuronales subyacentes en los modelos de inteligencia artificial generativa. Por ejemplo, han desarrollado técnicas para desactivar determinadas "neuronas", permitiendo maximizar o inhibir ciertos contenidos en las respuestas generadas por estos sistemas (ANTHROPIC, 2024). Este enfoque podría vislumbrar un futuro en el que la tecnología se alinee parcialmente con el derecho al olvido de los usuarios. Sin embargo, la viabilidad técnica sigue siendo extremadamente compleja. Por el momento, parece más factible avanzar hacia un nuevo marco regulatorio que contemple técnicas como la anonimización de datos y sistemas de seguridad diseñados para controlar los *outputs* relacionados con información personal. Sistemas que todavía no se contemplan en los marcos regulatorios actuales, que chocan con la falta de flexibilidad de los mismos ante un ecosistema tecnológico tan disruptivo y cambiante.

2.1.2. Data Governance Act (DGA) y Data Act

El Data Governance Act (DGA) y el Data Act complementan el GDPR al regular el intercambio seguro de datos en la UE, promoviendo un entorno de confianza para la reutilización de datos no personales. El Data Governance Act se enfoca en la creación de intermediarios de datos que garantizan la ética y la legalidad en el uso de información anonimizada.

El **Data Act**, que entró en vigor el 11 de enero de 2024 y derá de aplicación a apartir del 12 de septiembre de 2025, establece un marco para el acceso y la portabilidad de los datos generados por dispositivos conectados, según el **artículo 3**. Este reglamento será especialmente relevante para proyectos como **NeuroBlock**, que utilizan datos generados por los usuarios de internet mientras navegan, por el momento a través de extensiones de navegador, lo que podría estar sujeto a nuevas obligaciones de portabilidad y acceso.

Como expone PÉREZ-UGENA (2022), la interoperabilidad de los datos en Europa está alcanzando niveles críticos, y regulaciones como el *Data Act* intentan "asegurar que el control sobre los datos permanezca en manos de los usuarios, sin obstaculizar la innovación tecnológica". Desde el punto de vista tecnológico, están surgiendo nuevos paradigmas que buscan alinearse con este principio de soberanía del dato, fundamental para el desarrollo de una economía justa, lejos del monopolio de las grandes tecnológicas en los planos económico y político.

Proyectos como Web5, desarrollado por TBD, una subsidiaria de Block Inc. liderada por Jack Dorsey, fundador de la antigua Twitter, proponen un ecosistema descentralizado en el que una red de nodos reemplaza a los servidores centralizados tradicionales (TBD, n.d.). Este sistema se basa en principios como los Identificadores Descentralizados (Decentralized Identifiers, DIDs) y los Nodos Web Descentralizados (Decentralized Web Nodes, DWNs), permitiendo que los usuarios gestionen sus propios datos de manera encriptada y sin depender de terceros. En este modelo, los usuarios son los únicos con acceso a las claves que controlan su información, decidiendo quién puede acceder a ella y con la capacidad de revocar dicho acceso en cualquier momento (TBD, n.d.). Este enfoque promueve una gestión completamente autogestionada y portátil de la información personal, así como de los datos generados en plataformas digitales.

Este paradigma resulta de gran relevancia al proponer un escenario donde el acceso a la información sea verdaderamente permisivo y controlado. En este modelo, individuos, empresas, organismos y otras instituciones podrían ejercer control directo sobre su información. Aunque actualmente parece una solución algo lejana, desde un punto de vista técnico es una de las más prometedoras para garantizar que los modelos de inteligencia artificial accedan de forma lícita y justa a los datos, tanto para su entrenamiento como para su aplicación en sistemas de generación aumentada (Retrieval-Augmented Generation, RAG). Estos sistemas generan respuestas en tiempo real a partir de información recuperada de internet, asegurando mayor transparencia y equidad en el manejo de los datos.

En este sentido, el legislador debería poner el foco, de la mano de expertos en el campo tecnológico, en este tipo de nuevos paradigmas que surgen, alejándose de los marcos regulatorios centrados en realidades anacrónicas muy alejadas del camino que estás nuevas tecnologías están tomando y que inevitablemente, van marcar nuestros caminos.

2.1.3. Relación entre IA y los marcos de protección de datos

El Artificial Intelligence Act clasifica los sistemas de IA según su nivel de riesgo, estableciendo medidas específicas para cada categoría. Los sistemas de alto riesgo, que incluyen aquellos que procesan grandes volúmenes de datos personales, deben cumplir con estrictas

obligaciones de transparencia y supervisión humana, tal como se detalla en el artículo 14 del reglamento. Este enfoque basado en el riesgo divide los sistemas de IA en las siguientes categorías:

- **Riesgo inaceptable:** Incluye sistemas prohibidos debido a su potencial para causar daños significativos, como aquellos que manipulan el comportamiento humano de forma perjudicial o implementan la puntuación social por parte de gobiernos (artículo 6, apartado 1, AI Act).
- **Riesgo alto:** Abarca sistemas que afectan derechos fundamentales o la seguridad, como los utilizados en infraestructuras críticas, empleo, educación o aplicación de la ley. Estos deben cumplir con requisitos estrictos, incluyendo evaluaciones de conformidad y medidas para mitigar riesgos (artículos 8-29, y especificados en el Anexo III).
- **Sistemas con obligaciones específicas de transparencia:** Corresponde a sistemas que requieren transparencia hacia los usuarios, como los chatbots, que deben informar claramente que son máquinas y no humanos (artículo 50).
- **Riesgo mínimo o nulo:** Se refiere a sistemas que no presentan riesgos significativos para los derechos o la seguridad y que están exentos de obligaciones adicionales bajo el AI Act.

El objetivo principal del AI Act es garantizar que los sistemas de IA respeten los derechos fundamentales y la privacidad, estableciendo requisitos como la transparencia, la explicabilidad y la seguridad de los sistemas. Además, el reglamento busca fomentar la innovación tecnológica mediante un enfoque normativo que equilibre la protección de derechos con el apoyo a pequeñas y medianas empresas, facilitando su integración en el ecosistema tecnológico europeo.

En relación con la protección de datos, el reglamento refuerza la coherencia con normativas como el Reglamento General de Protección de Datos (GDPR), asegurando que el procesamiento de datos personales por sistemas de IA sea lícito y justo (esto se trata a lo largo del reglamento en varios artículos, especialmente en las disposiciones sobre calidad de datos y sesgos). Tecnologías como los Identificadores Descentralizados (DIDs) y los Nodos

Web Descentralizados (DWNs) representan soluciones técnicas viables para cumplir con estos principios. Estas herramientas, propuestas en proyectos como Web5, permiten a los usuarios gestionar de manera autónoma sus datos personales, controlando el acceso y las condiciones de uso.

En definitiva, el AI Act plantea un marco regulatorio que no solo establece directrices para la seguridad y la transparencia, sino que también refuerza la soberanía del usuario sobre sus datos personales. Este enfoque se alinea con principios fundamentales de los marcos de protección de datos en Europa, promoviendo un equilibrio entre la innovación tecnológica y la protección de los derechos individuales.

2.2. Marcos regulatorios en Estados Unidos

En Estados Unidos, la regulación de la inteligencia artificial (IA) sigue un enfoque sectorial y fragmentado. La ausencia de un marco normativo unificado ha dado lugar a que diferentes agencias federales y estatales desarrollen regulaciones específicas para abordar los riesgos y oportunidades que presenta la IA en diversos sectores. A continuación, se detallan los principales marcos regulatorios en el país:

2.2.1. Federal Trade Commission (FTC)

La **Federal Trade Commission (FTC)** desempeña un papel central en la supervisión del uso de la IA en relación con la protección del consumidor y la privacidad. A través de leyes como la **Fair Credit Reporting Act (FCRA)** y la **FTC Act**, la comisión tiene la autoridad de actuar contra prácticas comerciales desleales y engañosas que involucren el uso de datos personales mediante tecnologías de IA.

Uno de los casos más notables es el de **Everalbum, Inc.**, donde la FTC sancionó a la empresa por usar fotos de usuarios sin su consentimiento para entrenar un algoritmo de reconocimiento facial. La sanción no solo incluyó la eliminación de los algoritmos creados con esos datos, sino que también representó una acción firme contra la obtención de beneficios injustos mediante la recopilación indebida de datos personales (MCSWEENY, 2022). Este caso marcó un precedente sobre la transparencia y el consentimiento en el uso de IA en tecnologías biométricas.

En otro caso significativo, la **FTC** impuso sanciones a **Flo Health, Inc.**, una aplicación que compartió datos de salud sensibles de sus usuarias con terceros como **Facebook** y **Google**, incumpliendo sus promesas de privacidad. Además de las sanciones económicas, la empresa se vio obligada a notificar a los usuarios afectados, lo que supuso un avance en las regulaciones de transparencia en la gestión de datos personales por sistemas automatizados (SWEIJS, 2024).

Asimismo, la **FTC** ha emitido propuestas adicionales sobre la regulación de sistemas automatizados de toma de decisiones y el uso de **datos biométricos**, promoviendo la **transparencia** y equidad en la toma de decisiones algorítmicas, garantizando que los sistemas sean justos y comprensibles para los usuarios (MCSWEENY, 2022).

2.2.2. National Institute of Standards and Technology (NIST)

El National Institute of Standards and Technology (NIST) ha desarrollado el AI Risk Management Framework (RMF), un conjunto de directrices diseñadas para ayudar a las organizaciones a gestionar los riesgos relacionados con la IA. Aunque el RMF no es vinculante, se ha convertido en un estándar de referencia para el desarrollo de tecnologías de IA que sean seguras, transparentes y equitativas.

El **RMF** incluye cuatro funciones clave: **Mapear, Medir, Gobernar** y **Gestionar**. Estas funciones no solo promueven la transparencia en el diseño de sistemas de IA, sino que también fomentan una cultura organizacional de responsabilidad y supervisión de riesgos. El marco recomienda pruebas continuas para garantizar que los algoritmos no generen decisiones sesgadas o discriminatorias, un tema alineado con los principios de la **OCDE** sobre el desarrollo responsable de IA (NIST, 2023).

El documento **NIST AI 100-1** y su versión extendida, **NIST AI 600-1**, subrayan la importancia de la evaluación continua de los modelos de IA en la detección y mitigación de riesgos que puedan comprometer la confianza pública y los derechos de los individuos (NIST AI 600-1, 2024).

2.2.3. Food and Drug Administration (FDA)

La **Food and Drug Administration (FDA)** desempeña un papel crítico en la regulación de tecnologías de IA en el sector de la salud, especialmente en dispositivos médicos y herramientas de soporte clínico. El objetivo principal de la **FDA** es asegurar que estas tecnologías sean seguras y eficaces antes de su aprobación para su uso por los pacientes.

En particular, la **FDA** regula el uso de **Clinical Decision Support Tools (CDS)**, herramientas que ayudan a los médicos a interpretar datos clínicos complejos. La normativa exige que estas herramientas sean explicables y que los algoritmos no introduzcan errores que puedan comprometer la seguridad del paciente. Además, la **FDA** ha lanzado el **Digital Health Center of Excellence**, una iniciativa que promueve la innovación en salud digital, asegurando que las tecnologías basadas en IA cumplan con los estándares de seguridad y eficacia necesarios para proteger a los pacientes (FDA, 2023).

2.2.4. Iniciativas presidenciales

En octubre de 2023, el presidente Joe Biden firmó la Orden Ejecutiva 14110, titulada *Desarrollo y Uso Seguro, Seguro y Fiable de la Inteligencia Artificial* (*Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*), para garantizar el desarrollo seguro y transparente de la IA en Estados Unidos. Esta orden establece principios clave, como la protección de la privacidad, la equidad en los sistemas de IA y la supervisión en áreas críticas como la biotecnología y la ciberseguridad.

La **Orden Ejecutiva 14110** también promueve la cooperación internacional, buscando alinear las políticas de IA en Estados Unidos con un marco global que respete los derechos civiles y asegure la seguridad pública. Además, la iniciativa destina recursos a las agencias federales para supervisar los sistemas de IA en sectores estratégicos, reforzando la importancia de una regulación coherente y basada en estándares éticos.

2.2.5. Normativas estatales

A nivel estatal, varios estados han implementado regulaciones específicas sobre el uso de IA.

Estos marcos estatales complementan la ausencia de una regulación federal integral y abordan cuestiones específicas relacionadas con la privacidad y el uso ético de la IA.

- **California** fue pionera con la **California Consumer Privacy Act (CCPA)**, que regula el uso de datos personales en sistemas automatizados y exige que los chatbots notifiquen a los usuarios cuando interactúan con un sistema no humano. Esta legislación ha sentado un precedente en la protección de los consumidores contra el uso indebido de sus datos personales (MAREMONTI, 2024).
- En **Illinois**, se aprobó en 2019 la **Artificial Intelligence Video Interview Act**, que regula el uso de IA en entrevistas laborales. Esta ley garantiza que los candidatos sean informados sobre el uso de la IA y se les otorgue el derecho a rechazar su uso en los procesos de selección.
- En Colorado, se han implementado leyes que restringen el uso de modelos predictivos de IA en el sector asegurador, con el objetivo de evitar la discriminación basada en datos sensibles como la raza o el género. En mayo de 2024, el gobernador **Jared Polis** firmó la **Ley SB 205**, convirtiendo a Colorado en el primer estado en regular de forma integral el uso de IA en decisiones de alto riesgo. La ley exige la **notificación previa** a los usuarios sobre decisiones automatizadas, la **explicación del fundamento** de estas decisiones y la creación de **mecanismos para corregir información o apelar decisiones**. Además, requiere que las empresas realicen **evaluaciones de impacto** para mitigar riesgos de discriminación y divulguen detalles sobre los datos utilizados para entrenar sus modelos predictivos, garantizando así la **transparencia** y la **equidad** en sectores críticos como el asegurador (SB 205, 2024).

El panorama regulatorio de la IA en Estados Unidos es fragmentado pero dinámico, con esfuerzos tanto a nivel federal como estatal para abordar los desafíos que plantea esta tecnología emergente. Las agencias como la **FTC**, el **NIST** y la **FDA** desempeñan un papel clave en la supervisión y regulación del uso de la IA, mientras que las iniciativas presidenciales y las normativas estatales complementan estas acciones, creando un marco regulatorio más amplio que busca equilibrar la innovación tecnológica con la protección de derechos fundamentales.

Este enfoque fragmentado ofrece tanto oportunidades como desafíos para el desarrollo de la IA en Estados Unidos. A medida que la tecnología evoluciona, es posible que el país avance hacia una regulación más coherente e integral, alineada con los marcos internacionales que priorizan la ética y los derechos civiles. No obstante, y pese a la complejidad que pueda presentar tal dispersión, parece más realista para una tecnología tan disruptiva y novedosa como la IA generativa basada en arquitecturas de transformadores optar por un enfoque regulatorio flexible. Este enfoque, menos rígido y monolítico que el europeo, permite una mayor adaptabilidad y casuística, especialmente en etapas tan tempranas donde la tecnología requiere un espacio para maniobrar y florecer, en lugar de enfrentarse a una legislación férrea e intransigente.

2.3. Marcos regulatorios en China

China ha adoptado un enfoque centralizado y de control estatal para la regulación de la inteligencia artificial (IA), con el objetivo de posicionarse como líder mundial en esta tecnología para 2030. Su estrategia se delineó en el **Plan de Desarrollo de IA de Próxima Generación** (2017-2030), que prioriza la seguridad nacional y el crecimiento económico basado en la IA (CONSEJO DE ESTADO DE CHINA, 2017).

2.3.1. Plan de desarrollo de IA de próxima generación (2017-2030)

El **Plan de Desarrollo de IA de Próxima Generación** se estructura en tres fases principales:

- **2020**: China debía equipararse a los niveles tecnológicos globales en IA, estableciendo un entorno competitivo para la industria. Para este punto, la IA debía ser una fuerza clave en sectores estratégicos como el militar y el económico (CONSEJO DE ESTADO DE CHINA, 2017).
- **2025**: Se espera que para esta fecha, China haya logrado avances significativos en teorías fundamentales de la IA y haya creado un marco normativo y ético robusto para regular su uso.
- **2030**: El objetivo final es que China se convierta en el centro global de innovación en IA, con un marco regulador maduro que garantice la seguridad, la ética y la privacidad

en el uso de la IA, asegurando que la tecnología no ponga en riesgo la seguridad nacional ni los derechos de los ciudadanos (CONSEJO DE ESTADO DE CHINA, 2017; MISKELLY, 2024).

El enfoque centralizado de China en el desarrollo de la inteligencia artificial refleja su ambición de liderar en sectores estratégicos clave, como se detalla en el análisis realizado en ***Deciphering China's AI Dream*** (DING 2018). Este estudio señala que China utiliza la IA no solo como motor económico para impulsar su competitividad global, sino también como una herramienta para fortalecer el control estatal, mediante la supervisión y gestión de su población, y como un instrumento de proyección de poder blando en el escenario internacional, promoviendo su modelo tecnológico y político en otros países.

2.3.2. Gobernanza y control estatal

El enfoque regulatorio de China se centra en garantizar que la IA esté alineada con los intereses de seguridad nacional y la estabilidad política. **Las Provisiones de Gestión de Servicios de Algoritmos de Recomendación** de 2021 son un ejemplo de cómo el gobierno restringe el uso de algoritmos que puedan afectar la libertad de expresión o la seguridad política. Esta ley regula el contenido generado por IA para evitar que se difunda información que contravenga los valores socialistas (CYBERSPACE ADMINISTRATION OF CHINA, 2021).

2.3.3. Leyes clave en la regulación de IA

En 2023, China implementó nuevas regulaciones clave para la IA generativa a través de las **Medidas Provisionales para la Gestión de los Servicios de Inteligencia Artificial Generativa**. Estas regulaciones, vigentes desde agosto de 2023, incluyen requisitos como:

- Respeto a los valores socialistas fundamentales.
- Prevención de la discriminación en los algoritmos y la selección de datos.
- Protección de la propiedad intelectual.
- Transparencia en los modelos de IA, asegurando que los contenidos creados sean confiables (CYBERSPACE ADMINISTRATION OF CHINA, 2021).

Asimismo, la **Ley de Protección de la Información Personal (PIPL)** y la **Ley de Seguridad de Datos** complementan este marco, imponiendo requisitos estrictos para el tratamiento de datos personales y el consentimiento explícito de los usuarios.

2.4.4. Innovación y crecimiento económico

A pesar de su enfoque restrictivo, China fomenta activamente la innovación en IA. Las autoridades han promovido la creación de iniciativas de compartición de recursos y acceso a datos de alta calidad, permitiendo que las empresas locales desarrollen IA avanzada. El enfoque de China hacia la regulación de la IA, centrado en la seguridad nacional y el control estatal, contrasta significativamente con los marcos más liberales de Europa y Estados Unidos.

Como expone MISKELLY (2024), la protección de derechos de autor para productos generados por IA es un reflejo de los esfuerzos de China por equilibrar el control estatal con el impulso al crecimiento económico. Este enfoque demuestra una clara diferenciación de los modelos regulatorios occidentales, especialmente en cuanto a la flexibilidad y el fomento de la innovación tecnológica en colaboración con el estado.

3. Caso de estudio: NeuroBlock

En el vertiginoso panorama del big data y la tecnología, **NeuroBlock** se presenta como un ejemplo paradigmático de los desafíos éticos y legales que enfrentan las startups dentro del emergente campo de la Inteligencia artificial generativa y el Big data que la acompaña. Como fundador de esta empresa, ubicada en el **Parque Científico de la Universidad Miguel Hernández de Elche**, trabajo en la intersección entre la innovación tecnológica y la compleja regulación que rige el uso de datos. Este desafío abarca tanto el entrenamiento de modelos de inteligencia artificial generativa como su aplicación personalizada a casos de uso específicos y su integración en sistemas de recuperación de datos en tiempo real mediante técnicas como **Retrieval-Augmented Generation (RAG),** técnicas de generación aumentada de respuestas mediante la integración de datos externos al modelo en los sistemas de IA al

momento de generar la respuesta. Este enfoque permite no solo generar respuestas precisas en tiempo real, sino también abordar casos prácticos con un balance entre innovación y cumplimiento normativo, ya que permite integrar datos de forma dinámica, de modo que los modelos IA puedan utilizar cierta información o no de forma dinámica, sorteando así la problemática de los datos con los que sentrema, y que muy difícilmente pueden extirparse del modelo una vez entrenado.

Nuestra propuesta de valor para el **mínimo producto viable (MVP)** que desarrollamos parece, a primera vista, sencilla: una extensión para el navegador web que permite a los usuarios de internet recopilar sus datos de navegación de forma anonimizada, quienes son recompensados con una criptomoneda nativa. Estos datos, combinados con información obtenida mediante **Web Scraping** (técnica consistente en la obtención de datos de internet mediante *crawlers* y otros sistemas automatizados), alimentan algoritmos de inteligencia artificial para generar estudios de mercado accesibles y personalizados para pequeñas y medianas empresas (PYMEs). Sin embargo, esta simplicidad superficial oculta un complejo entramado de desafíos regulatorios, técnicos y éticos que deben ser abordados con rigor para evitar que asfixien la innovación.

En **NeuroBlock**, nuestra visión a largo plazo está profundamente arraigada en una realidad ineludible: en el auge de la inteligencia artificial, donde los datos son un recurso fundamental, se hace imprescindible garantizar que su comercialización sea **justa, transparente y equitativa**. Nuestro propósito no solo es dar respuesta a las necesidades inmediatas de PYMEs y usuarios, sino también sentar las bases para una **economía del dato justa**, en la que los datos no sean explotados indiscriminadamente por grandes corporaciones, sino gestionados y monetizados directamente por sus legítimos propietarios: las personas, las instituciones y las pequeñas empresas.

Para lograrlo, creemos en un paradigma en el que los usuarios pasen de ser meros consumidores pasivos de servicios digitales a convertirse en **sujetos activos**, empoderados para decidir cómo se usan sus datos y cómo estos generan valor. Nuestra misión incluye el desarrollo de una **infraestructura tecnológica descentralizada** que permita a los usuarios gestionar sus datos de forma soberana, segura y rentable. Tecnologías como los **sistemas de identidad descentralizada** (*Decentralized Identifiers*, DIDs) y las **redes de nodos seguros**

desempeñarán un papel clave en este ecosistema, facilitando la portabilidad y la autogestión de los datos.

Este enfoque no solo busca democratizar el acceso a los beneficios del *big data*, sino también incentivar un modelo de negocio más equilibrado, en el que tanto las PYMEs como los usuarios individuales puedan competir en igualdad de condiciones con las grandes corporaciones tecnológicas. A través de la extensión de navegador y tecnologías complementarias en nuestra primera fase, queremos establecer un sistema en el que la recopilación, el intercambio y el análisis de datos se lleven a cabo bajo estrictos estándares éticos y legales, fomentando la innovación sin comprometer los derechos fundamentales.

En definitiva, NeuroBlock aspira a ser un puente entre la tecnología de vanguardia y un modelo económico más justo, donde los datos no sean solo una moneda de cambio para las grandes empresas, sino un recurso accesible y beneficioso para todos. Nuestra razón de ser es clara: transformar la forma en que se gestionan y comercializan los datos en la era de la inteligencia artificial, liderando un camino que combine innovación, ética y equidad.

A continuación, se analiza el proyecto bajo los diversos marcos regulatorios para analizar los posibles retos y soluciones

3.1 Impacto de los marcos regulatorios en Europa

La Unión Europea, como se ha adelantado al comienzo de este trabajo, ha desarrollado un entramado normativo único en el mundo para garantizar que los derechos fundamentales de sus ciudadanos estén protegidos en el entorno digital, al tiempo que se promueve una economía basada en datos. Este marco, compuesto por el Reglamento General de Protección de Datos (GDPR), el Data Governance Act (DGA) y el Data Act, tiene como objetivo equilibrar los derechos individuales con las exigencias de la innovación tecnológica. Sin embargo, para startups como NeuroBlock, cuya propuesta se basa en la inteligencia artificial generativa y el análisis masivo de datos, estas regulaciones pueden suponer tanto barreras como oportunidades para liderar avances tecnológicos respetando la privacidad.

A continuación, se analizan los desafíos más relevantes que estas normativas presentan, con especial énfasis en cómo tecnologías emergentes como el paradigma Web5, las Identidades Descentralizadas (DIDs), los Nodos Web Descentralizados (DWNs) y los datos sintéticos pueden ofrecer soluciones innovadoras que conjuguen el cumplimiento normativo con la sostenibilidad tecnológica.

3.1.1. GDPR: Retos para la innovación y la IA

El GDPR es ampliamente reconocido como el estándar global más avanzado en la protección de datos personales. Aunque esta normativa refuerza derechos como el consentimiento explícito, la minimización de datos y el derecho al olvido, su implementación plantea dificultades significativas para las tecnologías que dependen de grandes volúmenes de información, como la inteligencia artificial generativa.

3.1.1.1. Consentimiento explícito y la IA generativa

El artículo 6 del GDPR establece que el tratamiento de datos personales debe estar respaldado por una base jurídica válida, siendo el consentimiento explícito uno de los pilares fundamentales. Este consentimiento, según el artículo 7, debe ser libre, informado, específico y verificable. Sin embargo, en contextos como el de NeuroBlock, donde se procesan datos anonimizados y masivos para entrenar modelos de IA, garantizar este nivel de consentimiento presenta varios problemas:

- **Ambigüedad en la anonimización:** Aunque el GDPR exime a los datos anonimizados de su ámbito de aplicación (Considerando 26), demostrar que un conjunto de datos ha sido irreversiblemente anonimizado es complejo, especialmente en el contexto del big data, donde los algoritmos pueden reidentificar patrones.
- **Reutilización de datos:** La naturaleza dinámica de los modelos de IA, que suelen reutilizar datos para propósitos múltiples y en constante evolución, entra en conflicto con el principio de granularidad del consentimiento, que exige especificidad en su uso.
- **Transparencia técnica:** El artículo 12 del GDPR requiere que los usuarios comprendan claramente cómo se utilizan sus datos. Sin embargo, explicar el entrenamiento de

modelos de IA y el uso de datos masivos en términos accesibles es un desafío tanto para desarrolladores como para legisladores.

3.1.1.2. Sistema de consentimiento dinámico basados en Web5

Una posible solución se encuentra en el paradigma Web5, desarrollado por TBD de Jack Dorsey, ofrece un enfoque descentralizado donde los usuarios tienen control absoluto sobre sus datos. Este sistema combina Identidades Descentralizadas (DIDs) y Nodos Web Descentralizados (DWNs) para transformar la gestión del consentimiento en un proceso transparente y auditable.

- **DIDs:** Los usuarios gestionan sus datos a través de un wallet digital que almacena sus preferencias y permisos. Esto elimina la dependencia de intermediarios y garantiza la soberanía del usuario sobre sus datos.
- **DWNs:** Los nodos personales descentralizados permiten a los usuarios otorgar, revocar o modificar accesos en tiempo real, garantizando que los datos procesados siempre cuenten con permisos actualizados.

Esta arquitectura puede integrarse con la infraestructura de NeuroBlock mediante contratos inteligentes que registren los accesos y permisos en blockchain. Además, la combinación con técnicas como la anonimización dinámica permite procesar localmente los datos en los dispositivos de los usuarios (edge computing), minimizando riesgos legales y técnicos.

3.1.1.3. Minimización de datos como principio fundamental en conflicto con la IA

El principio de minimización de datos, consagrado en el artículo 5(1)(c) del Reglamento General de Protección de Datos (GDPR), establece que solo deben recopilarse los datos estrictamente necesarios para alcanzar un propósito legítimo. Este principio busca proteger la privacidad de los individuos y reducir la exposición de información personal, promoviendo un tratamiento de datos limitado y controlado.

Sin embargo, en el contexto de la inteligencia artificial generativa, este principio puede entrar en tensión con las necesidades técnicas. Los modelos de IA requieren grandes cantidades de datos para identificar patrones, mitigar sesgos y garantizar predicciones de alta calidad. Por ejemplo, en sectores como la medicina, donde la precisión es crítica, se necesitan datos

diversos y representativos para entrenar sistemas de diagnóstico. Del mismo modo, en aplicaciones como el procesamiento de lenguaje natural o la generación de imágenes, la amplitud de datos disponibles influye directamente en la utilidad del modelo.

Esta aparente contradicción entre los requerimientos normativos y las demandas tecnológicas plantea preguntas cruciales: ¿Cómo equilibrar el cumplimiento legal con el desarrollo de modelos robustos y eficientes? ¿Es posible entrenar IA sin comprometer la privacidad de los usuarios?

Una respuesta prometedora a este dilema radica en el uso de datos sintéticos. Estas herramientas permiten simular conjuntos de datos que mantienen la utilidad técnica, pero eliminan cualquier vínculo directo con datos personales, ofreciendo una solución innovadora al conflicto.

3.1.1.4. Uso de GANs para la creación de datos sintéticos

Los **Modelos Generativos Adversarios (GANs)** se han posicionado como una solución clave para abordar el desafío de la minimización de datos. Estas redes neuronales avanzadas son capaces de generar datos sintéticos que replican los patrones estadísticos de los datos reales sin incluir información personal identificable.

Las GANs operan mediante la interacción de dos componentes principales: un generador, que crea datos falsos pero realistas, y un discriminador, que evalúa la autenticidad de estos datos. Esta dinámica competitiva mejora progresivamente la calidad de los datos generados, hasta que son prácticamente indistinguibles de los originales. Este enfoque permite generar imágenes, textos o incluso conjuntos de datos completos que son útiles para entrenar modelos de IA, pero que no comprometen la privacidad.

El uso de datos sintéticos tiene múltiples beneficios:

- **Cumplimiento normativo:** Al no contener información personal, los datos sintéticos permiten a las organizaciones desarrollar modelos de IA sin infringir el GDPR. ● **Mitigación de sesgos:** Los conjuntos de datos pueden ajustarse deliberadamente para corregir desequilibrios, asegurando que los modelos sean más justos y representativos.
- **Reducción del riesgo de exposición:** Una vez generados los datos sintéticos, los

conjuntos originales pueden eliminarse, reduciendo significativamente el riesgo de fugas o accesos no autorizados.

En términos de implementación, los datos sintéticos pueden integrarse en infraestructuras descentralizadas como **Web5**, donde tecnologías como los **Nodos Web Descentralizados (DWNs)** y las **Identidades Descentralizadas (DIDs)** garantizan la trazabilidad y el control por parte de los usuarios. Por ejemplo, un hospital podría generar datos sintéticos a partir de historiales médicos y almacenarlos en DWNs, asegurando que los investigadores tengan acceso a información útil para entrenar modelos, sin comprometer la privacidad de los pacientes.

Esta combinación de GANs y tecnologías descentralizadas también permite avanzar en el desarrollo de sistemas más transparentes y confiables, donde los usuarios tienen mayor control sobre cómo se gestionan sus datos.

3.1.1.5. Derecho al olvido: Desaprendizaje federado y blockchain para la trazabilidad

El desaprendizaje federado es una propuesta técnica clave para abordar las complejidades del derecho al olvido, establecido en el artículo 17 del Reglamento General de Protección de Datos (GDPR). Este derecho otorga a los usuarios la capacidad de solicitar la eliminación de sus datos personales cuando ya no sean necesarios para los fines por los que fueron recopilados, el consentimiento sea retirado, o el tratamiento sea ilícito. Sin embargo, la implementación de este derecho en sistemas de inteligencia artificial presenta retos únicos, particularmente en modelos que han utilizado estos datos para entrenamiento.

El GDPR exige que, tras una solicitud válida del usuario, los datos personales sean eliminados "sin dilación indebida" (artículo 17(1)). Además, los responsables del tratamiento deben tomar medidas razonables para notificar a terceros que estén procesando los datos sobre esta solicitud (artículo 17(2)). En el contexto de la IA, esto incluye eliminar cualquier rastro de los datos que hayan contribuido al aprendizaje del modelo. Esto implica:

1. Garantizar que los datos eliminados no puedan influir en las decisiones futuras del modelo.
2. Documentar adecuadamente las operaciones realizadas para cumplir con el derecho al olvido.

3.1.1.6. Desaprendizaje federado

El desaprendizaje federado extiende las capacidades del aprendizaje federado al permitir la eliminación de la influencia de datos específicos en un modelo sin necesidad de reentrenar desde cero. Esta técnica es particularmente útil en entornos descentralizados, donde los datos se almacenan y procesan localmente en dispositivos o servidores.

En síntesis, el funcionamiento sería el siguiente:

1. **Identificación de datos relevantes:** El sistema identifica los datos que deben eliminarse del modelo, siguiendo la solicitud del usuario.
2. **Eliminación distribuida:** Los dispositivos locales que participan en el aprendizaje federado eliminan la influencia de estos datos en los modelos entrenados. 3. **Actualización del modelo global:** Las correcciones realizadas en los modelos locales se integran en el modelo central sin necesidad de reentrenamiento completo, reduciendo significativamente los costos y tiempos operativos.

Este enfoque cumple con los requisitos del artículo 17 del GDPR al:

- Reducir la dependencia de servidores centralizados, minimizando riesgos de privacidad (artículo 5(1)(f): integridad y confidencialidad).
- Ofrecer una solución eficaz y documentada para eliminar datos en sistemas distribuidos.

El uso de blockchain en este contexto permite registrar de manera inmutable y verificable las solicitudes de eliminación y las acciones realizadas para cumplir con ellas. Esto responde al principio de responsabilidad proactiva (artículo 5(2) del GDPR), al proporcionar un medio transparente para demostrar el cumplimiento del derecho al olvido.

Beneficios de integrar blockchain:

1. **Trazabilidad**: Permite registrar cuándo y cómo se eliminaron los datos, así como las actualizaciones realizadas al modelo.
2. **Evitar reutilización**: Los datos marcados para eliminación no podrán ser

reincorporados al sistema, asegurando su exclusión total.

3. **Auditoría**: Facilita las inspecciones de las autoridades de protección de datos y aumenta la confianza del usuario.

Imaginemos un sistema de IA en el sector sanitario. Un paciente solicita la eliminación de su historial médico del modelo que predice diagnósticos. Con el desaprendizaje federado, los hospitales participantes eliminan localmente la influencia de estos datos en los modelos. Paralelamente, una transacción en blockchain documenta que:

- La solicitud fue recibida.
- Los datos fueron eliminados.
- Los modelos fueron actualizados para reflejar este cambio.

Pese a todo, aunque el desaprendizaje federado y el blockchain pueden suponer avances importantes para garantizar el cumplimiento del derecho al olvido, su implementación enfrenta desafíos clave, como la complejidad técnica de coordinar eliminaciones en sistemas distribuidos con grandes volúmenes de datos y dispositivos heterogéneos. Además, es fundamental que estas tecnologías se diseñen en cumplimiento con el GDPR, particularmente con el artículo 25 sobre privacidad desde el diseño, y que la trazabilidad ofrecida por el blockchain sea admisible como prueba en posibles litigios, asegurando tanto la compatibilidad normativa como la robustez legal.

3.2. DGA y Data Act: Gobernanza y portabilidad ética en la economía del dato

El Data Governance Act (DGA) y el Data Act representan un pilar clave en la construcción de una economía digital europea basada en la confianza, la transparencia y la equidad en el manejo de datos. Estas normativas se conciben como herramientas para fomentar un entorno en el que los datos puedan ser compartidos de manera segura y eficiente entre entidades públicas y privadas, sin comprometer los derechos individuales ni la confidencialidad empresarial. Este enfoque busca desbloquear el potencial económico de los datos al tiempo que establece salvaguardias robustas para proteger la privacidad y la soberanía de los usuarios.

Para startups como NeuroBlock, cuyo modelo de negocio depende de la recopilación y análisis de datos anonimizados, estas regulaciones presentan una doble vertiente: por un lado, abren nuevas oportunidades de colaboración y confianza en un ecosistema digital ético; por otro, plantean desafíos técnicos y estratégicos que exigen una adaptación continua.

3.2.1. Data Governance Act (DGA):

El Data Governance Act (Reglamento (UE) 2022/868) establece un marco regulatorio para facilitar el intercambio de datos mediante intermediarios de confianza. Este reglamento no solo reconoce la importancia de los datos como recurso económico estratégico, sino que también subraya la necesidad de garantizar la soberanía de los usuarios y la neutralidad de los intermediarios. En esencia, el DGA establece un puente entre la innovación tecnológica y la protección de derechos fundamentales.

Pilares fundamentales del DGA:

1. **Neutralidad de los intermediarios:**

   El DGA requiere que los intermediarios actúen exclusivamente como facilitadores, sin explotar los datos para fines propios. Esto está regulado explícitamente en el artículo 11, que obliga a los intermediarios a operar bajo estrictos estándares de imparcialidad. Para NeuroBlock, cumplir con esta disposición implica diseñar sistemas que puedan ser auditados externamente para demostrar transparencia, asegurando que los datos procesados no sean reutilizados de manera indebida.

2. **Control y trazabilidad:**

   Según el artículo 5, el DGA otorga a los usuarios el derecho de rastrear y controlar el uso de sus datos incluso después de compartirlos. Este requisito exige la implementación de herramientas avanzadas, como contratos inteligentes y sistemas de auditoría basados en blockchain, que permitan a los usuarios monitorizar y revocar accesos en tiempo real. NeuroBlock podría aprovechar estas tecnologías para garantizar la trazabilidad y la confianza, alineándose con los valores del reglamento.

3. **Interoperabilidad entre dominios:**

La interoperabilidad es un objetivo central del DGA, tal como se establece en el artículo 12(i). Sin embargo, garantizar la interoperabilidad en un ecosistema digital diverso y masivo es un desafío técnico significativo. Para NeuroBlock, esto implica desarrollar arquitecturas que puedan integrar datos de múltiples sectores y dominios sin comprometer la seguridad ni la eficiencia operativa.

3.2.2. Plataforma de intercambio ético basada en Web5

El paradigma Web5, en este sentido, gracias a sus herramientas descentralizadas, se presenta como el núcleo de una solución para cumplir con el DGA. Algunas propuestas técnicas incluyen:

1. **Contratos inteligentes:** Automatizan el intercambio de datos asegurando que solo se utilicen para fines específicos previamente acordados por los usuarios. Estos contratos también pueden incluir cláusulas que permitan revocar accesos automáticamente si las condiciones cambian.
2. **Cifrado homomórfico:** Este tipo de cifrado permite realizar cálculos directamente sobre datos encriptados, garantizando que la información sensible nunca sea expuesta durante el proceso de análisis.
3. **Auditorías descentralizadas:** Al almacenar registros de acceso y uso de datos en blockchain, NeuroBlock puede ofrecer una trazabilidad inmutable que cumpla con los requisitos del DGA.

3.2.3. Data Act:

El Data Act (Reglamento (UE) 2022/868) representa un pilar esencial en la estrategia de la Unión Europea para garantizar que los datos generados en la economía digital sean accesibles, reutilizables y gestionados de manera equitativa. Este reglamento, enmarcado dentro de la política de la UE para una economía basada en datos, busca equilibrar los intereses de los usuarios, las empresas y los desarrolladores tecnológicos al fomentar la portabilidad de datos generados por dispositivos conectados y otros sistemas digitales. Su implementación refuerza el control de los usuarios sobre su información personal y

empresarial, al tiempo que promueve la competencia, la innovación y la interoperabilidad en los mercados digitales.

Principales retos:

- **Portabilidad de datos en sistemas propietarios:** Portabilidad de datos en sistemas propietarios: Los artículos 3 y 4 del Data Act establecen las obligaciones y derechos relacionados con el acceso a datos. El artículo 3 requiere que los productos conectados sean diseñados y fabricados para que los datos sean fácilmente accesibles para el usuario por defecto, mientras que el artículo 4 detalla los derechos específicos de los usuarios y las obligaciones de los data holders para hacer disponibles los datos. Para NeuroBlock, esto implica adaptar sus sistemas internos para garantizar la compatibilidad con plataformas y dispositivos de terceros, adoptando formatos estándar y protocolos abiertos que faciliten la interoperabilidad.
- **Protección de secretos empresariales:** El Data Act equilibra la apertura de datos con la protección de información sensible. Específicamente, el artículo 4(3) y el artículo 5(8) establecen que los secretos comerciales solo se divulgarán si se toman todas las medidas específicas necesarias para preservar su confidencialidad. Para NeuroBlock, esto significa implementar medidas técnicas y organizativas adecuadas, como el cifrado homomórfico o contratos inteligentes, para proteger la información sensible mientras cumple con las obligaciones de compartir datos.
- **Garantías de no discriminación:** Los artículos 8 y 13 del Data Act establecen las condiciones para un acceso justo y no discriminatorio a los datos. El artículo 8 requiere que los términos para compartir datos sean justos, razonables y no discriminatorios, mientras que el artículo 13 protege contra términos contractuales injustos impuestos unilateralmente. Para NeuroBlock, es fundamental implementar sistemas transparentes que aseguren un trato equitativo en el acceso y uso de los datos.

Para alinearse con los requisitos del Data Act y superar los desafíos asociados, NeuroBlock puede adoptar varias soluciones tecnológicas innovadoras:

1. **Interoperabilidad mediante estándares abiertos:**

Adoptar protocolos y formatos reconocidos internacionalmente, como los promovidos por el W3C o la ISO, permite que los datos compartidos sean comprensibles y reutilizables por múltiples sistemas. Esto no solo facilita la portabilidad, sino que también reduce las barreras técnicas y promueve la competencia en mercados saturados por soluciones propietarias.

2. **Segmentación y anonimización avanzada:**

   Las técnicas de anonimización y la generación de datos sintéticos mediante redes adversarias generativas (GANs) permiten procesar datos sensibles sin comprometer la privacidad. Además, el uso del edge computing permite realizar análisis de datos localmente en dispositivos conectados, reduciendo la necesidad de transferir grandes volúmenes de información a servidores centrales y garantizando la seguridad de los datos.

3. **Sistemas de gobernanza descentralizada:**

   Los nodos web descentralizados (DWNs) permiten gestionar permisos de acceso a los datos en tiempo real. Con estas tecnologías, los usuarios y las empresas pueden definir quién tiene acceso a sus datos, durante cuánto tiempo y bajo qué condiciones, asegurando un control efectivo y transparente.

4. **Auditorías basadas en blockchain:**

   Utilizar blockchain para registrar el acceso y uso de datos proporciona una trazabilidad inmutable que cumple con las disposiciones del Data Act. Esto no solo fortalece la confianza en los sistemas de datos, sino que también facilita la verificación del cumplimiento normativo por parte de las autoridades regulatorias.

3.2.4 Potencial de NeuroBlock en el marco del DGA y el Data Act

Ambas normativas representan tanto retos como oportunidades para startups como NeuroBlock. Si bien es necesario invertir en la adaptación tecnológica para cumplir con los requisitos legales, estas regulaciones también abren nuevas vías para generar confianza y establecer un modelo de negocio ético y competitivo. La implementación de un sistema basado en Web5 no solo facilita el cumplimiento normativo, sino que posiciona a NeuroBlock como un referente en la economía digital ética.

En suma, el marco regulatorio europeo, compuesto por el GDPR, el DGA y el Data Act, establece las bases para una economía digital basada en la confianza, la protección de datos y la interoperabilidad. Aunque estas normativas presentan desafíos significativos para las startups que dependen de tecnologías emergentes, también ofrecen la posibilidad de liderar en innovación ética y técnica.

NeuroBlock, con su enfoque en tecnologías descentralizadas como Web5, tiene la capacidad de superar estos retos y establecer un estándar pionero en la gestión de datos. Este modelo no solo asegura el cumplimiento normativo, sino que también impulsa la sostenibilidad tecnológica y el respeto a los derechos fundamentales en la economía del dato.

3.2 Marcos regulatorios en Estados Unidos, desafíos y estrategias para NeuroBlock

El panorama regulatorio en Estados Unidos se caracteriza por su fragmentación y complejidad. A diferencia de la Unión Europea, que ha adoptado un enfoque integral a través del GDPR, Estados Unidos carece de una legislación federal unificada en materia de protección de datos y regulación de la inteligencia artificial (IA). En su lugar, existe una combinación de regulaciones federales sectoriales, leyes estatales y la supervisión de agencias como la **Federal Trade Commission (FTC)**. Para una startup como **NeuroBlock**, que opera en el ámbito de la IA generativa y el análisis masivo de datos, navegar este laberinto legal es tanto un desafío como una oportunidad para innovar en prácticas éticas y de cumplimiento normativo.

3.2.1. La Federal Trade Commission (FTC)

La FTC es la agencia federal encargada de proteger a los consumidores y promover la competencia, con un mandato que se centra en prevenir "prácticas comerciales desleales o engañosas", según lo establecido en la **Sección 5 de la FTC Act**. Sin embargo, esta legislación, que data de 1914, no fue diseñada para abordar los complejos desafíos que presentan la IA y el big data en el siglo XXI. A medida que la tecnología avanza, la FTC ha intentado adaptar su autoridad para enfrentar nuevas situaciones, pero enfrenta limitaciones legales y

constitucionales.

La falta de un marco legal específico para la IA y la protección de datos obliga a la FTC a basarse en interpretaciones amplias de su mandato, generando inseguridad jurídica para las empresas. Esto se ve reflejado en casos como el de FTC v. EVERALBUM, INC. (2021), donde la FTC sancionó a la empresa por utilizar fotos y videos de usuarios para entrenar algoritmos de reconocimiento facial sin obtener su consentimiento explícito y por no cumplir con las promesas de eliminación de datos.

Este caso es particularmente relevante para NeuroBlock, ya que pone de manifiesto la importancia de:

- **Consentimiento informado y explícito**: Las empresas deben asegurarse de que los usuarios comprendan cómo se utilizarán sus datos, especialmente cuando se trata de entrenamiento de modelos de IA.
- **Transparencia en las prácticas de datos**: Es esencial que las políticas de privacidad reflejen con precisión las prácticas reales y que se cumplan rigurosamente. ● **Cumplimiento proactivo**: Anticiparse a posibles interpretaciones regulatorias y establecer estándares éticos más allá de las obligaciones legales mínimas.

Para abordar estos desafíos, NeuroBlock podría desarrollar un **Sistema de Transparencia Algorítmica** que permita a los usuarios tener un control granular sobre sus datos y comprender claramente cómo se utilizan. Este sistema incluiría:

- **Registro inmutable de transacciones de datos**: Utilizando tecnología blockchain, se crearía un historial transparente y auditable de todas las interacciones con los datos de los usuarios, fortaleciendo la confianza y facilitando el cumplimiento regulatorio.
- **Gestión automatizada del consentimiento**: A través de contratos inteligentes, se automatizaría el proceso de obtención y revocación del consentimiento, permitiendo a los usuarios modificar sus preferencias en tiempo real.
- **Interfaces de usuario intuitivas**: Herramientas que faciliten a los usuarios visualizar el uso de sus datos y ejercer sus derechos de manera sencilla, fomentando una relación más transparente y colaborativa.

Implementar este sistema no solo alinearía a NeuroBlock con las expectativas de la FTC, sino que también la posicionaría como líder en prácticas éticas y responsables en el manejo de datos, lo que puede ser un diferenciador competitivo en el mercado.

3.2.2. El enfoque de California hacia la privacidad del consumidor

La **California Consumer Privacy Act (CCPA)**, en vigor desde enero de 2020, es uno de los intentos más significativos a nivel estatal para proteger la privacidad de los consumidores en Estados Unidos. Aunque comparte similitudes con el GDPR europeo, existen diferencias clave que afectan cómo las empresas deben abordar el cumplimiento.

Una de las principales diferencias es el modelo de consentimiento. Mientras que el GDPR requiere un consentimiento explícito (opt-in) antes de procesar datos personales, la CCPA opera bajo un modelo de opt-out, presumiendo el consentimiento hasta que el usuario decide retirarlo. Esto refleja una filosofía distinta sobre la protección de datos y plantea desafíos particulares para las empresas que operan a nivel internacional.

Además, la CCPA se aplica a empresas que cumplen ciertos criterios específicos, como ingresos superiores a 25 millones de dólares, manejo de datos de más de 50,000 residentes de California o que obtienen más del 50% de sus ingresos de la venta de datos personales. Esto crea un escenario donde no todas las empresas están obligadas a cumplir con la CCPA, a diferencia del GDPR, que tiene un alcance más universal.

Para NeuroBlock, cumplir con la CCPA implica:

- **Implementar mecanismos claros de opt-out**: Incluir enlaces visibles como "No vender mis datos personales" y establecer procesos eficientes para gestionar estas solicitudes.
- **Gestionar solicitudes de acceso y eliminación de datos**: Estar preparados para responder a las solicitudes de los usuarios dentro de los plazos legales y asegurar la capacidad de eliminar o proporcionar los datos requeridos.
- **Segmentar a los usuarios por jurisdicción**: Identificar a los residentes de California para aplicar las disposiciones de la CCPA de manera adecuada, evitando tanto el

incumplimiento como la aplicación excesiva de restricciones.

Uno de los mayores desafíos es cómo eliminar datos personales que ya han sido integrados en modelos de IA sin afectar la funcionalidad y precisión de estos modelos. Para resolver este problema, se propone la implementación de un **Sistema de Datos Modulares**, que permitiría:

- **Desagregación de datos**: Almacenar datos de forma modular facilita la identificación y eliminación de información específica sin comprometer el conjunto de datos global.
- **Uso de datos sintéticos**: Reemplazar los datos eliminados con datos sintéticos generados mediante técnicas avanzadas como Redes Generativas Antagónicas (GANs), manteniendo así el rendimiento del modelo de IA.
- **Flexibilidad para futuras regulaciones**: Este enfoque modular facilita la adaptación a nuevas leyes y requisitos, evitando la necesidad de rediseñar la arquitectura de datos ante cambios legislativos.

Adoptar estas medidas no solo asegura el cumplimiento con la CCPA, sino que también demuestra un compromiso con la privacidad y el control del usuario, lo que puede mejorar la reputación de NeuroBlock y fortalecer su posición competitiva.

3.2.3. Regulaciones específicas y su impacto en NeuroBlock

El enfoque sectorial de la regulación en Estados Unidos añade una capa adicional de complejidad al panorama legal. Existen leyes específicas para sectores como la salud, la educación y las finanzas, cada una con sus propias definiciones, requisitos y obligaciones. Algunas de las regulaciones más relevantes incluyen:

- **Health Insurance Portability and Accountability Act (HIPAA)**: Protege la información médica y establece estándares para su manejo y transmisión.
- **Family Educational Rights and Privacy Act (FERPA)**: Regula el acceso y divulgación de información educativa, otorgando derechos específicos a estudiantes y padres.
- **Gramm-Leach-Bliley Act (GLBA)**: Requiere que las instituciones financieras protejan la información financiera no pública y expliquen sus prácticas de intercambio de información.

Para una empresa como NeuroBlock, que puede manejar datos de diferentes sectores, esto implica:

- **Necesidad de identificar y cumplir con múltiples regulaciones**: Dependiendo del tipo de datos procesados, se aplican diferentes leyes, lo que requiere una gestión cuidadosa y detallada del cumplimiento.
- **Riesgo de superposición y conflictos normativos**: Los datos pueden estar sujetos a más de una regulación, generando incertidumbre sobre qué requisitos aplicar en cada caso.
- **Carga administrativa significativa**: Mantenerse al día con las regulaciones de cada sector y adaptar los procesos internos puede ser costoso y complejo.

Para enfrentar estos desafíos, se propone el desarrollo de un **Sistema de Cumplimiento Adaptativo basado en IA**, diseñado para:

- **Analizar los datos en tiempo real**: Utilizar algoritmos de aprendizaje automático para clasificar automáticamente los datos según su naturaleza y determinar las regulaciones aplicables.
- **Ajustar dinámicamente las políticas de cumplimiento**: Una vez identificadas las regulaciones pertinentes, el sistema adaptaría automáticamente los protocolos de seguridad y manejo de datos para asegurar el cumplimiento.
- **Mantenerse actualizado con cambios legales**: Integrar un módulo de inteligencia regulatoria que monitoree cambios en las leyes y actualice las políticas internas en consecuencia.
- **Facilitar auditorías y reportes**: Generar registros detallados y reportes de cumplimiento que simplifiquen el proceso de auditoría y demuestren el compromiso con la conformidad legal.

Este enfoque permite a NeuroBlock manejar la complejidad regulatoria de manera eficiente, reduciendo el riesgo de incumplimiento y permitiendo una mayor flexibilidad para explorar oportunidades en diferentes sectores.

3.2.4. El desafío y la oportunidad de la ambigüedad legal

La falta de una legislación federal unificada en Estados Unidos presenta desafíos significativos para empresas tecnológicas, pero también ofrece una oportunidad para liderar en prácticas éticas y establecer estándares que podrían influir en futuras regulaciones. Adoptar una estrategia proactiva es esencial para navegar este entorno incierto.

NeuroBlock puede beneficiarse al:

- **Adoptar voluntariamente estándares internacionales estrictos**: Implementar prácticas alineadas con el GDPR y otras regulaciones rigurosas puede preparar a la empresa para futuras leyes y demostrar un compromiso sólido con la privacidad.
- **Participar activamente en el diálogo regulatorio**: Involucrarse en foros, colaboraciones con organismos reguladores y grupos de trabajo puede permitir a NeuroBlock influir en la dirección de futuras legislaciones y aportar perspectivas valiosas basadas en su experiencia.
- **Fomentar la transparencia y la educación**: Comunicar claramente a los usuarios cómo se utilizan sus datos y educarlos sobre sus derechos fortalece la confianza y promueve una relación más sólida con los clientes.

Al adoptar estas medidas, NeuroBlock no solo se protege frente a posibles riesgos legales sino que también se posiciona como líder en ética y responsabilidad, lo que puede traducirse en ventajas competitivas y de reputación a largo plazo.

En definitiva, el entorno regulatorio en Estados Unidos es indudablemente complejo y fragmentado, presentando desafíos significativos para empresas que operan en el ámbito de la IA y el manejo masivo de datos. Sin embargo, estos desafíos también abren la puerta a oportunidades para innovar en prácticas de cumplimiento y establecer estándares éticos elevados.

Al desarrollar soluciones como el **Sistema de Transparencia Algorítmica**, el **Sistema de Datos**

Modulares y el **Sistema de Cumplimiento Adaptativo basado en IA**, NeuroBlock puede convertir los retos legales en ventajas estratégicas. Estos sistemas no solo facilitan el cumplimiento con las regulaciones actuales, sino que también preparan a la empresa para adaptarse rápidamente a futuros cambios legales.

La combinación de innovación tecnológica con un compromiso genuino con la ética y la transparencia posiciona a NeuroBlock como un líder potencial en su sector. Al adelantarse a las exigencias legales y fomentar prácticas centradas en el respeto a los derechos individuales, la empresa contribuye a construir un ecosistema tecnológico más responsable y sostenible.

En un contexto donde la regulación de la IA y el manejo de datos están en constante evolución, adoptar una postura proactiva y ética no es solo una estrategia de cumplimiento, sino una inversión en el futuro éxito y relevancia de la empresa. NeuroBlock tiene la oportunidad de liderar este camino, estableciendo un ejemplo para la industria y ayudando a dar forma al futuro de la protección de datos y la ética en la inteligencia artificial en Estados Unidos y más allá.

3.3 China: Ley de Protección de la Información Personal (PIPL) y Ley de Seguridad de Datos (DSL)

En el vasto y complejo panorama legal de China, el marco regulatorio se distingue por su rigurosidad y centralización, reflejando un énfasis particular en la seguridad nacional y la soberanía de los datos. Las principales leyes que impactan directamente a empresas como **NeuroBlock** son la **Ley de Protección de la Información Personal (PIPL)** y la **Ley de Seguridad de Datos (DSL)**, ambas promulgadas en 2021. Estas normativas establecen un entramado legal detallado y exigente para el tratamiento, almacenamiento y transferencia de datos personales y sensibles, imponiendo responsabilidades significativas a las organizaciones que operan en el territorio chino o que procesan datos de ciudadanos chinos.

3.3.1. Ley de Protección de la Información Personal (PIPL): Derechos del usuario y obligaciones empresariales

La PIPL representa el primer esfuerzo integral de China para regular la protección de la información personal, y aunque guarda similitudes con el Reglamento General de Protección de Datos (GDPR) de la Unión Europea, presenta características únicas que reflejan las prioridades y estructuras del sistema legal y político chino.

La ley establece que cualquier entidad que procese información personal debe hacerlo de manera legal, justa y necesaria, y debe limitarse a los mínimos propósitos necesarios para cumplir con sus funciones. Esto implica que **NeuroBlock** debe revisar cuidadosamente sus prácticas de recopilación y procesamiento de datos para garantizar que se ajusten a estos principios fundamentales.

Uno de los pilares centrales de la PIPL es el **consentimiento explícito e informado**. Según el artículo 14, el procesamiento de información personal requiere el consentimiento claro y voluntario del individuo, quien debe ser plenamente consciente de los propósitos, métodos y alcance del procesamiento. Esto significa que **NeuroBlock** debe proporcionar políticas de privacidad y términos de uso que sean transparentes y comprensibles, redactados en un lenguaje claro y accesible, y disponibles en chino mandarín. Es esencial evitar términos ambiguos o generales; en su lugar, se debe especificar detalladamente cómo se utilizarán los datos, para qué fines y durante cuánto tiempo.

Además, la PIPL impone **restricciones estrictas a las transferencias transfronterizas de datos**. Los artículos 38 al 43 establecen que las empresas deben cumplir con ciertos requisitos antes de transferir información personal fuera de China. Estas medidas incluyen pasar evaluaciones de seguridad organizadas por la Administración del Ciberespacio de China (CAC), obtener certificaciones de protección de información personal o firmar contratos con el destinatario extranjero que cumplan con los estándares establecidos por el CAC. Para **NeuroBlock**, esto implica que cualquier transferencia de datos personales a servidores o entidades fuera de China debe ser cuidadosamente evaluada y posiblemente adaptada para cumplir con estos requisitos, lo que puede requerir cambios significativos en la infraestructura tecnológica y en

las políticas de gestión de datos.

La ley también otorga a los individuos una serie de **derechos sobre sus datos personales**, como acceder a ellos, solicitar correcciones o eliminaciones, y retirar el consentimiento en cualquier momento. Esto obliga a **NeuroBlock** a establecer mecanismos eficientes y fáciles de usar que permitan a los usuarios ejercer estos derechos sin obstáculos. Por ejemplo, se podrían desarrollar herramientas en línea que permitan a los usuarios ver qué datos se han recopilado sobre ellos, solicitar correcciones si detectan errores o eliminar su información si así lo desean. Además, es crucial que la empresa responda a estas solicitudes dentro de los plazos establecidos por la ley, demostrando así su compromiso con la protección de los derechos de los usuarios.

La **información personal sensible** recibe una atención especial bajo la PIPL. Esta categoría incluye datos biométricos, de salud, financieros y de localización, entre otros. El procesamiento de este tipo de información requiere un consentimiento separado y explícito, y la implementación de medidas de seguridad adicionales. **NeuroBlock** debe, por tanto, identificar si maneja este tipo de datos y, en caso afirmativo, asegurarse de que cumple con todas las obligaciones adicionales, lo que puede incluir cifrado avanzado, controles de acceso estrictos y procedimientos específicos de manejo.

Los desafíos legales que plantea la PIPL para **NeuroBlock** son significativos. Garantizar la claridad y especificidad del consentimiento puede ser particularmente complejo, ya que implica diseñar procesos que no solo cumplan con la ley sino que también sean prácticos y no obstaculicen la experiencia del usuario. Además, las restricciones en la transferencia de datos pueden requerir la creación de infraestructura local, como centros de datos en China, lo que supone una inversión considerable. También es esencial mantener un cumplimiento continuo, dado que las autoridades chinas pueden realizar auditorías y exigir informes en cualquier momento.

Para abordar estos desafíos, **NeuroBlock** puede implementar soluciones técnicas como un **Sistema de Consentimiento Granular**, que permita a los usuarios otorgar consentimiento específico para cada tipo de procesamiento, mediante interfaces intuitivas y en su idioma nativo. Esto no solo cumple con los requisitos legales sino que también mejora la transparencia y fortalece la confianza del usuario en la empresa.

Asimismo, establecer **infraestructura local de datos** es una estrategia crucial. Al crear centros de datos dentro de China, **NeuroBlock** puede asegurar que los datos personales de los usuarios chinos se almacenen y procesen localmente, evitando la necesidad de transferencias internacionales y cumpliendo con las restricciones impuestas por la PIPL.

Por último, desarrollar **herramientas para el ejercicio de los derechos de los usuarios** es fundamental. Al proporcionar módulos en la plataforma que permitan acceder, corregir y eliminar datos de manera sencilla, **NeuroBlock** no solo cumple con la ley sino que también demuestra un compromiso genuino con la privacidad y el control del usuario sobre su información personal.

3.3.2. Ley de Seguridad de Datos (DSL): Enfoque en la seguridad nacional y gestión de riesgos

La **Ley de Seguridad de Datos (DSL)** complementa a la PIPL al centrarse en la seguridad de los datos desde una perspectiva de seguridad nacional y económica. Establece un marco para la clasificación y gestión de datos según su importancia, imponiendo obligaciones adicionales a las empresas que manejan datos considerados críticos.

Una de las disposiciones clave de la DSL es la **clasificación y categorización de datos**. Según el artículo 21, las empresas deben clasificar los datos que procesan basándose en su nivel de importancia para la seguridad nacional, el interés público y los derechos individuales. Esto requiere que **NeuroBlock** implemente sistemas que permitan identificar y categorizar los datos de acuerdo con los criterios establecidos, lo que puede implicar análisis detallados y continuos de los tipos de datos que se manejan y cómo se utilizan.

Además, la ley exige que las empresas realicen **evaluaciones de riesgo periódicas** y adopten **medidas de seguridad adecuadas**, especialmente cuando manejan datos que son considerados importantes o críticos para la nación. Esto implica que **NeuroBlock** debe invertir en sistemas de seguridad avanzados y mantenerse al día con las mejores prácticas en ciberseguridad. No se trata solo de proteger los datos contra posibles ataques cibernéticos, sino también de demostrar a las autoridades chinas que se están tomando todas las precauciones necesarias para salvaguardar la información.

Otro aspecto crucial de la DSL es la **restricción a la exportación de datos**. El artículo 31 impone controles estrictos sobre la exportación de datos que puedan afectar la seguridad nacional o el interés público, incluyendo posibles prohibiciones o la necesidad de obtener una aprobación previa de las autoridades. Para **NeuroBlock**, esto significa que cualquier transferencia de datos que pueda considerarse sensible está sujeta a un escrutinio riguroso. Es posible que se requiera obtener permisos especiales o incluso que ciertas transferencias estén completamente prohibidas, lo que puede afectar significativamente las operaciones internacionales de la empresa.

Los desafíos legales y técnicos que plantea la DSL para **NeuroBlock** son considerables. Identificar con precisión qué datos son considerados sensibles o críticos según los estándares chinos puede ser complejo, especialmente porque las definiciones pueden ser amplias y sujetas a interpretaciones. Por lo tanto, es esencial adoptar un enfoque conservador y posiblemente consultar con expertos locales para garantizar el cumplimiento.

En términos técnicos, la implementación de **medidas de seguridad avanzadas** es ineludible. Esto puede incluir la encriptación de datos tanto en tránsito como en reposo, el uso de redes privadas virtuales (VPN) seguras, la implementación de sistemas de detección y prevención de intrusiones, y la realización de pruebas de penetración y evaluaciones de vulnerabilidades de manera regular. Además, es crucial establecer **protocolos claros de respuesta a incidentes**, que incluyan procedimientos para notificar a las autoridades y a los usuarios en caso de una brecha de seguridad.

Para gestionar eficazmente el cumplimiento y los reportes requeridos por la DSL, **NeuroBlock** debe establecer un equipo dedicado a la **gestión de cumplimiento normativo**. Este equipo sería responsable de mantenerse actualizado sobre cualquier cambio en las regulaciones, coordinar las evaluaciones de riesgo, preparar los reportes necesarios y servir como punto de contacto con las autoridades regulatorias. La documentación detallada y la transparencia son fundamentales en este proceso, ya que demuestran a las autoridades el compromiso de la empresa con el cumplimiento y la seguridad.

3.3.3. Integración del paradigma Web5 en el contexto regulatorio chino

El paradigma **Web5**, que promueve la descentralización y el control del usuario sobre sus datos, presenta tanto oportunidades como desafíos en el marco legal chino. Por un lado, algunas de sus características pueden ayudar a cumplir con los requisitos de la PIPL y la DSL, especialmente en lo que respecta al consentimiento explícito y la transparencia. Por otro lado, la naturaleza descentralizada de Web5 puede entrar en conflicto con el enfoque centralizado y el control estatal que caracteriza al sistema chino.

Una de las áreas donde **Web5** puede ser beneficioso es en el **empoderamiento del usuario y el consentimiento explícito**. Al permitir que los usuarios gestionen directamente sus datos y otorguen permisos específicos para su uso, se alinea con la exigencia de la PIPL de obtener un consentimiento informado y detallado. **NeuroBlock** puede utilizar tecnologías como los **Identificadores Descentralizados (DIDs)** y los **Nodos Web Descentralizados (DWNs)** para desarrollar interfaces que faciliten este control por parte del usuario.

Además, la **transparencia y trazabilidad** que ofrece la tecnología blockchain, componente fundamental de Web5, puede ser una herramienta valiosa para demostrar el cumplimiento de las regulaciones. Al registrar todas las transacciones y movimientos de datos de manera inmutable, se crea un historial que puede ser auditado tanto internamente como por las autoridades, lo que refuerza la confianza y la responsabilidad.

Sin embargo, existen **desafíos significativos**. El enfoque descentralizado de Web5 puede ser visto con recelo por las autoridades chinas, que priorizan el control estatal y la supervisión de las infraestructuras críticas y los datos. Es posible que se requiera que todos los nodos y sistemas estén alojados dentro de China y bajo jurisdicción china, lo que limita la descentralización y puede afectar la interoperabilidad con sistemas globales.

Además, la adopción de tecnologías no aprobadas o desconocidas puede ser restringida, especialmente si se perciben riesgos para la seguridad nacional. Por lo tanto, es esencial que **NeuroBlock** trabaje en estrecha colaboración con las autoridades y socios locales para asegurar que las tecnologías utilizadas sean aceptables y cumplan con las normativas.

Para superar estos desafíos, **NeuroBlock** puede considerar el desarrollo de una **versión adaptada de Web5**, que cumpla con los requisitos chinos y se integre de manera armoniosa en el ecosistema local. Esto podría implicar alojar todos los componentes dentro de China, utilizar estándares y protocolos aprobados, y posiblemente limitar ciertas funcionalidades para alinearse con las regulaciones.

Otra estrategia es la **colaboración activa con las autoridades y entidades locales**. Al establecer asociaciones con empresas chinas y mantener un diálogo abierto con los reguladores, **NeuroBlock** puede obtener una comprensión más profunda de las expectativas y requisitos, facilitando el cumplimiento y la aceptación de sus soluciones.

Finalmente, el desarrollo de **soluciones híbridas** que combinen elementos centralizados y descentralizados puede ser una vía viable. De esta manera, se pueden aprovechar las ventajas de Web5 en términos de control y transparencia, al tiempo que se satisfacen las necesidades de supervisión y control estatal.

3.3.4. Conclusión

El marco legal chino, representado por la **PIPL** y la **DSL**, presenta un conjunto de desafíos únicos para empresas internacionales como **NeuroBlock**. La prioridad otorgada a la seguridad nacional, el control estatal y la soberanía de los datos exige que las empresas adapten sus prácticas y tecnologías de manera significativa.

Sin embargo, estos desafíos también ofrecen una oportunidad para innovar y establecer nuevos estándares en el manejo ético y responsable de datos. Al comprender profundamente las disposiciones legales y adoptar soluciones técnicas avanzadas, **NeuroBlock** puede no solo cumplir con las regulaciones sino también fortalecer su posición en el mercado chino.

La implementación de sistemas de consentimiento granular, infraestructuras locales de datos, medidas de seguridad avanzadas y la adaptación del paradigma Web5 son pasos clave en esta dirección. Además, la colaboración con autoridades y socios locales es esencial para navegar el entorno regulatorio y cultural.

En última instancia, el éxito de **NeuroBlock** en China dependerá de su capacidad para integrar de manera armoniosa las exigencias legales con la innovación tecnológica, respetando las particularidades culturales y regulatorias del país. Este enfoque puede servir como modelo para operar en otros mercados con marcos legales estrictos, fortaleciendo la posición de la empresa como líder en ética, cumplimiento y excelencia técnica en el ámbito global.

4. Análisis crítico y propuestas de mejora de los marcos regulatorios en la UE, EE.UU. y China para NeuroBlock

En la era contemporánea, la inteligencia artificial (IA) y el Big Data han transformado radicalmente la manera en que las empresas operan y generan valor. Startups innovadoras como **NeuroBlock** se encuentran a la vanguardia de esta revolución tecnológica, aprovechando vastos volúmenes de datos para desarrollar soluciones avanzadas. Sin embargo, esta innovación no está exenta de desafíos, especialmente en el ámbito regulatorio. Las normativas sobre privacidad, propiedad intelectual, web scraping y gobernanza de la IA varían significativamente entre diferentes regiones del mundo, creando un panorama complejo que las empresas deben navegar con cautela. A continuación, se examinan las regulaciones en tres de las jurisdicciones más influyentes: la Unión Europea (UE), Estados Unidos (EE.UU.) y China, identificando problemas estructurales y proponiendo adaptaciones legislativas y estandarizaciones tecnológicas que faciliten el cumplimiento y fomenten la innovación.

4.1 Unión Europea: Desafíos del GDPR y propuestas

El **Reglamento General de Protección de Datos (GDPR)** de la UE ha sido pionero en establecer estándares rigurosos para la protección de datos personales. No obstante, esta robusta legislación presenta desafíos significativos para empresas como NeuroBlock que operan en el ámbito de la IA y Big Data.

Uno de los problemas estructurales más destacados del GDPR es su dependencia excesiva

del consentimiento explícito en cada fase del procesamiento de datos personales, conforme al **Artículo 6**. Este requisito, aunque esencial para la protección de la privacidad, resulta inviable en contextos donde los datos son anonimizados o utilizados en modelos agregados. Para NeuroBlock, esto implica una carga administrativa considerable al tener que obtener consentimientos explícitos repetidos, lo que puede obstaculizar la innovación y la eficiencia operativa.

Además, el **derecho al olvido** estipulado en el **Artículo 17** del GDPR presenta un desafío técnico. En el contexto de la IA, eliminar datos personales utilizados para entrenar modelos de aprendizaje automático sin comprometer la integridad del modelo es complejo. Este derecho, aunque fundamental para la protección de los derechos individuales, no ha sido adecuadamente adaptado a las realidades técnicas de los sistemas de IA, creando un vacío legal que puede llevar a interpretaciones ambiguas y sanciones onerosas para empresas como NeuroBlock.

Otra complicación surge de las obligaciones estrictas sobre la protección de datos desde el diseño y por defecto (Artículo 25), que exigen que las empresas integren medidas de protección desde las etapas iniciales de desarrollo. Sin embargo, esta disposición no aborda de manera específica la eliminación de datos una vez que han sido utilizados para entrenar modelos, dejando un área gris en la regulación que dificulta el cumplimiento efectivo.

Para abordar estos desafíos, es fundamental interpretar el GDPR de manera que equilibre la protección de datos con las necesidades operativas de la IA. Se propone la creación de una nueva categoría denominada "Tratamiento de Datos Anonimizados" dentro del GDPR. Esta adaptación legislativa eximiría la necesidad de consentimiento explícito siempre y cuando se cumplan criterios específicos de irreversibilidad y no posibilidad de reidentificación. De esta manera, NeuroBlock podría operar de manera más eficiente sin comprometer la protección de los datos personales.

Desde una perspectiva técnica, NeuroBlock podría desarrollar un **sistema de protección de datos multicapa** que garantice que la información personal sea imposible de identificar o reconstruir. Este sistema implementaría tres niveles avanzados de anonimización:

El **k-anonimato** funciona como primer filtro, asegurando que cada individuo sea

indistinguible de al menos otros k usuarios en la base de datos. Por ejemplo, cada combinación de características (edad, ubicación, género) aparecería en al menos 5 personas diferentes, imposibilitando la identificación individual.

La **l-diversidad** actúa como segundo filtro, garantizando suficiente variedad en los datos sensibles dentro de cada grupo similar. Esto evita que, aunque tengamos un grupo de personas con características comunes, se puedan inferir patrones individuales de comportamiento.

El **t-closeness** refina la protección asegurando que la distribución de valores sensibles en cada subgrupo refleje la distribución general de la base de datos completa, previniendo análisis estadísticos que pudieran comprometer la privacidad.

Todo el sistema estaría protegido mediante **cifrado de extremo a extremo**, asegurando los datos desde su recolección hasta su procesamiento final. **Auditorías independientes** verificarían regularmente la efectividad de estas medidas, certificando el cumplimiento del GDPR para datos anonimizados.

4.2 Estados Unidos: Fragmentación y necesidad de un marco federal

En EE.UU., el panorama regulatorio de privacidad y protección de datos es altamente fragmentado, con diferentes estados implementando sus propias leyes. La **California Consumer Privacy Act (CCPA)** ha establecido estándares más estrictos, pero la ausencia de un marco federal unificado genera incertidumbre jurídica y aumenta los costos de cumplimiento para empresas que operan a nivel nacional como NeuroBlock. La coexistencia de múltiples leyes estatales de privacidad complica el cumplimiento, ya que NeuroBlock debe adaptar sus prácticas a cada conjunto de normativas, incrementando significativamente los costos y recursos necesarios.

Esta fragmentación impide la creación de un estándar coherente para la protección de datos personales, lo que dificulta la innovación tecnológica y la expansión de empresas como NeuroBlock. Las empresas deben constantemente monitorear y adaptar sus prácticas a cambios regulatorios en diferentes estados, creando un ambiente de incertidumbre jurídica.

Para abordar esta problemática, es imperativo que los legisladores de EE.UU. promuevan una **Ley Federal de Protección de Datos Personales (FDPD)** que unifique las normativas estatales, ofreciendo un marco claro y coherente para todas las empresas. Esta ley debería establecer un estándar mínimo de protección para los consumidores, similar al GDPR, facilitando así el cumplimiento normativo y reduciendo la incertidumbre legal.

En términos técnicos, NeuroBlock debe preparar una **estrategia de cumplimiento modular** que le permita adaptar sus sistemas de gestión de datos a diferentes marcos estatales mientras se alinea con un posible marco federal unificado. Esto implica la adopción de estándares de seguridad y privacidad que sean compatibles con múltiples legislaciones, facilitando así la rápida adaptación una vez que se establezca una ley federal. Por ejemplo, diseñar sistemas que puedan fácilmente actualizarse para cumplir con diferentes normativas estatales mediante módulos configurables, implementar estándares reconocidos internacionalmente como ISO 27001 y utilizar herramientas de gestión de cumplimiento que automaticen la adaptación a diferentes normativas, reduciendo el esfuerzo manual y minimizando errores.

Además, NeuroBlock podría integrar **plataformas de cumplimiento automatizadas** que utilicen inteligencia artificial para monitorear y ajustar las políticas internas conforme a las regulaciones estatales y federales, asegurando un cumplimiento continuo y eficiente.

4.3 China: Equilibrio entre innovación y control Estatal

China presenta un marco regulatorio altamente centralizado y estricto, con un enfoque claro en la seguridad de los datos y la soberanía nacional. Las principales leyes que afectan a NeuroBlock son la **Ley de Protección de la Información Personal (PIPL)** y la **Ley de Seguridad de Datos (DSL)**. Estas leyes imponen restricciones significativas sobre el procesamiento y la transferencia de datos personales, así como sobre el uso de tecnologías de IA. La PIPL exige un consentimiento explícito e informado para el procesamiento de datos personales y restringe la transferencia internacional de datos sin una revisión de seguridad gubernamental. Por su parte, la DSL establece un sistema de clasificación de datos y requisitos de localización estrictos, obligando a las empresas a almacenar datos sensibles dentro de China.

Esta regulación centralizada y estricta no solo limita la flexibilidad operativa de las empresas, sino que también impone costos significativos para cumplir con los requisitos de localización y seguridad de datos. Para facilitar la operación de NeuroBlock en China sin comprometer la seguridad de los datos y la soberanía nacional, se propone la introducción de un marco regulatorio que permita excepciones claras para el tratamiento de datos anonimizados y el uso de IA en contextos específicos. Este marco debería armonizar la protección de datos con la necesidad de innovación tecnológica, proporcionando directrices claras sobre cuándo y cómo se pueden transferir datos fuera de China sin comprometer la seguridad nacional.

Una forma de abordar esto sería permitiendo el tratamiento de datos personales anonimizados sin la necesidad de consentimiento explícito, siempre que se cumplan criterios estrictos de irreversibilidad y no posibilidad de reidentificación. Además, se podría autorizar el uso de IA para aplicaciones específicas como la trazabilidad de productos y la gestión de contratos inteligentes, sin requerir la transferencia de datos personales fuera de China. Para asegurar el cumplimiento, las empresas deberán implementar auditorías periódicas realizadas por terceros certificados que verifiquen las técnicas de anonimización y las aplicaciones autorizadas de IA.

Desde una perspectiva técnica, NeuroBlock podría adoptar una **arquitectura de datos segregada** que le permita operar con dos sistemas paralelos: uno completamente localizado en China para datos de usuarios chinos y otro global para el resto del mundo. Esto no solo cumpliría con la ley china, sino que también podría servir como modelo para otras empresas que buscan navegar aguas regulatorias similares en otros países con políticas de localización de datos. Esta arquitectura incluiría centros de datos dentro de China, algoritmos que clasifiquen automáticamente los datos según su sensibilidad y sistemas que realicen análisis y generen insights a partir de datos locales sin necesidad de transferir datos sensibles fuera de China.

Además, para cumplir con las regulaciones específicas sobre algoritmos de recomendación implementadas en marzo de 2022, NeuroBlock debe desarrollar un **marco de IA ética adaptativa**. Este marco permitiría ajustar dinámicamente los parámetros de los algoritmos según las directrices estatales, manteniendo al mismo tiempo los principios éticos

fundamentales de NeuroBlock. Técnicamente, esto podría lograrse mediante una arquitectura modular que permita la separación de componentes ajustables sin afectar el núcleo del modelo, la implementación de herramientas de supervisión que monitoreen el comportamiento de los algoritmos en tiempo real y la utilización de sistemas que registren automáticamente todas las modificaciones realizadas en los algoritmos, generando reportes detallados para auditorías regulatorias.

5. Web scraping y propiedad intelectual en la IA generativa

El **web scraping** es una técnica que permite la extracción automatizada de datos disponibles en la web, frecuentemente utilizados por desarrolladores y empresas para alimentar aplicaciones y modelos de inteligencia artificial generativa. Aunque esta práctica puede parecer una simple herramienta tecnológica, plantea complejas cuestiones jurídicas, especialmente en lo que respecta a la **propiedad intelectual**. Muchas bases de datos y contenidos online están protegidos por derechos de autor o por la Directiva sobre Bases de Datos de la Unión Europea, lo que genera tensiones entre el derecho de los titulares de proteger sus creaciones y la necesidad de grandes volúmenes de datos para entrenar algoritmos de IA generativa. Estas tecnologías disruptivas, que transforman datos en conocimiento con aplicaciones que van desde la creación de textos hasta la generación de imágenes, no solo amplían el potencial de la innovación, sino que también desafían los marcos legales tradicionales, obligando a reconsiderar el equilibrio entre los derechos individuales y el acceso a los datos como bien común.

5.1 Regulaciones aplicables al web scraping en la Unión Europea: Un análisis de la propiedad intelectual y su relación con la inteligencia artificial generativa

El marco normativo de la Unión Europea en materia de web scraping se encuentra en un punto de inflexión. Las normativas existentes, diseñadas para proteger los derechos de los titulares de bases de datos y obras creativas, están siendo desafiadas por la expansión de tecnologías disruptivas como la inteligencia artificial generativa. Estas herramientas, cuya capacidad para transformar datos en conocimiento tiene un potencial transformador inmenso, dependen de grandes volúmenes de información para entrenar modelos

complejos. Sin embargo, su funcionamiento entra en conflicto con la regulación tradicional de la propiedad intelectual, planteando cuestiones críticas sobre cómo equilibrar el derecho a la protección individual con el bien común de la innovación.

5.1.1. La Directiva 96/9/CE y la protección de las bases de datos

La Directiva 96/9/CE es uno de los pilares fundamentales en la regulación de bases de datos en la Unión Europea. Establece un marco dual de protección: por un lado, los derechos de autor protegen aquellas bases cuya estructura refleja un esfuerzo creativo original, y por otro, el derecho sui generis protege la inversión sustancial en la obtención, verificación o presentación de datos, independientemente de si estos son creativos o no.

5.1.2. Derechos de autor y bases de datos originales

Según el artículo 3 de la Directiva, una base de datos que evidencie originalidad en su selección o disposición de datos está protegida por derechos de autor. Esto implica que no solo los datos individuales, sino también la estructura que los organiza, gozan de protección. Para empresas como NeuroBlock, esta protección puede convertirse en un desafío si los datos recopilados mediante scraping están organizados de manera creativa.

El caso Fixtures Marketing Ltd v. Organismos (C-444/02) subrayó que, para que una base sea protegida por derechos de autor, debe demostrar un esfuerzo creativo en su disposición. Esto significa que el simple hecho de organizar datos en una tabla no siempre garantiza protección. NeuroBlock podría aprovechar esta distinción para operar dentro de los márgenes legales, limitando el scraping a datos cuya organización no evidencie originalidad.

5.1.3. Derecho sui generis: La inversión como criterio protector

El derecho sui generis, regulado en los artículos 7 a 11 de la Directiva, otorga protección a las bases de datos que representan una inversión sustancial en términos financieros, temporales o de esfuerzo. Este derecho prohíbe la extracción de partes sustanciales de dichas bases sin autorización, incluso si los datos individuales no están protegidos.

El caso British Horseracing Board Ltd v. William Hill Organization Ltd (C-203/02) destacó que

la extracción repetida de partes aparentemente no sustanciales puede, en conjunto, constituir una infracción si afecta al valor económico de la base de datos. Esto es particularmente relevante para NeuroBlock, que utiliza técnicas automatizadas para recopilar datos de múltiples fuentes. En este sentido, se deberían implementar tecnologías que fragmenten y minimicen la extracción de datos, reduciendo el riesgo de superar el umbral de lo permitido.

5.1.4. La Directiva 2019/790: Derechos de autor y minería de texto y datos (TDM)

En respuesta a los desafíos planteados por el avance tecnológico, la Directiva 2019/790 introduce un marco regulatorio innovador para la minería de texto y datos, reconociendo su papel fundamental en el desarrollo de la inteligencia artificial. La directiva establece dos niveles de excepciones que modernizan el enfoque tradicional de los derechos de autor.

El Artículo 3 contempla una excepción específica para organismos de investigación e instituciones del patrimonio cultural, permitiéndoles realizar minería de datos con fines de investigación científica, siempre que cuenten con acceso lícito a las obras. Esta disposición viene acompañada de la obligación de implementar medidas de seguridad adecuadas para el almacenamiento y verificación de resultados, garantizando así la integridad del proceso investigador.

Por su parte, el Artículo 4 amplía el alcance de las excepciones al establecer una provisión más general, aplicable a cualquier finalidad. Esta excepción resulta particularmente relevante para empresas como NeuroBlock, ya que permite la minería de datos sobre contenidos accesibles legalmente, con la única limitación de que los titulares de derechos no hayan implementado restricciones técnicas mediante medios de lectura mecánica. Este enfoque representa un cambio significativo respecto a las interpretaciones tradicionales, ya que los términos de servicio o condiciones de uso generales no son suficientes para prevenir la minería de datos.

Sin embargo, la directiva presenta limitaciones significativas en el contexto de la IA generativa. No aborda específicamente las particularidades del entrenamiento de modelos capaces de generar nuevo contenido, dejando en una zona gris legal la distinción entre la

simple minería de datos y el aprendizaje que permite replicar patrones creativos. Además, no proporciona orientación sobre los derechos de autor de las obras generadas por IA basadas en el aprendizaje de obras protegidas.

La distinción entre uso comercial y científico también plantea desafíos prácticos, especialmente cuando la colaboración entre empresas e instituciones académicas difumina las líneas entre investigación y aplicación comercial. Esta complejidad se intensifica en un contexto global donde el procesamiento y uso de datos trasciende fronteras.

Particularmente crítica resulta la ausencia de regulación específica sobre atribución y transparencia en el contexto de la IA generativa. La directiva no establece guías claras sobre cómo garantizar el reconocimiento de los creadores originales cuando un modelo genera contenido basado en múltiples fuentes, ni determina requisitos de transparencia sobre las fuentes de entrenamiento.

Estas limitaciones sugieren la necesidad de una evolución normativa que aborde específicamente los desafíos de la IA generativa, estableciendo un marco legal que equilibre la innovación tecnológica con la protección adecuada de los derechos de autor en esta nueva era digital.

5.1.5. La IA generativa como catalizador de una nueva interpretación del derecho de autor

La irrupción de la inteligencia artificial generativa ha provocado una profunda transformación en la relación entre la innovación tecnológica y la propiedad intelectual, exponiendo las limitaciones del marco normativo actual y planteando interrogantes fundamentales sobre la naturaleza misma de la creación y la autoría. Si bien la Directiva 2019/790 representa un avance significativo en la modernización del derecho de autor europeo, especialmente en lo referente a la minería de datos, su enfoque aún refleja una concepción pre-IA de la creación y el uso de obras protegidas.

El proceso de entrenamiento de modelos de IA generativa desafía las categorías tradicionales del derecho de autor de maneras que trascienden la simple distinción entre reproducción y comunicación pública. Cuando un modelo analiza grandes corpus de texto o imágenes

protegidas, no está simplemente extrayendo información o realizando copias, sino que está desarrollando una comprensión abstracta de patrones, estructuras y relaciones que posteriormente utilizará para generar contenido original. Este proceso plantea una pregunta fundamental: ¿en qué momento el aprendizaje de patrones se convierte en una forma de apropiación que requiere autorización?

La jurisprudencia europea ofrece algunas pistas para abordar esta cuestión. El caso British Horseracing Board Ltd v. William Hill Organization Ltd estableció principios importantes sobre la sustancialidad en el uso de bases de datos, mientras que SAS Institute Inc. v. World Programming Ltd reconoció que la funcionalidad de un programa no está protegida por derechos de autor. Estos precedentes sugieren que el derecho europeo es capaz de distinguir entre la protección de la expresión concreta y la libertad de aprender y replicar conceptos abstractos.

Sin embargo, la IA generativa introduce una complejidad adicional: la capacidad de producir obras que, aunque técnicamente originales, derivan de patrones aprendidos de miles de obras protegidas. La Directiva 2019/790, con sus excepciones para minería de datos científica y comercial, no aborda adecuadamente esta realidad. La distinción entre uso comercial y científico se vuelve especialmente problemática en un contexto donde la colaboración entre academia y empresa es cada vez más frecuente y necesaria para el avance tecnológico.

La experiencia de empresas como NeuroBlock demuestra que el desarrollo responsable de IA generativa requiere un nuevo marco conceptual que reconozca tanto el valor transformativo de estas tecnologías como los derechos legítimos de los creadores originales. Este marco debería fundamentarse en principios de transparencia, atribución justa y compensación equitativa, sin imponer restricciones que sofoquen la innovación.

La solución podría encontrarse en un sistema híbrido que combine licencias colectivas obligatorias para el entrenamiento de IA con mecanismos de compensación basados en el impacto comercial de las aplicaciones resultantes. Esto permitiría un acceso amplio a datos de entrenamiento mientras asegura una remuneración justa para los creadores. La implementación de estándares técnicos para la trazabilidad y atribución de fuentes de entrenamiento podría complementar este enfoque, proporcionando transparencia sin

comprometer la eficiencia computacional.

La evolución del marco normativo europeo debe reconocer que la IA generativa no es simplemente una herramienta más de análisis de datos, sino un nuevo paradigma de creación que desafía nuestras concepciones tradicionales de autoría y originalidad. El objetivo no debe ser preservar un modelo de propiedad intelectual diseñado para una era anterior, sino desarrollar nuevos mecanismos que fomenten tanto la innovación tecnológica como la creatividad humana.

En última instancia, el éxito de esta transformación dependerá de nuestra capacidad para construir un ecosistema donde la protección de derechos y la innovación tecnológica se refuercen mutuamente. La experiencia europea en la armonización de derechos fundamentales con objetivos de mercado único digital ofrece una base sólida para este desarrollo, pero requerirá una voluntad política sostenida y una colaboración activa entre legisladores, creadores y desarrolladores de tecnología.

5.2 Regulaciones aplicables al web scraping en Estados Unidos: Propiedad intelectual y límites legales

El sistema legal estadounidense aborda el web scraping desde una perspectiva distintivamente pragmática, característica de su tradición jurídica basada en el common law. A diferencia del enfoque más prescriptivo y estructurado de la Unión Europea, Estados Unidos ha permitido que la interpretación judicial de leyes preexistentes y la evolución de la jurisprudencia definan los contornos legales de esta práctica tecnológica. Esta aproximación ha generado un panorama legal que, si bien es más flexible y adaptativo, también presenta desafíos significativos para las empresas que buscan navegar sus complejidades.

5.2.1 La evolución del Copyright Act en la era digital

El Copyright Act de 1976, concebido en una época pre-digital, ha demostrado una notable capacidad de adaptación frente a los desafíos tecnológicos contemporáneos. Su aplicación al web scraping ilustra cómo los principios fundamentales del derecho de autor estadounidense pueden evolucionar para abordar nuevas formas de uso y reproducción de contenido. La ley

protege no solo las expresiones creativas tradicionales, sino también las compilaciones de datos que exhiben un mínimo de creatividad en su selección y organización, un aspecto particularmente relevante en la era del big data.

La verdadera innovación en la aplicación del Copyright Act al contexto digital proviene de la doctrina del fair use. Este concepto jurídico, único del sistema estadounidense, ha demostrado ser sorprendentemente adaptable a los desafíos tecnológicos modernos. Codificada en la Sección 107 del Copyright Act, la doctrina del fair use establece cuatro factores fundamentales para evaluar si un uso particular constituye fair use:

1. **El propósito y carácter del uso**: Se evalúa si el uso es de naturaleza comercial o educativa, y si añade algún valor nuevo al contenido original.
2. **La naturaleza de la obra protegida**: Se considera si la obra es más factual o creativa. **3. La cantidad y sustancialidad de la porción utilizada**: Se analiza qué tan grande y significativa es la parte utilizada en relación con la obra completa.
4. **El efecto sobre el mercado potencial de la obra**: Se examina si el uso afecta negativamente el mercado existente o potencial de la obra original.

Estos factores permiten una evaluación flexible y contextualizada del uso de obras protegidas, adaptándose a las necesidades y realidades cambiantes de la tecnología y la sociedad.

5.2.2 Authors Guild v. Google Inc: Un punto de inflexión en el contexto digital

El caso Authors Guild v. Google Inc. representa un hito crucial en la interpretación de la doctrina del fair use en el contexto digital. En esta sentencia, la corte consideró que la digitalización masiva de libros por parte de Google constituía un uso justo. Este fallo estableció varios principios fundamentales:

- Uso transformativo: La corte determinó que el uso transformativo del contenido puede justificar incluso la copia completa de obras protegidas, siempre y cuando se añada un valor nuevo o se realice una transformación significativa.
- Creación de índices y bases de datos buscables: La decisión subrayó que la creación de índices y bases de datos que facilitan la búsqueda y acceso a la información puede considerarse transformativa, ya que proporciona una utilidad adicional a las obras

originales.

- Impacto en el mercado: La corte enfatizó que el impacto en el mercado debe evaluarse considerando los nuevos usos tecnológicos. En este caso, la digitalización no sustituyó las ventas de libros físicos, sino que añadió una nueva dimensión de accesibilidad y búsqueda.

Estos principios tienen implicaciones profundas para empresas como NeuroBlock, sugiriendo que el uso de contenido protegido para entrenar algoritmos de inteligencia artificial o para crear análisis agregados podría considerarse fair use bajo ciertas condiciones. Este enfoque permite una mayor flexibilidad y fomenta la innovación tecnológica, al tiempo que protege los derechos de los creadores originales.

El caso HiQ Labs v. LinkedIn constituye uno de los desarrollos más significativos en la regulación del web scraping en Estados Unidos. La decisión del Noveno Circuito de permitir el scraping de datos públicamente accesibles, incluso contra los deseos explícitos del propietario de la plataforma, establece una posición distintivamente estadounidense sobre el acceso a la información pública en la era digital.

La decisión en HiQ Labs v. LinkedIn merece un análisis profundo debido a sus implicaciones fundamentales. Este caso no solo abordó la legalidad del web scraping bajo la Computer Fraud and Abuse Act (CFAA), sino que también articuló una visión específicamente estadounidense sobre el acceso a la información en la era digital. La corte reconoció implícitamente que, en una economía cada vez más dependiente de los datos, el acceso a información públicamente disponible no puede quedar completamente bajo el control discrecional de las grandes plataformas tecnológicas.

El razonamiento del tribunal refleja una tensión fundamental entre dos principios del derecho estadounidense: la libertad contractual y la política pública contra las restricciones al comercio. Al permitir el scraping de datos públicos incluso cuando los términos de servicio lo prohíben, la corte sugiere que existe un interés público superior en mantener los datos públicamente accesibles que supera las restricciones contractuales privadas.

5.2.3.2 La Computer Fraud and Abuse Act en la era del big data

La interpretación de la CFAA en el contexto del web scraping representa una evolución significativa en el pensamiento legal sobre el acceso a sistemas informáticos. Originalmente concebida como una ley anti-hacking en 1986, la CFAA ha tenido que adaptarse a un entorno tecnológico radicalmente diferente. La decisión en HiQ Labs refleja un reconocimiento judicial de que el acceso automatizado a información pública no puede equipararse con la intrusión maliciosa en sistemas protegidos.

Esta evolución interpretativa tiene implicaciones profundas para empresas como NeuroBlock. En primer lugar, establece una distinción clara entre el acceso técnico no autorizado, que viola la CFAA, y el acceso que simplemente viola términos de servicio, el cual puede dar lugar a reclamaciones contractuales pero no penales. En segundo lugar, sugiere que las medidas técnicas de protección, más que las restricciones contractuales, son el estándar apropiado para determinar la autorización de acceso. Finalmente, reconoce implícitamente el valor social y económico del análisis automatizado de datos públicamente accesibles, promoviendo así la innovación y el desarrollo tecnológico.

5.2.4 La intersección con el derecho antimonopolio

Un aspecto frecuentemente pasado por alto en el análisis del web scraping es su relación con el derecho antimonopolio estadounidense. La capacidad de las grandes plataformas para restringir el acceso a datos públicamente disponibles plantea preocupaciones anticompetitivas significativas. Casos recientes sugieren una creciente disposición de los tribunales a considerar estas implicaciones.

En primer lugar, la restricción del acceso a datos puede constituir una barrera artificial a la entrada en mercados digitales. Al limitar el acceso a grandes conjuntos de datos, las plataformas pueden dificultar la entrada de nuevos competidores, manteniendo o incluso reforzando su posición dominante en el mercado. En segundo lugar, el control exclusivo sobre grandes conjuntos de datos públicos puede contribuir a la creación o mantenimiento de poder de mercado, impidiendo que otras empresas accedan a los recursos necesarios para competir de manera efectiva. Finalmente, las restricciones al scraping pueden limitar la

competencia en mercados adyacentes o derivados, restringiendo la innovación y la diversidad de ofertas en el ecosistema digital.

5.2.5 Implicaciones prácticas y estratégicas

Para las empresas tecnológicas que operan en este entorno legal, es crucial desarrollar estrategias que maximicen la seguridad jurídica mientras mantienen la viabilidad operativa. Esto requiere un enfoque multinivel que abarque diversas áreas de acción.

En primer lugar, es fundamental implementar una evaluación estructurada del fair use. Esto implica desarrollar marcos internos para evaluar sistemáticamente el carácter transformativo del uso de datos, documentar detalladamente cómo el procesamiento de datos añade valor nuevo y realizar análisis regulares del impacto en mercados existentes y potenciales. Esta evaluación rigurosa permite a las empresas anticipar y mitigar posibles riesgos legales asociados al uso de contenido protegido.

En segundo lugar, las empresas deben establecer sistemas de cumplimiento técnico sofisticados. Esto incluye la implementación de protocolos que respeten archivos como *robots.txt* y otras señales técnicas que indican las preferencias de acceso de las plataformas. Además, es esencial desarrollar sistemas de monitoreo de carga y respuesta adaptativa para gestionar el acceso automatizado a los datos de manera eficiente y respetuosa. El mantenimiento de registros detallados de accesos y procesamientos también es crucial para demostrar cumplimiento y transparencia en las operaciones de scraping.

Finalmente, es necesario adoptar estrategias de mitigación de riesgos proactivas. Diversificar las fuentes de datos puede reducir la dependencia de plataformas individuales, disminuyendo así el riesgo de enfrentar restricciones legales específicas. El desarrollo de relaciones directas con proveedores de datos clave también puede facilitar el acceso autorizado a información valiosa. Además, la implementación de técnicas avanzadas de anonimización y agregación de datos puede minimizar el impacto en los sistemas objetivo y proteger la privacidad de los datos utilizados.

5.2.6 Hacia un nuevo paradigma regulatorio

La evolución del marco legal estadounidense para el web scraping refleja una tensión más

amplia en la regulación de tecnologías emergentes. Aunque el enfoque basado en precedentes judiciales ofrece una mayor flexibilidad comparado con el sistema europeo, ha comenzado a mostrar sus limitaciones frente a la creciente complejidad del ecosistema digital. Esta realidad está impulsando una conversación sobre la necesidad de un marco regulatorio más coherente que pueda abordar varios desafíos emergentes.

Uno de los principales desafíos es la convergencia de la inteligencia artificial (IA) y el web scraping. El uso creciente de la IA en conjunto con técnicas de web scraping está creando nuevos desafíos legales que los tribunales apenas están comenzando a abordar. La capacidad de los sistemas de IA para procesar y transformar datos masivos de manera automatizada desafía las concepciones tradicionales de uso transformativo bajo la doctrina del fair use. Por ejemplo, cuando un modelo de IA aprende de millones de textos scrapeados, surge la pregunta de en qué punto este aprendizaje constituye una transformación suficiente para calificar como fair use. Los tribunales estadounidenses están en las etapas iniciales de tratar estas cuestiones, lo que subraya la necesidad de una adaptación continua del marco legal.

Otro aspecto crítico es el papel de los datos en la competencia. La jurisprudencia reciente sugiere una creciente conciencia de que el acceso a datos masivos no es simplemente una cuestión de propiedad intelectual o seguridad informática, sino un problema fundamental de competencia económica. Esta perspectiva está influyendo en la interpretación judicial de las restricciones al scraping. Los tribunales están cada vez más dispuestos a considerar el impacto anticompetitivo de restricciones excesivas al acceso a datos, reconociendo que los datos públicamente accesibles constituyen una infraestructura esencial para la innovación digital. Las decisiones judiciales están comenzando a equilibrar los derechos de propiedad con la necesidad de mantener mercados competitivos, lo que puede llevar a una reevaluación de las políticas de acceso a datos por parte de las grandes plataformas tecnológicas.

En cuanto a las tendencias emergentes y desarrollos futuros, varios movimientos indican la dirección futura de la regulación del web scraping en Estados Unidos. Existe un creciente énfasis en la implementación de estándares técnicos que prioricen medidas de protección sobre restricciones contractuales. Esto incluye el desarrollo de estándares industriales para el scraping ético y la adopción de protocolos automatizados para comunicar restricciones de

acceso de manera más clara y eficiente.

Además, la doctrina del fair use está en proceso de evolución para acomodar el análisis automatizado. Se está expandiendo el concepto de uso transformativo para incluir aplicaciones de inteligencia artificial y se está considerando de manera más amplia el valor social y económico del análisis de datos. Asimismo, se están desarrollando criterios específicos para evaluar el fair use en contextos de IA, lo que facilitará una aplicación más precisa y justa de la doctrina en estos nuevos escenarios tecnológicos.

Finalmente, hay una creciente convergencia con principios de privacidad. Las consideraciones de privacidad están siendo integradas de manera más sistemática en el análisis legal del scraping, con el desarrollo de estándares para la anonimización y agregación de datos scrapeados. Además, se está prestando mayor atención a los derechos individuales sobre datos personales públicamente accesibles, lo que refleja una tendencia hacia una mayor protección de la privacidad en el contexto del web scraping y la inteligencia artificial.

5.2.7 Implicaciones estratégicas para la industria

Para empresas como NeuroBlock, el panorama evolutivo en la regulación del web scraping sugiere la necesidad de adoptar un enfoque proactivo y multinivel. Esto implica varias estrategias clave:

Inversión en cumplimiento adaptativo: Las empresas deben desarrollar sistemas que puedan responder dinámicamente a cambios en el entorno legal. Esto incluye la implementación de protocolos de transparencia y documentación robustos, así como el establecimiento de mecanismos de revisión y actualización regular de prácticas para asegurar el cumplimiento continuo con las normativas vigentes.

Participación en el desarrollo de estándares: Colaborar con asociaciones industriales en el desarrollo de mejores prácticas es esencial para influir en la creación de estándares técnicos y éticos. Además, contribuir al diálogo sobre estos estándares y participar en iniciativas de autorregulación sectorial puede posicionar a las empresas como líderes responsables en el ámbito del web scraping y la inteligencia artificial.

Innovación en prácticas de scraping: Las empresas deben enfocarse en desarrollar tecnologías que minimicen el impacto en los sistemas objetivo, implementando técnicas

avanzadas de procesamiento y agregación que respeten las limitaciones técnicas y legales. Además, explorar modelos colaborativos con proveedores de datos puede facilitar el acceso autorizado a información valiosa, reduciendo así el riesgo de enfrentar restricciones legales.

5.2.8 Conclusión y perspectivas futuras

El marco legal estadounidense para el web scraping continúa en constante evolución, impulsado por la intersección de múltiples fuerzas: innovación tecnológica, necesidades comerciales, preocupaciones de privacidad y consideraciones competitivas. Para las empresas que operan en este espacio, el éxito dependerá no solo de la comprensión técnica y legal, sino también de la capacidad para anticipar y adaptarse a los cambios en el panorama regulatorio.

La experiencia estadounidense ofrece lecciones valiosas para el desarrollo global de marcos regulatorios de scraping, sugiriendo que el enfoque más efectivo puede ser una combinación de la flexibilidad del common law con principios claros que promuevan tanto la innovación como la protección de derechos legítimos. Esta evolución será fundamental para el futuro del ecosistema digital global, ya que permitirá un equilibrio adecuado entre el fomento de la innovación tecnológica y la salvaguardia de los derechos de los creadores originales.

En última instancia, la transformación del marco regulatorio estadounidense reflejará nuestra capacidad para construir un ecosistema donde la protección de derechos y la innovación tecnológica se refuercen mutuamente. Este equilibrio será crucial para asegurar un entorno digital dinámico y competitivo que beneficie tanto a las empresas como a los individuos en la era de la inteligencia artificial generativa.

5.3 La Ley de Seguridad Cibernética de China y su interacción con la propiedad intelectual

La Ley de Seguridad Cibernética de China (CSL), promulgada en 2016, constituye un hito en la regulación global de datos al establecer un modelo que prioriza la soberanía digital y la seguridad nacional por encima del libre flujo de información. Este enfoque, claramente diferenciado de los marcos occidentales previamente analizados, presenta implicaciones profundas para las actividades de web scraping y el desarrollo de la inteligencia artificial. En

este contexto, la CSL no solo redefine los límites del acceso y procesamiento de datos, sino que también refleja una visión más amplia de gobernanza digital que subordina intereses económicos a objetivos estratégicos y de seguridad.

5.3.1 La arquitectura regulatoria china

La CSL no opera de manera aislada. Forma parte de un ecosistema legislativo robusto que incluye la Ley de Protección de Datos Personales (PIPL) y la Ley de Seguridad de Datos (DSL). Este tríptico legislativo configura un marco integral que refuerza la visión china de soberanía digital, donde el control sobre los datos no se limita a proteger la privacidad individual, sino que se extiende a garantizar intereses esenciales de seguridad nacional. En este esquema, el web scraping no es solo una cuestión de propiedad intelectual o privacidad; se convierte en una actividad regulada bajo estrictos estándares que buscan proteger infraestructuras críticas y preservar la estabilidad del sistema.

Bajo este modelo, cualquier actividad que involucre la recopilación o transferencia de datos debe pasar por un riguroso escrutinio regulatorio. El artículo 27 de la CSL, por ejemplo, prohíbe cualquier acción que comprometa la seguridad de los sistemas digitales. Este enfoque preventivo, centrado en el potencial de disrupción, contrasta marcadamente con legislaciones occidentales como la Computer Fraud and Abuse Act (CFAA) de Estados Unidos, que se activa únicamente ante accesos no autorizados. Este umbral más bajo para la intervención regulatoria china impone retos significativos a empresas que deseen operar en este entorno.

5.3.2 Protección de sistemas y control de datos en relación con el web scraping

El marco normativo chino impone requisitos detallados y exigentes que afectan directamente las actividades de scraping y el procesamiento de datos masivos. Entre los elementos más destacados se encuentran las evaluaciones de impacto obligatorias, los requisitos de localización de datos y la influencia del sistema de crédito social corporativo.

5.3.2.1 Evaluación de impacto obligatoria

Las empresas deben realizar evaluaciones exhaustivas para determinar el posible impacto de sus actividades de scraping en los sistemas objetivo. Esta obligación no se limita a evaluar

riesgos técnicos; también incluye el análisis de posibles implicaciones económicas y sociales. Lo más significativo es que el umbral para considerar una actividad como "disruptiva" es notablemente más bajo que en otras jurisdicciones. La carga de la prueba recae sobre la empresa, que debe demostrar que sus actividades no comprometen la estabilidad o seguridad de los sistemas.

5.3.2.2 Requisitos de localización de datos

La exigencia de almacenamiento y procesamiento local de datos es una de las características distintivas del modelo chino de soberanía digital. Esto implica que los datos recopilados dentro de China deben permanecer dentro de sus fronteras, a menos que se obtenga una autorización explícita para su transferencia. Este requisito impone importantes desafíos arquitectónicos y operativos, obligando a las empresas a:

1. Establecer infraestructuras físicas en China que cumplan con estándares regulatorios locales.
2. Someterse a auditorías de seguridad y procesos de aprobación antes de transferir datos transfronterizos.
3. Implementar controles rigurosos para garantizar que los datos no sean accesibles desde fuera del territorio chino.

5.3.2.3 Sistema de crédito social corporativo

Además de las exigencias técnicas y regulatorias, las actividades de scraping pueden influir directamente en la calificación de crédito social de una empresa bajo el sistema de crédito social corporativo de China. Este sistema evalúa de manera continua el comportamiento corporativo, penalizando cualquier práctica que sea percibida como contraria a los intereses regulatorios o de seguridad. Una calificación baja puede limitar la capacidad de la empresa para obtener licencias, establecer alianzas comerciales o incluso continuar operando en el país.

5.3.3 Estrategias avanzadas para operar en el entorno regulatorio chino

Para empresas tecnológicas como NeuroBlock, navegar con éxito el ecosistema regulatorio chino requiere una estrategia multifacética que combine cumplimiento normativo, innovación tecnológica y relaciones estratégicas.

5.3.3.1 Colaboraciones estratégicas con actores locales

Las asociaciones con empresas tecnológicas chinas que ya operan dentro del marco regulatorio ofrecen ventajas significativas. Estas empresas no solo proporcionan conocimiento del entorno local, sino que también actúan como facilitadores para establecer relaciones con autoridades reguladoras y para participar en iniciativas gubernamentales de "innovación responsable".

5.3.3.2 Adaptación tecnológica

Es esencial desarrollar tecnologías específicamente adaptadas al mercado chino. Esto incluye sistemas de scraping que limiten la cantidad de datos recopilados, implementen controles granulares sobre el procesamiento y utilicen tecnologías avanzadas de anonimización. Además, la adopción de protocolos de privacidad federada, como el aprendizaje federado, permite entrenar modelos de inteligencia artificial sin necesidad de transferir datos en bruto fuera de los sistemas locales.

5.3.3.3 Cumplimiento proactivo y monitoreo continuo

Las empresas deben establecer sistemas que monitoreen en tiempo real los cambios regulatorios y respondan de manera inmediata a preocupaciones planteadas por las autoridades. Esto no solo reduce el riesgo de sanciones, sino que también fortalece la relación con las entidades reguladoras. Mantener un canal de comunicación abierto y regular con las autoridades puede facilitar la obtención de autorizaciones y mejorar la reputación corporativa.

5.3.4 Implicaciones para la inteligencia artificial generativa y el procesamiento de datos masivos

El marco regulatorio chino presenta desafíos únicos para el desarrollo de la inteligencia artificial generativa, particularmente en lo que respecta al entrenamiento de modelos. Las restricciones sobre la transferencia y el procesamiento de datos limitan significativamente la diversidad de los conjuntos de datos disponibles, lo que puede introducir sesgos en los modelos entrenados exclusivamente con datos locales. Asimismo, las exigencias para obtener autorizaciones explícitas antes de utilizar datos públicamente accesibles añaden obstáculos adicionales al desarrollo de sistemas de IA que requieren grandes volúmenes de información.

La localización obligatoria de datos también dificulta la implementación de técnicas de aprendizaje continuo, ya que los modelos que necesitan actualizarse constantemente con nuevos datos enfrentan barreras regulatorias importantes. Esto limita la capacidad de las empresas para desarrollar sistemas de IA adaptativos y ágiles.

5.3.5 Innovación en cumplimiento normativo: lecciones para el futuro

A pesar de estos desafíos, el marco regulatorio chino también ha impulsado la innovación en el cumplimiento normativo. Tecnologías como la privacidad federada y los sistemas de auditoría automatizada están emergiendo como soluciones viables para operar en este entorno restrictivo. Estas herramientas permiten entrenar modelos de IA sin transferir datos en bruto, cumplir con requisitos de localización y garantizar la trazabilidad de las actividades de procesamiento.

5.3.6 Conclusión

El enfoque chino hacia el web scraping y la regulación de datos representa más que un conjunto de restricciones: es un modelo alternativo de gobernanza digital que prioriza la seguridad nacional y la soberanía sobre la eficiencia económica. Para empresas globales, adaptarse a este modelo no es solo una necesidad para operar en China, sino una preparación esencial para un futuro donde más países podrían adoptar enfoques similares.

La experiencia china ofrece lecciones valiosas sobre cómo equilibrar la innovación tecnológica con el cumplimiento regulatorio estricto. Las soluciones desarrolladas para navegar este entorno complejo tienen el potencial de convertirse en estándares globales, a medida que las políticas de soberanía digital y seguridad nacional ganan tracción en otras regiones. El éxito en este contexto dependerá de la capacidad de las empresas para mantener la innovación mientras se adaptan a las demandas de un ecosistema regulatorio cada vez más complejo y restrictivo.

6. Conclusiones

Primera. Los marcos regulatorios actuales presentan lagunas significativas al abordar la inteligencia artificial generativa, especialmente en lo que respecta al tratamiento y control de los datos. Estas tecnologías, al operar en ecosistemas altamente interconectados, generan flujos de datos que escapan a las estructuras normativas tradicionales, permitiendo puntos de fuga donde los derechos fundamentales, como la privacidad y la autodeterminación informativa, quedan desprotegidos.

Segunda. La comparación entre la Unión Europea, Estados Unidos y China pone de manifiesto contradicciones regulatorias que obstaculizan la creación de un mercado global armonizado. Mientras que el GDPR de la Unión Europea enfatiza la protección de datos personales, su rigidez limita la implementación de soluciones innovadoras como los sistemas descentralizados. En contraste, el modelo estadounidense fragmentado genera incertidumbre jurídica al carecer de cohesión federal, y el enfoque chino, aunque eficiente para la innovación local, prioriza el control estatal, afectando la interoperabilidad internacional.

Tercera. La regulación actual subestima la interrelación entre los datos, la tecnología y los derechos fundamentales. La dependencia de estructuras centralizadas de datos, como los servidores tradicionales, perpetúa desigualdades y crea vulnerabilidades críticas. Este modelo resulta insuficiente frente a tecnologías como el aprendizaje federado, los Identificadores Descentralizados (DIDs) y los Nodos Web Descentralizados (DWNs), que permiten un control

directo y en tiempo real de los datos por parte de los ciudadanos.

Cuarta. Es crucial reorientar los marcos regulatorios hacia un paradigma descentralizado que garantice el control total de los usuarios sobre sus datos. Esto implica que las normativas no solo deben establecer derechos, sino también exigir el desarrollo y la implementación de tecnologías que materialicen estos derechos, como Web5. Dicho paradigma permite reducir puntos de fuga y alinear los intereses comerciales con la protección de derechos fundamentales.

Quinta. Las lagunas en la regulación de la IA generativa afectan directamente a los principios básicos de transparencia y responsabilidad. Sin normativas claras sobre cómo se utilizan los datos en el entrenamiento de modelos, el riesgo de perpetuar sesgos, monopolios informativos y abusos en la toma de decisiones algorítmicas es alarmante. Por ello, resulta indispensable un marco jurídico que imponga auditorías algorítmicas obligatorias y trace el flujo de datos a lo largo del ciclo de vida de los modelos.

Sexta. La adopción de estándares tecnológicos como los sistemas de identidad descentralizada y la trazabilidad basada en blockchain no solo protege los derechos de los ciudadanos, sino que también redefine las bases del mercado digital hacia un modelo más equitativo. La regulación debe incentivar estos enfoques para evitar que los derechos de privacidad y control sobre los datos queden supeditados a la voluntad de grandes actores tecnológicos.

Séptima. Los sistemas legales deben abordar de manera urgente la contradicción entre la velocidad de la innovación y la lentitud del proceso normativo. La creación de *sandbox regulatorias* puede ofrecer un espacio controlado donde empresas, legisladores y expertos evalúen el impacto de nuevas tecnologías, permitiendo ajustes normativos ágiles sin comprometer la seguridad jurídica.

Octava. Proyectos como NeuroBlock destacan por su potencial para demostrar cómo la integración de tecnologías descentralizadas puede garantizar derechos fundamentales en entornos digitales. Al basarse en sistemas como los DIDs y la generación de datos sintéticos, NeuroBlock no solo cumple con las normativas existentes, sino que también propone un

modelo sostenible que equilibra innovación, privacidad y control ciudadano.

Novena. El futuro de la regulación tecnológica pasa por un cambio de paradigma: de la mera mitigación de riesgos a la promoción activa de tecnologías que garanticen derechos. Esto implica superar la visión tradicional de las normas como restricciones, adoptando un enfoque que combine incentivos al desarrollo de soluciones éticas y sostenibles con sanciones para quienes incumplan principios básicos de transparencia y justicia.

Décima. La clave para un progreso jurídico duradero radica en una gobernanza digital multinivel, en la que se reconozca la interacción constante entre las esferas local, nacional e internacional. Solo a través de un esfuerzo conjunto que trascienda los intereses geopolíticos podrá garantizarse un entorno digital en el que los derechos de los ciudadanos no se erosionen, y en el que la innovación tecnológica sea un motor para el desarrollo humano.

7. Referencias bibliográficas

Bibliografía básica

1. ANTHROPIC. *Scaling Monosemanticity: Extracting Interpretable Features from Claude 3 Sonnet* [en línea]. 2024 [consulta: 9 enero 2025]. Disponible en: https://transformer-circuits.pub/2024/scaling-monosemanticity.
2. ASIA SOCIETY. *China's emerging approach to regulating general-purpose artificial intelligence* [en línea]. 2024 [consulta: 9 enero 2025]. Disponible en: https://asiasociety.org/policy-institute/chinas-emerging-approach-regulating-general-purpose-artificial-intelligence-balancing-innovation-and.
3. ASIA SOCIETY POLICY INSTITUTE. *China's Emerging Approach to Regulating General-Purpose Artificial Intelligence: Balancing Innovation and Control* [en línea]. 2024 [consulta: 9 enero 2025]. Disponible en: https://asiasociety.org/new-report-chinas-emerging-approach-regulating-general-purpose-artificial-intelligence-balancing.
4. CHINA LAW VISION. *China proposes national standards on generative AI security* [en línea]. 2023 [consulta: 9 enero 2025]. Disponible en: https://www.chinalawvision.com.
5. DING, J. *Deciphering China's AI Dream: The Context, Components, Capabilities, and Consequences of China's Strategy to Lead the World in AI*. Future of Humanity Institute, University of Oxford [en línea]. 2018 [consulta: 9 enero 2025]. Disponible en: https://www.fhi.ox.ac.uk/wp-content/uploads/Deciphering_Chinas_AI-Dream.pdf.
6. LATHAM & WATKINS. *China's new AI regulations* [en línea]. 2023 [consulta: 9 enero 2025]. Disponible en: https://www.lw.com/admin/upload/SiteAttachments/Chinas-New-AI-Regulations.pdf.
7. MISKELLEY, A. *Innovation and International Influence: China's Artificial Intelligence Pursuit in a Geopolitical Context*. American University [en línea]. 2024 [consulta: 9 enero 2025]. Disponible en:

https://aura.american.edu/articles/thesis/Innovation_and_International_Influence_China_s_Artificial_Intelligence_Pursuit_in_a_Geopolitical_Context/25814236.

8. NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *AI Risk Management Framework* [en línea]. 2023 [consulta: 9 enero 2025]. Disponible en: https://www.nist.gov/itl/ai-risk-management-framework.
9. NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (NIST). *Artificial Intelligence Risk Management Framework: Extended Version (AI 600-1)* [en línea]. 2024 [consulta: 9 enero 2025]. Disponible en: https://www.nist.gov.
10. TBD. *Web5: A Decentralized Web Platform* [en línea]. Sin fecha [consulta: 9 enero 2025]. Disponible en: https://developer.tbd.website/projects/web5.
11. VASWANI, A., SHAZEER, N., PARMAR, N., USZKOREIT, J., JONES, L., GOMEZ, A. N., KAISER, Ł. y POLOSUKHIN, I. *Attention is all you need*. Advances in Neural Information Processing Systems. 2017, vol. 30 [en línea]. Disponible en: https://doi.org/10.48550/arXiv.1706.03762.

Legislación citada

1. ASAMBLEA POPULAR NACIONAL DE CHINA. *Data Security Law (DSL)* [en línea]. 2021 [consulta: 9 enero 2025]. Disponible en: http://www.npc.gov.cn/englishnpc/c2759/c23934/202112/t20211209_385109.html.
2. ASAMBLEA POPULAR NACIONAL DE CHINA. *Personal Information Protection Law (PIPL)* [en línea]. 2021 [consulta: 9 enero 2025]. Disponible en: https://personalinformationprotectionlaw.com/.
3. CALIFORNIA STATE LEGISLATURE. *California Consumer Privacy Act (CCPA)* [en línea]. 2018 [consulta: 9 enero 2025]. Disponible en: https://oag.ca.gov/privacy/ccpa.
4. CONSEJO DE ESTADO DE CHINA. *Plan de Desarrollo de IA de Próxima Generación (2017-2030)* [en línea]. 2017 [consulta: 9 enero 2025]. Disponible en: https://flia.org/wp-content/uploads/2017/07/A-New-Generation-of-Artificial-Intelligence-Development-Plan-1.pdf.
5. LEY ORGÁNICA 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPDGDD). *Boletín Oficial del Estado*. 2018, núm.

294 [consulta: 9 enero 2025]. Disponible en: https://www.boe.es/buscar/act.php?id=BOE-A-2018-16673.

6. PARLAMENTO EUROPEO Y CONSEJO DE LA UNIÓN EUROPEA. *Directiva (UE) 2019/790 sobre los derechos de autor y derechos afines en el mercado único digital*. *Diario Oficial de la Unión Europea* [en línea]. 2019 [consulta: 9 enero 2025]. Disponible en: https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A32019L0790.

Jurisprudencia referenciada

1. COURT OF JUSTICE OF THE EUROPEAN UNION. *British Horseracing Board Ltd v. William Hill Organization Ltd* (Case C-203/02) [en línea]. 2004 [consulta: 9 enero 2025]. Disponible en: https://eur-lex.europa.eu/legal-content/ES/TXT/?uri=CELEX%3A62002CJ0203.
2. COURT OF JUSTICE OF THE EUROPEAN UNION. *Fixtures Marketing Ltd v. Organismos* (Case C-46/02) [en línea]. 2004 [consulta: 9 enero 2025]. Disponible en: https://eur-lex.europa.eu/legal-content/ES/ALL/?uri=CELEX%3A62002CJ0444.
3. UNITED STATES COURT OF APPEALS FOR THE SECOND CIRCUIT. *Authors Guild v. Google Inc.* [en línea]. 2015 [consulta: 9 enero 2025]. Disponible en: https://law.justia.com/cases/federal/appellate-courts/ca2/13-4829/13-4829-2015-10-16.html.
4. UNITED STATES COURT OF APPEALS FOR THE NINTH CIRCUIT. *HiQ Labs, Inc. v. LinkedIn Corporation* [en línea]. 2019 [consulta: 9 enero 2025]. Disponible en: https://cdn.ca9.uscourts.gov/datastore/opinions/2022/04/18/17-16783.pdf.

Anexos

Anexo 1: Glosario de conceptos tecnológicos clave

API (Application Programming Interface):

Una API es un conjunto de reglas y protocolos que permite que diferentes aplicaciones o servicios se comuniquen entre sí. Actúa como un intermediario que define cómo los desarrolladores pueden interactuar con un software o servicio, proporcionando funciones específicas y datos en respuesta a solicitudes predefinidas. Por ejemplo, una API de redes sociales permite a las aplicaciones externas publicar contenido o recuperar datos como perfiles o publicaciones. Las APIs son esenciales para la interoperabilidad entre sistemas y habilitan la integración de servicios, como pasarelas de pago, análisis de datos o conectividad entre aplicaciones móviles y servidores. Aunque son muy útiles, su diseño y uso deben considerar la seguridad y el control de acceso para evitar vulnerabilidades.

Aprendizaje federado:

El aprendizaje federado es una técnica de inteligencia artificial que permite entrenar modelos sin que los datos salgan del dispositivo o entorno en el que se generaron. Esto se logra mediante la transmisión de actualizaciones del modelo local (pesos y gradientes) a un servidor central que combina estas actualizaciones para mejorar el modelo global. Este enfoque evita la necesidad de recopilar datos en un único servidor, reduciendo riesgos de privacidad y cumplimiento normativo. Su uso es crítico en sectores como la salud, donde los datos médicos deben permanecer en los hospitales, o en la banca, donde los datos financieros son extremadamente sensibles. A pesar de sus ventajas, enfrenta retos técnicos como garantizar la seguridad en las actualizaciones y lidiar con datos heterogéneos o distribuciones no equilibradas entre los participantes.

Blockchain:

El blockchain es una base de datos distribuida que permite registrar transacciones de forma segura, transparente e inmutable. Cada bloque de datos está vinculado al anterior mediante

algoritmos criptográficos, formando una cadena que es validada por una red descentralizada de nodos. Además de su popular uso en criptomonedas, como Bitcoin y Ethereum, el blockchain permite aplicaciones como contratos inteligentes (acuerdos autoejecutables), gestión de identidades digitales (evitando intermediarios), y trazabilidad en cadenas de suministro, asegurando la integridad de los productos desde su origen hasta el consumidor final. Su integración en sistemas legales y comerciales promete mayor transparencia y confianza, aunque plantea desafíos como el elevado consumo energético y la regulación de redes descentralizadas.

Criptomoneda:

Las criptomonedas son activos digitales que emplean criptografía para garantizar la seguridad de las transacciones y la creación de nuevas unidades. Operan en redes descentralizadas basadas en blockchain, lo que elimina intermediarios como bancos. Más allá de su uso en transacciones, las criptomonedas permiten innovaciones como los contratos inteligentes y las finanzas descentralizadas (DeFi). Sin embargo, su adopción plantea preocupaciones regulatorias relacionadas con el lavado de dinero, la volatilidad y su impacto medioambiental debido al consumo energético asociado a la minería.

Desaprendizaje federado:

El desaprendizaje federado es una técnica avanzada en inteligencia artificial que permite eliminar información específica de un modelo entrenado sin necesidad de reentrenarlo desde cero. Esto se utiliza en cumplimiento del derecho al olvido, garantizando que los datos eliminados no influyan en las decisiones futuras del modelo. Para lograrlo, el modelo se ajusta mediante algoritmos que identifican y borran las huellas de los datos en cuestión. Aunque es una solución prometedora, presenta desafíos técnicos y legales, como la dificultad de garantizar su efectividad total en sistemas complejos y distribuidos.

GANs (Generative Adversarial Networks):

Las GANs son un tipo de red neuronal que consta de dos componentes principales: un generador, que crea datos falsos pero realistas, y un discriminador, que evalúa si los datos son reales o generados. Esta dinámica competitiva permite a las GANs producir contenido de alta calidad, como imágenes fotorrealistas o datos sintéticos. Estas redes tienen aplicaciones

en el arte, el diseño, la creación de deepfakes y la generación de datos para entrenar modelos de IA. No obstante, plantean riesgos éticos, como la desinformación o el uso indebido de contenido generado.

Identidades descentralizadas (DIDs):

Las DIDs representan un nuevo paradigma en la gestión de identidades digitales. A diferencia de los sistemas tradicionales, donde una autoridad central como Google o Facebook controla las credenciales, las DIDs permiten que los usuarios posean y gestionen su identidad digital de manera autónoma. Están respaldadas por tecnologías como blockchain y DWNs, asegurando que las interacciones sean verificables y privadas. Esto tiene aplicaciones en sectores como la banca, la salud y el comercio electrónico, y plantea nuevas oportunidades para abordar el robo de identidad y la vigilancia masiva.

Nodos web descentralizados (DWNs):

Los DWNs son una tecnología que permite almacenar y gestionar datos de forma segura y descentralizada. Funcionan como contenedores de datos que garantizan su privacidad y accesibilidad solo a partes autorizadas. En combinación con blockchain y DIDs, los DWNs habilitan sistemas como la autenticación sin contraseñas, la portabilidad de datos y la gestión de derechos digitales. Esta tecnología es crucial para aplicaciones de inteligencia artificial y cumplimiento normativo, donde la privacidad y la trazabilidad son esenciales.

Web scraping:

Es una técnica utilizada para extraer información de sitios web de manera automatizada. Aunque es una herramienta útil para análisis de mercado, investigación y recopilación de datos, plantea desafíos legales y éticos, especialmente cuando los datos recopilados incluyen información personal o están protegidos por derechos de autor. En el contexto jurídico, el web scraping se sitúa en un área gris, dependiendo de las condiciones de uso de los sitios web y las normativas de protección de datos aplicables.

Web5:

Web5 es un concepto propuesto por la plataforma TBD, que combina tecnologías descentralizadas para devolver a los usuarios el control total sobre sus datos. Integra

blockchain, identidades descentralizadas (DIDs) y nodos web descentralizados (DWNs), permitiendo a las personas interactuar en línea sin depender de plataformas centralizadas. Esto tiene implicaciones significativas para la privacidad y la soberanía digital, ya que elimina intermediarios que suelen monetizar los datos personales. Su adopción requiere resolver desafíos técnicos como la interoperabilidad y el acceso universal a infraestructuras descentralizadas.