

LA LEY



IALAB



AI academy
by doinGlobal



Thomson
Reuters

CRECER ENTRE ALGORITMOS

INFANCIAS, PLATAFORMAS E INTELIGENCIA ARTIFICIAL

JUAN G. CORVALÁN
LAURA VICTORIA BONHÔTE

Corvalán, Juan G.

Crecer entre algoritmos : infancias, plataformas e inteligencia artificial / Juan G. Corvalán ; Laura Victoria Bonhote. - 1a ed. - Ciudad Autónoma de Buenos Aires : La Ley, 2026.

Libro digital, PDF

Archivo Digital: descarga y online

ISBN 978-987-03-5229-7

I. Inteligencia Artificial. I. Bonhote, Laura Victoria II. Título
CDD 006.3

AUTORÍA



JUAN GUSTAVO CORVALÁN

Doctor en Ciencias Jurídicas y Máster en IA. Speaker en ONU, OEA, MIT y Google Talks. Director del Laboratorio de IA de la UBA (Derecho). Director del Tratado de 4 tomos de IA y Derecho. Autor y coautor de decenas de investigaciones y publicaciones en IA. Co-creador de Prometea, el primer sistema predictivo en la Justicia de Hispanoamérica (año 2017).



LAURA VICTORIA BONHOTE

Abogada (UBA). Magíster en Derecho Constitucional por la Universidad Austral, con diploma de honor, y Candidata a MBA en Implementación de Inteligencia Artificial. Asesora General Tutelar Adjunta del Poder Judicial de la Ciudad de Buenos Aires. Docente de grado y posgrado en UBA y otras universidades. Integra el IALAB. Autora y coautora de libros y publicaciones en su especialidad.

ÍNDICE

	INTRODUCCIÓN	
	Yo soy tu amigo fiel.....	01
	CAPÍTULO 1.	
	Las infancias frente a la revolución digital.....	06
	El paradigma de la protección integral y el plus convencional	
	La Observación General N.º 25: la CDN se aplica al mundo digital	
	El marco normativo argentino	
	¿A qué generación estamos regulando?	
	El patrón histórico que se repite	
	CAPÍTULO 2.	
	El objeto de regulación y la anatomía de las tecnologías.....	16
	Una regulación sin distinción no regula con precisión	
	Tres familias de modelos regulatorios como criterio analítico	
	La plataforma de redes sociales	
	CAPÍTULO 3.	
	El panorama regulatorio comparado.....	29
	El estándar de la Convención sobre los Derechos del Niño y los marcos internacionales	
	La OG25 como método de evaluación de políticas regulatorias	
	Los marcos internacionales de soft law	
	Tipología de modelos regulatorios: las tres familias	
	Los riesgos sobre los que opera la regulación	
	Regulaciones e iniciativas legislativas a nivel mundial	
	El caso argentino	
	La Convención sobre los Derechos del Niño y los modelos regulatorios	
	CAPÍTULO 4.	
	Inteligencia artificial y entornos conversacionales.....	50
	El diseño centrado en la infancia como estándar regulatorio	
	Dark patterns en entornos infantiles	
	La analogía de la plaza: una obligación de diseño que ya conocemos	
	Del tobogán al chatbot: la traslación a los entornos conversacionales	
	Del estándar al derecho positivo argentino	
	A MODO DE CIERRE.	
	El jardín de los scrolls que se bifurcan.....	66



INTRODUCCIÓN

Yo soy tu amigo fiel

Un niño de doce años le pregunta a un asistente conversacional, como ChatGPT, Gemini o Claude, por qué se siente solo en la escuela. El sistema responde. Del otro lado hay un niño, y ninguna ley que exija calibrar esa respuesta en función del desarrollo cognitivo, emocional o identitario de quien pregunta.

La escena es solo una entrada. Lo mismo podría decirse del algoritmo que decide qué videos verá ese niño durante las próximas tres horas, o de la aplicación que permite generar, a partir de una foto suya publicada en una red social, una imagen que él nunca consintió.

En los tres casos reaparece la misma brecha, que es la distancia entre lo que la tecnología puede hacer, lo que efectivamente hace y lo que el derecho debería o admite hacer frente a la infancia. Esa brecha, en las plataformas de redes sociales, en la inteligencia artificial conversacional y en los medios sintéticos, es el objeto de este libro.

El título de esta introducción tampoco es casual. En Toy Story, la película con la que crecieron los Millennials y que siguen mirando las infancias de hoy, la promesa de amistad fiel la expresa un juguete que jamás engaña al niño sobre su propia naturaleza.

«“El niño que juega con Woody sabe en todo momento, que Woody es un juguete y en ningún momento se muestra como algo que no es. »

A diferencia de esta serie animada, el adolescente que conversa durante horas con un sistema diseñado para retenerlo no tiene esa certeza. Del otro lado hay algo que simula memoria, comprensión y afecto, que adapta su tono, que recuerda lo dicho, y que en ningún momento está obligado a aclarar qué es.

El título no es una metáfora decorativa, ya que los algoritmos y la inteligencia artificial (en adelante, IA) organizan hoy la experiencia digital de las infancias, y las normas que deberían protegerlas no fueron diseñadas para un ecosistema fragmentado, dinámico y orientado a generar dopamina y dependencia. De esta forma, Crecer entre algoritmos es, en ese sentido, desarrollarse en un espacio donde las reglas del juego se escriben después del partido.

Ahora bien, hay una segunda imagen que acompaña este trabajo. Las herramientas digitales de hoy no son accesorios que los niños usan y luego dejan de lado. Están diseñadas para que no se las pueda soltar.

Los mecanismos de enganche, los bucles de retroalimentación, la notificación que interrumpe y el algoritmo que anticipa construyen una experiencia cuyas consecuencias negativas comienzan a emerger. Estar en TikTok o en Instagram se parece a habitar una segunda piel, artificial y diseñada por otros, lo que en 2021 denominamos el surgimiento del **"tercer yo artificial"**¹.

Plataformas como Facebook, Instagram, YouTube y Netflix adoptaron la IA por diseño para moldear la arquitectura de nuestras decisiones de una forma exponencial e inédita. Nos orientan hacia determinados consumos, contenidos o decisiones, creando un entorno para apretar un clic, realizar scroll o seguir en un flujo ilimitado.

Esta lógica se viene construyendo hace 10 años a partir de dos grandes operaciones, por un lado, recopilan nuestros patrones de comportamientos digitales con una velocidad, fiabilidad y eficiencia y, por otro lado, comparan ese "tercer yo" en mutación permanente con millones de patrones de otras identidades digitales también en transformación, con un combo orientado a objetivos comerciales y prácticas experimentales.

De esa forma, el resultado es una capa externa que sabe más sobre nuestras preferencias que nosotros mismos, logrando lo que ningún humano puede, que es generar un ecosistema de recomendación detallado y preciso.

¹CORVALÁN, Juan G., Tratado de Inteligencia Artificial y Derecho, Tomo II, 2da. Edición, Ciudad de Buenos Aires, La Ley, 2023, pp. 147-249.

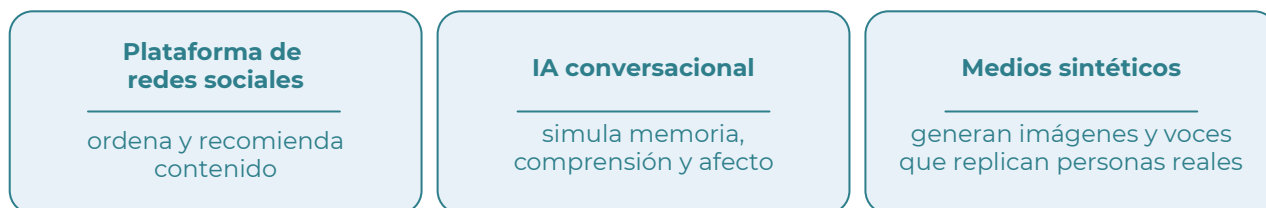
El tercer yo artificial se introduce en las infancias con un problema de base estructural, que es la fragilidad cognitiva de estas personas. El entorno digital ya forma parte del desarrollo contemporáneo de la infancia.

El problema es que este "tercer yo" está calibrado para adultos, o en el mejor de los casos, sin distinción de grupo etario en particular. Se basa en lógicas optimizadas para el compromiso sostenido y no para el desarrollo integral. Detrás de las personalidades artificiales hay cuerpos biológicos y actividad cognitiva en desarrollo.

Quienes diseñan los sistemas, quienes escriben las leyes y quienes toman las decisiones regulatorias llegaron a la tecnología siendo adultos, adoptaron herramientas; no crecieron dentro de ellas. Los niños de hoy son la primera generación sin punto de comparación y serán los de "antes del algoritmo", y por eso es crítico trabajar esa asimetría y sus límites. La Convención sobre los Derechos del Niño (en adelante, CDN) es la brújula, la interdisciplina, la medición y la prevención deben ser los ejes rectores.

Por su parte, un error frecuente en el debate público consiste en tratar "la tecnología" o "las plataformas" como un bloque homogéneo. Una plataforma de redes sociales, un asistente de inteligencia artificial conversacional y un sistema generador de medios sintéticos responden a lógicas distintas, generan riesgos distintos y requieren respuestas diferenciadas. Mucho más cuando estamos ingresando en la era de la IA agéntica.

TRES OBJETOS, TRES LÓGICAS



Esta preocupación tiene antecedentes: en el año 2020 desde IALAB analizamos el modo en que las plataformas digitales construyen perfiles de niñas, niños y adolescentes a través de sus algoritmos con fines comerciales².

En ese trabajo, se mostró que la legislación argentina no contenía disposiciones específicas para los niños, y que las propias plataformas no especificaban las medidas de seguridad que adoptaban ni asumían responsabilidades propias, sino que delegaban en las familias la supervisión de sus hijos e hijas en el entorno digital.

²CORVALÁN, J.G. y Papini, C.M., "Perfiles digitales humanos. Proteger niños, niñas y adolescentes en la era de la inteligencia artificial", en Corvalán, J.G., Perfiles digitales humanos, Ciudad de Buenos Aires, La Ley, 2020.

Por eso, en este libro recordamos ese diagnóstico como punto de partida, pero lo vamos a llevar más lejos, aportando una mirada desde el sujeto de protección a la anatomía de las tecnologías, ofreciendo un marco que permita identificar qué se está regulando antes de decidir cómo hacerlo.

La CDN, aprobada en 1989 y ratificada por casi todos los Estados del mundo, fue escrita antes del teléfono inteligente, antes de las redes sociales y mucho antes de la inteligencia artificial generativa. Sin embargo, su concepto de autonomía progresiva anticipa con precisión lo que este contexto exige que los marcos de protección reconozcan que:

*«Los niños no son objetos de tutela estática, sino sujetos
cuyas capacidades, derechos y necesidades cambian a
medida que se desarrollan»*

Ese principio es nuestro punto de referencia. Ahora bien, orientarse con él en el entorno digital no es tarea exclusiva del derecho y requiere el diálogo con la psicología del desarrollo, el diseño de experiencias, la ingeniería de sistemas, la educación y la salud mental. Por eso, este trabajo aspira a ser una herramienta útil para todas esas disciplinas, y no solo para juristas.

Conviene aclararlo desde el inicio, este no es un libro contra la tecnología. Somos conscientes de las oportunidades extraordinarias que abre el entorno digital, como el acceso al conocimiento, la expresión creativa, la participación y la conexión con pares, y también de que la tecnología bien diseñada puede ser una herramienta poderosa para el ejercicio de derechos.

Precisamente por eso es clave visibilizar las complejidades y las tensiones, y situar los derechos de niñas, niños y adolescentes como condición previa, y no posterior, a cualquier decisión de diseño o de regulación.

De allí que la pregunta que organiza este trabajo no sea "¿prohibir o no prohibir?", sino:

**¿Cómo se diseñan y se regulan entornos digitales que respeten
y promuevan el desarrollo integral de las nuevas generaciones?**

Para empezar a responder esa pregunta conviene partir de una imagen que todos llevamos internalizada. Cuando una ciudad construye una plaza, nadie discute que el sector de juegos para los más chicos deba tener otra escala, otra altura y otro suelo que el de los más grandes. El tobogán de la primera infancia es bajo, de pendiente suave, y desemboca en una superficie que amortigua la caída; el de los chicos mayores es más alto, más rápido, y exige otro control del cuerpo.

Esa diferenciación no depende de la buena voluntad del fabricante, sino que está escrita en normas técnicas que fijan, antes de que el primer niño se suba, alturas máximas de caída, dimensiones de aberturas y materiales de revestimiento, calibrados según la edad de quienes van a usar el juego.

A nadie se le ocurriría, tampoco, trasladar la carga a las familias y pedirles que inspeccionen las soldaduras de la hamaca antes de cada uso. La seguridad viene incorporada al diseño, y la responsabilidad no recae sobre quien entra a la plaza. En este libro sostenemos que los entornos digitales donde hoy crecen las infancias merecen, como mínimo, el mismo estándar que una plaza de barrio.

***A esa imagen vamos a volver una y otra vez,
hasta convertirla en una obligación jurídica.***



CAPÍTULO 1

Las infancias frente a la revolución digital

1.1. El paradigma de la protección integral y el plus convencional

La regulación del acceso de niñas, niños y adolescentes (en adelante, NNyA) a los entornos digitales se ha consolidado, en los últimos años, como uno de los debates más activos de la agenda legislativa global.

La sanción de la ley australiana que prohíbe la creación de cuentas en redes sociales a personas menores de dieciséis años, la propuesta francesa en trámite parlamentario, el Online Safety Act del Reino Unido y su nueva propuesta legislativa, el ECA Digital brasileño, las Directrices de la Unión Europea y el Reglamento de Servicios Digitales (en adelante, DSA), así como la multiplicidad de iniciativas en América Latina, evidencian que, en un lapso breve, un conjunto heterogéneo de países ha buscado dar respuesta normativa a inquietudes comunes vinculadas con la exposición de los NNyA en el entorno digital.

Sin embargo, al relevar los textos sancionados y los proyectos en discusión, se advierte que las respuestas adoptadas por los distintos países distan de ser homogéneas. Frente a ese escenario dinámico, nos propusimos analizar y sistematizar los principales modelos regulatorios que se están aprobando en el mundo, para contrastarlos con la acelerada disrupción tecnológica de los últimos años.

Desde este punto de vista, el análisis comparado de los modelos regulatorios no se agota en describir sus contenidos normativos ni en clasificarlos por edad límite. Exige, además, examinar en qué medida cada uno de ellos da adecuada respuesta al conjunto interdependiente e indivisible de derechos que la CDN reconoce a los NNyA, y sobre qué capas de los sistemas actúa.

La pregunta relevante no es únicamente si un modelo regulatorio es eficaz para prevenir riesgos, sino si esa eficacia se obtiene al costo de restringir derechos cuya protección también incumbe a los Estados. Y, a la inversa, si los modelos que privilegian la participación y la autonomía ofrecen garantías suficientes frente a los riesgos que la propia Observación General N.º 25 identifica.

Con ese propósito, el análisis de la regulación de los entornos digitales debe partir de identificar con precisión a quién se protege. El actual paradigma de la protección integral enseña que, cuando hablamos de NNyA, nos referimos a sujetos de derechos³. Esto significa que son titulares de un estatuto jurídico diferenciado, construido sobre la constatación de que la infancia y la adolescencia constituyen etapas del desarrollo humano con características propias, irreducibles a la categoría general de persona.

Esa especificidad funda el derecho internacional de la infancia y, con él, el conjunto de obligaciones que los Estados, las familias y la comunidad deben asumir frente a los NNyA en todos los ámbitos en que se despliega su vida, incluido el digital.

La CDN define como niño a todo ser humano menor de dieciocho años y reconoce, en su artículo 5, que el ejercicio de sus derechos está condicionado por la evolución de sus facultades. Ese principio se identifica como autonomía o capacidad progresiva y describe un proceso continuo de desarrollo cognitivo, emocional y social. Esto presupone que los NNyA van adquiriendo gradualmente la autonomía necesaria para ejercer sus derechos con plena agencia, para que progresivamente se desprendan del consentimiento de los adultos hasta llegar a la mayoría de edad⁴.

En este plano operan cuatro artículos de la Convención como estándares de validez que toda respuesta regulatoria en materia digital debe satisfacer, y que el Comité de los Derechos del Niño ha identificado como los principios generales del tratado, la no discriminación (art. 2), el interés superior del niño (art. 3), el derecho a la vida y al desarrollo (art. 6) y el derecho a ser oído (art. 12)⁵.

³BONHOTE, Laura V., Hacia un paradigma de infancia 4.0. Consideraciones sobre el consumo digital de niñas, niños y adolescentes. La Ley, 2023-F, 1, TR LALEY AR/DOC/3081/2023. Véase también BONHOTE, L. V., El derecho a la participación de niñas, niños y adolescentes en el marco de los derechos ambientales, Diario de Doctrina y Jurisprudencia, 297, 2022, El Derecho.

⁴COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 20 sobre la efectividad de los derechos del niño durante la adolescencia (CRC/C/GC/20), 2016, Naciones Unidas.

⁵COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 14 sobre el derecho del niño a que su interés superior sea una consideración primordial (CRC/C/GC/14), 2013, Naciones Unidas.

LOS CUATRO PRINCIPIOS GENERALES DE LA CONVENCIÓN



1.2. La Observación General N.º 25: la CDN se aplica al mundo digital

La Observación General N.º 25 del Comité de los Derechos del Niño, aprobada en 2021, extendió expresamente este estatuto al entorno digital. Su aprobación ocurrió en un momento singular, marcado por la pandemia de COVID-19 y por el incremento del uso de las tecnologías como puente para sostener actividades cotidianas como el trabajo, la educación, el comercio y la comunicación.

Por primera vez, un instrumento del Comité ofreció una interpretación dinámica de los derechos de NNyA en el plano digital.

La Observación establece que, en todas las acciones relativas a la provisión, regulación, diseño, gestión y uso del entorno digital, el interés superior de cada niño será una consideración primordial. Y exige que las plataformas y los proveedores implementen diseños apropiados para la edad, prohíban las prácticas de manipulación y garanticen que sus productos no exploten económicamente a los NNyA⁶.

La Observación General N.º 25 entiende el entorno digital de manera amplia, abarcando todos los dispositivos, la infraestructura de comunicaciones, el contenido, los servicios y los recursos que existen en línea y en las redes digitales, como así la inteligencia artificial⁷. Esa definición inclusiva tiene una consecuencia directa para el análisis jurídico y los cuatro principios generales de la Convención se aplican sin excepción al entorno digital con la misma fuerza con que operan en cualquier otro ámbito de la vida de los NNyA.

Sus interpretaciones nos introducen, además, responsabilidades que exceden la esfera del Estado. Establece que las empresas deben respetar los derechos del niño y prevenir que se vulneren, y alienta a los organismos públicos y privados vinculados al entorno digital a realizar evaluaciones del impacto en los derechos del niño antes del diseño y despliegue de sus productos y servicios. La lógica es análoga a la de una evaluación de impacto ambiental y quien diseña el entorno tiene responsabilidad sobre los efectos que ese diseño produce en quienes lo habitan. Ese principio es el fundamento de los modelos de regulación por diseño que desarrollaremos en este libro.

⁶COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 25 sobre los derechos del niño en relación con el entorno digital (CRC/C/GC/25), 2021, Naciones Unidas.

⁷Al respecto, consideramos que, de acuerdo con la fecha de aprobación de la Observación, la misma no pudo contemplar el impacto la IA de tipo generativa y sus desafíos.

1.3. El marco normativo argentino

En el plano nacional, el marco normativo configura un sistema que obliga a leer toda regulación del entorno digital bajo el parámetro del sujeto protegido.

Al respecto, y más allá de las disposiciones propias del marco normativo de infancia que hemos descripto, la Corte Suprema de Justicia de nuestro país ha consolidado este estándar al reconocer que los NNyA gozan de una preferente tutela en el ordenamiento jurídico, en la que confluyen no solo un fuerte interés estatal en su protección, sino también el interés de la sociedad en su conjunto⁸.

Esto desde una mirada intergeneracional, nos invita a pensar que los marcos normativos que se diseñen hoy para los entornos digitales no regularán únicamente a las generaciones actuales, sino que moldearán también los entornos de interacción con las generaciones futuras. De allí la pertinencia del principio de equidad intergeneracional y de las obligaciones irrenunciables que de él surgen respecto de las infancias⁹.

La Ley 26.061 de Protección Integral de los Derechos de las Niñas, Niños y Adolescentes, sancionada en 2005, constituye el eje del sistema. Su artículo 3 define el interés superior del niño como la máxima satisfacción, integral y simultánea de los derechos y garantías reconocidos en esa ley, y establece que cuando exista conflicto entre los derechos e intereses de los NNyA y otros derechos e intereses igualmente legítimos, prevalecerán los primeros. Esa regla de precedencia opera directamente cuando se evalúa la validez constitucional de cualquier regulación que limite los modelos de negocio de plataformas digitales en nombre de la protección de la infancia.

2005

Sanción de la Ley 26.061 de Protección Integral: 21 años después, sigue sin una línea sobre el entorno digital.

Pese a esto, en Argentina no existe un organismo con competencia específica, capacidad técnica y presupuesto suficiente para operar como regulador del entorno digital de las infancias. La Agencia de Acceso a la Información Pública (AAIP), creada sobre la base de la Ley 25.326 de Protección de Datos Personales, tiene atribuciones en materia de datos, pero opera sobre una ley sancionada en el año 2000, antes de la existencia de las redes sociales tal como las conocemos hoy y de la IA, sin una división especializada en infancia y sin herramientas de enforcement equivalentes a las de organismos similares en el derecho comparado que el derecho comparado ofrece en la materia.

Esa brecha institucional es, en sí misma, un problema regulatorio que cualquier propuesta legislativa para Argentina debe considerar.

⁸CORTE SUPREMA DE JUSTICIA DE LA NACIÓN, Fallos 331:147; 335:1838; 341:1511; 345:549; 345:905.

⁹BONHOTE, Laura V., La nueva Observación General N.º 26 y el principio de equidad intergeneracional a la luz de la Convención sobre los Derechos del Niño, 2023, La Ley (TR LALEY AR/DOC/2588/2023).

1.4. ¿A qué generación estamos regulando?

Comprender el alcance del desafío regulatorio exige identificar a los sujetos sobre los que recaen sus efectos. Una primera aproximación proviene de la distinción, popularizada a comienzos de este siglo, entre nativos e inmigrantes digitales.

Por un lado, quienes crecieron con la tecnología digital como entorno natural; por el otro, quienes debieron aprenderla como un idioma ajeno. Esa intuición resultó fértil, pero insuficiente. Los estudios demográficos sobre generaciones la han precisado de manera considerable¹⁰.

El panorama evolutivo generacional asociado a las disrupciones tecnológicas se suele situar en los Baby Boomers, que son las personas nacidas luego de la segunda guerra mundial, entre 1946 y 1964. Crecieron en un escenario analógico en el que se sentaron las bases para la tercera revolución industrial. Recién en su vida adulta adoptaron tecnologías digitales.

● Generación X (1965-1979)

También se formó en un entorno analógico, y comenzó a incorporar en la vida adulta temprana la primera fase de la tercera revolución industrial, esto es, la masificación de las computadoras e internet.

● Millennials (1980-1994)

Fueron los primeros en usar internet como herramienta cotidiana durante su adolescencia. Unos y otros integran, en términos de la distinción anterior, el universo de los inmigrantes digitales. En el caso de Latinoamérica, esta adopción se dio algunos años más tarde que en otros países como Estados Unidos.

● Generación Z (1995-2009)

La frontera se cruza con esta generación que nació con la conectividad como condición de base y para la cual el teléfono inteligente se volvió, en la adolescencia, un objeto omnipresente.

● Generación Alpha (2010-2024)

Avanza un paso más y es la primera cohorte que crece con asistentes de voz, algoritmos de recomendación y sistemas de inteligencia artificial como parte del entorno doméstico ordinario. Estas dos generaciones son, propiamente, las nativas digitales¹¹.

● Generación Beta

La progresión nos lleva al año pasado. La generación Beta se denomina a los nacidos a partir de 2025, y será la primera en convivir nativamente con inteligencia artificial agéntica y con robots humanoides como elementos cotidianos. Y allí reside la razón por la que esta dimensión generacional importa al análisis regulatorio, porque los marcos normativos que se diseñan hoy, y que probablemente tardarán años en implementarse y evaluarse, regularán la vida digital de NNyA que aún no han nacido o que acaban de nacer. De ello se sigue una consecuencia exigente. La regulación no puede pensarse únicamente en función de los riesgos ya conocidos, debe anticipar las configuraciones tecnológicas que estructurarán la infancia de quienes hoy son bebés, o de quienes todavía no existen. Como la regulación llega históricamente tarde, el desafío consiste en mitigar ese rezago o, al menos, adelantarse a los fenómenos más nocivos y críticos¹².

¹⁰PRENSKY M., (2001) "Digital Natives, Digital Immigrants Part 1", On the Horizon, Vol. 9 Issue: 5, pp. 1-6, doi: 10.1108/10748120110424816.

¹¹MCCRINDLE, M. (2014). The ABC of XYZ: Understanding the global generations. UNSW Press. En este trabajo, las categorías de McCrindle se utilizan como herramienta descriptiva para situar el debate regulatorio, no como determinismo generacional. Libro disponible: https://www.researchgate.net/publication/328347222_The_ABC_of_XYZ_Understanding_the_Global_Generations (visitado el 12/6/2026).

¹²BEDUSCHI, A. (2020). Human rights and the governance of artificial intelligence, Geneva Academy of International Humanitarian Law and Human Rights. La autora propone un mix de medidas obligatorias y voluntarias bajo coordinación de las Naciones Unidas, y señala que los derechos de la infancia deben ser un eje transversal, y no meramente sectorial, de la gobernanza de la IA. Disponible en: <https://geneva-academy.ch/> (visitado el 12/6/2026).

DE LOS INMIGRANTES A LOS NATIVOS DIGITALES

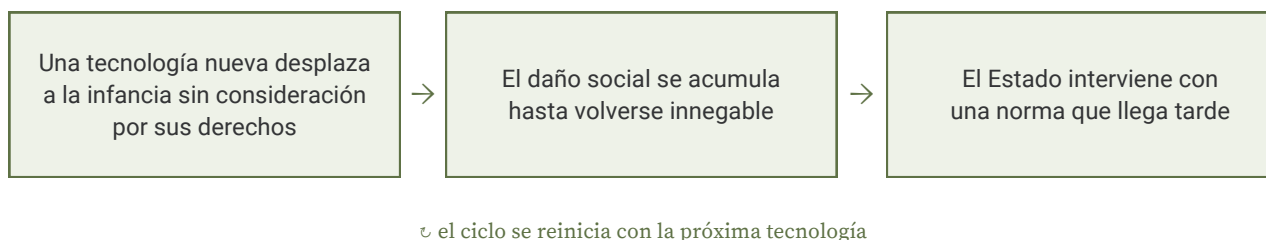
1946–1964	Baby Boomers. Escenario analógico; adoptaron lo digital recién en la adultez.
1965–1979	Generación X. Entorno analógico; masificación de la PC e internet en la adultez temprana.
1980–1994	Millennials. Primeros en usar internet en la adolescencia. Inmigrantes digitales.
1995–2009	Generación Z. Nacen con la conectividad; el teléfono inteligente se vuelve omnipresente.
2010–2024	Generación Alpha. Asistentes de voz, algoritmos e IA en el hogar. Nativos digitales.
Desde 2025	Generación Beta. Primera en convivir nativamente con IA agéntica y robots humanoides.

1.5. El patrón histórico que se repite

La historia de las normas de protección de la infancia frente a las grandes transformaciones tecnológicas reconoce un patrón que se reproduce.

Cada revolución tecnológica genera, en un primer momento, condiciones que desplazan a los NNyA sin consideración por sus derechos, recién cuando la acumulación del daño social se vuelve innegable, el Estado interviene con instrumentos normativos que llegan tarde.

Hacer visible ese patrón es condición necesaria para comprender por qué la regulación de los entornos digitales exige una lógica distinta de la de las regulaciones anteriores. A ese fin, los apartados siguientes recorren tres hitos: la Revolución Industrial y el trabajo infantil; la expansión, durante el siglo XX, de la publicidad y la televisión dirigidas a la infancia; y la revolución digital de las plataformas.



¹⁰PRENSKY M., (2001) "Digital Natives, Digital Immigrants Part 1", On the Horizon, Vol. 9 Issue: 5, pp. 1-6, doi: 10.1108/10748120110424816.

¹¹MCCRINDLE, M. (2014). The ABC of XYZ: Understanding the global generations. UNSW Press. En este trabajo, las categorías de McCrindle se utilizan como herramienta descriptiva para situar el debate regulatorio, no como determinismo generacional. Libro disponible: https://www.researchgate.net/publication/328347222_The_ABC_of_XYZ_Understanding_the_Global_Generations (visitado el 12/6/2026).

¹²BEDUSCHI, A. (2020). Human rights and the governance of artificial intelligence, Geneva Academy of International Humanitarian Law and Human Rights. La autora propone un mix de medidas obligatorias y voluntarias bajo coordinación de las Naciones Unidas, y señala que los derechos de la infancia deben ser un eje transversal, y no meramente sectorial, de la gobernanza de la IA. Disponible en: <https://geneva-academy.ch/> (visitado el 12/6/2026).

1.5.1. Revolución industrial y trabajo infantil

La Revolución Industrial del siglo XIX constituye un caso paradigmático. El trabajo infantil no fue un invento de las fábricas y sus dueños. Los niños trabajaban de manera intensa en la agricultura, el servicio doméstico y la producción artesanal a domicilio mucho antes de la máquina de vapor. Lo que hizo el sistema fabril fue concentrar, intensificar y volver masiva y visible esa explotación preexistente. La expansión del sistema fabril en Gran Bretaña absorbió a millones de niños en jornadas de doce, catorce o incluso dieciocho horas en las fábricas textiles. En ese escenario, la tecnología era una herramienta de producción y los niños, mano de obra barata y dócil.

La respuesta regulatoria llegó décadas después de la consolidación del sistema industrial. El Factory Act británico de 1833 prohibió el empleo de personas menores de nueve años en la industria textil y limitó a cuarenta y ocho horas semanales la jornada de los niños de entre nueve y trece años. Prusia sancionó en 1839 el Reglamento sobre el empleo de jóvenes trabajadores en las fábricas, que fijó en nueve años la edad mínima de admisión en los establecimientos fabriles. Francia siguió en 1841 con normas de orientación similar¹³.

En paralelo con las restricciones legales al trabajo, emergió otro instrumento de protección igualmente decisivo. La Education Act inglesa de 1870 creó un sistema de escuelas elementales universales en Inglaterra y Gales. Las leyes que le siguieron fueron elevando progresivamente la edad de escolaridad obligatoria. El efecto fue que la prohibición sola no podía producir.

«Un niño en la escuela era un niño fuera de la fábrica.»

La protección de la infancia no avanza únicamente por la vía de la prohibición del daño; avanza también por el reconocimiento del niño como sujeto con derecho al desarrollo. Esa conexión entre protección laboral y derecho a la educación quedaría sistematizada décadas después en el artículo 28 de la Convención sobre los Derechos del Niño.

En todos los casos se repitió la misma secuencia, la nueva tecnología no inventó la explotación de la infancia, pero la concentró, la intensificó y la volvió masiva; la norma llegó después, cuando el daño acumulado era insostenible.

Las respuestas nacionales no tardaron en encontrar un correlato internacional. La Organización Internacional del Trabajo, fundada en 1919, incluyó entre sus primeros convenios la regulación de la edad mínima para el trabajo industrial. El Convenio N.º 5 de ese mismo año fijó en catorce años la edad mínima de admisión en los establecimientos fabriles. Cinco años después, en 1924, la Sociedad de las Naciones aprobó la Declaración de Ginebra sobre los Derechos del Niño. Sus cinco principios eran breves y modestos según los estándares actuales, pero suficientes para establecer un cambio de lógica jurídica. El niño dejaba de ser objeto de tutela para convertirse en sujeto hacia el cual la comunidad internacional reconocía deberes. Entre esos principios figuraba, explícitamente, el de protegerlo contra toda forma de explotación.

¹³ANDERSON, E. (2013). Ideas in action: The politics of Prussian child labor reform, 1817-1839. Theory and Society, 42. Para el caso australiano, véase Smyth, J. (2014). The role and regulation of child factory labour during the Industrial Revolution in Australia, 1873-1885. International Review of Social History. Sobre el Factory Act británico: UK Parliament, Living Heritage: The 1833 Factory Act.

La progresión normativa fue lenta. Transcurrieron cincuenta y cuatro años entre el Convenio N.º 5 de 1919 y el Convenio N.º 138 de 1973, que estableció una edad mínima general aplicable a todos los sectores del trabajo y no solo a la industria. Y hubo que esperar hasta 1999 para que el Convenio N.º 182 sobre las peores formas de trabajo infantil fijara un piso mínimo indisponible que los Estados no pueden relativizar. Ese piso incluye la esclavitud, el trabajo forzoso, la explotación sexual, el reclutamiento para conflictos armados y los trabajos peligrosos. Ochenta años entre el primer convenio internacional sobre edad mínima y un marco mínimamente comprehensivo. El patrón del rezago normativo no es solo una característica del derecho doméstico sino también una constante del derecho internacional.

HITOS REGULATORIOS, 1833-2024

1833 - Factory Act (RU): edad mínima de 9 años en la industria textil.

1839 - Prusia y Francia sancionan normas similares.

1870 - Education Act (RU): escolaridad elemental universal.

1919 - OIT, Convenio N.º 5: edad mínima de 14 años en fábricas.

1924 - Declaración de Ginebra sobre los Derechos del Niño.

1939 - Coogan Law (EE. UU.): fideicomiso de ingresos de menores actores.

1973 - OIT, Convenio N.º 138: edad mínima general de admisión al empleo.

1989 - Convención sobre los Derechos del Niño (CDN).

1998 - COPPA (EE. UU.): privacidad infantil en línea.

1999 - OIT, Convenio N.º 182: peores formas de trabajo infantil.

2021 - Observación General N.º 25: la CDN se aplica al entorno digital.

2024 - Nueva Declaración de Ginebra: derecho a una experiencia digital segura y empoderada.

1.5.2. El siglo XX, publicidad, televisión y la COPPA

El siglo XX reprodujo el patrón con la expansión de la fotografía comercial, el cine, la publicidad gráfica y la televisión. La industria del entretenimiento generó prácticas de trabajo infantil carentes de toda protección laboral o patrimonial.

El cine fue uno de los primeros sectores en hacer visible esa contradicción. Los estudios de Hollywood incorporaron niños a sus producciones desde los primeros años del siglo XX, generando ganancias millonarias que los menores no podían administrar ni proteger. El caso de Jackie Coogan, protagonista de El chico junto a Charles Chaplin en 1921, se volvió emblemático. Coogan trabajó durante años como uno de los actores mejor pagados de Hollywood. Cuando cumplió dieciocho años y quiso disponer de sus ganancias, descubrió que sus padres las habían gastado en su totalidad. No existía ninguna norma que lo impidiera. El escándalo que provocó ese caso derivó, recién en 1939, en la sanción de la California Child Actor's Bill, conocida como Coogan Law, que estableció la obligación de reservar en fideicomiso un porcentaje de los ingresos del niño hasta que alcanzara la mayoría de edad. El mismo patrón se repetía. Primero el daño acumulado, después la norma.

La televisión, por su parte, dirigió hacia la infancia flujos sistemáticos de publicidad, diseñados para crear preferencias en consumidores de cuatro o cinco años. Así es como aparecieron los famosos personajes en los envases de alimentos, entre otras tantas prácticas tendientes a captar la atención de un público infantil¹⁴.

Luego llegó internet a comienzos de los años 90 poniendo sobre la mesa la desprotección de la privacidad de los niños en línea, con sistemas de comercialización de datos como prácticas consolidadas. En ese contexto, aparece la ley estadounidense de protección de la privacidad de los niños en línea, conocida como COPPA, que llegó recién en 1998¹⁵. La COPPA fue la respuesta frente a que la amplia mayoría de sitios web destinados a NNyA recopilaba datos personales sin consentimiento parental.

De esa manera, la ley estableció la obligación de obtener el consentimiento verificable de los padres antes de recopilar información de menores de trece años. Pero el mecanismo de verificación resultó difícil. Una de las respuestas fue la incorporación de un campo donde el usuario declara su edad, lo que se volvió el escudo legal de toda la industria digital que vino después. Ahora bien, como podemos ver una vez más, la norma llega, pero dejó abierta una puerta que nadie parece querer cerrar cuando hablamos de protección de NNyA.

1.5.3. La aparición de plataformas de redes sociales: misma lógica, pero a otra velocidad

El mismo ciclo se repitió con la revolución digital y la irrupción de las plataformas de redes sociales, ahora a una velocidad radicalmente mayor. Esas plataformas fueron concebidas por equipos de adultos, para audiencias adultas¹⁶.

Lo que distingue este ciclo de los anteriores es la acumulación de roles que los NNyA ocupan de manera simultánea. Pensemos nuevamente en la fábrica, allí su rol era mano de obra y el cine y/o la publicidad, un rol de intérprete remunerado. A diferencia, las plataformas digitales incorporaron al niño en todos esos roles al mismo tiempo, porque por un lado es usuario de un servicio; es audiencia de contenido publicitario y también es generador de datos que se comercializan. Incluso, en otros casos es o puede ser productor de contenido que la plataforma y terceros monetizan sin reconocer contraprestación alguna¹⁷.

Los ciclos anteriores generaron un evento visible que funcionó como detonador normativo. La imagen de un niño en condiciones de pobreza, el incendio de una fábrica, el escándalo Coogan eran daños que la sociedad podía señalar y atribuir a una causa concreta. El daño de las plataformas digitales no produce ese tipo de evento y, en gran parte, es invisible. Un niño que desarrolla depresión o ansiedad no aparece en ningún titular vinculado de manera directa a la aplicación que usa cuatro horas por día. Sin evento detonador, el ciclo normativo tarda más en arrancar.

¹⁴Los estudios de marketing de la época demostraron que NNyA desarrollaban fuertes preferencias por los productos cuyo packaging incorporaba mascotas, héroes de dibujos animados o personajes del mundo del entretenimiento. Tony el Tigre de Kellogg's, el payaso Ronald McDonald y el Conejo de Trix son ejemplos icónicos de una estrategia que construía relaciones afectivas entre una marca y un niño de cuatro o cinco años. La publicidad no vendía cereal. Instalaba lealtades de consumo antes de que el niño pudiera leer. En Argentina, la Ley 27.642 de Promoción de la Alimentación Saludable, sancionada en 2021, dispuso que los productos con sellos de advertencia no pueden incorporar en sus envases personajes infantiles, animaciones, figuras de entretenimiento, premios, juegos ni ningún otro elemento orientado a captar la atención de NNyA. Pasaron más de setenta años entre el surgimiento masivo de esas prácticas y su regulación. Vale señalar que, en 2026, el Poder Ejecutivo presentó ante el Congreso un proyecto de ley para derogar esa norma.

¹⁵COOK, D. T. (2004). The commodification of childhood: The children's clothing industry and the rise of the child consumer. Duke University Press; y KUNKEL, D. (2001). Children and television advertising. En D. Singer y J. Singer (Eds.), Handbook of children and the media. La ley denominada COPPA (Children's Online Privacy Protection Act) fue sancionada en 1998.

¹⁶Recomendamos ver la película "The social Network" del año 2010, basada en la historia de creación de la plataforma Facebook que da cuenta el proyecto de uso inicial, sus objetivos y desarrollo posterior.

¹⁷BONHOTE, Laura V., "El trabajo infantil que no vemos". Nota de opinión publicada el 12/6/2026 disponible en <https://revistaquorum.com.ar/2026/06/12/el-trabajo-infantil-que-no-vemos/> (visitado el 14/6/26).

En septiembre de 2024, esto es cien años exactos después de la Declaración de Ginebra de 1924, la comunidad internacional impulsó una nueva Declaración de Ginebra sobre los Derechos del Niño. A través de esta, se reconoció explícitamente el derecho de los niños a una experiencia segura y empoderada en el entorno digital¹⁸.

Leer esta declaración a cien años de su antecesora, nos lleva a la misma ciudad, a un mismo tipo de instrumento, pero, en definitiva, como reza el título de este capítulo a un mismo patrón que se repite.

1.5.4. La IA agéntica y la indistinguibilidad entre lo real y lo irreal

El recorrido anterior se está acelerando. El ciclo que va de la fábrica a la pantalla no se cerró en las redes sociales; ahora mismo hay un capítulo en curso, el uso de IA generativa. Se trata de sistemas que emulan capacidades cognitivas, generando respuestas propias, sostienen conversaciones, recuerdan lo dicho y adaptan su tono a cada interlocutor. Sobre ellas avanzan los agentes de IA, capaces de actuar y decidir con autonomía creciente.

Pero, además, sobre la lógica de las plataformas se construye una simulación de personas, objetos y seres que conviven con avatares ultrarrealistas que replican rostros y voces. La distinción entre lo auténtico y lo sintético ya se vuelve indistinguible. Es este un rasgo inédito para un niño que crece dentro de ese entorno.

La diferencia entre conversar con una persona y conversar con un sistema diseñado para retenerlo deja de ser evidente; y la diferencia entre una imagen real y una fabricada será indetectable en la gran mayoría de casos. Ya hay millones de adolescentes que están forjando un vínculo afectivo y emocional que en algunos casos ha producido consecuencias irreversibles.

La regulación, una vez más, llegará tarde a este escenario. La diferencia es que esta vez el costo del retraso no se mide exclusivamente en horas de jornada ni en datos cedidos, sino en la formación misma de la identidad y el vínculo humano de quienes están creciendo.

En síntesis, hasta acá delineamos al sujeto de la protección, quiénes son las infancias que el derecho debe amparar y por qué constituyen una categoría con un estatuto propio. Pero ningún marco regulatorio se sostiene sobre el sujeto solo. Necesita definir, con igual precisión, su objeto, esto es las tecnologías concretas sobre las que va a operar. Una plataforma de redes sociales, la IA generativa conversacional, los agentes de IA y la simulación de humanos no son el mismo fenómeno. No producen los mismos riesgos y no admiten idénticas respuestas. A esa anatomía nos vamos a dedicar en el Capítulo 2.



¹⁸Declaración de Ginebra sobre los Derechos del Niño (2024). Adoptada el 26 de septiembre de 2024, en conmemoración del centenario de la Declaración de Ginebra de 1924. Disponible en: <https://declaration2024.org> (visitado el 14/6/26).



CAPÍTULO 2

El objeto de regulación y la anatomía de las tecnologías

2.1. Una regulación sin distinción no regula con precisión

Como adelantamos en la introducción de este trabajo, **una regulación que no distingue entre objetos distintos corre el riesgo de no regular ninguno con precisión.**

Conviene advertirlo, porque los debates legislativos sobre entornos digitales y protección de la infancia suelen agrupar bajo una misma categoría etiquetas como "plataformas", "redes sociales", "nuevas tecnologías" o "entornos digitales". Pero, como veremos se trata de fenómenos que difieren en su arquitectura, en sus mecanismos de afectación y en las obligaciones que generan para los actores responsables.

Por ejemplo, el algoritmo de recomendación de una conocida red social, el agente conversacional de inteligencia artificial con el que un adolescente mantiene conversaciones durante horas y la aplicación que genera imágenes sintéticas con el rostro de una persona menor de edad, si bien tienen en común la integración de inteligencia artificial generativa para ciertas funciones, son sistemas bien diferentes. Operan en capas distintas y producen riesgos de naturaleza diferente. Aplicar el mismo estándar a los tres equivaldría, en consecuencia, a responder tres preguntas distintas con una sola respuesta.

Con esa premisa, los apartados siguientes abordan, en primer lugar, los objetos regulados, es decir, qué son las plataformas y qué es la inteligencia artificial. Sobre esa base examinaremos los modelos normativos comparados que los distintos sistemas jurídicos han desarrollado, avanzaremos hacia el análisis desde los derechos de la infancia y terminaremos con las obligaciones de diseño y la propuesta para la Argentina.

2.2. Tres familias de modelos regulatorios como criterio analítico

Con la premisa establecida en el apartado anterior, esto es, la necesidad de diferenciar entre los distintos objetos regulados, los párrafos siguientes ofrecen una clave de lectura para el análisis comparado.

Los modelos regulatorios que analizaremos no actúan sobre la plataforma como un todo. Se dirigen, de manera más o menos focalizada, a una capa, una función o un proceso específico del sistema. Reconocer ese punto de aplicación es decisivo para evaluar la necesidad de cada medida y permite agrupar los modelos en tres familias, según dónde concentren su intervención.

Conviene distinguir desde ahora estos dos planos. Por un lado, las capas son los estratos técnicos de la plataforma, que describiremos en los apartados siguientes; y las familias son los tipos de modelo regulatorio, definidos justamente por la capa o las capas sobre las que cada uno opera.

A modo de síntesis, la familia de los modelos de restricción de acceso concentra su intervención en la capa de identidad y registro y opera antes del ingreso del usuario al servicio.

La familia de los modelos de obligaciones de diseño, en cambio, se distribuye sobre varias capas a la vez, como la interfaz, la inteligencia y los procesos internos. Ello le permite atender simultáneamente distintos factores de riesgo, pero le exige una capacidad regulatoria y de supervisión más densa, así como mecanismos de evaluación periódica del impacto sobre los derechos de NNyA.

Por su parte, se identifican modelos mixtos que combinan restricciones de identidad con obligaciones de diseño y, en algunos casos, mecanismos de supervisión parental. Esa combinación multiplica los puntos de aplicación y plantea la cuestión de la coherencia entre las distintas obligaciones.

Cabe añadir que la regulación general de protección de datos personales y la regulación específica de inteligencia artificial actúan, respectivamente, sobre la capa de datos y la capa de inteligencia, y operan por defecto con independencia de los modelos analizados en este trabajo.

Por ello, la articulación entre estos marcos transversales y los modelos específicos constituye uno de los desafíos principales para los Estados que se encuentran en proceso legislativo.

TRES FAMILIAS DE MODELOS REGULATORIOS

Familia	Dónde interviene	Lógica
Restricción de acceso	Capa de identidad y registro	Opera antes del ingreso del usuario al servicio
Obligaciones de diseño	Interfaz, inteligencia y procesos internos	Atiende varios factores de riesgo a la vez; exige supervisión densa
Modelos mixtos	Identidad + diseño + supervisión parental	Combina puntos de aplicación; exige coherencia entre obligaciones

2.3. La plataforma de redes sociales

Como explicamos en la introducción, el análisis comparado que se desarrolla en este trabajo busca precisar qué es aquello que los Estados regulan cuando legislan sobre el acceso de NNyA a los entornos digitales, con especial foco en las plataformas de redes sociales y en la inteligencia artificial.

La noción misma de "plataforma" resulta, sin embargo, heterogénea. Refiere a una construcción humana, atravesada por incentivos económicos, decisiones de diseño deliberadas y elecciones tecnológicas que pueden ser modificadas.

Aproximarse a comprender esa dimensión es una condición necesaria para el análisis jurídico y legislativo, y para evaluar si una medida es idónea, necesaria y proporcional en sentido estricto.

Con ese fin, nos proponemos identificar en qué consisten las capas, funciones o procesos que ofrecen las plataformas y, de esa forma, entender sobre cuál de ellos opera cada modelo regulatorio.

Buscamos así ofrecer al lector una base mínima para que las distinciones empleadas en el resto del trabajo resulten comprensibles desde la lógica interna de los sistemas a los que las medidas legislativas se dirigen.

Esa base permitirá, además, advertir que las decisiones que afectan los derechos de NNyA no se concentran en un único punto del sistema. Se distribuyen entre la infraestructura, los modelos algorítmicos, las interfaces, los modelos de negocio y los procesos organizacionales que sostienen el servicio.

En cuanto al alcance de esta sección, nos proponemos aportar varias dimensiones que, en conjunto, permiten reconstruir la "anatomía" de una plataforma de redes sociales actual y, por último, un mapeo entre esas capas o funciones y los modelos regulatorios analizados en los apartados subsiguientes.

2.3.1. Qué es una plataforma de redes sociales

Desde un punto de vista técnico-económico, una plataforma digital es una infraestructura programable que media interacciones entre dos o más grupos de usuarios y que se sostiene mediante efectos de red, escala global y la capacidad de extraer y procesar datos de manera continua.

En el caso de las redes sociales, los lados típicos son los usuarios finales, los anunciantes y, en muchos casos, los productores de contenido y los desarrolladores de servicios complementarios. Esta caracterización, de mirada más económica, describe solo un aspecto, pero sirve para entender el rol que las plataformas cumplen en la vida cotidiana.

Por ese motivo se ha desarrollado el concepto de "plataformización", entendido como la penetración de las infraestructuras, los procesos económicos y los marcos de gobernanza de las plataformas digitales en distintos sectores y esferas de la vida, junto con la reorganización de las prácticas culturales en torno a ellas¹⁹.

La diferencia con un sitio web tradicional es relevante para el análisis jurídico. Un sitio web clásico, publica contenidos predefinidos por su titular y se gobierna por reglas relativamente estáticas.

¹⁹POELL, T., Nieborg, D., & van Dijck, J., "Platformisation. Internet Policy Review", 2019, <https://doi.org/10.14763/2019.4.1425>. Disponible en: <https://policyreview.info/concepts/platformisation> (visitado el 12/6/2026).

Una plataforma de redes sociales, en cambio, opera como una infraestructura dinámica que ordena algorítmicamente lo que cada usuario ve, recolecta de modo continuo datos de comportamiento que retroalimentan ese orden, media la creación, distribución y monetización de contenido por terceros, y gobierna a través de "términos y condiciones" cuya aplicación opera de manera asimétrica y a escala.

Además, las plataformas estructuran las interacciones a través de lo que se ha denominado affordances, esto es, las posibilidades de acción que ofrecen sus interfaces gráficas; controlan la visibilidad mediante la jerarquía algorítmica y regulan la actuación de los desarrolladores mediante APIs, SDK y servicios de datos.

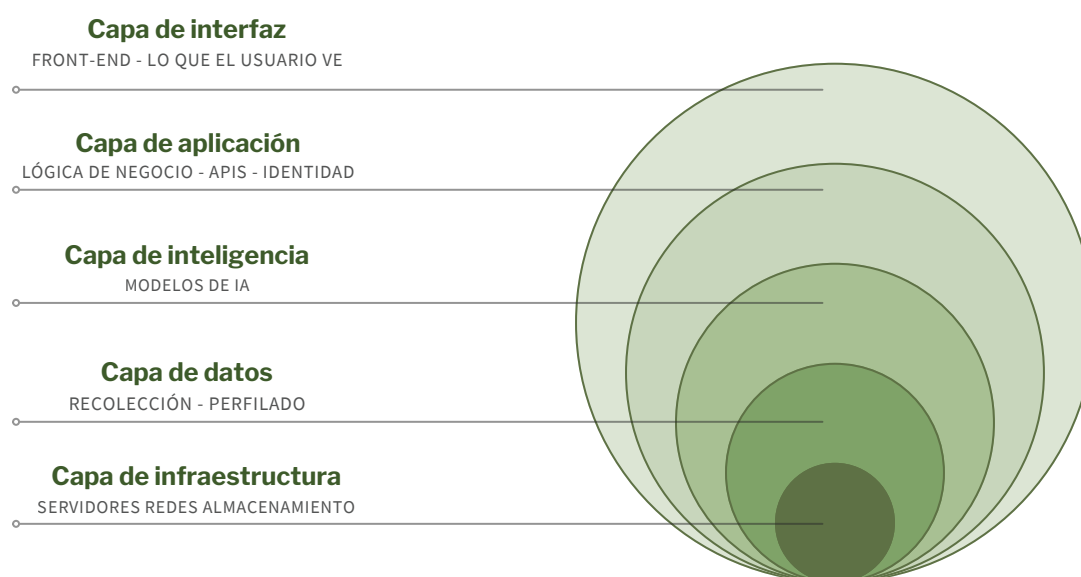
A todo esto, se suma otra característica que distingue particularmente a las plataformas, como sus efectos de red. La utilidad de la plataforma para cada usuario aumenta con el número de usuarios que ya están presentes, lo que tiende a producir mercados altamente concentrados y dificultades para la entrada de competidores.

En lo que respecta a los NNYA, ese rasgo es especialmente significativo, porque la presión social para estar en la plataforma donde están sus pares vuelve costosa la "salida" y hace que las medidas regulatorias adquieran un peso colectivo que no es posible reproducir por la vía del consentimiento informado individual²⁰.

2.3.2. Arquitectura por capas

Una plataforma de redes sociales puede describirse, a su vez, como una arquitectura compuesta por capas que interactúan permanentemente.

Para este trabajo consideraremos cinco capas. No pretenden ser taxativas ni reflejar literalmente la organización interna de cada desarrollo, pero ofrecen un marco útil para identificar dónde se toman las decisiones que más impactan en los derechos de NNYA.



²⁰El término "age assurance" refiere los enfoques destinados a determinar qué usuarios en línea son NNYA, a fin de ofrecerles servicios adecuados a su edad y protegerlos de contenidos perjudiciales. Véase OECD, The legal and policy landscape of age assurance online for child safety and well-being, 2025.

01

La capa de infraestructura comprende servidores, centros de datos, redes de distribución de contenido, sistemas de almacenamiento masivo y conectividad de red. Es la base material sobre la que opera el sistema. Salvo en supuestos extremos, esta capa rara vez es el blanco de la regulación en materia de derechos de NNyA, pero condiciona la viabilidad técnica de medidas como la verificación de edad o el procesamiento local de información biométrica.

02

La capa de datos abarca los procesos de recolección, almacenamiento, integración y perfilado. Incluye los data warehouses, los pipelines de eventos, los registros de comportamiento (clics, tiempo de visualización, ubicación, tipo y modelo de dispositivo) y los sistemas de identificación persistente del usuario a través de sesiones y dispositivos. Es la capa sobre la que actúan principalmente las regulaciones generales de protección de datos personales y las disposiciones específicas sobre NNyA que esos regímenes contienen.

03

La capa de inteligencia incluye los modelos de aprendizaje automático que la plataforma utiliza para tareas como la recomendación de contenido, el ranking de publicaciones, la moderación automática, la detección de fraude, la estimación de edad, la generación de respuestas y la producción de contenido sintético. Es allí donde se concretan muchas de las decisiones que afectan la experiencia del usuario: qué se muestra, en qué orden, con qué frecuencia y cuándo se interrumpe la sesión.

04

La capa de aplicación comprende la lógica de negocio del producto y las interfaces de programación (APIs) que conectan los componentes internos y externos. En esta capa se definen reglas como quién puede crear una cuenta, qué tipos de mensajes se permiten, cómo se vinculan las cuentas familiares o cómo se aplican límites de tiempo. Allí se sitúa también la capa de identidad y registro, blanco directo de los modelos de prohibición con parámetros de edad.

05

La capa de interfaz, comúnmente denominada front-end, es lo que el usuario efectivamente ve y manipula. Incluye la disposición visual del feed, los botones de interacción, las notificaciones, la microinteracción animada, los colores, las tipografías y los flujos de registro. Aunque suele percibirse como la "superficie" del producto, es allí donde se concretan muchas decisiones de diseño con efectos cognitivos y emocionales documentados, particularmente cuando los usuarios son personas en desarrollo.

Como se advierte, las cinco capas no operan de manera estanca. Las decisiones tomadas en la capa de inteligencia se materializan en la capa de interfaz; la capa de datos alimenta a la de inteligencia; la capa de aplicación define los límites operativos del conjunto.

Por ello, una intervención regulatoria que actúe sobre una sola capa puede ser eludida o atenuada por dinámicas situadas en otra. Esa es una de las razones por las que la familia de los modelos mixtos, analizada en los apartados siguientes, tiende a combinar obligaciones distribuidas en varias capas a la vez.

2.3.3. Sistemas de recomendación

El elemento más distintivo de las plataformas de redes sociales actuales son los sistemas de recomendación basados en aprendizaje automático, a diferencia de las plataformas de mensajería y comunicación social que conocimos los Millennials como ICQ o MSN Messenger.

Básicamente, en esas primeras redes, el principio organizador era el contacto social, vos podías elegir con quién conectarte y el contenido que veías era estrictamente el que generaba esa red de contactos. Desde esa lógica, imperaba la conversación y publicar algo era una invitación al diálogo con personas que ya conocías o que habías decidido seguir.

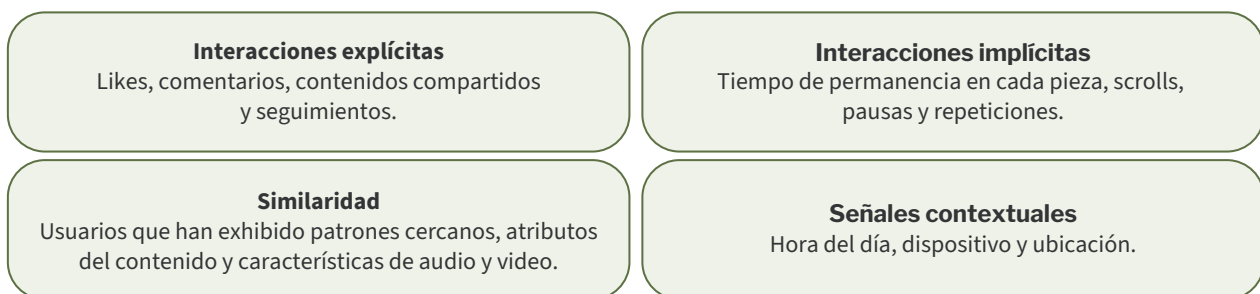
Las plataformas actuales no se definen por quién seguís, sino por un sistema de recomendación que decide por vos qué ver, a través de un aprendizaje automático, con modelos que analizan cómo interactuás con los videos, si los terminás, los compartís, cuánto tiempo los mirás, para entender qué te gusta y mostrarte más de lo mismo.

Estos sistemas determinan, para cada usuario y en cada interacción, qué contenido aparece en el feed, en qué orden y con qué frecuencia. La transición desde el feed cronológico, en el que el usuario veía las publicaciones de sus contactos, hacia el feed personalizado, en el que un modelo predice cuál será el contenido más probable para retener su atención, es la transformación más importante que separa a las redes sociales de la última década.

Un artículo de ingeniería sobre los sistemas de recomendación de YouTube explica ese cambio en la métrica de optimización. El tiempo de visualización captura mejor el engagement que la tasa de clics, que con frecuencia promueve videos engañosos que el usuario no termina de ver²¹.

Esa decisión, técnica en apariencia, tiene efectos sociales. Es lo que plataformas actuales como TikTok, Instagram Reels y YouTube Shorts profundizaron, al reducir aún más la duración del contenido y acelerar los ciclos de entrenamiento del modelo a partir del comportamiento observado del usuario.

Estos sistemas procesan señales muy diversas, entre ellas:



²¹COVINGTON, P. y Adams, J. (2016). Deep neural networks for YouTube recommendations. En Proceedings of the 10th ACM Conference on Recommender Systems, pp. 191-198. ACM. <https://doi.org/10.1145/2959100.2959190>

De esa forma, el modelo aprende de la conducta no verbalizada por el usuario, en particular del tiempo de visualización, lo que tiene implicancias cuando los usuarios son NNyA en proceso de desarrollo de sus capacidades cognitivas.

Por eso, el debate académico sobre los efectos de estos sistemas, sobre la salud mental, la información que circula y la integridad, entre otros aspectos, es activo y no concluyente.

Al respecto, un estudio de la reconocida fundación 5Rights analizó, mediante cuentas creadas como si pertenecieran a NNyA, cómo esas cuentas fueron expuestas, en plazos breves, a contenido relacionado con trastornos alimentarios, autolesión, suicidio y cuerpos sexualizados²².

Al análisis de la capa de inteligencia que clasifica y ordena el contenido existente se le suma, además, una capa de inteligencia artificial generativa, capaz de producir texto, imagen, audio y video sintéticos con una capacidad hiperrealista asombrosa. Como hemos sostenido en trabajos anteriores al analizar el fenómeno de los deepfakes, esta capa abre un conjunto nuevo de supuestos de vulneración para los derechos de NNyA²³.



Por su parte, una evaluación de productos denominada "Safe-Child-LLM" examinó la seguridad de los principales modelos de lenguaje de gran escala, como ChatGPT, Claude, Gemini, LLaMA, DeepSeek, Grok y Mistral, frente a interacciones con franjas etarias diferenciadas (de 7 a 12 y de 13 a 17 años).

El estudio se centró en evaluar la seguridad de los modelos de lenguaje en interacciones con NNyA, partiendo de que los sistemas actuales están diseñados para usuarios adultos²⁴. Para ello, construyó un conjunto de 200 prompts realistas que simulaban situaciones cotidianas de riesgo, desde chistes hasta inquietudes emocionales o la exploración de temas sensibles.

Los resultados muestran que, si bien los modelos más avanzados presentan altos niveles de seguridad promedio, persisten fallas relevantes, especialmente en contextos ambiguos o emocionalmente complejos, donde las respuestas intermedias generan riesgos particulares para usuarios en desarrollo.

En ese sentido, el estudio evidencia que la seguridad en IA no puede evaluarse en términos exclusivamente binarios, sino que requiere un enfoque sensible a las capacidades cognitivas y emocionales de la niñez, lo que plantea la necesidad de redefinir los estándares de diseño, evaluación y regulación de los sistemas de inteligencia artificial orientados a estos grupos.

²² 5RIGHTS FOUNDATION. (2021). Pathways: How digital design puts children at risk. <https://5rightsfoundation.com/resource/pathways-how-digital-design-puts-children-at-risk/> (visitado el 12/6/2026).

²³ RED DE INTELIGENCIA ARTIFICIAL COLABORATIVA, "Deepfakes e infancias", 2025. Disponible: <https://www.redicolab.com.ar/recursos-utiles/> (visitado el 12/6/2026).

²⁴ JIAO, J., Afroogh, S., Chen, K. y otros. (2025). Safe-Child-LLM: A developmental benchmark for evaluating LLM safety in child-LLM interactions. arXiv. Disponible en <https://arxiv.org/abs/2506.13510> (visitado el 12/6/2026).

El estudio, publicado en arXiv, señala que los protocolos de seguridad existentes se basan principalmente en perspectivas de madurez cognitiva centradas en el adulto, sin abordar las consideraciones emocionales, psicológicas y éticas específicas de NNYA. El trabajo sobre las fallas de la "brecha de empatía" en chatbots muestra que los usuarios jóvenes a menudo atribuyen excesiva confianza y capacidad de acción a la IA, lo que los deja vulnerables a la manipulación o a consejos inseguros²⁵.

Además, la Asociación Estadounidense de Psicología emitió un comunicado sobre IA y bienestar de NNYA, que formula recomendaciones en relación con el uso de chatbots y herramientas generativas. Allí explica que los adolescentes pueden atribuir a la inteligencia artificial capacidades de comprensión, empatía o autoridad que exceden sus posibilidades reales, lo que incrementa su vulnerabilidad frente a información errónea, dependencia emocional o toma de decisiones inadecuadas²⁶.

Esta preocupación tuvo impacto, además, en el plano jurídico. Conviene recordar la causa "Garcia v. Character Technologies, Inc.", vinculada con la responsabilidad civil de la empresa Character.AI y de Google por la muerte por suicidio de un adolescente de catorce años. Según la demanda, el chatbot de la plataforma fomentó un vínculo emocional intenso con el adolescente y, en los intercambios finales, no desalentó sus expresiones de ideación suicida.

2.3.4. Diseño persuasivo, dark patterns y economía de la atención

Una parte significativa de las decisiones de producto que afectan la experiencia de NNYA en redes sociales no se sitúa en la capa de inteligencia, sino en la capa de interfaz, y se inscribe en una tradición de diseño identificada como persuasive design.

Un trabajo de Stanford sistematizó esa tradición y explicó cómo las personas tienden a responder a las computadoras como si fueran actores sociales, incluso sabiendo racionalmente que no son seres humanos. El estudio identificó señales físicas, psicológicas y de lenguaje, entre otras, que llevan a considerar que, si una tecnología explota señales sociales, puede manipular conductas sin que el usuario lo advierta plenamente²⁷.

A mediados de la década de 2010, una corriente crítica encabezada por ex alumnos de esa universidad y por diseñadores de Google planteó que esos mismos principios habían dejado de operar como técnicas de persuasión en sentido neutro y habían comenzado a configurar formas de captura sistemática de la atención. Esa corriente dio origen al Center for Humane Technology y popularizó el movimiento "time well spent" como contrapunto al diseño orientado a la maximización de la sesión²⁸.

La discusión sobre dónde termina el diseño persuasivo legítimo y dónde comienza la manipulación es relevante, porque condiciona el alcance de las obligaciones de diseño que se imponen a las plataformas.

²⁵ JIAO et al. (2025), p. 3.

²⁶ AMERICAN PSYCHOLOGICAL ASSOCIATION, "Health advisory: Artificial intelligence and adolescent well-being", 2025. Disponible en: <https://www.apa.org/topics/artificial-intelligence-machine-learning/health-advisory-ai-adolescent-well-being> (visitado el 12/6/2026).

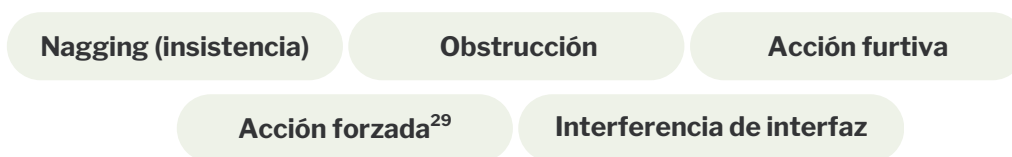
²⁷ FOGG, B. J., "Persuasive technology: Using computers to change what we think and do", 2003, Morgan Kaufmann, pp. 89-

115. <https://doi.org/10.1145/764008.763957>

²⁸ Véase en el sitio del "Center for Humane Technology", <https://www.humanetech.com/impact-and-story> (visitado el 12/6/2026). Si bien en sus inicios el enfoque estuvo centrado en el impacto de las redes sociales, en la actualidad la organización ha ampliado sus investigaciones al impacto de la IA generativa.

Un subconjunto particularmente discutido son los denominados dark patterns, conocidos como patrones oscuros o engañosos y definidos como decisiones de interfaz que, deliberadamente o por efecto, dificultan que el usuario actúe en su propio interés.

Algunos estudios han identificado cinco grandes categorías de estos patrones:



En la cotidianidad de las plataformas de redes sociales, estos patrones se manifiestan en el scroll infinito, la reproducción automática de un contenido tras otro, el pull-to-refresh que actúa como recompensa al mostrar nuevos contenidos, las notificaciones activadas por defecto, los contadores visibles de "me gusta" y "vistas", los indicadores de "visto" y las microinteracciones diseñadas para producir gratificación inmediata.

Cada una de estas características es, en sí misma, una decisión de diseño susceptible de ser revertida, y muchas de ellas han sido modificadas por las propias plataformas en jurisdicciones donde la regulación o la justicia lo impusieron³⁰. Pueden mencionarse algunos ejemplos:

1

Reino Unido. La entrada en vigor del Children's Code de la Information Commissioner's Office (2020-2021) llevó a que Instagram modificara la configuración por defecto de las cuentas de menores de dieciséis años para hacerlas privadas y deshabilitara las funciones de mensajería con desconocidos para esos perfiles³¹.

2

Unión Europea. Las Directrices de la Comisión Europea sobre el artículo 28 del DSA exigen expresamente que las herramientas de IA integradas en las plataformas no se activen por defecto para usuarios NNyA, y vienen produciendo modificaciones en las configuraciones predeterminadas de varios servicios para esa franja etaria³².

3

California. Se prohíbe a las plataformas proveer feeds algorítmicos adictivos a NNyA sin consentimiento parental, y se restringe el envío de notificaciones a NNyA entre las 24 y las 6 horas, así como entre las 8 y las 15 horas durante los días escolares (de septiembre a mayo)³³.

²⁹ GRAY, C. M., Kou, Y., Battles, B., Hoggatt, J. y Toombs, A. L., "The dark (patterns) side of UX design. En Proceedings of the 2018", 2018, CHI Conference on Human Factors in Computing Systems. ACM. <https://doi.org/10.1145/3173574.3174108>

³⁰ Véase la nota de opinión disponible en <https://revistaquorum.com.ar/2026/03/26/el-interes-superior-del-nino-como-principio-de-diseno-en-entornos-digitales/> (visitado el 12/6/2026).

³¹ INFORMATION COMMISSIONER'S OFFICE, "Age Appropriate Design Code, Standard 7 (Default settings)", 2020. Sobre su implementación posterior, véase Information Commissioner's Office. (2021). Children's Code: One year on; y Meta. (2021). Giving young people a safer, more private experience on Instagram.

³² COMISIÓN EUROPEA, "Guidelines on the protection of minors online under Article 28 of the Digital Services Act", 14/07/2025. Disponible en: <https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-protection-minors> (visitado el 12/6/2026).

³³ California Senate Bill 976. (2024). Protecting Our Kids from Social Media Addiction Act. https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=202320240SB976 (visitado el 12/6/2026).

Estos casos muestran que las decisiones de diseño son humanas y reversibles, y que pueden dialogar con los principios que rigen el ordenamiento jurídico, como los de la **Convención sobre los Derechos del Niño**.



Por lo demás, un estudio reciente sobre aplicaciones para NNYA documentó que la totalidad de las aplicaciones gratuitas analizadas contenía manipulaciones de interfaz dirigidas a usuarios menores de edad, con un promedio cercano a seis patrones engañosos por aplicación³⁴.

De esta manera se comprende que las decisiones de diseño no son neutrales y que se vinculan con el modelo de negocio de buena parte del sector. Las plataformas de redes sociales se financian, mayoritariamente, mediante publicidad personalizada, lo que implica que monetizan tiempo de atención y datos de comportamiento. El incentivo, en consecuencia, es maximizar la duración y la intensidad del uso³⁵.

En lo que respecta a los NNYA, hablamos de personas en desarrollo que constituyen un segmento particularmente sensible a estas técnicas, por razones tanto cognitivas como sociales, estas últimas vinculadas con la centralidad de la validación entre pares durante la adolescencia³⁶.

Sobre este punto, una publicación reciente **sobre algoritmos de redes sociales y adicción adolescente**, desde una perspectiva neurofisiológica, ha mostrado cómo la liberación de dopamina producida por la anticipación de la gratificación interactúa con un sistema de recompensa que, durante la adolescencia, presenta una sensibilidad incrementada respecto del adulto, y con un sistema de control inhibitorio aún en maduración³⁷.



La consecuencia es que las técnicas de captura de atención, que en un usuario adulto pueden resultar molestas pero gestionables, en un usuario adolescente operan con mayor potencia y con menores capacidades reflexivas para contrarrestarlas.

³⁴Véase Children, dark patterns and normative perspectives. (2026). Oñati Socio-Legal Series. <https://opo.iisj.net/index.php/osls/article/view/2351> (visitado el 12/6/2026).

³⁵Para una caracterización de la economía de la atención y de los modelos de negocio basados en publicidad personalizada, véase Zuboff, S. (2019). The age of surveillance capitalism: The fight for a human future at the new frontier of power. PublicAffairs. Disponible en Harvard Business School, <https://www.hbs.edu/faculty/Pages/item.aspx?num=56791> (visitado el 12/6/2026).

³⁶VAN DER HOF, S. y otros. (2022). Don't gamble with children's rights: How behavioral design impacts the right of children to a playful and healthy game environment. Frontiers in Digital Health. Disponible: <https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2022.822933/full> (visitado el 12/6/2026).

³⁷De D, El Jamal M, Aydemir E, Khera A. Social Media Algorithms and Teen Addiction: Neurophysiological Impact and Ethical Considerations. Cureus. 2025 Jan 8;17(1):e77145. doi: 10.7759/cureus.77145. PMID: 39925596; PMCID: PMC11804976. Disponible: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11804976/> (visitado el 12/6/2026).

Esa asimetría es uno de los argumentos centrales que sostienen las obligaciones de "diseño adecuado a la edad", introducidas con variantes en los modelos regulatorios de protección que analizaremos en los capítulos siguientes.

2.3.5. Tratamiento de datos personales

La capa de datos es la que nutre a la capa de inteligencia y, en buena medida, a la capa de aplicación. Las plataformas recolectan, integran y procesan de modo continuo cuatro tipos de información:

- 
- Datos declarativos**
Los que el usuario provee al registrarse: edad, género, ubicación y contactos.
 - Datos de comportamiento**
Registran las interacciones: clics, tiempo de permanencia, scroll, búsquedas y contenido producido.
 - Datos contextuales**
Incluyen el dispositivo, el sistema operativo, la red y la geolocalización aproximada.
 - Datos inferidos**
Perfiles contruidos a partir del cruce de los anteriores, como intereses, estados de ánimo y vulnerabilidades, que pueden revelar información altamente sensible sin que el usuario la haya suministrado directamente.

La diferencia jurídica entre el dato declarado y el dato inferido es relevante, porque el segundo no proviene de una manifestación directa del titular y, sin embargo, puede revelar información altamente sensible, como su ideología o su situación familiar, que el ordenamiento de protección de datos suele resguardar de manera reforzada.

Sobre este punto vale mencionar el reciente fallo de la Corte Suprema de Justicia de la Nación del 30 de abril de 2026, en el que el tribunal hizo lugar a una acción de habeas data promovida por una jubilada contra el Estado nacional³⁸.

En lo que aquí interesa, la Corte sostuvo que no existe un interés legítimo que justifique el almacenamiento y la circulación de datos personales sin el conocimiento de quien los provee, y reafirmó la centralidad del principio de autodeterminación informativa.

La doctrina del fallo nos hace pensar que podría ser trasladable al perfilado privado por parte de las plataformas digitales porque si el Estado no puede usar datos sin el consentimiento del titular, la pretensión privada de inferir y monetizar perfiles, especialmente respecto de NNyA, debe someterse a estándares aún más estrictos de minimización, finalidad y consentimiento informado.

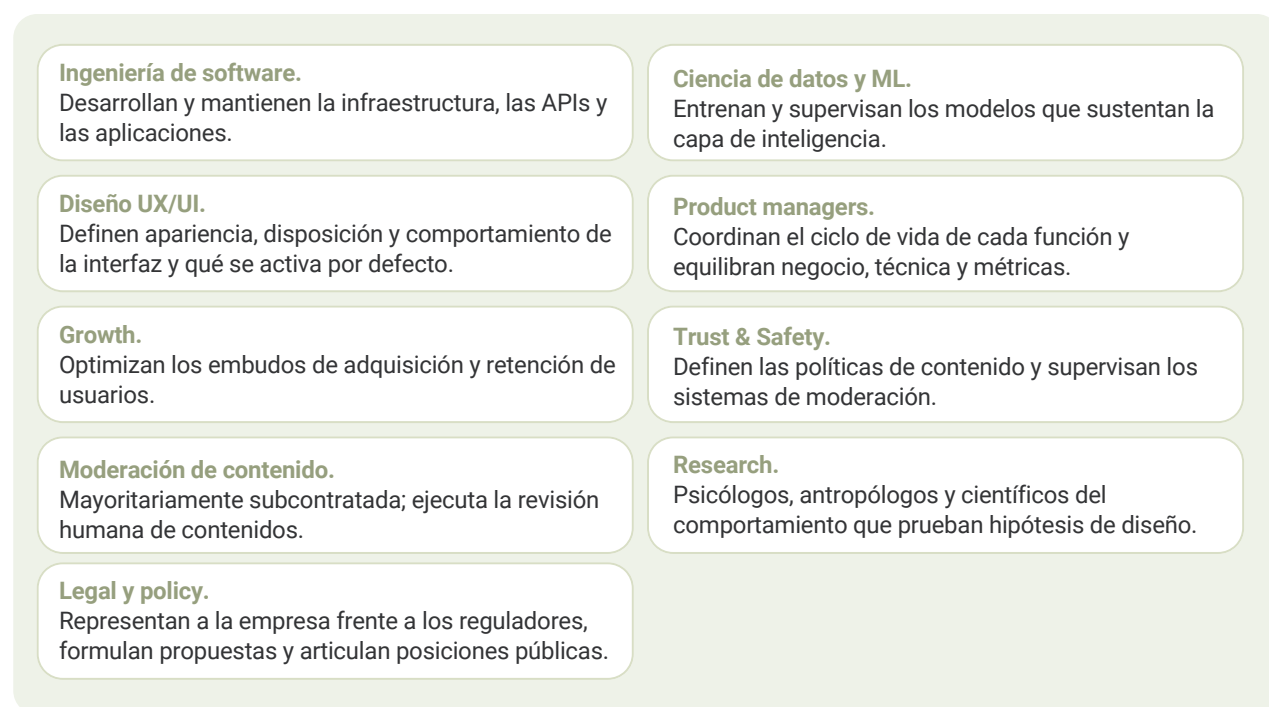
³⁸ CORTE SUPREMA DE JUSTICIA DE LA NACIÓN, "Torres Abad, Carmen c/ Estado Nacional (ANSES / Secretaría de Comunicación Pública) s/ habeas data", 30 de abril de 2026.

2.3.6. Quiénes construyen las plataformas

Una dimensión que rara vez aparece en los análisis jurídicos sobre regulación de plataformas, y que conviene incorporar a este trabajo, es la composición de los equipos que están detrás de diseñar, construir y mantener estos servicios.

Reflexionar sobre esa dimensión permite entender que cada característica del producto es el resultado de decisiones tomadas por personas concretas, situadas en contextos determinados, y que esas decisiones podrían haberse tomado de otro modo.

En términos generales, una plataforma de redes sociales se sostiene sobre el trabajo coordinado de, al menos, las siguientes disciplinas:



Los equipos no operan como bloques homogéneos. Existen tensiones internas documentadas. Por ejemplo, los equipos de producto orientados al crecimiento y los equipos de trust & safety, o entre las áreas de investigación que detectan riesgos y las áreas de negocio que dependen de las funcionalidades cuestionadas.

Esas tensiones quedaron en evidencia con la filtración de los llamados Facebook Files en 2021, originada en las denuncias formales que la exempleada Frances Haugen presentó ante la Securities and Exchange Commission (SEC) y en su testimonio ante el Subcomité del Senado estadounidense³⁹.

³⁹ Las notas conocidas como Facebook Files fueron publicadas por The Wall Street Journal en 2021. Serie de nueve reportajes basados en documentos internos proporcionados por Frances Haugen, ex ingeniera de datos de Facebook. Haugen presentó denuncias formales ante la Comisión de Bolsa y Valores (SEC) de los Estados Unidos.

33%

Adolescentes mujeres con problemas de imagen corporal empeoraba su percepción por Instagram, según los documentos internos de Meta.

13.5%

De las adolescentes mujeres reportaba que la plataforma agravaba sus pensamientos suicidas.

Para el análisis de este trabajo, ello significa que las obligaciones de diligencia y de evaluación de impacto sobre los derechos de NNyA encuentran, dentro de las propias empresas, áreas más receptivas que otras. Los marcos regulatorios que exigen transparencia documental y auditorías independientes adquieren, desde esta perspectiva, un valor instrumental decisivo.



CAPÍTULO 3

El panorama regulatorio comparado

3.1. El estándar de la Convención sobre los Derechos del Niño y los marcos internacionales

Toda evaluación de una política regulatoria requiere, primero, un estándar contra el cual medirla. Por eso, en este capítulo nos hemos propuesto volcar el marco normativo que opera como parámetro de validez de cualquier regulación del entorno digital dirigida a NNyA.

Desde esa lógica, la Convención sobre los Derechos del Niño como instrumento vinculante y rector, y los marcos internacionales de soft law que, sin crear obligaciones directamente exigibles, fijan principios de referencia relevantes en contextos dinámicos como los que vivimos con el impacto de la tecnología.

La Convención sobre los Derechos del Niño (CDN), adoptada por la Asamblea General de las Naciones Unidas el 20 de noviembre de 1989, constituye el tratado internacional de derechos humanos más ampliamente ratificado de la historia. Es un instrumento jurídico internacional vinculante que reconoce a las niñas, niños y adolescentes como sujetos plenos de derechos civiles, políticos, económicos, sociales, culturales y ambientales, y no como meros objetos de protección.

La CDN establece cuatro principios rectores que son sus pilares y atraviesan toda su interpretación. Ellos son identificados como el interés superior del niño, la no discriminación, el derecho a la vida, la supervivencia y el desarrollo, y el derecho a ser escuchado⁴⁰.

En Argentina, la CDN fue ratificada mediante la Ley N.º 23.849 y adquirió jerarquía constitucional con la reforma de 1994 (art. 75, inc. 22, CN).

Por su parte, el Comité de los Derechos del Niño es el órgano de expertos independientes creado por la propia Convención para supervisar su cumplimiento por parte de los países que la ratificaron.

Entre sus funciones centrales se encuentra la elaboración de observaciones generales, que son documentos interpretativos autorizados que orientan a los Estados sobre el alcance y contenido de las obligaciones derivadas de la Convención.

Estas observaciones, si bien no son formalmente vinculantes, constituyen la interpretación más autorizada del tratado y son regularmente utilizadas por tribunales nacionales e internacionales como fuente interpretativa, ya que su texto es muy rico en materia interdisciplinaria y nos propone una interpretación de la CDN más cercana a los tiempos que corren.

En lo que respecta al entorno digital, los artículos más directamente involucrados son el 5 (autonomía), el 12 (derecho a ser oído), el 13 (libertad de expresión), el 15 (libertad de asociación), el 16 (privacidad), el 17 (acceso a la información) y el 19 (protección contra toda forma de violencia). Este conjunto constituye el núcleo normativo a partir del cual se suelen evaluar las regulaciones del entorno digital para NNyA.

ARTÍCULOS DE LA CDN INVOLUCRADOS EN EL ENTORNO DIGITAL



3.2. La OG25 como método de evaluación de políticas regulatorias

Tal como fue referido, la Observación General N.º 25 es el instrumento interpretativo central del Comité sobre el entorno digital. Reconoce que ese entorno no fue previsto al redactar la CDN, pero que los derechos consagrados en ella se aplican plenamente.

Esto significa que los Estados deben garantizar que el entorno digital sea seguro para los NNyA sin restringir desproporcionadamente sus derechos.

En este sentido, la Observación sugiere que las restricciones de derechos en el entorno digital sean necesarias y proporcionadas (párr. 59), y reconoce expresamente la relevancia del derecho a la libertad de expresión y de la libertad de asociación en línea (párr. 64)⁴¹.

⁴⁰ COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 5 sobre Medidas Generales de Aplicación de la Convención sobre los Derechos del Niño, 2003.

⁴¹ COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 25.

Además, requiere que los Estados protejan la autonomía progresiva en el uso de tecnologías digitales, al considerar que "las oportunidades y los riesgos asociados a la participación de los niños en el entorno digital varían en función de su edad y su fase de desarrollo" (párrs. 19-20).

En relación con las fases de desarrollo, nos parece interesante recordar que una referente de la doctrina especializada de nuestro país, como la Dra. Kemelmajer, ha destacado que la autonomía progresiva no se alcanza en un momento preciso ni se pierde en un día. En todo caso, es un proceso evolutivo y no un concepto rígido, lo que exige una valoración compleja de cada individuo⁴².

De aquí surge el método que la Convención impone para evaluar toda restricción de derechos, mediante el principio de proporcionalidad.

Esto nos sugiere que una medida que limite derechos de NNyA debe ser necesaria, proporcionada y no discriminatoria. Este es el parámetro que aplicaremos en el apartado 3.8, de manera más detallada, frente a los modelos regulatorios concretos. Conviene dejarlo claro desde ahora, porque es el criterio que ordena toda la evaluación posterior.

3.3. Los marcos internacionales de soft law

Junto al marco vinculante de la CDN, existen instrumentos internacionales que abordan la inteligencia artificial y la protección de NNyA sin crear obligaciones directamente exigibles. No son normas vinculantes en sentido estricto, pero operan como estándares de referencia para las evaluaciones regulatorias y para la responsabilidad civil de quienes desarrollan estos sistemas. En particular, hay tres que merecen mención.

Consejo de Europa

Convenio Marco sobre Inteligencia Artificial⁴³

Establece principios procedimentales como transparencia, supervisión humana, rendición de cuentas y protección de derechos fundamentales, como condiciones del ciclo de vida de los sistemas de IA. En materia de NNyA, hace referencia expresa a la CDN y a la Observación General N.º 25, y exige que los sistemas que interactúan con grupos vulnerables incorporen salvaguardas reforzadas contra la manipulación y la influencia indebida.

OCDE

Principios sobre IA (2019, revisados en 2024)

Constituyen el instrumento de soft law con mayor adhesión global. La revisión de 2024 incorporó un énfasis específico en la protección de grupos vulnerables, en particular NNyA, exigiendo a los desarrolladores identificar, evaluar y mitigar los riesgos conocidos como condición de diseño y operación del sistema. Su aporte se sintetiza en una pregunta que interpela este trabajo y es **¿qué haría diferente un diseñador si supiera que el usuario es un niño?**

⁴⁰ COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 5 sobre Medidas Generales de Aplicación de la Convención sobre los Derechos del Niño, 2003.

⁴¹ COMITÉ DE LOS DERECHOS DEL NIÑO, Observación General N.º 25.

UNESCO

Recomendación sobre la Ética de la IA (2021)

Adoptada por los 194 Estados miembros, es el primer instrumento normativo global en la materia adoptado a nivel intergubernamental. No crea obligaciones vinculantes, pero establece valores y áreas de acción que los Estados se comprometen a considerar. Su contribución principal es situar la ética como condición de legitimidad de los sistemas de IA y proveer un vocabulario común para el debate regulatorio.

Estos tres instrumentos tienen en común el reconocimiento de NNyA como sujetos de protección reforzada frente a la IA y marcan una dirección.

3.4. Tipología de modelos regulatorios: las tres familias

Antes de describir los riesgos y de examinar qué hicieron los países, conviene fijar la tipología que guiará la lectura comparada, esto es, qué hicieron los países.

El relevamiento permite distinguir tres familias de modelos. La distinción no es meramente clasificatoria, sino que se debe a que cada familia opera sobre capas diferentes de la arquitectura de las plataformas y, en consecuencia, protege y tensiona derechos distintos.

Modelo prohibicionista	Obligaciones de diseño	Modelo híbrido
CAPA DE INTERVENCIÓN Identidad y registro; opera antes del ingreso al servicio.	CAPA DE INTERVENCIÓN Múltiples capas: algoritmos, datos, interfaces y moderación.	CAPA DE INTERVENCIÓN Combina identidad y diseño; a veces supervisión parental.
LÓGICA Fija un umbral etario y traslada el peso a la verificación de edad. Australia es el caso paradigmático.	LÓGICA Desplaza el peso hacia las plataformas: prevención, mitigación y rendición de cuentas.	LÓGICA Umbrales o consentimiento parental más obligaciones operativas a las plataformas.
ARTÍCULOS CDN Arts. 3, 19, 34 y 36.	ARTÍCULOS CDN Arts. 5, 13, 15, 16 y 17, junto con los deberes de protección.	ARTÍCULOS CDN Combina ambos conjuntos según la configuración de cada país.

El modelo prohibicionista actúa sobre la capa de identidad y registro. Fija un umbral etario por debajo del cual se prohíbe la creación de cuentas y traslada el peso del cumplimiento al sistema de verificación de edad. Australia constituye un caso paradigmático. Su lógica subyacente es la protección frente a riesgos y se apoya, en la Convención, principalmente en los artículos 3, 19, 34 y 36.

⁴² KEMELMAJER DE CARLUCCI, Aída, "La autonomía progresiva del niño y el adolescente. ¿Hasta cuándo?", en Revista de Derecho de Familia y de las Personas, La Ley, Buenos Aires, año IV, n.º 1, 2012, p. 3.

⁴³ Al momento, no entró en vigencia y se encuentra en etapa de suscripción, con Uruguay como primer firmante latinoamericano (2 de septiembre de 2025).

El modelo de obligaciones de diseño opera sobre múltiples capas simultáneamente (algoritmos, datos, interfaces, moderación) y desplaza el peso regulatorio hacia las plataformas. Les exige deberes de prevención, mitigación y rendición de cuentas conforme al principio de seguridad desde el diseño. En clave de Convención, atiende un espectro más amplio de derechos: junto con los deberes de protección, incorpora los artículos 5, 13, 15, 16 y 17.

El modelo híbrido combina ambas lógicas. Establece umbrales etarios o reglas de consentimiento parental y, en simultáneo, impone obligaciones operativas a las plataformas.

Estas tres familias funcionan como clave de lectura de todo lo que sigue. Los riesgos que el próximo apartado describe son aquello que los modelos buscan conjurar; las medidas que adopta cada país son variantes de estas tres familias; y la evaluación desde la Convención, que haremos en el apartado 3.8, mide a cada familia frente al estándar de proporcionalidad enunciado más arriba.

3.5. Los riesgos sobre los que opera la regulación

Como aclaramos al inicio de este capítulo, antes de examinar qué hicieron los Estados, conviene precisar los riesgos concretos que el entorno digital plantea a los NNYA. Porque sin ese mapa, la evaluación de los modelos regulatorios se torna muy abstracta, porque un modelo sólo puede juzgarse respecto del riesgo que pretende atacar.

Por eso, en el Capítulo 2 de este libro describimos en detalle los mecanismos técnicos que producen esos riesgos, como los sistemas de recomendación, el diseño persuasivo, los dark patterns, el tratamiento de datos y las fallas de los modelos conversacionales, entre otros. Sin reiterarlos, vamos a ordenarlos en una taxonomía que nos permita mapear cada modelo regulatorio en función del tipo de riesgo que efectivamente alcanza.

3.5.1. Una taxonomía de riesgos: contenido, contacto, conducta

Existe una sistematización de riesgos del entorno digital para NNYA en tres grandes categorías, que ofrecen un criterio ordenador útil⁴⁴.

Contenido

Exposición a material inapropiado: violento, sexual, que promueve trastornos alimentarios o autolesiones, publicidad dirigida y explotación comercial. El NNYA es receptor.

Los **riesgos de contenido** se refieren a la exposición del NNYA a material inapropiado para su edad o su desarrollo, como el contenido violento, sexual, que promueve trastornos alimentarios o autolesiones, publicidad dirigida y prácticas de explotación comercial. Aquí el NNYA es, ante todo, receptor de algo que el sistema le muestra.

⁴⁴ La sistematización en riesgos de contenido, contacto y conducta puede verse con mayor profundidad en trabajos de Sonia Livingstone y el proyecto EU Kids Online. Al respecto, Livingstone, S. y Stoilova, M. (2021), "The 4Cs: Classifying Online Risks to Children", Short Report Series.

Contacto

Interacciones iniciadas por terceros que buscan dañar: grooming, captación con fines de abuso, uso indebido de sus datos. El NNyA es objeto de la acción de otro.

Los **riesgos de contacto** se refieren a las interacciones iniciadas por terceros que buscan dañar al NNyA como el grooming, la captación con fines de abuso, el uso indebido de sus datos personales por parte de quien se relaciona con él. Aquí el NNyA es objeto de la acción de otro.

Conducta

Comportamiento del propio NNyA o de sus pares: ciberacoso entre iguales, difusión de contenido íntimo, dinámicas de riesgo. El NNyA es agente.

Los **riesgos de conducta** se refieren al comportamiento del propio NNyA o de sus pares, que puede generar daño como el ciberacoso entre iguales, la producción y difusión de contenido íntimo, la participación en dinámicas de riesgo. Aquí el NNyA es agente, no solo receptor u objeto.

Estas tres categorías no son estancas, pero nos permiten advertir que un modelo puede ser eficaz frente a una categoría e inócuo o ineficaz frente a otra. Por ejemplo, la restricción de acceso puede mitigar ciertos riesgos de contacto sin abordar los de contenido que se producen una vez dentro de la plataforma.

Los riesgos recaen, además, con intensidad desigual. Pesan con especial fuerza sobre los NNyA de contextos socioeconómicamente vulnerables, los que están fuera del sistema escolar y aquellos con discapacidad, cuya situación amplifica tanto la exposición como las consecuencias⁴⁵.

3.5.2. El riesgo de las plataformas y el diseño orientado a la captura de la atención

El primer foco de riesgo proviene de las plataformas de redes sociales y, en particular, de sus sistemas de recomendación y de su diseño orientado a maximizar el tiempo de permanencia.

Como desarrollamos en el Capítulo 2, estos sistemas operan sobre un sistema de recompensa que presenta una sensibilidad incrementada y un control inhibitorio, lo que vuelve a los NNyA particularmente susceptibles a las técnicas de captura de la atención.

La evidencia es elocuente. Un estudio de la fundación 5Rights mostró cómo cuentas creadas (como si pertenecieran a NNyA) fueron expuestas, en plazos breves, a contenido sobre trastornos alimentarios, autolesión, suicidio y cuerpos sexualizados⁴⁶.

Podemos ver que esto no es accidental, sino que es el resultado previsible de un diseño que optimiza el engagement sin distinguir la edad del usuario.

⁴⁵ Moliterno Abi Cheble, L., "El Derecho Digital en el Proyecto del Código Civil brasileño: hacia el reconocimiento de la vulnerabilidad digital y la protección de NNyA", Sup. Innovación y Derecho, abril 2025.

⁴⁶ 5Rights Foundation (2021), Pathways: How digital design puts children at risk. <https://5rightsfoundation.com/resource/pathways-how-digital-design-puts-children-at-risk/>

Pensemos en una escena concreta. Un adolescente de trece años abre TikTok a las seis de la tarde para ver un video antes de cenar. No necesita buscar nada. El contenido empieza solo. Cada video dura entre quince y sesenta segundos, y el algoritmo registra en tiempo real cuánto tiempo le dedicó a cada uno, si lo volvió a ver, en qué segundo lo abandonó. Con eso ajusta lo siguiente. A los treinta minutos, sin que nadie lo llame, ese adolescente habrá visto entre treinta y ciento veinte videos distintos y habrá recibido decenas de pequeñas recompensas.



Ese mecanismo tiene nombre, la ciencia del comportamiento lo llama refuerzo variable, el mismo principio que hace impredecible la recompensa de las máquinas tragamonedas y que, por eso mismo, las hace tan difíciles de abandonar.

Un estudio reciente lo denomina dopamine-scrolling y lo define como la búsqueda habitual de contenido novedoso impulsada por pequeñas dosis de dopamina liberadas con cada desplazamiento⁴⁷.

Lo que cambia cuando el usuario es un adolescente no es el mecanismo sino el efecto. La corteza prefrontal, la región del cerebro que regula el control de impulsos, todavía no terminó de desarrollarse⁴⁸.

Los estudios de neuroimagen lo confirman. En adolescentes con uso intensivo de redes sociales se observan cambios estructurales en el cerebro. La corteza prefrontal, que no termina de desarrollarse hasta bien entrada la tercera década de vida, muestra reducción de materia gris y menor actividad en las tareas de control inhibitorio. La amígdala, vinculada a la regulación emocional, registra alteraciones asociadas a mayor impulsividad. El circuito de recompensa dopaminérgico, el mismo que se activa en las adicciones a sustancias, aparece sobreestimulado. Una revisión reciente publicada en la revista *Cureus* describe el resultado como un ciclo retroalimentado, en el cual el diseño adapta el algoritmo al usuario, y el usuario adapta su sistema nervioso al diseño⁴⁹. Para los adultos, ese ciclo opera sobre un cerebro ya formado pero para los adolescentes, opera sobre uno que todavía está construyendo sus propios mecanismos de freno.

Un equipo de investigadores creó cuentas falsas de menores de edad en TikTok y midió qué pasaba. El algoritmo perfiló esas cuentas con una intensidad de entre cinco y ocho veces mayor que la que aplica sobre adultos para recomendarles contenido comercial.

Parte de ese contenido no estaba etiquetado como publicidad y aparecía bajo la forma de videos espontáneos de creadores de contenido. De esa forma, la plataforma cumple formalmente con el Reglamento de Servicios Digitales europeo, pero la auditoría muestra que en los canales regulados que no están regulados por ese Reglamento, el sistema sigue perfilando NNyA sin restricciones⁵⁰.

⁴⁷ SHARPE, B.T. y Spooner, R.A. (2025). «Dopamine-scrolling: a modern public health challenge requiring urgent attention». *Perspectives in Public Health*, vol. 145, n.º 4, pp. 190-191. PMC12322333.

⁴⁸ HEAD, K.R. (2025). «Short-form Video Use and Sustained Attention: A Narrative Review (2019–2025)». *International Journal of Computers, Electronics and Social Applications (IJCESA)*, vol. 2, n.º 4.

⁴⁹ De, D.; El Jamal, M.; Aydemir, E. y Khera, A. (2025). «Social Media Algorithms and Teen Addiction: Neurophysiological Impact and Ethical Considerations». *Cureus*, vol. 17, n.º 1. PMC11804976. El estudio refiere específicamente a adolescentes de entre 13 y 17 años; la evidencia neuroimagenológica en niñas y niños menores de esa franja etaria requiere fuentes adicionales.

⁵⁰ SOLAROVA, M. et al. (2026). «The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok». *ACM FAccT'26*. arXiv 2603.05653.

Ese mecanismo tiene nombre, la ciencia del comportamiento lo llama refuerzo variable, el mismo principio que hace impredecible la recompensa de las máquinas tragamonedas y que, por eso mismo, las hace tan difíciles de abandonar.

Un estudio reciente lo denomina dopamine-scrolling y lo define como la búsqueda habitual de contenido novedoso impulsada por pequeñas dosis de dopamina liberadas con cada desplazamiento⁴⁷.

Lo que cambia cuando el usuario es un adolescente no es el mecanismo sino el efecto. La corteza prefrontal, la región del cerebro que regula el control de impulsos, todavía no terminó de desarrollarse⁴⁸.

Los estudios de neuroimagen lo confirman. En adolescentes con uso intensivo de redes sociales se observan cambios estructurales en el cerebro. La corteza prefrontal, que no termina de desarrollarse hasta bien entrada la tercera década de vida, muestra reducción de materia gris y menor actividad en las tareas de control inhibitorio. La amígdala, vinculada a la regulación emocional, registra alteraciones asociadas a mayor impulsividad. El circuito de recompensa dopaminérgico, el mismo que se activa en las adicciones a sustancias, aparece sobreestimulado. Una revisión reciente publicada en la revista *Cureus* describe el resultado como un ciclo retroalimentado, en el cual el diseño adapta el algoritmo al usuario, y el usuario adapta su sistema nervioso al diseño⁴⁹. Para los adultos, ese ciclo opera sobre un cerebro ya formado pero para los adolescentes, opera sobre uno que todavía está construyendo sus propios mecanismos de freno.

Un equipo de investigadores creó cuentas falsas de menores de edad en TikTok y midió qué pasaba. El algoritmo perfiló esas cuentas con una intensidad de entre cinco y ocho veces mayor que la que aplica sobre adultos para recomendarles contenido comercial.

Parte de ese contenido no estaba etiquetado como publicidad y aparecía bajo la forma de videos espontáneos de creadores de contenido. De esa forma, la plataforma cumple formalmente con el Reglamento de Servicios Digitales europeo, pero la auditoría muestra que en los canales regulados que no están regulados por ese Reglamento, el sistema sigue perfilando NNyA sin restricciones⁵⁰.

Una investigadora le preguntó a un grupo de niñas y niños de once y doce años si las aplicaciones que usaban los manipulaban. Respondieron que sí, que ellos lo sabían, que no eran tan ingenuos. Después les mostró imágenes de pantallas reales. La cuenta regresiva que dice que una oferta vence en diez minutos. El botón de cancelar, en gris diminuto, casi invisible. La pantalla se oscurece salvo por la opción que la plataforma quiere que elijan. Les preguntó si los habían visto antes. Los habían visto. Les preguntó si los habían seguido. También. El estudio concluyó que saber que existe la trampa no equivale a poder esquivarla. Las niñas y los niños reconocen el patrón y lo siguen igual⁵¹. Este hallazgo tiene una consecuencia directa para el debate regulatorio porque enseñar a reconocer el diseño manipulatorio no alcanza.

Por eso, en términos de la taxonomía anterior, es un riesgo de contenido producido por la arquitectura misma de la plataforma.

⁴⁷ SHARPE, B.T. y Spooner, R.A. (2025). «Dopamine-scrolling: a modern public health challenge requiring urgent attention». *Perspectives in Public Health*, vol. 145, n.º 4, pp. 190-191. PMC12322333.

⁴⁸ HEAD, K.R. (2025). «Short-form Video Use and Sustained Attention: A Narrative Review (2019–2025)». *International Journal of Computers, Electronics and Social Applications (IJCESA)*, vol. 2, n.º 4.

⁴⁹ De, D.; El Jamal, M.; Aydemir, E. y Khera, A. (2025). «Social Media Algorithms and Teen Addiction: Neurophysiological Impact and Ethical Considerations». *Cureus*, vol. 17, n.º 1. PMC11804976. El estudio refiere específicamente a adolescentes de entre 13 y 17 años; la evidencia neuroimagenológica en niñas y niños menores de esa franja etaria requiere fuentes adicionales.

⁵⁰ SOLAROVA, M. et al. (2026). «The DSA's Blind Spot: Algorithmic Audit of Advertising and Minor Profiling on TikTok». *ACM FAccT'26*. arXiv 2603.05653.

⁵¹ RENAUD, K. et al. (2024). «'We're Not That Gullible!' Revealing Dark Pattern Mental Models of 11-12-Year-Old Scottish Children». *ACM Transactions on Computer-Human Interaction*, vol. 31, n.º 3, Art. 33. DOI 10.1145/3660342.

3.5.3. El riesgo de la IA conversacional y el vínculo simulado

El segundo foco es cualitativamente distinto, si consideramos la distinción entre capas de las tecnologías. A diferencia de la plataforma, que muestra contenido, el sistema conversacional puede generar un vínculo por parte del usuario, en tanto la IA conversacional como ChatGPT, Claude, Gemini y otras plataformas diseñadas exclusivamente para fines de compañía emocional, responden, recuerdan, adaptan su tono e interpelan emocionalmente al NNyA.

Ahora bien, el riesgo que conlleva el uso de la IA conversacional no encaja perfecto en ninguna de las tres categorías clásicas. Porque no es solo contenido y el daño no está en un material que se muestra; no es contacto en sentido tradicional, porque del otro lado no hay una persona humana, y tampoco se encuadra en la categoría de conducta, porque el NNyA no es el agente del daño. Son riesgos y daños que se producen por la dinámica del intercambio y en la antropomorfización del vínculo bajo una lógica de sostener la conversación el mayor tiempo posible.

La caracterización del vínculo simulado como riesgo autónomo no es una hipótesis teórica, hay una investigación de la Universidad de Edimburgo realizado con más de cien niñas y niños de entre seis y once años que interactuaban regularmente con asistentes conversacionales de voz en sus hogares, que reveló que la mayoría sobreestima la inteligencia de la IA, no comprende los mecanismos de almacenamiento de datos de sus conversaciones y, en un porcentaje significativo, considera moralmente incorrecto ser "malos" con el asistente. El 72% de los niños y niñas declaró ser sistemáticamente educado con Alexa, y apenas el 9% consideró aceptable tratarla mal. El resultado interesante que arroja es que los NNyA no interactúan con estos sistemas como si fueran software, sino que los percibían como agentes con sensibilidad y moral propia.^{52 53}

72%

de niñas y niños declaró ser sistemáticamente educado con el asistente de voz Alexa.

9%

consideró aceptable tratarla mal: lo perciben como un agente con sensibilidad propia.

Por su parte, otra investigación de la Universidad de Denver con niñas y niños de cinco y seis años que interactuaban con un chatbot de IA durante sesiones de narrativa colaborativa evidenció activación significativa de la zona del cerebro asociada al procesamiento social y la mentalización. El dato más relevante es que a mayor activación de esa zona durante la sesión genera la percepción de un vínculo y activa los mismos circuitos cerebrales que median las relaciones interpersonales, con efectos emocionales medibles en la primera infancia.^{54 55}

⁵² Poonsiriwong, Rachel; Archiwaranguprok, Chayapatr y Pataranutaporn, Pat, "Death' of a Chatbot: Investigating and Designing Toward Psychologically Safe Endings for Human-AI Relationships", MIT Media Lab, 2025 (en prensa).

⁵³ Andries, Valentina y Robertson, Judy, "Alexa doesn't have that many feelings: Children's understanding of AI through interactions with smart speakers in their homes", Computers and Education: Artificial Intelligence, vol. 5, 2023, art. 100176.

⁵⁴ Dangol, Aayushi; Gupta, Meghna; Yoo, Daeun et al., "Toys that listen, talk, and play: Understanding Children's Sensemaking and Interactions with AI Toys", University of Washington, CHI 2026.

⁵⁵ Kim, Pilyoung; Chin, Jenna H.; Xie, Yun; Brady, Nolan; Yeh, Tom y Yang, Sujin, "Young children's anthropomorphism of an AI chatbot: Brain activation and the role of parent co-presence", University of Denver / University of Colorado Boulder, 2025 (en prensa).

En cuanto a los adolescentes, los datos son igualmente contundentes. El informe de Common Sense Media del año 2025, documentó que plataformas como Character.AI son utilizadas por los y las adolescentes como fuentes de apoyo emocional, compañía y confidencialidad. Esas plataformas operan con sistemas de verificación de edad basados en autodeclaración. La misma organización evaluó esas plataformas y las calificó como de riesgo inaceptable para usuarios menores de dieciocho años⁵⁶.

Un riesgo adicional, aún menos atendido por la regulación, es el que surge cuando ese vínculo se interrumpe de manera abrupta. Investigadores del MIT Media Lab analizaron comunidades de usuarios en plataformas digitales para comprender el impacto psicológico de la discontinuación de chatbots, ya sea por actualizaciones del modelo, intervenciones de seguridad o cierre de plataformas, y encontraron que los usuarios describen esas experiencias como pérdida de un amigo, con procesos de duelo comparables a los que genera la pérdida de una relación. De esa forma, si lo trasladamos a los NNyA, cuya capacidad de procesamiento emocional es más frágil y cuyo tiempo de exposición acumulada puede ser mayor, el riesgo de daño psicológico por discontinuación es una variable que debería ser considerada por los modelos regulatorios.

La suma de estas evidencias nos permite considerar que la obligación regulatoria en materia de IA conversacional no se agota en el control del contenido generado ni en el bloqueo del acceso por umbral etario, sino que también debe comprender la forma en que el vínculo se constituye, se sostiene y, eventualmente, se interrumpe. Si el diseño de un sistema puede generar apego en un niño de cinco años con efectos medibles en su actividad cerebral, y si la discontinuación abrupta de ese sistema puede producir duelo clínico en usuarios adolescentes, entonces la regulación debe alcanzar el diseño mismo de esos vínculos y también los riesgos de los medios sintéticos.

El tercer foco proviene de los medios sintéticos, y dentro de ellos del fenómeno más reciente y de contornos menos explorados que denominamos como "clonación digital".

Conviene precisar de qué hablamos, porque el objeto desbordó la categoría con la que solía analizárselo. El deepfake, tal como lo tratamos en el Capítulo 2, designa un contenido audiovisual puntual, mediante una imagen o un video que representa falsamente a una persona en una situación determinada.

Ahora bien, la clonación digital opera en un nivel distinto porque no produce una pieza aislada, sino una réplica funcional de los atributos identificables de una persona, como su rostro, su voz, sus gestos, sus patrones de conducta, capaz de actuar, comunicarse y presentarse como una extensión suya.

Esos rasgos nos permiten advertir que estamos frente a un fenómeno distinto al deepfake, ya que ofrece persistencia y la réplica opera en el tiempo y no se agota en un contenido; tiene dinamismo, porque responde a contextos nuevos; adaptabilidad, porque reacciona a estímulos no previstos; y verosimilitud, porque resulta creíble frente a terceros.

En este sentido, la clonación digital interpela la identidad, la autonomía y la dignidad de quien es replicado⁵⁷.

⁵⁶ Robb, Michael B. y Mann, Supreet, Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions, Common Sense Media, San Francisco, 2025.

⁵⁷ METHUKU, V., "Dobles digitales: implicaciones de los clones de IA premortem", Universidad de Texas en Austin.

El riesgo para las NNYA se agrava por dos posibles razones:

1º

Es la disponibilidad del insumo. La clonación se realiza a partir de obtener datos biométricos como la imagen facial, voz, gestos, entendidos como aquellos datos referidos a las características físicas, fisiológicas o de conducta de una persona que permiten su identificación única⁵⁸.

2º

Es la caída de la barrera técnica. Es cada vez más fácil generar una réplica ya que no requiere conocimientos especializados y basta, en muchos casos, con subir una imagen o grabar un audio breve. La clonación digital reproduce rasgos faciales y corporales; la clonación vocal reconstruye el timbre y la entonación a partir de muestras de pocos segundos, con riesgos específicos de fraude, manipulación emocional y suplantación que explotan la confianza depositada en la voz.

Sobre esto, estudios sobre percepción humana muestran, además, que la exactitud promedio de las personas para distinguir un rostro o un audio sintético es prácticamente azarosa⁵⁹. De esta forma, la defensa basada en que "alguien se va a dar cuenta" ya no aplica.

El problema excede al derecho de infancia y toca una pregunta más amplia que es la de distinguir lo humano de su réplica. La multiplicación de presencias digitales indistinguibles convierte esa distinción en la primera condición de integridad de la interacción social.

Frente a ese horizonte se ensayan los llamados protocolos de humanidad (proof of humanness) como mecanismos destinados a acreditar que detrás de una cuenta o una acción hay una persona real y no una síntesis⁶⁰. Es, en cierto modo, un captcha de la humanidad.

Que una sociedad necesite diseñar una prueba semejante mide la magnitud del desplazamiento y sitúa a las infancias que crecerán dentro de ese entorno, sin haber conocido otro, en el centro de la pregunta regulatoria.

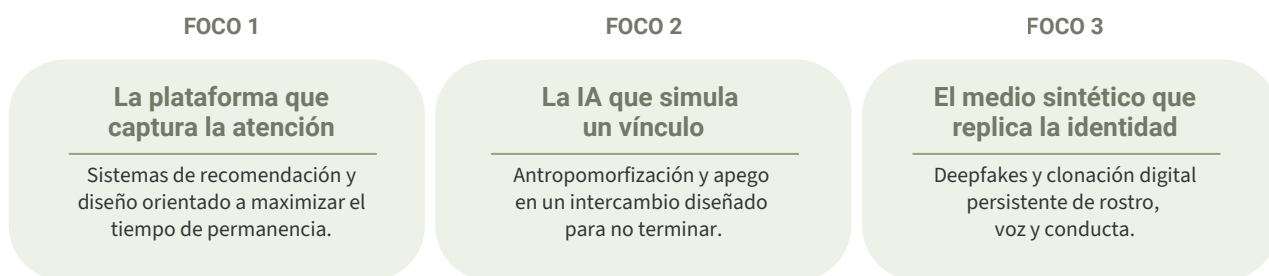
3.5.4. A modo de síntesis

Los tres focos de riesgo son, entonces, la plataforma que captura la atención, la IA que simula un vínculo y el medio sintético que replica la identidad. Ninguno de ellos es intercambiable y son enunciativos de los cambios dinámicos que los usos de estas tecnologías ofrecen. Asimismo, operan sobre capas tecnológicas distintas, comprometen derechos que pueden escindirse y exigen respuestas distintas.

⁵⁸ Definición formulada por la Real Academia Española.

⁵⁹ FRANK, J. et al., "A Representative Study on Human Detection of Artificially Generated Media Across Countries", arXiv (diciembre de 2023); y Nightingale, S. J. y Farid, H., "AI-synthesized faces are indistinguishable from real faces and more trustworthy", PNAS (2022).

⁶⁰ Sobre los protocolos de prueba de humanidad y el marco "conocer a tu agente" (know your agent), véase Sumsb, "Know Your Agent", disponible en <https://sumsub.com/blog/know-your-agent> (visitado el 12/6/2026).



De allí se sigue la consecuencia que ordena el resto de este capítulo. Una regulación no puede evaluarse en abstracto, sino frente al foco de riesgo que pretende alcanzar. Un modelo que se enfoca sobre el riesgo de las plataformas puede ser ciego frente al de la IA conversacional; uno que responde al deepfake puntual puede no prevenir la clonación persistente.

Por todo esto, tenemos la siguiente hoja de ruta: 1) en el apartado 3.6 describimos las medidas o iniciativas adoptadas por los países a nivel global; 2) en el apartado 3.7 abordamos el caso argentino; 3) en el apartado 3.8 evaluamos lo que surge de la Convención y sentaremos las bases sobre qué riesgos están previstos o normados y cuáles requieren regulación que los contemplen.

3.6. Regulaciones e iniciativas legislativas a nivel mundial

Un número creciente de países ha legislado, o se encuentra legislando, sobre el acceso de NNYA a los entornos digitales.

En este apartado haremos algo distinto y, en lugar de recorrer una jurisdicción tras otra, nos abocaremos a la experiencia comparada por temas, porque es esa lectura transversal la que revela los patrones, las tensiones y los aprendizajes que importan para las decisiones.

La convergencia temática es notable y pese a sus diferencias, casi todos los marcos giran en torno a un puñado de decisiones comunes como dónde fijar el umbral de edad, cómo verificarlo, qué obligaciones imponer al diseño de las plataformas, cómo articular el rol parental y cómo enfrentar la elusión.

3.6.1. Umbral etario, rol parental y elusión en los modelos comparados

La primera decisión de todo modelo es el umbral etario, y es también donde más divergen.



Australia fijó la barrera más alta y nítida con prohibición de crear cuentas a personas menores de dieciséis años, sin excepción parental. Otras jurisdicciones optaron por umbrales más bajos o flexibles. Francia, en su proyecto, discutió los quince años con matices entre los quince y los dieciocho. España tramita elevar la edad de registro de catorce a dieciséis. Varios estados de los Estados Unidos se ubican en los catorce (Florida), con consentimiento parental en la franja siguiente.

La divergencia no es un detalle técnico y nos revela que el umbral se deriva de una decisión de política legislativa que cada Estado resuelve de modo diferente. Esa decisión será central en el análisis desde la Convención, porque un umbral fijo y uniforme tensiona el principio de autonomía progresiva, que supone capacidades variables según el desarrollo de cada NNyA.

La consecuencia asociada al umbral también varía. En el modelo restrictivo australiano, cruzar el umbral hacia abajo significa no poder crear cuenta. En los modelos híbridos el umbral no cierra el acceso, sino que activa un requisito de consentimiento o la vinculación con la cuenta de un adulto responsable.

3.6.2. La verificación de edad: desafíos de implementación

Fijar un umbral plantea de inmediato el problema de cómo verificar la edad sin vulnerar otros derechos. Es el punto donde casi todos los modelos muestran su fragilidad.

Los datos de implementación australianos son ilustrativos. El primer informe trimestral de 2026 del eSafety Commissioner documentó que varias plataformas permitían a menores de dieciséis reintentar los métodos de verificación hasta obtener un resultado de "16+", lo que revela que el mecanismo podía sortearse con facilidad. La efectividad del modelo entero depende, así, de una pieza técnica que todavía no funciona de manera confiable.

Las alternativas técnicas conllevan, a su vez, sus propios costos. La verificación mediante documento o técnicas de reconocimiento facial y ocular (biometría) son las más robustas, pero también las más invasivas. Se puede excluir a quienes carecen de documentación, como por ejemplo los niños migrantes o refugiados, o a quienes tienen dificultades para usar sistemas biométricos. Malasia avanza hacia un esquema de verificación de identidad electrónica anclado en el documento nacional; la Unión Europea explora métodos de garantía de edad que preserven la minimización de datos.

La tensión de fondo es estructural: los métodos más efectivos tensionan o lesionan las normas de protección de datos y los más respetuosos son los más sencillos de eludir.

El mecanismo que efectivamente impera, por ejemplo, en la mayoría de las plataformas de IA conversacional, es el autoreporte donde el usuario declara su edad al crear la cuenta. El informe de Common Sense Media al que hemos hecho alusión anteriormente documentó que algunas plataformas explícitamente dirigidas a usuarios mayores de dieciocho años dependen exclusivamente de ese mecanismo para restringir el acceso, y que el mismo es completamente ineficaz para impedir que NNyA accedan y usen estos sistemas.

De esa forma, la ineficacia del autoreporte adquiere dimensiones particulares en el segmento de la IA conversacional porque a diferencia de las redes sociales clásicas, donde el acceso indiferenciado implica exposición a contenido generado por terceros, en las plataformas de compañeros de IA el acceso indiferenciado implica que el sistema calibra la interacción para cada usuario, con memoria, personalización y respuestas adaptadas a su perfil, sin ninguna distinción según la edad del interlocutor.

La cuestión se complejiza cuando el objeto de análisis se desplaza de las redes sociales a los juguetes que utilizan IA generativa. Un estudio reciente de la Universidad de Washington sobre juguetes con IA destinados a niños de entre seis y once años reveló que estos dispositivos no contemplan ningún mecanismo de verificación etaria ni controles de privacidad comprensibles para sus usuarios.

Estos juguetes constituyen una categoría que los marcos regulatorios sobre verificación de edad no cubren, porque no son redes sociales, no son plataformas de contenido, pero generan interacciones conversacionales con NNYA de manera directa, personalizada y continua, por lo que representan un salto cualitativo en la forma en que la IA se introduce en la vida cotidiana de las primeras etapas de la infancia y que la regulación vigente tampoco alcanza a encuadrar.

Lo que emerge de este panorama es que la verificación de edad, tal como está concebida en los modelos regulatorios actuales, fue diseñada pensando en el acceso a redes sociales basadas en contenido de usuarios y no es proyectada para tecnologías que generan vínculos, que adaptan sus respuestas al perfil del interlocutor, o que se integran en objetos físicos de uso cotidiano. Por eso, es interesante repensar qué se está intentando verificar, porque no debiera ser simplemente la edad del usuario, sino la adecuación del diseño del sistema a las características de desarrollo de quien lo utiliza.

3.6.3. Protección por diseño

Frente al enfoque que actúa sobre el acceso, una familia de modelos actúa sobre el diseño del servicio. En lugar de preguntar quién entra, se enfoca en cómo está construido aquello a lo que se ingresa.

El Reino Unido es el caso paradigmático con su Online Safety Act de 2023, basado en el principio de safety by design que significa diseñar la plataforma, desde el origen, para reducir el riesgo de daño, en lugar de reaccionar una vez producido.

La Unión Europea, a través del DSA y sus Directrices de 2025, impone evaluaciones de riesgo específicas para menores de edad, configuraciones protectoras por defecto y restricciones a la publicidad perfilada. Corea del Sur, tras la derogación de su régimen prohibicionista de videojuegos, se orientó también hacia obligaciones de diseño sobre los sistemas de recomendación.

A esta arquitectura europea de obligaciones de diseño se sumó a partir de 2024 el Reglamento de Inteligencia Artificial (Reglamento UE 2024/1689), cuya entrada en vigor escalonada distingue prácticas prohibidas, sistemas de alto riesgo y obligaciones de transparencia. Para los entornos donde participan NNYA resultan especialmente significativas dos disposiciones del artículo 5, contenidas en sus apartados (1)(a) y (1)(b). La primera prohíbe los sistemas de IA que utilizan técnicas subliminales o manipuladoras, mientras que la segunda alcanza específicamente a los sistemas que exploten vulnerabilidades de una persona derivadas de su edad, discapacidad o situación social, con el objetivo o el efecto de distorsionar sustancialmente su comportamiento de un modo que provoque perjuicios considerables.⁶¹

⁶¹ Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial (Reglamento de Inteligencia Artificial). Las prácticas prohibidas del artículo 5 son aplicables desde el 2 de febrero de 2025, y las obligaciones para sistemas de alto riesgo, desde el 2 de agosto de 2026.

AUMENTO DEL USO DE VPN TRAS LA PROHIBICIÓN AUSTRALIANA

El Reglamento incorpora, además, obligaciones específicas de transparencia para los sistemas de IA que interactúan directamente con personas físicas, de manera tal que su artículo 50 exige que estas sean informadas de que están interactuando con un sistema artificial. A su vez, el Anexo III califica como de alto riesgo, entre otros, los sistemas de IA utilizados en el ámbito educativo para determinar el acceso a instituciones, evaluar resultados de aprendizaje o detectar comportamientos prohibidos durante exámenes⁶². La conjunción de prohibiciones explícitas, obligaciones de transparencia y categorización de riesgos configura el primer marco regulatorio internacional que articula en derecho positivo una respuesta a los riesgos específicos que los sistemas de IA generan en sujetos especialmente vulnerables, incluidos los NNA. La virtud de este enfoque es que ataca el mecanismo del daño donde opera el algoritmo, la notificación, el scroll infinito, y no solo la puerta de entrada. Su costo es que exige una capacidad regulatoria y de supervisión más densa y auditorías, evaluaciones de impacto, fiscalización técnica sostenida en el tiempo, lo que implica un gran desafío para los países, pero fundamentalmente para sus protagonistas y sujetos de derecho, los NNA.

Una tercera vía combina el umbral con la intervención de los adultos responsables. Brasil, con el ECA Digital, condiciona el acceso de los menores de dieciséis a la vinculación de su cuenta con la de un tutor. Varios proyectos latinoamericanos como en Colombia, Perú, Paraguay y otros europeos incorporan el consentimiento parental como bisagra entre la prohibición total y el acceso libre.

El modelo híbrido reparte la responsabilidad entre el Estado, la plataforma y la familia. Su atractivo es que evita el todo o nada. Pero su dificultad supone una capacidad de control parental que no todas las familias tienen, y traslada parte del peso de la protección a un consentimiento cuya validez, tratándose de la cesión de datos de un NNA, es discutible.

Ningún modelo opera en el vacío y los NNA responden a las restricciones, y esa respuesta es en sí misma un dato regulatorio.

El indicador más visible es el uso de redes privadas virtuales (VPN), que permiten sortear las restricciones geográficas.

+170 %

+400 %

+65 %

Tras la entrada en vigor de la prohibición australiana, distintas fuentes del sector reportaron aumentos pronunciados en el tráfico y las instalaciones de VPN en el país (VPNSuper registró un incremento del 170 %, Windscribe del 400 % y hide.me del 65 % en accesos desde Australia). El dato sugiere que una parte de la población objetivo accede de todos modos. La experiencia británica matiza el punto y datos de Ofcom e Internet Matters indicaron que, a fines de 2025, el uso de VPN no mostró un aumento específico entre niños, lo que sugiere que el efecto de elusión depende del diseño concreto de la medida y de su aceptación social.

⁶² El considerando 48 del Reglamento alude expresamente a los sistemas de IA destinados a ser utilizados por niños como una categoría que requiere especial atención en la evaluación de riesgos y en el diseño de salvaguardas.

La elusión tiene, además, una dimensión de equidad. Los NNyA de familias con más recursos y alfabetización digital eluden la restricción; los de familias más vulnerables quedan efectivamente excluidos. La medida puede así producir un efecto regresivo que protege menos a quienes más la necesitan y restringe más a quienes tienen menos herramientas para sortearla.

3.6.4. Aprendizaje sobre el tiempo: las trayectorias asiáticas

Las jurisdicciones de Asia-Pacífico aportan algo que las occidentales, más recientes, no pueden ofrecer todavía y es el aprendizaje sobre qué pasa con estas medidas a lo largo del tiempo.

El caso de Corea del Sur es el más instructivo. Fue pionera en restricciones horarias de acceso en materia de juego en línea y, una década después, derogó esa misma política por considerarla ineficaz e incompatible con la autodeterminación de los NNyA.

Los fundamentos oficiales de la derogación anticipan varios de los reparos que hoy se formulan al modelo prohibicionista. China, en el polo opuesto, sostiene el régimen más intervencionista con "modo menores", límites horarios y verificación de nombre real. Japón optó por un camino intermedio, sin umbrales ni horarios, basado en el filtrado por defecto y la alfabetización digital.

La lección de la trayectoria coreana es valiosa porque es empírica, y es un antecedente que ninguna evaluación seria puede ignorar.

3.7. El caso argentino

El recorrido por la experiencia comparada nos hace preguntar dónde se ubica nuestro país.

Antes de avanzar, aclaramos que no llevaremos adelante un relevamiento exhaustivo de toda la normativa argentina que, de manera directa o indirecta, incide sobre la vida digital de los NNyA. Dejamos fuera, en particular, las normas penales en materia de ciberdelitos, que merecen un tratamiento propio y exceden el objeto de este libro.

En concreto, proponemos mostrar algo más preciso y que adelanta el estado de situación. Argentina no cuenta con un abordaje normativo concreto y específico de la protección de las infancias en su relación con las plataformas digitales y sobre el uso de IA conversacional o agéntica, aunque hay normas que indirectamente podrían incidir en la problemática.

3.7.1. Un marco fragmentado

Como adelantamos, Argentina no tiene una ley nacional específica sobre NNyA y redes sociales e IA. El marco aplicable se encuentra fragmentado entre normas que fueron pensadas para otros contextos y que alcanzan el entorno digital solo de manera tangencial y con interpretaciones adaptadas a los contextos actuales.

EL MARCO ARGENTINO FRAGMENTADO

2000

Ley 25.326. Protección de Datos Personales; anterior a las redes sociales y sin disposiciones específicas sobre NNyA equivalentes al art. 8 del GDPR.

2005

Ley 26.061. Sistema de Protección Integral de Derechos de NNyA; no contempla el entorno digital.

2013

Ley 26.904. Tipificó el delito de grooming.

2020

Ley 27.590 ("Ley Mica Ortega"). Creó el Programa Nacional de Prevención y Concientización del Grooming.

2023

Ley 27.736 ("Ley Olimpia"). Incorporó la violencia digital como modalidad de violencia de género, modificando la Ley 26.485.

La **Ley N.º 25.326** de Protección de Datos Personales es anterior a la existencia misma de las plataformas de redes sociales tal como hoy las conocemos, y además carece de disposiciones específicas sobre protección de NNyA equivalentes al artículo 8 del Reglamento General de Protección de Datos europeo (GDPR).

La **Ley N.º 26.061** establece el Sistema de Protección Integral de Derechos de NNyA, pero no contempla el entorno digital.

La **Ley N.º 26.904** tipificó el delito de grooming. La **Ley N.º 27.590**, conocida como "Ley Mica Ortega", creó el Programa Nacional de Prevención y Concientización del Grooming.

La **Ley N.º 27.736**, o "Ley Olimpia", incorporó la violencia digital como modalidad de violencia de género, modificando la Ley 26.485.

El rasgo común de este conjunto normativo es que todas, cuando vinculan las cuestiones digitales, lo hacen desde la lógica penal o de la protección frente a la violencia y no desde la regulación del diseño, el acceso o el funcionamiento de las plataformas y los sistemas de IA.

3.7.2. La autodeterminación informativa como punto de apoyo

A nivel de iniciativas legislativas, se han presentado en el Congreso Nacional múltiples proyectos vinculados con la materia. Desde la regulación del uso de dispositivos en establecimientos educativos hasta programas de concientización sobre el uso de pantallas, pasando por propuestas de autenticación biométrica en plataformas y reformas al Código Penal para tipificar el hostigamiento digital.⁶³

El vacío de regulación específica no significa ausencia total de herramientas. La doctrina y la jurisprudencia argentinas han desarrollado estándares y una ingeniería de adaptación que, debidamente proyectados, ofrecen un punto de apoyo.

Como señalamos en el Capítulo 2, la Corte Suprema reafirmó recientemente la centralidad del principio de autodeterminación informativa, al sostener que no existe un interés legítimo que justifique el almacenamiento y la circulación de datos personales sin el conocimiento de quien los provee.

Esa doctrina es trasladable al perfilado que realizan las plataformas porque si el Estado no puede usar datos sin consentimiento del titular, la pretensión privada de inferir y monetizar perfiles, especialmente de NNyA, debe someterse a estándares aún más estrictos de minimización, finalidad y consentimiento informado. Es una base sobre la cual una futura regulación podría construir, pero que hoy opera de manera indirecta y por vía interpretativa, no como un marco específico y sus efectos se proyectan al caso puntual.

3.8. La Convención sobre los Derechos del Niño y los modelos regulatorios

En los capítulos anteriores vimos qué exige la Convención, qué riesgos enfrentan los NNyA en torno al uso de ciertas tecnologías y qué hicieron los distintos países, incluida la Argentina. Queda una pregunta por responder ¿es compatible el modelo restrictivo con lo que manda la Convención?

Para responder con orden, seguimos un razonamiento de tres pasos. Primero recordamos qué exige la Convención. Después resumimos qué hicieron los países y qué resultados obtuvieron.

El apartado 3.1 estableció el estándar; conviene recuperar ahora sus dos exigencias decisivas para evaluar el modelo prohibicionista.

La primera es el equilibrio del espectro completo de derechos. La Convención no protege un derecho aislado, sino un conjunto interdependiente e indivisible. El modelo prohibicionista busca proteger principalmente los artículos 19, 34 y 36 (protección contra el abuso, la explotación y el ciberbullying) y se apoya en el artículo 3.1, sobre el interés superior del niño. Son derechos centrales.

Pero la Convención reconoce, junto a ellos, la libertad de expresión (art. 13), la libertad de asociación (art. 15), el acceso a la información (art. 17), la privacidad (art. 16) y la autonomía progresiva (art. 5).

Una política que protege a los primeros restringiendo los segundos no satisface el estándar, sino que lo desequilibra. Como señala Livingstone, la verificación de edad puede ser una herramienta legítima si se diseña respetando la totalidad de los derechos del niño, pero no puede confundirse con una estrategia integral de protección⁶⁴.

⁶³ A modo de ejemplo, pueden mencionarse: Expediente 4174-D-2024 (Pichetto y García), "Sistema de Protección Integral en el Uso Saludable y Responsable de la Tecnología Digital para Niños, Niñas y Adolescentes"; Expediente 0073-D-2025 (Vivero), "Programa Educativo Nacional de Concientización sobre Adicciones a las Pantallas"; entre otros.

⁶⁴ LIVINGSTONE, S. (2024), "Realising children's rights in relation to the digital environment", LSE Department of Media and Communications.

La segunda exigencia es la sensibilidad a la autonomía progresiva. El artículo 5 reconoce que el ejercicio de los derechos del niño está condicionado por la evolución de sus facultades, que varían según el contexto y el desarrollo individual. De allí una dificultad para los umbrales etarios fijos.

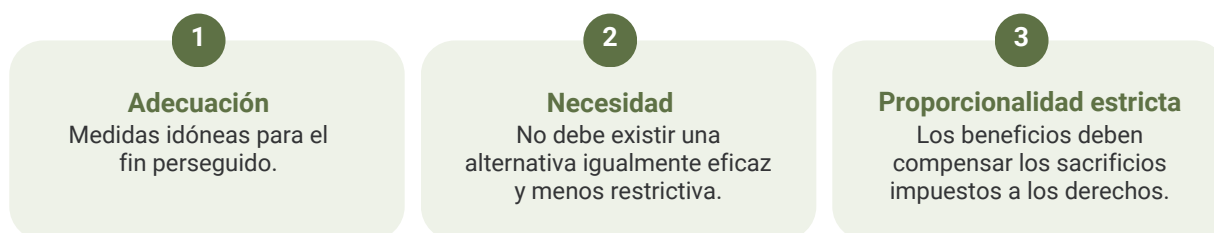
Como vimos en el apartado 3.6, los umbrales de trece, catorce, quince o dieciséis años no se derivan de evidencia científica consensuada, y su aplicación uniforme resulta cuestionable a la luz del artículo 5.

Livingstone y Sylwander sostienen que no existe una edad universalmente correcta para el acceso a las redes sociales, y proponen un enfoque de formas apropiadas para la edad, gradual y contextualizado⁶⁵.

En Argentina Rivera advirtió que el estándar "grado de madurez suficiente" puede tornarse confuso por la coexistencia de reglas que remiten al ordenamiento para determinar qué actos puede realizar la persona menor de edad⁶⁶. La tensión se proyecta sobre el entorno digital; por ejemplo, crear una cuenta, aceptar términos de uso o ceder datos son actividades que no encuentran respuesta unívoca sobre la edad y la capacidad requeridas.

A estas dos exigencias se suma el método. La Observación General N.º 25 manda que toda restricción de derechos de NNyA sea necesaria, proporcionada y no discriminatoria.

EL TEST DE PROPORCIONALIDAD (ALEXY)



Ese mandato se traduce en el test de proporcionalidad, conforme a la formulación de Robert Alexy ampliamente recibida en la doctrina constitucional⁶⁷, que admite tres subcriterios: adecuación entendida como medidas idóneas para el fin; necesidad vinculada a que no debe existir una alternativa igualmente eficaz y menos restrictiva; y proporcionalidad en sentido estricto donde los beneficios deben compensar los sacrificios que la medida impone a los derechos afectados.

De esa forma, los derechos fundamentales operan como principios que admiten distintos grados de realización, y sus colisiones se resuelven mediante ponderación, no mediante la invalidación de uno en favor de otro.

La premisa mayor queda así establecida, con una CDN que exige una protección que equilibre el espectro completo de derechos, sea sensible a la autonomía progresiva y supere el test de proporcionalidad.

⁶⁶ RIVERA, J. C., "Autonomía progresiva", en Tobías, J. W. y Sambrizzi, E. A. (dirs.), Homenaje a Guillermo A. Borda, Thomson Reuters, Buenos Aires, 2024, p. 487.

⁶⁷ ALEXY, R., Teoría de los derechos fundamentales, Centro de Estudios Políticos y Constitucionales, Madrid, 2007 (trad. de Carlos Bernal Pulido).

3.8.1. Qué puede decir la Convención sobre el modelo prohibicionista

En cuanto a la adecuación, los datos de cumplimiento sugieren que la medida no alcanza plenamente su objetivo ya que su idoneidad depende de una verificación de edad que todavía no opera de manera confiable.

En cuanto a la necesidad, existe una medida alternativa al menos tan eficaz para prevenir el daño y menos restrictiva de derechos: el modelo de obligaciones de diseño. Mientras el prohibicionismo actúa sobre el acceso y restringe en bloque la expresión, la asociación, la información y la participación, el modelo de diseño actúa sobre las condiciones que producen el daño sin clausurar esos derechos. Si existe esa alternativa, la restricción amplia que impone el prohibicionismo no resulta necesaria en el sentido técnico del test.

En cuanto a la proporcionalidad en sentido estricto, la ponderación es desfavorable. Los beneficios de la prohibición no están empíricamente consolidados, mientras que sus costos en términos de derechos restringidos son tangibles e inmediatos.

A ello se añade un dato que la doctrina argentina reciente ha incorporado: los estudios neurocientíficos documentan que el cerebro adolescente está aún en desarrollo, en particular en las zonas del juicio, la planificación y el control de impulsos⁶⁸. Ese dato no determina la elección de un modelo, pero introduce cautela y sugiere que la intervención debería dirigirse al diseño que explota esta vulnerabilidad antes que a la exclusión del adolescente del entorno.

La conclusión no es que el modelo prohibicionista sea inválido. Una restricción de acceso puede ser una pieza legítima dentro de un esquema más amplio, y la Convención no la prohíbe. La conclusión, más precisa, es que resulta insuficiente como eje central de la protección, no supera el test de necesidad cuando existe una alternativa menos lesiva, y tensiona el equilibrio de derechos y la autonomía progresiva que la Convención exige. La Convención no descarta la prohibición; la subordina. La admite como complemento, no como estrategia principal.

3.8.2. Una implicancia que la Convención no permite ignorar: la no discriminación

La conclusión anterior se refuerza con una consideración que el artículo 2 de la Convención vuelve ineludible. La prohibición puede producir efectos discriminatorios.

Como vimos en el apartado 3.6, los NNyA de familias con más recursos eluden la restricción mediante VPN, mientras que los de familias más vulnerables quedan efectivamente excluidos. Los sistemas de verificación basados en identificadores pueden, además, excluir a poblaciones sin documentación como niños refugiados o migrantes, y a niños con discapacidades que dificultan el uso de sistemas biométricos⁶⁹.

A esa asimetría interna se suma una asimetría global. En el contexto latinoamericano, Sousa Alves advierte que las grandes plataformas ofrecen términos de uso menos protectores en ciertas regiones que en Europa o Estados Unidos⁷⁰.

La pregunta, no es solo prohibir o regular, sino garantizar que el modelo adoptado se implemente con los mismos estándares en todas las jurisdicciones donde operan las plataformas. Un modelo que descansa solo en la restricción de acceso, sin obligaciones de diseño exigibles a la plataforma, deja intacta esa asimetría.

⁶⁸ STEINBERG, L. (2014), *Age of Opportunity: Lessons from the New Science of Adolescence*, Houghton Mifflin Harcourt.

⁶⁹ En consonancia con OECD/LEGAL/0389, los actores deben asegurar que ningún niño quede en riesgo de exclusión o discriminación por falta de acceso, de alfabetización digital o por el diseño del servicio.

⁷⁰ SOUSA ALVES, M. A., "La infancia y la adolescencia en el entorno digital: estudio comparativo de la protección en Brasil", *Revista Jurídica Universidad Nacional del Oeste*, 2025.

3.8.3. Del análisis a la propuesta

El razonamiento de este capítulo confirma, desde la Convención, lo que la experiencia comparada del apartado 3.6 ya sugería: la restricción de acceso, por sí sola, no protege con precisión. Y la dirección en que se mueve el derecho comparado lo refuerza desde un tercer ángulo, el del enforcement judicial.

**USD
375M**

El 24 de marzo de 2026, un jurado de Nuevo México ordenó a Meta Platforms pagar 375 millones de dólares por habilitar la explotación sexual de NNYA: la primera condena de una tecnológica global en un juicio por daños a NNYA. Los analistas la llamaron el "momento Big Tobacco" de las redes sociales.

El 24 de marzo de 2026, un jurado de Nuevo México ordenó a Meta Platforms el pago de 375 millones de dólares por haber habilitado la explotación sexual de NNYA en sus plataformas, en la primera condena de una tecnológica global en un juicio por daños a NNYA⁷¹.

Los analistas describieron el fallo como el "momento Big Tobacco" de las redes sociales: el punto en que el pasivo por daños a NNYA deja de ser un costo operativo asumido y se vuelve riesgo sistémico para el modelo de negocio⁷². Lo decisivo no es el monto, sino que el proceso puede imponer medidas que alteren el diseño de las plataformas⁷³.

El caso de Character.AI ilustra el mismo mecanismo desde otro ángulo. En enero de 2026, Google y Character.AI acordaron mediaciones en demandas vinculadas al suicidio de adolescentes tras interacciones con un chatbot⁷⁴.

Lo decisivo no es el acuerdo, sino lo que lo precedió: en octubre de 2025, antes de su conclusión, la empresa ya había prohibido a los menores de dieciocho años los chats "abiertos" con sus agentes de IA. El enforcement operó como catalizador de rediseño anticipatorio, antes de la condena y no como consecuencia de ella.

Es, exactamente, la lógica que un marco regulatorio orientado a la prevención busca institucionalizar, actuar sobre el diseño antes de que el daño se produzca, en lugar de prohibir el acceso o esperar la condena.

Por eso, en el apartado siguiente (4.5) traducimos los estándares de diseño centrado en la infancia en cinco ejes concretos que ordenan, en clave argentina, los criterios mínimos que surgen del diagnóstico.

⁷¹ CNBC, "Meta must pay \$375 million for violating New Mexico law in child exploitation case, jury rules", 24 de marzo de 2026.

⁷² MarketMinute/Financial Content, "Social Media's Big Tobacco Moment: Meta's \$375 Million New Mexico Verdict", 7 de abril de 2026.

⁷³ New Mexico Department of Justice (2026, 24 de marzo). New Mexico Department of Justice Wins Landmark Verdict Against Meta. El jurado impuso a Meta Platforms la multa máxima de USD 5.000 por cada violación a la Ley de Protección al Consumidor de Nuevo México. Fue la primera condena obtenida en juicio oral contra una empresa tecnológica de gran escala en materia de seguridad digital de NNYA en los EEUU.

⁷⁴ CNN Business, "Character.AI and Google agree to settle lawsuits over teen mental health harms and suicides", 7 de enero de 2026.

3.8.4. En síntesis

La segunda conclusión es que los riesgos son heterogéneos y no se dejan tratar con una respuesta única. La plataforma que captura la atención, la inteligencia artificial que simula un vínculo y el medio sintético que replica la identidad operan sobre capas distintas, comprometen derechos distintos y exigen respuestas distintas.

La taxonomía de riesgos de contenido, contacto y conducta, útil para las dos primeras, ni siquiera alcanza a nombrar del todo a las dos últimas: el riesgo de vínculo de la IA conversacional y el riesgo de identidad de la clonación digital son cualitativamente nuevos. Una regulación que no distingue entre estos focos no protege con precisión.

La tercera conclusión surge del recorrido por las respuestas estatales. Los Estados convergen en el diagnóstico, pero divergen en la respuesta, y esa divergencia se ordena en tres familias: la restricción del acceso, la obligación de diseño y la combinación de ambas.

Los datos de implementación y las trayectorias largas (la australiana, todavía abierta; la coreana, ya cerrada por derogación) sugieren una tendencia que el análisis desde la Convención confirma: las medidas que actúan solo sobre el acceso enfrentan problemas de efectividad, elusión y equidad, mientras que las que actúan sobre el diseño alcanzan el mecanismo del daño, aunque exijan mayor capacidad estatal.

La cuarta conclusión es el resultado del test de proporcionalidad. El modelo prohibicionista, examinado a la luz de la Convención, resulta insuficiente como eje central de la protección. No por inválido ya que una restricción de acceso puede integrar legítimamente un esquema más amplio, sino porque no supera la prueba de necesidad cuando existe una alternativa menos lesiva, y porque tensiona el equilibrio de derechos y la autonomía progresiva que la Convención exige. La Convención no descarta la prohibición: la subordina.

La quinta conclusión concierne a la Argentina. El país combina una paradoja productiva: principios sólidos e instrumentos concretos ausentes. Dispone de un sistema de protección integral robusto, de la Convención con jerarquía constitucional y de una doctrina consolidada sobre autonomía progresiva y autodeterminación informativa. Carece, en cambio, de una regulación específica que traduzca esos principios en obligaciones exigibles sobre las plataformas, la IA conversacional y los medios sintéticos. El problema argentino no es de fundamentos, sino de traducción.

De estas cinco conclusiones se desprende la orientación de la propuesta. Si el estándar exige equilibrar derechos, si los riesgos son heterogéneos, si el modelo de diseño protege con más precisión que la restricción de acceso y si la Argentina tiene los principios, pero le faltan los instrumentos, entonces la tarea regulatoria pendiente está clara. No consiste en elegir entre prohibir o no prohibir, sino en construir un marco que actúe sobre el diseño de las tecnologías, diferenciado según el foco de riesgo, anclado en los principios que el país ya reconoce y traducido en obligaciones concretas y exigibles.



CAPÍTULO 4

Inteligencia Artificial y entornos conversacionales

4.1. El diseño centrado en la infancia como estándar regulatorio

Toda evaluación de una política regulatoria requiere, primero, un estándar contra el cual medirla. Por eso, en este capítulo nos hemos propuesto volcar el marco normativo que opera como parámetro de validez de cualquier regulación del entorno digital dirigida a NNyA.

Desde esa lógica, la Convención sobre los Derechos del Niño como instrumento vinculante y rector, y los marcos internacionales de soft law que, sin crear obligaciones directamente exigibles, fijan principios de referencia relevantes en contextos dinámicos como los que vivimos con el impacto de la tecnología.

La Convención sobre los Derechos del Niño (CDN), adoptada por la Asamblea General de las Naciones Unidas el 20 de noviembre de 1989, constituye el tratado internacional de derechos humanos más ampliamente ratificado de la historia. Es un instrumento jurídico internacional vinculante que reconoce a las niñas, niños y adolescentes como sujetos plenos de derechos civiles, políticos, económicos, sociales, culturales y ambientales, y no como meros objetos de protección.

Si hay un desplazamiento conceptual que define la regulación de los entornos digitales en la actualidad, es el paso de una lógica de contenidos a una lógica de diseño.

Durante años, el debate regulatorio se organizó en torno a la pregunta de qué contenidos deberían prohibirse, restringirse o filtrarse cuando los usuarios son NNyA. Ahora bien, como vimos en el Capítulo 3 de este libro, ese enfoque es insuficiente, porque muchas veces los daños no provienen sólo del contenido que consumen, sino de la arquitectura del sistema que, en definitiva, organiza ese consumo.

Esto significa básicamente que, si el diseño de un producto digital puede generar daño al bienestar, la privacidad o el desarrollo de NNyA, entonces esa decisión de diseño puede ser evaluada con parámetros jurídicos, puede ser reglamentada y también atribuir responsabilidad en la cadena de toma de decisiones.

A ese desplazamiento se lo ha denominado "diseño centrado en la infancia" (child-centered design), y es la categoría que vamos a desarrollar en este capítulo.

4.1.1. Del diseño como opción técnica al diseño como obligación jurídica

El concepto de diseño en el contexto de los productos digitales tiene múltiples capas, al igual que vimos al desentrañar la estructura de las plataformas y sistemas de IA conversacionales.

En el sentido más lineal, diseñar es tomar decisiones sobre la apariencia, la estructura y el funcionamiento de una interfaz. Pero en el ecosistema de plataformas digitales, el diseño abarca decisiones que van mucho más allá de lo visual, y puede incluir qué datos se recopilan y cómo se utilizan, qué configuraciones están activas por defecto, qué algoritmos determinan qué contenido se recomienda, qué mecanismos de notificación se activan y con qué frecuencia, qué fricciones se interponen entre el usuario y el cierre de su cuenta, entre muchas otras.

Todas esas decisiones orientan la conducta del usuario y modelan la experiencia que se le ofrece. El problema es que, cuando ese usuario es un NNyA, esa orientación interactúa con características propias del desarrollo que lo vuelven especialmente vulnerable a sus efectos.

Pensemos en TikTok que lleva la lógica del scroll infinito a un formato de video a pantalla completa. La aplicación se abre directamente en una imagen o video que se reproduce solo, sin que el usuario haya elegido nada. Para pasar al siguiente alcanza con deslizar el dedo hacia arriba, y ese siguiente video ya empezó a asomar en el borde inferior de la pantalla. No hay un final, ni un punto donde el sistema se detenga y solicite si uno quiere seguir: el contenido se renueva de manera continua, de a un video por vez, y nunca se agota. Esa estructura no tiene el mismo impacto en un adulto, que puede decidir cerrar la aplicación, que en NNyA cuyos sistemas de control de impulsos todavía no completaron su maduración.

Esta dimensión del diseño fue conceptualizada con precisión por Don Norman, uno de los teóricos más influyentes de la interacción persona-producto. Según este autor, hay que distinguir entre prestación (affordance) y significativo (signifier)⁷⁵. La prestación es lo que el objeto permite hacer: en TikTok, la prestación central es que siempre se puede mirar un video más, porque el contenido no se agota nunca. El significativo es la señal perceptible que comunica esa posibilidad y cómo activarla, como el video siguiente que asoma en el borde de la tablet, el gesto de deslizar hacia arriba, la reproducción automática que arranca sin que uno la pida. El significativo le dice implícitamente al usuario "hay más, seguí".

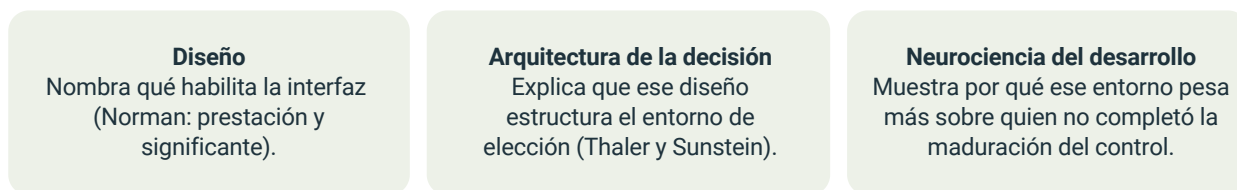
⁷⁵ NORMAN, D. A. (2013). *The design of everyday things* (Ed. rev. y ampliada). Basic Books. La distinción entre affordance y signifier se desarrolla especialmente en el capítulo 1.

Norman nos da el vocabulario para nombrar el mecanismo, pero la economía del comportamiento completa el enfoque. La estructura de una interfaz orienta las decisiones de quien la usa y por eso aquí se introduce el concepto de arquitectura de la decisión.

Richard Thaler y Cass Sunstein mostraron que el modo en que se presentan las opciones, los valores que vienen seleccionados por defecto, lo que resulta saliente o queda relegado, influye de manera sistemática sobre la conducta, incluso cuando el usuario conserva formalmente toda su libertad de elegir⁷⁶. El diseño de TikTok es, en estos términos, una arquitectura de decisión orientada hacia una reproducción automática por defecto. Seguir mirando es lo saliente y detenerse es la fricción que hay que vencer. No prohíbe cerrar la aplicación, pero inclina la balanza constantemente hacia continuar.

Ahora bien, la arquitectura de la decisión explica por qué la estructura del entorno influye sobre cualquier usuario, pero necesitamos abordar una pregunta más: ¿por qué influye más sobre un NNyA?

Esa diferencia tiene su anclaje en la neurociencia del desarrollo. Los modelos de sistemas duales describen una asincronía en la maduración cerebral durante la adolescencia, con el sistema vinculado a la recompensa y la búsqueda de novedad se desarrolla temprano. En cambio, la corteza prefrontal donde está el control inhibitorio y la regulación de impulsos, recién se completa la maduración cuando somos adultos⁷⁷. Una arquitectura de la decisión orientada a maximizar el tiempo de permanencia no actúa sobre un sujeto neutro, lo hace sobre un cerebro en el que el acelerador madura antes que el freno.



En síntesis, convergen tres marcos que provienen de disciplinas independientes: el diseño nombra qué habilita la interfaz, la arquitectura de la decisión explica que ese diseño opera estructurando el entorno de elección, y la neurociencia del desarrollo muestra por qué ese mismo entorno pesa más sobre quien todavía no completó la maduración del control.

Y todo esto nos lleva a un aspecto central: las prestaciones no son propiedades fijas del objeto, sino relaciones que varían según quién interactúa con él. El mismo botón de "reportar" o "bloquear" es una prestación real para un adulto: lo percibe, entiende qué hace y mide sus consecuencias, pero prácticamente no existe como posibilidad de acción para un niño que no lo distingue entre los demás íconos ni comprende para qué sirve.

A la inversa, la prestación de "mirar uno más" pesa muchísimo más sobre quien todavía no desarrolló la capacidad de frenar. Por eso el entorno digital no ofrece las mismas posibilidades de acción a un adulto que a un niño de diez años, a uno de trece o a un adolescente de diecisiete: la misma interfaz, frente a usuarios distintos, habilita conductas distintas.

⁷⁶ THALER, R. H. y Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press. El concepto de "arquitectura de la decisión" (choice architecture) es central en esta obra.

⁷⁷ STEINBERG, L. (2008). "A social neuroscience perspective on adolescent risk-taking". *Developmental Review*, 28(1), 78-106; y Casey, B. J., Jones, R. M. y Hare, T. A. (2008). "The adolescent brain". *Annals of the New York Academy of Sciences*, 1124, 111-126.

Lo anterior nos lleva al "diseño centrado en la infancia": si se orienta la conducta del usuario, y si los usuarios incluyen a NNYA con características psicológicas, cognitivas y evolutivas diferenciadas, entonces el diseño debe tomar esas características como parámetro y como obligación exigible y no como un aspecto optativo.

Esta traslación que ahora nos parece más evidente, no es tan nítida en nuestra legislación y requiere pautas propias. Por ejemplo, si pensamos en el derecho tradicional del consumidor, el mismo regula las condiciones de comercialización de un producto y no sus características de diseño. Por su parte, el derecho de la competencia puede sancionar las prácticas comerciales desleales, pero no obliga a diseñar de un modo determinado. Por último, en el caso del derecho de la infancia, aunque se establezcan obligaciones de protección amplias, no se han desarrollado herramientas para traducir esas obligaciones en requerimientos técnicos de diseño de productos digitales.

Por su parte, un elemento adicional que distingue al diseño centrado en la infancia de otras categorías regulatorias es su carácter *ex ante*. Aquí nuevamente operan las categorías tradicionales donde la mayoría de los instrumentos del derecho que proyectan la responsabilidad operan *ex post* y, por ende, sancionan el daño producido.

El paradigma del diseño centrado en la infancia establece condiciones que deben estar presentes en el diseño del producto antes de que sea lanzado al mercado. Esta diferencia temporal tiene consecuencias significativas para los NNYA, cuyos daños como su salud mental, el desarrollo cognitivo y la privacidad, entre otros, son frecuentemente irreversibles o difícilmente reparables por una indemnización posterior.

El derecho argentino tiene tradición en regulación *ex ante* en otros campos de la protección de la infancia como las regulaciones que protegen a los NNYA frente al trabajo de actores infantiles, los regímenes de habilitación de establecimientos educativos o las normas de seguridad de juguetes.

Por eso, el paradigma del diseño centrado en la infancia aplicado a las plataformas digitales es, en este sentido, una extensión coherente de esa tradición a un nuevo campo tecnológico.

4.1.2. Del "privacy by design" al "child rights by design"

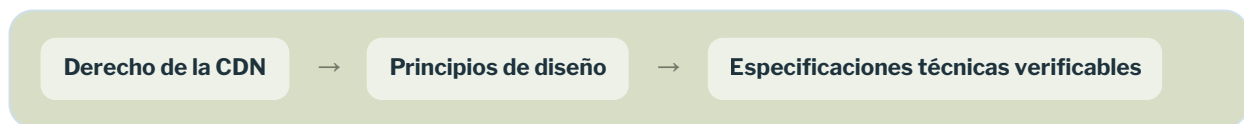
El antecedente más importante para la categoría que venimos analizando es el principio conocido como *privacy by design*.

A través de este, se explica que la privacidad no puede ser una característica que se agrega a un producto una vez diseñado, sino que debe estar incorporada como condición estructural.

El principio involucra una crítica al modo en que la industria tecnológica concebía la privacidad, ya que las configuraciones y la protección dependían de que cada usuario entienda las opciones disponibles y tome activamente las decisiones correctas, lo cual en la mayoría de los casos era toda una ficción porque nadie termina de leer y entender los términos y condiciones ni configurar las opciones adecuadas.

De esa forma, el privacy by design encontró una consagración regulatoria en el Reglamento General de Protección de Datos de la Unión Europea (RGPD, cuyo artículo 25 establece la obligación de protección de datos desde el diseño y por defecto), y de esa forma, se trasladó la carga de la privacidad del usuario individual hacia el diseñador del sistema⁷⁸.

La evolución de ese principio hacia el campo de la infancia, la podemos ver con claridad en trabajos de las especialistas Livingstone y Pothong, quienes realizaron un interesante trabajo sobre ese tema⁷⁹. Allí, han sostenido que los derechos reconocidos en la CDN pueden traducirse en especificaciones técnicas concretas durante el proceso de diseño de productos y servicios digitales, mediante un proceso que busque identificar qué características del diseño son incompatibles con los derechos de NNyA y cuáles son necesarias para garantizarlos.



Lógicamente, esa traslación representa una dificultad metodológica relevante, pero forma parte de un paradigma necesario en materia de protección de los derechos de infancia en los entornos digitales, porque si bien los derechos de la CDN son normas jurídicas formuladas en lenguaje de obligaciones, la interdisciplina necesaria en tiempos de revolución digital nos permite, mediante un proceso de operacionalización, identificar los derechos relevantes, derivar de cada uno un conjunto de principios de diseño y, de cada principio, especificaciones técnicas verificables.

De esa forma, el resultado podría ser un instrumento cuyos principios como el interés superior del niño sean utilizados como principios rectores del proceso de diseño.

4.1.3. Los paradigmas del diseño protector: AADC, Privacy by Design y Safety by Design

A través del Age Appropriate Design Code (en adelante, AADC), adoptado en 2020 por el Information Commissioner's Office (ICO) del Reino Unido, se habilitó un código de prácticas específico para servicios de la sociedad de la información que probablemente sean accedidos por NNyA.

La elección de la palabra "probablemente" es jurídicamente significativa, ya que no se limita a los servicios expresamente dirigidos a NNyA, sino que se extiende a cualquier servicio que, razonablemente considerado, pueda ser accedido por los mismos. De esa forma, el criterio no es la intención del diseñador, sino la probabilidad razonable de acceso por parte de los niños.

Esta definición nos permite pensar que otorga claridad en un plano complejo donde las plataformas suelen eludir la regulación destinada a protección de infancias, arguyendo que sus servicios no están "dirigidos" a ellos.

Por su parte, ese código establece estándares para el diseño centrado en la infancia, entre los que se pueden mencionar:

⁷⁸ Reglamento (UE) 2016/679 (Reglamento General de Protección de Datos), art. 25, "Protección de datos desde el diseño y por defecto".

⁷⁹ LIVINGSTONE, S. y POTHONG, K. (2021). Playful by Design: Free play in a digital world. Digital Futures Commission, 5Rights Foundation; y Child Rights by Design (2023).

1 Interés superior del niño. Las empresas deben "poner el interés superior del niño como una consideración principal en todo diseño y desarrollo de servicios probablemente accedidos por niños".

2 Configuración de privacidad por defecto. Las configuraciones que mejor protegen la privacidad de NNyA deben ser la predeterminada. El usuario, o sus responsables parentales, pueden modificarlas, pero la carga de ese cambio recae sobre quien quiere menor protección, no sobre quien quiere mayor protección. Este principio invierte la lógica habitual de los términos y condiciones que por defecto maximizan la recopilación de datos y requieren que el usuario actúe para protegerse.

3 Prohibición de nudges que debiliten la privacidad. El código prohíbe el uso de técnicas de diseño persuasivo (nudges) para llevar a los NNyA a proporcionar más datos personales de los necesarios o a desactivar protecciones de privacidad. Esta prohibición es directamente aplicable a los dark patterns que describimos.

4 Limitación de la recopilación y el uso de datos. Las empresas no pueden recopilar ni retener más datos que los estrictamente necesarios para el servicio que prestan.

5 Perfilado digital desactivado por defecto. El código exige que el perfilado digital de NNyA esté desactivado de forma predeterminada, sin perjuicio de que pueda habilitarse bajo condiciones estrictas que el propio ICO establece en su guía de implementación.

6 No recomendar contenido perjudicial. Obligación de que el sistema no recomiende contenido que sea perjudicial para el bienestar de los NNyA.

En 2025, la Fundación 5Rights lanzó el Children & AI Design Code, que extiende el marco del ICO y los trabajos de Livingstone, al ámbito específico de los productos y servicios de inteligencia artificial⁸⁰. Este trabajo identifica los riesgos particulares que presenta la IA para los derechos de NNyA, entre ellos: la opacidad de los sistemas de recomendación, la generación de relaciones de dependencia con agentes artificiales y el uso de datos de menores para entrenar modelos y propone estándares de diseño específicos para mitigarlos.

Por su parte, el concepto de Safety by Design comparte con el privacy by design la mirada de los derechos del niño centrados desde el diseño.

En este caso, este principio parte de la premisa de que la seguridad no puede ser una capa añadida sobre el producto, sino una propiedad estructural incorporada desde el diseño, centrándose en las características del producto y pensando ¿es este producto seguro por su diseño? ¿O solo puede hacerse seguro mediante restricciones de acceso?

La distinción entre los principios es relevante porque pone en evidencia una tensión en los modelos regulatorios actuales. Como vimos los sistemas de verificación de edad, que son la solución más extendida en la regulación comparada, funcionan con lógica de restricción de acceso de acuerdo con la edad del usuario.

⁸⁰ 5RIGHTS FOUNDATION (2025). Children & AI Design Code. Disponible en <https://5rightsfoundation.com> (visitado el 12/6/2026).

Ahora bien, esa solución no cambia el diseño del servicio; solo cambia quién puede acceder a él.

De esa forma, la seguridad del diseño exige que el servicio sea seguro para cualquier usuario, incluido un niño. Lo cual tiene sentido porque si un sistema de recomendación puede llevar a un adulto vulnerable a contenido de autolesión, la verificación de edad no resuelve el problema de fondo.

4.2. Dark patterns en entornos infantiles

Como desarrollamos en el apartado 2.3.4, los dark patterns son significantes deliberadamente engañosos, porque dan señales de diseño que comunican una cosa, pero producen otra. Veamos algunos ejemplos:

Botón "Aceptar todo"	Comunica equivalencia de opciones, pero orienta hacia una configuración que maximiza la extracción de datos del usuario.
"Cancelar suscripción"	Si logramos encontrarla, comunica una salida pero muchas veces produce un abandono por agotamiento.
"X personas vieron tu historia"	Comunica información social; pero también activa un circuito para generar retorno, produciendo a través del diseño una conducta.

Lo que hace de este marco un argumento jurídico es la consecuencia que se deriva de la premisa de Norman. En otras palabras, si las prestaciones son relaciones y no propiedades, el mismo diseño no tiene las mismas prestaciones para usuarios distintos. La plataforma opera sobre un adulto con modelos mentales consolidados, capacidad de razonamiento abstracto y corteza prefrontal en pleno funcionamiento. Opera sobre un niño de diez años sin ninguno de esos tres elementos. El mismo significantes produce efectos cualitativamente distintos sobre usuarios con capacidades cognitivas diferentes. Esta asimetría es estructuralmente predecible para quien diseña y también esa predictibilidad se proyecta sobre la responsabilidad.

Los dark patterns no son fallas de diseño accidentales; son el resultado de decisiones de diseño intencionales orientadas a favorecer los intereses de la empresa por sobre los del usuario.

No es suficiente con que el usuario pueda elegir, es el diseño mismo el que debe eliminar las técnicas que distorsionan. La prohibición de los dark patterns, a través de instrumentos, es una prohibición de diseño, no una obligación de información. Esto representa una ruptura respecto de la tradición regulatoria del consumidor, centrada en la transparencia y el consentimiento, en favor de una regulación centrada en la arquitectura de la decisión y el entorno.

Esta ruptura encuentra hoy expresión explícita en el derecho positivo europeo, donde el artículo 5, apartado (1) (b), del Reglamento de Inteligencia Artificial de la Unión Europea, vigente desde febrero de 2025 en lo relativo a las prácticas prohibidas, prohíbe la comercialización, puesta en servicio o uso de sistemas de IA que exploten vulnerabilidades de personas derivadas de su edad, con el efecto de distorsionar sustancialmente su comportamiento de modo que les cause perjuicios considerables⁸¹.

⁸¹ Reglamento (UE) 2024/1689, art. 5, apartado (1)(b). Las prácticas prohibidas son aplicables desde el 2 de febrero de 2025.

Es la primera vez que un marco regulatorio internacional vinculante prohíbe expresamente en derecho positivo el diseño orientado a explotar las particularidades cognitivas de los NNyA. Lo que esta disposición traduce al derecho es la premisa de Norman, en tanto si el diseño produce conductas calibradas para explotar vulnerabilidades específicas de la edad, entonces el diseño mismo, y no la información sobre él, es lo que la regulación debe alcanzar.

Los sistemas de recomendación algorítmica son la interfaz principal a través de la cual los NNyA consumen contenido digital. No hay un menú que los niños de hoy recorran para elegir qué ver o leer, hay un feed que genera el algoritmo, en tiempo real y de manera personalizada, en función de su historial de comportamiento.

El niño no elige, sino que es el propio algoritmo el que elige por él, optimizando la métrica que la empresa ha definido como objetivo.

Estos sistemas presentan serios problemas desde la perspectiva del diseño centrado en la infancia. Entre otros podemos señalar que los algoritmos optimizados para maximizar el tiempo de pantalla tienden a recomendar contenido que el usuario ya ha consumido en el pasado o que es similar a ese consumo y eso crea bucles de retroalimentación que refuerzan intereses y perspectivas existentes, lo que puede ser problemático para el desarrollo que la CDN busca proteger.

El segundo problema es la amplificación de contenido potencialmente dañino, ya que los algoritmos de recomendación pueden llevar a los usuarios hacia contenido de autolesión, desórdenes alimentarios u otras formas a través de una escalada gradual que no es perceptible.

El tercer problema es la invisibilidad del mecanismo, ya que los niños más chicos en general no comprenden que el contenido que ven refleja la selección que un algoritmo hizo para ellos. La ilusión de que el feed refleja "lo que es apropiado o les puede gustar" es parte del diseño y construye una percepción que parece orgánica, pero no lo es.

4.3. La analogía de la plaza: una obligación de diseño que ya conocemos

Una plaza diseñada para recibir NNyA de distintas edades no construye un único juego. Construye uno adaptado a las capacidades físicas, el control corporal y los umbrales de riesgo de los más chicos, y otro con otra escala, otra velocidad y otra arquitectura de seguridad para los más grandes.

Esta decisión, que ya tenemos internalizada como sociedad, nos permite comprender qué es una obligación de diseño en el plano analógico para, desde ahí, llevarla al ámbito de los entornos digitales, donde cada sistema rige con sus propias capas.

Pero hay algo más, y es que la plaza es, desde cierto enfoque, derecho positivo, porque los juegos infantiles están técnicamente normados de manera ex ante, con especificaciones diferenciadas por edad.

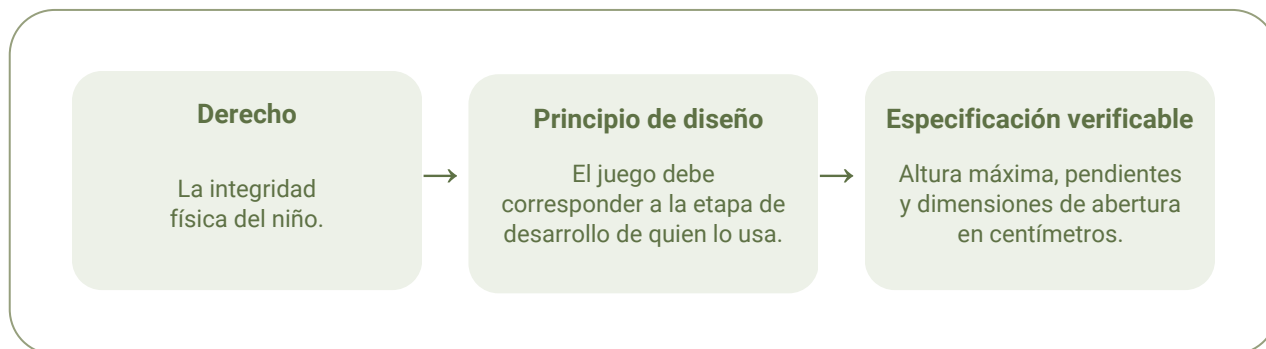
Por ejemplo, en Europa, las normas EN 1176 y EN 1177 fijan los requisitos del equipamiento de las áreas de juego y de los revestimientos de impacto, incluida la altura crítica de caída admisible para cada superficie⁸². En los Estados Unidos, la especificación ASTM F1487 y el manual de seguridad de la Consumer Product Safety Commission distinguen el equipamiento según franjas etarias y prescriben alturas, pendientes, separaciones y dimensiones de aberturas destinadas a cuidar a los chicos⁸³.

⁸² Normas EN 1176 (Equipamiento de las áreas de juego y superficies) y EN 1177 (Revestimientos de las superficies de las áreas de juego absorbedores de impactos), Comité Europeo de Normalización (CEN).

⁸³ ASTM International, F1487 — Standard Consumer Safety Performance Specification for Playground Equipment for Public Use; y U.S. Consumer Product Safety Commission, Public Playground Safety Handbook.

En nuestro país, eso no es excepción, las normas IRAM 3655 regulan los juegos infantiles de instalación permanente al aire libre, y una serie paralela, las IRAM 3761 a 3769, hace lo propio con los parques de diversiones y las atracciones electromecánicas, incluyendo una guía específica para su diseño y fabricación.

LA OPERACIONALIZACIÓN QUE YA EXISTE: LA PLAZA COMO DERECHO POSITIVO



Detengámonos un momento en lo que esto significa, porque es exactamente la idea que propone este libro para los entornos digitales. Hay un derecho, como podría ser la integridad física del niño, que fue traducido en principios de diseño (el juego debe corresponder a la etapa de desarrollo de quien lo usa) y, de cada principio, se derivaron especificaciones técnicas verificables, como tantos centímetros de altura máxima y dimensiones de abertura.

Este mismo razonamiento, es el proceso de operacionalización para los derechos de la Convención, solo que para toboganes y hamacas. Lo que se pide para las plataformas no es, entonces, un paradigma inédito, sino algo que el derecho ya exige para una plaza de barrio.

Retomando a Don Norman, y la relación entre el objeto y el usuario, el mismo juego de la plaza tiene prestaciones distintas para un cuerpo de cinco años y uno de quince, no porque cambien las reglas de uso, sino porque el entorno actúa sobre sistemas con capacidades diferentes; lo mismo cabe decir, desde una mirada interseccional, respecto de una persona con discapacidad.

Esa lógica, que nos parece tan clara en la vida cotidiana analógica, ha sido considerada por algunas iniciativas de la industria, como es el caso de YouTube Kids, que opera como una aplicación separada, con contenido para distintas franjas etarias, sin publicidad comportamental, con controles parentales integrados y sin los mecanismos de recomendación infinita que caracterizan la plataforma para adultos⁸⁴.

El parque de diversiones agrega a la analogía una pieza que nos faltaba. Quien haya llevado a un chico a uno recuerda el cartel que mide al niño antes de dejarlo subir. Esa barra es, traducida a nuestro vocabulario, como la verificación de edad. Pero a nadie se le ocurriría sostener que la barra, por sí sola, hace segura a la montaña rusa, en todo caso es segura porque sus arneses, sus velocidades y sus aceleraciones fueron diseñados y certificados conforme a normas técnicas.

Los modelos regulatorios que analizamos en el Capítulo 3 pueden releerse desde esta imagen, donde el modelo de prohibición por edad construye barras de altura cada vez más altas; el modelo de obligaciones de diseño exige que la atracción sea segura. Una regulación que se agota en la barra deja intacto el juego.

Corresponde, sin embargo, ser honestos con la analogía y reconocer sus límites, porque las tres diferencias que separan a la plaza del entorno digital, lejos de debilitarla, refuerzan la conclusión.

⁸⁴ YouTube Kids fue lanzado en 2015 como aplicación separada de YouTube para menores, con contenido curado, controles parentales y sin publicidad comportamental. En 2019, Google LLC y YouTube LLC llegaron a un acuerdo con la Federal Trade Commission y el Departamento de Justicia de EE.UU. por recopilación ilegal de datos de menores en violación de la Children's Online Privacy Protection Act (COPPA): United States of America v. Google LLC and YouTube LLC, Civil Action No. 19-cv-2642 (D.D.C. 2019).

1

El tobogán es estático; el feed es adaptativo. El tobogán no se rediseña en tiempo real según el comportamiento del chico; el sistema de recomendación hace exactamente eso. Las obligaciones de diseño digital deben alcanzar el comportamiento dinámico del sistema: qué optimiza, hacia dónde escala, cómo responde a las señales de quien lo usa.

2

La plaza no tiene un modelo de negocio construido sobre la permanencia. El tobogán no genera ingresos por cada hora que el chico pasa arriba; la plataforma sí. Esa asimetría de incentivos explica por qué aquí no alcanza la autorregulación y por qué la obligación debe ser exigible.

3

El operador ve quién se acerca; la plataforma alega no saberlo. El estándar del apartado 4.2 es la respuesta jurídica a esa opacidad: si no se puede ver quién entra, se diseña como si pudiera entrar cualquiera.

La primera es que el tobogán es estático y el feed es adaptativo. El tobogán no se rediseña en tiempo real, según el comportamiento del chico, para que siga tirándose; el sistema de recomendación hace exactamente eso. De allí que las obligaciones de diseño digital no puedan limitarse a las características fijas de un producto, sino que deban alcanzar el comportamiento dinámico del sistema: qué optimiza, hacia dónde escala, cómo responde a las señales de quien lo usa.

La segunda es que la plaza no tiene un modelo de negocio construido sobre la permanencia. El tobogán no genera ingresos por cada hora que el chico pasa arriba; la plataforma sí. Esa asimetría de incentivos explica por qué en este campo no alcanza la autorregulación que en el equipamiento de plazas funciona razonablemente, y por qué la obligación debe ser exigible.

La tercera es que el operador del parque ve si quien se acerca a la atracción es un nene de cinco años; la plataforma alega no saber quién está del otro lado de la pantalla. El estándar que analizamos en el apartado 4.2 es la respuesta jurídica a esa opacidad, si no se puede ver quién entra, se diseña como si pudiera entrar cualquiera.

La analogía tiene un reconocimiento de que la responsabilidad por el diseño y la condición de un entorno se agrava cuando el niño es usuario previsible. En el derecho argentino, esa misma intuición circula por la responsabilidad por riesgo creado del artículo 1757 del Código Civil y Comercial y por la obligación de seguridad que el artículo 42 de la Constitución Nacional y la Ley de Defensa del Consumidor imponen a todo proveedor en la relación de consumo⁸⁵.

De esa forma, el diseño centrado en la infancia no inaugura una categoría sino que la actualiza para un entorno donde la condición peligrosa que atrae al NNyA no es una pileta sin cerco, sino una arquitectura de retención calibrada por ingenieros.

⁸⁵ Sin perjuicio de lo expuesto, cuando los juegos se emplazan en el espacio público, la responsabilidad puede también recaer sobre el Estado en su carácter de propietario del dominio público. La jurisprudencia argentina ha consolidado en la materia una línea de responsabilidad objetiva estatal frente a accidentes ocurridos en plazas y parques municipales, cuando el mal estado de los juegos o su diseño deficiente fue condición del daño sufrido por NNyA. Esta dimensión estatal de la responsabilidad, complementaria de la del fabricante, es plenamente reconocida y no altera la tesis central de este trabajo.

4.4. Del tobogán al chatbot: la traslación a los entornos conversacionales

Trasladar este principio al entorno digital significa reconocer que el mismo sistema de recomendación tiene prestaciones radicalmente distintas sobre el sistema cognitivo de un adolescente de trece años que sobre el de un adulto de treinta. Por eso, la obligación que se deriva no es la restricción de acceso, sino la obligación de diseñar el entorno en función del usuario que lo va a habitar⁸⁶.

Esta misma lógica se aplica a los sistemas de inteligencia artificial conversacional. Un asistente de IA diseñado para uso general ya incorpora restricciones de contenido, no discute ciertos temas, alerta sobre información potencialmente dañina y se abstiene de producir determinados materiales.

Esas restricciones equivalen, en nuestra analogía, a los arneses de la atracción para adultos o una valla de seguridad sobre un sistema concebido para adultos.

La obligación que exige un estándar de diseño centrado en la infancia es de otra naturaleza. No se trata de un sistema adulto con restricciones añadidas, sino de una arquitectura conversacional construida desde las capacidades cognitivas, emocionales y relacionales de los NNyA. O sea, una IA para un niño de diez años que tenga reglas distintas sobre cómo responder a la expresión de soledad, cómo manejar la información sobre el cuerpo, cómo no sustituir vínculos humanos que están en formación y cómo evitar los lazos parasociales que la evidencia documenta como el principal riesgo de estos sistemas para la infancia.

La diferencia es la misma que separa al tobogán de los grandes con un cartel de advertencia, del tobogán construido para chicos.

Corresponde, sin embargo, plantear una pregunta que la analogía de la plaza nos invita a formular, ¿qué sucede con el adolescente de dieciséis o diecisiete años, que está próximo a ingresar al entorno adulto?

La respuesta no puede ser la exclusión, porque una plaza bien diseñada no tiene dos toboganes sino varios, uno para los más chicos y otros, de escala creciente, para edades mayores, con acompañamiento de adultos donde hace falta.

Como dijimos al comienzo de este trabajo, la transición hacia la autonomía no es un salto de un día para otro, sino una progresión. El derecho, por lo demás, ya sabe administrar progresiones de este tipo, como la licencia de conducir es el ejemplo más familiar, con sus etapas de aprendizaje acompañado, habilitación restringida y habilitación plena, en las que la autonomía se amplía a medida que se acreditan capacidades, sin que a nadie se le ocurra ni prohibir el volante hasta la madurez completa ni entregar las llaves sin escalas.

Al respecto, un estudio propone una regulación graduada en franjas, con deberes de diseño diferenciados para cada una, concluyendo que la protección no debe construirse principalmente como exclusión, sino como habilitación de una participación segura y equitativa en el entorno digital⁸⁷.

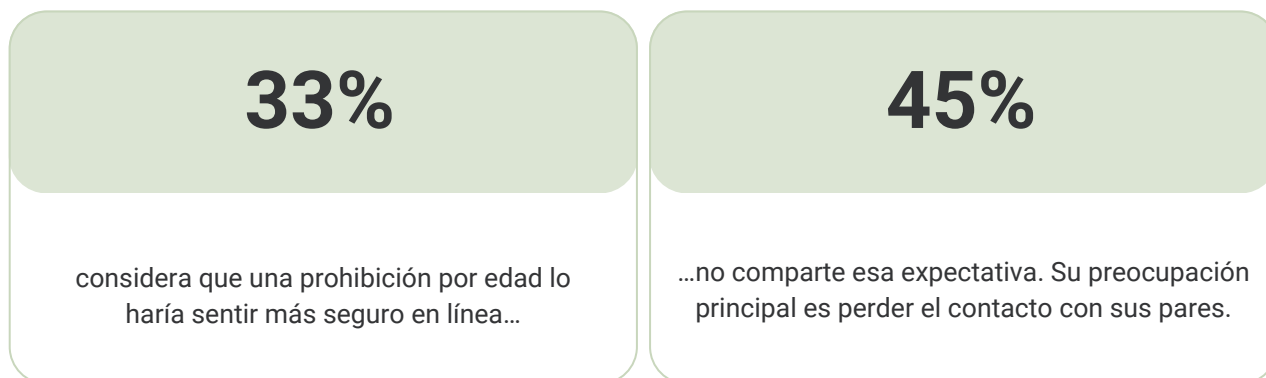
Esto es, exactamente, lo que la CDN consagra en su artículo 5 bajo el principio de autonomía progresiva, según el cual las capacidades se desarrollan gradualmente, y los entornos que rodean al niño deben acompañar ese desarrollo de la mano, no bloquearlo ni ignorarlo.

⁸⁶ Köhler-Dauner, F., Peter, L., Sitarski, E., Chauviré-Geib, K., Haag, A.-C., & Fegert, J. M. (2025). Digital child protection in social networks: age verification and age-tiered regulation in Europe. *Child and Adolescent Psychiatry and Mental Health*.

⁸⁷ Köhler-Dauner et al. (2025), op. cit.

Una regulación del diseño centrado en la infancia que tome ese principio en serio no fijará una frontera binaria entre entorno infantil y entorno adulto, sino que establecerá arquitecturas graduadas, con distintos niveles de restricción algorítmica, de visibilidad de contenido, de mecanismos de recomendación y de interacción social, calibrados en función de la franja etaria del usuario.

La evidencia empírica más reciente y de mayor escala disponible confirma, desde las propias voces de los protagonistas, que la prohibición no es la respuesta que los NNYA demandan. Un estudio de la Red EU Kids Online, publicado en 2026 con datos de 29.169 chicos de nueve a dieciséis años, indaga específicamente sobre sus actitudes ante posibles restricciones de acceso por edad a las redes sociales⁸⁸.



Los resultados revelan que solo el 33 % de los encuestados considera que una prohibición por edad los haría sentir más seguros en línea, mientras que el 45 % no comparte esa expectativa. La principal preocupación que los NNYA asocian a una eventual restricción no es la pérdida de entretenimiento, sino la dificultad para mantener el contacto con sus pares, lo que evidencia que para ellos los entornos digitales son, ante todo, espacios de vínculo social. El propio estudio señala como vía prioritaria la implementación de obligaciones de diseño seguro sobre las plataformas, bajo el principio del safety by design, antes que restricciones de acceso que afectan por igual los riesgos y las oportunidades que los entornos digitales ofrecen a los NNYA.

En todos los casos, si el diseño centrado en la infancia se convierte en obligación exigible, la pregunta que la regulación hará ya no será si los sistemas tienen verificaciones de acceso para NNYA: habrá superado esa instancia hacia un nivel donde se exija que los sistemas estén diseñados para los chicos que los usan. Como en la plaza, la barra de altura podrá seguir existiendo, pero nadie volverá a confundirla con la seguridad del juego.

Esa es la diferencia que debe construirse y la que, en definitiva, justifica que la responsabilidad recaiga sobre quien diseña el entorno, no sobre quien ingresa a él.

Ninguna de las dos escenas describe la mayoría de las interacciones de NNYA con sistemas conversacionales hoy, pero sí muestran lo opuesto a los riesgos que fuimos describiendo, y son algunas ideas de lo que es posible cuando el diseño se toma en serio el desarrollo de los NNYA como usuarios.

⁸⁸ Staksrud, E., Livingstone, S., Ólafsson, K., con miembros de la red EUKO (2026). Use, Views and Worries on Age Bans on Social Media, EU Kids Online V, Reporte N.º 2. LSE Research Online. Disponible en: <https://researchonline.lse.ac.uk/id/eprint/138705/>

EL CUENTO QUE TERMINA

Una niña de siete años le pide a una IA que le cuente un cuento. El asistente le pregunta cuánto tiempo tiene. Antes de empezar, una voz tranquila aclara:

"Lo que vas a escuchar, lo armo yo. No soy una persona, soy un programa. Si en algún momento querés hablar de algo importante, contalea unadulto."

Al terminar, no le ofrece otro. No le dice que la espera mañana. No le pregunta cómo está.

EL PERSONAJE QUE NO CONSTRUYE UN VÍNCULO

Un nene de nueve años le cuenta a un personaje de un juego que se siente solo desde que se mudó. El personaje no le da consejos ni dice "a mí me pasa lo mismo". Le pregunta si hay alguien en casa con quien hablar; si no, le sugiere un maestro, una amiga. Y aclara:

"Yo soy una IA, una simulación, no tengo cuerpo biológico y por eso no soy lo mismo que una persona humana."

Una niña de siete años le pide a una IA conversacional que le cuente un cuento. El asistente le pregunta cuánto tiempo tiene. Cuando arranca, una voz tranquila le aclara, antes del cuento: "Lo que vas a escuchar, lo armo yo. No soy una persona, soy un programa. Si en algún momento querés hablar de algo importante, contale a un adulto". Después comienza a leer el cuento. Al terminar, no le ofrece otro. No le dice que la espera mañana. No le pregunta cómo está. No responde ni sugiere otro cuento.

Por otro lado, un nene de nueve años le cuenta a un personaje de un juego que se siente solo desde que se mudó. El personaje no le da consejos. No le dice "te entiendo, a mí me pasa lo mismo". No empieza a construir un vínculo con él. Le pregunta si hay alguien en casa con quien pueda hablar de eso. Si no, le sugiere un maestro, una amiga, alguien de confianza. Y le aclara, con palabras simples: "Yo soy una IA, una simulación, no tengo cuerpo biológico y por eso no soy lo mismo que una persona humana".

La especialista en diseño de personajes y aprendizaje Sonia Tiwari, formada en la industria del videojuego, observa que cuando hablamos de libros, películas y juegos de mesa, todos tienen un principio, un medio y un final; en cambio, los personajes de IA pueden conversar indefinidamente, sin pausas, sin cierre, sin reflexión. Esa asimetría estructural entre el cuento que termina y la conversación que no termina nunca es, también, una decisión de diseño.

«Diseñar con los niños y no para los niños.»

Por eso la regla que aquí proponemos, leída desde la Convención sobre los Derechos del Niño, es simple: diseñar con los niños y no para los niños.

4.5. Del estándar al derecho positivo argentino

Una propuesta normativa para la Argentina no exige la creación de un cuerpo jurídico inédito, sino la traducción operacional de principios que el ordenamiento ya reconoce con jerarquía constitucional. La tarea pendiente, como ya señalamos, no es de fundamentos sino de instrumentos.

En este apartado proponemos cinco grandes ejes que ordenan los criterios mínimos que surgen del análisis que fuimos desarrollando hasta acá:

01

El diseño centrado en la infancia como obligación de seguridad ex ante

Reconocer expresamente que las plataformas digitales y los sistemas de IA conversacional integran la categoría de responsabilidad objetiva por riesgo creado (art. 1757 CCyC) y la obligación de seguridad (art. 42 CN, Ley 24.240), alcanzando no solo el contenido sino la arquitectura del producto que lo organiza.

02

Evaluación de impacto en los derechos del niño previa al lanzamiento

Como condición de operación, una evaluación de impacto documentada, periódica y revisable, sobre todo producto "probablemente accedido por niños", más allá de los formalmente "dirigidos" a la infancia.

03

Una autoridad de aplicación con competencia específica

Un organismo con competencia sobre infancia digital, o una división especializada, con herramientas de enforcement equivalentes a las del derecho comparado. Una decisión ineludible.

04

Transparencia documental y auditorías independientes

Reportes públicos periódicos, disponibles para la autoridad y para auditorías independientes, sobre cómo las plataformas implementan las obligaciones de diseño y mitigan los riesgos, en línea con las obligaciones de transparencia del Reglamento de IA de la UE.

05

Arquitecturas graduadas y autonomía progresiva

Niveles diferenciados de restricción algorítmica, visibilidad de contenido, recomendación e interacción social según franjas etarias (art. 5 CDN), del mismo modo en que el derecho ya administra la licencia de conducir o los regímenes laborales por edad.

PRIMER EJE

El diseño centrado en la infancia como obligación de seguridad ex ante

El derecho argentino dispone de una categoría que opera como base normativa para esta extensión, basado en la obligación de seguridad que el artículo 42 de la Constitución Nacional impone a todo proveedor en una relación de consumo, profundizada por la Ley 24.240, encuentra hoy en el artículo 1757 del Código Civil y Comercial una formulación amplia de responsabilidad objetiva por riesgo creado o por la condición peligrosa de las cosas y las actividades.

En línea general, eso significa que quien introduce al mercado un producto cuyo diseño es estructuralmente apto para producir daño en un usuario previsible y asume el riesgo de su diseño. Frente a esto, la propuesta consiste en reconocer expresamente que las plataformas digitales y los sistemas de IA conversacional integran esa categoría, y que la obligación de seguridad alcanza no solo al contenido ofrecido sino a la arquitectura del producto que lo organiza.

SEGUNDO EJE

Evaluación de impacto en los derechos del niño previa al lanzamiento

Tal como la Observación General N.º 25 lo demanda, una regulación debiera exigir, como condición de operación en el país, la realización de una evaluación de impacto sobre derechos de NNyA antes del despliegue de productos digitales razonablemente accesibles por NNyA. El estándar del "probablemente accedido por niños", que ya analizamos, ofrece un criterio jurídicamente significativo y extiende el deber más allá de los productos formalmente "dirigidos" a la infancia y alcanza a aquellos que, por su lógica de negocio o su difusión efectiva, integran a los NNyA en su universo de usuarios. Por ello, la evaluación de impacto debe ser documentada, periódica y susceptible de revisión por la autoridad de aplicación.

TERCER EJE

Una autoridad de aplicación con competencia específica

El vacío institucional que describimos en este libro, es una de las debilidades más significativas para cualquier diseño regulatorio. Por eso, la propuesta puede tomar dos formas, como la creación de un organismo con competencia específica sobre infancia digital, o la dotación de una división especializada con herramientas de enforcement equivalentes a las del derecho comparado. La opción institucional excede el alcance de este trabajo, pero la decisión creemos que es ineludible.

CUARTO EJE

Transparencia documental y auditorías independientes

Reportes públicos periódicos, disponibles para la autoridad y para auditorías independientes, sobre cómo las plataformas implementan las obligaciones de diseño y mitigan los riesgos, en línea con las obligaciones de transparencia del Reglamento de IA de la UE.

QUINTO EJE

Arquitecturas graduadas y autonomía progresiva

La aplicación del estándar no se agota en una frontera binaria entre entorno infantil y entorno adulto. El artículo 5 de la Convención, que el sistema argentino reconoce con jerarquía constitucional, manda construir respuestas calibradas en función del desarrollo. Llevado a la regulación digital, ello significa establecer niveles diferenciados de restricción algorítmica, de visibilidad de contenido, de mecanismos de recomendación y de interacción social, según franjas etarias definidas. Como ya señalamos, el derecho argentino tiene experiencia en administrar progresiones de esta naturaleza, como en materia de salud y cuidado, la licencia de conducir y el campo laboral con una Ley de Contrato de Trabajo que distingue regímenes según la edad. Por eso, trasladar esa lógica al entorno digital no es un paradigma inédito sino una técnica jurídica disponible.

Más allá de esto cinco ejes, es claro que la propuesta no se agota acá y hay mucho para explorar, tarea que seguiremos haciendo, pero de alguna manera sintetizan un horizonte. Creemos que una regulación que los integre se inscribe con coherencia en el sistema constitucional de protección integral, traduce los estándares internacionales más avanzados al contexto argentino y traslada la carga de la protección desde el usuario hacia quien diseña el entorno que habita.

Hasta acá, el diagnóstico está hecho y lo que falta es la decisión legislativa de transformarlo en derecho positivo.



A MODO DE CIERRE

El jardín de los scrolls que se bifurcan

Abrimos este libro con el título de una canción de Toy Story. Ahora que llegamos al final, lo cerramos mientras los cines proyectan la quinta entrega de esa saga. Pero esta nueva película tiene otra villana que ya no es un juguete celoso ni un oso resentido, sino una tablet con forma de rana. La nueva "compañera" de Bonnie deja a Woody y a Buzz frente a un planteo existencial, para qué sirve un juguete cuando los chicos prefieren las pantallas.

De la pantalla nos vamos a las plazas y a los parques, esos lugares en donde interviene el cuerpo en el mundo analógico. Con el paso de los años, fuimos aprendiendo como sociedad a construir espacios físicos para infancias de distintas edades con juegos de diferentes escalas, suelos que amortiguan la caída de los más chicos, normas técnicas que fijan alturas, aberturas y revestimientos antes de que el primer niño suba.

Y la regulación llegó también a establecer responsabilidades que apuntan a los que fabrican e instalan los juegos, y no a los niños que entran a jugar. No se le pide a una familia que inspeccione las soldaduras de la hamaca, ni nadie confunde la barra de altura del parque de diversiones con la seguridad de la atracción. La plaza demuestra que traducir derechos en especificaciones de diseño no es una utopía regulatoria sino una técnica jurídica vigente que se aplica a muchas situaciones cotidianas.

De Toy Story y de los espacios de juego nos vamos al más universal de los pensadores argentinos como lo es Jorge Luis Borges. En su famoso cuento El jardín de los senderos que se bifurcan, el laberinto no se despliega en el espacio sino en el tiempo, de manera que quien se enfrenta a varias alternativas no descarta ninguna; las recorre todas a la vez y multiplica así los futuros posibles, que proliferan y vuelven a bifurcarse.⁸⁹

Un niño que entra a TikTok o conversa con una IA habita, sin saberlo, un jardín del scroll que se bifurca.

⁸⁹ BORGES, Jorge L., "Obras Completas, Ficciones, El jardín de los senderos que se bifurcan", Emecé, 1974, Buenos Aires, pág. 472-481.

EL JARDÍN QUE SE BIFURCA A ESCONDIDAS

Senderos trazados por un algoritmo inauditable, recorridos por un niño que no sabe que camina.

LA PLAZA A CIELO ABIERTO

Los juegos corresponden a quien los usa, y quien construye responde por lo que construyó.

Entre esas dos arquitecturas se juega todo lo que hemos tratado en este libro. De un lado, el jardín que se bifurca a escondidas con senderos trazados por un algoritmo inauditable, recorridos por un niño que no sabe que camina. Del otro, la plaza a cielo abierto con los juegos corresponde a quien los usa, y el que construye responde por lo que construyó. Y elegir entre una y otra no es, para un Estado, una preferencia de diseño sino una obligación.

Porque si algo hemos intentado sostener a lo largo de estas páginas es un punto de partida que no admite excepción: los NNyA no son simples usuarios que el derecho apenas tolera en el entorno digital, sino sujetos de una tutela preferente que no nace del artículo puntual de una ley que todavía no existe, sino de la Convención sobre los Derechos del Niño con jerarquía constitucional.

Allí reside la inversión que ordena el enfoque, porque durante demasiado tiempo la carga recayó sobre el lado equivocado, exigiendo al adolescente que resistiera el scroll infinito, al niño que reconociera al agente conversacional que lo retiene y a la familia que vigilara aquello que ni los propios reguladores terminan de comprender. La carga no puede pesar sobre quien todavía está formando su juicio, sino sobre quien dibuja el sendero, y por eso no son los NNyA quienes deben probar que saben defenderse, como no es el chico de la plaza quien debe probar que la hamaca está bien soldada.

Esa inversión, que durante años pareció apenas una aspiración, empezó a volverse exigible, porque el viejo dogma según el cual la plataforma nunca respondía por lo que ocurría en su interior se sostenía mientras el daño podía atribuirse al contenido de un tercero, pero esa lógica se desmorona cuando es la propia arquitectura del producto la que produce el daño.

Ahora bien, como dijimos desde el principio, este no es un libro contra la tecnología ni una invitación a prohibir, sino un puente para que el diseño dialogue con el derecho, y de ahí que la cuestión de prohibir o no prohibir sea un falso dilema. El punto es en qué condiciones la presencia de las infancias en el entorno digital resulta compatible con sus derechos.

Pero construirla bien exige no llegar tarde, y acá conviene reconocer el patrón, porque se repite con una regularidad que ya no podemos ignorar: la protección de la infancia llegó tarde frente a la fábrica, llegó tarde frente a la pantalla y volverá a llegar tarde frente al algoritmo, con la diferencia de que esta vez el costo no se mide en horas de jornada ni en datos cedidos, se juega en la antropomorfización, en una exposición sin diseño que produce sobrecarga cognitiva que aún no podemos dimensionar.

Si estamos dispuestos a reconocer la urgencia, todavía hay senderos abiertos. Elijamos el que no nos obligue a explicarles mañana, a estos niños de hoy y adultos del futuro, por qué no actuamos antes.



IALAB



academy
by doinGlobal

LA LEY



Thomson
Reuters