

La inteligencia artificial en el derecho penal suizo

Potencial de eficiencia y límites del Estado de derecho

Kaspar Brönnimann · Web Edition 2026 · Switzerland · kaspar@kaspair.ch · kaspair.ch

ABSTRACT

Este trabajo examina si, y en qué medida, la inteligencia artificial puede mejorar la eficiencia de la justicia penal suiza sin debilitar las garantías fundamentales del Estado de derecho. Distingue entre funciones de investigación, sanción y decisión judicial, y analiza la explicabilidad, los sesgos, el sesgo de automatización, el principio de culpabilidad individual, las garantías de un proceso justo y la atribución de responsabilidad. El trabajo defiende un modelo híbrido en el que la IA actúe como herramienta de apoyo a la decisión mientras la responsabilidad jurídica última permanezca en una persona identificable.

Nota editorial: este trabajo se basa en un texto preparado en el contexto del curso «AI & Law» ofrecido por la Universidad de Lund en Coursera. La versión entregada para el curso se redujo para cumplir los límites de la actividad. Esta edición web independiente recupera un aparato de fuentes e incorpora pequeñas actualizaciones técnicas y bibliográficas. No es una publicación oficial de la Universidad de Lund ni ha sido avalada por ella. Este texto no constituye asesoramiento jurídico.

1 Introducción

La inteligencia artificial (IA) se utiliza cada vez más en procesos de decisión pública, especialmente en ámbitos que implican grandes volúmenes de datos o tareas predictivas complejas. El derecho penal, tradicionalmente caracterizado por la atribución individual de culpabilidad, la independencia judicial y estrictas garantías del Estado de derecho, también se ve afectado por esta evolución.

Para valorar el potencial de eficiencia de la IA en la justicia penal es necesario distinguir entre las funciones de los distintos actores estatales. El uso de IA por autoridades investigadoras puede servir a una averiguación rápida y eficaz de los hechos, mientras que su uso por autoridades que juzgan o sancionan está sometido a estándares jurídicos más estrictos.

Una particularidad del procedimiento penal suizo es que las decisiones sancionadoras no están reservadas exclusivamente a los tribunales. En el procedimiento de orden penal, el Ministerio Público puede determinar la culpabilidad e imponer una pena, ejerciendo materialmente una función decisoria. En estos casos, las consideraciones de eficiencia pierden claramente peso.

De ello deben distinguirse las infracciones menores tratadas mediante el procedimiento de multas de orden, basado en un catálogo estandarizado. El uso de sistemas automatizados o asistidos por IA es menos problemático en este contexto, siempre que se mantenga garantizada la posibilidad de formular oposición y trasladar así el asunto al procedimiento penal ordinario.

La actividad de los tribunales, en cambio, se orienta a la determinación individual de la culpabilidad y de la pena. En este ámbito, la rapidez no tiene el mismo peso que la transparencia, el deber de motivación, la independencia judicial y la protección jurídica efectiva.

El presente trabajo examina, mediante un análisis jurídico-doctrinal, si y en qué medida el uso de IA en el derecho penal suizo puede generar ganancias de eficiencia jurídicamente admisibles sin comprometer principios fundamentales del derecho penal ni garantías constitucionales y de derechos humanos.^{7 6 3}

2 Hechos y tecnología: la IA en el contexto penal

El uso de IA en la justicia penal comprende aplicaciones técnicas diversas con consecuencias jurídicas diferentes. En Suiza, la IA no adopta autónomamente decisiones penales, pero puede apoyar la decisión humana e influir de facto en determinadas orientaciones de investigación o priorización.^{7 6}

2.1 Formas técnicas relevantes para el derecho penal

Los sistemas de IA relevantes para la justicia penal incluyen herramientas de análisis y predicción, sistemas de clasificación, búsqueda y priorización. El aprendizaje automático deriva patrones o funciones de decisión a partir de datos en lugar de ejecutar únicamente reglas programadas de forma explícita; un modelo desplegado en producción no necesariamente continúa adaptándose de manera autónoma.

Una preocupación central es la explicabilidad limitada. Cuando la base de un resultado algorítmico no puede reconstruirse suficientemente por el decisor humano, pueden verse afectados el deber de motivación y el derecho a ser oído.^{7 8}

2.2 Uso por las autoridades de investigación en Suiza

En la persecución penal, la IA puede apoyar el trabajo de investigación, por ejemplo mediante el análisis de grandes conjuntos de datos o de material visual. Estos sistemas tienen principalmente una función preparatoria, pero pueden producir decisiones previas de hecho cuando sus recomendaciones condicionan el curso posterior de la investigación.⁶

2.3 Uso por tribunales y autoridades sancionadoras

El uso de IA por los tribunales sigue siendo comparativamente limitado en Suiza. Una decisión autónoma de culpabilidad o pena por parte de una IA no está prevista ni es compatible con la estructura de la responsabilidad penal individual. Sin embargo, los sistemas algorítmicos pueden apoyar procesos administrativos, investigación jurídica o la estructuración de expedientes extensos.

En otros ordenamientos se han utilizado herramientas de evaluación de riesgo para apoyar decisiones sobre detención, libertad condicional o determinación de la pena. Estos ejemplos muestran las tensiones que aparecen cuando las predicciones algorítmicas se aproximan al núcleo de la decisión judicial.

Desde el punto de vista jurídico, estos usos son especialmente sensibles porque pueden afectar directamente a derechos fundamentales, a la independencia judicial y al principio de culpabilidad.^{6 11}

2.4 Riesgos técnicos y límites estructurales

Los sistemas de IA utilizados en contextos de justicia penal pueden reproducir sesgos presentes en los datos, producir efectos discriminatorios y reducir la trazabilidad de las decisiones. A ello se añade el sesgo de automatización: el riesgo de que los decisores humanos otorguen un peso desproporcionado a recomendaciones automatizadas.

Estas características son centrales para la evaluación jurídica porque condicionan la posibilidad de una revisión efectiva, de una impugnación real y de una responsabilidad humana identificable.^{8 9 10}

3 Análisis jurídico: desafíos del uso de IA en el derecho penal

3.1 Principios fundamentales del derecho penal suizo

El derecho penal suizo se basa en el principio de culpabilidad individual. El art. 47 CP exige una determinación individualizada de la pena, teniendo en cuenta las circunstancias personales del autor y las circunstancias concretas del delito.

Los sistemas de IA, en cambio, operan con frecuencia mediante correlaciones estadísticas y probabilidades. Pueden detectar patrones o asignar puntuaciones de riesgo, pero no pueden valorar por sí mismos la culpabilidad moral ni asumir responsabilidad jurídica. Aparece así una tensión estructural cuando la inferencia estadística se aproxima a la valoración de la culpabilidad individual.²

3.2 Garantías de un proceso justo y transparencia

Las garantías esenciales del Estado de derecho se derivan, entre otras normas, del art. 29 de la Constitución Federal y del art. 6 del Convenio Europeo de Derechos Humanos: procedimiento justo, derecho a ser oído y acceso efectivo a recursos.

Cuando una decisión penalmente relevante se ve materialmente influida por un proceso algorítmico opaco, la persona afectada puede tener dificultades para comprender o impugnar las razones reales de la decisión. La explicabilidad y la documentación pasan así a ser exigencias procesales, y no meramente técnicas.¹⁴⁸

3.3 Igualdad ante la ley y discriminación

El art. 8 de la Constitución Federal garantiza la igualdad ante la ley. Los sistemas de IA pueden poner en riesgo este principio cuando los datos de entrenamiento o de uso reflejan desigualdades sociales existentes y producen efectos indirectamente discriminatorios.

La discriminación algorítmica puede ser difícil de detectar cuando surge mediante efectos estadísticos y no mediante distinciones explícitas, lo que complica el control judicial y la protección jurídica efectiva.¹⁹

3.4 Responsabilidad y atribución

Otro problema fundamental es la atribución de responsabilidad. En una decisión asistida por IA, las contribuciones técnicas y organizativas se distribuyen entre desarrolladores, autoridades, operadores y decisores. El ejercicio del poder público, sin embargo, debe seguir siendo atribuible a una autoridad identificable y responsable.

Si en una injerencia en la libertad asistida por IA deja de estar claro quién responde por la decisión, se debilita una exigencia básica del Estado de derecho. La gobernanza debe, por tanto, hacer visible la responsabilidad en lugar de permitir que se diluya dentro del sistema.⁶¹³

4 Discusión: eficiencia frente al Estado de derecho

La IA puede analizar rápidamente grandes cantidades de datos, favorecer una mayor consistencia y aliviar procesos sobrecargados. Estas ventajas potenciales son especialmente relevantes en funciones investigadoras y administrativas.

Sin embargo, deben ponderarse frente a riesgos importantes para el Estado de derecho. Las decisiones que implican graves injerencias en derechos fundamentales —como la detención, la declaración de culpabilidad o la imposición de una pena— requieren juicio humano,

transparencia y una responsabilidad claramente atribuible. En estos contextos, la IA solo puede desempeñar, como máximo, una función de apoyo.

Una posible vía consiste en un modelo híbrido: la IA apoya el análisis y la preparación, mientras que la decisión final y la responsabilidad permanecen en manos humanas. Este modelo requiere reglas claras sobre transparencia, documentación, revisabilidad y supervisión humana. El Reglamento europeo de IA también clasifica diversos usos en la aplicación de la ley y la administración de justicia como de alto riesgo y los somete a requisitos de gobernanza y supervisión.^{5 6 10}

5 Conclusión

El uso de inteligencia artificial en el derecho penal suizo ofrece un potencial de eficiencia, pero al mismo tiempo plantea importantes desafíos para los principios del derecho penal y del Estado de derecho.

Una delegación completa de las decisiones penales a la IA es incompatible con la culpabilidad individual, las garantías de un proceso justo y una administración de justicia cuya responsabilidad pueda atribuirse. Por ello, la IA solo puede justificarse como herramienta auxiliar, integrada en un marco jurídico claro y sometida a una supervisión humana efectiva.

La conclusión central no es que la IA no tenga cabida en la justicia penal. Es, más bien, que la eficiencia no puede comprarse al precio de la responsabilidad. Los derechos individuales, la transparencia, la posibilidad de revisión y la atribución de responsabilidad deben mantenerse intactos cuando los sistemas digitales intervienen en decisiones públicas de consecuencias graves.^{1 3 4 5}

FUENTES Y BASES JURÍDICAS

¹ Constitution fédérale suisse (Cst.), art. 8 et 29 / Constitución Federal Suiza, arts. 8 y 29.

<https://www.fedlex.admin.ch/eli/cc/1999/404/en>

² Code pénal suisse (CP), art. 47 / Código Penal Suizo, art. 47.

https://www.fedlex.admin.ch/eli/cc/54/757_781_799/en

³ Code de procédure pénale suisse (CPP), art. 352-354 / Código de Procedimiento Penal Suizo, arts. 352-354.

<https://www.fedlex.admin.ch/eli/cc/2010/267/en>

- 4 Convention européenne des droits de l'homme, art. 6 / Convenio Europeo de Derechos Humanos, art. 6.
<https://www.echr.coe.int/en/web/echr/european-convention-on-human-rights>
- 5 Règlement (UE) 2024/1689 sur l'intelligence artificielle / Reglamento (UE) 2024/1689 de Inteligencia Artificial.
<https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- 6 CEPEJ, Charte éthique européenne d'utilisation de l'IA dans les systèmes judiciaires (2018).
<https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment>
- 7 Harry Surden, "Machine Learning and Law", 89 Washington Law Review 87 (2014).
<https://digitalcommons.law.uw.edu/wlr/vol89/iss1/5/>
- 8 Jenna Burrell, "How the machine 'thinks': Understanding opacity in machine learning algorithms", Big Data & Society 3(1) (2016).
<https://doi.org/10.1177/2053951715622512>
- 9 Solon Barocas & Andrew D. Selbst, "Big Data's Disparate Impact", 104 California Law Review 671 (2016).
<https://doi.org/10.15779/Z38BG31>
- 10 Raja Parasuraman & Victor Riley, "Humans and Automation: Use, Misuse, Disuse, Abuse", Human Factors 39(2), 230-253 (1997).
<https://web.mit.edu/16.459/www/parasuraman.pdf>
- 11 Tania Sourdin, "Judge v Robot? Artificial Intelligence and Judicial Decision-Making", 41(4) UNSW Law Journal 1114 (2018).
<https://www.unsw.edu.au/content/dam/pdfs/law/unsw-law-journal/2010-2019/Vol-No-41-4-Sourdin.pdf>
- 12 Conseil fédéral suisse, « Artificial Intelligence and International Rules » (2022).
<https://www.admin.ch/gov/en/start/documentation/media-releases.msg-id-88019.html>
- 13 Office fédéral de la justice, aperçu de la réglementation de l'IA en Suisse / Oficina Federal de Justicia, regulación de IA en Suiza.
<https://www.bj.admin.ch/bj/en/home/staat/gesetzgebung/kuenstliche-intelligenz.html>