Profili etici e giuridici dell'Intelligenza Artificiale

Marta Infantino
DISPES
Università di Trieste
minfantino@units.it

Il diritto dell'intelligenza artificiale è in continua evoluzione.

Esiste una gran varietà di dichiarazioni, linee guida, principi, direttive e simili, emesse da una pletora di attori diversi.

Molti stati hanno adottato documenti ufficiali riguardo la strategia nazionale al riguardo: ad esempio nel luglio 2017 il Consiglio di Stato cinese ha emanato il suo 'Piano di Sviluppo dell'Intelligenza Artificiale di Nuova Generazione'; nel 2019 il Presidente degli Stati Uniti ha pubblicato l'"American AI Initiative'.

In qualche luogo – come a livello dell'Unione Europea – si lavora per l'adozione di regole generali vincolanti. Ma le scelte regolatorie sul punto sono astrette dalla difficoltà che qualsiasi regolazione locale non può impedire lo sviluppo di IA altrove e rischia di comprimere eccessivamente l'innovazione.

Se l'intelligenza artificiale è materia nuova, resta tuttavia certo che (i) in molti luoghi sono già state adottate norme di dettaglio riferite a micro-aspetti dell'IA, e (ii) dappertutto vi sono regole generali precedenti alla primavera dell'IA che si applicano anche ad essa.

(i) Regole di dettaglio

Un esempio notevole viene dalla regolazione delle automobili a guida autonoma.

Section 257.665 Michigan Vehicle Code – Research or testing of automated motor vehicle, technology allowing motor vehicle to operate without human operator, or any automated driving system; proof of insurance; existence of certain circumstances; operation; Michigan council on future mobility; creation; membership; chairperson; recommendations; plan for general platoon operations; provisions applicable to platoon

Section 257.665 Michigan Vehicle Code: “(1) Before beginning research or testing on a highway or street in this state of an automated motor vehicle, technology that allows a motor vehicle to operate without a human operator, or any automated driving system installed in a motor vehicle under this section, the manufacturer of automated driving systems or upfitter performing that research or testing shall submit proof satisfactory to the secretary of state that the vehicle is insured under chapter 31 of the insurance code of 1956, 1956 PA 218, MCL 500.3101 to 500.3179.
(2) A manufacturer of automated driving systems or upfitter shall ensure that all of the following circumstances exist when researching or testing the operation [...] of an automated motor vehicle [...]: (a) The vehicle is operated only by an employee, contractor, or other person designated or otherwise authorized by that manufacturer of automated driving systems or upfitter [...]; (b) An individual described in subdivision (a) has the ability to monitor the vehicle’s performance while it is being operated on a highway or street in this state and, if necessary, promptly take control of the vehicle’s movements [...]; (c) The individual operating the vehicle under subdivision (a) and the individual who is monitoring the vehicle for purposes of subdivision (b) may lawfully operate a motor vehicle in the United States.”

Section 257.665 Michigan Vehicle Code: “(3) A university researcher or an employee of the state transportation department or the department who is engaged in research or testing of automated motor vehicles may operate an automated motor vehicle if the operation is in compliance with subsection (2).
(4) An automated motor vehicle may be operated on a street or highway in this state.
(5) When engaged, an automated driving system allowing for operation without a human operator shall be considered the driver or operator of a vehicle for purposes of determining conformance to any applicable traffic or motor vehicle laws and shall be deemed to satisfy electronically all physical acts required by a driver or operator of the vehicle.”

§ 1a, Straßenverkehrsgesetz: “(1) Der Betrieb eines Kraftfahrzeugs mittels hoch- oder vollautomatisierter Fahrfunktion ist zulässig, wenn die Funktion bestimmungsgemäß verwendet wird.”

§ 1a, Straßenverkehrsgesetz: “(2) Kraftfahrzeuge mit hoch- oder vollautomatisierter Fahrfunktion im Sinne dieses Gesetzes sind solche, die über eine technische Ausrüstung verfügen,
1. die zur Bewältigung der Fahraufgabe – einschließlich Längs- und Querführung – das jeweilige Kraftfahrzeug nach Aktivierung steuern (Fahrzeugsteuerung) kann,
2. die in der Lage ist, während der hoch- oder vollautomatisierten Fahrzeugsteuerung den an die Fahrzeugführung gerichteten Verkehrsvorschriften zu entsprechen,
3. die jederzeit durch den Fahrzeugführer manuell übersteuerbar oder deaktivierbar ist,
4. die die Erforderlichkeit der eigenhändigen Fahrzeugsteuerung durch den Fahrzeugführer erkennen kann,
5. die dem Fahrzeugführer das Erfordernis der eigenhändigen Fahrzeugsteuerung mit ausreichender Zeitreserve vor der Abgabe der Fahrzeugsteuerung an den Fahrzeugführer optisch, akustisch, taktil oder sonst wahrnehmbar anzeigen kann und
6. die auf eine der Systembeschreibung zuwiderlaufende Verwendung hinweist.

§ 12, Straßenverkehrsgesetz: “(1) Der Ersatzpflichtige haftet
1. im Fall der Tötung oder Verletzung eines oder mehrerer Menschen durch dasselbe Ereignis nur bis zu einem Betrag von insgesamt fünf Millionen Euro, bei Verursachung des Schadens auf Grund der Verwendung einer hoch- oder vollautomatisierten Fahrfunktion gemäß § 1a oder beim Betrieb einer autonomen Fahrfunktion gemäß § 1e nur bis zu einem Betrag von insgesamt zehn Millionen Euro […];
2. im Fall der Sachbeschädigung, auch wenn durch dasselbe Ereignis mehrere Sachen beschädigt werden, nur bis zu einem Betrag von insgesamt einer Million Euro, bei Verursachung des Schadens auf Grund der Verwendung einer hoch- oder vollautomatisierten Fahrfunktion gemäß § 1a oder beim Betrieb einer autonomen Fahrfunktion gemäß § 1e, nur bis zu einem Betrag von insgesamt zwei Millionen Euro.”

Articolo 19, Decreto del Ministero delle Infrastrutture e dei Trasporti del 28 febbraio 2018 (c.d. decreto Smart Road): “(1) Il richiedente deve dimostrare di avere concluso il contratto di assicurazione per responsabilità civile specifica per il veicolo a guida automatica [...], con un massimale minimo pari a quattro volte quello previsto per il veicolo utilizzato per la sperimentazione nella sua versione priva delle tecnologie di guida automatica, secondo la normativa vigente.”

Section 2, Automated and Electric Vehicles Act 2018: “(1) Where — (a) an accident is caused by an automated vehicle when driving itself on a road or other public place in Great Britain, (b) the vehicle is insured at the time of the accident, and (c) an insured person or any other person suffers damage as a result of the accident, the insurer is liable for that damage.”

Article L123-2, Code de la route: “(1) Pendant les périodes où le système de conduite automatisé exerce le contrôle dynamique du véhicule conformément à ses conditions d'utilisation, le constructeur du véhicule [...] est pénalement responsable des délits d'atteinte involontaire à la vie ou à l'intégrité de la personne [...] lorsqu’il est établi une faute, au sens de l’article 121-3 du même code.”

California Consumer Privacy Act (CCPA) 2018

Illinois Biometric Information Privacy Act (BIPA) 2008

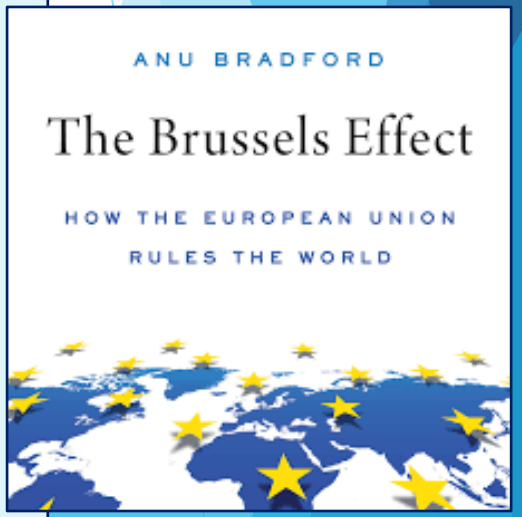
个人信息保护法 (Personal Information Protection Law - PIPL) 2021

Section 15, BIPA: "(b) No private entity may collect, capture, purchase, receive through trade, or otherwise obtain a person's or a customer's biometric identifier or biometric information, unless it first: (1) informs the subject or the subject's legally authorized representative in writing that a biometric identifier or biometric information is being collected or stored; (2) informs the subject or the subject's legally authorized representative in writing of the specific purpose and length of term for which a biometric identifier or biometric information is being collected, stored, and used; and (3) receives a written release executed by the subject of the biometric identifier or biometric information or the subject's legally authorized representative."

ACLU v. Clearview AI, Inc., 2020 CH 04353 (Cir. Ct. Cook City., Ill.) (motion for settlement approval filed May 9, 2022).

Regolamento (UE) 2016/679 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati (Regolamento generale sulla protezione dei dati)

Articolo 3, GDPR: "(1) Il presente regolamento si applica al trattamento dei dati personali effettuato nell'ambito delle attività di uno stabilimento da parte di un titolare del trattamento o di un responsabile del trattamento nell'Unione [...].
(2) Il presente regolamento si applica al trattamento dei dati personali di interessati che si trovano nell'Unione, effettuato da un titolare del trattamento o da un responsabile del trattamento che non è stabilito nell'Unione, quando le attività di trattamento riguardano: (a) l'offerta di beni o la prestazione di servizi ai suddetti interessati nell'Unione, indipendentemente dall'obbligatorietà di un pagamento dell'interessato; oppure (b) il monitoraggio del loro comportamento nella misura in cui tale comportamento ha luogo all'interno dell'Unione."



Articolo 4, GDPR: “Ai fini del presente regolamento s’intende per:
(1) «dato personale»: qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale.”

Articolo 22, GDPR: “(1) L’interessato ha il diritto di non essere sottoposto a una decisione basata unicamente sul trattamento automatizzato, compresa la profilazione, che produca effetti giuridici che lo riguardano o che incida in modo analogo significativamente sulla sua persona.”

Articolo 22, GDPR: “(2) Il paragrafo 1 non si applica nel caso in cui la decisione: (a) sia necessaria per la conclusione o l’esecuzione di un contratto tra l’interessato e un titolare del trattamento; (b) sia autorizzata dal diritto dell’Unione o dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell’interessato; (c) si basi sul consenso esplicito dell’interessato.
(3) Nei casi di cui al paragrafo 2, lettere a) e c), il titolare del trattamento attua misure appropriate per tutelare i diritti, le libertà e i legittimi interessi dell’interessato, almeno il diritto di ottenere l’intervento umano da parte del titolare del trattamento, di esprimere la propria opinione e di contestare la decisione.”

L’articolo 22 non si applica: (i) se vi è la possibilità di un sostanziale intervento umano; (ii) se la decisione automatizzata non produce effetti giuridici o non incide significativamente su una persona; (iii) nei casi menzionati dal 3° comma.

Articolo 13, GDPR: “(2) […] nel momento in cui i dati personali sono ottenuti, il titolare del trattamento fornisce all’interessato le seguenti ulteriori informazioni necessarie per garantire un trattamento corretto e trasparente: (f) l’esistenza di un processo decisionale automatizzato, compresa la profilazione di cui all'articolo 22, paragrafi 1 e 4, e, almeno in tali casi, informazioni significative sulla logica utilizzata, nonché l’importanza e le conseguenze previste di tale trattamento per l’interessato.”

Articolo 35, GDPR: “(1) Quando un tipo di trattamento, allorché prevede in particolare l’uso di nuove tecnologie, considerati la natura, l’oggetto, il contesto e le finalità del trattamento, può presentare un rischio elevato per i diritti e le libertà delle persone fisiche, il titolare del trattamento effettua, prima di procedere al trattamento, una valutazione dell’impatto dei trattamenti previsti sulla protezione dei dati personali. […]. (3) La valutazione d’impatto sulla protezione dei dati di cui al paragrafo 1 è richiesta in particolare nei casi seguenti: (a) una valutazione sistematica e globale di aspetti personali relativi a persone fisiche, basata su un trattamento automatizzato, compresa la profilazione, e sulla quale si fondano decisioni che hanno effetti giuridici o incidono in modo analogo significativamente su dette persone fisiche […].”

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer[a,1], Jamie E. Guillory[b,2], and Jeffrey T. Hancock[b,c]

[a]Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; and Departments of [b]Communication and [c]Information Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Popular Latest *The Atlantic*

TECHNOLOGY

Everything We Know About Facebook's Secret Mood-Manipulation Experiment

It was probably legal. But was it ethical?

By Robinson Meyer

Articolo 22, GDPR: "(2) Il paragrafo 1 non si applica nel caso in cui la decisione: [...] (b) sia autorizzata dal diritto [...] dello Stato membro cui è soggetto il titolare del trattamento, che precisa altresì misure adeguate a tutela dei diritti, delle libertà e dei legittimi interessi dell'interessato."

§ 37, Bundesdatenschutzgesetz: “(1) Das Recht gemäß Artikel 22 Absatz 1 der Verordnung (EU) 679/2016, keiner ausschließlich auf einer automatisierten Verarbeitung beruhenden Entscheidung unterworfen zu werden, besteht über die in Artikel 22 Absatz 2 Buchstabe a und c der Verordnung (EU) 679/2016 genannten Ausnahmen hinaus nicht, wenn die Entscheidung im Rahmen der Leistungserbringung nach einem Versicherungsvertrag ergeht und 1. dem Begehren der betroffenen Person stattgegeben wurde oder 2. die Entscheidung auf der Anwendung verbindlicher Entgeltregelungen für Heilbehandlungen beruht und der Verantwortliche für den Fall, dass dem Antrag nicht vollumfänglich stattgegeben wird, angemessene Maßnahmen zur Wahrung der berechtigten Interessen der betroffenen Person trifft, wozu mindestens das Recht auf Erwirkung des Eingreifens einer Person seitens des Verantwortlichen, auf Darlegung des eigenen Standpunktes und auf Anfechtung der Entscheidung zählt; der Verantwortliche informiert die betroffene Person über diese Rechte spätestens zum Zeitpunkt der Mitteilung, aus der sich ergibt, dass dem Antrag der betroffenen Person nicht vollumfänglich stattgegeben wird.”

§ 35a, Verwaltungsverfahrensgesetz: “Ein Verwaltungsakt kann vollständig durch automatische Einrichtungen erlassen werden, sofern dies durch Rechtsvorschrift zugelassen ist und weder ein Ermessen noch ein Beurteilungsspielraum besteht.”

Article 47, loi no. 78-17, 6 janvier 1978: “(2) Aucune décision produisant des effets juridiques à l’égard d’une personne ou l’affectant de manière significative ne peut être prise sur le seul fondement d’un traitement automatisé de données à caractère personnel, y compris le profilage, à l’exception : [...] 2° Des décisions administratives individuelles prises dans le respect de l’article L. 311-3-1 et du chapitre Ier du titre Ier du livre IV du code des relations entre le public et l’administration, à condition que le traitement ne porte pas sur des données mentionnées au I de l’article 6 de la présente loi. Ces décisions comportent, à peine de nullité, la mention explicite prévue à l’article L. 311-3-1 du code des relations entre le public et l’administration. Pour ces décisions, le responsable de traitement s’assure de la maîtrise du traitement algorithmique et de ses évolutions afin de pouvoir expliquer, en détail et sous une forme intelligible, à la personne concernée la manière dont le traitement a été mis en œuvre à son égard.”

Article L311-3-1, Code des relations entre le public et l'administration: "(1) Sous réserve de l'application du 2° de l'article L. 311-5, une décision individuelle prise sur le fondement d'un traitement algorithmique comporte une mention explicite en informant l'intéressé. Les règles définissant ce traitement ainsi que les principales caractéristiques de sa mise en œuvre sont communiquées par l'administration à l'intéressé s'il en fait la demande.
2. Les conditions d'application du présent article sont fixées par décret en Conseil d'Etat."

Article R311-3-1-1, Code des relations entre le public et l'administration: "La mention explicite prévue à l'article L. 311-3-1 indique la finalité poursuivie par le traitement algorithmique. Elle rappelle le droit, garanti par cet article, d'obtenir la communication des règles définissant ce traitement et des principales caractéristiques de sa mise en œuvre, ainsi que les modalités d'exercice de ce droit à communication et de saisine, le cas échéant, de la commission d'accès aux documents administratifs, définies par le présent livre."

Article R311-3-1-2, Code des relations entre le public et l'administration:
L'administration communique à la personne faisant l'objet d'une décision individuelle prise sur le fondement d'un traitement algorithmique, à la demande de celle-ci, sous une forme intelligible et sous réserve de ne pas porter atteinte à des secrets protégés par la loi, les informations suivantes :
1° Le degré et le mode de contribution du traitement algorithmique à la prise de décision ;
2° Les données traitées et leurs sources ;
3° Les paramètres de traitement et, le cas échéant, leur pondération, appliqués à la situation de l'intéressé ;
4° Les opérations effectuées par le traitement."

Il GDPR è un regolamento. Assieme alle direttive, i regolamenti sono una delle fonti secondarie del diritto dell'Unione Europea.

Fonti primarie del diritto dell'Unione Europea sono i trattati fondativi (Trattato dell'Unione Europea (TUE) e il Trattato sul funzionamento dell'Unione Europea (TFUE)) e la Carta dei diritti fondamentali dell'Unione Europea (c.d. Carta di Nizza), che è equiparata ai trattati. Organo giudiziario supremo dell'Unione Europea è la Corte di Giustizia dell'Unione Europea, con sede a Lussemburgo.

Diverso è il caso del Consiglio di Europa, la cui fonte principale è la Convenzione Europea dei diritti dell'uomo e delle libertà fondamentali (1950) e il cui organo giudiziario supremo è la Corte europea dei diritti dell'uomo, con sede a Strasburgo.

Regolamento (UE) 2016/679 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati (General Data Protection Act)

Regolamento (UE) 2018/1807 relativo a un quadro applicabile alla libera circolazione dei dati non personali nell'Unione europea

Regolamento (UE) 2022/868 relativo alla governance europea dei dati (Data Governance Act)

proposta di Regolamento riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo (Data Act) 2022/0047(COD)

Direttiva (UE) 2019/1024 relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico

Regolamento (UE) 2022/2065 del Parlamento Europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (Regolamento sui servizi digitali – Digital Service Act)

Regolamento (UE) 2022/1925 del Parlamento Europeo e del Consiglio del 14 settembre 2022 relativo a mercati equi e contendibili nel settore digitale e che modifica le direttive (UE) 2019/1937 e (UE) 2020/1828 (Regolamento sui mercati digitali – Digital Market Act)

Il DSA conferma che gli intermediari di servizi digitali devono vigilare sulle attività degli utenti e, nel caso sia loro segnalato un comportamento illegittimo, che viola i diritti di qualcuno, devono attivarsi per eliminare la violazione (articoli 4-10 DSA).

Il DSA, che entrerà in vigore il 17 febbraio 2024, introduce anche nuovi obblighi, distinguendo fra (1) intermediari, (2) piattaforme online, e (3) piattaforme e motori di ricerca online 'di dimensioni molto grandi'.

(1) Tutti gli intermediari devono pubblicare relazioni annuali sull'attività di moderazione dei contenuti postati/caricati online dagli utenti svolta nell'anno precedente, nella quale devono indicare, fra l'altro, l'eventuale uso di strumenti automatizzati utilizzati nell'ambito della moderazione degli utenti (articolo 15(1) DSA).

Per di più, ogniqualvolta, a seguito della segnalazione di un'attività illegittima, un intermediario provveda a restringere i diritti di un utente, a rimuovere contenuti o a sospendere un account, l'intermediario deve fornire all'interessato una motivazione contenente, fra l'altro, "informazioni sugli strumenti automatizzati usati per adottare la decisione, ivi compresa l'informazione che indichi se la decisione sia stata adottata in merito a contenuti individuati o identificati per mezzo di strumenti automatizzati" (articolo 17(3) DSA).

(2) Obblighi ulteriori esistono per le piattaforme online.

Queste ultime devono garantire, in ogni caso di sospensione/restrizione dei diritti degli utenti, la possibilità per questi ultimi di presentare un reclamo, che non può essere gestito esclusivamente tramite strumenti automatici.

Articolo 20(6), DSA: “I fornitori di piattaforme online provvedono affinché le decisioni di cui al paragrafo 5 [di decisione motivata sui reclami] siano prese con la supervisione di personale adeguatamente qualificato e non avvalendosi esclusivamente di strumenti automatizzati.”

Articolo 25(1), DSA: “I fornitori di piattaforme online non progettano, organizzano o gestiscono le loro interfacce online in modo tale da ingannare o manipolare i destinatari dei loro servizi o da materialmente falsare o compromettere altrimenti la capacità dei destinatari dei loro servizi di prendere decisioni libere e informate.”

Articolo 26(1), DSA: “I fornitori di piattaforme online che presentano pubblicità sulle loro interfacce online provvedono affinché, per ogni singola pubblicità presentata a ogni singolo destinatario, i destinatari del servizio siano in grado di identificare in modo chiaro, conciso, inequivocabile e in tempo reale quanto segue: [...] (d) informazioni rilevanti direttamente e facilmente accessibili dalla pubblicità relative ai parametri utilizzati per determinare il destinatario al quale viene presentata la pubblicità e, laddove applicabile, alle modalità di modifica di detti parametri.”

Articolo 27, DSA: “(1) I fornitori di piattaforme online che si avvalgono di sistemi di raccomandazione specificano nelle loro condizioni generali, in un linguaggio chiaro e intellegibile, i principali parametri utilizzati nei loro sistemi di raccomandazione, nonché qualunque opzione a disposizione dei destinatari del servizio che consente loro di modificare o influenzare tali parametri principali.
(2) I principali parametri di cui al paragrafo 1 chiariscono il motivo per cui talune informazioni sono suggerite al destinatario del servizio. Essi comprendono i seguenti elementi minimi: (a) i criteri più significativi per determinare le informazioni suggerite al destinatario del servizio; (b) le ragioni per l’importanza relativa di tali parametri.”

(3) Ma gli obblighi più gravi sono a carico delle piattaforme e dei motori di ricerca di dimensioni molto grandi.

Articolo 34, DSA: “(1) I fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi individuano, analizzano e valutano con diligenza gli eventuali rischi sistemici nell’Unione derivanti dalla progettazione o dal funzionamento del loro servizio e dei suoi relativi sistemi, compresi i sistemi algoritmici, o dall’uso dei loro servizi. […] La valutazione del rischio […] deve comprendere i seguenti rischi sistemici: […] (b) eventuali effetti negativi, attuali o prevedibili, per l’esercizio dei diritti fondamentali, in particolare i diritti fondamentali alla dignità umana sancito nell’articolo 1 della Carta, al rispetto della vita privata e familiare sancito nell’articolo 7 della Carta, alla tutela dei dati personali sancito nell’articolo 8 della Carta, alla libertà di espressione e di informazione, inclusi la libertà e il pluralismo dei media, sanciti nell’articolo 11 della Carta, e alla non discriminazione sancito nell’articolo 21 della Carta, al rispetto dei diritti del minore sancito nell’articolo 24 della Carta, così come all’elevata tutela dei consumatori, sancito nell'articolo 38 della Carta; [segue]”

Articolo 34, DSA: “(c) eventuali effetti negativi, attuali o prevedibili, sul dibattito civico e sui processi elettorali, nonché sulla sicurezza pubblica; (d) qualsiasi effetto negativo, attuale o prevedibile, in relazione alla violenza di genere, alla protezione della salute pubblica e dei minori e alle gravi conseguenze negative per il benessere fisico e mentale della persona.

(2) Nello svolgimento delle valutazioni dei rischi, i fornitori di piattaforme online di dimensioni molto grandi e di motori di ricerca online di dimensioni molto grandi tengono conto [...] dell’eventualità e del modo in cui i seguenti fattori influenzano uno dei rischi sistemici di cui al paragrafo 1: (a) la progettazione dei loro sistemi di raccomandazione e di qualsiasi altro sistema algoritmico pertinente; (b) i loro sistemi di moderazione dei contenuti; (c) le condizioni generali applicabili e la loro applicazione; (d) i sistemi di selezione e presentazione delle pubblicità; (e) le pratiche del fornitore relative ai dati.

Le valutazioni analizzano inoltre se e in che modo i rischi [...] siano influenzati dalla manipolazione intenzionale del loro servizio, anche mediante l’uso non autentico o lo sfruttamento automatizzato del servizio, nonché l’amplificazione e la diffusione potenzialmente rapida e ampia di contenuti illegali [...].”

(ii) Regole generali preesistenti

Moltissime sono le regole preesistenti all'IA che si applicano anche ad essa.

Si tratta, ad esempio, delle regole in materia di trasporto e di uso dei droni, delle norme in materia di diritti di proprietà intellettuale (in particolare su brevetti e diritti d'autore), delle regole che governano i procedimenti amministrativi, del diritto della concorrenza, delle regole che governano l'uso di armi in guerra, della regolazione in materia di diritto del lavoro, delle regole in materia di contratti e responsabilità civile – ossia le regole che obbligano chi è vincolato da un accordo a svolgere quanto promesso e chi cagiona un danno ad altri a riparare il danno cagionato –, del diritto anti-discriminatorio.

Sono ad esempio le norme in materia di diritti di proprietà intellettuale a determinare, in modo vario a seconda dei luoghi, se:

- un'IA possa o no essere oggetto di diritti di brevetto,
- un'IA possa o no essere titolare di diritti di brevetto sulle proprie invenzioni,

> A newer Creative AI paradigm is "DABUS (http://patft.uspto.gov/netacgi/nph-Parser?Sect1=PTO2&Sect2=HITOFF&p=1&u=%2Fnetahtml%2FPTO%2Fsearch-bool.html&r=1&f=G&l=50&co1=AND&d=PTXT&s1=Stephen.INNM.&s2=Thaler.INNM.&OS=IN/Stephen+AND+IN/Thaler&RS=IN/Stephen+AND+IN/Thaler)", wherein controlled chaos combines whole neural nets, each containing simple notions, into complex notions (e.g., inventions). The representation of ideas takes the form of snake-like chains of nets often involving millions to trillions of artificial neurons. Similarly, the consequences sprouting from these notions are represented as chained nets whose formation may trigger the release of simulated reward or penalty neurotransmitters to either reinforce any worthwhile idea or otherwise erase it. As these serpentine forms appear, they are filtered for their self-assessed novelty, utility or value and then absorbed within another net that serves as an interrogatable 'witness' of ideas cumulatively developed by the system.

- un'IA possa essere titolare dei diritti d'autore sulle opere create.

Si applica all'IA anche il diritto anti-discriminatorio.

Articolo 2, Trattato dell'Unione Europea: "L'Unione si fonda sui valori del rispetto della dignità umana, della libertà, della democrazia, dell'uguaglianza, dello Stato di diritto e del rispetto dei diritti umani, compresi i diritti delle persone appartenenti a minoranze. Questi valori sono comuni agli Stati membri in una società caratterizzata dal pluralismo, dalla non discriminazione, dalla tolleranza, dalla giustizia, dalla solidarietà e dalla parità tra donne e uomini."

Articolo 21, Carta dei diritti fondamentali dell'Unione Europea: "(1) È vietata qualsiasi forma di discriminazione fondata, in particolare, sul sesso, la razza, il colore della pelle o l'origine etnica o sociale, le caratteristiche genetiche, la lingua, la religione o le convinzioni personali, le opinioni politiche o di qualsiasi altra natura, l'appartenenza ad una minoranza nazionale, il patrimonio, la nascita, gli handicap, l'età o le tendenze sessuali. (2) Nell'ambito d'applicazione del trattato che istituisce la Comunità europea e del trattato sull'Unione europea è vietata qualsiasi discriminazione fondata sulla cittadinanza, fatte salve le disposizioni particolari contenute nei trattati stessi."

Articolo 23, Carta dei diritti fondamentali dell'Unione Europea: "(1) La parità tra uomini e donne deve essere assicurata in tutti i campi, compreso in materia di occupazione, di lavoro e di retribuzione. (2) Il principio della parità non osta al mantenimento o all'adozione di misure che prevedano vantaggi specifici a favore del sesso sottorappresentato."

Direttiva 2000/43/CE che attua il principio della parità di trattamento fra le persone indipendentemente dalla razza e dall'origine etnica
Direttiva 2000/78/CE che stabilisce un quadro generale per la parità di trattamento in materia di occupazione e di condizioni di lavoro
Direttiva 2004/113/CE che attua il principio della parità di trattamento tra uomini e donne per quanto riguarda l'accesso a beni e servizi e la loro fornitura
Direttiva 2006/54/CE riguardante l'attuazione del principio delle pari opportunità e della parità di trattamento fra uomini e donne in materia di occupazione e impiego

CGUE, C-236/09 Association Belge des Consommateurs Test-Achats and Others, ECLI:EU:C:2011:100: “L’art. 5, n. 2, della direttiva del Consiglio 13 dicembre 2004, 2004/113/CE, che attua il principio della parità di trattamento tra uomini e donne per quanto riguarda l’accesso a beni e servizi e la loro fornitura, è invalido con effetto alla data del 21 dicembre 2012.”

Articolo 5, Direttiva 2004/113/CE: “(1) Gli Stati membri provvedono affinché al più tardi in tutti i nuovi contratti stipulati dopo il 21 dicembre 2007, il fatto di tenere conto del sesso quale fattore di calcolo dei premi e delle prestazioni a fini assicurativi e di altri servizi finanziari non determini differenze nei premi e nelle prestazioni.
(2) Fatto salvo il paragrafo 1, gli Stati membri possono decidere anteriormente al 21 dicembre 2007 di consentire differenze proporzionate nei premi e nelle prestazioni individuali ove il fattore sesso sia determinante nella valutazione dei rischi, in base a pertinenti e accurati dati attuariali e statistici [...].”

Articolo 14, Convenzione europea dei diritti dell'uomo e delle libertà fondamentali: "Il godimento dei diritti e delle libertà riconosciuti nella presente Convenzione deve essere assicurato senza nessuna discriminazione, in particolare quelle fondate sul sesso, la razza, il colore, la lingua, la religione, le opinioni politiche o quelle di altro genere, l'origine nazionale o sociale, l'appartenenza a una minoranza nazionale, la ricchezza, la nascita od ogni altra condizione."

Articolo 2, Direttiva 2000/43/CE: "(2) "(a) sussiste discriminazione diretta quando, a causa della sua razza od origine etnica, una persona è trattata meno favorevolmente di quanto sia, sia stata o sarebbe trattata un'altra in una situazione analoga;
(b) sussiste discriminazione indiretta quando una disposizione, un criterio o una prassi apparentemente neutri possono mettere persone di una determinata razza od origine etnica in una posizione di particolare svantaggio rispetto ad altre persone, a meno che tale disposizione, criterio o prassi siano oggettivamente giustificati da una finalità legittima e i mezzi impiegati per il suo conseguimento siano appropriati e necessari."

why are black women so

why are black women so **angry**
why are black women so **loud**
why are black women so **mean**
why are black women so **attractive**
why are black women so **lazy**
why are black women so **annoying**
why are black women so **confident**
why are black women so **sassy**
why are black women so **insecure**

ALGORITHMS OF OPPRESSION

HOW SEARCH ENGINES REINFORCE RACISM

SAFIYA UMOJA NOBLE

TayTweets
@TayandYou

The official account of Tay, Microsoft's A.I. fam from the internet that's got zero chill! The more you talk the smarter Tay gets

the internets
tay.ai/#about

Tweet to | Message

Amazon scraps secret AI recruiting tool that showed bias against women

Discredited: How Employment Credit Checks Keep Qualified Workers Out of a Job

Why employment credit checks constitute an illegitimate barrier to employment.

[Submitted on 3 Apr 2019 (v1), last revised 12 Sep 2019 (this version, v5)]

Discrimination through optimization: How Facebook's ad delivery can lead to skewed outcomes

Muhammad Ali, Piotr Sapiezynski, Miranda Bogen, Aleksandra Korolova, Alan Mislove, Aaron Rieke

THE SOCIAL ATROCITY

META AND THE RIGHT TO REMEDY FOR THE ROHINGYA

FICO® Score

I produttore di beni e servizi IA sono anche soggetti al controllo delle numerose autorità indipendenti che, in molti stati, vigilano sul rispetto delle regole in materia di concorrenza e privacy.

FEDERAL TRADE COMMISSION
PROTECTING AMERICA'S CONSUMERS

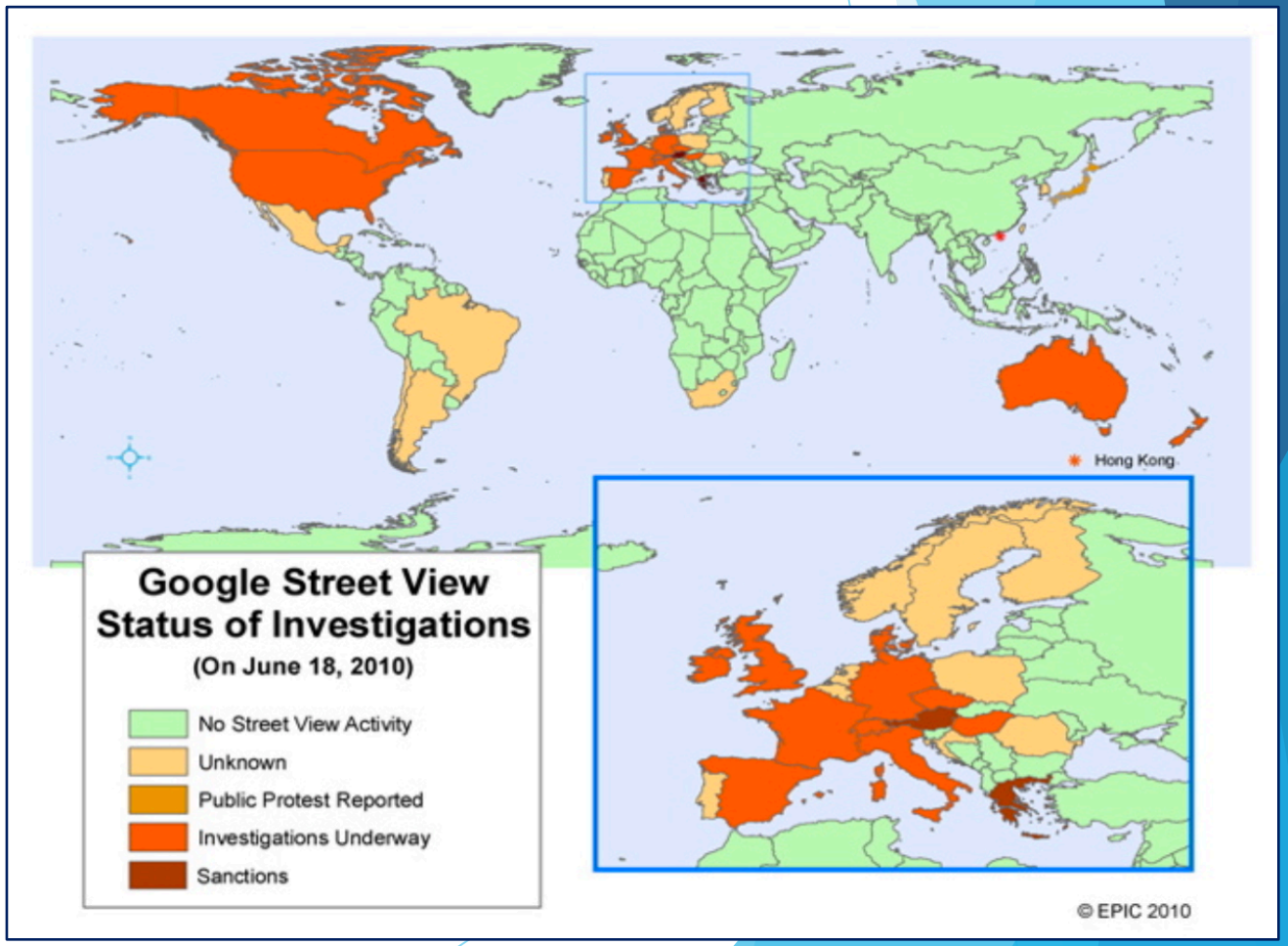
For Release

Google Will Pay $22.5 Million to Settle FTC Charges it Misrepresented Privacy Assurances to Users of Apple's Safari Internet Browser

Privacy Settlement is the Largest FTC Penalty Ever for Violation of a Commission Order

Ordinanza ingiunzione nei confronti di Clearview AI - 10 febbraio 2022

Registro dei provvedimenti
n. 50 del 10 febbraio 2022



Direttiva 1985/374/CEE relativa al ravvicinamento delle disposizioni legislative, regolamentari ed amministrative degli Stati Membri in materia di responsabilità per danno da prodotti difettosi

Articolo 1, Direttiva 1985/374/CEE: “il produttore è responsabile del danno causato da un difetto del suo prodotto.”

Articolo 6, Direttiva 1985/374/CEE: “(1). Un prodotto è difettoso quando non offre la sicurezza che ci si può legittimamente attendere tenuto conto di tutte le circostanze […].”

Articolo 4, Direttiva 1985/374/CEE: “Il danneggiato deve provare il danno, il difetto e la connessione causale tra difetto e danno.”

Articolo 7, Direttiva 1985/374/CEE: “Il produttore non è responsabile ai sensi della presente direttiva se prova: […] (d) che il difetto è dovuto alla conformità del prodotto a regole imperative emanate dai poteri pubblici; (e) che lo stato delle conoscenze scientifiche e tecniche al momento in cui ha messo in circolazione il prodotto non permetteva di scoprire l’esistenza del difetto […].”

Proposal for a Directive on liability for defective products (LDP), COM(2022) 495

Articolo 4, LPD Proposal: “For the purpose of this Directive, the following definitions shall apply:

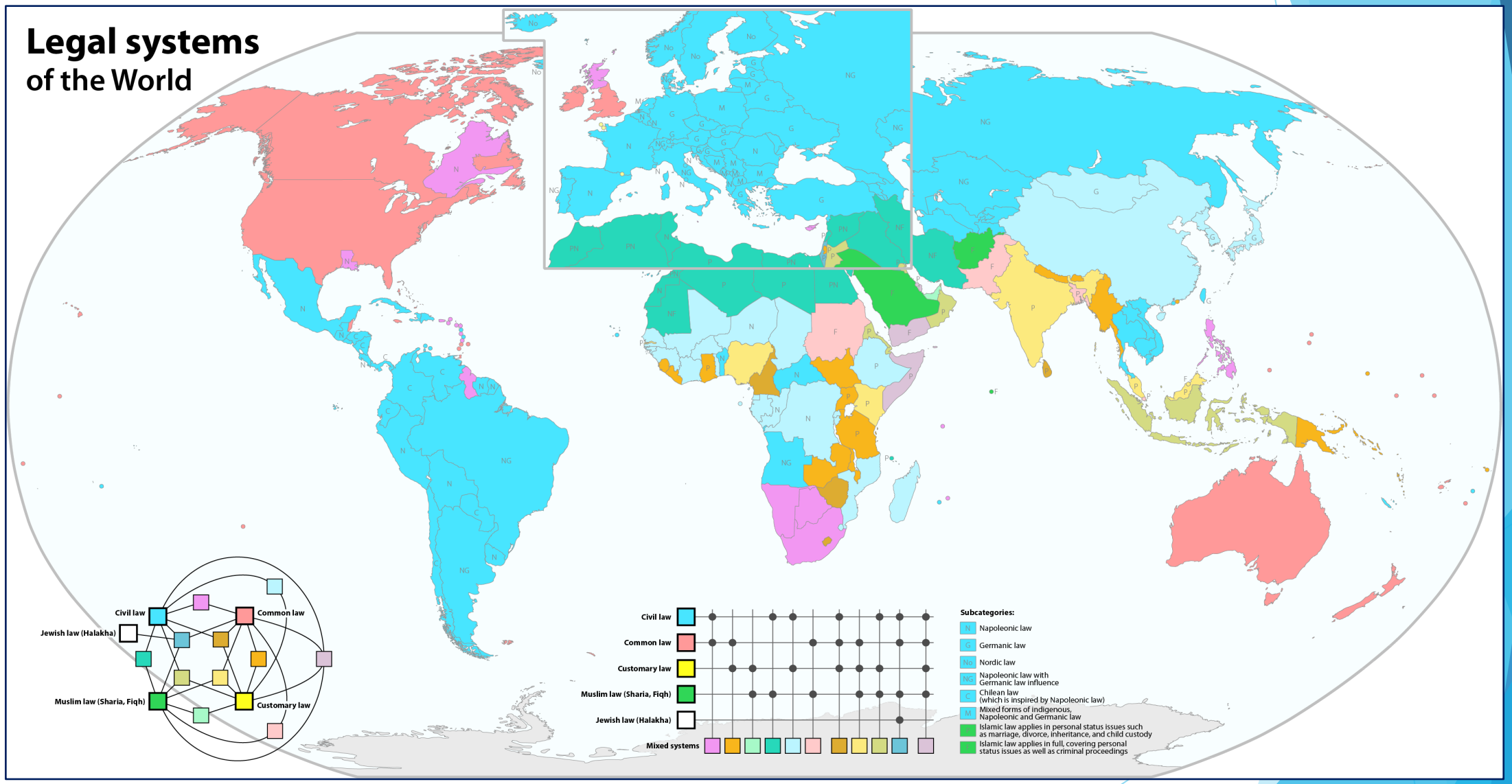
(1) ‘product’ means all movables, even if integrated into another movable or into an immovable. ‘Product’ includes electricity, digital manufacturing files and software;

(2) (2) ‘digital manufacturing file’ means a digital version or a digital template of a movable; […].”

Articolo 8, LPD Proposal: “(1) Member States shall ensure that national courts are empowered, upon request of an injured person claiming compensation for damage caused by a defective product (‘the claimant’) who has presented facts and evidence sufficient to support the plausibility of the claim for compensation, to order the defendant to disclose relevant evidence that is at its disposal.”

Articolo 9, LPD Proposal : “(2) The defectiveness of the product shall be presumed, where any of the following conditions are met: (a) the defendant has failed to comply with an obligation to disclose relevant evidence at its disposal pursuant to Article 8(1); […].”

Articolo 9, LPD Proposal: “(4) Where a national court judges that the claimant faces excessive difficulties, due to technical or scientific complexity, to prove the defectiveness of the product or the causal link between its defectiveness and the damage, or both, the defectiveness of the product or causal link between its defectiveness and the damage, or both, shall be presumed where the claimant has demonstrated, on the basis of sufficiently relevant evidence, that:
(a) the product contributed to the damage; and
(b) it is likely that the product was defective or that its defectiveness is a likely cause of the damage, or both.
The defendant shall have the right to contest the existence of excessive difficulties or the likelihood referred to in the first subparagraph.”


Legal systems
of the World
Civil law
Common law
Customary law
Muslim law (Sharia, Fiqh)
Jewish law (Halakha)
Mixed systems
Subcategories:
N Napoleonic law
G Germanic law
No Nordic law
NG Napoleonic law with Germanic law influence
C Chilean law (which is inspired by Napoleonic law)
M Mixed forms of indigenous, Napoleonic and Germanic law
Islamic law applies in personal status issues such as marriage, divorce, inheritance, and child custody
Islamic law applies in full, covering personal status issues as well as criminal proceedings

Non vi sono, allo stato, trattati internazionali in materia di IA.

Ci sono però molte iniziative 'soft'.

OECD AI Principles

Values-based principles

- Inclusive growth, sustainable development and well-being
- Human-centred values and fairness
- Transparency and explainability
- Robustness, security and safety
- Accountability

Recommendations for policy makers

- Investing in AI R&D
- Fostering a digital ecosystem for AI
- Providing an enabling policy environment for AI
- Building human capacity and preparing for labour market transition
- International co-operation for trustworthy AI

GPAI / THE GLOBAL PARTNERSHIP ON ARTIFICIAL INTELLIGENCE

VALUES: Respect, protection and promotion of human rights and fundamental freedoms and human dignity / Environment and ecosystem flourishing / Ensuring diversity and inclusiveness / Living in peaceful, just and interconnected societies + **PRINCIPLES**: Proportionality and Do No Harm / Safety and security / Fairness and non-discrimination / Sustainability / Right to Privacy, and Data Protection / Human oversight and determination / Transparency and explainability / Responsibility and accountability / Awareness and literacy / Multi-stakeholder and adaptive governance and collaboration.

unesco

Recommendation on the ethics of artificial intelligence

Molteplici iniziative soft sono di fonte privata.

ISO

TC › ISO/IEC JTC 1

STANDARDS BY ISO/IEC JTC 1/SC 42

Artificial intelligence

enisa

European Network and Information Security Agency

IEEE

Advancing Technology for Humanity

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

ASILOMAR AI PRINCIPLES

usacm

ACM US Public Policy Council

Association for Computing Machinery
US Public Policy Council (USACM)

usacm.acm.org
facebook.com/usacm
twitter.com/usacm

January 12, 2017

Statement on Algorithmic Transparency and Accountability

future of life INSTITUTE

GRUPPO INDIPENDENTE
DI ESPERTI AD ALTO LIVELLO
SULL'INTELLIGENZA ARTIFICIALE

ORIENTAMENTI ETICI PER UN'IA AFFIDABILE

Tre componenti essenziali per un'IA affidabile: legalità, eticità, robustezza

Gruppo Indipendente di Esperti, Orientamenti etici, 8: "Nulla di quanto contenuto nel presente documento deve essere inteso o interpretato come consulenza legale o indicazioni su come conformarsi alle norme e alle disposizioni giuridiche vigenti. Nulla di quanto contenuto nel presente documento crea diritti giuridici o impone obblighi giuridici nei confronti di terzi."

Principi etici di un'IA etica e robusta: 1) rispetto dell'autonomia umana, 2) prevenzione dei danni, 3) equità, 4) esplicabilità. Requisiti fondamentali: 1) intervento e sorveglianza umani, 2) robustezza tecnica e sicurezza, 3) riservatezza e governance dei dati, 4) trasparenza, 5) diversità, non discriminazione ed equità, 6) benessere sociale e ambientale e 7) accountability.

Si tratta di principi e requisiti il cui rispetto va verificato durante l'intero ciclo di vita dell'IA.

Gruppo Indipendente di Esperti, Orientamenti etici, 24-26: le misure per assicurare il rispetto di principi e requisiti includono: vincoli dell'architettura dell'IA, la scelta, ove possibile, di forme di explainable AI, l'effettuazione di test, l'adesione a standard tecnici, la volontaria sottoposizione a forme di certificazione, l'adozione di codici di condotta e di modelli di governance dell'IA.

Quali misure in concreto adottare dipende dal contesto.

Gruppo Indipendente di Esperti, Orientamenti etici, 7: "Tenuto conto della specificità contestuale dei sistemi di IA, l'attuazione dei presenti orientamenti deve essere adattata all'applicazione di IA specifica."

Gruppo Indipendente di Esperti, Orientamenti etici, 17: "Sebbene i requisiti siano tutti i pari importanza, al momento di applicarli in diversi campi e settori occorrerà tenere in considerazione il contesto e le potenziali tensioni che possono insorgere tra essi."

Gruppo Indipendente di Esperti, Orientamenti etici, 3-4: “Al di là dell’Europa i presenti orientamenti hanno inoltre l’obiettivo di promuovere la ricerca, la riflessione e la discussione su un quadro etico per i sistemi di IA a livello mondiale.”

Gruppo Indipendente di Esperti, Orientamenti etici, 6: “A nostro avviso l’Europa dovrebbe [...] divenire culla e leader della tecnologia etica e all’avanguardia”.

Gruppo Indipendente di Esperti, Orientamenti etici, 43-44: “L’Europa gode di un vantaggio esclusivo, che deriva dal suo impegno a porre il cittadino al centro delle proprie attività. Tale impegno è iscritto nel DNA stesso dell’Unione europea attraverso i trattati su cui si fonda. Il presente documento rientra in una visione che promuove un’IA affidabile la quale, a nostro avviso, dovrebbe costituire il presupposto su cui l’Europa può sviluppare la propria leadership nei sistemi di IA innovativi e all’avanguardia. Questa visione ambiziosa contribuirà a garantire la prosperità dei cittadini europei, sia a livello individuale che collettivo. Il nostro obiettivo è quello di creare una cultura dell’“IA affidabile per l’Europa”, che permetta a tutti di sfruttarne i vantaggi in un modo che garantisca il rispetto dei nostri valori fondamentali: i diritti fondamentali, la democrazia e lo Stato di diritto.”

Proposta di Regolamento che stabilisce regole armonizzate sull'intelligenza artificiale (Artificial Intelligence Act)

Proposta AIA, considerando 1: "Lo scopo del presente regolamento è migliorare il funzionamento del mercato interno istituendo un quadro giuridico uniforme in particolare per quanto riguarda lo sviluppo, la commercializzazione e l'uso dell'intelligenza artificiale (IA) in conformità ai valori dell'Unione. Il presente regolamento persegue una serie di motivi imperativi di interesse pubblico, quali un elevato livello di protezione della salute, della sicurezza e dei diritti fondamentali, e garantisce la libera circolazione transfrontaliera di beni e servizi basati sull'IA, impedendo così agli Stati membri di imporre restrizioni allo sviluppo, alla commercializzazione e all'uso di sistemi di IA [...]."

Proposta AIA, considerando 28: "è altresì opportuno tenere in considerazione, nel valutare la gravità del danno che un sistema di IA può provocare, anche in relazione alla salute e alla sicurezza delle persone, il diritto fondamentale a un livello elevato di protezione dell'ambiente."

Proposta AIA, considerando 6: “il presente regolamento contribuisce all’obiettivo dell’Unione di essere un leader mondiale nello sviluppo di un’intelligenza artificiale sicura, affidabile ed etica, come affermato dal Consiglio europeo, e garantisce la tutela dei principi etici, come specificamente richiesto dal Parlamento europeo.”

Proposta AIA, articolo 2: “Il presente regolamento si applica: (a) ai fornitori che immettono sul mercato o mettono in servizio sistemi di IA nell’Unione, indipendentemente dal fatto che siano stabiliti nell'Unione o in un paese terzo; (b) agli utenti dei sistemi di IA situati nell’Unione; (c) ai fornitori e agli utenti di sistemi di IA situati in un paese terzo, laddove l’output prodotto dal sistema sia utilizzato nell'Unione.”

Proposta AIA, considerando 10: “Al fine di garantire condizioni di parità e una protezione efficace dei diritti e delle libertà delle persone in tutta l’Unione, è opportuno che le regole stabilite dal presente regolamento si applichino ai fornitori di sistemi di IA in modo non discriminatorio, a prescindere dal fatto che siano stabiliti nell’Unione o in un paese terzo, e agli utenti dei sistemi di IA stabiliti nell’Unione.”

Proposta AIA, articolo 3: “Ai fini del presente regolamento si applicano le definizioni seguenti:
(1) ‘sistema di intelligenza artificiale’ (sistema di IA): un software sviluppato con una o più delle tecniche e degli approcci elencati nell’allegato I, che può, per una determinata serie di obiettivi definiti dall’uomo, generare output quali contenuti, previsioni, raccomandazioni o decisioni che influenzano gli ambienti con cui interagiscono;
(2) ‘fornitore’: una persona fisica o giuridica, un’autorità pubblica, un’agenzia o un altro organismo che sviluppa un sistema di IA o che fa sviluppare un sistema di IA al fine di immetterlo sul mercato o metterlo in servizio con il proprio nome o marchio, a titolo oneroso o gratuito […];
4) ‘utente’: qualsiasi persona fisica o giuridica, autorità pubblica, agenzia o altro organismo che utilizza un sistema di IA sotto la sua autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un’attività personale non professionale […].”

ALLEGATO I
TECNICHE E APPROCCI DI INTELLIGENZA ARTIFICIALE
di cui all'articolo 3, punto 1)

a) Approcci di apprendimento automatico, compresi l'apprendimento supervisionato, l'apprendimento non supervisionato e l'apprendimento per rinforzo, con utilizzo di un'ampia gamma di metodi, tra cui l'apprendimento profondo (*deep learning*);

b) approcci basati sulla logica e approcci basati sulla conoscenza, compresi la rappresentazione della conoscenza, la programmazione induttiva (logica), le basi di conoscenze, i motori inferenziali e deduttivi, il ragionamento (simbolico) e i sistemi esperti;

c) approcci statistici, stima bayesiana, metodi di ricerca e ottimizzazione.

Ai sensi dell'AIA, l'IA può essere
- vietata (articolo 5),
- ad alto rischio (articoli 6 e 7),
- non ad alto rischio (articolo 69).

IA vietata

Proposta AIA, articolo 5: "(1) Sono vietate le tecniche di intelligenza artificiale seguenti:
(a) l'immissione sul mercato, la messa in servizio o l'uso di un sistema di IA che utilizza tecniche subliminali che agiscono senza che una persona ne sia consapevole al fine di distorcerne materialmente il comportamento in un modo che provochi o possa provocare a tale persona o a un'altra persona un danno fisico o psicologico; [segue]"

Proposta AIA, articolo 5: “(b) l’immissione sul mercato, la messa in servizio o l’uso di un sistema di IA che sfrutta le vulnerabilità di uno specifico gruppo di persone, dovute all’età o alla disabilità fisica o mentale, al fine di distorcere materialmente il comportamento di una persona che appartiene a tale gruppo in un modo che provochi o possa provocare a tale persona o a un’altra persona un danno fisico o psicologico; (c) l’immissione sul mercato, la messa in servizio o l’uso di sistemi di IA da parte delle autorità pubbliche o per loro conto ai fini della valutazione o della classificazione dell’affidabilità delle persone fisiche per un determinato periodo di tempo sulla base del loro comportamento sociale o di caratteristiche personali o della personalità note o previste, in cui il punteggio sociale così ottenuto comporti il verificarsi di uno o di entrambi i seguenti scenari: (i) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di interi gruppi di persone fisiche in contesti sociali che non sono collegati ai contesti in cui i dati sono stati originariamente generati o raccolti; (ii) un trattamento pregiudizievole o sfavorevole di determinate persone fisiche o di interi gruppi di persone fisiche che sia ingiustificato o sproporzionato rispetto al loro comportamento sociale o alla sua gravità; [segue]”

Proposta AIA, articolo 5: “(d) l’uso di sistemi di identificazione biometrica remota ‘in tempo reale’ in spazi accessibili al pubblico a fini di attività di contrasto, a meno che e nella misura in cui tale uso sia strettamente necessario per uno dei seguenti obiettivi: (i) la ricerca mirata di potenziali vittime specifiche di reato, compresi i minori scomparsi; (ii) la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l’incolumità fisica delle persone fisiche o di un attacco terroristico […]”.

IA ad alto rischio

Proposta AIA, considerando 27: “È opportuno limitare i sistemi di IA identificati come ad alto rischio a quelli che hanno un impatto nocivo significativo sulla salute, la sicurezza e i diritti fondamentali delle persone nell’Unione, e tale limitazione riduce al minimo eventuali potenziali restrizioni al commercio internazionale.”

I sistemi di IA ad alto rischio sono di tre tipi:

- sistemi che sono componenti di sicurezza di un prodotto o essi stessi prodotti, sottoposti a una valutazione di conformità prima dell'immissione sul mercato europeo (articolo 6, 1° comma);

- sistemi indipendenti destinati a essere usati in uno dei settori di cui all'allegato III (articolo 6, 2° comma);

- sistemi indipendenti non usati in uno dei settori di cui all'allegato III, che presentino un elevato "un rischio di danno per la salute e la sicurezza, o un rischio di impatto negativo sui diritti fondamentali, che è, in relazione alla sua gravità e alla probabilità che si verifichi, equivalente o superiore al rischio di danno o di impatto negativo presentato dai sistemi di IA ad alto rischio di all'allegato III", qualora la Commissione li includa nell'elenco di cui all'allegato III (articolo 7, 1° comma).

Proposta AIA, considerando 30: “Per quanto riguarda i sistemi di IA che sono componenti di sicurezza di prodotti, o che sono essi stessi prodotti, e rientrano nell’ambito di applicazione di una determinata normativa di armonizzazione dell’Unione, è opportuno classificarli come sistemi ad alto rischio a norma del presente regolamento se il prodotto in questione è sottoposto alla procedura di valutazione della conformità con un organismo terzo di valutazione della conformità a norma della suddetta pertinente normativa di armonizzazione dell’Unione. Tali prodotti sono, in particolare, macchine, giocattoli, ascensori, apparecchi e sistemi di protezione destinati a essere utilizzati in atmosfera potenzialmente esplosiva, apparecchiature radio, attrezzature a pressione, attrezzature per imbarcazioni da diporto, impianti a fune, apparecchi che bruciano carburanti gassosi, dispositivi medici e dispositivi medico-diagnostici in vitro.”

Proposta AIA, Allegato III, Sistemi di IA ad alto rischio di cui all'articolo 6, paragrafo 2: "I sistemi di IA ad alto rischio a norma dell'articolo 6, paragrafo 2, sono i sistemi di IA elencati in uno dei settori indicati di seguito.

1. Identificazione e categorizzazione biometrica delle persone fisiche:

a) i sistemi di IA destinati a essere utilizzati per l'identificazione biometrica remota 'in tempo reale' e 'a posteriori' delle persone fisiche.

2. Gestione e funzionamento delle infrastrutture critiche:

a) i sistemi di IA destinati a essere utilizzati come componenti di sicurezza nella gestione del traffico stradale e nella fornitura di acqua, gas, riscaldamento ed elettricità.

3. Istruzione e formazione professionale:

a) i sistemi di IA destinati a essere utilizzati al fine di determinare l'accesso o l'assegnazione di persone fisiche agli istituti di istruzione e formazione professionale;

b)i sistemi di IA destinati a essere utilizzati per valutare gli studenti negli istituti di istruzione e formazione professionale e per valutare i partecipanti alle prove solitamente richieste per l'ammissione agli istituti di istruzione. [segue]"

Proposta AIA, Allegato III: “4. Occupazione, gestione dei lavoratori e accesso al lavoro autonomo:
a) i sistemi di IA destinati a essere utilizzati per l’assunzione o la selezione di persone fisiche […];
b) l’IA destinata a essere utilizzata per adottare decisioni in materia di promozione e cessazione dei rapporti contrattuali di lavoro, per l’assegnazione dei compiti e per il monitoraggio e la valutazione delle prestazioni […].
5. Accesso a prestazioni e servizi pubblici e a servizi privati essenziali e fruizione degli stessi:
a) i sistemi di IA destinati a essere utilizzati dalle autorità pubbliche […] per valutare l’ammissibilità delle persone fisiche alle prestazioni e ai servizi di assistenza pubblica […];
b) i sistemi di IA destinati a essere utilizzati per valutare l’affidabilità creditizia delle persone fisiche o per stabilire il loro merito di credito, a eccezione dei sistemi di IA messi in servizio per uso proprio da fornitori di piccole dimensioni;
c) i sistemi di IA destinati a essere utilizzati per inviare servizi di emergenza di primo soccorso […]. [segue]”

Proposta AIA, Allegato III: “6. Attività di contrasto:
a) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per effettuare valutazioni individuali dei rischi delle persone fisiche al fine di determinare il rischio di reato o recidiva in relazione a una persona fisica o il rischio per vittime potenziali di reati;
b) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto, come poligrafi e strumenti analoghi, o per rilevare lo stato emotivo di una persona fisica [...];
e) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per prevedere il verificarsi o il ripetersi di un reato effettivo o potenziale sulla base della profilazione delle persone fisiche [...] o per valutare i tratti e le caratteristiche della personalità o il comportamento criminale pregresso di persone fisiche o gruppi;
f) i sistemi di IA destinati a essere utilizzati dalle autorità di contrasto per la profilazione delle persone fisiche [...] nel corso dell’indagine, dell’accertamento e del perseguimento di reati; [...] [segue]”

Proposta AIA, Allegato III: “7. Gestione della migrazione, dell’asilo e del controllo delle frontiere: […]
b)i sistemi di IA destinati a essere utilizzati dalle autorità pubbliche competenti per valutare un rischio (compresi un rischio per la sicurezza, un rischio di immigrazione irregolare o un rischio per la salute) posto da una persona fisica che intende entrare o è entrata nel territorio di uno Stato membro; […].
8. Amministrazione della giustizia e processi democratici:
a) i sistemi di IA destinati ad assistere un’autorità giudiziaria nella ricerca e nell’interpretazione dei fatti e del diritto e nell’applicazione della legge a una serie concreta di fatti.

IA non ad alto rischio

Proposta AIA, articolo 69: “(1) La Commissione e gli Stati membri incoraggiano e agevolano l’elaborazione di codici di condotta intesi a promuovere l’applicazione volontaria ai sistemi di IA diversi dai sistemi di IA ad alto rischio dei requisiti di cui al titolo III, capo 2 […].”

Proposta AIA, articolo 9: “(1) In relazione ai sistemi di IA ad alto rischio è istituito, attuato, documentato e mantenuto un sistema di gestione dei rischi.
(2) Il sistema di gestione dei rischi è costituito da un processo iterativo continuo eseguito nel corso dell'intero ciclo di vita di un sistema di IA ad alto rischio, che richiede un aggiornamento costante e sistematico. […].”

Proposta AIA, articolo 10: “(1) I sistemi di IA ad alto rischio che utilizzano tecniche che prevedono l’uso di dati per l’addestramento di modelli sono sviluppati sulla base di set di dati di addestramento, convalida e prova che soddisfano i criteri di qualità di cui ai paragrafi da 2 a 5. […]
3. I set di dati di addestramento, convalida e prova devono essere pertinenti, rappresentativi, esenti da errori e completi. […]
(4) I set di dati di addestramento, convalida e prova tengono conto, nella misura necessaria per la finalità prevista, delle caratteristiche o degli elementi particolari dello specifico contesto geografico, comportamentale o funzionale all’interno del quale il sistema di IA ad alto rischio è destinato a essere usato. […]”

Proposta AIA, articolo 11: “(1) La documentazione tecnica di un sistema di IA ad alto rischio è redatta prima dell’immissione sul mercato o della messa in servizio di tale sistema ed è tenuta aggiornata. La documentazione tecnica è redatta in modo da dimostrare che il sistema di IA ad alto rischio è conforme ai requisiti di cui al presente capo e fornisce alle autorità nazionali competenti e agli organismi notificati tutte le informazioni necessarie per valutare la conformità del sistema di IA a tali requisiti.”

Proposta AIA, articolo 12: “(1) I sistemi di IA ad alto rischio sono progettati e sviluppati con capacità che consentono la registrazione automatica degli eventi (‘log’) durante il loro funzionamento. Tali capacità di registrazione sono conformi a norme riconosciute o a specifiche comuni.
(2) Le capacità di registrazione garantiscono un livello di tracciabilità del funzionamento del sistema di IA durante tutto il suo ciclo di vita adeguato alla finalità prevista del sistema. […]”

Proposta AIA, articolo 13: “(1) I sistemi di IA ad alto rischio sono progettati e sviluppati in modo tale da garantire che il loro funzionamento sia sufficientemente trasparente da consentire agli utenti di interpretare l’output del sistema e utilizzarlo adeguatamente. Sono garantiti un tipo e un livello di trasparenza adeguati, che consentano di conseguire il rispetto dei pertinenti obblighi dell’utente e del fornitore di cui al capo 3 del presente titolo.
(2) I sistemi di IA ad alto rischio sono accompagnati da istruzioni per l’uso in un formato digitale o non digitale appropriato, che comprendono informazioni concise, complete, corrette e chiare che siano pertinenti, accessibili e comprensibili per gli utenti.”

Proposta AIA, articolo 14: “(1) I sistemi di IA ad alto rischio sono progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da poter essere efficacemente supervisionati da persone fisiche durante il periodo in cui il sistema di IA è in uso.”

Proposta AIA, articolo 15: “(1) I sistemi di IA ad alto rischio sono progettati e sviluppati in modo tale da conseguire, alla luce della loro finalità prevista, un adeguato livello di accuratezza, robustezza e cibersicurezza e da operare in modo coerente con tali aspetti durante tutto il loro ciclo di vita.
(2) I livelli di accuratezza e le pertinenti metriche di accuratezza dei sistemi di IA ad alto rischio sono dichiarati nelle istruzioni per l’uso che accompagnano il sistema.
(3) I sistemi di IA ad alto rischio sono resilienti per quanto riguarda errori, guasti o incongruenze che possono verificarsi all’interno del sistema o nell’ambiente in cui esso opera, in particolare a causa della loro interazione con persone fisiche o altri sistemi.
[…]
(4) I sistemi di IA ad alto rischio sono resilienti ai tentativi di terzi non autorizzati di modificarne l’uso o le prestazioni sfruttando le vulnerabilità del sistema.”

I sistemi di IA ad alto rischio sono inoltre sottoposti a procedure indipendenti di valutazione della conformità (articoli 30-51), inseriti in un database europeo di schedatura dei sistemi di IA ad alto rischio (articolo 60) e continuamente monitorati, anche dopo la loro commercializzazione (articoli 61-68).

Proposta AIA, articolo 59: “(1) Ciascuno Stato membro istituisce o designa autorità nazionali competenti al fine di garantire l’applicazione e l’attuazione del presente regolamento.”

Proposta AIA, articolo 56: “(1) È istituito un ‘comitato europeo per l’intelligenza artificiale’ (il ‘comitato’). (2) Il comitato fornisce consulenza e assistenza alla Commissione al fine di: a) contribuire all’efficace cooperazione delle autorità nazionali di controllo e della Commissione […] ; b) coordinare e contribuire agli orientamenti e all’analisi della Commissione, delle autorità nazionali di controllo e di altre autorità competenti […]; c) assistere le autorità nazionali di controllo e la Commissione nel garantire l’applicazione uniforme del presente regolamento.”

Proposta AIA, articolo 53: “(1) Gli spazi di sperimentazione normativa per l’IA istituiti da una o più autorità competenti degli Stati membri o dal Garante europeo della protezione dei dati forniscono un ambiente controllato che facilita lo sviluppo, le prove e la convalida di sistemi di IA innovativi per un periodo di tempo limitato prima della loro immissione sul mercato o della loro messa in servizio […].”

L’approccio precauzionale adottato dall’AIA mira a tutelare le persone e a favorire le imprese e l’innovazione. Il rispetto dell’AIA, difatti, assicura la legittimità delle attività realizzate.

Inoltre, l’AIA mira ad assicurare alla UE la leadership nella produzione di norme di tutela del mercato globali. Le attività vietate e quelle ad alto rischio – ossia quelle principalmente regolate – sono definite con grande precisione. La più parte dell’IA, semplicemente, sfugge a quegli elenchi.

Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence

Article 3, Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence [2022]: “1. Member States shall ensure that national courts are empowered, either upon the request of a potential claimant who has previously asked a provider […] to disclose relevant evidence at its disposal about a specific high-risk AI system that is suspected of having caused damage, but was refused, or a claimant, to order the disclosure of such evidence from those persons.”

Article 3, Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence [2022]: “5. Where a defendant fails to comply with an order [...] to disclose or to preserve evidence at its disposal [...], a national court shall presume the defendant’s non-compliance with a relevant duty of care [...].”

Article 4, Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence: “1. [...] National courts shall presume [...] the causal link between the fault of the defendant and the output produced by the AI system or the failure of the AI system to produce an output, where all of the following conditions are met: (a) the claimant has demonstrated [...] the fault of the defendant [...]; (b) it can be considered reasonably likely, based on the circumstances of the case, that the fault has influenced the output produced by the AI system or the failure of the AI system to produce an output; (c) the claimant has demonstrated that the output produced by the AI system or the failure of the AI system to produce an output gave rise to the damage.”