CJN

Diritto Penale Contemporaneo

RIVISTA TRIMESTRALE

REVISTA TRIMESTRAL DE DERECHO PENAL
A QUARTERLY REVIEW FOR CRIMINAL JUSTICE

2/2022

Editore Associazione "Progetto giustizia penale", c/o Università degli Studi di Milano,
Dipartimento di Scienze Giuridiche "C. Beccaria" - Via Festa del Perdono, 7 - 20122 MILANO - c.f. 97792250157
ANNO 2022 - CODICE ISSN 2240-7618 - Registrazione presso il Tribunale di Milano, al n. 554 del 18 novembre 2011.
Impaginazione a cura di Chiara Pavesi

Diritto penale contemporaneo – Rivista trimestrale è un periodico on line ad accesso libero e non ha fine di profitto. Tutte le collaborazioni organizzative ed editoriali sono a titolo gratuito e agli autori non sono imposti costi di elaborazione e pubblicazione. La rivista, registrata presso il Tribunale di Milano, al n. 554 del 18 novembre 2011, è edita attualmente dall'associazione "Progetto giustizia penale", con sede a Milano, ed è pubblicata con la collaborazione scientifica e il supporto dell'Università Commerciale Luigi Bocconi di Milano, dell'Università degli Studi di Milano, dell'Università di Roma Tre, dell'Università LUISS Guido Carli, dell'Universitat de Barcelona e dell'Università Diego Portales di Santiago del Cile.

La rivista pubblica contributi inediti relativi a temi di interesse per le scienze penalistiche a livello internazionale, in lingua italiana, spagnolo, inglese, francese, tedesca e portoghese. Ogni contributo è corredato da un breve abstract in italiano, spagnolo e inglese.

La rivista è classificata dall'ANVUR come rivista scientifica per l'area 12 (scienze giuridiche), di classe A per i settori scientifici G1 (diritto penale) e G2 (diritto processuale penale). È indicizzata in DoGI e DOAJ.

La rivista può essere citata in forma abbreviata con l'acronimo: *DPC-RT*, corredato dall'indicazione dell'anno di edizione e del fascicolo.

La rivista fa proprio il Code of Conduct and Best Practice Guidelines for Journal Editors elaborato dal COPE (Committee on Publication Ethics).

La rivista si conforma alle norme del Regolamento UE 2016/679 in materia di tutela dei dati personali e di uso dei cookies (clicca qui per dettagli).

Ogni contributo proposto per la pubblicazione è preliminarmente esaminato dalla direzione, che verifica l'attinenza con i temi trattati dalla rivista e il rispetto dei requisiti minimi della pubblicazione.

In caso di esito positivo di questa prima valutazione, la direzione invia il contributo in forma anonima a due revisori, individuati secondo criteri di rotazione tra i membri dell'Editorial Advisory Board in relazione alla rispettiva competenza per materia e alle conoscenze linguistiche. I revisori ricevono una scheda di valutazione, da consegnare compilata alla direzione entro il termine da essa indicato. Nel caso di tardiva o mancata consegna della scheda, la direzione si riserva la facoltà di scegliere un nuovo revisore.

La direzione comunica all'autore l'esito della valutazione, garantendo l'anonimato dei revisori. Se entrambe le valutazioni sono positive, il contributo è pubblicato. Se una o entrambe le valutazioni raccomandano modifiche, il contributo è pubblicato previa revisione dell'autore, in base ai commenti ricevuti, e verifica del loro accoglimento da parte della direzione. Il contributo non è pubblicato se uno o entrambi i revisori esprimono parere negativo alla pubblicazione.

La direzione si riserva la facoltà di pubblicare, in casi eccezionali, contributi non previamente sottoposti alla procedura di peer review. Di ciò è data notizia nella prima pagina del contributo, con indicazione delle ragioni relative.

I contributi da sottoporre alla Rivista possono essere inviati al seguente indirizzo mail: editor.criminaljusticenetwork@gmail.com. I contributi che saranno ritenuti dalla direzione di potenziale interesse per la rivista saranno sottoposti alla procedura di peer review sopra descritta. I contributi proposti alla rivista per la pubblicazione dovranno rispettare i criteri redazionali scaricabili qui.

Diritto penale contemporaneo – Rivista trimestrale es una publicación periódica *on line*, de libre acceso y sin ánimo de lucro. Todas las colaboraciones de carácter organizativo y editorial se realizan gratuitamente y no se imponen a los autores costes de maquetación y publicación.
La Revista, registrada en el Tribunal de Milan, en el n. 554 del 18 de noviembre de 2011, se edita actualmente por la asociación "Progetto giustizia penale", con sede en Milán, y se publica con la colaboración científica y el soporte de la *Università Commerciale Luigi Bocconi* di Milano, la *Università degli Studi di Milano,* la *Università di Roma Tre*, la *Università LUISS Guido Carli*, la *Universitat de Barcelona* y la *Universidad Diego Portales de Santiago de Chile.*
La Revista publica contribuciones inéditas, sobre temas de interés para la ciencia penal a nivel internacional, escritas en lengua italiana, española, inglesa, francesa, alemana o portuguesa. Todas las contribuciones van acompañadas de un breve abstract en italiano, español e inglés.
El lector puede leer, compartir, reproducir, distribuir, imprimir, comunicar a terceros, exponer en público, buscar y señalar mediante enlaces de hipervínculo todos los trabajos publicados en "Diritto penale contemporaneo – Rivista trimestrale", con cualquier medio y formato, para cualquier fin lícito y no comercial, dentro de los límites que permite la licencia *Creative Commons - Attribuzione - Non commerciale 3.0 Italia* (CC BY-NC 3.0 IT) y, en particular, debiendo mantenerse la indicación de la fuente, el logo, el formato gráfico original, así como el autor de la contribución.
La Revista se puede citar de forma abreviada cn el acrónimo *DPC-RT*, indicando el año de edición y el fascículo.
La Revista asume el Code of Conduct and Best Practice Guidelines for Journal Editors elaborado por el COPE (*Comitte on Publication Ethics*).
La Revista cumple lo dispuesto en el Reglamento UE 2016/679 en materia de protección de datos personales (clica aquí para los detalles sobre protectión de la privacy y uso de cookies).
Todas las contribuciones cuya publicación se propone serán examinadas previamente por la Dirección, que verificará la correspondencia con los temas tratados en la Revista y el respeto de los requisitos mínimos para su publicación.
En el caso de que se supere con éxito aquella primera valoración, la Dirección enviará la contribución de forma anónima a dos evaluadores, escogidos entre los miembros del *Editorial Advisory Board*, siguiendo criterios de rotación, de competencia por razón de la materia y atendiendo también al idioma del texto. Los evaluadores recibirán un formulario, que deberán devolver a la Dirección en el plazo indicado. En el caso de que la devolución del formulario se retrasara o no llegara a producirse, la Dirección se reserva la facultad de escoger un nuevo evaluador.
La Dirección comunicará el resultado de la evaluación al autor, garantizando el anonimato de los evaluadores. Si ambas evaluaciones son positivas, la contribución se publicará. Si alguna de las evaluaciones recomienda modificaciones, la contribución se publicará después de que su autor la haya revisado sobre la base de los comentarios recibidos y de que la Dirección haya verificado que tales comentarios han sido atendidos. La contribución no se publicará cuando uno o ambos evaluadores se pronuncien negativamente sobre su publicación.
La Dirección se reserva la facultad de publicar, en casos excepcionales, contribuciones que no hayan sido previamente sometidas a *peer review*. Se informará de ello en la primera página de la contribución, indicando las razones.

Si deseas proponer una publicación en nuestra revista, envía un mail a la dirección editor.criminaljusticenetwork@gmail.com. Las contribuciones que la Dirección considere de potencial interés para la Revista se someterán al proceso de *peer review* descrito arriba. Las contribuciones que se propongan a la Revista para su publicación deberán respetar los criterios de redacción (se pueden descargar aquí).

Diritto penale contemporaneo – Rivista trimestrale is an on-line, open-access, non-profit legal journal. All of the organisational and publishing partnerships are provided free of charge with no author processing fees. The journal, registered with the Court of Milan (n° 554 - 18/11/2011), is currently produced by the association "Progetto giustizia penale", based in Milan and is published with the support of Bocconi University of Milan, the University of Milan, Roma Tre University, the University LUISS Guido Carli, the University of Barcelona and Diego Portales University of Santiago, Chile.

The journal welcomes unpublished papers on topics of interest to the international community of criminal scholars and practitioners in the following languages; Italian, Spanish, English, French, German and Portuguese. Each paper is accompanied by a short abstract in Italian, Spanish and English.

For citation purposes the journal's abbreviated reference format may be used: *DPC-RT,* indicating year of publication and issue.

The journal strictly adheres to the Code of Conduct and Best Practice Guidelines for Journal Editors drawn up by COPE (Committee on Publication Ethics).

The journal complies with the General Data Protection Regulation (EU) 2016/679 (GDPR) (click here for details on protection of privacy and use of cookies).

All articles submitted for publication are first assessed by the Editorial Board to verify pertinence to topics addressed by the journal and to ensure that the publication's minimum standards and format requirements are met.

Should the paper in question be deemed suitable, the Editorial Board, maintaining the anonymity of the author, will send the submission to two reviewers selected in rotation from the Editorial Advisory Board, based on their areas of expertise and linguistic competence. The reviewers are provided with a feedback form to compile and submit back to the editorial board within an established timeframe. If the timeline is not heeded to or if no feedback is submitted, the editorial board reserves the right to choose a new reviewer.

The Editorial Board, whilst guaranteeing the anonymity of the reviewers, will inform the author of the decision on publication. If both evaluations are positive, the paper is published. If one or both of the evaluations recommends changes the paper will be published subsequent to revision by the author based on the comments received and verification by the editorial board. The paper will not be published should one or both of the reviewers provide negative feedback.

In exceptional cases the Editorial Board reserves the right to publish papers that have not undergone the peer review process. This will be noted on the first page of the paper and an explanation provided.

If you wish to submit a paper to our publication please email us at editor.criminaljusticenetwork@gmail.com. All papers considered of interest by the editorial board will be subject to peer review process detailed above. All papers submitted for publication must abide by the editorial guidelines (download here).

Contents

Contents

Contents

L'intelligenza artificiale tra diritto e processo penale
La inteligencia artificial entre derecho y procesal penal
Artificial Intelligence between Criminal and Procedural Law

La responsabilità "penale" tra persona fisica e *corporation* alla luce della Proposta di Regolamento sull'Intelligenza Artificiale

La responsabilidad penal entre las personas físicas y jurídicas a la luz de la Propuesta de Reglamento sobre Inteligencia Artificial

"Criminal" Liability Between Human Beings and Corporations in Light of the Proposal of a Regulation on Artificial Intelligence

Camilla Minelli
Dottoressa in Giurisprudenza
camillaminelli@yahoo.it

Intelligenza artificiale, Principio di precauzione, Responsabilità da reato degli enti

Inteligencia artificial, Principio de precaución, Responsabilidad penal personas jurídicas

Artificial Intelligenc, Precautionary principle, Corporate Criminal Liability

Abstracts

La diffusione dell'Intelligenza artificiale impone una riflessione sui rischi connessi all'operare di sistemi autonomi, in ordine alla possibile verificazione di eventi dannosi. Se da un lato, infatti, sono destinati a emergere i limiti di tenuta del tradizionale diritto penale d'evento rispetto a eventuali ascrizioni di responsabilità colposa, dall'altro si avverte l'esigenza di arginare la verosimile tendenza a una diffusione della responsabilità lungo la catena di approvvigionamento del prodotto "intelligente". In tale contesto, si auspica la valorizzazione di un diritto di stampo proattivo e multilivello che sia in grado di intercettare i rischi governandoli, secondo un'impostazione che ispira anche la recente Proposta di Regolamento sull'Intelligenza artificiale adottata dalla Commissione europea nell'aprile 2021. Quest'ultima, pur non spiegando diretta efficacia in materia penale, sembra ritagliare aree di "rischio consentito" secondo una logica di precauzione "moderata".

La difusión de la Inteligencia Artificial exige reflexionar sobre los riesgos relacionados con el funcionamiento de los sistemas autónomos y la posible causación de hechos dañosos. En este tipo de casos seguramente se discutirá la concurrencia de los tradicionales delitos imprudentes. Sin embargo, existe la necesidad de establecer ciertos límites a la probable tendencia de extender la responsabilidad penal a lo largo de toda la cadena de suministro del producto "inteligente". En este contexto, sería deseable un derecho proactivo y multinivel que sea capaz de interceptar y controlar los riesgos, tal como lo hace la reciente Propuesta de Reglamento sobre Inteligencia Artificial adoptada por la Comisión Europea en abril de 2021. Esta última, aunque no es directamente eficaz en materia penal, parece delimitar zonas de "riesgo permitido" según una lógica de precaución "moderada".

The spread of Artificial Intelligence requires to reflect on the risks related to automation, as to potentially harmful outcomes. On the one hand, the features related to the traditional negligence-based result crimes are surely going to emerge; on the other hand, there is the need to limit the likely trend towards spreading liability in the supply chain of "intellingent" products. In the said framework, the paper sustains the idea of a proactive and multilevel regulation, to intercept and govern risks, in the same way as the recent Proposal of a Regulation on Artificial Intelligence adopted by the EU Commission in April 2021. Such a proposal, albeit without any direct effect in domestic criminal law matters, seems to cut out areas of "allowed risk" adopting a "moderate" precaution approach.

Sommario

1. Premessa: l'esigenza di una disciplina proattiva e multilivello.

La capacità di soddisfare aspettative di comportamento, alla base della rapida diffusione dell'intelligenza artificiale in quasi tutti i settori di vita quotidiana (a titolo di esempio, quello medico, giuridico e dei trasporti), si accompagna sempre alla possibilità di una loro delusione[1].

Proprio su tale prospettiva si innestano le criticità associate all'immissione nell'ambiente di sistemi autonomi in grado di apprendere e ai corrispondenti profili di responsabilità colposa, per danni da essi cagionati, ascrivibile ai soggetti a vario titolo coinvolti nel ciclo di vita dell'IA, laddove quest'ultima operi in qualità di vero e proprio agente adattivo e interattivo[2], piuttosto che in qualità di mero strumento agito dall'uomo a fini criminosi[3].

Attesa l'irrinunciabilità del ricorso al diritto penale a presidio di beni giuridici fondamentali, quali la vita o l'integrità fisica, potenzialmente vulnerati dalla verificazione di eventi riconducibili all'operare di sistemi autonomi, occorre dunque garantire che l'intero processo di imputazione sia informato al rispetto delle garanzie costituzionali al fine di scongiurare quelle "flessibilizzazioni" delle categorie classiche del reato già rilevate, come noto, con riguardo allo stravolgimento del paradigma colposo nelle dinamiche imposte dalla moderna "società del rischio"[4].

Da un lato, infatti, il fenomeno di *black box causale* che tradizionalmente marca il terreno della responsabilità per danno da prodotto è vieppiù aggravato dall'irriducibile opacità di alcuni "prodotti intelligenti", la cui imprevedibilità (non solo in situazioni per cui non è stata programmato un *output* adeguato, ma anche in risposta all'"esperienza" da essi maturata), costituisce tecnicamente "*a feature and not a bug*"[5]; dall'altro, per come la categoria della colpa è concepita dal diritto vivente, il programmatore risponderebbe per non aver previsto l'evento genericamente inteso e adottato le opportune cautele per evitarlo, purché esso costituisca un evitabile "logico sviluppo" (ad. es morte o lesioni) di tale imprevedibilità[6].

In tale ottica, se, da un lato, tecniche sanzionatorie dipendenti dalla concretizzazione di un evento lesivo potrebbero pregiudicarne l'effettività[7], dall'altro lato neppure si impone una retrocessione dell'evento a mera condizione obiettiva di punibilità, con eccessiva anticipazione

[1] Russell e Norvig (2009), p. 36 s.
[2] Pagallo e Quattrocolo (2018), p. 385.
[3] Cfr. il "considerando" AB della *Risoluzione del Parlamento europeo del 16 febbraio 2017 recante raccomandazioni alla Commissione concernenti norme di diritto civile sulla robotica* (2015/2103(INL)). Per un approfondimento si vedano Delogu (1974); Riondato (2014), p. 601; Magro (2019), p. 1207; Cappellini (2018) p. 505 s.; Gless *et al.* (2016), p. 412 s.
[4] L'espressione "flessibilizzazione" si rinviene per la prima volta a proposito dell'applicazione del principio di precauzione, nel processo relativo al petrolchimico di Porto Marghera, in Piergallini (2005), p.1684. *Contra,* Brusco (2012), p. 399 s.
[5] Millar e Kerr (2013), p. 107.
[6] Il riferimento è in particolare alla costruzione della cd. "colpa eventuale" operato dalla giurisprudenza che, tipicamente nei procedimenti legati a lesioni fisiche prodotte da esposizione professionale a sostanze tossiche e ad eventi megalici, avrebbe "bypassato" il pur complicato problema della causalità e scardinato l'impronta nomologica della riconoscibilità del rischio, rimodellando la prevedibilità e prevenibilità dell'evento e la sua ridescrizione. Cfr. Piergallini (2017), p. 235 s., 251 s; Magro (2019), p. 1207 s.; Magro (2020), p. 19. Sul punto si veda Piva (2022), p. 7, secondo cui «l'imprevedibilità pre-programmata dell'agente artificiale non consentirebbe di muovere un rimprovero per colpa (specie sotto forma di imperizia e sia pur eventualmente limitata alle ipotesi di gravità secondo il modello dell'art. 2236 c.c.), per lo più analogo a quello imposto dalla normativa sui diversi casi di divergenza tra voluto e realizzato (es. artt. 55, 83 o 116 c.p.)». Di questo avviso anche Massi (2022), p. 677, secondo cui, rispetto al produttore o programmatore, in posizione di garanzia del funzionamento del sistema, la rimozione dell'ostacolo alla concreta capacità di previsione di un certo evento di reato, costituito da meccanismi di disimpegno morale, non può essere raggiunto «accedendo all'idea di una "colpa eventuale"», la quale implicherebbe «l'accertamento di una prevedibilità di rischi esorbitanti rispetto a quelli consentiti *secondo un modello di prevedibilità puramente astratta,* con un conseguente impoverimento dell'accertamento dei contenuti soggettivi della colpa».
[7] Marinucci (2005), p.55 s. Specificamente, tale mancanza di effettività sussisterebbe «non solo e non tanto per i notori problemi di prova della causazione dell'evento […], quanto soprattutto perché l'evento […] fortunatamente si verifica ogni tanto, ma il *pericolo di tante altre morti o tante altre lesioni perdura minaccioso nel tempo*».

delle soglie di tutela, potendosi invece prospettare, come si vedrà, la valorizzazione di un diritto di stampo proattivo e multilivello in grado di intercettare i rischi governandoli[8].

Fermo restando che una valutazione dell'idoneità degli esistenti modelli di responsabilità a fronteggiare tali rischi non potrà fondarsi su «divieti *tout court* di produzione di agenti intelligenti» che rischierebbero di paralizzare l'innovazione nel settore, «se non quando il pericolo sia tale da superare a monte ogni beneficio atteso», ma presupporrà «un'articolata delimitazione di aree di "rischio consentito" presidiate da un apparato di sanzioni adeguato e funzionale alla tutela degli interessi coinvolti»[9].

2. La proposta di Regolamento sull'Intelligenza artificiale: la logica della precauzione "moderata".

In tale direzione, all'esito di un dibattito etico e giuridico da tempo avviato in seno alle istituzioni europee e internazionali sui principi guida dello sviluppo dell'intelligenza artificiale, sembrerebbe essersi mossa la Commissione europea con la recente Proposta di Regolamento sull'Intelligenza artificiale[10] (d'ora in poi, per brevità, anche proposta di Regolamento) che, pur non spiegando diretta efficacia in materia penale, sembra ritagliare aree di "rischio consentito" secondo una logica di precauzione per così dire "moderata"[11].

Da un punto di vista oggettivo, ai fini dell'applicazione delle nuove regole concernenti l'immissione sul mercato, la messa in servizio e l'utilizzo di sistemi IA, l'art. 3 fornisce una definizione di "sistema di intelligenza artificiale" in termini di «*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».

In secondo luogo, l'art. 2 circoscrive soggettivamente l'ambito di applicazione ai fornitori[12] che immettono sul mercato o mettono in servizio sistemi di IA all'interno dell'Unione (indipendentemente dal fatto che essi abbiano o meno la propria sede nel territorio UE ovvero in uno Stato terzo), agli utenti[13] di sistemi di IA collocati nell'Unione, nonché a fornitori e utenti di sistemi che, pur non sviluppati o prodotti all'interno dell'Unione, generano un *output* impiegato nell'UE (par. 1).

La proposta introduce una disciplina speciale nel quadro normativo sulla sicurezza generale dei prodotti (cd. GPSD o *General Product Safety Directive*) che costituisce attualmente la cd. legislazione orizzontale in materia, di natura integrativa rispetto a quella verticale[14] e che, secondo quanto affermato dalla stessa Proposta di Regolamento, continuerà ad applicarsi come «rete di sicurezza» al fine di garantire che i «sistemi di IA collegati a prodotti che non sono ad alto rischio in conformità al presente regolamento e che pertanto non sono tenuti a rispettare i requisiti ivi stabiliti siano comunque sicuri al momento dell'immissione sul mercato o della messa in servizio»[15].

[8] La Vattiata (2022), p. 710.
[9] Così Piva (2022), p. 683 s.
[10] *Proposta di Regolamento del Parlamento e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale (legge sull'intelligenza artificiale) e modifica alcuni atti legislativi dell'Unione*, 21.4.2021, COM (2021) 206 final.
[11] In tema di vedano Giunta (2006), p. 227 s.; Brusco (2012), p. 400.
[12] Per "fornitore" si intende «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» (art. 3 par. 2).
[13] Per "utente" si intende «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» (art. 3 par. 4).
[14] Si vedano ad esempio le direttive in materia di sicurezza dei presidi medico-chirurgici (dir. 2007/47/CE); farmaci (dir. 2001/83/CE); materiale elettrico (dir. 2006/95/CE); autoveicoli (dir. 2007/46/CE); motoveicoli a due e tre ruote (dir. 2002/24/CE); pneumatici (dir. 1992/23/CE; dir. 2001/43/CE; dir. 2005/11/CE); giocattoli (dir. 2009/48/CE). La GPSD, oggetto peraltro di una recentissima proposta di Regolamento da parte della Commissione europea (cd. GPSR), prevede infatti che, ove i prodotti siano soggetti ai requisiti di sicurezza prescritti da specifica normativa comunitaria, l'applicabilità della direttiva in esame sia limitata unicamente agli aspetti (quali il richiamo o il ritiro dei prodotti) non regolati da normative specifiche e per i rischi (o categorie di rischi) non previsti e disciplinati dai suddetti requisiti di sicurezza. Così Carnevali (2006), p. 2881 s., 2888 s.; Al Mureden (2017), p. 10;
[15] Cfr. l'82º "considerando" della Proposta di Regolamento. Nello stesso senso, la Proposta di Regolamento sulla sicurezza dei prodotti che, nella sezione dedicata a "eventuali sinergie con altri strumenti pertinenti", dispone: «Quadro orizzontale sull'intelligenza artificiale (IA): mira a concentrarsi sulle applicazioni ad alto rischio. Di conseguenza, e per quanto riguarda la sicurezza dei prodotti, funzionerà come la legislazione settoriale, stabilendo requisiti specifici per le applicazioni di IA, e la presente proposta fungerà da "rete di sicurezza" per i prodotti e gli aspetti non contemplati da altre legislazioni settoriali, per fornire una base giuridica per il ritiro di tali prodotti al fine di garantire una

Occorre dunque premettere che la strategia impiegata dal legislatore europeo nel perseguimento della sicurezza dei prodotti si fonda sulla determinazione, a livello di direttive, dei requisiti essenziali di sicurezza che i prodotti devono soddisfare per poter circolare liberamente nel mercato europeo[16], con relativo *onus probandi* in capo al fabbricante. È dunque introdotta una presunzione di conformità a tali requisiti in caso di prodotto conforme alle norme tecniche armonizzate (o, in mancanza di queste, alle norme tecniche nazionali)[17], su cui le industrie si basano per produrre e immettere in mercato prodotti dotati dei requisiti essenziali fissati dalle direttive.

Ebbene, pur dovendosi ancora adeguatamente esplorare i profili di reciproca integrazione tra le disposizioni di un possibile Regolamento e il quadro normativo nazionale e sovranazionale che eventualmente venga in considerazione[18], l'adozione della Proposta di Regolamento segna un'inversione di tendenza rispetto all'approccio finora seguito nella (non) regolamentazione dell'intelligenza artificiale[19], contrassegnata da una preponderanza di strumenti di *soft law*[20] rispetto alla ben più ridotta produzione di norme vincolanti, riservata peraltro a settori molto specifici[21]. Ciò in quanto tali strumenti, così come le norme etiche e tecniche, non appaiono di per sé sufficienti a garantire l'armonizzazione e il conseguimento di beni e obiettivi comuni, lasciando inalterata l'esigenza di ricorrere all'*hard law*, seppur limitatamente ad una funzione di "guida" sovranazionale che riservi un margine di apprezzamento statale, al fine di apprestare una cornice normativa unica a livello europeo che favorisca la creazione di un ecosistema di fiducia attorno all'intelligenza artificiale, secondo uno *human-centric approach*[22].

Nel disciplinare i vari aspetti del ciclo di vita dei sistemi IA, la Proposta di Regolamento adotta così un approccio *risk-based*, vietando innanzitutto l'immissione sul mercato, la messa in servizio o l'uso di sistemi che presentino rischi "inaccettabili"[23]. Mentre dunque il regolamento non trova applicazione in contesti di rischio "minimo" (in cui attualmente rientra, a ben vedere, gran parte dei sistemi AI), in presenza di un rischio "limitato" (per sistemi IA come i *chatbot*) si impongono specifici obblighi di trasparenza in termini di segnalazione dell'utilizzo di IA nell'interazione con un essere umano, sollecitandosi l'adozione su base volontaria di codici di condotta.

Particolare attenzione è riservata alla cd. IA ad alto rischio, individuata in quei sistemi IA utilizzabili come componenti di sicurezza[24] di prodotti disciplinati dalla normativa di ar-

protezione efficace dei consumatori».

[16] Intesi sia come caratteristiche costruttive (ad esempio, il montaggio delle protezioni fisse della macchina attraverso sistemi che richiedano l'impiego di strumenti per l'apertura) che come rischi da evitare (ad esempio, una costruzione che eviti o riduca la formazione di cariche elettrostatiche pericolose). Così Carnevali (2006), p. 2885.

[17] Con "norme tecniche" si intendono quelle regole di carattere tecnico-scientifico elaborate da organismi privati denominati «enti di normalizzazione» (UNI e CEI in Italia, DIN in Germania, BSI in Gran Bretagna ecc…), che costituiscono "lo stato dell'arte" e sono costantemente aggiornate secondo lo stato delle conoscenze tecniche e scientifiche, e nel cui ambito si distinguono le norme tecniche armonizzate, ossia quelle elaborate da organismi di normalizzazione europei (CEN e CENELEC) e a loro volta trasposte in norme tecniche nazionali a cura dei vari enti nazionali di normalizzazione. Così Carnevali (2006), p. 2883 s.

[18] In tal senso La Vattiata (2021), p. 15.

[19] Sottolineava la mancanza di «un unico e coerente quadro legislativo che disciplini il campo della progettazione, costruzione e impiego di sistemi d'IA, posta anche la difficoltà di individuare una base etica condivisa di standard rilevanti per la sicurezza» Magro (2019), p. 1207 s.

[20] Si vedano, a livello ultrastatale, le comunicazioni, risoluzioni e piani di azione approvati dalle istituzioni dell'UE, nonché i principi approvati dall'OCSE, ovvero le linee guida adottate da organi come l'European Data Protection Board per sviluppare le disposizioni del GDPR; a livello nazionale, si consideri la diffusione di atti di *soft law* espressivi di programmi e indirizzi comprensivi, in Italia tradottisi nel Libro Bianco Agid del 2018 e nelle "Proposte per una Strategia italiana per l'Intelligenza Artificiale" elaborate nel 2019 dal Gruppo di esperti del MISE sull'intelligenza artificiale. In tema si veda, *amplius,* Mobilio (2020), p. 401 s., 422.

[21] Quale il settore dei mercati finanziari, in cui la Direttiva 2014/65/UE impone obblighi specifici in capo agli operatori che svolgono "negoziazioni algoritmiche" (art. 17); ovvero quello della sicurezza aerea (Reg. 2018/1139/UE) per quanto concerne specificamente la registrazione, certificazione e regole generali di condotta per gli operatori di droni; ancora, indirettamente, quello della protezione dei dati personali, da ultimo disciplinato con il GDPR che regola il trattamento delle principali categorie di dati che alimentano i sistemi di IA.

[22] Così si esprimono il Gruppo di esperti ad alto livello sull'intelligenza artificiale (2018); Mobilio (2020), p. 424; Fiorella (2022), p. 652 s.

[23] Vi rientrano, in particolare, sistemi e applicazioni di IA in grado di manipolare il comportamento umano attraverso tecniche subliminali nell'inconsapevolezza dell'utente o di sfruttare la vulnerabilità di gruppi specifici (ad esempio giocattoli che utilizzano l'assistenza vocale per incoraggiare i comportamenti pericolosi del minore), ovvero quelli che consentono ai governi di attribuire un "punteggio sociale", nonché i sistemi di identificazione biometrica remota il cui utilizzo in tempo reale ai fini di contrasto in spazi accessibili al pubblico è in linea di principio vietato, salvo poche eccezioni rigorosamente individuate e disciplinate (ad esempio, ove strettamente necessario per cercare un minore scomparso, prevenire un'imminente minaccia terroristica o rilevare, identificare o perseguire autori o sospettati di un reato grave). Tale uso è comunque subordinato all'autorizzazione di un organo giudiziario o di altro organo indipendente e a limitazioni temporali, geografiche e personali.

[24] Con "componente di sicurezza di un prodotto o di un sistema" si intende «un componente di un prodotto o di un sistema che svolge una funzione di sicurezza per tale prodotto o sistema o il cui guasto o malfunzionamento mette in pericolo la salute e la sicurezza di persone o beni».

monizzazione dell'Unione elencata nell'allegato II e soggetti a valutazione di conformità *ex ante* da parte di terzi, nonché in altri sistemi con forti implicazioni sui diritti fondamentali elencati nell'allegato III, salvo integrazioni (art. 7)[25]: si tratta di sistemi IA i cui rischi si siano già concretizzati ovvero è probabile che si concretizzino in un futuro prossimo e per i quali l'acquisizione della marcatura CE, a seguito di valutazione di conformità ai fini dell'immissione in commercio, è subordinata all'osservanza di requisiti "minimi", di qualità dei *data-set* che alimentano il sistema (art. 10) [26], documentazione (art. 11), registrazione degli eventi (art. 12) e trasparenza (art. 13), funzionale alla valutazione dei rischi *ex ante* (art. 9) e alla sorveglianza umana *ex post* (art. 14), oltre che alla prevenzione di discriminazioni, nonché di affidabilità (art. 15). Ciò in un più ampio contesto di adeguamento a obblighi cogenti (artt. 16-29) sotto il controllo delle competenti autorità nazionali di vigilanza del mercato, coadiuvate da un Comitato europeo per l'IA di prossima istituzione[27].

In tale ambito, si invitano i fornitori a istituire e certificare un sistema di gestione del rischio (art. 9), basato sullo stato dell'arte generalmente riconosciuto e comprensivo di un *post-market monitoring system*[28], destinato ad assumere portata decisiva nell'assicurare un efficace e tempestivo contrasto dei rischi emergenti in contesti di apprendimento continuo anche a seguito dell'immissione in commercio, ferma restando la necessità, ove intervengano mutamenti sostanziali nel ciclo di vita del sistema IA, di rinnovare la valutazione di conformità[29].

Tale meccanismo garantisce una raccolta, documentazione e analisi sistematica dei dati rilevanti complessivamente ricavati sulle *performance* dei sistemi ad alto rischio durante il ciclo di vita, consentendo una valutazione costante della loro conformità ai requisiti stabiliti nel Capo 2 del Titolo III e, in caso, l'immediata attivazione di misure correttive, compresi il ritiro dal commercio e la notifica alle autorità di sorveglianza degli Stati membri di eventuali incidenti o malfunzionamenti[30].

In conclusione, la Commissione ha invitato gli Stati membri a garantire il rispetto delle previsioni del Regolamento introducendo un sistema che, in caso di violazioni dello stesso, contempli l'irrogazione di sanzioni (anche amministrative) effettive, proporzionate e dissuasive, pur tenendo particolarmente in conto gli interessi delle PMI in un'ottica di sostenibilità economica (art. 71).

[25] Si tratta, in particolare, di quelli in cui l'IA è impiegata nell'identificazione e categorizzazione biometrica delle persone fisiche, istruzione e formazione professionale, occupazione, gestione dei lavoratori e accesso al lavoro autonomo, accesso a prestazione e servizi pubblici e servizi privati essenziali e relativa fruizione, attività di contrasto, gestione della migrazione, dell'asilo e del controllo delle frontiere e amministrazione della giustizia e processi democratici. Si noti che nella relazione trasmessa al Parlamento ai sensi dell'art. 6, comma 4, della legge n. 234 del 2012, il Governo, nell'ambito delle prospettive negoziali e delle eventuali modifiche ritenute necessarie od opportune, ha sottolineato l'elasticità del perimetro di applicazione del nuovo regime (in quanto modificabile attraverso atti delegati) e la necessità di valutare il rischio di incertezza giuridica e di "delega in bianco" alla Commissione. Cfr. *Legge sull'intelligenza artificiale*, Camera dei deputati, Ufficio rapporti con l'UE, *Dossier* 57, 12 novembre 2021.

[26] A tal riguardo, la Proposta distingue tra "dati di addestramento" (i dati utilizzati per addestrare un sistema di IA adattandone i parametri che può apprendere, compresi i pesi di una rete neurale), "dati di convalida" (utilizzati per fornire una valutazione del sistema di IA addestrato e per metterne a punto, tra l'altro, i parametri che non può apprendere e il processo di apprendimento, al fine di evitare l'eccessivo adattamento ai dati di addestramento, cd. o*verfitting*, considerando che il set di dati di convalida può essere un set di dati distinto o essere costituito da una partizione fissa o variabile del set di dati di addestramento) e "dati di prova" (utilizzati per fornire una valutazione indipendente del sistema di IA addestrato e convalidato al fine di confermarne le prestazioni attese prima della sua immissione sul mercato o messa in servizio).

[27] Quest'ultimo avrà compiti di assistenza e consulenza nei confronti della Commissione che includeranno la raccolta e condivisione delle conoscenze e migliori pratiche tra gli Stati membri oltre che l'emanazione di pareri, raccomandazioni, consulenze o orientamenti su questioni relative all'attuazione del Regolamento, comprese le specifiche tecniche o le norme esistenti per quanto riguarda i requisiti stabiliti nel Regolamento. Cfr. artt. 56-58 della Proposta di Regolamento, nonché il 76° "considerando".

[28] Ex art. 3 (25) per "monitoraggio successivo all'immissione sul mercato" si intendono «tutte le attività svolte dai fornitori di sistemi di IA al fine di raccogliere e analizzare in modo proattivo l'esperienza maturata tramite l'uso dei sistemi di IA che immettono sul mercato o che mettono in servizio, al fine di individuare eventuali necessità di immediate azioni correttive o preventive». Si vedano anche l'84° "considerando" nonché l'art. 71.

[29] Si noti, peraltro, il limite apposto dal 66° "considerando" all'apprendimento continuo dopo l'immissione nel mercato con riguardo a eventuali mutazioni degli algoritmi e delle loro performance rispetto allo status presente al momento della valutazione di conformità.

[30] Cfr. il 78° "considerando", nonché gli artt. 62-67. Tale notifica dovrà essere inoltrata non appena sia stato individuato un nesso causale tra il sistema di AI e l'incidente o il malfunzionamento ovvero una probabilità ragionevole che tale nesso sussista e, in ogni caso, non più tardi di 15 giorni da quando il *provider* ne sia venuto a conoscenza. Una volta ricevuta la notizia di una violazione degli obblighi imposti dall'ordinamento comunitario l'autorità di sorveglianza del mercato dovrà informare le autorità pubbliche nazionali o gli organismi di cui all'art. 64.

3. Dal *machina delinquere potest* al paradigma della responsabilità da prodotto difettoso: l'ipotesi del *black box* decisionale.

A fronte di eventi dannosi riconducibili a un difetto del prodotto "intelligente", si è già percorsa la prospettiva, sviluppatasi sulla scorta del superamento del dogma *societas delinquere non potest,* di una responsabilizzazione diretta della macchina[31], ove si adotti un approccio di tipo funzionalistico che intenda la responsabilità penalistica quale prodotto dell'interazione sociale rivolto, tramite irrogazione di sanzioni, a soddisfare aspettative di stabilizzazione della società[32]. Tale soluzione, tuttavia, fondando un'eventuale imputazione sul riconoscimento normativo di una vera e propria "capacità" della macchina di assecondare o deludere aspettative normative, sembrerebbe, almeno allo stato, non praticabile, in quanto postulerebbe un grado di pervasività e di accettazione dell'IA ancora difficilmente riscontrabile.

Viceversa, ove si considerino i soggetti umani "intorno" alla macchina come autori "mediati" di fatti alle stesse riconducibili, sono destinati a emergere, da un lato, i limiti di tenuta del tradizionale diritto penale d'evento rispetto a eventuali ascrizioni di responsabilità colposa[33] e, dall'altro, l'esigenza di arginare la verosimile tendenza a una diffusione della responsabilità lungo la catena di approvvigionamento[34]. Al riguardo, è impossibile non notare come, se in passato il diritto penale si è confrontato col danno da prodotto difettoso sulla scorta delle rinomate categorie di derivazione civilistica del difetto da progettazione, da fabbricazione e da informazione[35], rispetto ai *software* sia effettivamente fuorviante parlare di vera e propria "difettosità"[36], risultando forse più appropriato il riferimento ai concetti di affidabilità, ossia «la misura di quanto un sistema sia in grado di comportarsi secondo ciò che è stabilito nelle sue specifiche»[37] e di "esito disvoluto del processo di *machine learning*", distinguibile dal "tradizionale" difetto da rischio di sviluppo[38].

Peraltro, l'azione causativa di un danno in contesti di intelligenza artificiale potrebbe essere ricondotta tanto all'utilizzatore, quanto a un difetto di programmazione, di costruzione, o di informazione, eventualmente interagenti *pro quota* alla stregua di concause (art. 41, commi 1

[31] In tema si vedano Hallevy (2015); Hallevy (2013); Hallevy (2010); Chopra e White (2011), p. 169.

[32] In questi termini si esprimono Piva (2022), p. 690 s. e, precedentemente, già Simmler e Markwalder (2019). *Contra,* Gless *et al.* (2016) nonché, nella dottrina italiana, Borgogno (2022), p. 741.

[33] In tal senso D. Piva (2022), pp. 684 s., secondo cui «Le più evolute tecniche di *machine learning* sembrano tendenzialmente incompatibili col tradizionale paradigma della responsabilità vicaria o indiretta che postula la riferibilità al programmatore di una condotta realizzata tramite algoritmi preimpostati in un mero strumento, nella logica dell'autore mediato (al pari di quella cui rispondono, ad esempio, gli artt. 46, 54 u.c., 86, 111 c.p.)».

[34] Definita «*The problem of many hands*» in Expert Committee on human rights dimensions of automated data processing and different forms of artificial intelligence (MSI-AUT), *Council of Europe Study. Responsibility and IA,* in www.coe.it, DGI(2019)05; così anche Beck (2018), p. 41 s.; Beck (2017), p. 245.

[35] Per un'accurata ricostruzione dei difetti dei prodotti, seppur in riferimento all'art. 5 del D.P.R 224/1988, si veda Carnevali (1999), p. 938 s.; Piergallini (2004), p. 46 s.

[36] Di particolare complessità si rivela la definizione di "*software fault*". Segnatamente, il fallimento (*failure*) di un sistema consiste nella non corrispondenza dei servizi offerti dallo stesso alle specifiche definite in fase di sviluppo, mentre l'errore (*error*) è la parte del sistema che causa il fallimento, a sua volta provocato da un guasto (*fault*) (distinguibile in attivo ove generi un errore, altrimenti dormiente). È importante ricordare che un fallimento va sempre rapportato a quella che è la funzione del sistema e non a quanto scritto nelle specifiche, per cui è possibile che il sistema, nonostante soddisfi le specifiche iniziali, non soddisfi le aspettative del cliente (o, per meglio dire, le aspettative di chi l'ha creato). Per un approfondimento si vedano Randell *et al.* (1978), p. 123 s.; Gray (1991), p. 39 s.; Munsona *et al.* (2006), p. 327 s.; Chiara (2019).

[37] Cit. Randell *et al.* (1978). Tale obiettivo può essere perseguito tramite le cd. tecniche di tolleranza ai guasti, tipicamente finalizzate a correggere errori generalmente derivanti da un'errata fase di progettazione del *software,* nonché a implementare la robustezza del sistema, ossia la capacità del sistema di comportarsi ragionevolmente anche a fronte di situazioni inaspettate. A tal proposito, la Proposta di Regolamento prevede che i sistemi garantiscano un adeguato livello di accuratezza, cibersicurezza e robustezza (*i.e.* resilienza rispetto sia ai rischi connessi alle limitazioni del sistema, quali errori, guasti, incoerenze, situazioni impreviste, sia alle azioni dolose che possono comprometterne la sicurezza e determinare comportamenti dannosi o altrimenti indesiderati), con una specifica attenzione ai sistemi che proseguono il loro apprendimento a seguito dell'immissione nel mercato, il cui sviluppo dovrà garantire che gli output potenzialmente distorti a causa dei cd. *feedback loops* siano oggetto di appropriate misure di attenuazione. Cfr. *amplius*, art. 15 e i "considerando" 49-51.

[38] Così, Piva (2022), p. 686. In particolare, tali esiti disvoluti possono derivare da una duplice serie di fattori in grado di pregiudicare l'affidabilità del sistema. Innanzitutto, dal fatto che il modello non sia in grado di svolgere bene il compito in condizioni considerate normali per l'uomo; in secondo luogo, dalla circostanza che il modello, pur funzionando bene, presenti delle vulnerabilità che possono condurre a malfunzionamenti *fisiologici* tra cui spiccano il cd. *overfitting* (ossia situazioni in cui il modello non apprende alcun *pattern* significativo, ma si limita a memorizzare i dati di *input,* riducendo notevolmente il potere di generalizzazione del modello), in merito al quale si sottolinea l'importanza di una validazione esterna, indipendente dalla fase di *training*; ovvero rischi di *bias* nello spettro, ossia la presenza di esempi nel *dataset* che non riflettono la diversità e la complessità delle situazioni. A tal proposito la Proposta di Regolamento fissa degli standard di qualità per i *datasets* impiegati per l'addestramento e la validazione dei sistemi, in termini di sufficiente completezza, rappresentatività e immunità da errori, al fine di garantire un rendimento sicuro e conforme alla destinazione d'uso. Cfr. più ampiamente l'art. 10.

e 3 c.p.)[39], nonché a fattori ambientali esterni[40] (si pensi all'ipotesi in cui si sia verificata una condizione di scarsa visibilità sia per i sensori di una *self-driving* car che per il guidatore), nel contesto di una vera e propria *web of causation*[41] attraverso cui la ridescrizione dell'evento così come occorso *hic et nunc* e la ricostruzione delle rispettive sfere di responsabilità potrà semmai essere agevolata dall'implementazione di "scatole nere" nel sistema[42].

Ove si accerti un "esito disvoluto del processo di *machine learning*", tuttavia, è da escludersi l'idoneità del cd. fattore robotico ad assurgere a manifestazione di caso fortuito o forza maggiore (artt. 45-46 c.p.)[43] ovvero a integrare un fattore interruttivo del nesso causale tra la condotta del programmatore e l'evento dannoso ai sensi dell'art. 41 cpv. c.p., che abbia rilasciato un rischio qualitativamente diverso da quello iniziale[44] nei termini dell'attivazione di «un rischio nuovo e incommensurabile, del tutto incongruo rispetto al rischio originario attivato dalla prima condotta»[45]; tanto più ove si consideri la tradizionale reticenza mostrata dalla giurisprudenza nel riconoscere cause sopravvenute "autosufficienti", «che assumano, cioè, interamente su di sé il "peso" della determinazione dell'evento»[46]. Inoltre, anche a voler qualificare il comportamento di un sistema di IA come "comportamento di un terzo" che si sia interposto tra la condotta del produttore e l'evento penalmente rilevante, è anche noto che la frapposizione della condotta colposa del terzo nel decorso dell'iter causale non determina «*tout court* l'interruzione del rapporto causale come conseguenza di una netta recisione del nesso eziologico per la creazione non volontaria di un rischio autonomo […]»[47].

Ad ogni modo, deve preliminarmente rilevarsi la tendenziale risolvibilità del problema di cd. *black box* decisionale[48] (*i.e.* una condizione, determinata dall'inaccessibilità di elementi dell'algoritmo o del modello, di opacità dei processi interni che hanno condotto, dato un determinato *input*, all'*output* osservabile) attraverso le soluzioni tecniche fornite dal settore e secondo un approccio di cd. interpretabilità *post-hoc* di modelli a scatola nera[49]: cosicché,

[39] Mutuando dai più recenti studi in materia di Mezzetti (2021), p. 30 s., secondo cui «il legislatore del 1930 evidentemente pensava ad un autore singolo o che avesse agito eventualmente in concorso, allorché dettava le norme sul rapporto di causalità […] nell'attuale selva degli antecedenti ed interferenze causali la (con-)causazione è fenomeno normale che bene descrive la multifattorialità e complessità degli antecedenti causali nella moderna società del rischio». In tale contesto, se è vero che «il progresso delle scoperte scientifiche ha contribuito a "specializzare" l'osservazione di tutti i possibili antecedenti causali di un determinato evento concretamente realizzatosi, incrementando il novero delle "possibili", "astratte" cause, e naturali e normative, dell'evento tipico, cosa evidentemente fuori dalla portata euristica di un osservatore di qualche decennio fa, ciò non toglie che la ricerca continua dell'esclusività dell'antecedente causale «da *solo* sufficiente a determinare l'evento» rimane un problema aperto di sconcertante attualità».

[40] Con la precisazione che i sistemi IA ad uso "civile", a differenza da quelli ad uso militare, sono pensati per operare in ambienti cooperativi, che favoriscono la prevedibilità e dunque la controllabilità dei loro comportamenti. In tal senso, Amoroso e Tamburrini (2019), p. 37 s.

[41] L'espressione è tratta da Piergallini (2020), p. 1762.

[42] Si veda in tal senso l'art. 12 della Proposta di Regolamento, che introduce un obbligo di sviluppo dei sistemi autonomi secondo tecniche che consentano una registrazione automatica degli eventi (cd. *log*), al fine di garantire la tracciabilità del sistema durante l'intero ciclo di vita. Tuttavia, non è consentito, per ciò solo, ritenere il problema causale risolto, se si considera, come rilevato da Preziosi (2021), p. 301, che nelle situazioni in cui conosciamo esattamente la serie degli avvenimenti «riteniamo, erroneamente (dal punto di vista causale), che ogni antecedente della predetta serie sia una condizione necessaria dell'evento. Ma questo discende unicamente dalla circostanza che, conoscendo la serie degli antecedenti, sono i fatti che sembrano incaricarsi di indicare la causa, non è la *causa* a farci conoscere i fatti, come avviene, invece, quando il decorso causale si presenta oscuro, non lo conosciamo in modo evidente».

[43] Così Magro (2019), p. 1179; Magro (2018).

[44] In tal senso Piergallini (2020), p.1760. Da ultimo, in tema di interruzione del nesso causale, si vedano Mezzetti (2021), p. 4, secondo cui «le cause sopravvenute sono in rapporto di dipendenza dalle antecedenti, di cui rappresentano uno sviluppo *esorbitante* ed *incontrollato* verso *quell'*evento che si è concretamente realizzato. Il concorso di cause può aversi sia che le cause stesse siano «indipendenti», che «dipendenti» (altrimenti non si spiegherebbe la locuzione «*anche* se indipendenti»; viceversa, in assenza di una specificazione nel cpv. dell'art. 41, le cause sopravvenute sono senz'altro interdipendenti da quelle antecedenti»; Preziosi (2021), p. 329, secondo cui «l'efficacia interruttiva delle cause sopravvenute è limitata a quelle condizioni il cui valore causale sia ben maggiore di quelle che possono venire in considerazione in generale come condizioni causali per fondare il rapporto eziologico fra condotta ed evento, poiché la causa interruttiva non basta che sia condizione necessaria di una condizione sufficiente ma deve essere *da sola sufficiente*».

[45] In questi stessi termini Cass. pen., Sez. IV, 5.5.2015, n. 33329, in *C.E.D. Cass.,* n. 264365. In tal senso, ancor più recentemente, Cass. pen. Sez. IV, 13.5.2019, n. 20270, in *C.E.D. Cass.,* n. 276238.

[46] Mezzetti (2021), p. 14 s. Per un'approfondita disamina delle sentenze, in verità assai ridotte in numero, che hanno riconosciuto l'operatività di tale fattore di esclusione della responsabilità, si veda Blaiotta (2007), p. 365 s.

[47] Mezzetti (2021), p. 178 s.

[48] Il concetto richiama quello di *black box causale,* largamente impiegato nel contesto della responsabilità da prodotto difettoso per indicare la circostanza che «si sa che il prodotto […] ha un nesso con determinati danni, si sa che non vi sono fattori estranei di causazione del danno, ma non si conosce qual è il fattore dannoso all'interno del prodotto», alla base di modelli di imputazione alternativi a quelli di tipo condizionalistico-nomologico. Tali ricostruzioni si fonderebbero principalmente sulla rilevazione di una mera connessione temporale tra gli eventi dannosi e l'utilizzazione del prodotto, nonché sulla cd. causalità negativa, ossia sull'insussistenza di decorsi causali alternativi. Per un esame approfondito di alcuni "accertamenti causali pionieristici" quali il caso Contergan, Lederspray, Holzschulzmittel e della frode alimentare dell'olio di colza cfr. Piergallini (2004), p. 189 s.

[49] Essenzialmente, si interroga selettivamente il modello per rivelarne alcune proprietà. Ad esempio, le tecniche di spiegazione delle decisioni assunte da un classificatore nelle attività di visione artificiale includono generalmente l'individuazione dell'area di interesse che il classificatore

a ben vedere, «un problema causale o non si pone affatto o, pur ponendosi, è tecnicamente risolvibile»[50].

Ciononostante, in ipotesi di irriducibile opacità del sistema[51] persiste il rischio che l'assenza di un solido corredo nomologico in materia legittimi il ricorso a categorie già sviluppatesi nel campo della responsabilità penalistica per danno da prodotto e, segnatamente, a modelli imputativi alternativi a quello condizionalistico: in questi casi, si verserebbe in un'anomia causale «difficilmente riconducibile allo schema di una legge di copertura, che, interpretando il programma e gli algoritmi utilizzati, consenta di affermare, *ex post,* che dato l'impiego di un determinato sistema di IA, secondo una predefinita regolarità nomologica, ne sarebbe seguito, con ragionevole certezza, l'evento effettivamente verificatosi»[52].

Al riguardo, attesa la distanza concettuale che separa "causalità *ex ante*" e "prova particolaristica del nesso causale", non può che richiamarsi l'esigenza di un'attenta valutazione di quest'ultima ad opera della giurisprudenza, alla luce delle irrinunciabili garanzie dell'"oltre ogni ragionevole dubbio"[53] nonché delle possibili evidenze fornite dalla tecnologia stessa, che segnerebbero l'ingresso nel processo della cd. *machine evidence* (*i.e* l'evidenza sull'interazione tra l'uomo e la macchina prodotta dal sistema IA stesso, sulla cui scorta il giudice possa affermare la credibilità razionale dell'ipotesi circa la produzione dell'evento)[54].

Così eventualmente accertata la sussistenza dell'elemento causale, la delimitazione della responsabilità dell'utilizzatore o del produttore richiederà peraltro un rigoroso accertamento della colpevolezza.

3.1. *La punibilità dell'utilizzatore tra (legittimo) affidamento nei sistemi di valutazione di conformità e distorsioni dell'automazione.*

Guardando ai sistemi connotati da orientabilità umana, la figura dell'utilizzatore risalta come "catalizzatore" di responsabilità, ossia come immediato e disponibile centro di accollo degli eventi dannosi, nonché come meccanismo di salvaguardia in grado di impedire che un malfunzionamento della macchina cagioni danni altrimenti evitabili (cd. *fail-safe mechanism*) nelle ipotesi in cui permanga in capo allo stesso un potere-dovere di intervento, cd. di *override*, sulle scelte della stessa[55].

A tal proposito, si pone il quesito circa l'opportunità di una disciplina che, mutuando dall'art. 590 *sexies* c.p., escluda in primo luogo la punibilità del "controllore" ove risultino rispettate le regole che definiscono l'area dei rischi connessi all'impiego di IA; in secondo luogo, attesa la speciale difficoltà tecnica della materia e muovendo dall'art. 2236 c.c., ne circoscriva la responsabilità alle ipotesi di colpa grave[56]. Fermo restando che «l'utente che manovri strumenti supportati da meccanismi di intelligenza artificiale, senza cognizione adeguata dei rischi che ciò comporti, versi in una situazione in cui si assume un rischio per lui non controllabile, riproducendo lo schema della cd. "colpa per assunzione"»[57].

ha considerato rilevante ai fini della decisione. Cfr. Hamon (2020), p. 13. Più in generale, Ordish *et al.* (2019), p. 23 s.; Biran e Cotton (2017), p. 8 s.

[50] La Vattiata (2022), p. 709.

[51] Sebbene, infatti, non tutti i modelli di apprendimento automatico siano di per sé indecifrabili, persistono talvolta delle ragioni, specialmente in termini di accuratezza, a favore di un minor grado di interpretabilità del sistema. In tema si veda Hamon (2020), p. 14.

[52] Preziosi (2022), p. 721.

[53] In particolare, sono state manifestate serie perplessità circa la possibilità che l'accertamento *ex post* del decorso eziologico, ai fini di un'ineludibile "certezza processuale", avvenga «secondo quelle leggi (o nozioni) scientifiche "di copertura" cui la giurisprudenza penale ormai si riferisce secondo il duplice parametro della probabilità statistica e soprattutto di quella logica o di elevata credibilità razionale: tanto più che, sulla base dei futuribili sistemi di *cloud computing,* la macchina si ritroverà a dover processare informazioni che, in quanto elaborate da una pluralità di applicativi e trasmesse in tempo reale alla rete, non si prestano al raggiungimento dello standard dell'oltre ogni ragionevole dubbio, *ex* artt. 533 e 530 o 546 lett. *e)*, cpv. c.p.p., in ordine all'esclusione di decorsi causali alternativi ipotetici». Così Piva (2022), p. 686 s.

[54] In tema si vedano Gless (2020); Canzio (2021); Canzio (2019), p. 45 s.; La Vattiata (2022), p. 703. Più in generale, sull'ingresso di strumenti "intelligenti" nel processo penale e, in particolare, nelle dinamiche probatorie, si vedano, *ex multis,* Luparia e Fiorelli (2022).

[55] Così, Amoroso e Tamburrini (2019), p. 51. Ad esempio, nel caso delle auto a guida semi-automatica si parla di *override button.* Cfr. Douma e Palodichuk (2012), pp. 1157-1162.

[56] Di tale avviso Fiorella (2022), il quale teorizza l'introduzione di una disposizione dedicata ai «*Limiti dell'autonomia ammissibile dell'IA. La colpa punibile del garante*» che escluda "oggettivamente" la colpa per imperizia del controllore «*quando sono rispettate le raccomandazioni previste dai protocolli sull'IA come definiti e pubblicati ai sensi di legge ovvero, in mancanza di questi, secondo le buone pratiche operative, sempre che le raccomandazioni previste dai predetti protocolli o le buone pratiche risultino adeguate alle specificità del caso concreto*».

[57] In tal senso Massi (2022), p. 678. Sulla colpa per assunzione si vedano, in successione cronologica, Marinucci (1965), p. 203; Forti (1990), p. 291 s.; Giunta (1993), p. 235 s.; Mantovani (1997), p. 341; Bartoli (2005), pp. 177-180 e 203-205; Castronuovo (2009), p. 601 (spec.

In secondo luogo, l'utilizzatore emerge quale soggetto maggiormente esposto a nutrire un forte affidamento[58] sia sulla certificazione di conformità del sistema agli standard di sicurezza garantiti, cui l'immissione in commercio è subordinata, che sulla correttezza degli *output* della macchina stessa, per effetto di quella che è stata definita come "distorsione dell'automazione".

Come accennato, la proposta di regolamento distingue tra sistemi IA indipendenti, di cui all'Allegato III, e sistemi IA impiegati come componenti di sicurezza di prodotti disciplinati conformemente al nuovo quadro normativo e soggetti a valutazione di conformità *ex ante* da parte di terzi[59], quali macchine[60] o dispositivi medici[61]. Per i sistemi indipendenti viene istituito un nuovo sistema di conformità e applicazione mediante verifiche di controllo interno da parte dei fornitori (fatto salvo il caso dei sistemi di identificazione biometrica a distanza, che saranno soggetti alla valutazione della conformità da parte di un terzo) in combinazione con un forte *enforcement ex post* che «potrebbe costituire una soluzione efficace e ragionevole […] considerato che l'intervento normativo è in fase iniziale e che il settore dell'IA è molto innovativo e solo ora si stanno maturando le competenze di audit»[62]. I secondi, invece, saranno soggetti ai medesimi requisiti di conformità e applicazione *ex ante* ed *ex post* dei prodotti di cui sono componenti, con la novità che tali meccanismi assicureranno la conformità non solo ai requisiti stabiliti dalla normativa settoriale ma anche a quelli fissati dal Regolamento.

A prescindere da eventuali profili di responsabilità dell'organismo notificato, si pone allora il quesito circa il valore da annettere a tali garanzie di sicurezza a fronte di una lesione cagionata da un dispositivo intelligente rivelatosi "difettoso" e, conseguentemente, in merito alla possibilità che un'applicazione del principio di affidamento, quale limite all'imputazione dell'evento a titolo di colpa, discenda dalla certificazione del sistema ad alto rischio ad opera dell'organismo terzo che dovrebbe garantirne la conformità agli standard fissati dal Capo 2 della proposta di regolamento[63].

Sotto diverso profilo, il possibile affidamento sulla correttezza degli *output* della macchina è messo in luce dall'articolo 29 della proposta, che impone agli utenti di IA ad alto rischio il rispetto di obblighi specifici, primo tra tutti quello di utilizzare tali sistemi conformemente alle istruzioni per l'uso che accompagnano tali sistemi (par. 1)[64], impregiudicati gli altri obblighi previsti dal diritto unitario o nazionale e ferma restando la discrezionalità dell'utente

nota 328); Mezzetti (2010), p. 513 s.; Pisani (2012); Piva (2020), p. 1139 s.

[58] In tema di affidamento, si vedano Marinucci e Dolcini (2021), p. 420 s.; Mezzetti (2020), p. 407; Piergallini (2017), p. 238 s.

[59] Il cd. "organismo notificato", ossia «un organismo di valutazione della conformità designato in conformità al presente regolamento e ad altre pertinenti normative di armonizzazione dell'Unione» [Articolo 3 (22)].

[60] Secondo la nuova Proposta di Regolamento sui *machinery products* «un ulteriore aspetto di semplificazione è costituito dalla complementarità tra le proposte legislative sull'intelligenza artificiale e sulle macchine, nell'ambito delle quali il regolamento sull'intelligenza artificiale delega la valutazione della conformità a quello sulle macchine affinché la valutazione dei rischi per la macchina completa con i sistemi di intelligenza artificiale venga effettuata una volta soltanto attraverso il futuro regolamento sui prodotti macchina».

[61] Il Regolamento UE/2017/745 suddivide i c.d. *software as medical devices* (o SaMD), in differenti classi di rischio. In particolare, la conformità dei dispositivi di classe superiore alla I viene valutata anche da un organismo notificato e attestata tramite certificazione, in relazione alla quale è stata in passato riconosciuta efficacia scusante nei confronti del medico che abbia impiegato un dispositivo difettoso da cui siano scaturiti omicidio o lesioni (cfr. Cass. Pen, Sez. IV, 10.11.2011, n. 40897, in *DeJure*), quale applicazione del principio di affidamento, il cui limite coinciderebbe con la natura riconoscibile del difetto del sistema. In tema, Lagioia (2016), p. 82 s.

[62] Cfr. Lusardi e Ferrari (2021).

[63] Sul punto si veda Borgogno (2022), p. 738, secondo cui la "virtuosa" standardizzazione delle misure cautelari da adottare in relazione ad ogni attività pericolosa è imposta «dalla necessità di adeguarsi ai più alti canoni di prudenza e di cautela noti nel momento dell'azione e richiesti al fine di evitare il prodursi di eventi dannosi. Ove pertanto un soggetto, per adeguarsi ai suddetti *standards,* utilizzi un sistema di intelligenza artificiale regolarmente autorizzato e testato con esito positivo dalle competenti autorità regolatorie si ridurranno evidentemente gli spazi per l'affermazione di una sua responsabilità penale a titolo di colpa, anche nel caso ove dall'esercizio della suddetta attività derivino conseguenze lesive non volute, a meno che non risulti provato che l'evento lesivo sia chiaramente riconducibile ad un suo errore nell'installazione, nell'uso o nella manutenzione del sistema».

[64] La precisione degli obblighi informativi rileverebbe anche in relazione al ruolo potenzialmente svolto dal comportamento scorretto o incauto della vittima-utilizzatore del sistema ai fini della responsabilità per il delitto colposo del produttore, il cui rapporto con il consumatore finale sembra improntato a un "modello a precauzione bilaterale" (in contrapposizione al "modello a precauzione unilaterale" predominante, invece, in settori come quello della tutela ambientale), connotato da reciprocità, essendo entrambe le parti chiamate a contribuire alla prevenzione del danno. Più specificamente, tale "reciprocità" si manifesta come possibilità, per l'utilizzatore, di confidare nel rispetto degli *standard* di sicurezza da parte del produttore nonché sulla correttezza delle istruzioni e informazioni fornite le quali, sotto tale aspetto, non potranno utilmente fondare un affidamento, dal lato del produttore, sull'"uso normale o ragionevolmente prevedibile" del sistema ove, più in generale, la formulazione delle informazioni rilasciate renda prevedibile il mancato adeguamento anche da parte di un consumatore avveduto giacché, ad esempio, non sono state adeguatamente specificate le possibili conseguenze discendenti dall'utilizzo improprio. Cfr. Castronuovo (2009), p. 328; Castronuovo (2005), p. 333 s. Nel contesto dell'immissione in commercio di sistemi IA, tale aspetto sembra indirettamente valorizzato dalla Proposta di Regolamento nel momento in cui il sistema di gestione dei rischi contemplato dall'art. 9 richiede espressamente, oltre all'«identificazione e all'analisi dei rischi noti e prevedibili associati a ciascun sistema di IA ad alto rischio», anche la «stima e valutazione dei rischi che possono emergere quando il sistema di IA è usato conformemente alla sua finalità prevista e in condizioni di uso improprio ragionevolmente prevedibili».

nell'organizzare le proprie risorse e attività al fine di attuare le misure di sorveglianza umana indicate dal fornitore (par 2)[65]. Più in generale, per quanto concerne l'aspetto della sorveglianza umana, i sistemi ad alto rischio devono essere progettati e sviluppati, anche con strumenti di interfaccia uomo-macchina adeguati, in modo tale da essere efficacemente supervisionati durante l'utilizzo al fine di prevenire o ridurre al minimo i rischi connessi[66].

Tali misure sono funzionali non solo ad una piena comprensione di capacità e limiti del sistema di IA ad alto rischio, il cui funzionamento deve essere monitorato per cogliere possibili segnali di disfunzione, anomalie o prestazioni inattese ed eventualmente consentire interventi sul funzionamento o l'attivazione di procedure di arresto, ma anche «a restare consapevole della possibile tendenza a fare automaticamente affidamento o a fare eccessivo affidamento sull'*output* prodotto da un sistema di IA ad alto rischio (cd. distorsione dell'automazione)»; in particolare per quelli impiegati «per fornire informazioni o raccomandazioni per le decisioni che devono essere prese da persone fisiche» (art. 14, par. 4, lett. b).

In conclusione, può senz'altro condividersi l'opinione per cui una responsabilità dell'utilizzatore non sia configurabile per il solo utilizzo del prodotto, ove ciò sia avvenuto conformemente alle informazioni fornite dal produttore[67], ma presupponga un'attiva modificazione ovvero, atteso l'obbligo di conservarne le funzionalità, un'omessa o insufficiente custodia del sistema; ancora, in capo all'utilizzatore che mantenga un dominio sulle funzioni esercitate dalla macchina o quantomeno un potere di intervento sulle stesse, a rilevare sarebbe un'omessa o insufficiente supervisione[68]: un addebito di responsabilità, rispettivamente a titolo omissivo e attivo, ben potrebbe configurarsi in caso di omessa attivazione[69] a fronte di un prevedibile fallimento del sistema che renda inoperante l'affidamento dell'utente circa il funzionamento dello stesso in conformità agli standard di diligenza[70] ovvero a causa di un errore commesso al

[65] Il riferimento è in particolare all' art. 14, par. 3, che contempla due diverse tipologie di misure per garantire la sorveglianza umana: oltre a quelle individuate dal fornitore e attuabili dall'utente (lett. b), quelle individuate e integrate nel sistema IA dal fornitore prima della sua immissione in mercato o in servizio ove tecnicamente possibile (lett. a).

[66] Rischi che, secondo l'art. 14, potrebbero emergere quando il sistema IA «sia usato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile, in particolare quando tali rischi persistono nonostante l'applicazione di altri requisiti di cui al relativo capo». A tal fine, l'art. 14 prevede una serie di importantissimi requisiti riguardanti la sorveglianza umana dei sistemi, che dovranno essere programmati e sviluppati in modo tale da garantire, secondo modalità specificamente prescritte (v. *amplius* par. 4), la loro effettiva sorveglianza da parte di una persona fisica all'atto dell'utilizzo, al dichiarato fine (par. 2) di prevenire o minimizzare i rischi per la salute, la sicurezza o i diritti fondamentali, che possono emergere in relazione sia alla destinazione d'uso, sia ai prevedibili usi impropri del *device* (si veda, inoltre, il 48º "considerando").

[67] Ove l'impiego di sistemi autonomi o semi-autonomi determinasse un'anticipazione dell'accertamento della colpa alla fase di implementazione e scelta di utilizzo del sistema, con una valutazione dell'eventuale negligenza dell'utilizzatore fondata sulla categoria dogmatica di estrazione civilistica della *cura in eligendo,* rilievo primario assumerebbe la scelta del sistema cui delegare determinate funzioni e la specifica attività delegata, specialmente sotto il profilo del rapporto instauratosi tra produttore e utilizzatore e delle informazioni fornite all'utente, che accompagnano la circolazione del sistema: così Bertolesi (2019), p. 236. Si noti che, secondo l'articolo 13 della Proposta di Regolamento, gli algoritmi impiegati dovranno essere trasparenti, in modo da consentire agli utenti l'interpretazione dell'*output* del sistema, e assicurare, pertanto, un uso corretto dello stesso; inoltre, ogni sistema dovrà essere corredato di una serie di istruzioni per l'uso concise, complete, corrette, chiare e formulate in termini comprensibili per gli *users* (informazioni che dovranno riguardare: il *provider* e/o un suo rappresentante; le caratteristiche, le capacità e i limiti del sistema; i cambiamenti attesi – pre-determinati – nel sistema e nel suo funzionamento; le misure di controllo umano, tra cui quelle di interpretazione dell'*output*; una stima della durata di funzionamento del sistema e ogni misura di manutenzione e cura, inclusi gli aggiornamenti) (si veda, inoltre, il 47º "considerando"). Circa gli obblighi di custodia, l'art. 13 include espressamente tra le informazioni che devono essere fornite in sede di immissione in commercio del sistema quelle relative a «tutte le misure di manutenzione e cura necessarie per garantire il corretto funzionamento del sistema, anche per quanto riguarda gli aggiornamenti software».

[68] Profilo, quest'ultimo, maggiormente esposto a criticità derivanti dalla possibilità che una ridotta capacità di attenzione, maturata a causa di una prolungata assenza di attività, determini il venir meno dell'effettiva possibilità di intervento del soggetto sull'azione già innescata dal sistema quasi totalmente autonomo. Tale aspetto tuttavia non potrebbe, verosimilmente, essere correttamente valorizzato in sede giudiziale attraverso un accertamento eccessivamente individualizzante sul versante della cd. misura soggettiva della colpa, che finirebbe per devolvere al giudice una valutazione eccessivamente discrezionale. Cfr. Piergallini (2017), p. 244. Ad analoga conclusione sembra condurre il quesito circa le potenziali ricadute di una deresponsabilizzazione generata dall'interazione con sistemi di IA sulla rimproverabilità dell'agente, sotto il profilo della componente psicologica della colpa. È stato rilevato che «il disimpegno morale, nella misura in cui determini un disinteresse verso gli sviluppi dell'impiego di strumenti supportati dall'IA, può ostacolare un compiuto giudizio in rapporto alla previsione di un certo evento di reato, quale conseguenza dell'attivazione dell'IA, impedendone la prevedibilità in concreto da parte del soggetto. Resta però dubbio che ciò escluda l'addebito per colpa, visto che l'imprevedibilità sarebbe frutto di un atteggiamento colpevole nascente da un atteggiamento riprovevole quale è appunto il disimpegno morale […]» e, ad ogni modo, si esclude che tale scelta possa essere rimessa alla valutazione del giudice. Così Massi (2022), p. 679.

[69] Gli effetti derivanti dall'impossibilità di avvalersi del principio di affidamento differiscono, come noto, a seconda della sussistenza di un vero e proprio obbligo di impedimento dell'evento, gravante sul soggetto in posizione di garanzia, ovvero di un mero obbligo di vigilanza o di attivarsi al verificarsi di determinati presupposti. Solo nel primo caso potrà infatti applicarsi la clausola dell'equivalenza di cui all'art 40 cpv. c.p. Cfr. Mezzetti (2020), p. 410.

[70] Come fa notare Beck (2016), pp. 138-141, la circostanza secondo cui «*the user can rely on the lawfulness of the actions of programmer, producer etc. does not mean that he can absolve himself fully from responsibility by arguing that the machine who made the decision should have worked properly*».

momento di riassumere il controllo[71].

3.2. *La responsabilità per il "tipo" e per il "modo" di produzione tra posizioni di garanzia e colpa di organizzazione.*

L'implementazione di IA sconta un'elevata complessità dovuta alla specificità tecnica della materia e alla varietà di organizzazioni coinvolte: a fronte di una società responsabile per lo sviluppo di parti di un algoritmo poi assemblato da una seconda società, altre potrebbero intervenire in sede di *testing*, nonché di vendita ad ulteriori *corporations* che implementino tale soluzione in un *hardware* a sua volta prodotto da terzi[72].

Sono così destinate ad accrescersi le difficoltà tradizionalmente riscontrate nell'appuntare un giudizio di responsabilità in capo alla persona fisica chiamata a governare la cellula funzionale da cui è scaturito il difetto concretizzatosi in danno[73] e che, tuttavia, non potrebbero giustificare, al fine di sopperire ad un'eventuale debolezza dell'impianto probatorio in materia causale, il ricorso a posizioni di garanzia "da ingerenza" sulla cui scorta riconoscere la responsabilità del produttore (o del programmatore) per omesso impedimento dell'evento lesivo[74], ovvero ad affermazioni di "fungibilità" tra condotte attive e condotte omissive per aggirare le difficoltà che notoriamente connotano l'accertamento del reato omissivo improprio, connotato da un giudizio controfattuale di secondo grado[75].

Più in generale, in tema, è stato rilevato come «sebbene si parli correntemente di causalità commissiva ed omissiva, non è affatto chiaro dove finisca l'una e inizi l'altra»[76]: il problema concerne, notoriamente, l'esatta individuazione del confine tra azione e omissione basata su parametri non naturalistici bensì normativi, che attengono al significato della condotta consistente, rispettivamente, nell'attivazione ovvero nell'omesso impedimento di un decorso causale già in atto[77].

Ebbene, nel contesto in esame, dominato da una pluralità di attori chiamati a confrontarsi con una «quantità di scenari […] vastissima e potenzialmente indefinita»[78] e da asperità probatorie ipoteticamente insormontabili, può solo richiamarsi l'esigenza di accertare, in capo al produttore, l'effettiva presenza di quei «poteri e doveri corrispondenti a dati ruoli»[79] che

[71] Bertolesi (2019), p. 240.
[72] Diamantis (2021), p. 5.
[73] La Vattiata (2022), p. 705; Piergallini (2020), p.1755.
[74] In particolare, secondo Piergallini (2004), p. 242 s., «la giurisprudenza individua nella messa in circolazione del prodotto un comportamento preliminare (prodromico) a rischio crescente, la cui riprovevolezza deriverebbe non dalla pericolosità del prodotto ma dalla verificazione dei primi casi di danno, la cui insorgenza legittimerebbe *ex post* la prognosi di pericolosità del prodotto *genericamente* formulabile *ex ante*», così provocando un'irreparabile fenditura nella categoria del rischio consentito che, notoriamente, comprende quella sfera di attività ammesse o tollerate dall'ordinamento pur nella prevedibilità che dalle stesse possano derivare eventi dannosi.
[75] Grasso (1983), p. 385 s. Nel noto caso Lederspray, ad esempio, «se, di regola, la responsabilità per danno da prodotto si atteggia come omissiva (coincidendo con il mancato ritiro del prodotto pericoloso)», la giurisprudenza, nell'impossibilità di risalire al comparto produttivo di origine del danno e di individuare le sfere di competenza con le correlate posizioni di garanzia gravanti sui singoli soggetti, ha valorizzato la decisione del consiglio di amministrazione di lasciare il prodotto sul mercato, giungendo a coinvolgere «tutti i responsabili del vertice dell'organizzazione complessa per aver partecipato, *collegialmente*, alla decisione di non ritirare il prodotto dal mercato». Così Piergallini (2004), pp. 234-238.
[76] Così Blaiotta (2010), p. 318. L'Autore ricorda che «non è neppure chiaro, in molti casi, se il giudizio controfattuale che si è chiamati a compiere in ordine all'evitabilità dell'evento attenga alla causalità o alla colpa. Si tratta di una questione fondamentale, mai sufficientemente approfondita in giurisprudenza, ove si riscontra una grande confusione. Il tema, d'altra parte, presenta rilevantissime implicazioni applicative. […] In molti campi e particolarmente in quello della responsabilità medica ed in quello di attività pericolose poste in essere senza l'osservanza delle doverose cautele si è in presenza, assai spesso, di condotte attive nelle quali - però - assumono un ruolo preminente momenti omissivi, che risultano essere la chiave di volta per la spiegazione degli accadimenti». In dottrina v., altresì, Masera (2006), p. 499, il quale ha osservato che «la giurisprudenza maggioritaria che qualifica da sempre in termini omissivi la relazione eziologica, tanto nel campo dell'esposizione a sostanze tossiche che in quello medico, confonde il concetto di omissione causale ai sensi dell'art. 40, comma 2, c.p. con la naturale componente omissiva propria di ogni condotta colposa. L'erroneità dell'indirizzo emerge con evidenza - si afferma- nell'ambito dell'esposizione a tossici. Ritenere omissiva la responsabilità dell'imprenditore perché ha omesso di adottare le cautele per ridurre entro i limiti consentiti il contatto dei lavoratori con la sostanza pericolosa, significa dimenticare la previa condotta positiva dell'imputato, che ha predisposto la struttura produttiva al cui interno è avvenuta l'esposizione della vittima all'agente tossico. In breve, si ha sempre causalità commissiva nell'ambito di attività pericolose come quella industriale o della circolazione stradale».
[77] Come osservato da Piva (2011), p. 65 a proposito della colpa infortunistica del datore di lavoro, al quale si contesta «tanto un elemento "positivo", consistente nella predisposizione di una struttura carente o nella selezione di personale inadeguato, quanto un elemento "negativo", traslato dallo schema di accertamento dell'illecito colposo, alla stregua di mancata adozione di cautele doverose».
[78] Così Cappellini (2019), p. 326 s.; Piergallini (2020), p. 1755.
[79] Pulitanò (2019), p. 193. Secondo la più recente dottrina, la Proposta di Regolamento definirebbe l'ambito di tali poteri e doveri attraverso le disposizioni di cui agli artt. 16-29, che individuerebbero i campi personali di responsabilità per l'osservanza degli obblighi relativi ai requisiti

convergerebbero verso la realizzazione di una forma rafforzata di protezione nei confronti di diritti ritenuti fondamentali e che solo potrebbero far sorgere quello «speciale vincolo di tutela»[80] tra beni e garanti tipico delle posizioni di garanzia, sulla base del quale opererebbe l'art. 40 cpv. c.p.[81].

Sorge peraltro il quesito se, ove il danno sia causalmente riferibile a un difetto del sistema dovuto alla violazione di *standard* di diligenza e non si individuino gli autori dell'illecito, possa o meno configurarsi un'ipotesi di responsabilità della *societas* ai sensi dell'art. 8 d.lgs. 231/2001, la cui applicabilità presupporrebbe, tuttavia, una modifica *de lege ferenda* che includa nel catalogo dei reati presupposto della responsabilità dell'ente i reati di omicidio e lesione colposi realizzati in violazione delle norme sulla sicurezza dei prodotti[82].

Una corresponsabilizzazione dell'ente, quantomeno per le contravvenzioni in materia di sicurezza dei prodotti[83], sarebbe in primo luogo giustificata dalla considerazione per cui, il più delle volte, gli interessi e motivazioni che inducono a scelte lesive di beni giuridici rilevanti siano esterni al soggetto in carne ed ossa e riconducibili alla "politica di impresa", derivandone un'esigenza di prevenire punizioni eccessive nei confronti delle persone fisiche. A ciò si aggiunge l'idea secondo cui l'art. 8 sarebbe pensato soprattutto in relazione a reati colposi di evento, potendo dunque rivelarsi un valido aiuto per far fronte ad ipotesi di difficile individuazione dei soggetti responsabili all'interno dell'impresa, soprattutto in contesti di successione diacronica di garanti e di difficile individuazione dell'arco temprale in cui il bene difettoso sia stato realizzato, sebbene accertato il nesso di causalità tra difetto del prodotto e danno[84].

Ad ogni modo, al fine di impedire che il rimprovero devii verso forme di responsabilità oggettiva per rischio e di natura circolare (che collega automaticamente la mancata individuazione dell'autore ad una carenza organizzativa, per cui la prima proverebbe l'esistenza della seconda), va tenuta ferma la necessità che l'evento da cui scaturisce l'addebito costituisca pur sempre un fatto colpevole per l'ente, ossia che il rimprovero non esaurisca il suo contenuto nella mancata individuazione del reo, dovendo accertarsi una correlazione funzionale tra carente organizzazione e singolo reato-presupposto[85].

Ciò premesso, una prima delimitazione dell'area di responsabilità ascrivibile al produttore del sistema autonomo può discendere dalla distinzione, ben nota per il danno da prodotto, tra responsabilità per il "tipo" e responsabilità per il "modo" di produzione[86] che, proiettati nel settore di nostro interesse, si specificano in funzione della pericolosità del prodotto: non integralmente contenibile ricorrendo a misure cautelari o adottando determinate tecniche costruttive nel primo caso, eliminabile o minimizzabile attraverso una funzionalizzazione del processo produttivo e un'adeguata informativa per l'utilizzatore nel secondo[87].

Al riguardo, sulla scorta della proposta di regolamento, una responsabilità per il "tipo" di

che un sistema ad alto rischio deve rispettare per poter essere immesso nel circuito economico, identificando specificamente i destinatari di tali obblighi, non rivolti dunque alla generalità dei consociati. Si veda, inoltre, il 53° "considerando", ove viene enfatizzata la posizione del *provider* (fornitore), destinatario principale, sia sotto il profilo quantitativo, sia sotto quello qualitativo, degli obblighi: «È opportuno che una specifica persona fisica o giuridica, definita come il fornitore, si assuma la responsabilità dell'immissione sul mercato o della messa in servizio di un sistema di IA ad alto rischio, a prescindere dal fatto che tale persona fisica o giuridica sia la persona che ha progettato o sviluppato il sistema». Così La Vattiata (2021), p. 13. Sul punto Fiorella (2022), p. 659 che, nel richiamare gli artt. 16-29, sottolinea l'esigenza di una chiara definizione dell'area del controllo umano, giacché «dalla necessità di un telaio molto preciso di regole cautelari e/o permissive, dovrebbero scaturire nell'immediato futuro protocolli chiari e conclusivi *con precisa ripartizione di competenze* nella gestione dell'IA».

[80] Così Fiandaca e Musco (2019), p. 650 s.

[81] Così La Vattiata (2021), p. 13. In tema, si vedano *amplius* Sgubbi (1975); Fiandaca (1979); Grasso (1983).

[82] In questi termini si esprime Piergallini (2020), p. 1756 s.

[83] Nell'ambito del "sistema integrato di disciplina" disegnato dal codice del consumo attraverso disposizioni di diritto civile, amministrativo e penale, nel quadro di una tutela improntata a cadenze tipiche del modello ingiunzionale, l'art. 112 contempla delle figure contravvenzionali che introducono reati di mera condotta applicabili in via sussidiaria, ove il fatto non costituisca più grave reato, in ipotesi di immissione di prodotti pericolosi in violazione del divieto posto dalle competenti autorità amministrative (comma 1), immissione in commercio di prodotti pericolosi (comma 2), inottemperanza ai provvedimenti adottati dalle autorità amministrative (comma 3). Così Consulich (2007), p. 2977; Castronuovo (2012), p. 20. Per le diverse ricostruzioni delle fattispecie in esame si vedano Consulich (2007), p. 2981; Gargani (2013), p. 692 s.

[84] Bertolesi (2019), p. 194 s.

[85] Piergallini (2020), p. 1756 s. Per un approfondimento sulla ricostruzione del nesso di ascrizione dell'art. 8 d.lgs 231/2001, si veda Paliero (2008), p. 1541 s.

[86] Si noti come, sebbene distinguibili in linea teorica, le responsabilità per il tipo e il modo di produzione tendano talvolta a sovrapporsi nella pratica: è possibile che, a fronte del verificarsi di danni irreparabili riconducibili ad un tipo di produzione che avrebbe dovuto essere vietata *ab initio,* si contesti al produttore la violazione di norme inerenti piuttosto al modo di produzione, sostanzialmente "recuperandosi" eventuali lacune attinenti alla fase dei controlli amministrativi (in sede di selezione dei tipi di produzione da autorizzare), sul piano delle modalità di svolgimento della produzione. Cfr. Bertolesi (2019), p. 90. Vedi più approfonditamente Perini (2002), pp. 389-391; Bricola (1978), p. 75 s; Bricola (1997), p. 1231 s.; Piergallini (2004), p. 45.

[87] Piergallini (2020), p. 1754; Piergallini (2004), p. 45.

produzione (ossia concernente la qualità stessa del rischio connesso all'attività) potrà configurarsi solamente in corrispondenza degli impieghi di IA vietati dalla normativa, giacché per le rimanenti applicazioni dell'IA la cui immissione in commercio sia stata autorizzata previa valutazione di conformità ai requisiti richiesti, l'attività è da considerarsi consentita (fatte salve le condotte decettive per eludere i controlli a tal fine previsti), pur residuando profili di rischio "accettabili"[88] a fronte del rispetto di regole cautelari da parte del produttore[89].

Piuttosto, rilevato un difetto del sistema, sorgerebbe il quesito circa la sussistenza di una responsabilità per il "modo" di produzione dello stesso che si collochi al di fuori di quell'area di rischio consentito su cui il produttore fa affidamento[90], potendo contare su una regolamentazione cautelare che orienti la sua condotta tramite *standard* predefinibili *ex ante,* adeguati a «disinnescare occasioni di rischio o a ridurne l'incidenza dannosa»[91] e funzionali a evitare indebite ricostruzioni della fattispecie colposa *ex post facto*. In ciò si evidenza come la regola precauzionale non abbia, rispetto all'illecito colposo, unicamente funzione cautelare rivolta a prevenire la verificazione di eventi dannosi, quanto soprattutto una funzione permissiva, di individuazione di un'area di rischio permesso oltre la quale "non può esserci responsabilità penale".

A tal fine, risulta imprescindibile il riferimento a criteri normativi e specialmente a regole cautelari codificate[92], la cui capacità di individuare la condotta doverosa del soggetto lascia talvolta spazio a valutazioni di prevedibilità ed evitabilità dell'evento secondo le circostanze del caso concreto da parte dell'agente[93], chiamato ad avvalersi di parametri normativi (come protocolli o linee guida), informazioni di carattere tecnico-scientifico[94] (eventualmente cristallizzati in "regole tecniche" o autorizzazioni amministrative a contenuto prescrittivo, che subordinano l'esecuzione di un'attività a determinate condizioni[95]) ovvero usi invalsi nel settore.

Tali parametri, nel complesso, valgono a delimitare l'ampiezza dell'obbligo di previsione

[88] Art. 9, comma 4 «Le misure di gestione dei rischi di cui al paragrafo 2, lettera d), sono tali che qualsiasi rischio residuo associato a ciascun pericolo nonché il rischio residuo complessivo dei sistemi di IA ad alto rischio sono considerati accettabili, a condizione che il sistema di IA ad alto rischio sia usato conformemente alla sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile. Tali rischi residui sono comunicati all'utente».

[89] In tal senso La Vattiata (2020), p.10 s.

[90] Il "rischio consentito" è notoriamente un concetto chiave nella teoria generale della colpa penale, la cui collocazione dogmatica è tuttavia dibattuta in dottrina. L'orientamento dominante, che ravvisa nel rischio consentito il livello di rischio autorizzato in via generale dall'ordinamento per lo svolgimento di una determinata attività socialmente utile, lo colloca alternativamente sul piano della tipicità ovvero della colpevolezza. Alla prima opzione sono riconducibili sia la corrente maggioritaria, che tratta il rischio consentito nella misura oggettiva della colpa e quest'ultima nella tipicità; sia quella che aderisce alla teoria dell'imputazione oggettiva dell'evento [*ex multis,* Militello (1988), p. 55 s.]. Alla seconda opzione è invece riconducibile l'opinione, minoritaria, seppur autorevole, che pure colloca il rischio consentito nella misura oggettiva della colpa, ma considera quest'ultima come un elemento del reato interamente afferente alla colpevolezza [Marinucci e Dolcini (2021), p. 402 s.]. Infine, l'orientamento minoritario in dottrina e diffuso nella giurisprudenza sulle lesioni sportive, che concepisce il rischio consentito come frutto di bilanciamenti da effettuarsi in concreto, gli attribuisce natura di causa di giustificazione, collocandolo nell'antigiuridicità. In tema si veda ampiamente Zirulia (2018), p. 405 s.; nella manualistica cfr. Fiandaca e Musco (2019), p. 586 s. Circa il rilievo del rischio consentito nel settore dell'IA si veda Gless *et al.* (2016), p. 431, che lo qualificano come "*margin of tolerance*". In tal senso anche Beck (2018), p. 54 s.

[91] Si assiste, a tal proposito, in settori quali la sicurezza sul lavoro o la sicurezza dei prodotti, a un intenso fenomeno di "positivizzazione" di regole preventive combinate a forme di "procedimentalizzazione" o "proceduralizzazione" della sicurezza, tramite una valutazione dei fattori di rischio in vista di una sua riduzione al minimo. Ad esempio, la direttiva 2001/95 si caratterizza per l'integrazione di forme di etero-regolazione a opera di un sistema comprensivo di fonti normative pubblicistiche statali ed europee (quali «leggi, regolamenti, norme di uniformizzazione delle caratteristiche di sicurezza dei prodotti, "specifiche tecniche", "standard di sicurezza", "requisiti generali e specifici di sicurezza", "limiti tabellari" o "valori soglia" circa la composizione dei prodotti ecc…»), nonché di etero-controllo da parte delle Autorità amministrative («sistemi di autorizzazioni e controlli, ingiunzioni di sicurezza contenenti, a loro volta, divieti o prescrizioni ecc…»), con modelli di autodisciplina che, in un'ottica di parziale de-istituzionalizzazione e privatizzazione dei compiti legati alla sicurezza, impongono doveri (o oneri o facoltà) di auto-regolazione e auto-controllo rimessi, almeno parzialmente, agli stessi operatori economici coinvolti nello svolgimento di attività rischiose. Così Castronuovo (2005), p. 320 s.; in tal senso anche Mantovani (2017), p. 333. Nel campo dell'IA, si veda l'articolo 9 della proposta di regolamento, relativo al sistema di gestione dei rischi e, precisamente, lett. b) e c), relative alla stima e valutazione dei rischi che possono emergere quando il sistema di IA ad alto rischio è usato conformemente alla sua finalità prevista e in condizioni di uso improprio ragionevolmente prevedibile; valutazione di altri eventuali rischi derivanti dall'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato di cui all'articolo 61.

[92] Zirulia (2018), p. 363 s.; cfr. Marinucci e Dolcini (2021), p. 404 s.

[93] Piergallini (2017), pp. 222-228.

[94] Circa il rilievo del sapere scientifico nella creazione delle regole cautelari in generale si veda Forti (1990), pp. 206-210, 515-522. La dottrina ritiene che il livello di sapere doveroso coincida con quello tipico della cerchia di riferimento dell'agente, ossia dell'*homo eiusdem condicionis et professionis,* e respinge il criterio della "migliore scienza ed esperienza del momento storico" «poiché così facendo si rinuncerebbe totalmente a quel momento normativo insito nella delimitazione del rischio consentito connesso all'esercizio di attività pericolose: la regola cautelare coincidente con quella ricavabile dalla migliore scienza ed esperienza, infatti, […] impedirebbe quella valutazione ponderata tra utilità sociale e pericolosità dell'attività che è invece resa possibile dal ricorso appunto a un "modello" (dunque, necessariamente normativo) di agente». Cfr. Palazzo (2016), p. 328. Nello stesso senso, Forti (1990) p. 261 s.; Zirulia (2018), p. 367.

[95] Cfr. Militello (1988), p. 91 s.; Forti (1990), p. 309 s.

del produttore agente concreto[96] che non potrà, salvo le dovute precisazioni (su cui si veda *infra*), essere chiamato a rispondere degli eventi lesivi occorsi nonostante l'osservanza di regole cautelari, giacché «accontentarsi della prevedibilità del decorso causale porterebbe a ravvisare la colpa generica anche rispetto a un evento che è concretizzazione di un rischio residuale consentito», mentre l'agire di un agente modello si connota tipicamente per il mantenimento del rischio entro il livello tollerato nell'ambito dell'attività considerata, nel rispetto di una concezione autenticamente normativa della colpa[97].

A tal riguardo, pur nell'attuale mancanza di *standard* tecnici adeguati[98], giacché i competenti organismi sembrano impegnati a capire quali siano i rischi da evitare ancor prima di stabilire come evitarli[99], un primo quadro orientativo di obblighi specificamente concernenti l'introduzione di sistemi IA in commercio, che andranno ad aggiungersi a quelli specifici del settore in considerazione (si pensi a quelli previsti dalla direttiva macchine)[100], sono forniti dagli artt. 8-15 della proposta di regolamento.

In merito, si noti come le fonti europee, pur nell'impossibilità di interventi diretti in materia penalistica, siano nondimeno in grado di esercitare un condizionamento "indiretto" sulla legislazione interna[101] laddove, all'interno di fattispecie colpose di evento, si innestino normative cautelari di derivazione comunitaria concernenti settori normativi complementari di rilevanza centrale nella moderna società del rischio. Ciò è particolarmente vero per gli atti regolamentari, in grado di produrre un'europeizzazione "immediata", con potenziali ricadute sulla configurabilità della responsabilità "per colpa", soprattutto specifica, in caso di violazione di obblighi di sicurezza in esso eventualmente contenuti[102].

Ebbene, alla luce del limite di "sufficiente determinatezza" della fattispecie legale ritagliato dalla giurisprudenza costituzionale per l'operatività di una confluenza della normativa primaria con fonti non legislative[103], l'ammissibilità di etero-integrazioni dovrà essere vagliata sulla base dell'incidenza riconosciuta alla regola cautelare nella costruzione del "tipo colposo"[104], giungendosi a legittimare il ricorso alla normazione sub-legislativa in funzione (para)cautelare «solo se la cornice preventiva e la base valoriale su cui dovrebbe plasmarsi il contenuto

[96] Alla base la concezione per cui l'ambito del rischio illecito è necessariamente più ristretto di quello del rischio riconoscibile. In tal senso, Militello (1988), pp. 56-60, 130-132, 140 s. Più specificamente, si è affermato che il rischio consentito coinciderebbe con «il rischio obiettivamente prevedibile in base all'insieme delle conoscenze nomologiche ed ontologiche disponibili ex ante, ma non prevedibile dal punto di vista della figura modello o, ancorché da questa prevedibile, non tale da influire sulle sue modalità di condotta». Cit. Forti (2006a), p. 945 s., 952; più diffusamente Forti (1990), p. 249 s.; Bertolesi, (2019), p. 111.

[97] Zirulia (2018), p. 358 s.; Militello (1998), p. 140 s.; Giunta (1999), p. 86. Sulla figura dell'agente modello cfr. Mantovani (2017), p. 371.

[98] Per quanto concerne specificamente il campo della robotica, fino al 2014 l'elaborazione di standard, limitata alla robotica industriale (ISO 10218-1: 2006; ISO 10218-2: 2011), era essenzialmente finalizzata a ridurre l'interazione tra robot ed esseri umani. Solo successivamente (a partire dagli standard ISO 13482: 2014), sulla scorta dell'esigenza di regolare i cd. *service robots*, si è dato avvio a una serie di lavori, oltre che a livello internazionale da parte dell'ISO e dell'IEEE, anche a livello UE attraverso CEN e CENELEC. In tema, si veda *amplius* Mobilio (2020), p. 413 s.

[99] Beck (2016), p.139; Beck (2018), p. 47.

[100] La direttiva 2006/42/CE, che disciplina le fasi di progettazione e costruzione delle macchine: nella versione presentata dalla Commissione europea con una recente Proposta di Regolamento, essa sarà volta a garantire un'integrazione sicura del sistema di IA nella macchina nel suo complesso che non ne comprometta la sicurezza (Cfr. art. 9 della relativa Proposta di Regolamento). Sulla rilevanza della direttiva macchine per i robot quali artefatti meccanici si veda Turano (2020).

[101] In tema di effetti indiretti esercitati dal diritto comunitario sul diritto penale interno si veda Caianello e Manes (2020), p. 22 s.

[102] Quanto all'efficacia diretta *in malam partem* delle disposizioni comunitarie, la giurisprudenza interna, come noto, l'ha già più volte ammessa nel caso di contrasto con una norma nazionale più favorevole, dovendosi tuttavia distinguere il caso in cui l'eterointegrazione incida soltanto sulla definizione del fatto da quello in cui contribuisca a definire il precetto: cfr., ad esempio, Cass. pen., Sez III, 23.2.2011, n. 107, in *www.lexambiente.it*; nonché già Sez. III, 3 luglio 2007, n. 39345, in *pa.leggiditalia.it,* secondo cui, ove l'eterointegrazione incida sulla mera specificazione di elementi della fattispecie già definiti nel nucleo significativo essenziale delle scelte valutative della legge penale e il conflitto si manifesti in forma di incompatibilità evidente «il giudice è tenuto [..] a non applicare la disposizione contrastante con quella di fonte comunitaria». Ai nostri fini, sull'inquadramento delle regole cautelari come *standard* comportamentali la cui inosservanza delinea il volto oggettivo e la dimensione lesiva dell'azione, svolgendo quindi una funzione di integrazione del precetto dal punto di vista della tipizzazione obiettiva del comportamento sanzionato, v., tra i tanti, Giunta (1993), pp. 15-17; Giunta (2012), p. 581; Piergallini (1997), pp. 1492 s.; Greco (2016), p. 133 s. *A contrario*, disconoscono la funzione tipizzante delle regole in questione per affermarne la rilevanza sul piano della colpevolezza, De Francesco (2012), p. 666; Donini (2013), p. 132 s.

[103] Inaugura l'indirizzo Corte cost., 23.3.1966, n. 26, in *Giur. cost*., 1966, 255, cui hanno fatto seguito le sentenze n. 168 del 1971 e n. 282 del 1990. Sul punto si veda Manes (2010), p. 101.

[104] A tal proposito, Preziosi (2022), p. 720 s. ha rilevato che occorre «tener conto della possibilità di una vera e propria *trasformazione di mere regole tecniche cautelari in regole strettamente e normativamente permissive*» che, ritagliando un'area di rischio consentito avente ad oggetto l'impiego di sistemi di IA evoluti che non consentono un completo dominio umano sui meccanismi decisionali della macchina, escluderebbero in radice un eventuale problema causale e dunque la stessa tipicità del fatto (comprensivo di condotta ed evento lesivo). In definitiva, «la funzione tipizzante della regola precauzionale e del correlativo dovere di diligenza oggettiva, non opera sulla condotta del reato colposo, definendone la sua rilevanza sul piano della colpevolezza colposa quale condotta antidoverosa, bensì opera sul piano del nesso eziologico condotta-evento».

della norma abbiano già formato oggetto di valutazione (in termini di rischio consentito/non consentito) da parte di una cautela primaria»[105]: circostanza, peraltro, che non potrà essere affermata in termini assoluti, dovendosi piuttosto ragionare in termini di prevalenza dell'uno o dell'altro aspetto ("modale" o "assiologico") nella regola cautelare positivizzata[106].

Quanto finora affermato richiede tuttavia delle precisazioni.

In primo luogo, persistono dei limiti all'affidamento che l'agente può nutrire in ragione dell'osservanza delle regole cautelari scritte che non varrebbero ad annullare *in toto* l'ambito di applicabilità della colpa generica laddove, ad esempio, per determinate attività pericolose, potrebbero concretamente adottarsi misure precauzionali più efficaci. Tali limiti (i cd. limiti del limite), possono essere fatti coincidere con le situazioni in cui risulti prevedibile, in astratto (a causa di difetti intrinseci della regola cautelare stessa) o in concreto (per circostanze di fatto anomale) il fallimento della regola cautelare, non potendo operare il limite del rischio consentito ove le cause del fallimento fossero conosciute o riconoscibili da parte di un agente modello. A tal fine, il momento a partire dal quale diventa prevedibile il fallimento della regola cautelare e dunque può pretendersi dall'agente il riconoscimento dei rischi connessi all'attività è individuabile nel momento in cui le conoscenze sui rischi sono disponibili a un agente modello appartenente alla medesima cerchia di riferimento dell'agente concreto.

In secondo luogo, le riflessioni sopra condotte non devono indurre alla conclusione che il rischio consentito coincida con la diligenza obiettiva, in quanto mentre «la diligenza dovuta esprime le modalità di una condotta che sono richieste per limitarne le potenzialità offensive», il rischio consentito «caratterizza la pericolosità dell'azione dopo l'applicazione delle precauzioni necessarie». Del resto, si è osservato che un'identificazione di tali categorie concettuali in un sistema che contempla la colpa generica porrebbe le premesse perché il concetto di rischio consentito perda «la sua funzione di limite ai giudizi di prevedibilità ed evitabilità in concreto, risultandone piuttosto a sua volta "annacquato"»[107].

Alla luce di tali considerazioni, il generico obbligo gravante in capo al produttore "tradizionale" di «immettere soltanto prodotti sicuri»[108] risulterà specificato dal rinvio a leggi e normative tecniche emanate da enti di normalizzazione internazionali o europei, nonché a eventuali provvedimenti amministrativi (autorizzazioni, prescrizioni ecc…) che, ad esempio, stabiliscano ulteriori condizioni per la distribuzione, secondo meccanismi di etero-regolazione[109], e accompagnato dalla previsione di obblighi successivi al rilascio in mercato del prodotto (monitoraggio, informazioni aggiuntive, richiamo o ritiro del prodotto)[110] in relazione ai quali problematici rimangono sia l'individuazione del momento in cui, riconosciuto il cd. rischio di sviluppo, sorga in capo al produttore l'obbligo di assumere l'iniziativa, sia l'alternativa tra il ritiro del prodotto dal mercato e l'adozione di altri obblighi cautelativi[111].

[105] Greco (2016), p. 132.

[106] In particolare, l'aspetto "tecnico" prevarrebbe nelle regolamentazioni cautelari in materia di circolazione di macchinari, apparecchiature o attrezzature che introducono, in capo ai diversi soggetti coinvolti nella loro produzione e distribuzione (quali il fabbricante, l'eventuale mandatario, i fornitori ecc.), particolari procedure cautelari c.d. di valutazione del rischio, obblighi informativi, obblighi di dotare i prodotti in questione di determinati dispositivi di sicurezza e così via. In questi casi, infatti, sebbene le norme regolamentari specifichino alcune variabili ricomprese nella valutazione di accettabilità del rischio, l'aspetto tecnico, "modale" sembrerebbe prevalente. Viceversa, alla categoria della prevalenza "assiologica", sarebbero probabilmente riconducibili le regole in materia ambientale che introducono, svolgendo un bilanciamento di interessi, elenchi di sostanze ritenute inquinanti o pericolose. Ciò ferma restando la necessità di valutare se, a fondamento di tali regole, stiano scelte tipicamente politiche, "utra-prudenziali", piuttosto che valutazioni di tipo prettamente tecnico-scientifico. Di tale avviso Buzio (2020), p. 11 s.

[107] Zirulia, (2018), p. 363 s., 378 s.

[108] Art. 104, comma 1 d.Lgs 206/2005 (Codice del consumo). L'obbligo, di per sé a valenza programmatica e limitata alla progettazione e costruzione dei prodotti, è a sua volta precisato in una serie di misure indicate dal comma 4, che comprendono ad esempio controlli a campione ed esame dei reclami.

[109] Bertolesi (2019), pp. 191, 229.

[110] Si richiama a tal proposito quanto già detto circa la rilevanza annessa al sistema *post-market monitoring system* dalla proposta di Regolamento. In particolare, l'art. 21, dedicato alle misure correttive, prevede espressamente che «i fornitori di sistemi di IA ad alto rischio che ritengono o hanno motivo di ritenere che un sistema di IA ad alto rischio da essi immesso sul mercato o messo in servizio non sia conforme al presente regolamento adottano immediatamente le misure correttive necessarie per rendere conforme tale dispositivo, ritirarlo o richiamarlo, a seconda dei casi. Essi informano di conseguenza i distributori del sistema di IA ad alto rischio in questione e, ove applicabile, il rappresentante autorizzato e gli importatori».

[111] Bertolesi (2019), p. 233.

3.3. *Gli obblighi di conformità del fornitore tra deregulation e self-regulation di settore.*

Il fornitore di sistemi di IA è tenuto, ex art. 16, a rispettare i requisiti stabiliti dal capo 2 con la specificazione che «le soluzioni tecniche precise atte a conseguire la conformità ai requisiti essenziali richiesti dal capo 2 per l'immissione in commercio di sistemi IA ad alto rischio, possono essere previste mediante norme o altre specifiche tecniche o altrimenti essere sviluppate in conformità alle conoscenze ingegneristiche o scientifiche generali, a discrezione del fornitore del sistema di IA». Tale flessibilità è infatti ritenuta «particolarmente importante in quanto consente ai fornitori di sistemi di IA di scegliere il modo in cui soddisfare i requisiti che li riguardano, tenendo conto dello stato dell'arte e del progresso tecnologico e scientifico nel settore».

In secondo luogo, il fornitore è chiamato a dotarsi di un sistema di gestione della qualità, volto a garantire la conformità ai requisiti del regolamento e comprensivo di una serie di aspetti "minimi" (art. 17). In sostanza, al produttore non verrebbe lasciata libertà decisionale in merito all'*an,* bensì al *quomodo* dell'autonormazione, tramite l'imposizione di vincoli che circoscrivono la facoltà di adattamento delle prescrizioni sovraordinate alle specificità della realtà organizzativa[112].

Il primo obbligo richiama la considerazione per cui sempre più, nell'attuale panorama giuridico, siano proprio le fonti autonormate a offrire un significativo contributo nella determinazione del comportamento tenuto dai soggetti attivi in settori a rischio[113]. E ciò è particolarmente evidente nelle fattispecie colpose, contraddistinte da una "tipicità evanescente" che spesso necessita di un'etero-integrazione che colmi gli spazi di indeterminatezza del precetto statale[114]. Risulta pertanto evidente come le regole prodotte da «formazioni intermedie rappresentative di esigenze o interessi omogenee», traducibili in soluzioni quali linee guida o *best practices*, presentino un duplice rilievo: *ex ante* nei confronti dei destinatari operanti in contesti di rischio consentito sulla scorta di una valutazione di utilità sociale ed *ex post* nei confronti del giudice ai fini dell'accertamento della sussistenza dell'elemento soggettivo del reato contestato[115].

Più in generale, anche lo Stato trae una certa utilità dagli *standard* di settore[116] laddove, soprattutto nella disciplina di fenomeni transnazionali, ne recepisca in sede di normazione i contenuti che siano già stati sperimentati dagli operatori stessi (come è avvenuto, seppur a livello di istituzioni europee e di principi generali, per la proposta di regolamento)[117] contribuendo in tal modo a colmare le lacune di conoscenze a livello tecnico[118]. Ebbene, specificamente in un settore come quello dell'intelligenza artificiale, il cui sviluppo promette effetti potenzialmente dirompenti, è evidente come occorra abbandonare il pregiudizio insito nella presunta "neutralità" della tecnologia[119] per delineare una regolamentazione eterogenea che contempli norme etiche, tecniche e strumenti di *self-regulation*[120].

A tal riguardo, può senz'altro apprezzarsi la scelta di non affidare la materia a un'autonormazione "pura" in un contesto di completa *deregulation* pubblica e di rimpiazzo del diritto im-

[112] Gargani (2021), p. 51.
[113] *Ibidem.*
[114] *Ibidem*. L'Autore ricorda infatti come il diritto penale non sia più solo un «imperativo statuale», specialmente nei settori di tutela più tecnicizzati e specializzati, in cui si assiste a una proliferazione delle fonti sia verso l'alto (europeizzazione e internazionalizzazione delle norme, anche penali) che verso il basso (ossia contemplando apporti creativi dai destinatari finali), quest'ultima tendenza esprimendo la «matrice privata e autonormata delle discipline di settore» che nel complesso contribuisce a quella che viene definita come la «conformazione reticolare dell'ordinamento giuridico». In tema si veda anche Bianchi (2019).
[115] È in tal senso che può leggersi l'affermazione secondo cui «la compressione del principio di legalità dal punto di vista della riserva di legge statale è così bilanciata da un incremento di determinatezza e capacità orientativa del precetto». Cfr. Gargani (2021), p. 52.
[116] Intendiamo qui il termine "standard" nella sua dimensione più ampia, ossia come comprensivo di norme, obiettivi, finalità e regole «attorno alle quali un regime regolatorio è organizzato, tramite procedimenti di adozione che coinvolgono attori pubblici e privati, a livello nazionale e sovranazionale, rendendone così condivisa la responsabilità». Così Scott (2010), p. 104 s.
[117] Specificamente, la proposta afferma che «i requisiti minimi proposti costituiscono già lo stato dell'arte per numerosi operatori diligenti e rappresentano il risultato di due anni di lavoro preparatorio, derivato dagli orientamenti etici del gruppo di esperti ad alto livello sull'intelligenza artificiale, guidato da più di 350 organizzazioni. Tali requisiti sono altresì in gran parte coerenti con altre raccomandazioni e altri principi internazionali, circostanza questa che assicura che il quadro dell'IA proposto sia compatibile con quelli adottati dai partner commerciali internazionali dell'UE».
[118] Mobilio (2020), p. 412.
[119] Mobilio (2020), p 406.
[120] Piergallini (2020), p. 1773.

perativo dello Stato con forme di *self-regulation* che, invece di assicurare un bilanciamento tra esigenze di flessibilità e di sostegno al progresso tecnico scientifico ed economico da un lato, e di tutela di diritti e libertà fondamentali dall'altro, produrrebbe uno squilibrio a sicuro vantaggio delle prime. In tal senso, va salutata con favore l'opzione da parte della Commissione per lo strumento regolamentare, particolarmente incisivo dal punto di vista dell'armonizzazione che si giustifica proprio in ragione della considerazione della natura potenzialmente "ubiquitaria" dell'intelligenza artificiale ormai proiettata verso la libera circolazione nel mercato europeo attraverso l'integrazione in prodotti e servizi di ogni genere, tale da rendere ormai anacronistico il tentativo di regolarne gli sviluppi tramite strumenti meno penetranti[121].

In particolare, imponendo al fornitore l'adozione di un sistema di gestione della qualità[122] teso all'individuazione, valutazione e comunicazione alle competenti autorità dei rischi connessi all'attività nonché all'eventuale adozione di misure di contenimento dei medesimi, il nuovo quadro normativo potrebbe assumere la fisionomia di una "meta-normazione" che tenti di comporre in un insieme equilibrato interesse pubblico e autonomia privata tramite un'imposizione di autonormazione. Si introdurrebbero, così, veri e propri obblighi di autodisciplina la cui inosservanza comporta un'irrogazione di sanzioni[123] di varia entità e natura, a prescindere dalla commissione di un illecito penale, nel quadro di quella che potrebbe essere definita come una "*enforced self-regulation*"[124].

4. Prospettive *de iure condendo*: lo spostamento verso modelli ingiunzionali e funzioni di *cooperative compliance*.

La "prevenzione mediante autonormazione" eventualmente imposta dal legislatore, pur passando necessariamente da un'attività gruppale, potrà coinvolgere individui in carne ed ossa a monte o a valle del procedimento, con la conseguenza per cui il disallineamento tra destinatario dell'obbligo/onere e gruppo (ossia il protagonista per definizione del processo autoregolativo) non configurerà una (inammissibile) responsabilità per fatto altrui solo ove il soggetto abbia un potere giuridico e fattuale di influire effettivamente sull'organizzazione[125].

Tuttavia, la complessità della realtà affrontata non può che suggerire - come accennato - un progressivo spostamento dell'attenzione circa l'individuazione dei centri di responsabilità per danni da IA verso le organizzazioni complesse[126], vale a dire le entità che meglio possono gestire anticipatamente i rischi connessi a tali attività nonché adempiere agli eventuali oneri imposti in termini di adozione di misure conformative e di sostenimento dei relativi costi, operanti peraltro in contesti ove lo stesso scenario *one-corporation-per-algorithm* risulta ormai sempre più anacronistico[127].

Non esiste, cioè, un approccio "*one size fits all*", peraltro osteggiato anche dal Parlamento europeo[128], quando si tratta di rispondere a interrogativi in punto di ascrizione della responsabilità penale, considerata l'estrema eterogeneità delle tecnologie IA e dei loro possibili impieghi[129].

A tal riguardo, merita richiamarsi una proposta maturata nella nostra dottrina[130], specifi-

[121] La Vattiata (2021), p. 16.
[122] Art. 17 della Proposta di Regolamento.
[123] Si veda a tal proposito l'art. 71.
[124] Braithwaite (1982), p. 1466 s. Un esempio di tale modello di regolamentazione, in ambito nazionale, potrebbe essere individuato nel documento di valutazione dei rischi per la salute e la sicurezza dei lavoratori che il datore è tenuto ad adottare ai sensi del d.lgs. 81/2008 (Cfr. Gargani (2021), p. 54). In ipotesi come quella del diritto penale del lavoro, in cui il datore è obbligato all'analisi e al trattamento dei rischi per la salute e sicurezza dei lavoratori, lo Stato chiede ai soggetti privati di predisporre regole autonormate finalizzate a scongiurare la concretizzazione di eventi offensivi che, ove verificatisi, possono andare a integrare fattispecie criminose direttamente imputabili agli stessi soggetti destinatari della delega normativa. Il mandato dello Stato, tuttavia, potrebbe anche essere non a costruire regole e dispositivi di controllo atti a escludere o mitigare il rischio di verificazione di eventi avversi (*incidents*), ma piuttosto ad autodisciplinarsi e automonitorarsi in modo tale da impedire o rendere più difficoltosa la perpetrazione di fatti criminosi nel contesto gruppale (cd. *occupational crimes*). Così Bianchi (2019), p. 1504.
[125] Come nel caso del datore di lavoro rispetto al sistema di gestione della sicurezza e salute sul luogo di lavoro. Bianchi (2019), p. 1511.
[126] Diamantis (2021), p. 4 s.
[127] Così Infantino e Wang (2019).
[128] Parlamento europeo, Risoluzione *Una politica industriale europea globale in materia di robotica e intelligenza artificiale*, 2019, punto 116.
[129] Constantine (2021), p. 10.
[130] La proposta, avanzata da Piergallini (2020), p. 1773, costituisce fondamentalmente un'evoluzione di quella già elaborata da Forti, che contempla un'ipotesi di responsabilità penale colposa per omessa comunicazione alle pubbliche autorità di informazioni rilevanti per

camente rivolta alle organizzazioni complesse e secondo cui il diritto penale sarebbe in primo luogo funzionale a presidiare, tramite sanzioni indirizzate alla *corporation*, gli oneri imposti all'ente in termini non solo di condivisione del *know how* rilevante ai fini della valutazione e prevenzione dei rischi non ancora concretizzatisi, ma anche di comunicazione alle pubbliche autorità dei danni verificatisi e del relativo tasso di frequenza, al fine di valorizzare la superiorità informativa di cui i cd. *Big tech* sono dotati in ordine alle possibili manifestazioni e all'eziologia dei rischi[131].

In secondo luogo, la medesima ottica preventiva dovrebbe favorire, sulla scorta di una «condivisibile ripulsa nei confronti di una pena indirizzata sempre e solo verso la persona fisica»[132], l'elaborazione di un diritto penale incentrato sulla persona giuridica e calibrato non più sul consueto schema «se fai A allora B», bensì su quello tipico del modello ingiunzionale, secondo cui «ti ingiungo di fare A, e se non fai A scatterà una sanzione»[133], già sperimentato in ambito infortunistico e ambientale, nonché da ultimo teoricamente generalizzato dalla riforma operata con la legge delega 27 settembre 2021, n. 134 (cd. riforma Cartabia) 134, in attesa di attuazione. In tal senso, al verificarsi di eventi dannosi, si prospetta la possibilità di un'ingiunzione da parte delle autorità del settore di adottare misure conformative (ri-programmatorie) ovvero definitive (disattivatrici) che, in caso di inosservanza, potrebbero stimolare il ricorso al diritto penale, il quale in tal modo verrebbe a svolgere una funzione, per così dire, di *cooperative compliance*[135].

A ciò si aggiunga l'ipotesi di un ricorso all'art. 8 del d. Lgs. 231/2001 non solo per il caso in cui «si sa che un reato è stato commesso, ma non lo si può accertare perché non si riesce a scovare/identificare il suo autore», bensì nella direzione di una responsabilizzazione diretta dell'ente coinvolto nella progettazione dell'IA, laddove «si sa in radice che un reato *non* è stato commesso» proprio per la difficoltà di muovere un rimprovero di colpevolezza all'autore che sia stato eventualmente identificato. *De iure condendo*, una forma di responsabilità «diretta e originaria della *societas*» costituirebbe «nel suo sviluppo la punta estrema dell'autonomizzazione della responsabilità dell'ente fondata sulla colpevolezza organizzativa», ferma restando, come già sottolineato, la necessità di «vincolare il giudizio di colpevolezza a uno stringente accertamento della correlazione funzionale tra il fatto illecito e la lacuna organizzativa»[136].

Da ultimo, valorizzandosi la possibilità di una più proficua integrazione tra diritto penale, civile e amministrativo ispirata al principio di sussidiarietà, potrebbe applicarsi un sistema sanzionatorio "multilivello" modellato in base al grado di lesività della condotta rispetto ai beni giuridici oggetto di tutela e, specificamente, immaginare sanzioni amministrative irrogate dall'autorità di regolazione del settore per "illeciti minori" in grado di incidere su beni giuridici diversi dalla vita o dall'integrità fisica e sanzioni penali irrogate dal giudice per illeciti "gravi" lesivi di questi ultimi ovvero lesivi in misura rilevante di altri beni[137].

Tale approccio dovrà peraltro completarsi con quelle «*'technical' or 'technological' solutions*», a integrazione o in sostituzione delle norme che andrebbero nel complesso a contribuire a

la gestione del rischio nel quadro di una *preventive governance* fondata sulla proceduralizzazione del processo decisorio tramite scambio continuo di informazioni tra autorità e privato che eserciti un'attività produttiva potenzialmente lesiva di beni giuridici tutelati. In tal modo, risulterebbero escluse dal novero delle attività "consentite" quelle «foriere di conseguenze che la figura modello non avrebbe potuto prevedere, ma che sarebbero state prevedibili dall'autorità pubblica ove questa avesse potuto disporre di un bagaglio di conoscenze (nomologiche e fattuali) non inferiore a quello di cui disponeva l'agente; che sarebbero state prevedibili, potremmo anche dire, da una sorta di "agente modello collettivo", portatore del *know-how* nomologico-fattuale risultante dal doveroso scambio con l'istanza di controllo delle cognizioni supplementari rilevanti possedute o nel frattempo acquisite dall'agente». Cfr. Forti (2006b), p. 219.

[131] Così Bianchi (2019), p. 1500.

[132] La Vattiata (2022), p. 711; così anche Piergallini (2007), p. 1129.

[133] Marinucci (2005), p. 55 s.

[134] Contenente una delega al Governo a prevedere una causa di estinzione delle contravvenzioni, destinata a operare nella fase delle indagini preliminari a seguito del tempestivo adempimento di apposite prescrizioni impartite dall'organo accertatore. In sede di attuazione della delega, il Governo è chiamato ad individuare le contravvenzioni per le quali sarà ammessa la causa di estinzione, che dovranno necessariamente essere suscettibili di elisione del danno o del pericolo mediante condotte ripristinatorie o risarcitorie. Per un approfondimento sulla riforma del processo e del sistema sanzionatorio penale si vedano Gatta (2021); Palazzo (2021).

[135] Piergallini (2020), p.1773. Così anche Piva (2022), p. 692.

[136] Tripodi (2022), p. 752 s. Sul punto, già Paliero (2004), p. 30; Mongillo (2018), p. 400, nt. 255.

[137] Di questo avviso La Vattiata (2022), p. 711; in una prospettiva di *common law*, La Vattiata (2020), p. 15. Per una valorizzazione del diritto civile anche Borgogno (2022), p. 737. Così anche Constantine (2021), p. 2, secondo cui il quesito circa l'imponibilità di una responsabilità penalistica a fronte di eventi dannosi inerenti l'uso di "RAI" (robot autonomi e intelligenza artificiale) e l'individuazione del soggetto su cui farla ricadere dovrà essere sciolto in funzione della gravità e del rischio di danni attuali o potenziali inerenti l'utilizzo del sistema nei relativi contesti; del livello di autonomia del sistema (in particolare, all'aumentare di questa si prospetta l'opportunità di definire espressamente in via legislativa o regolamentare chi possa considerarsi come utilizzatore e i corrispondenti obblighi di supervisione o controllo); del grado di supervisione sul sistema e di coinvolgimento nelle sue decisioni.

quella che viene definita come una *Law 3.0* che, a fronte dei nuovi rischi generati dalla tecnologia non si affidi unicamente all'evoluzione e all'adattamento delle norme[138]. In tal senso, è proprio attraverso il codice inteso alla stregua di architettura o *design* del sistema che sarebbe possibile operare scelte in grado di rendere determinati comportamenti più difficili, costosi o addirittura impossibili, in ciò palesandosi una cogenza analoga a quelle di vere e proprie regole giuridiche, non limitata dunque a una rilevanza di natura meramente valoriale e morale[139] ("*ethics by design*")[140].

Le stesse norme di diritto verrebbero sempre più immesse nel codice dei dispositivi tecnologici al fine di arginare i problemi posti dall'innovazione per mezzo della tecnologia stessa[141], lasciando emergere il ruolo preponderante esercitabile dal diritto non solo in contesti di regolazione *della* tecnologia (cd. *regulation of technology*), ma anche di regolazione *attraverso* la tecnologia (*regulation through technology*)[142]. Sebbene infatti l'idea di integrare nel "codice" una forma di regolazione nasca con internet e il cyberspazio, è nel campo delle più moderne tecnologie, specialmente quelle di intelligenza artificiale, che essa è destinata a rivelare appieno il suo potenziale[143], sulla scorta di una consapevolezza, ormai diffusa a livello sia europeo che internazionale secondo quanto emerge dai documenti ufficiali, che il buon *design* dei sistemi IA consenta di assicurare che i sistemi intelligenti operino all'interno dei parametri prefissati fornendo i risultati attesi[144].

Bibliografia

Al Mureden, Enrico (2017): "La responsabilità del fabbricante nella prospettiva della standardizzazione delle regole sulla sicurezza dei prodotti", in Al Mureden, Enrico, *La sicurezza dei prodotti e la responsabilità del produttore: casi e materiali* (Torino, Giappichelli).

Amoroso, Daniele e Tamburrini, Gabriele (2019): "I sistemi robotici ad autonomia crescente tra etica e diritto: quale ruolo per il controllore umano", *Rivista di BioDiritto,* 1, pp. 33-51.

Bartoli, Roberto (2005): *Colpevolezza: tra personalismo e prevenzione* (Torino, Giappichelli).

Beck, Susanne (2016): "Intelligent agents and criminal law – Negligence, diffusion of liability and electronic personhood", *Robotics and Autonomous Systems*, pp. 138-143.

Beck, Susanne (2017): *"Google Cars, Software Agents, Autonomous Weapons Systems –New Challenges for Criminal Law"*, in Hilgendorf, Eric e Seidel, Uwe (a cura di) *Robotics, Autonomics and the Law* (Baden-Baden, Germany, Nomos), pp. 227-251.

Beck, Susanne (2018): "Robotics and Criminal Law. Negligence, Diffusion of Liability and Electronic Personhood", in Hilgendorf, Eric (a cura di), *Digitization and the Law* (Baden Baden, Germany, Ed. Nomos), pp. 41-55.

Bertolesi, Riccardo (2019): *Intelligenza artificiale e responsabilità penale per danno da prodotto*, (tesi di dottorato di Diritto penale nell'ambito del Corso di dottorato di ricerca in Scienze Giuridiche "Cesare Beccaria" dell'Università degli Studi di Milano – Curriculum di diritto penale e processuale penale - XXXII ciclo).

Bianchi, Davide (2019): "Appunti per una teoria dell'autonormazione penale", *Rivista italiana di diritto e procedura penale,* 3, pp.1477-1525.

[138] Brownsword (2020), p. 2.
[139] Sul punto, vedi più approfonditamente Dommering e Asscher (2006).
[140] Di recente, Dignum *et al.* (2018), p. 60 s.
[141] Pagallo (2014), p. 4.
[142] Cfr. Brownsword e Yeung (2008), i quali distinguono "*Technology as a Regulatory Target*" da "*Technology as a Regulatory Tool*". Tale approccio peraltro trova già espresso riconoscimento nel GDPR declinandosi nel cd. principio del *data protection by design* che dovrebbe orientare l'intero ciclo di attività di un sistema tecnologico (ricerca, progettazione, sviluppo, implementazione e utilizzo pratico) attraverso l'integrazione della tutela della privacy e la protezione dei dati sfruttando il design dello stesso. Cit. Mobilio, (2020), p. 417.
[143] Come suggerisce anche De Vanna (2018), p. 395.
[144] Bryson e Theodorou, (2019), p. 310.

Biran, Or e Cotton, Courtenay (2017): "Explanation and Justification in Machine Learning: A Survey", *IJAI-17 Workshop on Explainable AI*, 8.

Blaiotta, Rocco (2007): "La causalità giuridica alla luce della teoria del rischio", Intervento al Convegno "*causalità e imputazione oggettiva dell'evento*", Università "La Sapienza" di Roma, org. da Stile, 24.11.2005, in *Cass. pen,* 47, 1, pp. 365-400.

Blaiotta, Rocco (2010): *Causalità giuridica* (Torino, Giappichelli).

Borgogno, Roberto (2022): "La responsabilità penale nei processi ad elevata automazione", in Giordano, Rosaria, Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale. Persona, Mercato, Amministrazione, Giustizia* (Milano, Giuffrè), pp. 727-743.

Braithwaite, John (1982): "Enforced Self-Regulation: A New Strategy for Corporate Crime Control", *Michigan Law Review*, 80, 7, pp. 1466-1507.

F. Bricola (1978): "Responsabilità penale per il tipo e per il modo di produzione", in AA. VV., *La responsabilità dell'impresa per i danni all'ambiente e ai consumatori* (Milano, Giuffrè), pp. 75 ss.; poi in Bricola, Franco (1997), *Scritti di diritto penale*, vol. I, tomo II (Milano, Giuffrè), pp. 1231 ss.

Bricola, Franco (1997): *Scritti di diritto penale,* vol. I, tomo II (Milano, Giuffré).

Brownsword, Roger (2020): *Law 3.0: rules, regulation and technology* (Abingdon, New York, Routleged).

Brownsword, Roger e Yeung, Karen (a cura di) (2008): *Regulating Technologies: Legal Futures, Regulatory Frames and Technological Fixes* (Oxford, Hart Publishing).

Brusco, Carlo (2012): "Rischio e pericolo, rischio consentito e principio di precauzione. La c.d. "flessibilizzazione delle categorie del reato"", *Criminalia*, pp. 383- 412. Bryson, Joanna e Theodorou, Andreas (2019): "How Society Can Maintain Human-Centric Artificial Intelligence", in Toivonen, Marja e Saari, Eveliina (a cura di), *Human-Centered Digitalization and Services*, (Singapore, Springer), pp. 305-323.

Buzio, Carolina (2020): "Regole cautelari eurounitarie e colpa penale. Spunti per una riflessione", *Discrimen*, pp. 37-50.

Caianello, Michele e Manes, Vittorio (2020): *Introduzione al diritto penale europeo. Fonti, metodi, istituti, casi* (Torino, Giappichelli).

Canzio, Giovanni (2019): "La motivazione della sentenza e la prova scientifica: "reasoning by probabilities"", in Amato, Alessandro, Flora, Giovanni e Valbonesi, Cecilia (a cura di), *Scienza, diritto e processo penale nell'era del rischio* (Torino, Giappichelli), pp. 45-60.

Canzio, Giovanni (2021): "Intelligenza artificiale, algoritmi e giustizia penale", *Sistema penale,* 8, pp. 1-7.

Cappellini, Alberto (2018): "Machina delinquere non potest? Brevi appunti su intelligenza artificiale e responsabilità penale?", *Criminalia,* pp. 499-520.

Cappellini, Alberto (2019): "Profili penalistici delle self-driving cars", *Diritto penale contemporaneo*, 2, pp. 325-353.

Carnevali, Ugo (1999): "Responsabilità del produttore", in *Enc. del dir.* Aggiornamento, II (Milano, Giuffrè), pp. 936-950.

Carnevali, Ugo (2006): "Prevenzione e risarcimento nelle direttive comunitarie sulla sicurezza dei prodotti", in *Studi in onore di Giorgio Marinucci*, II, (Milano, Giuffrè), pp. 2881-2901.

Castronuovo, Donato (2005): "Responsabilità da prodotto e struttura del fatto colposo", *Rivista italiana di diritto e procedura penale*, 1, pp. 301-340.

Castronuovo, Donato (2009): *La colpa penale* (Milano, Giuffrè).

Castronuovo, Donato (2012): *Principio di precauzione e diritto penale. Paradigmi dell'incertezza nella struttura del reato* (Roma, Aracne Editrice).

Chiara, Pier Giorgio (2019): "Software e responsabilità da prodotto: il caso boeing 737 max 8", *CyberLaws.*

Chopra, Samir e White, Laurence (2011): *A Legal Theory For Autonomous Artificial Agents* (Ann Arbor, Michigan, University of Michigan Press).

Constantine, Simon (2021): *Report on Criminal Liability, Robotics and AI Systems*, Singapore Academy of Law, Law Reform Committee.

Consulich, Federico (2007): "Tutela del consumatore (voce)", in Palazzo, Francesco Carlo e Paliero, Carlo Enrico (diretto da), *Commentario breve alle leggi penali complementari* (Padova, Cedam), pp. 2969-2986.

De Francesco, Giovannangelo (2012): "Colpa e prevenzione del rischio nel campo delle malattie professionali", in *Diritto penale e processo*, 6, pp. 665-669.

De Vanna, Francesco (2018): "Diritto e nuove tecnologie: il nodo (controverso) della regolazione giuridica", *Lo Stato*, 11, pp. 387-402.

Delogu, Tullio (1974): "Lo "strumento" nella teoria generale del reato", in *Rivista italiana di diritto e procedura penale.*

Diamantis, Mihailis (2021): "Vicarious liability for AI", *U Iowa Legal Studies Research Paper*, 27, pp. 1-18.

Dignum, Virginia *et al.* (2018): "Ethics by Design: necessity or curse?", *AIES 2018. Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 60-66.

Dommering, Egbert e Asscher, Lodevijk (a cura di) (2006): *Coding Regulation. Essays on the Normative Role of Information Technology,* vol. 12 (L'Aia, T.M.C. Asser Press).

Donini, Massimo (2013): "L'elemento soggettivo della colpa. Garanzie e sistematica", in *Rivista italiana di diritto e procedura penale*, 1, pp.124-156.

Douma, Frank e Palodichuk, Sarah Aue (2012): "Criminal Liability Issues Created by Autonomous Vehicles", *Santa Clara Law Review,* 52, 4, pp. 1157-1169.

Fiandaca, Giovanni (1979): *Il reato commissivo mediante omissione*, Milano.

Fiandaca, Giovanni e Musco, Enzo (2019): *Diritto penale. Parte generale*, VIII ed., (Bologna, Zanichelli).

Fiorella, Antonio (2022): "Responsabilità penale del Tutor e dominabilità dell'Intelligenza Artificiale. Rischio permesso e limiti di autonomia dell'Intelligenza Artificiale", in Giordano, Rosaria, Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale*, cit., pp. 651-663.

Forti, Gabrio (1990): *Colpa ed evento nel diritto penale* (Milano, Giuffrè).

Forti, Gabrio (2006a): voce *Colpa (dir. pen)*, in Cassese, Sabino, *Dizionario di diritto pubblico*, II (Milano, Giuffrè), pp. 945 ss.

Forti, Gabrio (2006b): *Accesso alle informazioni sul rischio e responsabilità*, *Criminalia,* pp.155-225.

Gargani, Alberto (2013): "Reati contro l'incolumità pubblica. Reati di comune pericolo mediante frode", in Grosso, Carlo Federico e Padovani, Tullio e Pagliaro, Antonio (diretto da), *Trattato di diritto penale,* Parte speciale, vol. IX (Milano, Giuffrè).

Gargani, Alberto (a cura di) (2021): "Illeciti punitivi in materia agro-alimentare", in Palazzo, Francesco Carlo e Paliero, Carlo Enrico e Pelissero, Marco (diretto da), *Trattato teorico pratico di diritto penale* (Torino, Giappichelli).

Gatta, Gian Luigi (2021): "Riforma della giustizia penale: contesto, obiettivi e linee di fondo della 'Legge Cartabia'", *Sistema penale* online, 15 ottobre.

Giunta, Fausto (1993): *Illiceità e colpevolezza nella responsabilità colposa* (Padova, Cedam).

Giunta, Fausto (1999): "La normatività della colpa penale. Lineamenti di una teorica", in *Rivista italiana di diritto e procedura penale,* pp. 86-115.

Giunta, Fausto (2006): *Il diritto penale e le suggestioni del principio di precauzione*, in *Criminalia*, pp. 227-250.

Giunta, Fausto (2012): "Il reato colposo nel sistema delle fonti", *Giustizia penale*, 2012, II, 11, pp. 577-592.

Gless, Sabine e Silverman, Emily e Weigend, Thomas (2016): "If robots cause harm, who is to blame? Self-driving cars and criminal liability", *New Criminal Law Review*, 19(3), pp. 412-436.

Gless, Sabine (2020): "AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials", *Georgetown Journal of International Law*, 51/2, pp.195-253.

Gray, Jim e Siewiorek, Daniel P. (1991): "High-Availability Computer Systems", *IEEE Computer*, vol. 24 (9), pp. 39-48.

Grasso, Giovanni (1983): *Il reato omissivo improprio. La struttura obiettiva della fattispecie* (Milano, Giuffrè).

Greco, Eliana (2016): "Eterointegrazione cautelare e successione di leggi nelle cadenze strutturali dell'illecito colposo. In particolare: il microsistema degli spettacoli musicali, cinematografici e teatrali delineato dal «decreto palchi»", in *Diritto penale contemporaneo – Rivista trimestrale,* 1, pp. 126-155.

Gruppo di esperti ad alto livello sull'intelligenza artificiale (2018), *Orientamenti etici per un'IA affidabile*.

Guidi, Dario (2010): "Regime sanzionatorio e cause di estinzione degli illeciti sulla sicurezza del lavoro", in Giunta, Fausto e Micheletti, Dario (a cura di), *Il nuovo diritto penale della sicurezza nei luoghi di lavoro,* Milano, pp. 935-966.

Hallevy, Gabriel (2010): "The Criminal Liability of Artificial Intelligence Entities – from Science Fiction to Legal Social Control", *Akron Intellectual Property Journal*, 4, 2, Art. 1.

Hallevy, Gabriel (2013): *When Robots Kill: Artificial Intelligence Under Criminal Law* (Boston, Northeastern University press).

Hallevy, Gabriel (2015): *Liability of Crimes Involving Artificial Intelligence System*, (Berlino, Springer).

Hamon, Ronan e Junklewitz, Henrik e Sanchez, Ignacio (2020): "Robustness and Explainability of Artificial Intelligence. From technical to policy solutions", *JRC Technical Report, European Commission.*

Infantino, Marta e Wang, Weiwei (2019): "Algorithmic torts: A perspective Comparative Overview", *Transnational Law & Contemporary Problems,* 29, 1, pp. 280-332.

Lagioia, Francesca (2016): *Responsabilità penale e automazione nell'E-Health*, [Dissertation thesis], Alma Mater Studiorum Università di Bologna. Dottorato di ricerca in Diritto e nuove tecnologie, 28 Ciclo. DOI 10.6092/unibo/amsdottorato/7697.

La Vattiata, Federico Carmelo (2020): "Artificial Intelligence in Healthcare: Risk Assessment and Criminal Law", *Diritto Penale e Uomo*, 12.

La Vattiata, Federico Carmelo (2021): "Brevi note "a caldo" sulla recente Proposta di Regolamento UE in tema di intelligenza artificiale", *Diritto Penale e Uomo*, 6.

La Vattiata, Federico Carmelo (2022): "La responsabilità penale per danni da intelligenza artificiale alla prova del processo", in Giordano, Rosaria, Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale,* cit., pp. 695-712.

Luparia, Luca e Fiorelli, Giulia (2022): "Diritto probatorio e giudizi criminali ai tempi dell'Intelligenza Artificiale", in Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale,* cit., pp. 779-797.

Lusardi, Giacomo e Ferrari, Alessandro (2021): "Regolamento UE sull'Intelligenza Artificiale: uno strumento articolato per gestire il rischio", *Quotidiano giuridico*, 3 giugno.

Magro, Maria Beatrice (2018): "A.I.: la responsabilità penale per la progettazione, la costruzione e l'uso dei robot", *Il Quotidiano giuridico,* 12 giugno.

Magro, Maria Beatrice (2019): "Robot, cyborg e intelligenze artificiali", in Cadoppi, Alberto e Canestrari, Stefano, e Manna, Adelmo e Papa, Michele (a cura di), *Trattato di diritto penale - Cybercrime*, (Torino, Utet Giuridica), pp.1179-1212.

Magro, Maria Beatrice (2020): "Decisione umana e decisione robotica, un'ipotesi di responsabilità da procreazione robotica", *La legislazione penale,* 10 maggio.

Manes, Vittorio (2010): "L'eterointegrazione della fattispecie penale mediante fonti subordinate, tra riserva "politica" e specificazione "tecnica"", *Rivista italiana di diritto e procedura penale*, 1, pp. 84-114.

Mantovani, Marco (1997): *Il principio di affidamento nella teoria del reato colposo* (Milano, Giuffrè).

Mantovani, Marco (2017): *Diritto penale. Parte generale* (Padova, Cedam).

Marinucci, Giorgio (1965): *Colpa per inosservanza di leggi* (Milano, Giuffrè).

Marinucci, Giorgio (2005): "Innovazioni tecnologiche e scoperte scientifiche: costi e tempi di adeguamento delle regole di diligenza", *Rivista italiana di diritto e procedura penale,* 1, pp. 29-59.

Marinucci, Giorgio e Dolcini, Emilio (2021): *Manuale di Diritto Penale. Parte generale* (Milano, Giuffrè).

Masera, Luca Mario (2006): "Il modello causale delle Sezioni Unite e la causalità omissiva", *Diritto penale e processo,* pp. 493-502.

Massi, Silvia (2022): "Affidamento sull'intelligenza artificiale e "disimpegno morale" nella definizione dei presupposti della responsabilità penale", in Giordano, Rosaria, Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale. Persona, Mercato, Amministrazione, Giustizia,* (Milano, Giuffré), pp. 665-679.

Mezzetti, Enrico (2010): "Colpa per assunzione", in S. Vinciguerra, F. Dassano (a cura di), *Scritti in memoria di G. Marini* (Napoli, Edizioni Scientifiche Italiane), pp. 513-536.

Mezzetti, Enrico (2020): *Diritto penale. Dottrina, casi e materiali*, III ed. (Bologna, Zanichelli).

Mezzetti, Enrico (2021): "Autore del reato e divieto di «regresso» nella società del rischio", in Donini, Massimo e Mezzetti, Enrico e Pelissero, Marco e Seminara, Sergio (diretto da), *Diritto penale in evoluzione* (Napoli, Jovene editore).

Militello, Vincenzo (1988): *Rischio e responsabilità penale* (Milano, Giuffrè).

Millar, Jason e Kerr, Ian (2013): *Delegation, Relinquishment and Responsibility: The Prospect of Expert Robots*, 107, *SSRN Electronic Journal,* pp. 102-127.

Mobilio, Giuseppe (2020): "L'intelligenza artificiale e i rischi di una "disruption" della regolamentazione giuridica", *Rivista di BioDiritto*, 2, pp. 401-424.

Mongillo, Vincenzo (2018): *La responsabilità penale tra individuo ed ente collettivo* (Torino, Giappichelli).

Munsona, John C. *et al*. (2006): "Software faults: A quantifiable definition", *Advances in Engineering Software,* 37, 5, pp. 327-333.

Ordish, Johan, Murfet, Hannah e Hall, Alison (2019): *Algorithms as medical devices* (Cambridge, PHG Foundation).

Pagallo, Ugo (2014): *Il diritto nell'età dell'informazione* (Torino, Giappichelli).

Pagallo, Ugo e Quattrocolo, Serena (2018): "The impact of AI on criminal law, and its twofold procedures", in Barfield, Woodrow e Pagallo, Ugo (a cura di), *Research Handbook on the Law of Artificial Intelligence* (Cheltenham-Northampton, Edward Elgar Publishing), pp. 385-409.

Palazzo, Francesco Carlo (2016): *Corso di diritto penale. Parte generale,* VI ed. (Torino, Giappichelli).

Palazzo, Francesco Carlo (2021): "I profili sostanziali della riforma penale", *Sistema penale* online, 8 settembre.

Paliero, Carlo Enrico (2004): "La responsabilità penale della persona giuridica: profili strutturali e sistematici", in De Francesco, Giovannangelo (a cura di), *La responsabilità degli enti: un nuovo modello di giustizia "punitiva"* (Torino, Giappichelli).

Paliero, Carlo Enrico (2008): "La Società punita: del come, del perché e del per cosa", in *Rivista italiana di diritto e procedura penale*, pp. 1516-1545.

Perini, Chiara (2002): "Rischio tecnologico e responsabilità penale. Una lettura criminologica del caso Seveso e del caso Marghera", in *Rassegna Italiana di Criminologia,* pp. 389-412.

Piergallini, Carlo (1997): *Attività produttive e imputazione per colpa: prove tecniche di diritto "penale del rischio",* in *Rivista italiana di diritto e procedura penale,* pp. 1473-1495.

Piergallini, Carlo (2004): *Danno da prodotto e responsabilità penale, profili dogmatici e politico-criminali* (Milano, Giuffrè).

Piergallini, Carlo (2005): "Il paradigma della colpa nell'età del rischio: prove di resistenza del tipo", *Rivista italiana di diritto e procedura penale*, 48, pp. 1684-1703.

Piergallini, Carlo (2007): "La responsabilità del produttore: una nuova frontiera del diritto penale?", *Diritto penale e processo,* 9, pp. 1125-1130.

Piergallini, Carlo (2017): voce *Colpa,* in *Enc. Dir., Annali*, vol. X, pp. 222-265.

Piergallini, Carlo (2020), "Intelligenza artificiale: da mezzo ad autore del reato?", in *Rivista italiana di diritto e procedura penale,* 4, pp.1745-1774.

Pisani, Nicola (2012): *La "colpa per assunzione" nel diritto penale del lavoro. Tra aggiornamento scientifico e innovazioni tecnologiche* (Napoli, Jovene Editore).

Piva, Daniele (2011): *La responsabilità del «vertice» per organizzazione difettosa nel diritto penale del lavoro* (Napoli, Jovene editore).

Piva, Daniele (2020): "Spunti per una riscoperta della colpa per assunzione", in Bondi, Alessandro e Fiandaca, Giovanni *et al.* (a cura di), *Studi in onore di Lucio Monaco,* (Urbino, Urbino University Press), pp. 1275-1282.

Piva, Daniele (2022): "Machina discere, (deinde) delinquere et puniri potest", in Giordano, Rosaria, Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale. Persona, Mercato, Amministrazione,* Giustizia (Milano, Giuffrè), pp. 681-693.

Preziosi, Stefano (2021): *La causalità penale nell'orizzonte della "scienza nuova"* (Napoli, Jovene editore).

Preziosi, Stefano (2022): "La responsabilità penale per eventi generati da sistemi di IA o da processi automatizzati", in Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale,* cit., pp. 713-726.

Pulitanò, Domenico (2019): *Diritto penale*, VIII ed. (Torino, Giappichelli).

Randell, Brian, Lee, Pete e Treleaven, Philip (1978): "Reliability Issues in Computing System Design", *ACM Computing Surveys*, 10, vol. 2, pp. 123-165.

Riondato, Silvio (2014): "Robotica e diritto penale (robot, ibridi, chimere, "animali tecnologici")", in Provolo, Debora, Riondato, Silvio e Yenisey, Feridun, *Genetics, Robotics, Law, Punishment*, (Padova, Cedam), pp. 589-609.

Russell, Stuart e Norvig, Peter (2010): *Artificial Intelligence: A Modern Approach,* III ed. (Hoboken, New Jersey, Prentice Hall).

Scott, Colin (2010): "Standard-Setting in Regulatory Regimes", in Cave, Martin e Baldwin, Robert e Lodge, Martin (a cura di), *The Oxford Handbook on Regulation,* (Oxford, Oxford University Press), pp. 104-119.

Sgubbi, Filippo (1975): *Responsabilità penale per omesso impedimento dell'evento* (Padova, Cedam).

Simmler, Monika e Markwalder, Nora (2019): "Guilty robots - Rethinking the nature of culpability and legal personhood in an age of artificial intelligence", *Criminal Law Forum,* 30, pp. 1–31.

Tiraboschi, Michele (a cura di) (2008): *Il testo unico della salute e sicurezza nei luoghi di lavoro. Commentario al decreto legislativo 9 aprile 2008, n. 81* (Milano, Giuffrè).

Tripodi, Andrea Francesco (2022): "Abusi di mercato e trading algoritmico", in Giordano, Rosaria, Panzarola, Andrea, Police, Aristide, Preziosi, Stefano, Proto, Massimo (a cura di), *Il diritto nell'era digitale,* cit., pp.745- 755.

Turano, Alessandro (2020): "Robotica e roboetica: questioni e prospettive nazionali ed europee", in Alpa, Guido (a cura di), *Diritto e intelligenza artificiale. Profili generali, soggetti, contratti, responsabilità civile, diritto bancario e finanziario, processo civile* (Pisa, Pacini Editore), pp.125-160.

Zirulia, Stefano (2018): *Esposizione a sostanze tossiche e responsabilità penale* (Milano, Giuffrè).

Diritto Penale Contemporaneo

RIVISTA TRIMESTRALE

REVISTA TRIMESTRAL DE DERECHO PENAL

A QUARTERLY REVIEW FOR CRIMINAL JUSTICE

http://dpc-rivista-trimestrale.criminaljusticenetwork.eu