Diritto Penale Contemporaneo

RIVISTA TRIMESTRALE

REVISTA TRIMESTRAL DE DERECHO PENAL
A QUARTERLY REVIEW FOR CRIMINAL JUSTICE

1/2023

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 2023 - 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

Intelligenza artificiale e diritto penale

Inteligencia artificial y derecho penal

Artificial Intelligence and Criminal Law

Criminal compliance e nuove tecnologie 1
Criminal compliance ***y nuevas tecnologías***
Criminal Compliance and New Technologies
Luca D'Agostino

La responsabilità penale del produttore di sistemi di intelligenza artificiale 26
La responsabilidad penal del fabricante de sistemas de inteligencia artificial
The Criminal Liability of Artificial Intelligence System Manufacturers
Beatrice Fragasso

AI and Criminal Liability. Algorithmic Error and Human Negligence in the Context of the European Regulation 46
IA e responsabilità penale. Errore dell'algoritmo e colpa della persona fisica nel contesto della regolamentazione europea
IA y Responsabilidad Penal. Error de algoritmo y culpa de la persona natural en el contexto de la regulación europea.
Marta Giuca

La responsabilità penale al tempo di ChatGPT 70
La responsabilidad penal en la era de ChatGPT
Criminal Liability in the Era of ChatGPT
Leonardo Romanò

Speciale su "sicurezza dello stato e poteri investigativi paralleli"

Especial sobre "seguridad del estado y facultades investigativas paralelas"

Special on "State Security and Parallel Investigative Powers"

Speciale su "Sicurezza dello Stato e poteri investigativi paralleli". Premessa 92
Especial sobre "Seguridad del Estado y facultades investigativas paralelas". Premisa
Special on "State security and parallel investigative powers". Introduction
Donatella Curtotti

Agenzia per la _cybersicurezza_ nazionale, sicurezza della Repubblica e investigazioni dell'Autorità giudiziaria 97
Agencia Nacional de Ciberseguridad, Seguridad de la República italiana e investigación judicial
National Cybersecurity Agency, Security of Italian Republic and Judicial Investigation
Federico Niccolò Ricotta

Contents

Le indagini d'*intelligence* e gli strumenti d'intercettazione preventiva 114
Investigaciones de inteligencia y herramientas de interceptación preventiva
Intelligence Investigations and Preventive Interception Tools
Wanda Nocerino

Le inchieste dell'agenzia nazionale per la sicurezza del volo e i limiti all'attività della polizia giudiziaria 134
Las investigaciones de la Agencia de Seguridad Aeronáutica y los límites a la actividad de la policía judicial
Investigations by the National Agency for Flight Safety and the Limits to the Activity of the Judicial Police
Ottavia Murro

Securitizzazione dell'Unione europea e poteri concorrenti. Dall'investigazione, alla prevenzione, all'osservazione 145
Securitización y competencias concurrentes en la Unión Europea. De la investigación a la observación y prevención
Securitization and Competing Powers in the European Union. From Investigation to Observation and Prevention
Angela Procaccino

Il focus su...
Focus sobre...
Focus On...

Il rinvio pregiudiziale in ambito penale e i problemi posti dalle sentenze interpretative della Corte di Giustizia 172
La remisión prejudicial en materia penal y los problemas que generan las sentencias interpretativas del Tribunal de Justicia
The Preliminary Reference in Criminal Matters and the Issues Raised by Interpretative Judgments of the Court of Justice
Alessandro Bernardi

The Crime of Money Laundering: A Touchstone for The Principles of *Il Manifesto del diritto penale liberale e del giusto processo*
***Il reato di riciclaggio: un banco di prova per i principii del* Manifesto del diritto penale liberale e del giusto processo**
***El delito de lavado de activos: una prueba para los principios del* Manifiesto del derecho penal liberal y del debido proceso** 213
Matthias Jahn, Federica Helferich

"Gimme Shelter": The Right to Silence for Silenced Migrant Victims
"Gimme Shelter"*: il diritto al silenzio per le vittime migranti silenziate*
"Gimme Shelter"*: el derecho al silencio por las víctimas migrantes silenciadas* 227
Sara Bianca Taverriti

Intelligenza artificiale e diritto penale
Inteligencia artificial y derecho penal
Artificial Intelligence and Criminal Law

1 ***Criminal compliance* e nuove tecnologie**
Criminal compliance ***y nuevas tecnologías***
Criminal Compliance and New Technologies
Luca D'Agostino

26 **La responsabilità penale del produttore di sistemi di intelligenza artificiale**
La responsabilidad penal del fabricante de sistemas de inteligencia artificial
The Criminal Liability of Artificial Intelligence System Manufacturers
Beatrice Fragasso

46 **AI and Criminal Liability. Algorithmic Error and Human Negligence in the Context of the European Regulation**
IA e responsabilità penale. Errore dell'algoritmo e colpa della persona fisica nel contesto della regolamentazione europea
IA y Responsabilidad Penal. Error de algoritmo y culpa de la persona natural en el contexto de la regulación europea.
Marta Giuca

70 **La responsabilità penale al tempo di ChatGPT**
La responsabilidad penal en la era de ChatGPT
Criminal Liability in the Era of ChatGPT
Leonardo Romanò

AI and Criminal Liability. Algorithmic Error and Human Negligence in the Context of the European Regulation

IA e responsabilità penale. Errore dell'algoritmo e colpa della persona fisica nel contesto della regolamentazione europea

IA y Responsabilidad Penal. Error de algoritmo y culpa de la persona natural en el contexto de la regulación europea

Marta Giuca
Ph.D. in Criminal Law, University of Catania
marta.giuca@unict.it

Artificial Intelligence, EU Law | Intelligenza artificiale, Diritto UE | Inteligencia artificial, Derecho UE

Abstracts

Drawing on European interventions in the field of Artificial Intelligence (in particular the Proposal for a Regulation of April 2021 (AI Act)), the article reflects on the apportionment of responsibilities between the manufacturer and the user of AI systems when a negligence offence occurs due to an error of the algorithm, defined here as "artificial negligence". It is argued that the manufacturer's liability could be assessed for non-compliance with rules established by written norms (case of "specific negligence") or through the reasonable man standard (case of "generic negligence"). For this purpose, a notion of "artificial diligence" is given since it is argued that the reasonable manufacturer parameter will be modelled on the product that complies with the characteristics specified by law and is safe at the same time. Then, discussing the hypothesis of artificial negligence, a distinction between cases of errors that are ex ante foreseeable and unforeseeable is offered in order to address the manufacturer's liability and to define his duty of care. As far as the user is concerned, the duty of information, vigilance and to intervene are investigated, to conclude that compliance with the duty of "human oversight" in the assessment of negligence should be ascertained *in concreto*, in order to evaluate whether to exclude culpability or even the objective dimension of negligence, according to the reasonable man standard.

Traendo spunto dagli interventi europei in tema di Intelligenza Artificiale (in particolare la Proposta di Regolamento dell'aprile 2021, (AI Act)), il contributo offre una riflessione sulla riparto delle responsabilità tra il produttore e l'utente di sistemi di IA, allorché sia commesso un reato colposo di evento a causa di un errore dell'algoritmo, ipotesi qui definita come "negligenza artificiale". Si sostiene che la responsabilità del produttore può essere fondata sull'inosservanza di regole scritte (ipotesi di colpa specifica) o stabilita attraverso il parametro del produttore modello (ipotesi di colpa generica). A tal fine, viene fornita una nozione di "diligenza artificiale", in quanto si sostiene che il parametro del produttore modello sarà calibrato sulle caratteristiche del prodotto che rispetta i requisiti prescritti dalla legge ed è allo stesso tempo sicuro. In seguito, discutendo l'ipotesi di negligenza artificiale, si propone una distinzione tra casi di errori ex ante prevedibili e imprevedibili, al fine di affrontare la responsabilità del produttore e definire il contenuto del suo dovere di diligenza. Quanto alla posizione dell'utente, si indagano i doveri di informazione, vigilanza e intervento, per concludere che l'adempimento del dovere di "sorveglianza umana" nella valutazione della responsabilità colposa deve essere accertato in concreto, al fine di valutare l'opportunità di escludere la colpevolezza o anche la dimensione oggettiva della colpa, secondo lo standard dell'agente modello..

Basándose en las intervenciones europeas en el ámbito de la Inteligencia Artificial (en particular, la Propuesta de Reglamento de abril de 2021 (AI Act)), el artículo reflexiona sobre la división de la responsabilidad entre el productor y el usuario de sistemas de IA cuando se produce un delito imprudente debido a un error del algoritmo, aquí definido como "negligencia artificial". Se argumenta que la responsabilidad del productor podría fundarse en el incumplimiento de las normas establecidas por estándares específicos de conducta ("negligencia específica") o mediante el estándar de la persona razonable ("negligencia general"). Para ello, se aporta una noción de "diligencia artificial", en la que se sostiene que el estándar del productor razonable tendrá como modelo el producto que cumple las características especificadas por la ley y que, al mismo tiempo, es seguro. A continuación, discutiendo la hipótesis de la negligencia artificial, se propone una distinción entre los casos de error que eran previsibles desde una perspectiva *ex ante*, y aquellos que eran imprevisibles, con el fin de abordar la responsabilidad del productor y definir su deber de diligencia. Por parte del usuario, se investigan los deberes de información, vigilancia e intervención, para concluir que el cumplimiento del deber de "vigilancia humana" en la apreciación de la negligencia debe constatarse *in concreto*.

Sommario

1. Introduction. – 2. The field of investigation: criminal negligence. – 3. The position of the manufacturer. – 3.1. Artificial diligence. – 3.2. Ontological and nomological basis of artificial foreseeability. – 3.3. Artificial Negligence. – 4. The position of the user. – 4.1 Duty of information. – 4.2. Duty of vigilance and duty to intervene – 4.2.1. Situations actualizing the duty to intervene. – 4.3. When will the user be liable? Insufficiency of the breach of the duty of care. – 5. Conclusion.

1. Introduction.

That of AI is a topic that draws attention on the relationship between law and technology. The legal issues that emerge are many, from the problem of regulating the production of AI systems, which is called into a confrontation with the right of free economic initiative, to that of transparency, which clashes with intellectual property rights protected by secrecy, stretching onto the issue of liability. What can be seen, in the face of these technological innovations, is a 'thirst for law', now that man has created a new world, he cannot escape what is implied by the creative act of a new reality: ordering it and giving it laws[1].

In the European context, the reaction to the new technological phenomenon of AI was quite immediate[2]. The institutions quickly realised the importance of dealing with a sector that offers new opportunities but also brings with it new risks[3].

Certainly, the EU already has a solid regulatory framework that lends itself to regulating certain aspects of new technologies[4]. Despite such legislative apparatus, the damage caused by AI systems has been considered a field that needs regulatory intervention, which would take into account the characteristics of the new technological products. Indeed, AI applications present themselves as *complex, opaque, open, autonomous, unpredictable systems, in a dependent relationship with data and vulnerable*[5]. It follows that these peculiarities open up new scenarios in terms of liability for damage caused by AI systems, before which the European legislator certainly did not remain silent[6]. As it was observed, the European approach to AI is a "regulatory" one, whose purpose is to establish norms for new technological phenomena, with the ambition to render the European one a model to be imitated by other geopolitical regions[7].

In general, the idea that emerges from the European Institutions' interventions on the topic of AI is that of regulation both upstream, aimed at guiding production, and downstream, aimed at protecting those who suffer damage. Such an approach is fully in line with the dual role of liability, which must certainly guarantee fair compensation to those who suffer damage, but at the same time must constitute an incentive to avoid causing damage or harm *ab origine*[8].

It has rightly been observed that the issue of liability in the field of AI is perhaps the most

[1] Cotta (1968), p. 82.

[2] Already in 2017 the European Parliament adopted the Resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)), *www.europarl.europa.eu*.

[3] *Report from the Expert Group on Liability and New Technologies - New Technologies Formation*, *Liability for AI and other emerging technologies,* (European Commission 2019), p. 32. For a Critical discussion of such a Report, see Bertolini, Episcopo (2021), 644-659.

[4] C*ommunication from the Commission to the European Parliament, the Council, the European Economic and Social Committee, and the Committee of the Regions "Building Trust in Human-Centric Artificial Intelligence"* COM(2019) 168 final, 8.04.2019, par. 2, p. 2. In particular, data protection, by design, is guaranteed by the General Data Protection Regulation (GDPR); as for non-personal data, their free movement and processing in Europe is guaranteed by the Regulation on the free movement of Data (Regulation (UE) 2018/1807, 14 November 2018, on the free movement within the EU for non-personal data) and finally, the Regulation on cybersecurity (Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013) helps to create a climate of confidence in the operations carried out online.

[5] *Report from the Expert Group on Liability and New Technologies*, p. 32 et seq., but also COM (2020) 64 final del 19.02.2020, which refers to the document of the High-Level Expert Group.

[6] The Resolution of the European Parliament of 2017 was then followed by some Communication from the European Commission (see, for example, *Artificial Intelligence for Europe,* COM (2018) 237 final, 25 April 2018; *Coordinated Plan on AI,* COM (2018) 795 final, 7 December 2018; COM(2019) 168 final, 8 April 2019), the White Paper of the European Commission "*An European approach to excellence and trust"* of 19 February 2020, the *European Parliament resolution of 20 October 2020 with recommendations to the Commission on a civil liability regime for artificial intelligence* (2020/2014(INL)), the *European Commission Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain union legislative acts*, of April 202, COM/2021/206 final, and the *Commission Proposal for a Directive on adapting non-contractual civil liability rules to artificial intelligence*, of September 2022. For an in-depth analysis of the AI Act Proposal, see Camardi (ed.) (2022).

[7] Is it the so-called "Brussels effect", *see* Bradford (2020).

[8] See par. A., *Motion for a European Parliament Resolution with recommendations to the Commission on a civil liability regime for artificial intelligence (2020/2014(INL)).*

delicate from a legal point of view since it also involves profiles linked to economic growth and the development of research[9]. And indeed, the AI sector is a tester for the coexistence of law and technology, a coexistence in which limits must be drawn not only on technology by law, but also on law by technology. A balance must therefore be found between regulation and industrialization. New rules must not constitute an obstacle in developing and using new technologies but should rather serve as a means of orienting the production towards the pursuit of improving people's living conditions, providing that, in some cases, AI systems can even reduce the exposure to the danger of interests protected by law[10].

Therefore, the unprecedented scenario of AI systems opens the field to new areas of regulation, but not necessarily to an entirely new regulation. New laws do not always appear indispensable in the face of a new phenomenon. This is the direction taken by the debate in the European institutions on liability for AI systems, which always comes down to recognising the relevance of the regulations already in place, especially those concerning defective products, which need to be complemented by new regulations based on the new characteristics of AI products[11].

In line with such an approach, this article explores the field of criminal liability for the production and the use of AI systems, applying the traditional categories of criminal law, leaving apart, at least in such a context, the idea of criminal liability of the AI system, in order to reflect on possible subjects of intervention for national legislators.

2. The field of investigation: criminal negligence.

In dealing with the categories of "the criminal law of the result" in the face of the new risks created by the production and use of intelligent systems, the investigation plan chosen here is that of negligent offence, which is the privileged field of application of studies on the criminal law of risk[12].

The renewed vitality of the negligent offence model has definitively asserted itself with the transition to the risk society and has found further confirmation in this current historical era which is characterized by what doctrine often refers to as the fourth revolution[13], a time in which man is no longer the only "informational organism", but is assisted in his activities by intelligent artificial agents to perform certain tasks and duties.

In this context, the positions that come to the fore are those of the manufacturer and the user and the subject of product damage knows a new extension that pushes scholars to go beyond the traditional 'human perspective' of approaching the subject and to "measure themselves against the performance of the *new entry AI*"[14]. If the actions in which an AI system is involved become criminally relevant, they will be so mainly in the guise of the negligent offence, commissive and omissive, on which this investigation will focus. Whereas, it is likely that the issues raised by an intentional act on the part of the agent, be it the producer or the user of the intelligent system, will not differ from those that have traditionally been addressed in the study of intentional offences[15] at least in case of a perfect correspondence between the

[9] Bifulco (2018), p. 389.
[10] Ruffolo, Al Mureden (2019), 7, p. 1704 et seq. The Authors refer to driverless cars, which would guarantee higher security standards, significantly reduce road traffic accidents, and would ensure other virtuous effects, such as greater access to traffic also for disabled people.
[11] An example of such an approach is the product liability regime, since, as stated by the Commission, "The Product Liability Directive covers producer's no-fault liability for defective products, leading to compensation for certain types of damages, mainly suffered by individuals", whereas the recent proposal for a Directive "covers national liability claims mainly based on the fault of any person with a view of compensating any type of damage and any type of victim", so that the two legislative interventions "complement one another to form an overall effective civil liability system" (See the AI Liability Directive, p. 3).
[12] But it is well known that reflections on permissible or appropriate risk also apply to intentional offences, on the point. See Militello (1988), p. 55, p. 205 et seq.; Donini (2010), p. 646 et seq. On the topic of AI and criminal negligence offences, see for example Beck (2016), 138-143; *Id.* (2017), 227-252.
[13] Floridi (2017a), p. 99 et seq. The philosopher identifies the first revolution in the discovery of heliocentrism by Nicholas Copernicus, with the publication of his treatise *Sulle rivoluzioni dei corpi celesti* in 1543, a moment from which man would cease to believe in his centrality, at least with reference to planet Earth. The second revolution is made to coincide with the publication of Charles Darwin's *The Origin of Species* in 1859, which pushes man to renounce his centrality in the animal world, having to confront the idea that every living species derives from common ancestors through a process of natural selection. The third revolution came with the psychoanalytic work of Freud, who disproved the idea of the mind thought of as a box that can be known simply by looking inside, showing that many of our actions are the result of the unconscious. The fourth revolution questions man as the only being unsurpassed in intelligence.
[14] Piergallini (2020), p. 1745-1774, particularly p. 1749.
[15] On the topic of intentional offence, see Gless, Silverman, Weigend (2016), p. 425, who conclude that the case does not raise particular

conduct conceived by the author and that occurred *in concreto*[16].

The case of negligent offence is different. Here, questions arise concerning the identification of objective rules of diligence on a praxeological basis or on a legal basis, with particular reference to the role of the criteria of foreseeability and avoidance of the result, and the content of the duty of care with specific regards to the position of the producer and the user. The apportionment of liability between these two figures also appears problematic, especially in the presence of the self-learning mechanisms of AI systems.

In the background remains the broader problem of balancing the performance of a dangerous but socially useful activity, such as the production of intelligent systems, and the protection of the interests threatened by it, which prompts us to question the extent of the area of "acceptable risk" (*Erlaubtes Risiko*)[17].

In production activities, the "acceptable risk" involves two kinds of responsibility: "for the type of production" and "for the mode of production"[18]. With specific regard to the producer's criminal liability for product damage, scholars have observed that this is a "*transversal" liability*, embracing both the type and the mode of production[19].

3. The position of the manufacturer.

In the context of this study, the duty of care will be addressed to the manufacturer and the user and may be a written rule, thus setting the groundwork for the assertion of "specific negligence", for which it is unnecessary to establish the violation of the duty of care according to the reasonable person parameter[20].

The tendency to establish written rules of conduct is more and more pronounced in the technological age, while unwritten rules are characteristic of "more technologically tranquil and restful eras"[21]. This trend is not exempt from the context of the production of AI technologies, which already has European regulations that set the criteria for the production chain.

The proposal of the AI Act of April 2021[22] establishes rules regulating the production of high-risk AI systems with a precautionary purpose, and thus intended for producers. They are contained in Chapters II and III of Title I. However, as this is only a proposal for a regulation, at present such rules are not yet in force.

Chapter II, in setting out the "requirements for high-risk AI systems", already identifies quite specific obligations addressed to the producer. For instance, it is stipulated that high-risk intelligent products must be equipped with risk management systems for the entire life cycle of the system (Art. 9), which allows not only to identify known and foreseeable risks but also to take appropriate measures to manage them (para. 2 (d)). The manufacturer is then obliged to eliminate or reduce risks by means of adequate design and development (para. 4(a)) and to put in place measures to mitigate and control risks that cannot be eliminated (para. 4(b)). Hence, according to the meaning of this provision, it would seem that the European legislator requires the manufacturer to introduce into the system forms of emergency control of control-

problems; the conclusion, among others, is also shared by Vagliasindi (2021), p. 375-76; Salvadori (2021), p. 100; Borsari (2019) p. 264; in this regard, according to Basile (2019), "We must, in short, prepare ourselves for an era in which the commission of crimes with the tool of AI could become very frequent and incisive, also because of the increased vulnerability of certain aspects of human life connected to uses of artificial intelligence" and therefore the A. asks "Is it necessary, then, to field new criminal offences (or to remodel existing ones) in order to make them applicable to the implementation of criminal conduct through the AI tool, thus offering protection to legal assets also from this new source of attacks?", pp. 26-27.

[16] On the contrary, a problematic scenario could be that of intentional criminal use of the AI system with an unexpected development of the factual dynamic.

[17] Fundamental works on the topic include Preuss (1974); Roeder, (1962); Hilgendorf (1993); Prittwitz (1993). In the Italian doctrine, Militello, (1988); Forti (1990), pp. 250 et seq.; Gallo (1960) p. 638; Marinucci (1965), p. 210 et seq.; Donini (1989), p. 588; Consulich (2021), p. 1102 et seq.

[18] Bricola (1978), pp. 75 et seq..

[19] Piergallini (2004), p. 46.

[20] The specific negligence is so-called since the judgment on foreseeability and avoidance is established by a written rule laid down by the legislator, the authority, or even a private person (e.g. the owner of a firm), and this makes it unnecessary to establish the violation of the duty of care according to the reasonable person parameter. See *ex multis,* Bartoli (2021) pp. 519 et seq.; Canestrari (2013), p. 144; Mantovani (1988), p. 306.

[21] Mantovani (1988), p. 306.

[22] The Commission's proposal of the EU AI Act will become law once both the Council and the European Parliament agree on a common version of the text. At the time of writing this contribution, the European Parliament's Internal Market Committee and the Civil Liberties Committee adopted a draft negotiating mandate on the AI Act proposal (*www.europarl.europa.eu*).

lable risks and forms of mitigation of uncontrollable risks. In this way, indications are given on the characteristics of the product that will then be launched on the market, which must always be accompanied by a sort of "first aid kit" provided by the manufacturer to the user, useful in the event of risks arising during the use of the intelligent product[23].

Production criteria are also introduced concerning training data sets, which must meet certain quality standards (Art. 10). In addition, detailed technical documentation is required (Art. 11), and the design shall be sufficiently transparent to enable users to interpret the system's output and use it appropriately (Art. 13). High-risk systems must also be robust, *i.e.* resistant to attacks by third parties aimed at modifying their use or performance by exploiting system vulnerabilities (Art. 15(4)).

Rules on production are also contained in Chapter III. One example is the provision of a conformity assessment procedure for systems before they are placed on the market (Art. 19).

From the Proposal for a regulation emerge many rules of diligence that could be classified as second-rate, *i.e.* aimed at preventing risks and not specific events (most of the rules set out above, in my opinion, can be brought into this category, starting from Article 9 on the risk management system, then moving on to the one on data set requirements in Article 10, and concluding with those in Article 13 on transparency and interpretability of output and Article 15 on the robustness of systems), as well as rules with a monitoring purpose and rules with an administrative attitude and a precautionary purpose only in a mediated way (this is the case of the certification and conformity assessment regime).

According to the doctrine[24], such characteristics of the duty of care make necessary the ascertainment of the generic negligence in the light of the reasonable man standard. This represents a guarantee for the defendant: his negligence must be ascertained *in concreto*, according to the parameter of *homo eiusdem professionis et condicionis,* since it is not sufficient to state that he behaved contrary to a written diligence rule with generic or technical-administrative content, and which is not intended to prevent a specific result[25].

In the case of criminal product liability, the reasonable manufacturer parameter will be modelled on the product that complies with the characteristics specified by law and is safe at the same time. Then to ascertain negligence we will ask whether the product meets the standards of the defect-free product.

Having thus drawn our line of enquiry into product liability, we need to dwell on the characteristics that the intelligent product must have to be defect-free.

3.1. *Artificial diligence.*

Since intelligent systems are characterised by their ability to make decisions, we could consider that technological devices function according to a certain "behaviour", which is a response-output to a particular input[26]. This "behaviour" of the system is not left to chance, but is guided by the "information" fed into it during the design and training phase. It may be that such information coincides with social norms, *i.e.* rules of actions used to constrain the AI system's behaviour[27], in a way that guarantees the successful coexistence of multiple programs[28], so that AI systems are defined as Normative Multi-Agent Systems (NMAS)[29]. Indeed, according to normative computing theories, the system's (agent) behaviour is guided by norms, which are encoded by the designer, in an off-line design approach, or inferred from

[23] There will be negligence on the part of the developer or programmer for not providing manual intervention on the system in emergency conditions, see Salvadori (2020), p. 103.
[24] Di Giovine (2003), pp. 391 ff.
[25] However, such a conclusion is not always shared by Italian case law. It is then considered that the failure to comply with a duty of care imposed by law, regulation, order, or discipline is sufficient to prove negligence, provided that the event that occurred is attributable to the type of event that the duty of care is intended to prevent (see, *ex multis,* Cas. Pen. sez. IV, 01/12/1989, n.1501; Cass. pen. sez. IV, 08/11/2022, (ud. 08/11/2022, dep. 01/02/2023), n.4155; Cass. Pen. sez. IV, 17/05/2022, (ud. 17/05/2022, dep. 23/05/2022), n.20035).
[26] Lagioia, Sartor (2020), p. 434, arguing that "under an appropriate level of abstraction", AI systems have cognitive attitudes (intentions, beliefs, awareness) which are relevant for the realisation of *mens rea* and might be taken into consideration to appropriately react to their harmful behaviour.
[27] Hollander, Wu, (2011) , par. 1.3 and 2.18.
[28] Shoham, Tennenholtz (1992), p. 276.
[29] Andrighetto, Governatori, Noriega, der Torre (2013).

the environment through machine learning techniques, in the bottom-up approach[30].

In the case of 'agent' products, *i.e.* those that move in a real or virtual environment[31] and make decisions, as intelligent systems do, we could then say that the free-defect good is that artefact that "acts with diligence"[32]. It is clear that this diligence cannot be equated with the human one, because the machine, though intelligent, is different from humans[33]. This is an "artificial" diligence, in the sense that it derives from the knowledge that the AI system has acquired during the production phases, or from the information learned during its training or/and use. In order to establish the diligence behaviour requested from the machine, the parameter will not be that of the reasonable man, but that of the reasonable algorithm, with a consequent transition from the *homo eiusdem professionis et condicionis* to the *machina eiusdem fabricationis et condicionis.*

As a result, social norms aimed at governing an agent's behaviour may consist in rules of conduct that, just as they direct the behaviour of human agents, are intended to orient the behaviour of the artificial agent called upon to perform a function typically intended for humans. One need only think of the conduct of driving on the road or that of medical diagnosis. The driver, before driving, is required to know the rules of the highway code. Similarly, the doctor, before making a diagnosis, is required to know the rules that allow a correct diagnosis to be made and to behave in essence in the manner prescribed by the *leges artis*.

When these tasks are delegated to an AI system, such a claim to knowledge of the rules of the road and the *leges artis* is no longer addressed to the human agent, but to the artificial one[34]. It follows that, in this context, the duty of care expressed by a rule of conduct becomes a cognitive element that must be incorporated into the intelligent system through training, and that contributes to forming its knowledge[35], in line with the theory of normative multi-agent systems (NMAS).

Thus, cognitive competencies of AI systems are preconditions for criminal liability[36], as will be discussed in the following paragraphs.

3.2. *Ontological and nomological basis of artificial foreseeability.*

Some scholars argue that AI systems can achieve "situation awareness", which consists of three steps: the perception of the elements of the environment where the systems act, the comprehension of the current situation through the integration of all the disjointed information collected during the perception step, the projection of future action in the environment[37].

The process that leads to situation awareness characterises the first moment of the artificial systems' "behaviour", which ends with a decision corresponding to the output. Therefore, a distinction between an "internal behaviour", identified in the achievement of the situation

[30] Savarimuthu, Cranefield (2009), p. 6. The A. describe also a third approach defined as norm-entrepreneur. On the topic of translating law into the algorithm and the two approaches (top-down and bottom-up) ungern-sternberg (2018), p. 262.

[31] As Lagioia, Sartor (2020), p. 441, correctly underline, both AI systems with or without a physical presence (think of robots, for the former, and software agents and bots for the latter), can fulfill the conducts requirement of an *actus reus*.

[32] The idea of a standard of care referring directly to the artificial system is taken into consideration in the area of civil liability, where it is customary to refer to the so-called 'reasonable algorithm' with regard to ML algorithms, which autonomously make decisions. It is understood that the 'reasonable algorithm' standard will find a safe place of application when algorithms are recognised as having legal personality, thus being equated with humans (*see* Abbot (2020), p. 69) it is held that it may be relevant even where the algorithm is not considered an independent center of imputation, since such an assessment serves to direct the judgment on the producer's conduct towards a form of liability requiring a different burden of proof from that required in product liability; rather, the producer's diligence should be assessed in light of the reasonable algorithm standard, Chagal-Feferkorn, (2018), pp. 111-148. However, the foreign doctrine does not agree on the subject. Some opinions emphasise that it is not necessary to examine algorithmic reasonableness (diligence), given that algorithms are not comparable to human agents, but are still tools that can be used by humans, Colonna (2012); Balkin (2017), underlines that *"there is no little person inside the program"* and that algorithms take the decisions for which they have been programmed so that there is no point in examining algorithmic reasonableness independently of that of the programmers.

[33] La Vattiata (2023), p. 492.

[34] Ungern-Sternberg (2018), p. 252 with regards to the case of autonomous cars.

[35] As it emerges from the *Commission Implementing Regulation (EU) 2022/1426, of 5 August 2022, laying down rules for the application of Regulation (EU) 2019/2144 of the European Parliament and of the Council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles cit.,* providing that automated driving systems (ADS) shall comply with traffic rules. Such knowledge of traffic rules changes according to the level of autonomy, as explained by Heikoop, Hagenzieker, Mecacci, Calvert, Santoni De Sio, Arem (2019). On the topic, see also Ungern-Sternberg (2018), p. 262.

[36] Lagioia, Sartor (2020), p. 435.

[37] *Ibidem,* p. 441.

awareness, and an "external behaviour", corresponding to the output reflected to the external environment (which in NMAS should rely on a given norm of conduct) can be made.

At this point, we could define as *diligent* the so-called *norm-abiding behaviour*[38], *i.e.* the behaviour of the system that recognises a given situation through the correct interpretation of the input data (internal behaviour) and that produces an output that conforms to that required by a norm (external behaviour). On the contrary, the behaviour of the system that errs in the interpretation of the input data, therefore it does not correctly recognise the situation in which it is called to intervene and produces an output that differs from the one that would be required by the observance of the norm, or the system that, although correctly interpreting the input data, produces an output that is inconsistent with the interpretation of the input due to a defect of internal functioning, will be *negligent*.

If the error of the system produces or contributes to the production of a risk that results in an offence to one or more legally protected interests, this negligence will be causal with respect to the production of an injury or endangerment of interests. This opens up the scenario of the producer's criminal liability for having produced and placed on the market a defective and unsafe good.

Let us consider the case of the system that errs in the interpretation of the input data, thus not correctly recognising the situation in which it is called upon to intervene, and produces an output that differs from what would normally be required. The question we must ask ourselves is first of all what factors enable the system to recognise a given piece of information and avoid producing an output that is harmful to a certain legal interest.

According to criminal scholars, *recognisability* and *avoidance* are the parameters of the probabilistic judgement of the negligent fact used for the *ex ante* verification of the adequacy between the conduct carried out and the result[39]. We can then draw on the doctrine's reflections concerning the parameter of recognisability to develop them on the terrain of the responsibility of the producer of the intelligent system.

The starting point is that under the theory of the "double measure of negligence"[40], the *Erkennensollen* relevant to the objective measure of negligence is composed of a *nomological* basis and an *ontological* one, constructed in relation to a reasonable person[41]. In particular, causal laws make up the nomological component, while the ontological one encompasses all the factual elements capable of connecting nomological knowledge to the concrete situation, in the sense that the presence of certain initial conditions enables the agent to read the concrete situation in the light of causal laws.

This scheme could be re-proposed for the intelligent agent who is required to recognise the factual situation in order to avoid reaching a decision that endangers legal interests.

We will say, then, that in a NMAS the nomological basis of recognisability consists in the duty to possess norms or models of knowledge representation that allow causal correlations between input and output to be identified, while the ontological basis is substantiated in the duty to understand the elements of the factual situation in which the system operates through the techniques of knowledge acquisition and the sensors, in case of an artificial physical system, or the instruments for tracking activities, in case of software, that make the external environment perceptible.

The nomological component of artificial recognisability requires us to dwell further on the characteristics of artificial learning. When designing AI systems, the idea of building thinking machines, mechanical brains that emulate human cognitive capacities, has been abandoned[42]. Nowadays, in fact, science is not yet able to explain how exchanges between neurons translate into 'consciousness', and consequently technology is not capable of building a system with its own thinking and creativity[43].

[38] Conte, Castelfranchi (1993), p. 40.
[39] Exner (1910), p. 137 ff.; In the Italian doctrine, see, *ex multis,* Forti (1990); Castronuovo (2009). Scholars highlight the difference between the concept of recognisability (*Erkennbarkeit*) and foreseeability (*Voraussehbarkeit*), *see,* for example, Schroeder (2003), §16, par. 128, (68)-(69); Gallo (1960), p. 638.
[40] This theory highlights the distinction between the objective element and the subjective element of negligence, the first concerning the breach of the duty of care according to the standard of the reasonable agent, the latter dealing with the personal capabilities of the agent and the specific circumstances of the case. See, for example, in the traditional German doctrine, Hippel (1908), p. 568 et seq.; Engisch (1995), p. 349 et seq.; Jescheck, Weigend (1996), p. 561 et seq.; among the Italian Authors, see De Francesco (1977-78), pp. 275 et seq.
[41] Forti (1990), p. 211 et seq., 233.
[42] For a clear explanation of the technique of machine learning, see Searle (1990), p. 26-32.
[43] Amato (2020), p. 86.

The machine's lack of reasoning capacity means that it is not able to understand the meaning of a piece of data, rather it merely establishes a correlation[44], associating the data with a certain result that may consist of a behaviour to be adopted. For example, the algorithm learns to recognise the 'stop' signal through the data fed into the system and the instructions given by the trainer; however, this will not be sufficient to command the autonomous car to drive: the algorithm will also need to be instructed to stop as soon as it recognises the signal in question[45]. A correlation is then established between the input (the 'stop' signal) and the output (the vehicle's stopping behaviour), and to achieve such a result the algorithm must be 'fed' not only with data but also with norms describing behavioural response practices to the input classification. Stopping in front of a stop signal will thus be achieved without the system understanding the reasons for such a behavioural command. Therefore, the system is unable to develop argumentative reasoning as a human brain does. Machines are not required critical thinking, they are required to work properly[46].

I do not think that this can be totally contradicted by the ability of AI systems to acquire 'situation awareness'. After all, the notion of 'situation awareness' does not include, nor is identifiable with that of 'situation comprehension'. Being aware of something means "knowing that something exists, or having knowledge or experience of a particular thing", while comprehension implies "understanding something completely", "knowing the meaning of something"[47], something that machines are not yet capable of doing. This is in line with the idea that AI is not "about coupling artificial agency and intelligent behaviour into new artefacts", but the opposite: "AI is about decoupling successful problem solving from any need to be intelligent", because "[i]t is only the outcome that matters, not whether the agent or its behaviour is intelligent. Thus, AI is not about reproducing human intelligence, it is about doing without it"[48]. AI systems have an "unconscious knowledge" that could be better defined as mere "information"[49].

All this leads us to conclude that 'artificial diligence' does not imply that the algorithm understands the meaning of the instructions given during the learning phase[50]. The direct consequence of this is that the 'thinking mind' during the operation of an AI system always remains the human one, more precisely, that of the trainer in the training phase, while in the moment of deployment, it is that of the user.

3.3. *Artificial Negligence.*

These premises on the structure of human and artificial recognisability of a factual situation are useful for the study of pathologies in the functioning of the AI system that can give rise to liability for what scholars call "AI-Crime" (AIC)[51]. We can therefore identify two main reasons that hinder perfect learning on the part of the AI system: I) faults in the moment of knowledge representation (nomological basis of machine learning) and II) faults in the moment of knowledge acquisition (ontological basis of machine learning), and then there are III) hybrid situations, involving both (I) one and the other (II) basis of machine learning.

I) Since the representation capacity is strictly dependent on the system's programming and the settings chosen during its design, including the set of norms encoded by the designer of a normative "top-down" MAS, the defects found in the knowledge representation phase can be blamed on the producer[52]. According to the classification proposed above, these defects can be

[44] *See* Searle (1980) and Id., (1990); Gigerenzer (2022), p. 122 et seq., and p. 143 et seq.
[45] *See* Gigerenzer (2022), p. 96, and p. 102.
[46] Amato (2020), p. 90.
[47] Definitions from *dictionary.cambridge.org.*
[48] Floridi (2017b), p. 126.
[49] Faggin (2022), p. 54.
[50] See, for this purpose, the interesting example of the school bus proposed by Gigerenzer (2022), p. 100 et seq., and the reflections on the concept of consciousness proposed by Faggin (2022), pp. 128 et seq.
[51] King, Aggarwal, Taddeo, Floridi (2021), pp.195-227
[52] Consider the cases of overfitting and underfitting. If in the training phase, the system learns to recognise the input data, providing an output that corresponds perfectly to expectations, the result will not be satisfactory, because the model is not able to generalise well and will produce errors if fed with data other than those entered in the training phase. In the case of underfitting, the model will present inaccuracies already in the training phase, because it is unable to recognise the input data, and it will continue to err even when it is faced with data other than those used at the time of learning.

classified as construction or manufacturing defects, depending on the fact that they affect the entire production series or only some elements of it.

II) The case of defects at the time of knowledge acquisition is different. An erroneous acquisition of external data may depend, for example, on a sensor malfunction caused by a production defect (being then construction or manufacturing defects[53]), imputable to the manufacturer, or by poor maintenance attributable to the user.

III) Then there are hybrid situations or ones that are more difficult to classify, in which the damaging event is the result of a mixture of limitations (and not defects) of the system found in the ontological and nomological phase. This is the case if we imagine that an incomplete (and not erroneous) acquisition of knowledge can be determined by the peculiarities of the concrete case, where circumstances arise that make the factual situation completely unique and unpredictable and such as to be characterised by an *eccentric risk*. Just think of the case of the sudden crossing of a pedestrian on a street at night that could not be perceived in time by the sensor[54]. Or the case in which a certain piece of data is perceived by the sensors, therefore acquired, but the system is then unable to process it because the unique character of the situation made it unforeseeable by the producer, who therefore did not 'train' the system to recognise such a situation[55].

To avoid flaws of this kind in the knowledge representation and acquisition phase, the producer should be required during the designing or training stage to describe all possible situations of risk exposure of legal interest, a goal that is difficult to achieve[56] because it is impossible to predict the factual details of every single situation in which the system will operate[57]. It is impossible to predict the future.

This, moreover, is in line with the idea that the 'warning signals' that make a certain danger recognisable "must be seen, not foreseen; they are a matter of detection, not of foresight" and that recognisability is of a reconnaissance nature and not of an investigative one, in that the agent is not required to explore all factors that make the offence abstractly possible[58].

Such statements seem to be confirmed by the recent *Commission Implementing Regulation (EU) 2022/1426, of 5 August 2022, laying down rules for the application of Regulation (EU) 2019/2144 of the European Parliament and of the Council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles*. Annex II provides the *Performance requirements in different traffic scenarios*, distinguishing among nominal, critical and failure scenarios. In the case of critical traffic scenarios, it is stated that the automated driving system shall be able to perform the dynamic driving task for *all reasonably foreseeable critical traffic scenarios* in the operational design domain[59]. It follows that the manufacturer's duty of care does not comprehend traffic scenarios which are not reasonably foreseeable, therefore *ex ante* unpredictable[60].

The inexistence of such a demand on the producer becomes apparent if we move from the field of AI applied to cars to that of AI applied to medical science. Think of the case of *new medical knowledge*, in which a new virus manifests itself after a robot has been placed on the market to diagnose infectious diseases. The virus will not be recognised by the system because it is not included among the possible infections and will therefore not be diagnosed. For such

[53] The distinction is elaborated by Piergallini, (2004), p. 46 et seq.
[54] Similarly Ungern-Sternberg (2018), p. 258
[55] As observed by the Panasonic Professor of robotics at MIT, Rodney Brooks, "A lot of technologists think if you do a demo, then that's it. But scaling is what kills you". "You run into all sorts of things that didn't happen at a smaller scale" (commenting on the news of Cruise's robot cars blockade in the streets of San Francisco, reported by Marshall in Wired, Jul. 8, 2022, *www.wired.com/story/cruises-robot-car-outages*). On the issues of criminal liability and self-driving cars, see, *ex multis*, Lohmann (2016); Crane, Logue, Pilz (2017), in the Italian doctrine, Picotti (2021), 813-837; Cappellini (2019), 325-353.
[56] Faggin (2022), p. 72, refers to the case of autonomous driving, arguing that many problems can never be completely solved.
[57] As noted by Selbst (2020), p. 1324, with regards to risks correlated to driverless cars: "Autonomous vehicles will face unexpected changes: detours from road construction, drivers who break traffic laws or stop very suddenly, or other drivers misapprehending what the automated vehicle itself will do and reacting badly. Each of these will be unique in some way—the timing, the type of stimulus—such that the machine cannot possibly be trained on all of them", p. 1324. See also Ruffolo (2020), p. 165.
[58] Giunta (2019), p. 16.
[59] Annex II, par. 2.1.
[60] See Annex III, Appendix I, para. 1 describing the "Generation and classification of scenarios", providing that: "From a qualitative perspective, scenarios can be classified into Nominal/Critical/Failure and correspond to normal or emergency operation. For each of these categories, a data-based approach and a knowledge-based approach can be used to generate corresponding traffic scenarios. A knowledge-based approach utilizes expert knowledge to identify hazardous events systematically and create scenarios. A data-based approach utilizes the available data to identify and classify occurring scenarios. Scenarios shall be derived from the ODD of the fully automated vehicle". These scenarios, once identified, must be assessed through simulation or physical testing (par. 5).

an algorithmic limitation, no blame can be laid at the door of the manufacturer, since the discovery of the virus is subsequent to the programming of the algorithm. On the other hand, a fault may be found if, following the discovery of the virus, the manufacturer does not provide a software update that covers the new pathology[61].

We have thus drawn an initial distinction between situations in which it is possible to provide an *ex ante* description in terms of danger, and situations in which such an *ex ante* description of danger is nearly impossible (case of the *eccentric risk* and *new medical knowledge*), since they are related to concrete dynamics that are difficult to foresee, therefore not imagined in the algorithm's training phase and proposed as examples of the training set.

It should also be considered at this point that the unpredictability of the output result may be, not much due to the inability to condense all the concrete variables in algorithmic programming, but rather to the evolution of the algorithm as a result of self-learning during the utilisation phase[62] (*self-learning* case). Such circumstances are usually defined as "emergent behaviour"[63], since the artificial agent acts beyond the original expectation. That is, while the tests at the design stage reveal a relatively simple behaviour, during its deployment the AI system acts in a more sophisticated way, and might even misalign his conduct with the original design, with possible criminal implications[64].

The emergent behaviour could also result from the interaction among AI systems through cloud computing, a technology that brings together AI systems operating in different environments and leads to an exponential increase in initial knowledge[65].

In all these cases, the knowledge implementation may lead to decision outcomes that differ from those assessed during the manufacturer's validation tests, which are therefore unpredictable[66], with the result that the decision of the ML algorithm is not always attributable to the programming that preceded its release[67].

It can then be observed that "artificial negligence" does not always automatically lead to producer liability[68]. Rather, three typical situations could be identified: a) cases in which the system had to act diligently and did so; b) cases of artificial negligence in which the system did not act as expected; c) cases in which the system did not act as expected but could not have done otherwise, which we could call "temperate artificial negligence" or unattainable diligence. These options are reproduced in the summary table below.

Case a)	Diligent producer	Diligent Algorithm
Case b)	Negligent Producer	Negligent Algorithm
Case c)	Diligent Producer (Acceptable risk)	Negligent Algorithm (Unattainable diligence)

The cases of algorithmic error include the last two type-situations, but the outcomes in terms of liability will be different. If artificial negligence gives rise to producer liability (hypothesis *sub b)*), the same cannot be said for hypotheses of the third type, *i.e.* for cases in which the producer, on the basis of the technical-scientific knowledge available, is unable to design and train the algorithm to cope with the concrete situation in the best way possible so as to avoid a certain risk (cases of *eccentric risk, new medical knowledge* and *self-learning*). This is a margin of risk that the law tolerates when it authorises an activity of production of goods in which the zero margin of error does not exist[69], because the error is inherent in deci-

[61] Chagal-Feferkorn (2018), p. 136.
[62] Hubbard (2014), p. 1851.
[63] *Ibidem;* King, Aggarwal, Taddeo, Floridi, (2021), p. 6.
[64] King, Aggarwal, Taddeo, Floridi, (2021), p. 6.
[65] Borsari (2019), p. 265.; Severino (2020), p. 533; Hubbard (2014), p. 1851, notes that this leads to considerable probations difficulties at trial.
[66] Chagal-Feferkorn (2018), p. 133, emphasises how complexity increases if one considers that many machine learning algorithms improve their knowledge through interactions with the network, as they are online-based (also p. 135 on this point). Furthermore, the author notes that the programmer could certainly set limits to the self-learning capacity, and insert a data selection mechanism, as well as subject any changes to the algorithm's decision-making process to its pre-approval. Nonetheless, this last characteristic could represent an impediment to the algorithm's functionality and usefulness: think of the repercussions such a system would have on self-driving cars, which would cease to function and be usable if they had to wait for the programmer's approval every time they were confronted with new situations (*ibidem,* p. 134).
[67] Cappellini (2022), p. 9.
[68] Borsari (2019), p. 265.
[69] As noted by Selbst (2020), p. 1331: "Because AI will not prevent all accidents, the promise of AI is to reduce—not eradicate—errors. Thus,

sion-making systems that proceed by correlations and generalisations from a set of input data to produce an output[70]. We are therefore faced with the area of *permissible or acceptable risk*[71], in which the producer-agent is not liable for negligence with regard to the damage that could be *in abstracto* foreseeable, but occurred despite the faithful observance of technical rules[72], and which strikes a balance between the opposing needs to protect the threatened legal interests and to carry out useful but inherently dangerous activities, identified by grading the intensity of protection of the exposed protected legal interests according to their value[73].

To hold the manufacturer liable in such cases, it would mean, for the legal system, to contradict itself, on the one hand, authorising the production of high-risk AI systems and, on the other hand, imputing to the manufacturer any foreseeable harmful consequences (if one considers that AI systems based on ML techniques are not by their very nature zero-risk systems) and avoidable only by refraining from the risky activity.

Of course, the conclusion is different in the case of a legislative intervention aimed at criminalizing endangerment, therefore punishing conduct which is not linked to a result, with the consequence of anticipating criminal usage[74]. In any case, it should not be forgotten that the exclusion of the manufacturer's criminal liability is not an obstacle to the application of other sanctions that do not have a criminal nature, but that can be effective in guaranteeing the protection of the victim.

I have therefore emphasised here how, from the point of view of traditional categories of negligent crime, the impossibility of avoiding the occurrence of the negligent act is already relevant in the context of the objective duty of care if the conduct required from the producer in order to avoid the situation of damage or danger falls outside the compendium of duties outlined in the light of the *Maßfigur* instead of the concrete agent. Moreover, the response that we have imagined on the part of the legal system in these cases would be the same as that which occurs when other professional figures, such as doctors, are involved. Just as the duty to save a terminally ill person is not imposed on the doctor by law because there is no treatment that can avoid the inauspicious outcome, neither can the duty be imposed on the manufacturer to produce a system that does not realise risks for which no technology exists that can avoid them.

What is already unattainable on the level of the manufacturer's duty of care, *e.g.* the prediction of the individual and peculiar concrete case when training the system, may however leave room for a claim of a duty of care from the human agent. That is why we must now analyse the position of the user and the content of the duty of care addressed to him/her so that the table we have drawn above can be enriched with a further column, the one describing the position of the user.

4. The position of the user.

The analysis of duties of care and their respective holders in the context of intelligent product damage must at this point turn to the figure of the user.

One of the first observations focuses on the fact that the offence invoking the remedy offered by criminal law will reasonably take place during the user's use of the AI system. It is then precisely from the context of the use of the intelligent product that the investigation of the user's position must begin with observing how the use of the product implies the establishment of a "relationship" between the user and the AI system. Together with what has already been defined as "artificial diligence", special concern must also be given to the human diligence of the user, which has different contents.

when AI is used, there will still be some errors that result in harm". See also Gless *et al.* (2016), p. 426.

[70] Such a feature is taken into consideration in the European Commission's AI Act proposal when stating that a risk management system shall be established, implemented, documented, and maintained in relation to high-risk AI systems (art. 9 (1)). It is then requested the adoption of suitable risk management measures (2 (d)) which "shall be such that any residual risk associated with each hazard as well as the overall residual risk of the high-risk AI systems is judged acceptable, provided that the high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse" (4).

[71] Severino (2020), p. 536

[72] Mantovani (1988), p. 311, on the acceptable risk.

[73] Forti (1990), p. 459 et seq.

[74] Consulich (2022), p. 1051.

4.1. *Duty of information.*

The user's duty of care consists first and foremost of information obligations.

Negligence due to 'failure to inform' naturally presupposes the existence of an apparatus of rules capable of guiding the agent's behaviour in the face of a dangerous situation. Well then, in the context of the activities of production and use of goods, this apparatus of rules may include the conditions of use of the product, even an intelligent one. What follows is quite obviously a reflection on the role that negligence for violation of duties of information has in the panorama of intelligent systems.

As the European approach to AI teaches us, for regulatory purposes it is always good to distinguish between high-risk and low-risk systems, and it is within the first case that scope can be found for user negligence due to failure to inform. The user must be responsible and must realise the complexity of the system in order not to leave its use to improvisation. An indication in this sense comes once again from the proposal of the AI Act of April 2021, which states in Article 29 para. 1 that "users of high-risk AI systems shall use such systems in accordance with the instructions for use accompanying the systems". Indeed, high-risk systems must always be accompanied by instructions for use containing "concise, complete" and "clear" information that is "accessible and comprehensible to users" (Art. 13), in the absence of which negligence on the part of the manufacturer will certainly be found because the product has an information defect[75].

But on closer inspection, in addition to this specific provision concerning the user's obligation of information provided for by the draft regulation for high-risk systems, this obligation to inform can be traced more generically to the duties of social solidarity, given that the user who decides to use a high-risk system creates an area of potentially harmful effects not only for himself but also for third parties who might suffer harm[76].

The user's duty to inform is not new in the product liability scenario; the user is put in a position to comply with it as soon as he is provided with an instruction manual by the manufacturer, which should indicate the potential risks that may arise when using the good. If, therefore, this duty of information exists for all products, it is, however, true that in the case of high-risk intelligent systems, unlike for other products, the duty of information is certainly more stringent because of their greater potential for harm[77]. Moreover, it must be considered that this duty of information can also be reminded to the user by means of software update messages sent through the system by the manufacturer. The user's duty to inform is therefore also characterised by a *duty to update* him/herself, which is added to the initial duty to know the conditions of use of the product, given the capacity of these intelligent systems to "evolve". This is a condition that especially involves machine learning systems, which are subject to change as a result of self-learning.

4.2. *Duty of vigilance and duty to intervene.*

In his or her relationship with the AI system, the user acts as a "human controller", whose supervision of the system's operation is twofold. On the one hand, as a safeguard mechanism, aimed at preventing damage resulting from the system's malfunctioning and on the other hand, as a liability catalyst, *i.e.*, as the subject to whom any avoidable damage is to be attributed[78].

To this twofold function of human control correspond two duties of care: the *duty of vigi-*

[75] In this sense, *see Commission Implementing Regulation (EU) 2022/1426 of 5 August 2022, cit.,* Annex II, para. 11 "Operating manual". On the topic of information defect, see Piergallini, (2004), p. 47-48.

[76] The duty of solidarity may have a constitutional base, as it is in the Italian legal system, where the duty of social solidarity is stated in Article 2 of the Constitution.

[77] In favour of specific training for the use of self-driving cars Barré (2022); differently Leiman (2021), p. 262, concerning the case in which the duty to inform results in "extra training" for the user of the intelligent product as compared to the user of a product without such features. In particular, the consideration is developed concerning drivers of self-driving cars, to argue that requiring them to undergo additional training beyond that required to obtain a driving license would be unreasonable since such a requirement is not imposed by law; moreover, it could lead to the consequence that a higher standard of diligence would be required of those who decide to drive cars with automated systems, which are known to be safer, and not of those who drive traditional vehicles that expose members of the public to greater risk.

[78] Piergallini (2020), p. 1757-58.

lance and the *duty of intervention*. The latter both represent fundamental requisites of the user's virtuous example and act as a mechanism for safeguarding legal interests. Should any violation be ascertained, it would result in the catalysing of liability on the person of the user[79].

Beginning with the examination of the *duty of vigilance*, it should be noted that human control seems to be unavoidable at present and the European vision of AI is also oriented in this direction. Such an approach can only be endorsed[80], also in light of the fact that "weak AI systems" have limited computational cognition, i.e. they move well within a given perimeter of knowledge but do not go beyond the thematic field of knowledge set up in the production phase[81].

Today, there are highly intelligent forms on delimited topics, nonetheless, an Artificial General Intelligence (AGI) -also called strong artificial intelligence- which is an intelligence of the same level as the human one, has not yet been created[82].

The result we draw from this in terms of the user's duties is that the duty of vigilance is made current at the time of the system's start-up and remains throughout its use.

The duty of vigilance is however susceptible to change in character, and to become a *duty to intervene*.

4.2.1. *Situations actualizing the duty to intervene.*

In order to identify the circumstances in which the user's duty of vigilance turns into a duty to intervene, we can refer to some examples to make our reflection easier.

Let us first think of a robot-doctor used to recognise only a limited number of pathologies. The diagnosis referring to one of those pathologies will certainly be more precise and will probably also arrive more quickly than a human doctor can. Nevertheless, the system will only be "intelligent" within that field (*e.g.* in the diagnosis of tumour pathologies of the eye), but once we have ruled out the occurrence of one of those pathologies, we cannot consider the patient to be healthy (the machine will not be able to recognise an infection symptomatic of a neurological pathology because that type of disease does not belong to its "knowledge").

Let us now imagine another scenario, in which diligent conduct is performed in fulfilment of a rule describing a duty of care with a very generical content. This rule could be of social source, then being the basis of generic negligence, or a written rule which makes reference to the duty of care of common experience, now for the determination of diligent conduct, now for the identification of the factual conditions for the rule's applicability[83].

Let us take as example the rule "drive carefully" taken from Article 140 of the Italian Road Traffic Act, according to which "road users must behave in such a way that they do not constitute any danger or hindrance to traffic and in such a way that road safety is in any case safeguarded"[84]. In this case, we have a written rule, nevertheless, its content is determined by reference to the social duty of care[85].

[79] Selbst (2020), p. 1345: "The human-in-the-loop aspect of the technologies that still rely on negligence law ensure that this type of wildly unexpected AI injury cannot happen, or in fact, the human would be reasonably blamed for it". The A. discusses the interesting example proposed by Lemley, Casey (2019), p. 1311 et seq. The case is that of a drone trained to reach the centre of a circle. The drone, after the first few attempts during which it received positive reinforcement for its success in the assigned task, began to behave differently: when it got close to the edge of the circle, it would suddenly move away from it. The trainers then switched him off and placed him back in the centre of the circle to start the experiment again. After various investigations, the programmers discovered that the drone had realised that if it moved away from the edge, it would somehow find itself 'teleported' to the centre of the circle, thus achieving its goal. Selbst (2020), p. 1345, then observes that if the operation of the drone had been supervised by a human being, as soon as it started to move away from the circle, the human being could have taken control of it again and prevented it from moving away from the centre.

[80] But it is worth noting the observations that the American doctrine makes on this point when it emphasises how in certain situations the need for human intervention could create additional risks over and above those that would result from a fully automated operating process that does not require human oversight at all. The example is related to the presence of the vehicle's automatic steering system, whereby it is observed that once the driver perceives the concrete possibility of an imminent collision, he might decide to take back control of the steering but intervene, due to the panic situation, in a way that aggravates the situation. It is concluded, then, that *"At some point, then, removing the human entirely from active driving may be safer than managing the "mushy middle" of shared human-machine operation"*. Smith (2017), p. 49.

[81] And this is also true in the case of self-learning machines: the algorithm will increase its knowledge but always in order to accomplish the task it has been given. If this consists of the recognition of tumour diseases of the eye, the expansion of knowledge cannot occur towards neuronal diseases for which the machine was not designed to recognise.

[82] See Searle (1990), pp. 26-32; Selbst (2020), p. 1344; Calo (2017), p. 432.

[83] See Giunta (1999), p. 92, indications on art. 140 of the Italian road traffic Act.

[84] The same principle is stated in art. R. 412-6 of the French Traffic Road Act, see www.legifrance.fr

[85] Giunta (1999), p. 92.

It is precisely its essence as a social rule, and not a legal one, that makes the rule "drive carefully" comprehensible to a human agent, who is capable of declining and framing it in its realistic scenario because he is endowed with critical capacity. Such a rule will be, for example, declined in the following *dictum*: "in the presence of a green traffic light, make sure in any case that there are no cars coming from the side before continuing to drive", or again: "in the presence of road signs whose recognisability has been hindered by stickers or graffiti, slow down to make sure you have understood their content". The reasoning capacity of the human mind makes it possible to deduct from a generic or elastic rule (drive carefully) a peculiar rule to a given context (check that there are no cars in the intersection even when you have the right of way; slow down when road signs are covered with stickers or graffiti in order to check their content).

Let us now turn to the perspective of the artificial agent. A rule such as "drive carefully" would be a "silent" rule, from which the normative "top-down" MAS alone is unable to extrapolate generalisations to be translated into rules suited to the concrete context presented. This means that the system is unable to read the factual situation of potential danger in the light of the general rule "drive carefully", and that all the possible situation of danger should be pre-codified by the manufacturer[86].

It could be argued that a NMAS based on a bottom-up technology would be able to act in compliance with the general norm "drive carefully" thanks to the observation of the environment, and learning from the experience[87]. However, two observations are urged in dealing with normative "bottom-up" MAS. The first is that learning from experience does not necessarily imply the comprehension of the rule "drive carefully", since there will always be new empiric conditions in which that rule will still be "silent" for the AI system, which must be first experienced by it in order to be detected as dangerous situations[88]. The second implication is that the legal system cannot be open to the idea of authorising AI agents that, when acting in real environments, could cause harm to legal goods in order to learn from the experience. Such a case must be part of the unpermitted risk and no more doubts should arise.

In conclusion, it seems that the algorithm cognition will remain limited in the presence of norms with general content, as is the case of the norm "drive carefully".

Also, the first example, that of the robot-doctor, reveals limited artificial knowledge. Hence, we can conclude that *limited algorithmic cognition* emerges either when the algorithm is confronted with a concrete phenomenology that is outside its field of knowledge (this is the case of eye disease and the newly discovered virus), or when a general rule of diligence intervenes to avoid exposing the legal interest to danger and the consequent offence (these are the cases in which generic negligence would normally come into play, or specific negligence in which the diligent behaviour is described by a generical rule, see Art. 140 of the Road Traffic Act), which while understandable for a human driver, will be much more difficult to "understand" for an artificial driver.

Such phenomena of *limited algorithmic cognition* fall within those type-situations that, in dealing with the manufacturer's position, we have classified as hypotheses in which the system has not produced the expected result, but could not have done otherwise (case *sub c)*). This is a situation type in which, according to the proposed reasoning, it does not seem possible to contest the manufacturer's liability, since his action would be within the area of the acceptable risk (but see the clarifications developed *supra*). We may now add that, within this area of acceptable risk, the user's liability may be found, provided that the circumstances of the concrete case are such as to transform the user's duty of vigilance over the AI system into a duty to intervene which, if disregarded, lays the foundations for his negligence. Under what conditions does the user's duty of vigilance exist is the question we shall now address.

Let us begin by examining situations that are outside the knowledge of the algorithmic, taking the example of medical conditions not recognised by the robotic doctor. In this situation, the duty to intervene is actualized by the limited computational knowledge of the algo-

[86] The example proposed here referred to the give way rule in case of green light could be considered just a textbook case since it is reasonable to argue that the producer will encode into the system the rule "check that there are no cars in the intersection even when you have the right of way"; notwithstanding, a similar conclusion is more difficult for all the other dangerous situations arising from the act of driving.

[87] In such a case, after having realised that the give-way rule is not always respected by the other agents (more likely human agents rather than artificial ones), the driverless car will comply with the right of way rule only when assured that no cars are coming from the side before continuing to drive

[88] This concept is explained by Faggin (2022), pp. 140-41, when he compares the learning activities of the child and that of the algorithm, concluding that human knowledge is characterised by the intuitive aspect of understanding.

rithm, which is designed to respond to specific tasks (recognition of eye tumours)[89].

The user's duty to intervene is then a direct consequence of the limited field of knowledge of the algorithm, which operates, and even better than man, only within a restricted cognitive environment[90]. These "intelligent" machines are designed to react to situations foreseen by those who create them, but not to those that are beyond the programmer's cognition[91].

Let us now move on to the other situation-type of limited algorithmic knowledge, the one in which the elastic rule of diligence presents content that is not comprehensible, *i.e.* characterised by reference to a rule from a social source that the system is unable to assimilate. Once again, we note that these limits characterise the AI alone and not just the human agent, who is a conscious part of that society which elaborates unwritten social rules through generalisations and which constitute the source of generic negligence and supplement the generical rules from which specific negligence is derived. Think, for example, of the rule "verify that there are no other cars at the crossroad even when you have priority" which, although not codified, can in any case be presumed by the driver from the more general rule "drive carefully". Therefore, the law's expectation of the user's duty of diligence does not stop at compliance with the rules on road circulation regulated by traffic lights or road signs to give way but also extends to compliance with the more general principle "drive carefully" of traffic regulations, which, in the specific case, requires verifying that there are ideal conditions for continuing to travel despite having road priority.

Peculiar is the hypothesis of traffic signs covered with stickers or graffiti. A group of researchers demonstrated that such disturbing elements are able to mislead the algorithm, which interprets a "stop" sign covered with stickers as a speed limit sign of 45 km/h[92]. Here too, then, the "drive carefully" rule fails for the algorithm, which does not perceive the sticker, errs in its interpretation of the road sign and fails to realise its content in the concrete situation.

There are other circumstances which implement the user's duty to intervene. That is the case for context in which it would be more appropriate for the protection of legal goods not to comply with the duty of care[93]. This happens especially when the duty of care is not generic, rather it has a rigid structure, and the damage cannot be neutralised except when the duty of care is violated. The function of prevention is then reversed. The rigid structure of the duty of care is transformed from safe protection into an occasion for offending the legal good[94]. In these situations, it could be problematic for AI systems to recognise that the concrete situation requires non-compliance with the duty of care[95].

Such a case is taken into consideration in the recommendations of the European Commission Expert Group established to advise on specific ethical issues raised by driverless mobility for road transport[96]. Recommendation 4 (Consider revision of traffic rules to promote safety of CAVs and investigate exceptions to comply with existing traffic rules by CAVs [*i.e.* Connected and Automated Vehicles]) provides that "Traffic rules are a means to road safety, not an end in themselves. Accordingly, the pursuit of greater road safety may sometimes require non-compliance with traffic rules", therefore "Policymakers and researchers should

[89] Another example from which the difference between human intelligence and artificial "intelligence" emerges, reported by Chagal-Feferkorn (2018), p. 137-138, is that of a woman who goes to the emergency room accompanied by her husband for injuries she claims to have sustained when she crashed into the door. The doctor, while the woman describes the incident, notices an introverted attitude of the woman, who avoids meeting her husband's eyes; he also notes a lack of empathy in the man and this makes him suspect that it is a case of domestic violence for which the intervention of a social worker is necessary, as indicated by the protocols of the health authority.
If the same case were presented to a doctor-robot, endowed with innumerable computational capabilities, even with the ability to deduce psychological states from the tone of voice, he might not process the hypothesis of possible domestic violence, thus not activating the appropriate procedure for treating the case.

[90] See Floridi (2017a), p. 155. According to the A., the most efficient AI systems are those that operate within an environment that is conformed around their limits. The environment must be adapted to the robot to make sure it can operate in it successfully; he notes that "the real difficulty for the AI system is to deal with the unpredictability of the world out there [...]. This is known as the frame problem, which relates to how a context-situated agent can represent to itself a changing environment and interact with it over time in an efficient manner", to conclude that (p. 163) AI systems "are not getting smarter while making us dumber. Instead, it is the world that is becoming an infosphere increasingly suited to [their] limited capabilities".

[91] Piergallini (2020), p. 1759.

[92] Eykholt *et al.* (2018).

[93] Such a situation is frequent in automated vehicle contexts (not just cars, but also airplanes) as highlighted by Ruffolo (2020), pp. 161-162.

[94] Marinucci (1965), p. 248.

[95] As stressed by Ungern-Sternberg (2018), ("[autonomous cars] will unconditionally obey all legal norms duly reflected in the driving algorithms. Unlike human drivers (…) autonomous cars can be programmed not to violate traffic law"), p. 257.

[96] *Horizon 2020 Commission Expert Group to advise on specific ethical issues raised by driverless mobility (E03659). Ethics of Connected and Automated Vehicles: recommendations on road safety, privacy, fairness, explainability, and responsibility. 2020. Publication Office of the European Union: Luxembourg*, p. 29.

use data provided by manufacturers and deployers to identify contexts in which it would be more appropriate to (a) change a traffic rule so that CAVs can act safely without engaging in non-compliance, (b) have the CAV handover control so that a human can make the decision to not comply with a traffic rule, or (c) allow the CAV to not comply with a traffic rule if it can explain why it made this decision and leave it to the justice system to decide whether this non-compliance was justified by the pursuit of greater safety".

Hence, there are three solutions devised for the exception to comply with traffic rules. Among them, the latter is presented with caution by the Expert Group, which indeed highlights in the same Recommendation 4 that "Researchers should study the extent to which it is reasonable to expect that an intelligent non-human system is able to engage in the complex process of evaluation of the interpretation of a legal, ethical or societal norm and its balancing with another norm, value or principle. Researchers should also test the *ex-post* explainability of these decisions".

The second solution devises the human-in-the-loop situation, in which the decision not to comply with the traffic rule is referred to the driver. The human agent therefore, might be able to realise that the concrete situation requires the breach of the duty of care and would be required to resume control of the system, adapting his/her behaviour to a more general rule of conduct thus allowing peculiarities of the concrete situation to prevail over the rigidly structured rule[97]. It means that the focus should be on the manufacturers[98] and, ultimately, on the users, whose involvement may be necessary due to uncertainty in the interpretation of the concrete situation.

The time has therefore come to add another column to the table presented above, in which the user's position is taken into account, indicating the content of his duty of care. It will be noted that the duties of information, vigilance and intervention are always present but what is being highlighted is that these duties are latent until factual circumstances capable of activating them arise. Whereas it is the use of the system that activates the duties of information and vigilance, it is the arising of peculiar situations, such as those we have attempted to identify in the previous examples, that activates the duty to intervene.

Case a)	Diligent producer	Diligent Algorithm	Diligent User (duty of information, vigilance and duty to intervene)
Case b)	Negligent Producer	Negligent Algorithm	User's duty of vigilance and of intervene
Case c)	Diligent Producer (acceptable risk)	Negligent Algorithm (unattainable diligence)	User's duty of vigilance and of intervene

[97] This aspect, analysed here in the context of criminal guilt and the duty of care, actually ties in with the problem of the ethicality of artificial intelligence systems which is much debated, and concerns dilemmas, i.e. critical situations in which, at a given point in time, the machine will inevitably cause harm to a group of individuals. *See* Anderson, Anderson (2011); Wallach, Allen (2009); Tafani (2020), pp. 83 et seq.. The issue is made very concrete through the well-known trolley problem, which if originally posed as a dilemma for the man driving the tram, with the advent of AI has been presented as a dilemma for the vehicle itself, called upon to make decisions in emergency situations. The example was first presented by the English philosopher Philippa Ruth Foot, in her essay on abortion entitled "The Problem of Abortion and the Doctrine of the Double Effect", published in the Oxford Review, V, 1967, pp. 5-15, and then taken up and re-proposed under the name of the 'trolley problem' by Thomson (1976). In essence, one imagines the driver of a tram out of control, who has the only possibility of diverting the tram onto a track other than the one it is running on, but here is the dilemma: five men are working on one track and only one man on the other, and whoever is on the track the tram will enter is doomed to be killed. The question then is: kill five men who are on the track that the tram is already on, or kill one man who is on the track to which the tram may be diverted?
Dilemmas and crash avoidance are taken into account in recommendation 6 of the Ethics of Connected and Automated Vehicles report.

[98] *See Discussion of Recommendation 6* of the Expert Group report, *cit.*, p. 33, providing that "Rather than defining the desired outcome of every possible dilemma, it considers that the behaviour of a CAV in a dilemma situation is by default acceptable if the CAV has, during the full sequence that led to the crash, complied with all the major ethical and legal principles stated in this report, with the principles of risk management arising from Recommendation 5 and if there were no reasonable and practicable preceding actions that would have prevented the emergence of the dilemma. This may be necessary to give manufacturers and deployers of CAVs the confidence to deploy their systems, with reduced speed and preventative manoeuvres always being the best solution to decrease safety risks".

4.3. *When will the user be liable? Insufficiency of the breach of the duty of care.*

At this point an objection might be formulated: whereas the area of manufacturer liability will be reduced through the theory of the acceptable risk, the same will not occur for the user, who bears the burden of a new science unable to produce perfect machines capable of avoiding all kinds of harms. Notwithstanding, it must be noticed that the user's liability could be excluded when in specific circumstances an abnormal or extraordinary situation occurs. These abnormal situations may meet the requirements of *unforeseeable circumstances* or *force majeure* that, in the Italian criminal system, exclude sanction according to art. 45 of the Italian crim. Code.

There may also be extraordinary circumstances that, while not presenting the characteristics of unforeseeable circumstances or *force majeure*, are such as to render the user's compliance with the duty of care unattainable[99]. In the latter case, culpability, thus the subjective dimension of negligence, will be excluded, even if a duty of care is breached.

Indeed, when ascertaining culpability, it is requested to take into account the cognitive and physical limitations of the user. It follows that "this transfer of responsibility should only occur if the human operator has sufficient time and information to make responsible control decisions and in no circumstance should the human operator be assigned a task for which humans are unsuited or for which they have not been sufficiently trained"[100].

On this aspect, it should be pointed out that some researches carried out in the field of human-computer interactions (HCI) [101], which take into account the psycho-physical limits of users, have shown that the driverless car greatly reduces the capabilities of the average user to retake the control of the vehicle[102] (an aspect that would therefore already be relevant in the objective dimension of guilt for the reasonable user's parameter). In particular, it is customary to refer to the so-called "handover problem" [103] in level 3 of assisted driving vehicles, in which the user is required to regain control of the car when the automatic system encounters criticalities in the course of its operation. Well, such research shows that, on average, the level of attention with which users monitor the guidance of the system is not constant due to their lack of direct involvement in the act of driving, and this negatively affects their level of alertness[104]. Therefore, it is paramount to reconsider the role of the user as a fall-back mechanism during automated driving[105].

We can ask ourselves whether such scientifically proven difficulties in the level of monitoring can be taken into account for the purposes of establishing negligent culpability. The answer seems to me to be affirmative, especially in the context of criminal liability[106]: if even cognitive-psychological research shows that the average human being is not capable of always having an "extraordinary" level of attention, this must be taken into account by the judge when assessing the circumstances of the concrete case in order to ascertain the subjective dimension of negligence. To achieve such a goal, a legislative intervention seems appropriate, in order to introduce a sort of "immunity clause" that codifies the general principle of unattainability.

Furthermore, if researches will demonstrate with certainty that in some fields is unattain-

[99] Giannini (2021), p. 24, with regards to the effects of automation on the attainability of compliance with the duty of care.
[100] *Discussion of Recommendation 4* of the Expert Group report, *cit.*, p. 30.
[101] Selbst (2020), p. 1346.
[102] Aria, Olstam, Schwietering (2016), p. 764.
[103] See American Association For Justice, *Driven To Safety: Robot Cars And The Future of Liability* (2017), p. 14, (*perma.cc*): *"Research shows that humans are not well adapted to re-engaging with complex tasks, like driving a vehicle in an emergency situation, once their attention has been allowed to wander. A 2015 study by the National Highway Traffic Safety Administration (NHTSA) found that it took test subjects an average of 17 seconds to respond to a request to regain control of their vehicle. That's enough time for a car traveling at 60 miles per hour to travel a quarter of a mile"*. Reference is made to the study *Human Factors Evaluation of Level 2 and Level 3 Automated Driving Concepts, National Highway Traffic Safety Administration (NHTSA)*, August 2015, *www.nhtsa.gov*.
[104] *See* the previous footnote. A list of concerns raised by AV is presented in the work of Aria, Olstam, Schwietering (2016), which includes the degrading driving skills of human drivers in the absence of practice and his/her tendency to become involved with secondary tasks. The idea of establishing the boundaries of the negligence of the user is also supported by Beck (2016), p. 141.
[105] Heikoop, *et al.* (2019), p. 7, par. 3.3; Cappellini (2022), p. 13.
[106] Panattoni (2021), in dealing with the topic of human control, defines the user's oversight as "an overly tight obligation". More generally on the efficacy and impacts of human oversight policies in automated decision-making systems, including the criminal justice one, see Green (2022) who proposes a shift from human oversight to institutional oversight. The A. affirms that "policymakers must stop relying on human oversight as a remedy for the potential harms of algorithms" (p. 2) and that "there must be evidence suggesting that people can oversee the algorithm" (p. 14).

able even for the reasonable agent to grant a prompt intervention on the system in order to retake control of it, it could be possible to argue that the duty of care (*i.e.* the objective dimension of negligence) will be influenced. It follows that the reasonable agent parameter should take into consideration the limited capabilities of the average user.

5. Conclusion.

This work has explored the capabilities of an intelligent system, ultimately concluding that the diligent behaviour of the artificial system can never be equated with human diligence.

As I have tried to demonstrate, "artificial diligence" is limited to the field of knowledge available to the system, which will not be endowed with a critical spirit, common sense, and the ability to make an overall assessment of the dangerous situation, which altogether allow the application of duties of care, especially the generic ones, in the concrete case. The artificial system will know how to calculate the best solution in order to conform its behaviour using its set of inbuilt knowledge. However, as soon as the dangerous situation goes beyond the perimeter of its artificial knowledge, the machine will have no predetermined answer and will have to proceed making generic assumptions based on precedent examples and stored data thus increasing the risk of error.

Of course, the limits of artificial diligence are relevant for the purposes of assessing the manufacturer and user's negligent conduct, who are addressees of duties of care inferable from European legislation.

With the purpose to delineate the apportionment of responsibilities between the producer and the user, I distinguished the situations of danger connected to the non-observance of duties of care in two categories: situations in which it is possible to provide *ex ante* a description in terms of danger, and situations in which such an *ex ante* description of danger is unattainable since they are correlated to concrete dynamics that are difficult to foresee and therefore not imagined in the algorithm's training phase and offered as examples during training. In other words, these are cases that cannot be encoded in the computational language to make them knowable in advance to the machine, because they cannot be predicted.

In the presence of such circumstances (enlisted in situation types *sub c)* as seen in the previous chart) the system will proceed with generalisations that may contain a margin of error that could not have been avoided by the manufacturer even if using different training, due to the unpredictability of the concrete dynamics.

It follows that the manufacturer cannot be considered negligent for a generalisation that is unsuitable for the concrete case if – and only if – an entirely unforeseeable risk is involved, originating from the combination of a series of factors occurring in the concrete case, which make the entire dynamic leading to the causation of the event absolutely peculiar and unique.

However, situations that cannot be codified in computer language *ex ante* by the manufacturer may, at least in some cases, be recognisable and avoidable by the user, who is called upon to monitor the correct operation of the system and intervene where necessary. At this point, the intervention of the human agent – whose semantic reading capacity enables him/her to understand the dangerous situation in its concreteness – will be necessary and he/she will adopt a conduct that is the most diligent not in abstract, but in concrete terms.

Nevertheless, it must be taken into account that it is not always attainable compliance to the duty of care from the user. In the presence of specific circumstances, his/her liability could be excluded for lack of culpability of the subjective dimension of negligence. In other cases, unforeseeable circumstances or *force majeure* can occur, which will make the fact criminally irrelevant. Moreover, even the objective dimension of negligence can be excluded when scientific evidences demonstrate how not even a reasonable man could have complied with the duty of care.

I have tried to show how a criminal liability can be founded with the traditional categories of criminal law, without considering the machine as the author of a crime. Nevertheless, the liability regime seems to attribute much more responsibility to the user, rather than to the manufacturer. Therefore, on one hand, a legislative intervention seems necessary in order to delimitate the user's liability only to cases in which an intervention is attainable. On the other hand, the law should enhance the role of the manufacturer whose responsibilities can be established according to other branches of the legal system different from the criminal one.

It follows that the response of the law in front of the new risks of AI shall be considered as a whole[107]. Probably in a context of such a high degree of uncertainty, the jurist, but also the community, will have to accept a marginal role for criminal law, so that technological progress will be a confirmation of the achievements of modern legal civilisation and not a context favourable to the application of security measures that entrust a pivotal role to the criminal sanction. Criminal law is and must remain the last remedy, the *ultima ratio*.

Bibliography

Abbot, Ryan (2020): *The Reasonable Robot. Artificial Intelligence and the Law*, (Cambridge, Cambridge University Press)

Allen, Colin, Wallach, Wendell, (2009): *Moral Machines: Teaching Robots Right from Wrong* (Oxford, Oxford University Press)

Amato, Salvatore (2020): *Biodiritto 4.0. Intelligenza artificiale e nuove tecnologie* (Torino, Giappichelli)

Anderson, Michael, Anderson, Susan L. (eds.) (2011): *Machine Ethics* (Cambridge, Cambridge University Press)

Andrighetto, Giulia, Governatori, Guido, Noriega, Pablo, Van der Torre, Leendert W.N. (2013) (eds.): *Normative Multi-Agent Systems* (Schloss Dagstuhl, Leibniz-Zentrum für Informatik, GmbH, Dagstuhl Publishing, Saarbrücken/Wadern, Germany)

Aria, Erfan, Olstam, Johan, Schwietering, Christoph (2016), "*Investigation of Automated Vehicle Effects on Driver's Behavior and Traffic Performance", Transportation Research Procedia*, Volume 15, pp. 761–770

Balkin, Jack M. (2017): "*The Three Laws of Robotics in the Age of Big Data", 78 Ohio St. L.J. (Sep. 10*)

Barré, Jessy (2022) : « *Quelle formation pour les conducteurs de véhicules autonomes ? », Recherche, Transport, Sécurité,* February

Bartoli, Roberto (2021): "*voce Fonti della colpa", Enc. dir., I Tematici, Il reato colposo,* Milano, Giuffrè, pp. 519 et seq.

Basile, Fabio (2019): *"Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine", in Dir. pen. e uomo,* 29 September 2019

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

Beck, Susanne (2017): *"Google Cars, Software Agents, Autonomous Weapons Systems – New Challanges for Criminal Law*", in Hilgendorf, E., Seidel, U. (eds.), *Robotics, Autonomics and the Law*, Nomos, Baden-Baden, 2017, 227-252

Bertolini, Andrea, Episcopo, Francesca (2021): "*The Expert Group's Report on Liability for Artificial Intelligence and Other Emerging Digital Technologies: A critical assessment*", in *European Journal of Risk Regulation, 12*(3), 644-65, doi:10.1017/err.2021.30

Bifulco, Raffaele (2018): "*Intelligenza artificiale, internet e ordine spontaneo*", in Pizzetti, F.(ed.), *Intelligenza artificiale, protezione dei dati personali e regolazione*, Torino, pp. *383-400*

Borsari, Riccardo (2019): *"Intelligenza Artificiale e responsabilità penale: prime considerazioni"*, in *MediaLaws*, Novembre 20, 3

[107] Ruffolo (2020), p. 155 stresses how the law should not be such to undermine technological development, considering the benefits for society from AI systems (particularly, the A. refers to the estimated reduction of car accidents thanks to the introduction of automated vehicles).

Bradford, Anu (2020): *The Brussels Effect: How the European Union Rules the World*, (New York, Oxford University Press)

Bricola, Franco (1978): "*Responsabilità penale per il tipo e il modo di produzione*", in *La responsabilità dell'impresa per i danni all'ambiente e ai consumatori*, Milano, Giuffrè, pp. 75-90

Calo, Ryan (2017): "*Artificial Intelligence Policy: A Primer and Roadmap*", 51 *U.C. Davis l. Rev*. 399-435

Camardi, Carmela (ed.) (2022): *La via europea per l'intelligenza artificiale,* (Milano, Wolters Kluwer)

Canestrari, Stefano (2013): "*La colpa*", in *Trattato di diritto penale, Parte generale, vol. II, Il Reato*, a cura di Cadoppi, Alberto, Canestrari, Stefano, Manna, Adelmo, Papa, Michele, (Torino, UTET)

Cappellini, Alberto (2019): "*Profili penalistici delle self-driving cars*", in *Dir pen. cont.-Riv. trim.*, 2, 325-353

Cappellini, Alberto (2022): "*Reati colposi e tecnologie dell'intelligenza artificiale*", *Arch. Pen.,* 3, 2022, 1-19

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

Chagal-Feferkorn, Karni (2018): "*The Reasonable Algorithm*", in *University of Illinois Journal of Law (UIJL), Technology & Policy, no. 1,* pp. 111-148

Colonna, Kyle (2012): "*Autonomous Cars and Tort Liability*", *4 Case W. Res. J. L. Tech & Internet,* 81-130

Consulich, Federico (2021): "*voce Rischio consentito*", *Enc. dir., I Tematici, Il reato colposo,* (Milano, Giuffrè)

Consulich, Federico (2022): "Flash offenders. *Le prospettive di* accountability *penale nel contrasto alle intelligenze artificiali devianti*", in *Riv. it. dir. proc. pen.,* 3, 1015-1055

Conte, Rosaria, Castelfranchi, Cristiano (1993): *"Norms as mental objects. From normative beliefs to normative goals"*, in Castelfranchi, C., Müller, JP. (eds) *From Reaction to Cognition. MAAMAW 1993, Lecture Notes in Computer Science, vol 957,* Berlin, Heidelberg, Springer, online version: *aaai.org*

Cotta, Sergio (1968): *La sfida tecnologica* (Il Mulino, Bologna)

Crane, Daniel A., Logue Kyle D., Pilz, Bryce (2017): "*A survey of the legal issues arising from the deployment of autonomous and connected vehicles*", in *Michigan Telecom. And Tech. L. Rew.,* p. 191 et seq.

De Francesco, Vittorio (1977-78): "*Sulla misura soggettiva della colpa*", in *Studi Urbinati,* 1977-78, pp. 275 et seq.

Di Giovine, Ombretta (2003): *Il contributo della vittima nel delitto colposo* (Torino, Giappichelli)

Donini, Massimo (1989): "*Lettura sistematica delle teorie dell'imputazione oggettiva dell'evento*, P.te I", in *Riv. it. dir. proc. pen.*, 588-638

Donini, Massimo (2010): "(voce) *Imputazione oggettiva dell'evento (diritto penale)", in Enciclopedia del diritto, Annali*, vol. III, Milano, Giuffrè, p. 646 et seq.

Engisch, Karl (1995): *Untersuchungen über Vorsatz und Fahrlässigkeit im Strafrecht*, (Aalen, Scientia (reprint 1930 ed.))

Exner, Franz (1910): *Das Wesen der Fahrlässigkeit: eine Strafrechtliche Untersuchung*, (Leipzig, F. Deuticke)

Eykholt, Kevin, *et al.* (2018): "*Robust Physical-World Attacks on Deep Learning Visual Classification*", *arXiv.it*

Faggin, Federico (2022): *Irriducibile. La coscienza, la vita, i computer e la nostra natura* (Milano, Mondadori)

Floridi, Luciano (2017a): *La quarta rivoluzione. Come l'infosfera sta trasformando il modo* (Milano Raffaello Cortina Editore)

Floridi, Luciano (2017b): "*Digital's Cleaving Power and Its Consequences", Philosophy & Technology*, May

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

Gallo, Marcello (1960): "*voce Colpa penale (dir. vig.)", Enciclopedia del diritto*, vol. VII, 1960, Milano, Giuffrè

Giannini, Alice (2021): *"Intelligenza artificiale, human oversight e responsabilità penale: prove d'impatto a livello europeo", Criminalia*, 21.11.2021, discrimen.it

Gigerenzer, Gerd (2022): *Perché l'intelligenza umana batte ancora gli algoritmi* (Milano, Raffaello Cortina editore)

Giunta, Fausto (1999): *"La normatività della colpa penale. Lineamenti di una teorica",* in *Riv. it. dir. proc. pen.* p. 86-115

Giunta, Fausto (2019), "Culpa, culpae", in *disCrimen*, 4.06.2019

Gless, Sabine, Silverman, Emily, Weigend, Thomas (2016): *"If Robots Cause Harm, Who Is to Blame? Self-Driving Cars and Criminal Liability",* in *New Criminal Law Review*, 3, 19, 412-436

Green, Ben (2022): "*The flaws of policies requiring human oversight of government algorithms", Computer Law&Security rew.,* 45, 1-22

Heikoop, Daniel, Hagenzieker, Marjan, Mecacci, Giulio, Calvert, Simeon, Santoni De Sio, Filippo, van Arem, Bart (2019): *"Human behaviour with automated driving systems: A quantitative framework for meaningful human control",* in *Theoretical Issues in Ergonomics Science (TIES)*, March.

Hilgendorf, Eric (1993): *Strafrechtliche Produzentenhaftung in der "Risikogesellschaft"* (Berlin, Duncker und Humblot)

Hippel, Robert (1908): *„Vorsatz, Fahrlässigkeit, Irrtum"*, in Birkemeyer, Karl v. (ed.), *Vergleichende Darstellung des Deutschen und Ausländischen Strafrechts, Allgemeiner Teil*, III, Berlin, Liebmann, 374-599

Hollander, Christopher. D., Wu, Annie S. (2001): *"The Current State of Normative Agent-Based Systems", Journal of Artificial Societies and Social Simulation,* 14 (2), 6, 2011

Hubbard, F. Patrick (2014), " *"Sophisticated Robots": Balancing Liability, Regulation, and Innovation", 66 Fla. L. Rev.*, 1803 et seq.

Jescheck, Hans-Heinrich, Weigend, Thomas (1996): *Lehrbuch des Strafrechts. Allgemeiner Teil. 5 Auflage* (Berlin, Duncker-Humbolt)

King, Thomas, Aggarwal, Nikita, Taddeo, Mariarosaria, Floridi, Luciano (2021): *„Artificial intelligence crime: An interdisciplinary analysis of foreseeable threats and solutions", The 2020 Yearbook of the Digital Ethics Lab*, October, pp.195-227, *researchgate.net*

La Vattiata, Federico, Carmelo (2023): *„AI Systems Involved in Harmful Events: Liable Persons or Mere Instruments? An Interdisciplinary and Comparative Analysis"*, in *BioLaw Journal – Rivista di BioDiritto, n. 1,* 485-499

Lagioia, Francesca, Sartor, Giovanni (2020): "*AI Systems Under Criminal Law: a Legal Analysis and a Regulatory Perspective", Philosophy & Technology*, 2020, 33, 433–465

Leiman, Tania (2021): *"Law and tech collide: foreseeability, reasonableness and advanced driver assistance systems", Policy and Society*, 40:2, 250-271

Lemley, Mark A., Casey, Bryan (2019): "*Remedies for Robots", University of Chicago Law Review*, Vol. 86: Iss. 5, Article 3, p. 1311 et seq.

Lohmann, Melinda Florina (2016): "*Liability Issues Concerning Self-Driving Vehicles",* in *European Journal of Risk Regulation,* 7, 335 et seq.

Mantovani, Ferrando (1988): *"voce Colpa, in Dig. Disc. Pen.",* vol. II, Padova, Cedam

Marinucci, Giorgio (1965): *La colpa per inosservanza di leggi* (Milano, Giuffrè)

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

Panattoni, Beatrice (2021): "*AI and Criminal Law: The Myth of 'Control' in a Data-Driven Society*", in (eds.) Vermeulen, Peršak, Recchia, *Artificial Intelligence, Big Data and Automated Decision-Making in Criminal Justice*, Antwerpen, Maklu, *Rev. int. droit pén., vol.* 92, 125-142

Picotti, Lorenzo (2021): "*Profili di responsabilità penale per la circolazione di veicoli a guida autonoma",* in *Studi in onore di Antonio Fiorella*, Vol. I (Roma Tre Press, Roma), 813-837

Piergallini, Carlo (2004): *Danno da prodotto e responsabilità penale* (Milano, Giuffrè)

Piergallini, Carlo (2020): "*Intelligenza artificiale: da 'mezzo' ad 'autore' del reato?",* in *Riv. it. dir. proc. pen.*, 4, p. 1745-1774

Preuss, Wilhelm (1974): *Untersuchungen zum erlaubter Risiko im Strafrecht* (Berlin, Duncker und Humblot)

Prittwitz, Cornelius (1993): *Strafrecht und Risiko. Untersuchungen zur Krise von Strafrecht und Kriminalpolitik in der Risikogesellschaft* (Frankfurt a.M., Vittorio Klostermann)

Rehberg, Jürg (1962): *Zur Lehre vom „erlaubten Risiko"* (Zürich, Schulthess)

Roeder, Hermann (1969): *Die Einhaltung des sozialadäquaten Risikos* (Berlin, Duncker und Humblot)

Ruffolo, Ugo (2020): "*Intelligenza artificiale ed* automotive: *la responsabilità da veicoli* self-driving *e* driverless", in Ruffolo, Ugo (eds.) "*Intelligenza artificiale. Il diritto, i diritti, l'etica*", Milano, Giuffrè, 153-178

Ruffolo, Ugo, Al Mureden, Enrico (2019): "*Intelligenza artificiale e diritto. Autonomous vehicles e responsabilità nel nostro sistema e in quello statunitense",* in *Giur. It.*, 2019, 7, p. 1704 et seq.

Ruth Foot, Philippa (1967): *"The Problem of Abortion and the Doctrine of the Double Effect"*, *Oxford Review*, V, pp. 5-15

Salvadori, Ivan (2021): "*Agenti artificiali, opacità tecnologica e distribuzione della responsabilità penale*", in *Riv. it. dir. proc. pen. (RIDPP),* 1, 2021, 83- 118

Savarimuthu, Bastin Tony Roy, Cranefield, Stephen (2009): "*A categorization of simulation works on norms", Dagstuhl Seminar Proceedings 09121, Normative Multi-Agent Systems*, *drops.dagstuhl.de/opus/volltexte/2009/1905i.*

Schroeder, Friedrich-Christian (2003): *„§16 Irrtum über Tatumstände"*, in *Strafgesetzbuch. Leipziger Kommentar, Großkommentar, 11, neu bearbeitete Auflage*, hrsg. von Jähnke, Burkhard, Laufhütte, Heinrich Wilhelm, Odersky, Walter, (Berlin, De Gruyter Recht), Bd. 1, (10)-(104)

Searle, John R. (1980): "*Minds, brains, and programs*", in *Behavioral and Brain Sciences,* 3 (3), 417-457

Searle, John R. (1990): "*Is the Brain's Mind a Computer Program?*", in *Scientific American*, Jan., p. 26-32

Selbst, Andrew D. (2020): "*Negligence and AI's Human Users*", in *Boston University Law Review,* 100, 1315-1376

Severino, Paola (2020): *"Intelligenza artificiale e diritto penale"*, in Ruffolo, Ugo (eds.), *Intelligenza artificiale. Il diritto, i diritti, l'etica* (Milano, Giuffrè), 531-546

Shoham, Yoav, Tennenholtz, Moshe (1992), "*On the synthesis of useful social laws for artificial agent societies (preliminary report)", www.aaai.org*, 276-281

Tafani, Daniela (2020): "*Sulla moralità artificiale. Le decisioni delle macchine tra etica e diritto",* in *Riv. di filosofia*, 1, aprile, pp. 83 et seq.

Thomson, Judith J. (1976), "*Killing, letting die and the trolley problem*", in *The Monist*, LIX, pp. 204-217

Ungern-Sternberg, Antie (2018): "*Autonomous driving: regulatory challenges raised by artificial decision-making and tragic choices*", in Barfield, Woodrow, Pagallo, Ugo, (eds.), *Research handbook on the law of artificial intelligence,* Edward Elgar**,** pp. 253-277

Vagliasindi, Grazia Maria (2021) : « *Intelligence artificielle et droit pénal entre outils d'augmentation de l'efficacité préventive et punitive de l'État et risques et défis pour les droits humains et l'État de droit* », in Sedjari, A. (ed.), *Impact du numérique et de l'intelligence artificielle sur les transformations des gouvernances publiques*, 357-382

Diritto Penale Contemporaneo

RIVISTA TRIMESTRALE

REVISTA TRIMESTRAL DE DERECHO PENAL

A QUARTERLY REVIEW FOR CRIMINAL JUSTICE

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