Journal of Law, Market & Innovation

Journal of Law, Market & Innovation

3/2023

Editors: *Riccardo de Caria – Bianca Gardella Tedeschi – Gideon Parchomovsky - Cristina Poncibò*

TABLE OF CONTENTS

Bianca Gardella Tedeschi – Gideon Parchomovsky - Cristina Poncibò - Riccardo de Caria

FOREWORD TO ISSUE 3/2023
IN LOVING MEMORY OF ALBERTO M. MUSY

The present issue of the Journal of Law, Market & Innovation features a special section dedicated to the topic of torts and innovation.

It is dedicated to honouring our late colleague and friend Alberto Musy, in the tenth anniversary of his passing.

The special section covers, with a comparative perspective, a range of subjects, such as how tort doctrines (and private law doctrines, more generally) can be reconfigured to encourage innovation in our society; innovation within tort law itself; innovative approaches to regulation; and novel remedies that are capable of accommodating new products and processes.

The topic of how the rules and policies of tort law affect innovation was one of Alberto's many interests. As his colleagues and friends, we would like to offer here a brief summary of his multi-faceted academic career and personal life.

Alberto was a dear colleague and a loyal friend to all of us. He was lively and learned, loved life and everything that could in a clever way pick his infinite curiosity. He lived happily with his wife Angelica and four daughters: Isabella, Maria Luisa, Bianca and Eleonora. In March 2012, a despicable and heartless assassin shot Alberto in the courtyard of the apartment building where he was living with his beautiful family, early in the morning, after Alberto accompanied his daughters to school. Alberto entered in an irreversible coma and died in October 2013.

Alberto M. Musy graduated in law with honours from the University of Turin in 1990. He wrote a thesis in the history of law. In his thesis, he analysed divorce law in Piedmont during the Napoleonic era. After graduation, Alberto became interested comparative law. Hence, he continued his academic career as a comparativist. In 1995, he earned an LLM from the University of California, Boalt Hall School of Law (Berkeley).In 1998, he obtained his PhD in comparative law from the University of Trento. His doctoral dissertation entitled "The duty of pre-contractual information," offered an insightful analysis of disclosure duties in the pre-contractual stage. Subsequently, he spent a year as a Junior Boulton Fellow, at McGill University, Montréal.

In 2000, he was appointed Associate Professor of Comparative Law at the University of Piemonte Orientale. There, four years later, he became Full Professor, in the field of comparative law. Alberto taught a vast number of subjects pertaining to comparative law: Economic Analysis of Law, Comparative Private Law, Comparative Contract Law, International Contracts, and Comparative Corporate Law. . Alberto was a popular and deeply admired professor. Students have always shown great appreciation for his teaching because of his unparalleled ability to connect with them.

He was, in addition, an Adjunct Professor of Economic Analysis of Law in Rome, in the Scuola Superiore della Pubblica Amministrazione; Lecturer in Anglo-American Law in the University of Turin, Faculty of Law; and Lecturer in Comparative Private Law in Bocconi University. He built an international carrer a visiting professor: Professeur invité in Nantes, Faculté de droit; visiting professor in the Faculty of Law, Bar-Ilan University; visiting professor in the Cardozo School of Law in New York; and visiting professor in Fordham Law School, New York.

He wrote two important books. The first "The Duty to Inform: An Essay in Comparative Law" published in the Quaderni series of the Department of Legal Sciences of the University of Trento

(Il dovere di informazione: Saggio di diritto comparato, pubblicato nella collana dei Quaderni del Dipartimento di Scienze Giuridiche dell'Università di Trento). This topic had already been addressed by Alberto in his PhD dissertation and in an entry "Responsabilità precontrattuale" for the renowned encyclopedia *Digesto italiano* (sez. discipline privatistiche, XVII, Turin, 1997, 319 ff). The results of this research are accessible to the general public in the article "The Good Faith Principle in Contract Law and Precontractual Duty to Disclose a Comparative Analysis of New Differences in Legal Cultures," published in Global Jurist, Vol. I, Number 1, 2001 and in the Italian report included in the book "Mistake, Misrepresentation and Nondisclosure," edited by Ruth Sefton Green and John Cartwright, which appeared for the Cambridge University Press, 2004, as part of the series "The Common Core of European Private Law." In his research on pre-contractual liability, Alberto Musy combined his comparative expertise with the economic analysis of law. Alberto's scholarship has been and still remains a landmark of the literature in the field.

The second monographic work, "La comparazione giuridica nell'età della globalizzazione: riflessioni metodologiche e dati empirici sulla circolazione del modello nordamericano in Italia," published in 2004 in the Collana degli studi di diritto privato dell'Università statale di Milano, was preceded by the important article entitled "Alcune note sul pragmatismo nell'esperienza giuridica italiana", Rivista critica del diritto privato, 2002. In these studies, the circulation of the North American model in Italy is addressed as part of the broader global diffusion of U.S. patterns and models of law production, with particular emphasis on aspects related to trusts, the interactions between business choices and the regulatory environment, and the efficiency of civil justice.

In 2006, he published with Silvia Ferreri the volume on the Sale contract in the Treatise on Civil Law edited by Rodolfo Sacco. There, he had the opportunity to address and update some of the topics he held most dear. In the following years, he collaborated on several civil law works, including the Commentary to the Civil Code edited by Paolo Cendon, published for Giuffrè in 2009.

Alberto Musy's research work includes, in addition to what has already been mentioned, more than 30 scholarly articles and several contributions to collected works, published in Italy, the United Kingdom, Germany and the Netherlands.

Beside academia, Alberto Musy's research work was conducted at the Einaudi Center in Torino (https://www.centroeinaudi.it), a think tank where jurists and economists develop common knowledge. Alberto was a member of the Steering Committee and, in this capacity; he promoted crucial research on the functioning of the Italian judicial system and on the quality of law and administrative action in Italy.

At the University of Piemonte Orientale, Alberto Musy has been the promoter and organiser of the Stresa Lectures in Law. The meetings, which were held annually from 2003 to 2010, had among their guests Ian Ayres (Yale Law School), Henri Hansman (Yale Law School), Richard Espstein (University of Chicago Law School); Muriel Fabre-Magnan (La Sorbonne), Katharina Pistor (Columbia Law School), Gideon Parchomovsky (University of Pennsylvania) and James Boyd White (University of Michigan Law School), and helped create a scholarly community willing to continue working on the ideas and projects he shared with Alberto Musy.

Alberto M. Musy's intense scientific and teaching activities were made possible by his personal and human qualities. Alberto was first and foremost a devoted husband and father and a family man. He was also an amazing friend. His optimistic disposition towards life, his ability to look ahead and think outside the box, his kindness of spirit, cheerfulness and sense of humour formed the basis of the enthusiasm with which he pursued his initiatives and research. Alberto was also a great lover of Italy. He loved Italy with all his heart and was willing to make great sacrifices for it. We miss him every day.

B.G.T., G.P., C.P., R.d.C.

*Daniela Tafani**

INNOVATION LETTER

PREDICTIVE OPTIMISATION SYSTEMS DO NOT WORK, AND THEY INFRINGE RIGHTS. SO WHY DO WE USE THEM?

Abstract

In an increasing number of areas, judgments and decisions that have major effects on people's lives are now being entrusted to Machine Learning systems. The employment of these predictive optimisation systems inevitably leads to unfair, harmful and absurd outcomes: flaws are not occasional and cannot be prevented by technical interventions. Predictive optimisation systems do not work and violate legally protected rights. As Machine Learning is no excuse to break the law, the question is: why do we use such systems?

JEL CLASSIFICATION: D79

SUMMARY
1 Predictive optimisation systems – 2 The "AI ethics" narrative as cultural capture – 3 Machine learning is no excuse to break the law

1 Predictive Optimisation Systems

In an increasing number of areas, judgments and decisions that have major effects on people's lives are now being entrusted to Machine Learning (ML) systems that do not work.[1] In areas such as pre-trial risk assessment, financial services, education, social services and recruitment, the employment of ML systems in assessment and decision making has led to unfair, harmful and absurd outcomes,[2] as documented in an extensive

---

* Daniela Tafani is a researcher in political philosophy at the Department of Political Science of University of Pisa.

[1] Inioluwa Deborah Raji, I Elizabeth Kumar, Aaron Horowitz and Andrew Selbst, 'The Fallacy of AI Functionality' (Conference on Fairness, Accountability, and Transparency (FAccT 2022), Seoul, Republic of Korea, June 21–24 2022) <https://doi.org/10.1145/3531146.3533158> accessed 2 April 2023.

[2] Ali Alkhatib, 'To Live in Their Utopia: Why Algorithmic Systems Create Absurd Outcomes' (Conference on Human Factors in Computing Systems (CHI '21), Yokohama, Japan, May 8-13 2021) <https://ali-alkhatib.com/papers/chi/utopia/utopia.pdf> accessed 2 April 2023.

body of literature,[3] with consequences that can rumble on for a long time, sometimes years, in the lives of victims.

Such flaws are not occasional and cannot be prevented by technical interventions.[4] On the contrary, they are part and parcel of the normal functioning of ML systems,[5] which are misused for tasks they cannot adequately perform, or which are completely impossible (unless one believes, as in ancient divination activities, that each individual's future is already written and therefore readable[6]). Using ML systems to detect character traits or predict the actions of individuals has no grounding in science. The use of the term "prediction" is itself misleading: though an ML system can predict words in sequences of text strings, this in no way implies that it can predict the future, or, more specifically, future social outcomes or the actions of particular individuals.

The idea that ML systems are capable of such predictions stems from the notion – essential to superstition and ascribed, in the twentieth century, to the world of psychosis – that all connections are meaningful, regardless of the distinction of causal relationships, that all details are meaningful and everything explains everything.[7] Like faith in the predictions of astrology,[8] faith in these algorithmic predictions vanishes as soon as the modern scientific criteria of communicability and reproducibility are applied.[9] Closer examination reveals that such systems are unreliable in predicting individual events and actions, to the extent that some researchers have suggested using a lottery rather than ML systems to choose between eligible individuals when resources are scarce and it is not possible to use simple computational methods with relevant and explicit variables.[10] If gender predicts lower pay and skin colour predicts the likelihood of being stopped by the police, then in the transition from prediction to decision such social profiling becomes self-fulfilling, legitimising the biases embedded in the initial statistical description by

---

[3] Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (Broadway Books 2017); Safiya Umoja Noble, *Algorithms of oppression* (New York University Press 2018); Virginia Eubanks, *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (St. Martin's Press 2018); White House Office of Science and Technology Policy, *Making Automated Systems Work for the American People* (Blueprint for an AI Bill of Rights 2022) <https://www.whitehouse.gov/ostp/ai-bill-of-rights> accessed 2 April 2023.
[4] Meredith Broussard, *More than a Glitch. Confronting Race, Gender, and Ability Bias in Tech* (The MIT Press 2023).
[5] Louise Amoore, *Cloud Ethics. Algorithms and the Attributes of Ourselves and Others* (Duke University Press 2020) 115ff.
[6] Elena Esposito, 'The Future of Prediction: From Statistical Uncertainty to Algorithmic Forecasts' (*Artificial Communication,* 22 June 2022) <https://artificialcommunication.mitpress.mit.edu/pub/m8xpxiru> accessed 1 June 2023.
[7] Paolo Rossi, *Il tempo dei maghi. Rinascimento e modernità* (Raffaello Cortina 2006); Paolo Rossi, *The Birth of Modern Science* (Blackwell Pubishers 2001).
[8] Daniela Tafani, 'What's wrong with "AI ethics" narratives' [2022] Bollettino telematico di filosofia politica <https://commentbfp.sp.unipi.it/daniela-tafani-what-s-wrong-with-ai-ethics-narratives> accessed 2 April 2023.
[9] Sun-ha Hong, 'Predictions without futures' (2022) 61 History and Theory 369 <https://onlinelibrary.wiley.com/doi/10.1111/hith.12269> accessed 2 April 2023.
[10] Angelina Wang, Sayash Kapoor, Solon Barocas and Arvind Narayanan, 'Against Predictive Optimization: On the Legitimacy of Decision-Making Algorithms that Optimize Predictive Accuracy' (2022) <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4238015> accessed 2 April 2023.

virtue of the supposed objectivity of the algorithm. Prediction thus produces what it purports to predict.[11]

In areas such as justice, health, education or finance, decisions based on such systems are infallibly prejudicial to legally protected rights, since they operate by grouping individuals into classes determined on the basis of regularities detected in training data and are therefore automatically discriminatory. Rooted in the statistical nature of these systems, the characteristic of forgetting "margins" is structural:[12] it is not accidental and is not due to single, technically modifiable biases. A person can end up on the margins of algorithmic models of normality by virtue of characteristics that are totally irrelevant to the decisions being made.[13]

2 The "AI Ethics" narrative as cultural capture

Fearing a blanket ban, Big Tech have responded to the extensive documentation of the unfair, harmful and absurd consequences of such decisions with a discourse on ethics, financed in an obvious conflict of interest.[14] This is how AI ethics came into being, with the aim of making a merely self-regulatory regime seem plausible.[15] Lobbying thus includes a "cultural capture": by "colonising the entire space of scientific intermediation", [16] it succeeds in convincing regulators, rather than (or in addition to) capturing them through incentives, and labelling all those who express concern as retrogrades or Luddites. The nonsense of decision-making based on automated statistics is thus presented as a problem of single and isolated biases, amendable by algorithmic fairness, ie, by technical fulfilment.

The "AI ethics" narrative (or its fungible variants, such as "value alignment" or "algorithmic fairness") is therefore a commodity,[17] which researchers and universities are

---

[11] Dan McQuillan, *Resisting AI. An Anti-fascist Approach to Artificial Intelligence* (Bristol University Press 2022) 36.

[12] Abeba Birhane, Elayne Ruane, Thomas Laurent, Matthew S Brown, Johnathan Flowers, Anthony Ventresque and Christopher L Dancy, 'The Forgotten Margins of AI Ethics' (Conference on Fairness, Accountability, and Transparency (5th ACM Conference on Fairness, Accountability, and Transparency, (FAccT 2022), Seoul, Republic of Korea, June 21–24 2022) <https://doi.org/10.1145/3531146.3533157> accessed 2 April 2023.

[13] Frank Pasquale, *New Laws of Robotics. Defending Human Expertise in the Age of AI* (The Belknap Press of Harvard University Press 2020).

[14] Mohamed Abdalla and Moustafa Abdalla, 'The Grey Hoodie Project: Big Tobacco, Big Tech, and the threat on academic integrity' (Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES 2021), May 19–21, 2021, Virtual Event (ACM 2021)) <https://arxiv.org/abs/2009.13676v4> accessed 2 April 2023.

[15] Rodrigo Ochigame, 'The Invention of "Ethical AI". How Big Tech Manipulates Academia to Avoid Regulation' (*The Intercept*, 20 December 2019) <https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/> accessed 2 April 2023; Ben Wagner, 'Ethics As An Escape From Regulation. From "Ethics-Washing" To Ethics-Shopping?' in Emre Bayamlioglu , Irina Baraliuc , Liisa Albertha Wilhelmina Janssens and Mireille Hildebrandt (eds), *Being Profiled: Cogitas Ergo Sum* (Amsterdam University Press 2018) <https://www.degruyter.com/document/doi/10.1515/9789048550180-016/html> accessed 2 April 2023.

[16] Andre Saltelli, Dorothy J Dankel, Monica Di Fiore, Nina Holland and Martin Pigeon, 'Science, the endless frontier of regulatory capture' (2022) 135 Futures <https://doi.org/10.1016/j.futures.2021.102860> accessed 2 April 2023.

[17] Tao Phan, Jake Goldenfein, Declan Kuch and Monique Mann (eds), *Economies of Virtue: The Circulation of 'Ethics' in AI*, (Institute of Network Cultures 2022) <https://networkcultures.org/blog/publication/economies-of-virtue-the-circulation-of-ethics-in-ai/> accessed 2 April 2023.

interested in supplying as oil that "greases the wheels of collaboration" with large technology companies,[18] and which companies commission and buy because it is useful to them as reputational capital, generating a competitive advantage. In this discourse, universities have the role and autonomy of a megaphone; it is "the bait through which trust in technology is extracted from publics or users",[19] an advertising discourse that, because it is declaimed by others, does not even appear as such. Researchers thus become "suppliers of service in this new economy of virtue" and are induced to be "complicit with systems and actors that seek to operationalise ethics to protect their own self-interest",[20] turning ethics into a matter of procedural compliance with an "anemic set of tools" and technical standards.[21]

The function of this discourse is to protect and legitimise a surveillance business model, the core of which is to sell false promises of individual predictions based on algorithmic profiling.[22] This business model – which externalises the costs of labour, environmental impact and social harms[23] – actually exploits a "legal bubble",[24] ie, it takes place in violation of legally protected rights, betting on a subsequent legal rescue in the name of the inevitability of technological innovation.[25]

Since the framing of the discourse is determined by its function, AI ethics is peddled within the perspective of technological determinism and solutionism,[26] within the "logic of the *fait accompli*".[27] No consideration is ever given to the possibility of not building certain systems at all, or not using them for certain purposes, because "ethics discourses pre-empt questions regarding the rationale of AI development, positioning investment and implementation as inevitable and, provided ethical frameworks are adopted, laudable".[28]

---

[18] Michael Richardson, 'Military Virtues and the Limits of 'Ethics' in AI Research' in Phan, Goldenfein, Kuch, Mann (eds) (n 17).

[19] Sarah Pinker, 'Extractivist Ethics' in Phan, Goldenfein, Kuch, Mann (eds) (n 17) 39, 41.

[20] Tao Phan, Jake Goldenfein, Monique Mann and Declan Kuch, 'Economies of Virtue: The Circulation of 'Ethics' in Big Tech' (forthcoming) Science as Culture <https://ssrn.com/abstract=3956318> accessed 2 April 2023; Meredith Whittaker, 'The steep cost of capture' (2021) 28 Interactions 51 <https://interactions.acm.org/archive/view/november-december-2021/the-steep-cost-of-capture> accessed 2 April 2023.

[21] Jacob Metcalf, Emanuel Moss and Danah Boyd, 'Owning Ethics: Corporate Logics, Silicon Valley, and the Institutionalization of Ethics' (2019) 82 Social Research: An International Quarterly 449 <https://datasociety.net/wp-content/uploads/2019/09/Owning-Ethics-PDF-version-2.pdf> accessed 2 April 2023.

[22] Cory Doctorow, 'How to Destroy Surveillance Capitalism' (*OneZero*, 26 August 2020) <https://onezero.medium.com/how-to-destroy-surveillance-capitalism-8135e6744d59> accessed 2 April 2023.

[23] Kate Crawford, *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence* (Yale University Press 2021).

[24] Marco Giraudo, 'Legal Bubbles', *Encyclopedia of Law and Economics* (Springer 2022) <https://www.researchgate.net/publication/357702553> accessed 2 April 2023.

[25] Jack Stilgoe, *Who's Driving Innovation? New Technologies and the Collaborative State* (Palgrave Macmillan 2020).

[26] Evgeny Morozov, *To Save Everything, Click Here: The Folly of Technological Solutionism* (Public Affairs 2013).

[27] Catherine Tessier, 'Éthique et IA: analyse et discussion' in Olivier Boissier (ed), *CNIA 2021: Conférence Nationale en Intelligence Artificielle* (PFIA 2021) 22 <https://hal-emse.ccsd.cnrs.fr/emse-03278442> accessed 2 April 2023.

[28] Alexandra James and Andrew Whelan, ''Ethical' artificial intelligence in the welfare state: Discourse and discrepancy in Australian social services' (2022) 422 Critical Social Policy 22, 37 <https://journals.sagepub.com/doi/abs/10.1177/0261018320985463> accessed 2 April 2023.

In recent years, the work of demystifying the instrumental nature of the "AI ethics" narrative has been so effective that many have dismissed the whole moral philosophy as useless or harmful – as a toothless alternative to law or empty corporate rhetoric.[29]

3 Machine Learning is no excuse to break the law

As the "AI ethics" narrative unravels, the Stone Guest it was meant to keep out appears, and many are now arguing that there is an urgent need for drastic legislative intervention. The adoption of machine-learning systems for decision-making purposes in areas relevant to people's lives, such as the judiciary, education or social welfare, is in fact tantamount to creating, by administrative decisions, "almost human rights-free zones".[30]

The priority of individual rights specifically protected by law over a generic principle of innovation,[31] and the evidence of violations of such rights when using ML systems for activities that have a significant effect on people's lives, underpin Frank Pasquale and Gianclaudio Malgieri's proposal. High-risk artificial intelligence systems embedded in products and services should be governed by a regime of "unlawfulness by default": until proven otherwise, they should be considered illegal, and the burden of proof to the contrary should be on companies, ie, it should be up to the companies to prove, before deployment, that their systems meet "clear requirements for security, non-discrimination, accuracy, appropriateness, and correctability".[32] This would put an end to the general infringement of legally protected rights; indeed, predictive optimisation systems prevent people from accessing resources or exercising rights in ways that are in conflict with existing legal systems.

A step in this direction is the recent introduction of a "Fundamental rights impact assessment for high-risk AI systems" in the Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act)[33]. It remains to be seen whether the final draft will retain this novelty

---

[29] Elettra Bietti, 'From Ethics Washing to Ethics Bashing: A View on Tech Ethics from Within Moral Philosophy' [2021] SSNR Electronic Journal <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3914119> accessed 2 April 2023.

[30] Philip Alston, *The Digital Welfare State – Report of the Special Rapporteur on Extreme Poverty and Human Rights* (UNGA A/74/493, 11 October 2019) <https://daccess-ods.un.org/access.nsf/Get?OpenAgent&DS=A/74/493&Lang=E> accessed 2 April 2023.

[31] On the innovation principle as a mask behind which large economic actors claim the protection of their concrete interests, see Saltelli, Dankel, Di Fiore, Holland and Pigeon (n 16).

[32] Frank Pasquale and Gianclaudio Malgieri, 'From Transparency to Justification: Toward Ex Ante Accountability for AI' (2022) 8 (33) Brussels Privacy Hub Working Papers <https://brusselsprivacyhub.com/wp-content/uploads/2022/05/BPH-Working-Paper-vol8-N33.pdf> accessed 2 April 2023.

[33] Draft Compromise Amendments on the Draft Report Proposal for a regulation of the European Parliament and of the Council on harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts, (COM(2021)0206 – C9 0146/2021 – 2021/0106(COD)), May 16, 2023, <https://www.europarl.europa.eu/resources/library/media/20230516RES90302/20230516RES90302.pdf> accessed 1 June 2023.

or instead bear the marks of the intervention of the "lobbying ghost in the machine" of regulation.[34]

Other banning proposals are based on the technical analysis of the characteristics of ML systems. Predictive optimisation systems should be banned outright, where decisions have major consequences on people's lives, because they are based on false promises.[35] For the same reason, the narratives spun by those who claim, for commercial purposes, that they exist should be equated with misleading advertising. The US Federal Trade Commission reminds companies of this and makes explicit reference to predictive optimisation systems:

"Are you exaggerating what your AI product can do? Or even claiming it can do something beyond the current capability of any AI or automated technology? For example, we're not yet living in the realm of science fiction, where computers can generally make trustworthy predictions of human behaviour. Your performance claims would be deceptive if they lack scientific support or if they apply only to certain types of users or under certain conditions".[36]

In stark contrast to the position of large corporations, the Federal Trade Commission has stated that technology products are subject to existing regulations[37]. The inability, for technical reasons, to comply with legal requirements is not a reason to declare oneself exempt from the same laws and ask for new ones, but - as in any other sector faced with a finding of illegality - a reason not to market such products at all.

In areas such as justice, health, education or finance, where we are entitled to an explanation of the decisions that affect us, it should be mandatory to use systems that, unlike ML systems, are based on explicit models and interpretable variables, and "data supply chains" that are designed, generated and maintained, in each case, in a manner consistent with the system being built.[38]

Such operations would seem to be suggested by common sense, but they are not implemented, because the costs are greater than capturing huge amounts of data through surveillance mechanisms, and because transparent systems, without the magical aura of algorithmic clairvoyance, could not be sold as predicting the future. Companies therefore choose to include, among the costs to be externalised, those arising from the social harms produced by predictive optimisation systems.

---

[34] Corporate Europe Observatory, 'The lobbying ghost in the machine. Big Tech's covert defanging of Europe's AI Act', February 23, 2023, <https://corporateeurope.org/en/2023/02/lobbying-ghost-machine>.

[35] Wang, Kapoor, Barocas and Narayanan (n 10).

[36] Michael Atleson, 'Keep your AI claims in check' (*Federal Trade Commission business blog*, 27 February 2023) <https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check> accessed 1 June 2023.

[37] Alvaro M. Bedoya, 'Early Thoughts on Generative AI. Prepared Remarks of Commissioner Alvaro M. Bedoya, Federal Trade Commission Before the International Association of Privacy Professionals', April 5, 2023, <https://www.ftc.gov/news-events/news/speeches/prepared-remarks-commissioner-alvaro-m-bedoya-international-association-privacy-professionals> accessed 1 June 2023.

[38] Nello Cristianini, 'Shortcuts to Artificial Intelligence' in Marcello Pelillo and Teresa Scantamburlo (eds), *Machines We Trust. Perspectives on Dependable AI* (The MIT Press 2021) <https://philpapers.org/archive/CRISTA-3.pdf> accessed 2 April 2023; Idem, *The Shortcut. Why Intelligent Machines Do Not Think Like Us* (CRC Press 2023).

The divide is not, therefore, between respect for human rights and the principle of innovation. It is between respect for rights and the business model of the monopolies of intellectual capitalism.[39]

Moreover, a radical defence of individual rights through antitrust laws would also promote innovation because those same monopolies crush any beneficial, disruptive innovations that do not fit their business model and thus mainly promote toxic innovation that extracts or destroys value instead of producing it.[40]

In short, predictive optimisation systems do not work and violate legally protected rights.

So, I ask: why do we use them?

---

[39] Ugo Pagano, 'The Crisis of Intellectual Monopoly Capitalism' (2014) 38 Cambridge Journal of Economics 1409 <https://ssrn.com/abstract=2537972> accessed 2 April 2023; Tim Wu, *The Curse of Bigness. Antitrust in the New Gilded Age* (Columbia Global Reports 2018); Marco Giraudo, 'On legal bubbles: some thoughts on legal shockwaves at the core of the digital economy' (2022) 18 Journal of Institutional Economics 587 <https://doi.org/10.1017/S1744137421000473> accessed 2 April 2023; Shoshana Zuboff, 'Surveillance Capitalism or Democracy? The Death Match of Institutional Orders and the Politics of Knowledge in Our Information Civilization' (2022) 3 Organization Theory <https://doi.org/10.1177/26317877221129290> accessed 2 April 2023.

[40] Ariel Ezrachi and Maurice E Stucke, *How Big-Tech Barons Smash Innovation - And How To Strike Back* (HarperCollins 2022); Thomas Hoppner, 'From creative destruction to destruction of the creatives: innovation in walled-off ecosystems' (2022) 1 Journal of Law, Market & Innovation 10 <https://www.ojs.unito.it/index.php/JLMI/article/view/6951> accessed 2 April 2023.

Francesco Parisi *

LIABILITY OR NO LIABILITY?

PROMOTING SAFETY BY SHIFTING ACCIDENT LOSSES ONTO THIRD PARTIES

Abstract

In a recent article, Guerra et al.[1] considered the problem of liability for accidents caused by the activity of robots, proposing a novel liability regime, which they referred to as 'manufacturer's residual liability.' Under this regime, injurers (robot operators) and victims are liable for accidents due to their negligence (hence, they are incentivised to act diligently), and third-party robot manufacturers bear all remaining accident losses, even when the accident is not caused by a defect or malfunction of the robot. In this article, I explore the possibility of extending this framework of liability to other tort scenarios. I refer to this broader application of the rule as 'third-party residual liability.' This rule incentivises third parties to make optimal safety investments in environments that are under their control, beyond the incentives obtainable under negligence or products liability law. Third-party residual liability rules will lead to a reduction in the price of safer goods and services offered by the third party, driving unsafe technologies out of the market. Further, thanks to the percolation effect of third-party residual liability, injurers and victims would also be incentivised to limit their activity levels in risky environments. The ideal application of this rule is to accident situations where the risk is affected not only by the behaviour of injurers and victims, but also by external factors that are controlled by a third party.

JEL CLASSIFICATION: K13, K32

SUMMARY

[1] Alice Guerra and others, 'Liability for Robots II: An Economic Analysis' (2022) 18 Journal of Institutional Economics 553.

1 Introduction

In real life tort situations, several factors contribute to the probability of an accident occurring. From a law and economics perspective, traditional liability rules implicitly focus on two main factors: (*i*) the behaviour of prospective injurers, and (*ii*) the behaviour of prospective victims (hereinafter, I shall refer to prospective injurers and victims conjunctly as the 'parties'). The relevant elements of the parties' behaviour are their 'levels of care' (ie, the effort undertaken to carry out their activities in a careful manner) and their 'activity levels' (ie, how much activity the parties carry out, exposing themselves and others to the risk of an accident). This traditional framing of a tort problem omits to consider that, in the current age of safety technology, in several accident situations the probability of an accident is also affected—at times, greatly affected—by additional factors that are not under the control of the parties. For example, the risk of a car accident is affected not only by the parties' behaviour, but also by the weather, the road conditions, or the quality of the safety technology adopted by the car manufacturer. Although there is not much that tort law can do to bring good weather to us, liability rules can create incentives for third parties who have control over other risk factors.

Let us continue our analysis focusing on a scenario where the risk of an accident arises from the interaction of three parties: an injurer (he), a victim (she), and a third-party (it) who controls external risk factors. By extending the analysis of Guerra et al.,[2] under my proposed rule, when an accident is not attributable to the negligence of injurers and/or victims, the third-party who has the control of the external risk factors would bear the accident loss[3]. I shall refer to this liability-shifting regime as 'third-party residual liability.' Third-party residual liability rules can activate incentives that outperform those of traditional products liability or third-party liability rules – shown below in the example of a no-fault tort occurring at a pool. Most notably, shifting residual liability on third parties can create incentives for third parties, without diluting the injurer's and victim's precaution incentives. Under a third-party residual liability rule, the third-party would bear the accident loss regardless of any negligence on its part (ie, on a strict liability basis) for accidents that are causally attributable to the environmental risks that are under

---

[2] ibid.

[3] Alice Guerra and others, 'Liability for Robots I: Legal Challenges' (2022) 18 Journal of Institutional Economics 311; Alice Guerra and others (n 1) considered the legal and economic aspects of liability for robots. In their analysis, they considered a more specific scenario where the injurer was a robot 'injurer' (i.e. a human who utilised the robot to carry out an activity), causing harm to a victim. The injurer had some form of control over the robot—control which could be exercised in negligent manner. The 'victim' was the party who suffered the harm caused by a robot's decisions. The 'third party' controlling the riskiness of the parties' activities was the robot manufacturer (the authors considered a monolithic manufacturer, including those responsible for the development, programming, and production of the robot). Their analysis considered the application of a rule of 'manufacturer residual liability,' which imposed liability on robot manufacturers for the non-negligent accidents arising in the use of robots.

its control[4]. In this way, third parties would internalise the full benefit (and cost) of improving safety and reducing environmental risks. This distinguishes our rule of third-party residual liability from a rule imposing liability on a third party based on product defects, malfunctions, or other grounds of third-party liability based on negligence[5]. For example, traditional product liability uses the threat of liability to encourage manufacturers to produce well-functioning products that meet the quality standard needed to avoid product liability. However, with modern technology, there are several accidents that are not caused by any defect or malfunction, but are attributable to design limitations that could be corrected over time with additional investments in R&D. These design limitations cannot be legally construed as grounds for products liability, although further improvements in technology and additional investments in R&D might have avoided them. Such improvements are socially desirable but are not encouraged by existing legal rules. Rules of third-party residual liability would push incentives beyond the defect-avoidance threshold of products liability and lead to greater investments by third parties to reduce the environmental risks that are under their control: investments that may lead to the improvement of the safety of the goods and/or services that they produce, such as the quality of roads and the safety of environments that third parties control. For a related discussion regarding the sharing of liability among multiple tortfeasors, see Kornhauser and Resevz[6].

		Injurer	
		Negligent	*Diligent*
Victim	*Negligent*	Share*	Victim
	Diligent	Injurer	Third-Party

Figure 1: *Third-Party Residual Liability (*Under Comparative Negligence)*

---

[4] Hereinafter, I refer to the activities undertaken by the third parties as to include all activities that may affect the risks that are under their control, and which may affect the probability of an accident for other parties.

[5] For example, a car manufacturer would face products liability if a defect in the car causes an accident. Such liability generally arises for defects causing 'malfunctions.' In the case of a car injuring a pedestrian, the driver's (injurer's) liability would arise when the product defect can be identified as the main cause of the harm to a victim. Product liability law would allow victims to sue manufacturers directly or would allow drivers to sue the car's manufacturer in subrogation when they face direct liability under conventional tort law.

[6] Lewis Kornhauser and Richard Revezs, 'Sharing Damages Among Multiple Tortfeasors' (1989) 98 The Yale Law Journal 831.

In Figure 1, I illustrate the allocation of liability under a rule of third-party residual liability when such rule is applied in conjunction with a comparative negligence rule. As shown by Guerra et al[7], the choice of different allocations of the accident loss in cases of bilateral negligence—accident loss falling entirely on the injurer (under simple negligence), accident loss falling entirely on the victim (under contributory negligence), or accident loss shared between the parties (under comparative negligence)—does not affect the parties' care and activity-level incentives[8]. Therefore, in the following, I will conjunctly discuss the effects of third-party residual liability under all the variations of negligence liability.

2 The Functioning of Third-Party Residual Liability

2.1 The social objective

Let's denote the human injurer by *I*, the victim by *V*, and the third-party by *T*. Let *x* and *w* represent the injurer's care and activity levels, respectively, and *y* and *z* the victim's care and activity levels, respectively. The value of the activity is denoted by V_I *(w)* for the injurer, and by V_V *(z)* for the victim. Let *r* represent the third-party's investments to control the external risk affecting the probability of an accident (eg, the investment of a toll road company in maintaining roads in good conditions and with good lighting, a manufacturer's R&D investment for new safety technology to improve the safety of the external environment). The expected accident loss is represented by *L > 0*.

Following Singh[9], Dari-Mattiacci et al.[10], Carbonara et al.[11], and Guerra et al[12] I define the overall cost function (ie, the sum of expected accident costs and actual precaution

---

[7] Alice Guerra and others (n 1).

[8] Simple negligence rules are easier to administer because courts only need to evaluate the behaviour of injurers. Contributory negligence rules are generally used in products liability cases. Misuse of a product or a plaintiff's failure to follow clear instructions and/or warnings provided by the manufacturer are often construed as contributory negligence. Several jurisdictions in the U.S. have abandoned the contributory negligence rule in favour of comparative negligence rules in ordinary negligence cases (non-products liability), when bilateral negligence is established, since the former rule would bar victims from obtaining compensation even when the negligence of the tortfeasor is much more serious. Most civil law systems also utilise comparative negligence rules in apportioning damages in bilateral negligence cases. Alice Guerra and others (n 1) considered coupling with the robot's manufacturer residual liability: (a) simple negligence; (b) contributory negligence; and (c) comparative negligence. The only difference between the three rules is the way in which they allocate the accident loss when the injurer and victim are both negligent (top-left quadrant in Figure 1). Under all three rules, if the injurer is the only negligent party (bottom-left quadrant in Figure 1), the injurer bears the accident loss, and if the victim is the only negligent party (top-right quadrant in Figure 1), the victim bears the accident loss. Under all regimes, when both injurer and victim are diligent (bottom-right quadrant in Figure 1), the accident loss is shifted on the third party.

[9] Ram Singh, 'On the Existence and Efficiency of Equilibria Under Liability Rules' [2006] NBER Working Paper Series 12625.

[10] Giuseppe Dari-Mattiacci and Francesco Parisi, 'Loss-sharing Between Nonnegligent Parties' [2014] Journal of Theoretical and Institutional Economics 571.

[11] Emanuela Carbonara and Others, 'Sharing Residual Liability: The Cheapest Cost Avoider Revisited' (2016) 45 Journal of Legal Studies 173.

[12] Alice Guerra and others (n 1).

costs) as L (*x*, *y*, *w*, *z*, *r*) + *wx* + *zy* + *r*. The social optimization problem is thus Max $S = V_I(w)$ + $V_V(z)$ – L (*x*, *y*, *w*, *z*, *r*) – *wx* – *zy* – *r*. That is, the social objective is to maximise the value of the parties activities, $V_I(w)$ and $V_V(z)$, at the net of the expected accident costs, L (*x*, *y*, *w*, *z*, *r*) and the precaution costs, *wx*, *zy*, and *r*. The injurer's and victim's socially efficient care levels, *x*** and *y***, are identified by differentiating *S* with respect to the relevant control variables. Care levels are efficient when the marginal reduction in the expected accident loss equals the marginal cost of care, *x* and *y*. The parties' socially efficient activity levels, *w*** and *z*** are similarly identified by differentiating S with respect to the relevant variables. The optimality conditions will show that activity levels are socially efficient when the marginal benefit from an increase in activity level equals the marginal cost of the activity (the incremental expected accident cost caused by an increase in activity). The third-party's efficient investment in effort to increase the safety of the external environment, *r***, is similarly obtained by differentiating *S* with respect to *r*. The investment in *r* is socially efficient when the third-party's marginal cost from the research investments equals the social marginal benefit of the increased safety of the environment.

2.2 Injurer's and victim's incentives

To determine the actual incentives of the parties under a third-party residual liability rule, we can follow the derivation of the results in Guerra et al.[13] to show that rules of third-party residual liability maintain optimal care incentives for injurers and victims. That is, even though in equilibrium the accident loss is shifted on a third-party, the injurer's and victim's incentives to undertake precautions remain unaltered.

Similarly, this can be verified by deriving the Nash strategies that injurers and victims would undertake to maximise their payoffs in Figure 1. As it can be seen, both parties' dominant strategies lead them to act diligently, notwithstanding the fact that neither of them bears the accident loss in equilibrium[14]. Additionally, as it will be discussed below, a rule of third-party residual liability creates optimal incentives for the third-party's effort to reduce the riskiness of the external environment (ie, rules of third-party residual liability will lead to the optimal supply of safety by third-parties).

---

[13] Ibid.
[14] Although not yet formally established in the literature, the dominance of optimal care incentives should be robust to role-uncertainty conditions (i.e. situations where parties take 'precautions' without knowing the role they will find themselves in—as injurers or victims—in the event of an accident, as in Barbara Luppi and others, 'Double-edged Torts' (2016) 46 International Review of Law and Economics 43.

2.3 Third parties' incentives

Under a third-party residual liability rule, third parties will have incentives to invest optimally to mitigate the risks that are under their control. This result does not hinge upon the observability of the third parties' behaviour by prospective injurers or victims. Nor is it needed for consumers to be willing to pay a premium for safer goods or services offered by the third parties, as required in Hay and Spier[15]. Instead, competitive market forces will cause the safest goods or services to be in greater demand, if only because they will be offered at the lowest price. This follows from the fact that, under this rule, the total cost of the third parties' goods or services captures the expected residual liability that they face. Third parties will minimise total production costs, optimally balancing the marginal investments in r with the marginal reduction in their expected liability. Third parties will be investing in development, maintenance, and post-production improvements (eg. recalls, replacements, upgrades, etc.) to reduce the level of environmental risk faced by parties who utilise their goods or services (hence, to reduce their expected residual liability). Under third-party residual liability, third parties would make their production decisions accounting for the net present value of their future liability, rather than looking solely at the bare development and production costs that they would face in the absence of residual liability. This induces third parties to invest in safety, until they reach the socially optimal level of r^{**} as defined above[16]. By doing so, third parties who have better safety technology will also be able to offer their goods or services more affordably. When third parties face residual liability, they fully internalise the benefits of safety of their services or products. Once the good or service is supplied, and is in the hands of the injurer, the third party is unable to influence the risk of injury. Non-negligent accidents impose a cost of liability on the third party. The expected cost of future non-negligent accidents becomes part of the third party's cost of production. When determining their optimal total investment in quality and technology—which includes the ex-ante investments in R&D for safety, and the ex-post cost of updates, and maintenance plans to maintain or improve safety—third parties will balance the cost of these safety investments and reduction in their expected liability. The resulting choices will determine the price of their product in a competitive market[17].

---

[15] Bruce Hay and Kathryn E Spier, 'Third-party Liability for Harms Caused by Consumer to Others' (2005) 95 American Economic Review 1700.

[16] Making the third party internalise the full cost of the harm causes the price of the good or service to reflect its dangerousness, whereby more dangerous third-party's goods or services would be more expensive, and safer third-party goods or services would be less expensive. Similar to what was pointed out by Alice Guerra and others (n 1) in the context of market for robots, third parties would compete on price to sell their goods or services and by doing so they would compete on safety, producing and maintaining optimal levels of safety, minimising price. Even if injurers are not held liable for the non-negligent harm by their activity, competitive market forces would lead to the supply and demand of safer third-party goods or services, regardless of whether injurers are informed—or even care about safety as a social value—when making their purchasing decisions.

[17] Alice Guerra and others (n 1).

The fact that safer third-party goods or services will be less expensive would encourage consumers to acquire them, even in the absence of full information about safety on their part. The result of these supply and demand effects will generate prices that reveal the riskiness of third-party's goods and services and will lead (even uninformed) parties toward safer environments[18]. This risk-revelation mechanism will be particularly desiderable in our context because the opportunities and costs of developing safer technologies are not easily verifiable in court, and the alternative means of incentivizing safety through regulation and negligence rules would be hard to implement at the policy level. Regulators have no direct information to determine what would be the socially optimal advancement in safety technology and further it would be difficult to adopt safety standards, given the opacity of the relevant information for regulators. It would be equally difficult for courts to establish standards of "due safety development" and use them as benchmarks of liability, holding third parties responsible for not having optimally invested in safety development. Investment decisions in R&D are best delegated to third parties, who have direct information about the costs and benefits of safety development.

Under a rule of third-party residual liability, a gradual—and allocatively efficient—spread of safety technology will be obtained, with safer technology introduced in environments that are most exposed to the risk caused by the third party. For example, as pointed out by Guerra et al.[19], when the third parties are suppliers of goods that are used by prospective injurers, residual liability rules create optimal incentives for such third parties to provide upgrades of technology and safety solutions to individuals who already acquired their goods. Third-party manufacturers will initially invest in providing free upgrades or replacements to high-activity-level prospective injurers, and later to others. Adopting safer technology reduces the expected accident cost of carrying out an activity and injurers with high activity levels would benefit more from operating in safer environments or adopting safer third-party goods or services. Allocative efficiency will result and will lead to advances in safety technology that will be optimally allocated: newer and safer environments and products will first be made available and accessed by those who would use (and benefit from) them the most.

---

[18] In many products liability models, the belief that safer products will develop in the market rests on two fundamental assumptions: (i) that consumers are willing to pay a premium for safer products, and (ii) that product safety is perfectly observable to consumers when making their purchasing decisions, see Alan Mitchell Polinsky, 'Strict Liability vs. Negligence in a Market Setting' (1980) 70 American Economic Review 363; Hay and Spier (n 15). In a world of fast-evolving safety technology, neither of these assumptions is likely to hold. First, in the three-party scenario we consider, injurers are only interested in avoiding liability, which they can do by adopting due care in their activity. They would not be willing to pay a premium to operate in a safer environment, because any such investment in safety would reduce the risk of accidents, not their expected liability. Second, the safety of the environment or technology controlled by third parties is not observable by prospective injurers or victims prior to engaging in their activities (e.g. specific shortfalls in the maintenance of toll roads or level of advancement of the safety technology used by sporting facilities is only learned over time, through experience and use). These factors are unknown to the users, just as they may be unknown to the third party before the risk materialises.

[19] Alice Guerra and others (n 1).

3 Percolation Effects of Third Party Residual Liability

3.1 Third parties' incentives to monitor the care levels of injurers and victims

Under a negligence regime, victims have the burden of proving the negligence of their injurers to obtain compensation. Compensation is a powerful motivation for prospective victims to collect and produce evidence of the injurer's negligence. Third-party residual liability rules may reduce the efforts of the victim to prove the negligence of the injurer. A victim can bring an action and obtain compensation, even if she fails to prove the injurer's negligence: in the absence of proof of the injurer's negligence, the victim would still be able to obtain compensation from the third party.

However, fortunately the victim's reduced incentives to prove the injurer's negligence do not undermine the injurer's incentives to invest in optimal precautions. Third-party residual liability rules shift evidentiary incentives from plaintiffs to third-party defendants. The threat of residual liability incentivises third parties to monitor the activities of prospective injurers and victims that operate in their environments, because proving the negligence of injurers and/or victims would shield the third party from residual liability. Third parties will anticipate the victims' reliance on their residual liability to obtain compensation and will thus be incentivised to invest resources and to produce evidence to show that the accident was caused by the injurer's (or the victim's) negligence[20]. This would allow them to avoid liability in case of accidents attributable to the parties' negligent behaviour. In our running examples, the company that manages a toll road would have incentives to install video cameras or other evidence technology to determine if any accident was caused by the parties' negligence. Likewise, the owner of a sports facility would have incentives to monitor the users' activities. In this way, the victim's reduced litigation efforts may be (often, more than fully) offset by the third-party's efforts to establish the injurer's negligence.

In several circumstances third parties may have a comparative advantage over parties in documenting and proving the parties' violation of duties of due care. In these cases, shifting litigation incentives from victims to third parties would thus increase the probability that negligent injurers may face liability, thereby reinforcing their primary incentives to adopt due care[21].

---

[20] When the third-party's residual liability operates under rules of contributory or comparative negligence (as in Figure 1), third parties would have incentives to also monitor victims' care levels, since under this rule the finding of victims' negligence would equally shield third parties from liability.

[21] When such comparative advantage in obtaining evidence is not present, an inversion of legal presumptions, shifting the burden of proof on defendants would change the analysis above. See Alice Guerra and others (n 1). Possibly, creating legal presumptions of negligence (and contributory negligence) on injurers (and victims) would be a way to overcome the dilution of evidentiary incentives that third-party residual liability rules may create. For a related analysis of the incentives to invest in private evidence under alternative presumption and discovery rules, see Alice Guerra and Francesco Parisi, 'Investing in Private Evidence: The Effect of Adversarial Discovery' (2021) 13 Journal of Legal Analysis 657; Francesco Parisi and Giampaolo Frezza, 'Burdens of Proof in Establishing Negligence: A Comparative Law and Economics Analysis' (2023) 9 Italian Law Journal, forthcoming.

Under our legal regime, we should therefore expect third parties to adopt evidence technology (eg, to install video cameras or speed detectors on toll roads, or to incorporate other evidence technologies into their third-party goods or services), to produce evidence showing that any given accident was caused by the parties' negligence and should not fall under their blanket residual liability coverage. As pointed out by Guerra and Parisi[22], and Parisi and Frezza[23], exposure to such evidence technology will, in turn, strengthen the injurers' (and victims') incentives to adopt due care in environments controlled by the third party.

As an illustrative example of third-party residual liability, consider the case of a swimming pool or sporting facility. Here, a third party, the facility operator, controls an environment in which accidents may occur. Under a traditional tort regime, it is typical for the actions of only the injurer and victim to be considered. Under a traditional tort law and products liability regime, the swimming pool or sporting facility operators are incentivised to ensure that there are no unsafe areas on their premises that could cause an accident to occur, such as a faulty diving board, dangerous blind corners, or obstructed safety exits. However, when liability is extended through third-party residual liability, the operators would internalise the costs of all accident risk factors in circumstances where neither the injurer nor the victim acted negligently. The facilities operators would invest in the development of safety and monitoring procedures beyond those that would be incentivised by standard liability rules. A third-party residual liability rule would incentivise the facility operators to mitigate the environmental risks while monitoring the injurer's and victim's activities on their premises.

3.2 Correcting injurer's and victim's moral hazard through third-party residual liability

As suggested above, in situations where third parties control the safety of the environment in which prospective injurers and victims operate, an optimal liability regime might be one where the third party is the sole residual bearer of the accident loss. The logic is two-fold. First, when third parties face residual liability, they have optimal incentives to innovate and improve the safety of the environment under their control. This is especially important considering that the safety level of the environment is often undetectable by the users. Second, as well-known in the literature, negligence rules cannot incentivise non-verifiable precautions because non-verifiable precautions are undetectable by courts in determining negligence. However, incentives for 'non-verifiable precautions' can be created through the allocation of residual liability[24]. In our three-party scenario with injurers, victims, and third parties, by allocating the residual liability

---

[22] Alice Guerra and Francesco Parisi (n 21).
[23] Francesco Parisi and Giampaolo Frezza (n 21).
[24] Giuseppe Dari-Mattiacci and Francesco Parisi, 'The Economics of Tort Law' in J.G. Backhaus (eds), *The Elgar Companion to Law and Economics* (Edward Elgar Publishing 2005).

to third parties, our rule would create optimal incentives for third parties' 'non-verifiable' R&D investments, leading to safer third-party goods or services[25].

Because residual liability can only be placed on one party, as per Shavell's theorem[26], it may seem that we cannot simultaneously incentivise third parties to produce safer environments and at the same time encourage injurers and victims to mitigate their activity levels. In the absence of a price relationship between the parties, neither injurers nor victims internalise the risk caused by their non-negligent activities, with a resulting excessive activity level for both parties. This misalignment of incentives occurs because injurers (and victims) derive benefits from their activities and use third-party goods and services, disregarding the fact that their activities—albeit diligently carried out—increase the probability of accidents, with a resulting externality on the third party who bears residual liability. Most readers will recognise this problem as a three-party incarnation of Shavell's[27] activity-level theorem[28].

From a policy point of view, in the absence of a price mechanism a difficult choice would thus arise: should residual liability be allocated on injurers to incentivise them to undertake optimal activity levels, or should residual liability be allocated on third parties to incentivise them to optimally control the risks and the non-verifiable safety investments that they undertake? Fortunately, as notably pointed out by Guerra et al.[29], these two policy objectives are not necessarily mutually exclusive. Market mechanisms are likely to emerge and cause a percolation of the third-party's residual liability, into the injurers' (and their prospective victims') objective functions. For example, with current technology, third parties can keep track of the injurer's activity level (eg, toll roads can keep track of a car's usage of the road; manufacturers can verify the level of usage of their product, swimming pool and sports facilities operators can monitor the access and activity levels of users, etc.). In a competitive market, third parties would have strong incentives to develop a price mechanism to transfer the cost of the risk created by the activities of injurers and victims back to them. This pricing mechanism will induce parties to internalise the cost of the expected residual liability that they create on third parties, and in turn reduce their activities to socially optimal levels. As a simple illustration, a fee can be charged by the third party to the injurer based on his activity level (eg, a toll road company can charge a per-mile fee to cars; a third-party manufacturer can charge a maintenance or leasing fee based on usage, sports facilities could charge per-hour access

---

[25] In the standard tort model, residual liability should be imposed on the party whose non-verifiable precautions most effectively reduce the cost of accidents, see Emanuela Carbonara and others (n 11).
[26] Steven Shavell, 'Strict Liability versus Negligence' (1980) 9 Journal of Legal Studies 1.
[27] Ibid.
[28] Shavell's theorem holds that only the bearer of residual liability is incentivised to undertake precautions that are not incentivised by the negligence standard. This is because the party who does not bear residual liability only wants to avoid liability by showing that he adopted due care, whereas the bearer of residual liability wants to avoid causing harm tout court. For a more extensive explanation and discussion of this theorem, see 'Shavell's Activity Level Theorem' in Francesco Parisi, '*The Language of Law and Economics: A Dictionary*' (Cambridge University Press 2013).
[29] Alice Guerra and others (n 1).

fees, etc.). The third party will set the fee equal to the cost of the expected incremental liability generated by the injurer's activity level. As a result, the injurer will internalise the risk created by his (non-negligent) activity and will choose an activity level that aligns with the socially optimal level w^{**}, as defined above[30]. In the context of a swimming pool or sports facility operator, for example, a pricing mechanism would likely be adopted to induce users to internalise the cost of expected residual liability they may create on the facility operator. This pricing mechanism could be in the form of an added entrance fee to use the pool or sports facility, or possibly a fee based on time spent on the operators' premises. This percolation will lead to injurers and victims to perform at socially optimal levels of activity.

Additionally, as pointed out by Guerra et al.[31] if prospective injurers and victims are contractually related (eg, the injurer is using a toll road to offer a limousine service to the victim), the cost of the service would increase to reflect the extra fees charged by the third party to the injurer. Through these price adjustments, incentives would percolate from the third-party supplier of the service to prospective injurers, and from prospective injurers down to prospective victims, aligning the incentives of all parties (i.e. the toll road company would make optimal safety investments, r^{**}, and the limousine driver and his passenger would mitigate their activity levels, w^{**} and z^{**}, in response to the higher fees being charged).

4 Conclusions

To the extent that we want the assignment of residual liability to affect the incentives of as many parties as possible, it would be desirable assigning residual liability to the party who can more easily affect the incentives of others (e.g. if a third party can more easily affect the incentives of injurers and victims, compared to injurers and victims affecting the incentives of third parties, a third-party residual liability rule would be preferable). In three-party incentive-alignment problems, it may often be the case that the allocation of residual liability on third parties may be preferable to alternative allocations. If injurers or victims were assigned residual liability, the inverse percolation of incentives toward the third-party would not be equally practical. Contractual and market mechanisms can be more easily imagined transmitting residual liability incentives from third parties to injurers and victims, although they may also potentially be constructed to transmit incentives upstream from injurers and victims to third parties. The percolation effect of residual liability on activity levels will not undermine the third-party's incentives to produce safer third-party goods or services. Even when third parties can transfer the cost

---

[30] Several mechanisms could be implemented to transmit incentives from third parties to injurers through the price system. In a competitive market, we expect pricing mechanisms to shift the expected cost of non-negligent accidents associated with higher activity levels to injurers. In this way, the residual liability incentives faced by the third parties would percolate downstream to the injurers, leading them to undertake optimal activity levels.
[31] Alice Guerra and others (n 1).

of their expected residual liability back to injurers (eg, charging user fees equal to the residual liability associated with the parties' activity levels), their incentives to invest in safety would remain in place. Newer, safer third-party goods or services would in fact be cheaper and more attractive in the marketplace. Their higher level of safety would entail lower activity level externalities: the percolation effect of third parties' residual liability on users would be of lesser magnitude and would contribute to the financial attractiveness of the safer third parties' goods or services. This would thus be a contributing force that incentivises third parties to optimally reduce the risks that are under their control.

As a final note, third-party residual liability regimes are a second-best solution to our three-party incentive problem, and our analysis relied on the implicit assumption that third parties' investments in safety are more important factors in preventing accidents, compared to the parties' activity levels. In the absence of price mechanisms connecting all parties, there will always be some misaligned incentives, and the optimal allocation of the residual liability turns on the empirical question of which is the most important and cost-effective in reducing the risk of accidents.

Juanita Goicovici [*]

REBUTTABLE PRESUMPTIONS OF CAUSALITY AND REVERBERATIONS OF EVIDENCE DISCLOSURE, AS EPITOMIC PIECES IN THE PHYSIOGNOMY OF LIABILITY FOR DEFECTIVE AI

Abstract

The paper examines the typification of the rebuttable presumptions of causality the use of which has been introduced by the new "Proposal for a Directive of the Parliament and of the Council on adapting non-contractual civil liability rules to artificial intelligence" ("AILD Proposal") while insisting on the quintessence of the binomial set of presumptions aimed at facilitating the probatory efforts and resorting to a "disclosure of evidence mechanism" in tort litigation founded on AI deficiencies generating bodily harm and/or patrimonial losses to consumers. As resulting from the provisions of Art. 1, para 1 and 2 of the AILD Proposal, the 'hybridisation' of the conceptual nature of civil liability applied in matters concerning autonomous / embedded AI defectiveness does not remain deprived of potential consequences in terms of the reserved possibility of judicial courts, to assess the existence of the illegal action/omission of the responsible persons by referring to compliance with the transparency obligation incumbent on AI providers, as resulting from Art. 13 of the Draft Regulation (EU) on AI (Artificial Intelligence Act) (i) or to retain the civil liability of manufacturers/importers or AI providers for bodily/patrimonial damages caused to consumers, in the assumptions concerning the ignoring, at the time of AI system design or development, of the necessity of "human effective surveillance" for the time bars related to the AI usage (ii).

JEL CLASSIFICATION: K13, K39, K41

---

[*] Associate Professor PhD. at the Faculty of Law of the Babeș-Bolyai University of Cluj-Napoca.

pertinent use of presumptions of causality in cases concerning 'self-learning' / 'self-evolving' AI algorithmic categories – 6 Concluding remarks

1 Foreword and introductive observations

The contemporary efforts to scrutinise the 'classic' liability regimes' capacity to metabolise and absorb civil liability rules adapted to 'harm causing' artificial intelligence, are constantly intensifying, while the physiognomy of specific liability[1] for autonomous or incorporated AI defective products are considerably influenced by their 'regimenting' into 'risk categories' scaled from 'highly threatening' to 'lower degrees' of menacing.

Allogeneic for this type of civil liability remains the presence (and recourse to) rebuttable presumptions to fill evidentiary gaps that could make particularly vulnerable the procedural posture of the consumer injured by the defective autonomous or embedded AI action/omission or by the interaction with algorithmic systems showing design flaws/manufacturing defectiveness. As resulting from the provisions of Art. 1 para 1 and 2 of the AILD Proposal, the "hybridisation" of the conceptual nature of civil liability applied in matters concerning autonomous / embedded AI defectiveness does not remain deprived of potential consequences in terms of the reserved possibility of judicial courts, to assess the existence of the illegal action/omission of the responsible persons by referring to compliance with the transparency obligation incumbent on AI providers[2], as resulting from Art. 13 of the Draft Regulation (EU) on AI (Artificial Intelligence Act) (i) or to retain the civil liability[3] of manufacturers/importers or AI providers for bodily/patrimonial damages caused to consumers[4], in the assumptions concerning the ignoring, at the time of AI system design or development, of the necessity of "human effective surveillance" for the time bars related to the AI usage (ii).

Secondly, the "allopathic" conceptual approach, which appeals to the AILD Proposal by consecrating a "binary" set of presumptions of causation in the perimeter of civil liability for AI harmful deficiencies[5], is rendering permeable the border between 'classical' civil responsibility based on culpable behaviour / gross negligence and, on the other versant of the discussion, objective responsibility[6], completely detached from the element of

---

[1] Mihailis Diamantis, 'Who Pays for AI Injury?' (Oxford Business Law Blog, 04 May 2020) <https://blogs.law.ox.ac.uk/business-law-blog/blog/2020/05/who-pays-ai-injury> accessed 20 October 2023.

[2] Ajay Agrawal, Joshua Gans and Avi Goldfarb, *Máquinas predictivas. La sencilla economía de la inteligencia artificial* (Barcelona: Editorial Reverté 2019) 42, 48.

[3] María Luisa Atienza Navarro, *Daños causados por inteligencia artificial y responsabilidad civil* (Barcelona: Editorial Atelier 2022) 74, 81.

[4] Jan de Bruyne, Orian Dheu, and Charlotte Ducuing, 'The European Commission's Approach to Extra-Contractual Liability and AI – A First Analysis and Evaluation of the Two Proposals' [2022] SSNR Electronic Journal <https://ssrn.com/abstract=4239792> accessed on 15 July 2023.

[5] Hélène Christodoulou, 'La responsabilité civile extracontractuelle à l'épreuve de l'intelligence artificielle' (Lexbase Hebdo édition privée 2019 n 807. hal-03349668f) <https://hal.science/hal-03349668/document> accessed 12 July 2023.

[6] Jia Qing Yap and Ernest Lim, 'A Legal Framework for Artificial Intelligence Fairness Reporting' (2022) 81 (3) Cambridge Law Journal 610, 644.

culpable behaviour[7], due to the imperative to facilitate the administration of evidence in compensatory actions[8] initiated by the prejudiced consumer[9], whose interests have been harmed by interaction with deficient AI systems. Starting from the findings on the vulnerability of the procedural position of the plaintiff in 'classical' actions, the admitting of which would require proof of the connection relating the incidence of product defectiveness to the prejudicial effects that were generated, disclosure of elements of proof of "high-risk AI systems" seems to enable the claimant to plead for tort liability based on non-contractual fault (i), while placing the burden of proof on the shoulders of professional defendants, in the perimeter of providing access to AI systems whose complexity often exceeds (including) the predictions of its creators (for AI systems from self-learning / self-evolving AI taxonomy). Consistent with the desiderata of 'unburdening' consumers in providing relevant evidence on damage causation while embarking on non-contractual civil law actions (based on fault / culpable behaviour) related to AI system deficiencies[10], consecrating courts' possibilities to order disclosure of evidence on professional defendants[11] (ii) remains the key premise in understanding the innovative system proposed in the AILD Proposal.

Thirdly, instead of focusing on a more permeable or 'fluid' approach to the subjective element of culpable behaviour as a central pillar for retaining the producers' responsibility, or that of the importer or supplier of the deficient AI, the implications of using a presumption relative to the existence of imputable behaviour would allow the consumer, as claimant, to engage in remedial actions without being placed in the undesirable position of not meeting the particularly difficult evidentiary demands regarding AI behaviour as a source of bodily/patrimonial damage; the proposed regulatory norms are permitting professionals the reversing of the presumption of causation in B2C relations, starting from the premise that the relevant information regarding the nexus of (non)causality[12] is rather exclusively in their possession than in the possession of the

---

[7] Margo Bernelin, 'Intelligence artificielle: une proposition de directive sur la responsabilité civile extracontractuelle' Dalloz actualité (22 Nov. 2022) <https://www.dalloz-actualite.fr/flash/intelligence-artificielle-une-proposition-de-directive-sur-responsabilite-civile-extracontract> accessed 12 July 2023.

[8] Christoph Busch, 'Rethinking Product Liability Rules for Online Marketplaces: A Comparative Perspective' (The 49th Research Conference on Communication, Information and Internet Policy, 22-24 September 2021) <https://ssrn.com/abstract=3897602> accessed on 15 July 2023; Christoph Busch, 'When Product Liability Meets the Platform Economy: A European Perspective on Oberdorf v. Amazon' (2019) 8 (5) Journal of European Consumer and Market Law 173.

[9] Mehmet Fatih Burak, 'Effects of Artificial Intelligence on E-Commerce' in A N Özker (ed), *Reviews in Administrative and Economic Science Methodology, Research and Application* (Livre de Lyon Publishing 2022) 91, 100.

[10] Miriam Buiten, Alexandre de Streel and Martin Peitz, 'The law and economics of AI liability' (2023) 48 Computer Law & Security Review.

[11] Laurène Mazeau, 'Intelligence artificielle et responsabilité civile: Le cas des logiciels d'aide à la décision en matière médicale (2018) 1 Revue pratique de la prospective et de l'innovation 38, 42.

[12] Alan Butler, 'Products Liability, and the Internet of (Insecure) Things: Should Manufacturers Be Liable for Damage Caused by Hacked Devices?' (2017) 50 (4) University of Michigan Journal of Law Reform 913.

profane consumer[13]. The pertinence of causality presumptions is worth discussing, on the binary premises (that remain partly subjective, generated by the element of faulty conduct and partly objective, focused on the element of the presence of the design/ manufacturing defect of the AI system) of engaging in the civil liability of producers/importers and the contrast established by reference to 'classical' versions of subjective responsibility; especially, the discussion may be conducted from the perspective of consecrating a taxonomy of defects covered by the substantial sphere of civil liability, divided into manufacturing defects versus design flaws and informative deficiencies[14]. Deriving from the scope of penalising the non-fulfilment of the transparency obligation incumbent on the manufacturer/importer of the defective AI systems, to which the defects are added in an autonomous manner, there are four types of defective AI response that may reverberate on the civil liability regimen, including the one consisting in jeopardising the safety of the data uploaded by the consumer, as a distinct species of damage whose coverage will be envisaged by the adapted civil liability regime for the remediation of damages caused by defective AI conduct.

2 Binary taxonomies of rebuttable presumptions of causality under tort liability regimen

2.1 Presumption of causal nexus between defendant's faulty conduct and damageable AI result

Saliently, the mechanism described in Article 4 of the AILD Proposal, which provides for a reversible presumption of causal proximity between the defendant's fault or culpable conduct or inexcusable negligence, remains focused on presuming the existence of a causal nexus; the latter's applicability would be differentiated on grounds related to the range of potentially risky behaviour[15], under which the AI systems have been regimented. In the foreground of the discussion, when engaging in the evaluation of elementary premises for retaining the specific liability of the AI manufacturer/importer, the latter's faulty conduct could also be presumed[16] by national courts based on failure to comply with a judgment on judicial disclosure or preservation of evidence pursuant to the provisions of Art. 3 para 5 of the AILD Proposal. Although the presumption of causation

---

[13] Aileen Nielsen, 'How to measure and regulate the attention costs of consumer technology' (TechStream: Brookings Institute, 4 November 2021) <https://www.brookings.edu/techstream/how-to-measure-and-regulate-the-attention-costs-of-consumer-technology/> accessed 12 July 2023.

[14] Antonia Waltermann, 'On the legal responsibility of artificially intelligent agents. Addressing three misconceptions' (2021) 3 Technology and Regulation 35, 43 <https://techreg.org/article/view/10985/11959> accessed 12 July 2023.

[15] Tiago Sérgio Cabral, 'Robotics and AI in the European Union: opportunities and challenges' (2018) 4 (2) UNIO - EU Law Journal 135, 146.

[16] Tiago Sérgio Cabral, 'Liability and artificial intelligence in the EU: Assessing the adequacy of the current Product Liability Directive' (2020) 27 (5) Maastricht Journal of European and Comparative Law 615, 635.

gains in relevance[17] in hypotheses, in which the inexcusable omission or culpable action of the AI manufacturer/designer has influenced in a decisively obvious manner the poor performance of the AI system[18] (which would be assessed based on the circumstantial peculiarities of the case), the claimant is expected to produce evidence of existence of the damage. Facing the arduous task of bringing consistent evidence that, in consideration of the damageable results generated by the AI system collapse, or the failure of the AI to provide for an appropriate result, the consumer would still be expected to bring proof of the causal link between the registered damageable result and the alleged dysfunctionality of the AI system, in the perimeter of interactions with AI systems regimented in the medium/low-risk categories; reversibly, concerning damages generated by "high-risk AI systems", the text of Art. 4 para 4 of the AILD Proposal establishes an exception to the necessity of proof by the plaintiff consumer of said causal link, establishing the (relative) presumption of causality. Except in the (rare) cases where the professional defendant proves that, in absence of recourse to the presumption of causation, sufficient evidence and an appropriate level of expertise are reasonably accessible to the complaining consumer, the latter is seemingly enabled to prove the existence of the causal link between the faulty behaviour of the AI system and the recording of the physical/property damage.

Controversy is fuelled by the fact that the incidence of the exception from the sphere of incidence of the causality presumption is subject to the courts' decision involved in settling the dispute, which could find it excessively difficult for the claimant to provide evidence as to the causation of the damage even in cases of medium-risk AI use; obviously evidentiary difficulties will be assessed bearing in mind the characteristics of AI technologies, describable as autonomous conduct, or AI opaque behaviour, rendering the consumer's assignment of explaining the peculiarities of an AI system almost impossible to accomplish from the perspective of the availability of evidence, since the claimant is (almost invariably) facing difficulties to prove (in absence of such a reversible presumption) a sufficient nexus connecting the consumer's decision concerning the selected design parameters for the manufactured AI product[19] and the biological / patrimonial harm caused to the consumer in the interaction with the defective AI system[20]. In terms of the proportionality requirements for ordering the disclosure or preservation (by the defendant AI manufacturer / supplier) of evidence regarding the

---

[17] Ryan Calo, 'Artificial Intelligence Policy: A Primer and Roadmap' (2017) 51 UC Davis Law Review 399, 435.

[18] Tiago Sérgio Cabral, Iakovina Kindylidi, 'Sustainability of AI: the case of provision of information to consumers' (2021) 13 (21) Sustainability <https://doi.org/10.3390/su132112064> accessed 23 October 2023.

[19] Juanita Goicovici, 'Matricea răspunderii civile extracontractuale pentru prejudiciile cauzate de produsele cu defecte de manufacturare, între testul riscuri-beneficii și testul așteptărilor legitime ale consumatorului' [The Matrix of Non-Contractual Civil Liability for Damage Caused by Products with Manufacturing Defects, between the Risk-Benefit Test and the Test of Legitimate Expectations of the Consumer] (2022) 67 (1) Studia Universitatis Babes Bolyai-Iurisprudentia 106, 185.

[20] ENISA (European Union Agency for Cybersecurity), *Cybersecurity of AI and Standardisation*, March 2023, <https://www.enisa.europa.eu/publications/cybersecurity-of-ai-and-standardisation> accessed 12 July 2023.

compliance / dysfunctionality of high-risk AI systems[21], it is worth mentioning that the courts' assessment of the plausibility of the consumer's complaint remains essential in the scope of the analysis undertaken a priori by the courts, which may order that the defendant be bound to disclose pertinent evidence regarding suspected "high-risk AI systems", even if this evidence incriminates the debtor of the obligation, by way of derogation from the procedural *actor incumbit probatio* principle (given the fact that the lay consumer does not, most often, possesses evidence that is conclusive or relevant regarding the placement of AI system malfunctions at the time its design/manufacturing protocols were selected); the central question of the plausibility of the claim encapsulates the need, for national courts[22], to establish the existence of sufficient grounds for ordering the disclosure or preservation of evidence regarding "high-risk" AI systems the behaviour of which is suspected to be linked to the occurrence of the prejudice[23].

2.2 Presumption of non-observance of the AI provider's duty of care

The launching on the market of defective AI, encompassing AI design defectiveness, as well as manufacturing flaws, might represent per se a form of violating the duty of care incumbent on the AI providers, both at the manufacturing stage, as well as at the pre-contractual stage, in B2B or B2C contextualised relations; the machine-to-machine (M2M) or AI-to-AI contractual interaction might also raise questions on the pertinence of by design and by default selections operated by the AI manufacturer, in terms of consumer safety[24], or through the lens of the 'risks – benefits' balance. Exploiting the *ex-ante* mechanisms of the AI producer's 'duty of care' remains crucial for the success of the liability systems, both at national and transnational levels.

Moreover, in the continuous process of identifying solutions for entailing AI producers' or AI importers' liability for harmful defectiveness, the problems themselves are subject to evolving, thus requiring higher degrees of adaptability for 'classical' remedial paradigms.

The triptych of defect categories engloebed in the objective sphere of the specific liability of AI producers/importers can be distinguished on three levels (to which the fourth is added, consisting of 'loss/damage or alteration of consumer's data'): (i) the objective liability, covering AI manufacturing defects, as defects due to human errors,

---

[21] Michael P. Chatzipanagiotis, 'Product Liability Directive and Software Updates of Automated Vehicles' (Proceedings of SETN 2020 – 11th Hellenic Conference on Artificial Intelligence, 2020) <https://ssrn.com/abstract=3759910> accessed on 15 July 2023.

[22] Juanita Goicovici, 'The inapplicability of personal exceptions between joint debtors and creditors, under Romanian and French private law' in *Looking for New Paths in Comparative and International Law* (Societatea de Stiinte Juridice si Administrativa 2021) 85, 98.

[23] Geraint Howells, Christian Twigg-Flesner, 'Interconnectivity and Liability: AI and the Internet of Things' in Larry A DiMatteo, Cristina Poncibò and Michel Cannarsa (eds), *The Cambridge Handbook of Artificial Intelligence: Global Perspectives on Law and Ethics* (Cambridge University Press 2022) 179, 199.

[24] Juanita Goicovici, *Dreptul relaţiilor dintre profesionişti şi consumatori* (Hamangiu 2022) 183.

generating AI defects that appear or manifest in the 'design – production – distribution' chain (AI structural defects in the design of algorithmic systems, defects or failures of AI equipment etc.); (ii) types of extra-contractual liability for design defects of AI, or for excessive risk compared to the benefits, in which the respective product incorporating AI was designed in manners ignoring the balance of risks and benefits for consumption[25]; (iii) in the absence of deficiencies in the first two categories, extra-contractual liability can be generated by informational deficiencies[26], consisting of incomplete, inadequate or erroneous information regarding the product assuming autonomous or embedded AI, as described in Proposal COM/2021/206 ("Artificial Intelligence Act"). We argue that flawed or erroneous decisions generated by algorithmic systems, as well as low resistance to performance alterations of autonomous/embedded AI in products, can be preferentially treated as product malfunctions, thus allowing the consumer to implicate the manufacturer/importer in damages specific to AI systems (while preserving the duality of 'autonomous material damage vs. derivative material damage', as well as the dichotomous approach of the categories of 'liability for product security deficiencies' versus 'AI design defects').

The dynamics of adapting the 'classic' regime of "civil liability" for physical/bodily damages caused by products presenting design/manufacturing defects to the legal challenges triggered by the gradual generalisation of the use of autonomous/embedded AI are characterised, above all, by aspects such as the adaptation of the burden of proof, by placing some significant sequences of the provision of evidence in the charge of the professional (located, as a rule, in the position of the defendant in the action initiated by the injured consumer) (i) or even by postulating some (relative) presumptions of causality between the existence of the physical/property damage invoked by the claimant and the presence of the defect in the AI product/AI components (ii). In a carefully calibrated dosage, the incidence of the relative presumption of connection is to be confined to the perimeter of cases where, as highlighted in Recital (25) of the preamble of the Draft Directive on the adaptation of the civil liability regime (…), non-compliance with the due diligence obligations incumbent on the manufacturer or on the designer of the AI system, would set premises for retaining producers' responsibility for damage coverage.

---

[25] Mihailis Diamantis, 'Vicarious Liability for AI' [2021] SSNR Electronic Journal <https://ssrn.com/abstract=3850418> accessed 12 July 2023.

[26] Juanita Goicovici, 'The Traders' Liability for Lack of Conformity of the Digital Content and of the Digital Services, as Regulated by Directive (EU) 2019/770' (2020) 66 (1 Suppl.) Analele Științifice ale Universității Alexandru Ioan Cuza din Iași, seria Științe Juridice 79, 98.

3 Presuming causality in liability litigation pillared on damageable AI deficiencies / AI flaws

3.1 Administration of evidence in liability claims conducted towards the providers of "high-risks" AI systems

At the horizon of the special civil liability whose specific regime is outlined in Article 7 of the AILD Proposal remain the B2C relationships involving the use or interaction of the consumer (in the sense of natural person, using the disputed product in a predominantly extra-professional context) with actions/omissions of algorithmic systems or deficient autonomous/incorporated AI[27], without the need to establish the existence of a contractual link connecting the defendant and the plaintiff in the action brought before the courts; the extra-contractual nature of the liability remains decisive (in view of the fact that, in most cases, the injured consumer did not contract directly with the designer/ manufacturer of the defective AI). Nevertheless, being closer to a tort liability regime, the special liability the regulating of which is envisaged by the AILD Proposal does not completely distance itself from the subjective element of the fault of the (potentially) responsible subjects[28]. Despite the fact that it does not suddenly or dramatically alter the (preponderantly) subjective nature of the pre-existing liability regimes in the member states, at the level of repairing damages caused by the illegal/imputable or inexcusable action/omission, the harmonised set of rules on liability for AI defectiveness 'dilutes' the mentioned subjective nature accordingly, by consecrating a relative presumption of causality applicable against the responsible persons[29], in the presence of malfunctions of the AI systems whose origin can be placed at the time of AI designing / manufacturing of the algorithmic systems suspected to be at the origin of the damage. The absence/presence of a direct or indirect/ successive contractual link between the responsible person and the consumer injured by the action/omission of the AI system is not crucially relevant, since it will suffice to locate the defect at the time the producer decided on the launching of defective design/defective manufacturing of the harmful autonomous AI (including AI from the self-learning category[30])/ detrimental embedded AI, and even for prejudicial generative AI.

---

[27] Mikołaj Domagała, 'Threats associated with the introduction of autonomous vehicles as an example of the negative effects of the development of artificial intelligence' in Luigi Lai, Marek Świerczyński (eds), *Legal and technical aspects of artificial intelligence* (Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego 2021) 247, 264.

[28] Mark Coeckelbergh, 'Artificial Intelligence: Some ethical issues and regulatory challenges' (2019) 1 Technology and Regulation <https://doi.org/10.26116/techreg.2019.003> accessed 12 July 2023.

[29] Martin Ebers, 'Liability for Artificial Intelligence and EU Consumer', (2021) 12 Journal of Intellectual Property, Information Technology and Electronic Commerce Law 204 <https://www.jipitec.eu/issues/jipitec-12-2-2021/5289/ebers_pdf.pdf> accessed 12 July 2023.

[30] Adrien Bibal, Michael Lognoul, Alexandre de Streel, Benoît Frénay, 'Legal requirements on explainability in machine learning' (2020) 29 (2) Artificial Intelligence and Law 149, 169.

Would resorting to the subjective element of the fault of the producer/supplier of the AI (at the stage of designing/ manufacturing / distributing the defective AI) sabotage the effectiveness of the non-contractual civil liability remedies placed at the disposal of the injured consumers? Would it suffice, as a premise for the engagement of specific liability, to consecrate the objective element, of the existence of the manufacturing/ conception defect (such as design flaws, cybersecurity flaws, transparency flaws etc.) of the AI system suspected to be at the origin of the bodily/property damage that was subject to B2C litigation? The positioning of the liability of producers, suppliers, or importers of AI as types of extra-contractual liability was not doubled, in the text of the AILD Proposal, by eliminating the subjective element of the fault of the (potentially) responsible persons; therefore, using the term 'culpable' behaviour (of the AI producer/supplier) would not be inappropriate to describe the specific extra-contractual engaging of AI producers' liability. Congruently, it has been emphasised in the Preamble of the AILD Proposal that the new set of harmonised rules on liability for defective AI is meant to complement the existing liability (detached from fault) system applicable to producers and importers for damages caused to consumers, or for certain types of biological (physiological) or psychological[31] damages, as well as for patrimonial damages caused to consumers; the AILD Proposal would cover, at the antipode, the issues of compensating damages caused to various types of victims, although not falling under the definition of 'consumer'. These discrepancies in terms of personal sphere of incidence, as well as in terms of objective versus subjective criteria of liable conduct might be seen as complementary, thus offering a plethora of specific mechanisms in view of compensating the damages caused by defective AI.

Recital (3) of the Preamble of the AILD Proposal notes that the reference to the provider's fault (proven by the complaining consumer, whose interests were harmed by defective AI systems) in the light of the requirements applicable in the matter of the administration of evidence in order to engage in retaining the 'classic' civil liability in national legal systems would be a pernicious solution in the context of the permanently amplified technical complexity of AI systems. Thus, one solution would be to abandon the 'traditional' vision of producers' liability that enshrines the *actor incumbit probatio* principle; more pertinently, and strictly for the field of administration of evidence regarding AI defectiveness, it is crucial to consecrate the conveying, in a first stage, of a relative presumption of connectivity (or of the existence of the causal nexus) between the damage caused to the plaintiff consumer and the action/omission of the expected AI responses, the latter not being exempted from the requirements of proving bodily/property damages. Obviously, the current national rules on liability, especially

---

[31] Chiara Gallese, 'Legal Issues of the Use of Chatbot Apps for Mental Health Support' in *Highlights in Practical Applications of Agents, Multi-Agent Systems, and Complex Systems Simulation. The PAAMS Collection* (Springer Cham 2022) 258, 267.

those based on fault, are not adequate for the settlement of actions in liability for damages related to defective AI conduct.

The substantial changes brought to the 'classic' regime of civil liability of producers/importers or suppliers in the retailer category (as a backup solution), in order to adapt its rules to the legal complexity presented by situations where the damage is generated by the defectiveness of the autonomous/incorporated AI can be listed as follows:

(a) the identification of AI product categories using as benchmarks the degree of risks presented for the bodily or physiological safety of consumers or for the respect of their fundamental rights[32] (being targeted: respect for human dignity, combating discrimination[33] in the case of automated decisions which involve the total or partial algorithmising of the decision-making process, or respecting the right to privacy)[34];

(b) imposing the obligation, for AI in the 'high risk' category, to proceed to conducting an *ex-ante* conformity assessment procedure (which remains optional in the case of the other AI categories);

(c) reporting by the courts, in the event of litigation, of the relative legal presumption of causality between the defendant's culpable conduct and the harmful result by reference to the consumer's justified assumptions or the inability of the AI system to generate results on the adequacy of which the consumer could have reasonably relied, as emphasised in Art. 4 of the AILD Proposal; nevertheless, maintaining the possibility for the responsible defendants to bring evidence to the contrary, of the excusable character[35] of the behaviour that is the subject of the litigation, respectively to bring proof of the non-existence of fault for the imputed actions/omissions, referring to the prejudicial dysfunction of the AI systems, remains a major trait of the harmonised liability regime;

(d) establishing the possibility, for the courts, to refer in the assessment of the illegal act of the responsible persons, to the existence/non-existence and to the degree of adequacy of the measures employed by the responsible person (the AI producer or supplier) within the management system of the risks and the results achieved[36], where

---

[32] Mark Fenwick, Paulius Jurcys, 'From Cyborgs to Quantified Selves: Augmenting Privacy Rights with User-Centric Technology and Design' (2022) 13 (1) Journal of Intellectual Property, Information Technology and Electronic Commerce Law <https://www.jipitec.eu/issues/jipitec-13-1-2022/5512> accessed 12 July 2023.

[33] Florian Teleaba, Sorin Popescu, Marieta Olaru and Diana Pitic, 'Riscurile bias-urilor observabile și neobservabile în inteligența artificială utilizată în predicția alegerii consumatorului' [Risks of Observable and Unobservable Biases in Artificial Intelligence Predicting Consumer Choice] (2021) 23 (56) Amfiteatru Economic 104, 121 <https://www.amfiteatrueconomic.ro/temp/Abstract_2981.pdf> accessed 12 July 2023.

[34] It has become a truistical observation that the EU AI Liability Act is focused on dividing the product categories into 'high-risk AI', 'medium-risk AI', and 'low-risk AI' (initially, in the last category, were regimented the virtual assistants from the range of chatbots that assist consumers in online contracting).

[35] Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, Madhulika Srikumar, 'Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI' [2020] Berkman Klein Center for Internet & Society <http://nrs.harvard.edu/urn-3:HUL.InstRepos:42160420> accessed 12 July 2023.

[36] Enrico Francesconi, 'The winter, the summer, and the summer dream of artificial intelligence in law' (2020) 30 Artificial Intelligence and Law 147, 161.

the AI system uses techniques that involve training models[37], as mentioned in Art. 10 para 2 to 4 of the Draft Regulation (EU) on Artificial Intelligence;

(e) enshrining the possibility of the courts to assess the existence of an illegal action/omission by the responsible persons by referring to compliance with the transparency obligation incumbent on AI providers, as stated in Art. 13 of the Draft Regulation (EU) on Artificial Intelligence;

(f) consecrating the possibility of the courts to retain the civil liability of AI producers/importers or suppliers for bodily/property damages caused to consumers, and establishing the possibility for the courts to relate, in assessing the relevance of the AI producer/supplier's liability, to the degree to which the AI system was respecting accuracy or pertinency as emphasised in Art. 15 and Art. 16, let. a) of the Draft Regulation (EU) on Artificial Intelligence;

(g) establishing the possibility for the courts to refer to criteria such as the existence, non-existence, or relevance of the enactment, by the persons (potentially) responsible for the damage caused to consumers, of corrective actions, which needed to be taken in order to enhance the AI systems compliance with the conformity prerequisites, based on the provisions of Art. 16, let. g) and of Art. 21 of the Draft Regulation (EU) on AI (Artificial Intelligence Act);

(h) consecrating the possibility for courts, as mentioned in Art. 3 of the AILD Proposal, to impose on the producer, importer, or supplier of AI the disclosure of relevant evidence for settling the claims of the consumer/ claimant, even if the responsible person who would become the debtor of the mentioned procedural obligation would pose as a defendant since the evidence thus provided would facilitate the admission of the plaintiff's claims;

(i) enshrining the possibility for courts to operate with a reversible presumption of deficient AI systems' collapsing, suspected to be at the origin of the damage caused to consumers, while maintaining the possibility for the defendants to bring evidence of the objective performance bars of the AI system whose behaviour is subject to litigation.

3.2 Recourse to the presumption of causality in claims against non-professional AI users

3.2.1 Implications of the risks-benefits test for AI defectiveness

The contemporary physiognomy of specific liability for stand-alone or embedded AI defective products is considerably influenced by the 'confrontation for pre-eminence' waged between the two possible (classical) product defect assessment criteria, namely

---

[37] Luciano Floridi, 'AI as Agency Without Intelligence: On ChatGPT, Large Language Models, and Other Generative Models' (2023) 36 Philosophy and Technology <https://doi.org/10.1007/s13347-023-00621-y> accessed 12 July 2023.

the consumer's 'justified expectations' test versus the ratio of risks and benefits; thus, far from forming a true hierarchy or pre-eminence, each of these criteria provides a different role for national courts: reporting to an abstract standard while addressing the experiences of the average consumer or, conversely, taking into account typical product safety expectations and, on the other part, the engagement in the administration of the evidence from which the apparently (unreasonably) exorbitant weighting of the risks emerged in relation to the benefits of the consumption/use of the AI product. At one pole, the 'reasonable consumer expectations' test, according to which manufacturers/suppliers are liable if they have generated an embedded or autonomous AI product that is placed on the market in a defective and dangerous condition, exploits to a greater extent the criteria than would have been considered, through the lens of an ordinary degree of risk, by the average consumer who purchased the respective product; the mentioned criterion implies an assessment through the prism of the usual knowledge or the possession of common knowledge by the 'average consumer', with regard to the characteristics of the respective AI product, which implies engaging into an objective responsibility (as a form of strict liability[38], detached from the subjective element of fault / gross negligence) of manufacturers for AI products that present safety risks beyond the limit of reasonableness or beyond the safety threshold expected by consumers.

At the opposite pole lies the 'risks-benefits' test, under which the defendant's conduct is assessed by reference to the latter's capacity to have moderated or avoided harmful results by adopting reasonable alternatives (in the case of design defects when the product was designed based on technical parameters, according to which the benefits of consumption are exorbitantly exceeded by the security risks brought by the consumption of that product); the latter criterion uses the conceptual framework or the dogmatic paradigm of negligence imputable to the manufacturer, who accepted the initiation of the production process based on paradigms designed in an excessively risky manner referring to the benefits or the target utility (the so-called 'deficient AI design'). In practice, for products with digital content, for example, the courts might tend to oscillate between the two criteria, and to either intermittently adopt one of these criteria, or use both tests consecutively.

On another side of the discussion, important questions might be raised related to the findings by national courts of potential (yet non-manifested) defects of embedded or autonomous AI products introduced on the market, as a basis for engaging the specific liability of the producers/ the importers; these aspects fuel a constant aporia, targeting the emblematic issues of the specific liability of AI producers, designers or importers for potential AI defects whose manifestation varies on different time bars.

---

[38] Mark Geistfeld, 'Strict Products Liability 2.0: The Triumph of Judicial Reasoning over Mainstream Tort Theory' (2021) 14 (2) Journal of Tort Law 403, 467.

As mentioned in the previous sections, the notion of 'high-risk AI' has been circulated in the recent drafts of normative acts at the Union level to describe a risk that, following an *ex-ante* assessment of compliance[39], can be assessed as having the potential to intervene, causing physiological or patrimonial detrimental effects, or to be blamed for negatively impacting fundamental rights[40], when causing damage that requires judicial (remedial) solutions. Also, when assessing AI safety deficiencies, unfitness for consumption, or defectiveness of these products, it may become difficult or pernicious and inappropriate to develop synchronised standards for eclectic types of algorithms, as they require adaptation to the inherent risk valences of different types of AI products. It should be noted that, as follows from the provisions of Art. 9 para 2 of the AI Act Proposal, for products that involve autonomous/embedded AI with a high degree of potential risks for the bodily safety of users, the AI producer/supplier would be expected to implement a risk management system in the sense of continuous monitoring/intervention (as a permanent iterative process) for the entire usage time bars/exploitation cycle of the respective products.

It must also be observed that the 'fidelity' to the objective nature of specific liability for defective products has been kept intact in the text of Proposal 2021/0106 (Artificial Intelligence Act); in order to engage the producers' civil liability, it would suffice that an illegal action or omission was established (for example, the launching into circulation of an embedded AI or autonomous AI presenting design deficiencies), connected, or, at least, placed in a sufficiently characterised connection with the damage caused to the consumer, without being necessary that this action/omission was imputed to the AI producer. Thus, the plaintiff would not be expected to bring proof that the illegal action/omission was intentionally directed towards causing damageable results, since it is not necessary to prove, within the scope of the specific non-contractual liability, the gross fault, or the inexcusable negligence of the AI producer / AI importer.

3.2.2 AI design flaws, AI manufacturing flaws, AI implementing flaws and relevant taxonomies

It should be emphasised that not only strictly manufacturing defects are included in the substantial scope of application of the specific type of civil liability, which also includes design defects (so-called AI design defects and AI design inadequacy), respectively delivery deficiencies/omissions, incorrect and incomplete information offered to

---

[39] Gianclaudio Malgieri and Frank Pasquale, 'Licensing High-Risk AI: Towards Ex Ante Justification of a Disruptive Technology' [2023] SSNR Electronic Journal <https://ssrn.com/abstract=4346120> accessed 12 July 2023.
[40] Sara Gerke and Delaram Rezaeikhonakdar, 'Privacy Aspects of Direct-to-Consumer Artificial Intelligence / Machine Learning Health Apps' (2022) 6 Intelligence-Based Medicine <https://doi.org/10.1016/j.ibmed.2022.100061>accessed 12 July 2023.

consumers[41] (on technical characteristics and parameters, limited performance of the autonomous or embedded AI product, etc.), since all these types of deficiencies may per se represent a source of producers' liability[42]. The omission of adequate information regarding the foreseeable risks[43] is considered per se a defect of the product for which the civil liability of the manufacturer/importer may be engaged under EU regulations, particularly for physical/property damage caused to consumers through interaction with the AI defective product. Thus, the typical conceptual pillars include the malfunctioning of the AI systems that triggered the loss or compromise of the integrity of the consumer's data, which allows the consumer to recover, by way of a court action, the costs of data recovery/restoration: (i) the *stricto sensu* defectiveness, or unintended syncopation in the manufacturing process; in using these assumptions, the courts analyse the liability prerequisites by comparing the characteristics of the resulting AI product with the admissible safety standards, in an attempt to identify possible inadvertences of the resulting product; (ii) the design defects, when from the 'risks-benefits' ratio applied in these cases, it follows that, in hypotheses where the risks considerably exceed the benefits brought by the use of the AI product, diametrically opposed to the situations described as representing 'manufacturing defects' in a *stricto sensu* perception, the product obtained in the manufacturing process meets the technical parameters that the producer intended to achieve, while the manufacturer accepted the initiation of the production process based on design flaws which were initially assessed as adequately balancing the risks and benefits involved; in the interaction with the final consumer[44], the embedded or autonomous AI might appear excessively perilous for the users' physical/psychological safety; (iii) informational deficiencies regarding the instructions addressing the issues of integrating the AI product into the digital environment controlled by the consumer (where applicable); the informative omissions/erroneous or apparently incomplete information might represent grounds for retaining the producer's / importer's extra-contractual civil liability, in a similar manner to that which is used in the case of the above-mentioned categories of the AI defectiveness.

When assessing the damage caused by AI systems in interaction with consumers, it is worth emphasising the accent placed, in the AILD Proposal, on facilitating the administration of evidence regarding the existence of a connection, on the one hand, between the defectiveness of the AI system and the damage caused to the consumer and, on the other hand, of a causal link between the behaviour attributable to the

---

[41] Monika Namysłowska, Agnieszka Jabłonowska, Katarzyna de Lazari-Radek, 'AI-driven personalisation – a new challenge for Consumer law' in Luigi Lai and Marek Świerczyński (eds), *Legal and technical aspects of artificial intelligence* (Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego 2021) 95, 114.

[42] Olivier Musy and Bertrand Chopard, 'Market for Artificial Intelligence in Health Care and Compensation for Medical Errors' (2023) 75 (C) International Review of Law and Economics.

[43] Monica Navarro-Michel, 'Vehículos automatizados y responsabilidad por producto defectuoso' (2020) 7 (5) Revista de Derecho Civil 175, 223.

[44] Carlos Gómez Ligüerre and Tomás Gabriel García-Micó, 'Responsabilidad por daños causados por la Inteligencia Artificial y otras tecnologías emergentes' (2020) 1 InDret 501.

producer/importer or the supplier of the deficient AI system and the prejudice triggered for the claimant, as an aspect the importance of which becomes more visible in the light of the AI producer's/importer's failure to comply with an obligation of diligence. Encapsulated in the efforts of recognising the inherent procedural challenges in assessing the inadequacy of AI to acceptable safety standards[45] and, on the other side, the harmful results /prejudicial effects of the AI systems that generated relevant damages/harmful effects, the reversible presumption of causality integrated in the text of Art. 4 para 1 of the AILD Proposal would represent per se a significant elevator of consumers' chances in accessing pertinent evidence in damageable AI-related litigious contexts[46].

As mentioned in the previous paragraphs, in order to preserve the prerequisites of establishing the fault of the responsible persons, the defective behaviour of the AI systems might be linked to a possible failure to comply with the 'duty of care' placed on the AI producer/importer or on the retail distributor. It remains essential to note that the said failure to comply with a duty of care is expected to be assessed either in accordance with the provisions of the future Regulation (EU) on AI (i) or in accordance with other rules established at the EU level, namely those sets of nomothetic provisions that regulate the obligation of automatic and continuous monitoring of the behaviour of AI products launched on the market, in interactions with consumers or those that regulate, for example, the operation of unmanned aircraft or public means of transport without a human operator (ii).

3.3 Dichotomic approaches to the administration of evidence in liability cases concerning medium/low risk AI systems

In the background of the discussion, when engaging in the assessing of the elementary premises for maintaining the specific liability of the AI system producer/importer, although the presumption of causation gains in relevance when it can be considered sufficiently probable that the omission or action of the manufacturer/designer decisively influenced the deficient performance of the AI system, the claimant is expected to produce evidence of the existence of the prejudice. Faced, on the other hand, with the need to provide consistent evidence that the result generated by the medium /low-risk AI system has damageable effects, the consumer claimant remains required to bring evidence on the causal relationship between the claimed damage and the alleged malfunction of the system AI, at least in the perimeter of assessing the 'behaviour' of AI systems in the medium /low-risk categories; only in the case of "high risk" AI systems, as stated in the text of Art. 4 para 4 of the AILD Proposal, the courts might resort to

---

[45] Philipp Hacker, 'The European AI liability directives – Critique of a half-hearted approach and lessons for the future' (2023) 51 Computer Law and Security Review <https://doi.org/10.1016/j.clsr.2023.105871> accessed 16 Sept. 2023.
[46] Philipp Hacker, Ralf Krestel, Stefan Grundmann, Felix Naumann, 'Explainable AI under contract and tort law: legal incentives and technical challenges' (2020) 28 Artificial Intelligence and Law 415, 439 <https://doi.org/10.1007/s10506-020-09260-6> accessed 12 July 2023.

establishing an exception from the need for the consumer to establish the elements of causality; nevertheless, the reversible presumption of causation remains significantly functional, except in cases where the professional defendant demonstrates that, due to the absent recourse to the presumption of causation, sufficient evidence and an appropriate level of expertise were reasonably available to the consumer to prove 'on its own merits' the interrelatedness between the AI system's malfunction and the bodily injury/ property damage registered by the plaintiff.

The controversies are fuelled by the fact that the incidence of the reversible presumption is subject to the decision of the court invested in the settlement of the dispute, which may find that it would be excessively difficult for the plaintiff to provide evidence regarding the causation of the damage should no presumed interconnection be established. Evidently, other evidentiary difficulties will be assessed considering the characteristics of AI technologies, in terms of autonomy and opacity, which make explaining the inner workings of the AI system almost impossible from the perspective of the evidence available to the average consumer, the latter being (almost invariably) incapable of proving (without resorting to a reversible presumption) a causal link between the defendant's decision to launch the defective AI product on the market and the damage caused to users in the interaction with the AI system (as specified in Art. 4 para 7 of the AILD Proposal); the AI manufacturer/importer will be able to try to overturn the relative presumption of causation, while also having the possibility to invoke *ab initio* the elimination of the incidental resumption of causation, on the grounds that the circumstances of the case do not require recourse to 'invasive' procedural means.

In terms of the proportionality requirements for ordering the disclosure or preservation (by the defendant AI manufacturer/supplier) of evidence regarding the compliance/ dysfunctionality of "high-risk" AI systems, it is worth noting that the assessment of the plausibility of the claim remains essential within the scope of the analysis undertaken a priori by the courts, when ordering the defendant to disclose relevant evidence regarding the behaviour of suspected "high-risk" AI systems, even if this evidence incriminates the defendant, distancing from the procedural mechanism of *actor incumbit probatio* (given the fact that the lay consumer does not, most often, possess conclusive or relevant evidence regarding the placement of the dysfunctionalities of the AI system at the time of its launching into circulation). The said mission of the national court may come into tension with the requirements to consider the legitimate interests/fundamental rights[47] and confidential information (particularly information related to general security when interacting with AI systems). Recurring cyclically in the perimeter of the assessment of proportionality, the issue of respecting the legitimate interests of third parties is (jointly) related to the provision of adequate procedural guarantees against which the defendant,

---

[47] Bartłomiej Oręziak, 'Artificial intelligence and human rights' in Luigi Lai and Marek Świerczyński (eds), *Legal and technical aspects of artificial intelligence* (Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego 2021) 217, 230.

that was ordered to disclose or preserve the evidence regarding AI deficiencies, may challenge, under procedural terms, the legitimacy of evidence disclosure measures ordered by the court; the central question of the plausibility of the request encapsulates the need, for the national courts, to establish the existence of sufficient reasons for ordering evidence disclosure regarding the conduct of "high-risk" AI systems suspected of causing harmful effects.

In an aporia under which there is retaliatory bias regarding the delimitation of the sphere of responsible persons, it is important to note that the manufacturer/importer of products in the autonomous/incorporated AI categories will, in principle, be considered responsible for prejudicial reverberations, the solution mainly exploiting conceptual pillars such as: (i) the manufacturer is expected to exercise control over the selecting of technical procedures, compatibility with cyber security requirements in the design of the AI systems; (ii) when complying with the 'duty of care', the AI manufacturer/importer is expected to assume the risks resulting from the performance of these activities, to the extent that the manufacturer/designer and/or importer of the product (in the autonomous/incorporated defective AI categories) cannot resort to liability-exonerating contractual terms. On the other side, as it follows from the 'duty of care' obligation and responsibility related to the use of artificial intelligence technologies, liability effects are expected to fall on the actors most able to control certain specific risks; additionally, important attention must be paid to the fact that AI system providers within the EU's territorial settings are subject to the EU' regulatory framework, regardless of where the factual AI operator was based.

In particular, excessively risky AI design would represent legal grounds (in 'risks-benefits test failure' scenarios) for the specific liability of the manufacturer/designer/importer/supplier of a "high-risk AI system" for physical harm/ property harm caused to consumers, in each case where: (a) the harm was related to the use of an AI system that involved training models that would not meet the inherent safety criteria; (b) the courts would have consistent reasons to retain justified criticism regarding the manner under which the AI system was inconsistent with the transparency requirements provided for in the text of Art. 13 of the drafted (EU) AI Act; (c) the court would recognise the existence of some deficient characteristics of the AI system designed and developed by the responsible persons through the lens of the typical use for which the AI product was intended or through the lens of the typical purpose found in the absence of an adequate level of accuracy, robustness and cyber security as mentioned in Art. 15 and 16 of the Draft Regulation (EU) on artificial intelligence; (d) the manufacturer/importer did not immediately take relevant corrective actions, as applicable.

Obliquely mirrored in the structural requirements of liability, as pointed out, due to the complexity of the technologies involved, the element that emerges from the wording of the mentioned texts refers to the fact that it remains a challenge for the injured

consumer to correctly and pertinently identify the responsible 'AI producer', particularly in cases where the provider that 'trained' an AI application may be distinct from the original manufacturer and may have made a safety error that caused the harmful outcome in terms of bodily injury, biological harm, or patrimonial losses. B2C service providers using algorithmic systems[48] or other AI devices may face civil liability for negligent behaviour, especially when the latter neglect safety measures or instructions recommended by the manufacturer of the AI product[49].

The subrogation in the compensatory rights and the adjustment of the personal sphere of incidence of the extra-contractual civil liability for AI conduct would also represent significant progress, since recognising the quality of the "plaintiff" would not only be possible for the individual consumer who submits a claim for compensation, when a bodily/property damage was generated as a result of the action/omission of an AI system, but also for the person who was subrogated in the victim's compensatory rights or acting based on a mandate of joint representation, against the manufacturer / importer of the AI system or its authorised representative, as resulting from Art. 2 pt. (6) of the Proposal for Regulation (EU) on artificial intelligence. Similarly, in cases where the manufacturing defect of the AI product coexisted with elements related to the culpable act, unjustifiable omission, or culpable conduct[50] of a third party, and the manufacturer was held responsible for financial coverage of the damage caused to the consumer, the AI manufacturer/importer subsequently may pose as a plaintiff in a regress action against other actors involved upstream in the AI systems design/production/distribution chain.

Deciphered as 'central' or at least decisively prevalent in the liability regime of manufacturers/designers/importers of defective AI, the ordering (by the court) of the obligation (for the defendant) to make available to the court the evidence from which it would result in the involvement of the AI system in generating the damage invoked by the complaining consumer remains dependent on the requirements of proportionality; the granting of an order against the manufacturer to allow access to the relevant incriminatory evidence remains dependent on the outcome of the 'proportionality test', particularly in contexts where the 'disclosure order' was obtained against relevant third-party AI providers who are procedurally intervening in the dispute between the aggrieved consumer and the AI manufacturer / AI importer. When discussing the presence of recurring or, on the contrary, innovative aspects of AI-related specific liability, it becomes crucial to highlight how the involvement of autonomous/embedded AI systems can affect

---

[48] Ugo Pagallo, Marcelo Corrales, Mark Fenwick, Nikolaus Forgó, 'The Rise of Robotics & AI: Technological Advances & Normative Dilemmas' in Marcelo Corrales, Mark Fenwick, Nikolaus Forgó (eds), *Robotics, AI and the Future of Law* (Springer 2018).

[49] Paulo Henrique Padovan, Clarice Marinho Martins, Chris Reed, 'Black is the new orange: how to determine AI liability' (2023) 31 Artificial Intelligence and Law 133, 167 <https://doi.org/10.1007/s10506-022-09308-9> accessed 12 July 2023.

[50] Carsten Orwat, Jascha Bareis, Anja Folberth, Jutta Jahnel, Christian Wadephul, 'Normative Challenges of Risk Regulation of Artificial Intelligence and Automated Decision-Making' [2022] KIT Scientific Working Papers 197 <http://dx.doi.org/10.2139/ssrn.4274828> accessed 12 July 2023.

the assessment of the classical terms[51] of liability for damages caused to users, such as be the requirements relative to the causal relationship and the fault of the defendant professional; the premises of establishing liability for damages caused by defective AI products were strongly influenced by its dissociation from the element of fault or the concept of 'culpable actions/omissions' of the responsible persons; however, proof of a manufacturing defect related to the behaviour of the AI product in interaction with users remains necessary, conditional on establishing manufacturer/importer liability. In a conjugate way, the pair of relative presumptions, namely the presumption of causation between the damage caused to the consumer and the culpable behaviour of the AI manufacturer / designer, respectively the presumption of causation between the recorded damage and the defect manifested in the operating of the AI system represents a major step towards facilitating the administration of evidence or incriminating elements.

It is also remarkable that, in the AILD Proposal, no detailed attention was paid to the problem arising from establishing whether, for the assessment of AI product failure, the 'risks-benefits' ratio could present procedural pertinency or, on the contrary, adequate results could be obtained by using the test of consumers' legitimate (reasonable) expectations. The dilemmatic atmosphere of this regime of legal liability is completed by the fact that the pivotal element remains the condition of the existence of a manufacturing/design defect of the AI systems; obviously, the regime of extra-contractual liability will require a nuanced approach to the typology of covered defects, which includes their design deficiencies, regardless of whether a manufacturing defect, understood as syncope in the production process, was involved in causing the damage to the final consumer or, on the contrary, a design error or overestimation of the cyber security offered by the implementation of the respective AI system was involved, including informational deficiencies seen through the lens of the impossibility of recovering the compromised data[52], as a type of autonomous damage caused to consumers.

Another salient aspect is that, as resulting from Art. 4 para 2 of the AILD Proposal, the courts, in the event of a dispute having as its object the engagement of extra-contractual civil liability of the manufacturer, supplier or importer of the AI systems, would be able to assess the pertinency of the measures taken by the responsible persons within the applicability of internal protocols, and the results of these measures in the sense of 'tempering' the harmful effects for consumers[53]. Relevant for enumerating, in the text of Art. 4 para 3 of the AILD Proposal, the aggravating circumstances that could form the argumentation in front of the court invested with the settlement of the dispute that

---

[51] Jaap Hage, 'Theoretical foundations for the responsibility of autonomous agents' (2017) 25 Artificial Intelligence and Law 255, 271 <https://doi.org/10.1007/s10506-017-9208-7> accessed 12 July 2023.
[52] Leon Yehuda Anidjar, Nizan Geslevich Packin, Argyri Panezi, 'The Matrix of Privacy: Data Infrastructure in the AI-Powered Metaverse' (Faculty of Law Blogs, University of Oxford, 5 April 2023) <https://blogs.law.ox.ac.uk/oblb/blog-post/2023/04/matrix-privacy-data-infrastructure-ai-powered-metaverse> accessed 12 July 2023.
[53] Stefan Heiss, 'Artificial Intelligence Meets European Union Law' (2021) 10 (6) Journal of European Consumer and Market Law 252, 256.

justifies the retaining of extra-contractual civil liability for the defectiveness of the AI system, remains the claimant's capacity to reasonably access sufficient evidence and expertise to establish the existence of AI harmful conduct. From the taxonomy mentioned in the text of the normative act, it follows that this relative presumption of causality, although it may turn into a procedural reflex of the courts, would only be used in cases in which, when facing the opacity and accentuated complexity of the AI systems whose defectiveness was addressed, the consumer claimant would be deprived (due to the circumstances) of the real and effective possibility of procuring conclusive evidence for establishing the causal relationship connecting the defendant's culpable behaviour and the undesirable result produced by the AI system, thus generating a procedural vulnerability that the court can mitigate by resorting to the reversible presumption of causation.

It remains important to specify that the reversible presumption of causality referred to in the text of Art. 4 para 1 of the AILD Proposal would be plainly inapplicable in civil liability actions referring to AI systems in the medium risk/low-risk categories, and that the discussed remedial measures would become incidental only to the extent that the courts would consider it to present procedural pertinency, since, as stated *expressis verbis* in the text of Art. 4 para 5 of the AILD Proposal, in cases of requests for compensation related to the prejudicial conducts of medium-risk AI systems, the rebuttable presumption of causality would become applicable only if the national courts assess that it would be considerably pernicious for the plaintiff to attempt to prove the causality link through personal efforts. Concluding, it can be stated that, within the scope of civil liability actions for damages caused to the consumer by interaction with AI systems in the "high risks" category, the reversible presumption of causality referred to in the text of Art. 4 para 1 of the AILD Proposal represents the new procedural main option, the courts being expected to waive its application only to the extent relevant evidence would suffice in the absence of the 'rebuttable presumption' mechanism; on the contrary, in the perimeter of civil liability actions requesting for remedial measures for damages caused to consumers by the interaction with AI systems from the 'medium risks' or 'low risks' categories, the incidence of the reversible presumption of causation remains exceptional, intervening only to the extent that the courts would resort to the discussed presumption in order to strengthen the procedural position of the plaintiff who would not be presented the possibility of accessing conclusive evidence in the sense of establishing the originating of the alleged damage in the actions/omissions of the AI systems producer/importer or supplier.

The assumptions set out in Art. 4 para 6 of the AILD Proposal are related to the situations in which the defendant (the potential person responsible for covering the damage caused to the final consumer) is, in turn, a supplier who exploits the AI system for non-professional purposes, in which case the presumption provided for in para. (1) applies only if the defendant significantly intervened in the operating conditions, yet it

lacked pertinent measures oriented towards damage avoidance. The reversible, non-irrefragable nature of the discussed presumption of causality is expressly enshrined in the text of Art. 4 para 7 of the AILD Proposal, since in actions directed against AI suppliers/producers for the damages caused to consumers in interaction with AI systems regimented in the "high risks" categories (and, even more so, for those in the medium risks/low risks categories), the courts would be expected to allow the defendant to overturn the rebuttable presumption, the latter having the possibility to propose to the court the administration of evidence in this regard.

Innovative, in terms of the prerogatives that the court can dispose of in civil liability actions directed against AI suppliers/producers for the damages caused to the consumer by the interaction with deficient AI systems, the provisions of Art. 3 paras 1-4 of the AILD Proposal would enable national courts, either based on Art. 24 and Art. 28 para 1 of Regulation (EU) on AI, or on the AI user, to compel the defendant (the manufacturer, designer, or supplier of the AI system) to disclose the relevant evidence at the court's disposal regarding a certain system of "high-risk" AI that is suspected of causing harm to the complaining consumer. As resulting from Art. 3 para 1 of the AILD Proposal, it follows that the mentioned possibility, as a prerogative of the courts in managing the administration of conclusive evidence, remains confined to the perimeter described by the principle of proportionality of judicial measures taken during the trial, by referring, at the same time, to the plausibility (in a summary analysis, prima facie, of the relevance of the object of the requests made by the consumer/claimant); thus, proportionality requirements remain decisive for ordering (by the court) the disclosure or preservation (by the defendant AI manufacturer/ supplier) of evidence regarding the compliance/ malfunction of high-risk AI systems. The assessment of the plausibility of the claim for compensation remains essential within the scope of the analysis undertaken *a priori* by the courts[54] that may order the defendant to disclose the relevant evidence regarding the conduct of "high-risk" AI systems, even if this evidence incriminates the AI supplier[55]. The procedural context is characterised by the fact that the consumer would lack access to relevant evidence regarding the malfunctions of the AI system potentially identifiable at the time of its launching into circulation.

Additionally, the principle of moderating consumers' procedural disadvantage would be applied, as a distinct obligation incumbent on the defendant in such situations, since from the provisions of Art. 3 para 2 of the AILD Proposal, it follows that, when assessing the pertinency of a request for compensation, the national courts would resort to ordering evidence disclosure should the plaintiff have made all reasonable attempts to collect the

---

[54] Stanley Greenstein, 'Preserving the rule of law in the era of artificial intelligence (AI)' (2022) 30 Artificial Intelligence and Law 291, 323 <https://doi.org/10.1007/s10506-021-09294-4> accessed 12 July 2023.

[55] Regine Paul, 'The Politics of Regulating Artificial Intelligence Technologies: A Competition State Perspective' in Regine Paul, Emma Carmel, Jennifer Cobbe (eds), *Handbook on Public Policy and Artificial Intelligence* (Cheltenham: Edward Elgar forthcoming) <https://ssrn.com/abstract=4272867> accessed 12 July 2023.

relevant evidence from the defendants; nevertheless, when the quantitative or qualitative inadequacy of the evidence proposed by the plaintiff is due to the latter's inexcusable passivity, rather than the opacity or complexity of the AI system whose defectiveness is invoked, the defendant may challenge the pertinency of an evidence disclosure order.

Concerning the preventive procedural measures, as referred to in Art. 3 para 3 of the AILD Proposal, it is worth noting that, in order to prevent situations in which the defendant could be tempted by the idea of destroying the relevant evidence connected to the dysfunctionality of the litigious AI system, evidence that indicts the defendant, the court may order the defendant to preserve conclusive data and information for the period of time set by the court, with the prohibition to delete or destroy the data/information the defendant controls; this procedural measure might be ordered autonomously or simultaneously with obliging the defendant to disclose to the court / to 'declassify' relevant information regarding the security deficiencies / non-compliance of the AI systems invoked by the plaintiff.

As in the case of the first reversible presumption discussed (relating to the interconnectedness of the defendant's culpable conduct and the undesirable result produced by the AI system), the second relative presumption regarding the defendant professional's failure to comply with the requirements of the 'duty of care' (irrespectively of ordering the mandatory disclosure of relevant inculpatory evidence), the principle of proportionality of the ordered judicial measure remains applicable. As noted in Art. 3 para 4 of the AILD Proposal, national courts are expected to limit the issuing of a 'disclosure of evidence' order to cases where the 'proportionality test' has been passed by the proposed procedural measures.

As follows from the text of the mentioned normative act, by ordering the defendant IA manufacturer/ supplier to 'declassify' relevant information regarding the security deficiencies / non-compliance of AI systems, the courts are expected to 'calibrate' the mentioned procedural measure according to (i) the desiderata of preserving the commercial interests[56] of third parties involved directly or indirectly in complying with the measure ordered, especially in terms of the need to protect commercial secrecy or the autonomous rights of third parties; (ii) the concurrent recognition of the defendant's procedural prerogatives to challenge the fairness or appropriateness of the judicial measure ordered against the latter.

Similarly, the following elements are relevant for suitably understanding the cited provisions: (i) the reversible nature of the presumption of culpable failure to comply with the requirements of the duty of care implies, for the defendant involved, the possibility of overturning the presumption of inexcusable negligence invoked against the latter, while bringing evidence based on which the court would be able to assess the level of vigilance[57]

---

[56] Mehmet Fatih Burak (n 9).
[57] Maarten Herbosch, 'The Diligent Use of AI Systems: A Risk Worth Taking?' (2022) 11 (1) Journal of European Consumer and Market Law 14, 22.

manifested by the AI producer or AI supplier regarding security incidents; (ii) at the antipode of the presumption of good faith circulated in a litigious context by reference to the common regime of civil responsibility in most of the Member-States, in the perimeter of retaining AI manufacturers'/suppliers' liability in relation with the injured consumer, the courts would be able to operate with a relative presumption of faulty conduct applied to the recalcitrant behaviour of the defendant against whom the court ordered the obligation to divulge or preserve the evidence at the court's disposal pursuant to the issuing of a 'disclosure of evidence' order that the defendant deliberately chose to ignore.

It should be noted that this autonomous category of material damages consisting of the involuntary alteration/ compromise / deletion of data uploaded by / for the consumer does not include the compensations that could be separately requested from the data operators, for damages related to non-compliance to the General Regulation (EU) 2016/679 on data protection or to the ePrivacy Directive, as these rights of the data subjects remain the subject of separate actions without being affected by the possible introduction of liability actions directed towards recovering the consumer's data restoring costs (as expenses included in the category of damages recoverable from the defendant manufacturer/ supplier to whom the shortcomings of the AI systems were attributable, or the dysfunction of which was at the origin of the compromising of consumer's data).

The 'hybrid' nature of civil liability 'adapted' to prejudicial situations involving AI responses/omissions due to design/manufacturing defects is fuelled by the necessity of preserving the subjective element of the fault of the responsible persons, while resorting to a relative presumption of causality applied against the manufacturer/importer or the supplier of the defective AI, at the opposite of the presumption of good faith characteristic of the classic regimes of civil liability (i), seconded by the implications of a relative presumption of causality between the illegal action of launching the defective AI system on the market[58] and the existence of the design/ manufacturing defect and bodily/property damage caused to the consumer[59]. The considerably more 'fluid' approach to the subjective element of the supplier of defective AI's responsibility would depend on the 'classic' liability regimen to absorb procedural mechanisms involving a rebuttable presumption on the existence of imputable behaviour, that would allow the plaintiffs to engage in remedial actions without being placed in the position of facing insurmountable difficulties, particularly in complying to evidentiary requirements regarding the AI's (opaque[60]) behaviour, at the source of bodily/property damages. Thus, it would suffice for the plaintiff to prove the existence of the claimed damage, while

---

[58] Geraint Howells, Christian Twigg-Flesner (n 23).

[59] Christoph Busch, 'Rethinking Product Liability Rules for Online Marketplaces: A Comparative Perspective' (The 49th Research Conference on Communication, Information and Internet Policy, 22-24 September 2021) <https://ssrn.com/abstract=3897602> accessed 12 July 2023.

[60] Henry L. Fraser, Rhyle Simcock, Aaron J. Snoswell, 'AI Opacity and Explainability in Tort Litigation' (2022 ACM Conference on Fairness, Accountability, and Transparency (FAccT 2022), Seoul Republic of Korea, June 21 - 24, 2022) <https://dl.acm.org/doi/10.1145/3531146.3533084> accessed 12 July 2023.

allowing those responsible for designing/manufacturing or distributing the deficient AI the chance to overturn the reversible presumption of causation, starting from the premise that the relevant information[61] regarding the (non)causality nexus is rather exclusively in the latter's possession, than in the possession of the profane consumer.

These conclusions are reinforced by the specification, in the text of Art. 4 para 2 let. (a) of the AILD Proposal, of the possibility, for the courts, to refer to the existence/non-existence and degree of adequacy of the procedural measures in the assessment of the conduct of the responsible persons (AI producer or supplier) within the framework of the AI supplier's internal security-assessment protocols and the results achieved; at the same time, the courts remain free to assess the existence of the illegal action/omission of the responsible persons by referring to compliance with the transparency obligation incumbent on AI providers, as stated in Art. 13 of the Draft Regulation (EU) on AI. Therefore, the requirement of fault-proving is not necessarily enumerated among the conditions for engaging the liability of manufacturers/importers of defective AI products, except when the consumer accessed specific remedies for civil liability (in principle, non-contractual remedies) founded on the subjective element of the responsible person's faulty conduct; the existence of damage the extent of which can be proven by the complaining consumer remains indispensable in the perimeter of the specific types of civil liability for AI defectiveness, with the specification that the biological damages are also retained when transposed into a pecuniary equivalent by the courts.

Considering the fact that, in the hypothesis that the damage caused to the consumer by the defectiveness of the incriminated IA system is bodily damage, its compensation is difficult, in view of its uncertain or even fluctuating contours over time, the evolving nature of the value claim having as its owner[62] the consumer will involve an assessment of the damage at the time of the issuing of the court's decision while taking into account the established pillars in the scope of civil liability[63] (including the missing opportunities whose materialisation would have been foreseeable for the victim of future damage or in the form of diminishing the possibilities of valorising the current existential potential, for the injured consumer[64]). Generically, this type of damage can be divided into three levels, considering, as a rule, that it consists of: (i) the economic components related to the bodily damage[65], pecuniary assessable elements (which include medical expenses and loss

---

[61] Cristina Frattone, 'Reasonable AI and Other Creatures. What Role for AI Standards in Liability Litigation?' (2022) 1 (3) Journal of Law, Market & Innovation <https://www.ojs.unito.it/index.php/JLMI/article/view/7166> accessed 12 July 2023.

[62] Akanksha Bisoyi, 'Ownership, liability, patentability, and creativity issues in artificial intelligence' (2022) 31 (4) Information Security Journal: A Global Perspective 377, 386 <https://doi.org/10.1080/19393555.2022.2060879> accessed 12 July 2023.

[63] David Bomhard and Marieke Merkle, 'Regulation of Artificial Intelligence' (2021) 10 (6) Journal of European Consumer and Market Law 257, 261.

[64] Bartosz Brożek and Marek Jakubiec, 'On the legal responsibility of autonomous machines' (2017) 25 Artificial Intelligence and Law 293, 304 <https://doi.org/10.1007/s10506-017-9207-8> accessed 12 July 2023.

[65] Mark Anthony Camilleri and Ciro Troise, 'Live Support by Chatbots with Artificial Intelligence: A Future Research Agenda' (2023) 17 (1) Service Business 61, 80.

of income from lucrative activities), (ii) non-economic bodily harm (involving recreational harm, generated by the physical and mental suffering of the victim, direct/indirect psychological discomfort, etc.), (iii) patrimonial harm[66] consisting in the loss/alteration of consumer data, as due to deficient AI response/ omission of an adequate response of the AI system involved. Correlatively, it remains to be noted that, regarding the non-patrimonial damages related to the interaction with AI defective products, in a deductive reasoning, the assessment of damages for psychological incidents would be particularly difficult for national courts, when connected to the bodily harm suffered by the complaining consumer[67].

The requirement of the existence of a design / manufacturing defect affecting the functionality or responsiveness of the AI system involved in generating the damage refers to a defect that endangers the safety of the consumer and/or other active components (other than the defective AI product), in an approach that draws the lines of demarcation by reference to the notion of 'design defects' of the AI system. It can be noted that the latter includes, in the perimeter of the specific liability referred to in the drafted (EU) Artificial Intelligence Act, both the intrinsic and the extrinsic defectiveness of the product (particularly for the self-learning autonomous/embedded AI categories[68]), resulting from a syncope in the manufacturing processes, either from the use in the manufacturing process of excessively risky design parameters, by reference to the product's benefits, or from deficiencies in the level of adequate and complete consumer information (non-compliance to the requirements of the obligation of transparency, including the *ex-ante* evaluation of the conformity of AI systems); on the other side, the abnormal dangerousness of the AI product launched into circulation might be assessed based on elements resulting from the fact that it does not meet the safety standards that correspond to the legitimate expectations of the consumer, or by referring to the typical characteristics of the product in autonomous AI / embedded AI category.

---

[66] John Linarelli, 'Artificial Intelligence and Contract Formation: Back to Contract as Bargain?' in Stacy-Ann Elvy and Nancy Kim (eds), *Emerging Issues at the Intersection of Commercial Law and Technology* (Cambridge University Press forthcoming) <https://ssrn.com/abstract=4363410> accessed 12 July 2023.
[67] Chiara Gallese, 'Suggestions for a revision of the European smart robot liability regime' in Paul Griffiths and Caroline Stockman (eds), *Highlights in Practical Applications of Agents, Multi-Agent Systems, and Complex Systems Simulation. The PAAMS Collection* (ACPIL 2022) 29, 35.
[68] Blair Attard-Frost, 'Generative AI Systems: Impacts on Artists & Creators and Related Gaps in the Artificial Intelligence and Data Act' [2023] SSRN Electronic Journal <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4468637> accessed 12 July 2023.

4 Courts' ordering on the disclosure of relevant information on AI flaws and deficiencies

4.1 Proportionality of the judicial measure

The embodiment of the national courts' mission when deciding on the ordering of evidence disclosure may come into tension with the requirements to consider the legitimate interests/fundamental rights of the third parties concerned[69], especially the protection of trade secrets and confidential information. Recurring cyclically within the scope of the proportionality assessment, the issue of assessing the legitimate interests of third parties is (jointly) related to providing adequate procedural safeguards enabling the defendant, against whom an evidence disclosure was emitted (or who was ordered to preserve evidence regarding AI deficiencies), to challenge, procedurally, the legitimacy of measures ordered by the court in the perimeter of administrating relevant evidence.

Deliberate ignoring an evidence disclosure order constitutes legal premises for entailing the mechanism of the presumption of violation of the transparency obligations incumbent on the defendant; as resulting from the provisions of Art. 3 para 5 of the AILD Proposal, in cases in which the defendant culpably ignores a judicial order on "evidence disclosure", a national court is entitled to presume non-compliance by the defendant with a relevant "duty of care", especially in hypotheses similar to those described in Art. 4 paras 2 and 3 of the AILD Proposal, and that the compromised evidence was potentially crucial in deciphering the implications of the awarding of compensation for tortious conduct. The defendant might present a procedural interest in reversing the rebuttable presumption of culpable behaviour. Numerous conceptual and practical elements are relevant to the proper understanding of the functioning of the mentioned procedural mechanism:

(i) the rebuttable/reversible nature of the presumption of ignoring the requirements of the obligation of due diligence implies, for the defendant involved in the AI-design selection, the possibility of overturning the presumption of culpable conduct, while bringing evidence supporting its defences based on vigilant conduct;

(ii) situated at the antipode of using the presumption of good faith conveyed in a litigious context divided by reference to the 'classical' regime of tort liability, and within the scope of employing the (extra-contractual) liability[70] of manufacturers/suppliers of deficient AI systems in relation to the injured consumer's claim for compensation[71], the court will be able to operate with a relative presumption of fault applied to the

---

[69] Stamatis Karnouskos, 'Symbiosis with artificial intelligence via the prism of law, robots, and society' (2021) 30 Artificial Intelligence and Law 93, 115 <https://doi.org/10.1007/s10506-021-09289-1> accessed 12 July 2023.

[70] Fabiana Di Porto, 'Algorithmic disclosure rules', (2023) 31 Artificial Intelligence and Law 13, 51 <https://doi.org/10.1007/s10506-021-09302-7> accessed 12 July 2023.

[71] F Lagioia, A Jabłonowska, R Liepina, K Drazewski, 'AI in Search of Unfairness in Consumer Contracts: The Terms of Service Landscape' (2022) 45 Journal of Consumer Policy 481, 536 <https://doi.org/10.1007/s10603-022-09520-9> accessed 25 October 2023.

recalcitrant behaviour of the defendant against whom the court ordered the obligation to divulge or keep evidence at court's disposal pursuant to Art. 4 paras 1 and 2 of the AILD Proposal; moreover, it might be (reversibly) presumed that the defendant deliberately chose to ignore the content of the obligation established by the judicial order of evidence disclosure;

(iii) it follows, from the text of the mentioned normative act, that, by ordering the AI manufacturer / AI supplier to 'declassify' relevant information regarding security deficiencies / non-compliance of AI systems, the court is expected to 'calibrate' its ordered measures according to imperatives deducted from: (a) the desiderata of preserving the integrity of commercial interests of third parties directly or indirectly involved in compliance with the ordered measure, especially in view of the need to protect the confidential trade information that the respective third parties could invoke either through the lens of the protection of the autonomous procedural rights or in a particular procedural context; (b) concurrent recognition of the defendant's procedural prerogatives to challenge the fair or appropriate nature of the judicial measures on evidence disclosure, issued against the defendant.

Reminiscent from the procedural contradictory requirements arising from the principle of proportionality in assessing informational asymmetries, aspects such as the appropriateness of the judicial measures on evidence disclosure are also saliently important, as in the case of applying the rebuttable presumption regarding the culpable conduct and the undesirable results connected to the launching of defective AI systems; similar conclusions may be extracted in the case of the second relative presumption regarding 'failure' (in bad faith/from inexcusable fault) of the defendant professional's compliance with the requirements of the duty of care, for ordering the mandatory communication by the defendant of the relevant information, as noted in Art. 3 para 4 of the AILD Proposal. Additionally, the principle of moderating plaintiff's own procedural disadvantage will become applicable, as a distinct obligation incumbent on the consumer/claimant in these hypotheses, as resulting from the provisions of Art. 3 para 2 of the AILD Proposal; when the quantitative or qualitative inadequacy of the evidence proposed by the consumer/claimant is due to the latter's own inexcusable passivity (rather than to the opacity or complexity of the AI system whose defectiveness is invoked), the courts are expected to reject the request for issuing an evidence disclosure order against the defendant. From Art. 3 para 1 of the AILD Proposal, it follows that the mentioned possibility, as prerogative of the courts in managing the administration of conclusive evidence, remains confined to the perimeter described by the principle of proportionality of judicial measures taken during the litigious stages of evidence administration, by reference, simultaneously, to the plausibility (in a summary analysis, prima facie, on the relevance of the object of the requests made by the

consumer/claimant) or the pertinency of the plaintiff's claims[72]. The requirement of proportionality remains decisive for ordering the disclosure or preservation of evidence (by the defendant / AI manufacturer / AI provider) regarding compliance/ dysfunctionality of high-risk AI systems. Assessing the plausibility of the request for compensation remains essential within the scope of the analysis undertaken a priori by the courts of judgment that may order that the defendant was expected to disclose the relevant evidence regarding suspected "high-risk AI systems", even if this evidence incriminates the debtor of the obligation, as mentioned in the previous paragraphs while distancing from the exigencies of the *actor incumbit probatio* procedure. Thus, the procedural context is characterised by the fact that the consumer does not, most often, possess conclusive or relevant evidence placing the malfunctions of the AI system at the source of the prejudicial consequences.

Within the scope of addressing the civil liability for prejudicial effects generated by the interaction with defective AI systems included the "high-risk AI" categories, the rebuttable presumption of causality the incidence of which is addressed in the text of Art. 4 para 1 of the AILD Proposal would present pre-eminence, the courts waiving its application to the extent to which the defendant expressly requested the removal of the incidence of the presumption, on the grounds that the relevant evidence would be accessible to the complaining consumer which would render futile the use of the rebuttable presumption; on the contrary, in the perimeter of actions in civil liability for remedying damages caused to the consumer by the interaction with unsecured AI systems in the medium risk/low-risk category, the incidence of the relative presumption of causality is discussed as exceptionally available, intervening only to the extent that the court judges positively assessed its pertinency, meant to strengthen the procedural position of the plaintiff consumer who would be circumstantially granted the possibility of accessing conclusive evidence on the damageable results originating in the AI deficient response/omission to adequately respond. Exempted from the rebuttable presumption of causation between culpability and the undesirable result, upon request of the defending AI manufacturer/ importer / AI provider, remain the situations where, even in the absence of recourse to the substance of the said presumption in favour of the consumer/claimant, the latter would have access to sufficiently consistent expertise establishing the existence of the mentioned causal nexus.

---

[72] Ana Pošćić, Adrijana Martinović, 'Regulatory sandboxes under the Draft EU Artificial Intelligence Act: an opportunity for SMES?' (2022) 9 (2) InterEULawEast: Journal for the International and European Law, economics and market integrations 71, 117.

4.2 Temperament (concerning the disclosure of evidence order) pertaining to the consequences on third-party extra-patrimonial or commercial interests

When debating the preservation of the integrity of third parties' commercial interests the involvement of whom is directly or indirectly resulting in compliance with the ordered measure, the discussions refer to the necessity of protecting the confidential trade information that the respective third parties could resort to, both through the means of protecting the autonomous procedural rights to which the third parties might refer and of the recognition of the defendant's procedural prerogatives to challenge the appropriateness of the judicial measures on evidence disclosure. In order to consider the legitimate interests/fundamental rights[73] of the third parties concerned, especially the protection of trade secrets and confidential information[74], within the scope of the proportionality assessment, the courts would be expected to assess the legitimate interests of third parties while providing adequate procedural safeguards that would enable the defendant to challenge, procedurally, the legitimacy of measures ordered by the court in the perimeter of administrating relevant evidence. Thus, the AI supplier against whom an evidence disclosure order was emitted (or who was ordered to preserve evidence regarding AI deficiencies) might resort to challenging the pertinency of the judicial order, on the grounds of inexcusably ignoring the interests of third parties (including those of pre-existing B2B contractual arrangements).

5 Conceptual and practical interrogations on the pertinent use of presumptions of causality in cases concerning 'self-learning' / 'self-evolving' AI algorithmic categories

For the pertinent understanding of the 'concentric circles' involved in the subject of civil liability retained for AI products from the self-learning category, the evolutionary autonomy of these algorithmic systems remains central to the economy of the discussion regarding the initiation of actions against the AI provider. When 'dissipating' the nebulous possibility of invoking 'development risks', as a type of exonerating motif applicable in this matter, the text of Art. 6 para 1 let. (c) of the AILD Proposal retains *expressis verbis* that the manufacturer/importer can be held liable, in situations in which the defect manifested itself in the context of the self-evolving capacity of the AI product, specifying that an AI product might be defective based on its unpredictable self-learning capacities, which might evolve in 'maligned' AI behaviour.

---

[73] Dina Babushkina and Athanasios Votsis, 'Disruption, technology, and the question of (artificial) identity' (2022) 2 AI and Ethics 611, 622 <https://doi.org/10.1007/s43681-021-00110-y> accessed 12 July 2023.
[74] Alexandre Lodie, Stephanie Celis J. and Theodoros Karathanasis, 'Towards a new regime of civil liability for AI systems: comment on the European Commission's Proposals' (AI-regulation.com, 14 October 2022) <https://ai-regulation.com/eu-commission-proposals-on-ai-civil-liability/> accessed 12 July 2023.

On the opposite of the prerequisites that could be retained within the perimeter of 'classic' civil liability for the hidden defects of the goods that were subject to the sales contract, in which case the warranty owed to the buyer would cover only the pre-existing deficiencies at the time of contracting and the knowledge of which was not accessible to the buyer as a result of the examination of the goods with reasonable vigilance, different premises are to be observed in the case of the "extra-contractual liability" of the producers/suppliers of deficient AI systems, especially for those included in the self-learning or self-training category, or for AI systems that involve constant adaptation of responses as a result of 'learning' from their interaction with AI users.

At the same time, it remains crucial to note that the evaluation of the security deficiencies presented by self-training / self-learning AI systems[75] would be conjectural, by reference to the standard of a consumer's justified assumptions on AI product safety; as pertinent to assess the degree of dangerousness assumed by the interaction with an AI system whose responses (actions/omissions) had an evolutionary character (that are partially unpredictable, including for its creators), the consumer's 'legitimate expectations' test would sequentially focus on whether the manufacturer of each of the interconnected systems relinquished control over the AI product or algorithmic system[76], particularly in light of the general obligation on AI system providers to continuously monitor, as an itinerant and permanent obligation, the behaviour of the AI system in the interaction with consumers; the monitoring obligation represents a central pillar in the perimeter of civil liability for bodily/property prejudicial aspects affecting consumers by the malfunctioning of AI systems in the self-learning category.

6 Concluding remarks

The concept of the AI providers' 'culpable behaviour' might be seen as having, conjecturally, a multitude of meanings, in terms of transparency, devotion and fidelity towards the principles of precaution and preservation of consumers' safety, constancy and obedience to reasonable standards of probity. The mentioned semantic arborescence of the 'AI credibility' conceptual framework would find different forms of objectification in the national legal systems while placing rebuttable presumptions of causality in the vanguard of the efforts of adapting the 'classical' liability regime to the provoking aspects of liability for AI systems defectiveness.

We argued that the binary premises (partially subjective, generated by the element of fault and partly objective, focused on the element of the presence of the design/manufacturing defect of the AI system) in engaging the specific liability of the

---

[75] Jarosław Protasiewicz and Marek Michajłowicz, 'Agile development of intelligent systems: a case study' in Luigi Lai and Marek Świerczyński (eds), *Legal and technical aspects of artificial intelligence* (Wydawnictwo Naukowe Uniwersytetu Kardynała Stefana Wyszyńskiego 2021) 297, 310.

[76] Jaromír Šavelka and Kevin D Ashley, 'Legal information retrieval for understanding statutory terms' (2022) 30 Artificial Intelligence and Law 245, 289 <https://doi.org/10.1007/s10506-021-09293-5> accessed 12 July 2023.

pernicious AI systems producers/importers might be contrasted to the versions of the subjective liability 'classical' regimen, especially from the perspective of consecrating a taxonomy of defects covered by the material scope of the liability, compartmentalised into manufacturing defects versus design defects and information deficiencies; congruently, failure to comply with the transparency obligation incumbent on the manufacturer/importer of deficient AI systems would represent *per se* legal ground for retaining the AI supplier's responsibility for covering the damageable effects.

In terms of 'expanding' the material scope of the extra-contractual civil liability of manufacturers, importers or suppliers of defective AI systems, the most important innovative elements consist in the expansion of the scope of legal protection rules in order to include patrimonial damages caused to the consumer by loss or corrupting of consumer's data, in the interaction with the defective AI system, unless the data was used exclusively for professional purposes, without excluding the hypotheses of coverage for the mixed purpose (simultaneously professional and personal) of the compromised data. According to Recital 16 para 1 of Draft Directive COM/2022/495, in principle, data restoration costs must be included in the category of damages recoverable from the defendant (AI manufacturer/supplier) to whom the shortcomings of the AI system are attributable. Apparently, it is not relevant, for the coverage of these categories of damages, where the data was stored, as it does not matter to any extent whether the defendant (the AI manufacturer/supplier) directly controlled the storage procedures for the consumer's lost data. Thus, the (EU) AI Act Proposal seems to engage in a long-run debate outlined around the controversial concept of 'ownership of data' stored by smart technologies or through algorithmic applications. The focus is not placed on the possible recognition of a genuine category of 'property rights' over the compromised data, the damage of which would be the subject of a civil liability action against the designer/producer or supplier of the deficient AI system, but on their inclusion in the category of recoverable material damage, without simultaneously and necessarily engaging in recognising consumer 'ownership' over compromised data, as a prerequisite for damage coverage. Nevertheless, when assessing the pertinency of a request for compensation, the national courts would resort to ordering evidence disclosure should the plaintiff have made all reasonable attempts to collect the relevant evidence. In cases when the quantitative or qualitative inadequacy of the evidence proposed by the plaintiff is due to the latter's inexcusable passivity, rather than the opacity or complexity of the AI system, the defectiveness of which is invoked, the defendant may challenge the pertinency of the evidence disclosure order. Finally, the rebuttable/reversible nature of the presumption of ignoring the requirements of the obligation of due diligence implies, for the defendant, the possibility of overturning the presumption of inexcusable negligence in breaching the 'duty of care' incumbent on AI providers, while bringing evidence illustrating the level of vigilance manifested during the AI design-selecting process.

Maria Teresa Bartalena [*]

REASONABLE ROYALTY DAMAGES ACROSS DIFFERENT COUNTRIES AND THROUGH A LAW AND ECONOMICS LENS

THE IMPACT ON INNOVATION GOVERNANCE

Abstract

In most legal systems, a reasonable royalty represents the minimum compensation for patent infringement litigation, and it is undoubtedly the most used among the remedies at the courts' disposal. Therefore, its calculation is crucial in the intellectual property field and, more specifically, within the function of liability in this context to incentivise investments in R&D and promote innovation.

Nonetheless, some differences can be found between the methods used to calculate reasonable royalty awards in two largely developed countries, ie in the U.S., where judges are deemed to consider only the parties' *ex ante* information and in Germany, where *ex* post considerations are involved in calculation.

This article intends to analyse how these different approaches can affect investments in the field of technology and innovation. In this process, German and U.S. case law will be primarily investigated, but some references will also be taken from Chinese experience, a legal system for some aspects similar to civil law ones and in which patent infringement disputes are solved by specialised judges, such as in Germany, whose set of remedies, nonetheless, has been recently modified to include a measure that resembles much the treble or punitive damages of the U.S. Code.

JEL CLASSIFICATION: K13, O31, O34.

[*] PhD candidate at Università degli Studi di Padova and teaching assistant in Civil law at Università Cattolica del Sacro Cuore of Milan.

logic of torts overcome… - 5.3 …and the residuality of reasonable royalty awards - 5.4 The market-based approach - 5.5 The Chinese experience: is there a sound economic rationale? - 6 Conclusions

1 Introduction

It is common knowledge that damage awarding in patent litigations is no easy task. This is due to several reasons, one of which is surely that reality, as in most cases, is far more complex than what appears in legislative provisions.[1]

Among the parameters to take into consideration for damages liquidation in patent infringement cases, the "reasonable royalty" is often considered the most accessible one, thus generally the most used by courts.

This is the case in the U.S., where between 2007 and 2016, reasonable-royalty-only have been awarded in 61% of the judicial decisions, almost three times more than lost-profit-only awards;[2] even more in Germany where, dating back to 2005, they have reached a percentage of around 95%.[3]

In China, they are proposed as a (almost) last resort, when plaintiffs are unable to demonstrate a precise amount of damages through the lost profit approach or via restitutionary relief.

These considerations seem to suggest that reasonable royalty calculation is a relatively univocal and easy parameter to calculate damages.

Reality, once again, contradicts these impressions and the aim of this study will be to display the differences in assessing awards via reasonable royalty across two countries

---

[1] This statement is magisterially demonstrated in the example of the *12 Monkeys* movie in Roger D Blair and Thomas F Cotter, 'An Economic Analysis of Damages Rule in Intellectual Property Law' (1998) 39 (5) William & Mary Law Review 1586 at 1591 where it is affirmed that "*the analysis shows that any precise calculation of optimal damages is likely to be next to impossible in the real world*".

[2] *See* Pricewaterhousecoopers, '2017 Patent Litigation Study: Change in the Horizon?' (May 2017, PWC US) <https://www.ipwatchdog.com/wp-content/uploads/2017/05/2017-Patent-Litigation-Study_PwC.pdf> accessed 10 November 2023. Reasonable Royalty and lost profit combined provide for 19% of the cases in the same period. It is also shown that this trend was found in the previous time lapse, between 1997 and 2007, where reasonable-royalty-only were awarded in 60% of the cases, lost-profit-only accounted for a 26%, while a combination of the two for 14%.

[3] Hans Marshall, 'The Enforcement of Patent Rights in Germany' in Christopher Heath and Laurence Petit (eds) *Patent Enforcement Worldwide: a Survey of 15 Countries: Writings in Honour of Dieter Stauder* (2nd edn, Hart Publishing 2005) at 136, for the period antecedent to this year. Also Marcus Schönknecht, 'Determination of Patent Damages in Germany' (2012) 43 (3) International Review of Intellectual Property & Competition Law 309, 332 attests that licence analogy remains the most commonly used methodology for damages award; nowadays, though, the percentage should have decreased, due to the growing relevance of the restitutionary relief, that allows to reach higher damages measures (in this regard, see eg Tobias Wuttke and Florian Henke commenting the Decision of the German Supreme Court dated 26 March 2019, docket no. X ZR 109/16 in 'Patent Infringement in Germany - 10 years' liability for infringer's profits' (Meissner Bolte Milestones 2019) <https://www.meissnerbolte.com/en/meissner-bolte-newsletter/2019/german-patent-infringement-10-years-liability/> accessed 10 November 2023). In addition, the latest amendment of the German Patentgesetz, '*Zweites Gesetz zur Vereinfachung und Modernisierung des Patentrechts*' of 10 August 2021 has introduced §145a, whose *ratio* is to provide a higher level of confidentiality in infringement proceeding, where it is stated that "*All information introduced into the proceedings by the plaintiff and the defendant shall be deemed to be a trade secret within the meaning of […] the Act on the Protection of Trade Secrets*". In this way, the German legislator intends to spur the use of the lost sales methodology to calculate damages in patent disputes; but on the effectiveness of this measure and its relationship with the "Düsseldorf proceeding" *see*, critically, Luc Desaunettes-Barbero and Reto Hilty, Daria Kim, Matthias Lamping, Peter R Slowinski and Hanns Ullrich, 'Position Paper on the Envisaged Reform of the German Patent Act' [2020] Max Planck Institute for Innovation & Competition Research Paper No. 20-05, part II.

where this parameter is mostly used, Germany and the U.S., to try to assess if these can account for the significant divergence of damages measures that can be detected between the two.[4]

A separate paragraph will be dedicated to China, a country where the urgent need to overcome the perceived underdeterrence in property rights protection clashes with the serious obstacles faced by plaintiffs to satisfy the burden of proof required by courts for damages awards in patent litigation.

This analysis will constitute the basis to confront the reality that emerges from case law with the legal and economic theory on the matter and to try to understand the way a reasonable royalty award can better serve the purpose of fostering innovation.

The article, as a consequence, will open with a first paragraph dedicated to the interests that intellectual property rights aim at protecting and how torts fit in this framework. Later, a comparison between, first, the letter of the law on damage in three countries and then, the judicial application will be presented. Lastly, an overview of the law and economics analysis on the matter, to conclude with a paragraph dedicated to the considerations that result from the previous comparisons.

2 Torts and innovation

2.1 Static inefficiencies and dynamic rationale of intellectual property rights

The standing point to comprehend the role of torts in the intellectual property system is to consider the latter as the intersection of public and private interests.

As a matter of fact, knowledge and innovation constitute a public good, because their use is non rivalrous and non-exclusive.[5] In other words, from a static point of view, there are no marginal costs associated with the use of knowledge; therefore, ideally free distribution would be a more efficient solution than restricting its use. Indeed, the incremental character of inventions and of creative work in general[6] means that their value can be properly acquired when made available for society, thus fostering further developments.

---

[4] In the absence of a specific measure of the median average damage awards in Germany, this divergence can be deducted from WIPO, *An overview of patent litigation systems across jurisdictions*, Table "S1. Overview of the main characteristics of major patent litigation systems", where the "Level of damages" of Germany is "Average", while for the U.S. it is "High". Also, it should be added that while a large sample of the literature considered expresses concerns about the risk of overdeterrence in the U.S. patent litigation, this does not seem a problem perceived in the German one.

[5] Joseph E Stiglitz, 'Economic Foundation of Intellectual Property rights' (2008) 57 (8) Duke Law Journal 1693; Richard A Posner, 'Intellectual Property: The Law and Economics Approach' (2005) 19 (2) The Journal of Economic Perspectives 57. A more philosophical analysis of these characteristics has been conducted by Edwin C Hettinger, 'Justifying Intellectual Property' (1989) 18 (1) Philosophy & Public Affairs 31.

[6] Richard A Posner (n 5), where this argument is made not only for patents, but more intensively for creative works whose protection falls under copyright law.

Nonetheless, from a dynamic point of view, failure to provide inventors an adequate remuneration is deemed to cause underinvestment in innovation in the long term, since the fixed costs to create the novelty are indeed large.[7]

As a result, an efficient patent system is deemed to be the one that create a static inefficiency, however justified in the light of a dynamic efficiency, thus aiming at creating incentives for further innovation. Its mechanisms should, therefore, allow knowledge to reach and enrich the public domain as soon as possible, thus exploiting the incremental character of innovation and, at the same time, guarantee the inventor to recoup investments in R&D. In other words, the aim is to maximise the difference between the social value of the invention and its costs.

The static inefficiencies inherent to a patent regime, deriving from the monopoly gains granted to the patentee, and its high costs have led a part of the literature to challenge the role of IP rights in fostering innovation and its impact on the economic growth. As a matter of fact, some have highlighted how similar – or even better results – can be reached by a combination of IP rights with other measures, in which government should play a central role in financing R&D.[8]

2.2. The importance of torts in innovation governance

In this complex scenario, torts play a fundamental role in the enforcement and protection of the exclusive right attributed to the patent holder, as well as the relative monopoly gains. Consequently, they have attracted a growing interest both in the economic literature, for their role in fostering innovation,[9] and in the legal discourse, because the logic behind the patent damage award, in particular for what account of profits is concerned, seems to evade the strictly compensatory logic that pervades the ordinary torts system.[10]

---

[7] Roberto Moro Visconti, 'Funzioni del brevetto e rilevanza della valutazione economica' (2007) 6 Il Diritto Industriale 513.

[8] From the sharper position taken in Michele Boldrin and David Levine, *Against Intellectual Monopoly* (CUP 2008) to the more conciliatory solutions proposed in Joseph E Stiglitz (n 5), in particular, where the author proposes a mixed system, what he calls a "Portfolio approach to innovation", in which the remuneration for the inventor is not solely provided by intellectual property rights, but also by a prize system and innovation is fostered also by government funded research, in order to overcome the static inefficiency deriving from the patent system; *see* Thomas F Cotter, *Comparative Patent Remedies. A legal and Economic Analysis* (Oxford University Press 2013) 26,27, where the author, indeed, recognises that there are viable alternatives to patents, though not as efficient.

[9] *See* G Colangelo, *La proprietà delle idee* (Il Mulino 2015), who highlights how an effective and efficient torts system is decisive to promote innovation; but also David Encaoua, Dominique Guellec, Catalina MartÍnez, 'Patent system for encouraging innovation: Lessons from economic analysis' (2006) 35 (6) Research Policy 1423.

[10] This aspect was particularly discussed in civil law countries, where the compensatory logic is predominant in this sector: in Germany, *see* eg Maximilian Haedicke, 'Die Gewinnhaftung des Patentverletzers' (2005) 7 GRUR 529. In Italy *see, ex multiis*, A Nicolussi, 'Proprietà intellettuale e arricchimento ingiustificato: la restituzione degli utili nell'art. 45 Trips?' (2002) 4 Europa e Diritto Privato 1003, 1036; C Castronovo, 'La violazione della proprietà intellettuale come lesione del potere di disposizione. Dal danno all'arricchimento' (2003) 1 Il Diritto Industriale 7, 16; A Plaia, *Proprietà intellettuale e risarcimento del danno* (Giappichelli 2005); P Pardolesi, 'La retroversione degli utili nel nuovo codice della proprietà industriale' (2005) 1 Il Diritto Industriale 37; P Sirena, 'La restituzione del profitto ingiustificato (nel diritto industriale italiano)' (2006) 3 Rivista Del Diritto Civile 305; M Barbuto, 'Il risarcimento del danno da

The relatively recent interest in this mechanism stems from the fact that literature has primarily focused its attention on optimality in relation to the duration and scope of a patent.[11] These are characteristics that highly influence the effectiveness of a patent, but they are usually not disposable by parties: the first factor is generally fixed by legislators, while the second mainly involves an administrative control.

The result – that somebody considers absurd[12] – is that a standardised measure of these factors is used for different sectors and extremely diversified types of inventions. A solution has shown its pitfalls, especially in recent times, for sequential inventions and in the software industry.

Torts only represent a fragment of the balance and the trade-offs described above, the one that intervenes when a breach has occurred. It is – nonetheless- a system in which the continuous dialogue between private and public interests is clearly expressed.

The role of damages award in these two levels will be analysed in para 5, by providing a brief description of a model that deals mainly with the private incentives involved in the dispute[13] and one that intends to show its importance in attracting and fostering investments in innovation, thus benefiting the social welfare.[14] Both these theories will be analysed to understand whether legal provisions and case law are in line with the result shown in the models and to potentially suggest improvement to align the legal practise with the innovation goals inherent to the patent system.

3 A comparison between three normative provisions

3.1 Introduction: the choice of the U.S., German and Chinese legislative systems

As mentioned above, a comparison between the U.S., German and, partially, Chinese systems will be provided in the following sections. According to recent WIPO reports, these

---

contraffazione di brevetto e la restituzione degli utili' (2007) 1 Il Diritto Industriale 177; A Plaia, 'Allocazione contrattuale del rischio e tutela civile della proprietà intellettuale' (2008) 5 Danno e responsabilità 499; M S Spolidoro, 'Il risarcimento del danno nel codice della proprietà industriale. Appunti sull'art. 125 c.p.i.' (2009) 3 Rivista di diritto industriale 157.

[11] Among the many who dealt with the matter, *see* William D Nordhaus, 'The Optimal Life of a Patent' [1967] Cowles Foundation Discussion Papers No. 474; Richard Gilbert and Carl Saphiro, 'Optimal Patent Length and Breadth' (1990) 21 (1) The RAND Journal of Economics 106; Nancy T Gallini, 'Patent Policy and Costly Imitation' (1992) 23 (1) The RAND Journal of Economics 52.

[12] *See* eg Ted Sichelman, 'Innovation Factors for Reasonable Royalties' (2018) 25 Texas Intellectual Property Law Journal 277.

[13] *See* Roger D Blair and Thomas F Cotter (n 1); Thomas F Cotter (n 8).

[14] Nikolaos Papageorgiadisa and Wolfgang Sofka, 'Patent enforcement across 51 countries – Patent enforcement index' 1998–2017' (2020) 55 Journal of World Business, where the authors explain how enforcement can affect the trust in the patent system of a country; Keith N Hylton and Mengxi Zhang, 'Optimal remedies for patent infringement' (2017) 52 International Review of Law and Economics 44.

three countries present a high number of patented inventions per year and relatively high innovation rate.[15]

For what IP enforcement is concerned, judicial applications in these countries have led to very different results. Some authors claim that damages awarded in U.S. court for patent infringement are extremely high, leading to a risk of overdeterrence;[16] contrarily, the European and the U.S. Congress have often lamented the tangible underdeterrence stemming, among other factors, from the low damages awarded by Chinese courts.[17] Germany is also particularly interesting because of its relatively fast proceedings and the level of its highly trained judges, who contributed to a high general efficiency of the IP enforcement system.

Also, Chinese IP system is relatively recent and takes its cue from both other countries: patent infringement disputes are solved by specialised judges, such as in Germany, but at the same time the set of remedies has been recently modified to include a measure that resembles much the treble damages of the U.S. Code.

The first part of the comparison here proposed will be focused on the analysis of the respective normative texts, to verify if any difference can be detected at this level.

Even though the rules that deal with patent infringement remedies – *ie* U.S. Code, Title 35, §284;[18] Art. 71, of the recently amended Patent Law of the People's Republic of

---

[15] World Intellectual Property Organization (WIPO), *Global Innovation Index 2023: Innovation in the face of uncertainty* (Geneva 2023); for the number of patent filed and registered in each country: <https://www.wipo.int/en/ipfactsandfigures/patents> accessed 10 November 2023.

[16] *See* Roger D Blair and Thomas F Cotter (n 1).

[17] *See* eg, European Commission, 'Report on the Protection and Enforcement of Intellectual Property Rights (IPR) in third countries' SWD(2023) 153 final, <https://policy.trade.ec.europa.eu/news/commission-releases-its-report-intellectual-property-rights-third-countries-2023-05-17_en> accessed 10 November 2023 at 16; Congressional Research Service, 'China-U.S. Trade Issues' (updated 30 July 2018), <https://crsreports.congress.gov/product/pdf/RL/RL33536/156> accessed 10 November 2023 at 39; Congressional Research Service, 'Intellectual Property Violations and China: Legal Remedies' (report No R46532, 2020), and also World Trade Organization (WTO) 'Panel report: China – measures affecting the protection and enforcement of intellectual property rights' (WT/DS362/R 2009) <https://www.worldtradelaw.net/document.php?id=reports/wtopanels/china-iprights(panel).pdf> accessed 25 July 2023.

[18] 35 US Code §284, *Damages*:
"*Upon finding for the claimant the court shall award the claimant damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the infringer, together with interest and costs as fixed by the court.*
When the damages are not found by a jury, the court shall assess them. In either event the court may increase the damages up to three times the amount found or assessed. Increased damages under this paragraph shall not apply to provisional rights under section 154(d).
The court may receive expert testimony as an aid to the determination of damages or of what royalty would be reasonable under the circumstances".

China;[19] and Division 9, Section 139 of the German Patentgesetz[20] – do not dramatically differ from one another, the analogies and differences that emerge can be useful to individuate some of the traits that play a decisive role in damages determination.

3.2 The main analogies: vagueness of the terms and reasonable royalty's place in the remedial systems

For what here is more of interest, the letter of the laws is not particularly detailed when referring to the values, which parties must provide evidence for, and judges are supposed to take into consideration when awarding damages.

Indeed, there is no further explanation of the criteria to be included or else excluded in the liquidation phase, so that judges are left with quite an ample space for interpretation. Incidentally, this is particularly true for the "reasonableness" of the royalty that should constitute a parameter to measure the legality of the judicial damage award. Nowhere in the law a detailed definition can be found, so that it is necessary to search for it between the various courts' decisions and the literature on the matter.

Secondly, even if it is not expressly stated in the German law, the reasonable royalty award constitutes in all three countries a residual measure, the one to be used when no other means conducted to a clear measure of damages.

---

[19] Art. 71, Patent Law of the People's Republic of China:
"*The amount of compensation for patent right infringement shall be determined on the basis of the actual losses suffered by the right holder as a result of the infringement or the profits earned by the infringer as a result of the infringement. Where it is difficult to determine the losses suffered by the right holder or the profits earned by the infringer, the amount shall be reasonably determined by reference to the multiple of the amount of the royalties for the patent license. For intentional infringement of a patent right, if the circumstances are serious, the amount of compensation may be determined at not less than one time and not more than five times the amount determined in accordance with the above-mentioned method.*
Where it is difficult to determine the losses suffered by the right holder, the profits earned by the infringer and the royalties for the patent license, the people's court may determine the amount of compensation, which is not less than RMB 30,000 Yuan and not more than RMB 5,000,000 Yuan, in light of such factors as the type of the patent right, the nature and the circumstances of the infringing act.
The amount of compensation shall also include the reasonable expenses of the right holder paid for putting an end to the infringement [...]".
[20] Division 9, Section 139, Patentgesetz ("Patentgesetz in der Fassung der Bekanntmachung vom 16. Dezember 1980 (BGBl. 1981 I S. 1), das zuletzt durch Artikel 1 des Gesetzes vom 30. August 2021 (BGBl. I S. 4074) geändert worden ist"):
"*Any person who uses a patented invention contrary to sections 9 to 13 may, in the event of the risk of recurrent infringement, be sued by the aggrieved party for cessation and desistance. This right may also be asserted in the event of the risk of a first-time infringement. This right is ruled out if asserting it would, based on the particular circumstances of the individual case and the principle of good faith, lead to disproportionate, unjustified hardship for the injurer or third parties which is not justified by the exclusive right. In such cases, the aggrieved party is to be granted reasonable monetary compensation. The claim for compensation under subsection (2) remains unaffected thereby.*
Any person who performs such act intentionally or negligently is obliged to compensate the aggrieved party for the damage caused. When assessing the compensation, consideration may also be given to the profit which the infringer has obtained by infringing the right. The claim for compensation may also be calculated on the basis of the amount which the infringer would have been required to pay as equitable remuneration if the infringer had obtained permission to use the invention. [...]".

Chinese law is particularly strict on the matter, since it allows the court to assess damages based on a reasonable royalty only when the other two parameters, namely the patentee's lost sales and the infringer's profit are difficult to determine, so that no clear damage figure can be achieved with the other two methods.

As a matter of fact, art. 71 of the Chinese Patent Act provides for a clear hierarchy between the possible means of damages calculation: firstly, as alternative measures, patent losses and infringer's profits; secondly, in the event that the previous ones are difficult to determine, "*the amount shall be reasonably determined by reference to the multiple of the amount of the royalties for the patent license*". Lastly, when even this parameter does not lead to a clear result, statutory damages can be awarded.

Art. 27 of the *Interpretation (III) of the Supreme People's Court on Several Issues Concerning the Application of the Law in the Trial of Disputes over Infringement of Patent Rights*[21] further specifies that "*actual losses*" (the patent holder's lost profits) constitute the primary parameter and only subordinately parties and judges may resort to the infringer's profit measure.

It is likewise notable Chinese provisions allow judges to compensate the patentee with an award not equal to a reasonable royalty, but to a multiple of it. Similarly, Section 35 of the U.S. Code, § 284 states that damages ought to be "*adequate to compensate for the infringement, but in no event less than a reasonable royalty*"; so that the latter is thought to only represent the minimum compensation to be awarded in patent litigation.

It is indeed the lost profits method alone which is usually considered as the one that can effectively put the patentee in the same place he would have been absent the breach, with the implicit consequence that – taken alone – damages in the amount of royalties do not encompass the entire damages suffered by the patentee.[22] It will be interesting to see how this is expressed in judicial application as well as in the law and economics analysis.

3.3 Disgorgement of profits

The first main difference, immediately perceivable, is that U.S. legislation does not include the disgorgement of profits among the remedies available to calculate damages. As a matter of fact, the provision that contemplated the restitutionary relief was eliminated in 1946, due to the time, expenses and complexity that such a measure required.[23]

---

[21] Released on 1 January 2021.

[22] This is especially clear in U.S. case law, *see Rite-Hite Corp v Kelley Co Inc* 56 F. 3d 1538 (Court of Appeals Fed Circ1995); *Del Mar Avionics Inc v Quinton Instrument Co* 836 F. 2d 1320 (Court of Appeals Fed Circ 1987).

[23] *See* Thomas F Cotter (n 8) 95, 96, in the context of a wider and comprehensive analysis of the major reforms in the U.S. patent system. The restitutionary relief remains for design patent cases, where it is expressively contemplated by the law.

This is particularly interesting, since the common law systems are the ones where the disgorgement of profits theory was first elaborated and fully developed,[24] while this remedy is not so popular in civil law systems, where it is seen as quite incompatible with the compensatory nature of damages.[25]

It must be considered, though, that this provision was firstly introduced in civil law countries and in China on the thrust of the TRIPs agreement, an international treaty on intellectual property rights protection, whose art. 45 allowed adherent States to articulate their respective legislation as to include infringer's profit award in their remedy set.[26]

Actually, this measure, as we will see, is almost never used in China and less used than the *Lizenzanalogie* in Germany, even though it is becoming increasingly more popular in the latter. Recent judicial decisions have indeed affirmed the equivalence of account of profit and reasonable royalty award, since neither of them is designed to compensate the patentee for the harm suffered and have given a legal justification to this measure, to be functioned together with the rendering of accounts necessary to substantiate the claim.[27]

3.4 Intentional or negligent conduct: injunction and punitive damages

German law explicitly states that damages may be awarded whether infringement is the result of an intentional or negligent conduct[28] (Section 139, para 2 of the German Patent Act). This can be explained if we think, firstly, about the central role of injunctions

---

[24] Conceived as an equity remedy, mainly for fiduciary contracts in England, it has then found great fortune in the U.S. within the unjust enrichment doctrine. *See* J Edelman, 'Gain-Based Damages: Contract, Tort, Equity and Intellectual Property' (2002) 2 (2) Oxford University Commonwealth Law Journal 291.

[25] *See* Weoud Hondius, Andre Janssen (eds), *Disgorgement of profits: gain-based remedies throughout the World* (Springer 2015).

[26] World Trade Organization, Trade-Related aspects of Intellectual Property Rights, 15 April 1994, Art. 45:

1. The judicial authorities shall have the authority to order the infringer to pay the right holder damages adequate to compensate for the injury the right holder has suffered because of an infringement of that person's intellectual property right by an infringer who knowingly, or with reasonable grounds to know, engaged in infringing activity.

2. The judicial authorities shall also have the authority to order the infringer to pay the right holder expenses, which may include appropriate attorney's fees. In appropriate cases, Members may authorize the judicial authorities to order recovery of profits and/or payment of pre-established damages even where the infringer did not knowingly, or with reasonable grounds to know, engage in infringing activity.

[27] Firstly, the BGH Urteil vom 24.07.2012 X ZR 51/11 – Flaschenträger gave some general principles profits deriving from infringement, that as this stage is largely referred to the free judgement of the interpreter; while with the following BGH Urteil vom 26.03.2019 X ZR 109/16 the German Supreme Court gives a more solid bases to justify and to substantiate the account of profit claim.

[28] Section 139 at para 1, as lastly amended in 2021, now contemplates a hypothesis of damages award for unintentional conducts, in cases where an injunction cannot be granted for the disproportionality exception introduced. For a comment on the reform on the matter *see*, Martin Stierle and Franz Hofmann, 'The Latest Amendment to the German Law on Patent Injunctions: The New Statutory Disproportionality Exception and Third-Party Interests' (2022) 71 (12) GRUR International 1123; Ansgar Ohly and Martin Stierle, 'Unverhältnismäßigkeit, Injunction Gap und Geheimnisschutz im Prozess. Das Zweite Patentrechtsmodernisierungsgesetz im Überblick' (2021) 10 GRUR 1229.

and specific remedies not on only in patent infringement cases,[29] but also in the general liability system of the country.[30]

Indeed, the first paragraph of Section 139 of the German Patentgesetz deals with the order of cessation and desistence that can be issued when a risk of infringement occurs and when the infringement is unintentional, so irrespectively of the subjective status.

Also, the injunction of the first paragraph can be issued at the end of the first phase of patent litigation - the one that precedes the damage award one - in which the infringement is ascertained (the declaratory judgement). Therefore, the damage award takes place after this phase is concluded and the subsequent actions adopted by the court (eg at the end of the declaratory judgement, an order to access records can be issued as well).

In the U.S., the subjective status of the infringer is one of the factors to be taken into consideration for treble damages that is to say to allow the court to increase actual or compensatory damages up to three times.[31]

Something that resembles this last provision has been introduced in China in 2021,[32] to react and remedy to the worldwide accusations of underdeterrence and inefficiencies of the enforcement system. As it will be illustrated below in para 4.4, in order to embitter the sanctions against infringers and in light of the apparent difficulties in Chinese patent litigation to make use of the "traditional" methods, a provision has been introduced, to allow the judge to enhance by one to five times the damages assessed with the traditional methods, in case of intentional infringement and "*if circumstances are serious*" (Art. 71, Patent Law of the Republic of China).

3.5 Statutory damages

The latter also allows that, when the traditional methods are not suitable to satisfy the burden of prove for the losses directly deriving from the patent infringement, the judge can resort to a statutory measure of damages. It is indeed the law that establishes a minimum and a maximum, within which the court can determine the sum to be attributed to the patentee as compensation for the harm suffered. These values have been increased pursuant to the Fourth Amendment of 2021.

---

[29] The absolute primacy of the injunctive relief in IP rights litigation has been highlighted by WIPO, *An international Guide to Patent Case Management for Judges*, 2023 <https://www.wipo.int/patent-judicial-guide/en/full-guide> accessed 10 November 2023.

[30] Where the *Unterlassunganspruch* in §1004 has not only assumed the role of a general provision in the system, to protect primary rights.

[31] *See eg Halo Elecs Inc v Pulse Elecs Inc* 579 US 93 103-04 136 SCt 1923 195 LEd 2d 278 (2016).

[32] With the Fourth Amendment that becomes effective on 1 July 2021.

This is something that is not allowed in neither of two other countries, where courts are supposed to award damages in the amount of a reasonable royalty, even if the plaintiff fails to produce evidence of it.[33]

4 The law in practice

When awarding damages by means of a reasonable royalty, courts are required to conduct an investigation on the amount two rational operators would have agreed upon in a licence negotiation.

If this constitutes a common basis for damages calculation, the circumstances that surround the negotiation and their terms can vary in the judicial applications, with the result that widely different results can be obtain.

Something that clearly emerges from a comparison between the values of damage awards in Germany and the U.S., two great economies that attract patent seekers and holders.[34]

4.1 The German Lizenzanalogie

The main feature of German patent litigation is the unrivalled predictability of the decisions.[35]

Being the parameter that is mostly used by courts, the acquired experience in assessing damages via reasonable royalty and the clarity of the principle set by jurisprudence on the matter represent some of the factors that contribute to this primacy.

Other main characteristics of German patent litigation are: a) specialised courts; b) a bifurcated system, that is to say that IP rights validity and infringement judgements are carried out in separate trials, often coordinated by a suspension of the infringement one, while waiting for the decision on validity of the patent; c) the relatively shortness of infringement proceedings (at least compared to the validity ones); d) a separation between the part of the declaratory judgment, in which the infringement is ascertained, and an order to render accounts can be released and the one where damages are awarded, with a high percentage of settlements after the first decision.[36]

---

[33] In the U.S., this is clearly affirmed in the famous *Apple Inc v Motorola Inc* 757 F 3d 1286 (Court of Appeals Fed Circ 2014).

[34] *See* note 4.

[35] Julia Schönbohm and Bolko Ehlgen, 'Germany' in Trevork Cook (ed), *The Patent Litigation Law Review* (6th edn, Law Business Research Ltd 2022) 67.

[36] Katrin Cremers and Paula Schliessler, 'Patent litigation settlement in Germany: why parties settle during trial' (2015) 40 European Journal of Law and Economics 185; Katrin Cremers, 'Determinants of Patent Litigation in Germany' [2004] ZEW (Centre for European Economic Research) Discussion Paper No. 04-072, <http://dx.doi.org/10.2139/ssrn.604467> accessed 10 November 2023.

These factors and, in someone's opinion,[37] the "patentee-oriented" tendency registered in German patent litigation are some of the reasons that made this country particularly attractive to patent seekers and holder.

As initial approximation, it can be said that in German case law the reasonable royalty (or, better said, the *Lizenzanalogie*) represents the remuneration for the transfer of the invention exclusive use, object of a fictitious agreement between two reasonable parties.[38] Courts, indeed, refer to the objective value of the presumed right of use,[39] when setting what is considered to be the reasonable and customary royalty.

From the analysis of some of the recent decisions that engaged both the district and the federal courts in intellectual property litigation, there are some recurrent features, relevant to the discourse here developed.

The first interesting element to take into consideration is the perimeter of damages that can be awarded. As a matter of fact, it was recently stated[40] that the object of compensation to the patentee ought to equate the economic value of the invention.

This means that, at least in theory, courts should investigate the economic exploitability of the manufact or process and assess damages according to it. However, continues the Hamburg District Court, measuring the economic utility of an invention can be quite demanding, therefore judges use proxies of this value. Amongst these, the *Lizenzanalogie* is considered to be particularly suitable to this end, because it mimics a remuneration which the two parties that are now in dispute, would have agreed upon. The value derived from these setting needs, then, to be adjusted, in order to adapt the figure to the specificities of the present controversy.[41]

In addition to this, courts have determined the moment at which this hypothetical negotiation should be placed, namely when parties would have concluded an agreement for a licence, therefore before the breach has occurred.

However, it is generally accepted that some subsequent factors should be accounted for, namely the extent of the infringement and its duration,[42] so that what judges consider is what the parties who entered in negotiation could have foreseen as future development.

---

[37] *See* Brian P Biddinger, Paul M Schoenhard and Karen I Leyva-Drivin, 'International Patent Strategy: a focus on Germany and China' (*Bloomberg Law*, 21 June 2016) <https://news.bloomberglaw.com/ip-law/international-patent-strategy-a-focus-on-germany-and-china> accessed 10 November 2023, which deals with China and Germany attractiveness for patent application, while specifically on Germany: Martin Stierle, 'Provisional measures and the risk of patent invalidity—Phoenix Contact and the German approach to interlocutory injunctions' (2022) 17 (11) Journal of Intellectual Property Law & Practice 962; Michael C Elmer and Stacy D Lewis, 'Where to Win: Patent-Friendly Court Revealed' [2010] Managing Intellectual Property Magazine.

[38] This is the general definition of the *Lizenzanalogie* parameter for damage awards that can be found in virtually all the decisions in which it is used. *See, ex multiis,* for trademarks litigation LG Hamburg Urteil vom 14.03.2019 327 O 289/17; OLG Stuttgart 2021 Urteil vom 14.01.2021 2 U 34/20 2021.

[39] *See* OLG Stuttgart (n 38), OLG Düsseldorf Urteil vom 13.04.2017 I-2 17/15 and the judicial decision there cited.

[40] LG Hamburg Teilurteil vom 07.05.2020 – 327 O 146/18.

[41] The District Court of Hamburg for this part of the reasoning cites, as relevant precedent, BGH GRUR 2002 801,802 – Abgestustes Getriebe; BGH GRUR 2010 223 marginal no. 13 – Türinnenverstärkung. In addition to these, OLG Karlsruhe Urteil vom 09.12.2020 6 U 103/19 states that judges must take into account all circumstances of the individual case.

[42] This is clearly declared in OLG Karlsruhe (n 41) and OLG Düsseldorf, Urteil vom 133.04.2017 I-2 17/15, both citing BGH GRUR 1992 597, 598.

A step that, according to case law, is necessary guarantee that the infringer is placed in a position that is neither better nor worse with respect to the bona fide licensee,[43] meaning that the advantages of the licensee must be counterbalanced with the respective disadvantages, both deriving from the infringement.

The latter reasoning assuming also a decisive role also in the faculty, grant to judges, to put a "surcharge" to the reasonable and customary royalty, subordinately to the ascertainment of extraordinary circumstances, therefore in cases in which the correlation between advantages and disadvantages for the infringer is unbalanced in his favour.[44]

The main parameter to calculate the royalty rate, both for trademarks and patent rights litigation, is the return on sale that can generally be achieved in the respective industry[45], since it is what normally influence the measure of a licence royalty as well. In this respect, reference should be made not to the actual profit deriving from the infringement, but to the prospective return, the one the parties would expect from the licence agreement.[46]

The normal rate of return of the industry is usually just the starting point to determine the reasonable royalty rate to be applied to the infringer's turnover, since subsequent adjustment can be made by judges to reflect the circumstances of the case at law.

Specifically for patent litigations, factors that influence royalty rate calculation can be: the technical advantages of the inventions compared to the same or similar artifacts; the eventual monopoly position of the IP right holder and the presence of economically reasonable substitutes in the market, which would have allowed the infringer to avoid using the protected manufact.

Also, the amount resulting from a *Lizenzanalogie* can be influenced by the strength of the parties' respective negotiation power.

The measure of the reasonable royalty damages has to be assessed by the court in accordance with § 287 ZPO, taking into account all circumstances of the individual case according to its free conviction. Usually, though, German Courts use as main reference the existing licences, the ones conceded by the IP right holder – if any exist – and/or those usually practised in the respective market.

Even if this may vary across industries, Courts tend to refer to standard intervals, at which the royalty rate is generally fixed.[47] It is, therefore, the plaintiff that asks for a higher rate who is deemed to prove that a sufficient number of licence agreements have been concluded at that amount, even if it does not constitute the customary and reasonable rate of the specific market.[48]

---

[43] OLG Karlsruhe (n 41); OLG Düsseldorf (n 42).
[44] OLG Karlsruhe (n 41).
[45] *See* OLG Stuttgart (n 38) and OLG Karlsruhe (n 41).
[46] OLG Stuttgart (n 38).
[47] The damages award can be a lump-sum or a percentage on the net profits derived from the infringement, the latter more common. For trademark infringement it is set between 1-5% of the infringer's net turnover.
[48] OLG Düsseldorf (n 39) and LG Hamburg (n 38).

This figure is, then adjusted to the specificities of the case, when all the other factors are considered.

Recently, the BGH[49] has excluded as valid licence agreement, to take as a reference in a patent litigation, the ones concluded during a settlement or, equally, when the infringement has already occurred, pending possible legal actions. In this case, the price the parties agreed upon in the contract reflects not only the value of the future right of use, but other factors, that should not be considered when negotiating a license and when setting the appropriate contractual consideration. In particular, the royalty will include an amount corresponding to the waiver to the right to bring legal actions.

Therefore, highlights the BGH the licence agreements that should be taken as references when awarding damages are the ones that are concluded in the market.[50]

The method used by German courts to calculate reasonable royalty mainly relies on objective values arising from a market analysis (such as the market normal rate of return or comparable licenses), with the result of a high predictability of the damages measures.

4.2 U.S. hypothetical and analytical reasonable royalty

Recent literature decries the overdeterrence deriving from damage awards in patent litigation and attributes this effect mainly to the relatively high compensations that come from jury decisions.[51]

U.S. patent litigation system is indeed characterised by the fact that jury may be involved in the damages award proceeding,[52] so that their members are entrusted with the calculations involved. Also, judges called to decide patent infringement controversies are not specialised, an issue that probably is less relevant that it may appear, given the high sophistication and attention to the law and economics literature shown in some decisions.

Other important features to be mentioned are the wide discovery, typical of common law systems, for which parties are deemed to produce the evidence necessary to render

---

[49] BGH Urteil vom 18.06.2020 I ZR 93/19.

[50] *See* also, BGH Urteil vom 13.09.2018 I ZR 187/17.

[51] *See* Wei Hu, Tohru Yoshioka-Kobayashi and Toshiya Watanabe, 'Determinants of patent infringement awards in the US, Japan, and China: A comparative analysis' (2020) 60 World Patent Information <https://doi.org/10.1016/j.wpi.2019.101947> accessed 10 November 2023. The difference between bench and jury damages award is also shown in 2018 Patent litigation study (n 2), with median damages award of $ 1.9 for the first and $ 10.2 for the second in the years 2013-2017. The values of the previous periods' testimony that the gap is narrowing but remains significant. Others point out that, apart from this, it is also the availability of punitive damages that influences the measure of damages in patent litigations.

[52] It is instead excluded from the preliminary injunction judgment, being this an equity court proceeding, reserved to bench.

a decision,[53] as well as the length of the proceedings and their high costs, that some consider decisive in the high rate of settlement that has been registered.[54]

As for reasonable royalty calculation, two separate theories have been elaborated by courts: *a*) the "hypothetical" – which takes into consideration an "arm's length negotiation between a willing-licensor and a willing-licensee",[55] at the moment of the infringement and which is mainly based on the 15th *Georgia-Pacific* criteria;[56] and the "analytical" one, which actually derives from a further elaboration of the hypothetical one,[57] but has then acquired some kind of autonomy. It requires judges to calculate the infringer's expected profits at the time of the breach, and then decrease this result of the medium profit of the respective market.

---

[53] This is particularly relevant in patent litigation since parties may be asked to produce the accounting records and all the documents that might be necessary to assess damages.

[54] Around 95%, according to Branka Vauleta, '25 Patent Litigation Statistics - High-Profile Feuds about Intellectual Property' (*Legaljobs*, 20 May 2023) <https://legaljobs.io/blog/patent-litigation-statistics/> accessed 10 November 2023.

[55] *See* e.g. *TWM Mfg Co Inc* v *Dura Corp* 789 F 2d 895 (Court of Appeals Fed Circ 1986); *Hanson* v *Alpine Valley Ski Area Inc* 718 F2d 1075 (Court of Appeals Fed Circ 1983); *Tektronix Inc v United States* 552 F2d 343 (Ct Cl 1977).

[56] *Georgia-Pacific Co v US Plywood Co* 318 F Supp 1116 (SDNY 1970), modified 446 F2d 295 –(2d Cir 1971), 121-23 135*n*239. The 15 factors that courts can take into consideration in the calculation of the hypothetical reasonable royalty are:

1) The royalties received by the patentee for the licensing of the patent in suit, proving or tending to prove an established royalty.

2) The rates paid the licensee for the use of other patents comparable to the patent in suit.

3) The nature and scope of the license, as exclusive or non-exclusive; or as restricted or non-restricted in terms of territory or with respect to whom the manufactured product may be sold.

4) The licensor's established policy and marketing program to maintain his patent monopoly by not licensing others to use the invention or by granting licenses under special conditions designed to preserve monopoly.

5) The commercial relationship between le licensor and licensee, such as, whether they are competitors in the same territory in the same line of business; or whether they are inventor and promoter.

6) The effect of selling the patented specialty in promoting sales of other products of the licensee; the existing value of the invention to the licensor as a generator or sales of his non-patented items; and the extent of such derivative or convoyed sale.

7) The duration of the patent and the term of the license.

8) The established profitability of the product made under the patent; its commercial success; and its current popularity.

9) The utility and advantages of the patent property over the old modes or devices, if any, that have been used for working out similar results.

10) The nature of the patented invention; the character of the commercial embodiment of it as owned and produced by the licensor; and the benefits to those who have used the invention.

11) The extent to which the infringer has made use of the invention; and any evidence probative of the value of that use.

12) The portion of the profit or the selling price that may be customary in the particular business or in comparable businesses to allow for the use of the invention or analogous inventions.

13) The portion of the realizable profit that should be credited to the invention as distinguished from non-patented elements, the manufacturing process, business risks, or significant features or improvements added by the infringer.

14) The opinion testimony of qualified experts.

15) The amount that a licensor (such as the patentee) and a licensee (such as the infringer) would have agreed upon (at the time the infringement began) if both had been reasonably and voluntarily trying to reach an agreement; that is, the amount at which a prudent licensee-who desired, as a business proposition, to obtain a license to manufacture and sell a particular article embodying the patented invention – would have been willing to pay as a royalty and yet be able to make a reasonable profit and which amount would have been acceptable by a prudent patentee who was willing to grant a license.

[57] Mark Glick and David Mangum, 'The Economics of Reasonable Royalty Damages: The Limited, Proper Role of the So-Called "Analytical Method"' (2015) 49 (1) J. Marshall Law Review.

For a long time, courts and experts also referred to the so-called "rule of thumb", by which it was presumed that a reasonable royalty should correspond to the 25% of the infringer's profits deriving from the breach.[58] Such a rule was found to lack valid economic rational[59] and was finally rejected by the Federal Court,[60] so that no relevant traces can be found nowadays in judicial decisions.

It must be outlined that, irrespectively of the merits that some literature has found in this reasoning,[61] even before *Uniloc USA, Inc.* v. *Microsoft Corp.,*[62] the analytical approach was, in fact, seldom used in reasonable royalty calculations.

Nonetheless, a specific discourse on the matter may be useful, since it involves some market considerations and in order to highlight the differences with the German system.

The first traces of this theory can be found in the same *Georgia-Pacific Co.* v. *U.S. Plywood Co.* case, when revised by the Second Circuit,[63] and in *Tektronix, Inc. v. United* States,[64] but it was only later, in *TWM Mfg. Co., Inc. v. Dura* Corp.[65] that it was first isolated from the hypothetical negotiation scheme and specifically called "analytical approach", so that the greatest part of the literature prefers to trace back the origins of this methodology to the latter decision.[66]

In these first judgments, the courts use the analytical approach in order to correct the measure of the reasonable royalty award resulting from the hypothetical one, while later it acquired autonomy. In both cases, the measure of damages was the result of the subtraction from the infringer's projected net profits (therefore, the ones that the infringer expects to get thanks to the infringing item at the time of the infringement) the normal rate of return of the market. This theory lays on the assumption that the extra rate of return is attributable to the patent, therefore must be allocated to the damaged part. In this way, everything that falls above the standard return on sale is to be awarded to patentee, irrespectively of its actual sources. It is, in fact, possible that part of the profits is attributable to infringer's efficiency, and it is contrary to any efficiency consideration to disgorge this portion to the patentee, who will be unjustly overcompensated.

---

[58] *See* Robert Goldscheider, 'The Classic 25% Rule and the Art of Intellectual Property Licensing' (2011) 10 Duke Law & Technology Review.

[59] Thomas F Cotter (n 8) 122, 123.

[60] *See Uniloc USA Inc v Microsoft Corp* 632 F3d 1292 –(Fed Cir 2011).

[61] BGH (n 50).

[62] *See* note 53.

[63] The Second Circuit, as a matter of fact, revised the royalty rate determined by the Southern District Court of New York, finding that no reasonable licensee would have agreed upon an agreement that would have left him with virtually no profit margin, therefore in appeal the GP considered the normal margin profit of the market and corrected the amount of damages in order to leave the infringer with that return.

[64] *Tektronix Inc v United States* 552 F2d 343 (Ct Cl 1977).

[65] *TWM Mfg Co Inc v Dura Corp* (n 85).

[66] *See* Thomas F Cotter, 'The Economics of Reasonable Royalty Damages: the Limited, Proper Role of the So-called "Analytical Method"' (2015) 49 J. Marshall Law Review: it is in *Polaroid Corp v Eastman Kodak Co* No 76-1634-MA WL 324105 (D Mass Oct 12 1990) that the author finds the first assessment that the analytical approach constituted a different, separate and alternative methodology to the hypothetical one.

In Germany, courts, as well, consider the normal rate of return on sale of the market amongst as factor that influences the most the measure of reasonable royalty. Differently from the U.S., though, this constitutes the basis of the calculation, because it is presumed that the infringer would have been able to make an analogous profit from the breach. In this way, any return deriving from the infringer's efficiency should remain in his sphere, at least in great part. This solution seems to be much more in line with the compensatory logic of damages than the U.S. "analytical approach".

This highly neglected approach, though, has recently called increasing attention from the literature; in particular it has been thought as a solid parameter to put a cap on the excessive reasonable royalty damages awarded by courts.[67]

Another methodology is far more common in U.S. courts and has been subject to different interpretations and applications, thus leading to a large variance of results and strong critiques from the literature.[68]

In addition to *Georgia-Pacific Co.* v. *U.S. Plywood Co.*,[69] many precedents contributed to delineate profile and bounds of the hypothetic negotiation. In this section, we intend to look at some relevant judgements, in order to analyse how courts make use of the principles crystallised in previous decisions.

Prima facie, what emerges from these judgements is, on the one side, the acknowledgement of the difficulties in setting a reasonable royalty award, when it is affirmed that "any reasonable royalty analysis necessarily involves an element of approximation and uncertainty";[70] on the other hand, it is considered to be the duty of the judge to ascertain that the methodology used to determine the reasonable royalty is, in fact, sound.[71] Case law, therefore, mainly concerns the methodology to be used and the factual aspects of the controversy to be valued.

In determining the relevant factors in the reasonable royalty analysis, courts have distinguished it from the lost profit analysis. The latter, as a matter of fact, involves the reproduction of the so-called "but-for" market, that it to say judges are supposed to reconstruct how the market would have developed, absent the infringing product.[72]

On the contrary, the reasonable royalty theory seeks to compensate the patentee for its lost opportunity to obtain a reasonable royalty that the infringer would have been

---

[67] BGH (n 50).

[68] Thomas F Cotter (n 8), at 122-123 provides an overview of the major critiques to this approach, as it has been used by courts; but also, Michael Risch, ' (Un)Reasonable Royalties' (2018) 98 Boston University Law Review 187, at 218.

[69] Some courts underline how these criteria have been distorted in expert testimonies, but more importantly how most of these factors "*simply are not relevant*" in case involving RAND-encumbered patents, *see Ericsson Inc v D-Link Sys* 773 F3d 1201 1226 (Fed. Circ 2014). This is particularly interesting, since it opens to the possibility to see an evolution in the future, as for reasonable royalty awards is concerned.

[70] *Unisplay SA v Am Elec Sign Co* 69 F3d 512 517 (Fed. Circ 1995).

[71] *Trustees of Boston University v Everlight Electronics Co Ltd* (Fed Circ 2018).

[72] *Grain Processing v Am Maize-Products* 185 F3d 1341 (Fed Circ 1999).

willing to pay if it had been barred from infringing.[73] Therefore, the "but-for" analysis is something that courts consider extraneous to the reasonable royalty award calculation.

The latter, instead, ought to be done by courts by means of the hypothetic negotiation approach, whose main feature resides in the *ex ante* perspective,[74] clearly stated in the 15th Georgia-Pacific criteria and further specified in the principle stated in *Lucent Techs, Inc.* v. *Gateway, Inc.,*[75] to which courts nowadays often refer:[76] the hypothetical negotiation approach "*attempts to ascertain the royalty upon which the parties would have agreed had they successfully negotiated an agreement just before infringement began*".

This means that courts, when awarding damages, should mimic the conduct of reasonable licensors and licensees at the moment of the infringement and, more importantly, relying only on the information available to the parties at that moment, therefore irrespectively to the subsequent events, first of all the circumstances of the infringement itself, its extent and duration.

Among the information at the parties' disposal, recent decisions[77] have again recalled *Lucent Techs, Inc.* v. *Gateway, Inc.* where it is stated that "*The hypothetical negotiation also assumes that the asserted patent claims are valid and infringed*", so that, when mimicking the license agreement between the willing parties, the terms of the fictitious negotiation should not be influenced by the probability of the patent to be found invalid or of its use to be considered non-infringing.

Also, another important factor that can influence the reasonable royalty calculation is the presence of non-infringing alternatives, so that the user could have "designed around" the patented item in order to avoid the breach. When a non-infringing alternative exists, the royalty rate is normally reduced, for the value of the invention is limited.[78]

Another aspect that emerges from recent case law concerns the object and limits of the reward via reasonable royalty. What, in fact, should be compensated to the patentee is the incremental value of the invention, which can be determined by considering "*what the infringer, in a hypothetical pre-infringement negotiation under hypothetical conditions, would have anticipated the profit-making potential of use of the patented technology to be, compared to using non-infringing alternatives*".[79] In other words, the reasonable royalty should be calculated by comparing the infringer's expected profit with or without the use of the patented item. This measure should as well constitute the

---

[73] *AstraZeneca AB v Apotex Corp* 782 F3d 1324 1334 (FedCirc 2015) and *Lucent Techs Inc v Gateway Inc* 580 FSupp2d 1016 (SDCal 2008).
[74] The *ex ante* perspective can be found in the "analytical approach" as well and constitutes the main difference with the German approach.
[75] *Lucent Techs Inc v Gateway Inc* (n 73).
[76] *Aptiv Services 5 US Llc v. Microchip Technology Inc* No 19-1537 (Fed Circ 2020).
[77] *See eg Trustee of Boston University v Everlight Electronics Co Ltd et al* (n 71)*, Astrazeneka AB v Apotex Corp* (n 73).
[78] *Grain Processing v Am Maize-Products* (n 72) and *Riles v Shell Exploration & Prod Co* 289 F3d 1302 1312 (Fed Circ 2002), but also in *Atrazeneca AB v Apotex Corp* (n 73).
[79] *AquaShield Inc v Inter Pool Cover Team* 830 FSupp2d 1285 (2011).

maximum of the reasonable royalty award since no rational party would have entered in an agreement if the difference between the two values is negative.

This reasoning has been summarised in *Dowagiac Mfg. Co. v. Minnesota Plow Co.*[80] by affirming that "*As the exclusive right conferred by the patent was property and the infringement was a tortious taking of a part of that property, the normal measure of damages was the value of what was taken*". "*The value that was taken*", is the brocard used in recent case law to express this idea that only the value attributable to the invention should be compensated and that this value is to be calculated by comparing the profit deriving from patented item and the best-next alternative.

These considerations also played an important role in the apportionment theory (and its exception, the entire market rule), a subject of great interest in recent case law, due to the fact that technological advancement most of the time involve only small components and only rarely an entire product or process.

A complete dissertation of this theory will deviate from the aim of this study; therefore, we will focus on the reasonable royalty determination, when the patented item is only part of the entire product.

Case law seems to agree on the fact that the correct method requires the judge not to deduct the value of the single component by subtracting the value of the other components from that of the entire product, but to autonomously individuate and isolate it and from that calculate the reasonable royalty award.[81]

In this way, courts avoid the possible interferences that may derive from factors other than the effective incremental value of the invention.[82]

In reconstructing the hypothetical negotiation of the parties, similarly to what happens in Germany, one of the most reliable methods is to consider a so-called "comparable licence".

This can be done, according to courts,[83] when a prior agreement between the parties has been reached. In the absence of such an agreement, "*courts permit reasonable royalty damages only if the evidence provides sufficient reliable basis to calculate such damages*".[84]

The reliability of this method, therefore, lies on the particular attention U.S. courts have shown in selecting the previous license agreement to be used as a proxy for damage calculation.[85]

---

[80] *Dowagiac Manufacturing Co v Minnesota Moline Plow Co* 235 US 641 648 (1915).

[81] Amongst the many on the matter, *VirnetX Inc v Cisco Systems Inc* 767 F 3d 1308 (Fed Circ 2014) and *Laserdynamics Inc v Quanta Computer Inc* 694 F3d 51 (Fed Circ 2012), where the "smallest salable unit" principle is affirmed.

[82] *Grain Processing* v *Am Maize-Products* (n 72).

[83] *See Fortinet Inc v Fortanix Inc* 3:20-cv-06900 (District Court ND Cal 2022), citing *Marketquest Group Inc v BicCorp* 316 F Supp 3d 1234 1300 (SD Cal 2018).

[84] *Lucent Techs Inc v Gateway Inc* (n 75).

[85] In *Ericsson Inc v D-Link Sys* (n 69), cited in *Mondis Technology Ltd v LG Electronics Inc* 6 F 4th 1379 (Court of Appeals Fed Circ 2021), where the carefulness emerges clearly: "Prior licenses, however, are almost never perfectly analogous to the infringement action" *VirnetX* 767 F3d at 1330. For example, allegedly comparable licenses may cover more patents than are at issue in the action, include cross-licensing terms, cover foreign intellectual property rights, or, as

The "comparability" of a previous licence (not between the parties) is considered as a factual circumstance, to be proven: "[*T*]*here must be a basis in fact to associate the royalty rates used in prior licenses to the particular hypothetical negotiation at issue in the case*" (*Uniloc USA, Inc.* v. *Microsoft Corp.*), where a "*loose or vague comparability between different technologies or licenses does not suffice*", as it was affirmed in *LaserDynamics, Inc.* v. *Quanta Comput., Inc.*[86]

From these remarks we can derive that the main parameters to consider are the affinity of the technology protected with the patent as well as the structure of the license agreement. This is confirmed in some recent case law, where judges consider the particular sector at which the invention and the agreement pertain, as well as the exclusive or non-exclusive nature of the licence.[87]

4.3 The ex post and ex ante perspective: the U.S. "book of wisdom"

The preceding analysis highlighted the main difference between the hypothetical and the German approaches for reasonable royalty calculation, residing in the different time at which the negotiation is placed.

The *ex ante* perspective adopted by U.S. courts is justified in order to minimise the so-called "patent holdup" risk,[88] which can arise when royalties are based not exclusively on the patent value, but partly upon the infringer's sunk costs. These are the costs that must be borne in order to make use of the patent and to include these in the damage award means in most cases to strongly deter efficient users, in a measure that is considered excessive to a large part of the literature on the matter.[89]

On the other hand, *ex post* considerations allow for a more accurate estimation of the patent value, when used as index of the value that parties could have reasonably expected from the economic exploitation of the invention. To this end, courts start from the normal profit margin of the market and adjust it to the peculiarities of the case.

If this is the main difference between the case law approaches used, it must be outlined, however, that U.S. courts have recognised how in some cases it is more appropriate to include *ex post* considerations in the reasoning. This is what is called the

---

here, be calculated as some percentage of the value of a multi-component product. Testimony relying on licenses must account for such distinguishing facts when invoking them to value the patented invention. Recognizing that constraint, however, the fact that a license is not perfectly analogous generally goes to the weight of the evidence, not its admissibility."; but also *see The Coleman Co v Team Worldwide Corp* United States District Court Eastern District of Virginia 2:20-cv-351 (ED Va 2022).

[86] *LaserDynamics Inc v Quanta Comput* Inc 694 F3d 51 67–68 (Fed Circ 2012).

[87] *Lucent Techs Inc v Gateway Inc* (n 75).

[88] On the patent holdup risk, *see* Thomas F Cotter, Eric Hovenkamp and Norman Siebrasse, 'Demystifying Patent Holdup' (2019) 76 Washington & Lee Law Review,1501; Thomas F Cotter, 'Patent Holdup, Patent Remedies, and Antitrust Responses' (2009) 34 (4) Journal of Corporation Law 1151, 1207.

[89] *Dowagiac Manufacturing Co v Minnesota Moline Plow Co* (n 80), but also Thomas F Cotter (n 8).

"book of wisdom" theory, first elaborated in *Sinclair Refining Co.* v. *Jenkins Petroleum Process Co.*[90] and implemented in recent case law with alternate fortune.[91]

The "Book of Wisdom" doctrine allows judges to involve in the reasonable royalty calculation, post-hypothetical negotiation circumstances, with some important limitations. It is, in fact, specified that these post-infringement facts can only be used in order to correct the *ex ante "prophecy"*,[92] shedding light on some factors that were already present at the time of the hypothetical negotiation, because "*it is incorrect to replace* "the hypothetical inquiry into what the parties would have anticipated, looking forward when negotiating, with a backward-looking inquiry into what turned out to have happened"".[93]

This means that, contrarily to what happens in Germany, *ex post* facts can only be included in the reasoning if they were, in some way, already present in the hypothetical negotiation, so that it can be assumed that "*the hypothetical negotiators would have anticipated these later events and valued the technology accordingly during their negotiation*".[94]

It is, still, quite a big difference from what pacifically emerges in Germany case law, where courts are allowed to plainly consider the duration and the extent of the infringement, in order to reach an accurate estimate of the invention value.[95]

Still, it seems rather unlikely that this difference alone can encompass the huge difference in the average median values in damage awards that can be observed between two countries.

4.4 Chinese patent litigation: why reasonable royalty is seldom used

The sub-paragraphs above show the vivacity of the judicial reasoning on intellectual property rights litigation and, in particular, on the remedies for their infringement. This is partly the result of a strong cultural background on the topic and fertile ground for a legal system designed to vehicle innovation and quality control overall.[96]

---

[90] *Sinclair Refining Co v Jenkins Petroleum Process Co* 289 U.S. 689 (1933).

[91] *See* eg *mSIGNIA Incv InAuth Inc* US Dist (United States District Court for the Central District of California October 18 2018, Filed); *Aqua Shield v Inter Pool Cover Team* 774 F3d 766 US App 2 113 USPQ2D (BNA) 1347 (2014) WL 7239738 (United States Court of Appeals for the Federal Circuit December 22 2014, Decided).

[92] *Sinclair Refining Co v Jenkins Petroleum Process Co* (n 90) cited in *mSIGNIA Inc v InAuth Inc* (n 91).

[93] *Opticurrent LLC v Power Integrations Inc* US Dist(2018), citing *Aqua Shield v Inter Pool Cover Team* (n 91) .

[94] *mSIGNIA Inc v InAuth* Inc (n 91).

[95] That the methodology used in Germany allows reaching a more accurate value of the patent, with respect to the next-best alternative *see* Cotter (n 8).

[96] Think about the fact that in the U.S. intellectual property rights are protected in the Constitution, where art. 8, Clause 8 states: [*The Congress shall have power*] "To promote the progress of science and useful arts, by securing for limited times to authors and inventors the exclusive right to their respective writings and discoveries."

China is a different story, better still has a different history, one in which it has been necessary to justify the need of intellectual property rights in its economic system.[97] Also, the first legislation on the matter is relatively recent, since the Patent Law of the People's was promulgated in 1985, therefore after the Cultural Revolution, when Deng Xiaoping made his return to the public scene.

In order to make their market attractive to foreign investments and maintain competitiveness, it was indeed necessary to promptly and quite urgently catch up with the rest of the world on the matter and to, in some way, implant their intellectual property rights system and remedies in the country, mainly through the adhesion to international agreements.[98]

The result is that, even if art. 71 of the Chinese Patent Act contemplates three parameters suggested in art. 13 of TRIPs, which can also be found in the German legislation (therefore, both patentee's and infringer's profits and reasonable royalty), they are cumulatively used in just the 2% of all the patent litigation brought before the civil courts between 2006 and 2017, with reasonable royalty accounting for a 1%, corresponding to 33 cases.[99]

This means, conversely, that statutory damages were used in 98% of the case. What is interesting to notice is that, presumably coupled with statutory damages, injunctions were awarded in 97% of the cases where infringement was ascertained (with a win rate of 77%, it means that an injunction was granted in approx. 75% of the total). Also, from the CIELA analysis it emerges that courts award an average of 33% of the claimed damages, even though it is not indicated the variance of the awards.

Other specific features of the Chinese judicial system to be mentioned are that the procedure is derived from civil law countries' systems and, such as in Germany, courts are highly specialised.

The reason for the almost exclusive use of statutory damages lies in the fact that it is extremely difficult for plaintiffs to prove actual damages "*given the courts' strict proof requirement and the lack of a generous discovery*".[100] In particular, reasonable royalty awards are only admissible when a previous agreement on the patent right at issue is presented as evidence, this being existent and performed. Also, it is not to be neglected the "reasonableness" attribute, which courts generally require this to be substantiated

---

[97] Mark L Wu-Holson, 'A Commentary on China's new Patent and Trademark Laws' (1986) 6 Journal of International Law and Business 86 gives an overview of the genesis of the Chinese patent and trademark legislation and its function in a Marxist economy.

[98] Such as the already mentioned TRIPs, *see* note 26.

[99] CIELA report for 'Civil Infringement', 'Invention', 'All locations', 'All courts', 'All industries' <https://www.ciela.cn/en/> accessed 15 July 2023.

[100] Thomas F Cotter (n 8) 355; similarly *see* also Jingjing Hu, 'Determining Damages for Patent Infringement in China' (2016) 47 (1) International Review of Intellectual Property and Competition Law 5; Xiaowu Li and Don Wang, 'Chinese Patent Law's Statutory Damages Provision: The One Size That Fits None' (2017) 26 Washington International Law Journal 209.

with additional issues, such as the circumstances of the infringement, the comparability of the licence fee, and the analogies in the patent enforcement.

The restrictive interpretation, together with the absence of an extensive discovery, means that courts are unable to use one of the traditional remedies (lost profits, reasonable royalty and infringer's profit), with the result of a much-lamented underdeterrence.[101]

It is interesting to notice that the literature seemed inclined to work on the reasonable royalty theory to increase courts' capability to assess damages, but in the end the choice has been to introduce the possibility to increase the base damages from one to five times.[102] A provision that much resembles the treble and punitive damages that are typical of a common law system, rather than the civil law ones, from which the Chinese legal order took inspiration. It will probably take time to fully appreciate the effect of this provision on the efficiency of the Chinese patent remedies, but it is indeed something to look out for.

4.5 The main findings to be tested in a law and economics perspective

From the previous analysis it can be inferred that similar premises guide judges in three countries since reasonable royalty award is considered residual and just the minimum compensation. This is explainable by considering that the reasonable royalty approach is detached from the measure of the harm suffered by the patentee but is based on a hypothetical negotiation between the latter and the infringer.

As a consequence, U.S. and German case law show that judges rely on market-based approaches to calculate this value and in particular on the return on sale of the patented invention. Moreover, in the U.S. courts seem to require the measure of the reasonable royalty to be calculated by comparing the return on sale of the patented item with the next-best alternative.

The main difference in the U.S. and German approach can be found in the moment at which the negotiation between the two parties is imagined, if *ex ante* or *post* infringement, with all the consequences in terms of the information available to the parties at those two moments.

As far as China is concerned, the difficulties in satisfy the burden of proof by parties lead to the generalised awarding of statutory damages, so that at present it is not possible to find a theoretical framework of reasonable royalty awards in this country. Another relevant feature is that it appears that statutory damages are often coupled with an injunctive order.

---

[101] *See* note 17, but critically Chenguo Zhang and Jin Cao, 'How Fair is Patent Litigation in China? Evidence from the Beijing Courts' (2020) 241 The China Quarterly 247.

[102] *See* Chung-Lu Shen, 'Reasonable Royalties and the Calculation of Patent Damages: Reflections and Recommendations for a fair and Adequate Calculating Basis of Reasonable Royalties in Terms of Harmonization of China-Taiwan Regional Patent Laws' (2013) 12(1) Chicago-Kent Journal of Intellectual Property 156.

5 The law and economics analysis as a tool to investigate the courts' rationale

In this section we will try to verify if both the remedies designed in legislative provisions and the reasonable royalty judicial applications described above fit the law and economics framework.

In doing so, we will consider two fundamental theories that have been elaborated on optimal patent damages[103] and we'll try to derive from these some parameters that will then be used as benchmark for an evaluation of the IP adjudications considered in the previous paragraphs.

Given the potential breadth of the discussion, after having outlined some basic concepts of the law and economics analysis, we will focus on some main points, *in species* the residuality of the remedy under scrutiny and the market-based approach to calculate its value.

In addition to this, a separate sub-paragraph will be dedicated to introducing the injunctive relief in the reasoning, to try to verify whether it is possible to find some sound rationale in the Chinese method described above.

Eventually, in the following paragraph these points will be used in order to draw some conclusions.

5.1 Normative law and economics theories of optimal patent remedies

While the majority of the studies on patent remedies concerns the scope and length of patents[104] or, when dealing with torts, they usually test the economic efficiency of the existing remedies,[105] the theories that will be analysed in this subsection provide for a general model of optimal patent remedies and, therefore, may represent a useful tool to give an evaluation of the reasonable royalty damages of the previous analysis.

The main difference between two theories is their background, since, as will be better explained herein after, one was constructed in the property/liability rule framework,[106] while the most recently elaborated derives from a social welfare objective function.[107]

Amongst the main assumption from which it is necessary to start to comprehend the theory first elaborated by Cotter and Blair,[108] is that property rules – such as injunctions–

---

[103] Roger D Blair and Thomas F Cotter (n 1), and Thomas F Cotter (n 8); Band Keith N Hylton and Mengxi Zhang (n 14).

[104] Starting with the renowned William D Nordhaus (n 11) extensive literature can be found on the matter, *see* also note 11.

[105] *See* eg JJ Anton, DA Yao, 'Finding "Lost" profits: an equilibrium analysis of patent infringement damages' (2007) 23 (1) Journal of Law, Economics and Organization 186.

[106] The foundation of this theory can be found in Guido Calabresi and A Douglas Melamed, 'Property Rules, Liability Rules and Inalienability: One View of the Cathedral' (1972) 85 (6) Harvard Law Review 1089.

[107] *See* Keith N Hylton and Mengxi Zhang (n 14), where the authors explain why their study should be considered as the first to provide a general model for optimal patent remedies.

[108] Roger D Blair and Thomas F Cotter (n 1).

represent a more suitable remedy for patent infringements.[109] The ownership structure of IP rights and the exclusion faculty that they entail are indeed predominant factors, inevitably affecting the effectiveness of remedies. In addition to this, information on the value of a patent can be better assessed by parties (ie the patentee and the willing user) rather than a third one, such as courts.

Therefore, optimal damages award is the one that mimics the incentive structure that derives from a property rule, that is to say the one that leads to negotiations *ex ante* of the exclusive right, while at the same time preserving the innovative incentives for inventors.

In general terms and for what the measure of damages is concerned, it should be considered that patent value is relative. This means that the value of an invention can and should be measured in relation to the next-best alternative. It is, therefore, this surplus, directly attributable to the patent that should be awarded to the patentee via damages.

Given these premises and in the absence of: a) close substitutes to the patented item, b) price discrimination, c) transaction, litigation and information costs, this theory highlights how a reasonable royalty rate would indeed be the result of negotiations, in cases in which the infringer was more efficient than the patentee.

In the latter scenario, as a matter of fact, it is more efficient for the infringer to engage in the production of the patented item, since he is the more efficient of the two economic operators. If we then assume perfect information, the parties will reach an agreement for the use of the IP right by the infringer in exchange for a determined (and reasonable) royalty, rather than resort to litigation.[110]

The conditions sketched above are quite impossible to stand, let alone together, therefore to different conclusions leads the law and economic analysis when introducing such factors. Once these variables are introduced, though, reasonable royalty awards seem insufficient to give the right incentives to parties, so that the lost profit approach (whenever the infringer is more efficient) and the infringer's profit award (in the opposite situation) are considered as the respective optimal responses in infringement cases.

Apart from the strictly economic considerations, one of the reasons why these two remedies are considered preferable to a reasonable royalty award is that information regarding the patentee's lost profits or the infringer's profit are available to the parties, therefore should be more easily accessible.

As a consequence, a reasonable royalty award, in this context, is only justifiable when the less efficient patentee would have preferred to exclude the user from the market via injunction rather than licensing him, but proving the measure of lost profits is concretely difficult, so that reasonable royalty is an "imperfect substitute" for this figure.[111]

---

[109] Roger D Blair and Thomas F Cotter (n 1) and Thomas F Cotter (n 8).
[110] ibid.
[111] ibid.

It is particularly interesting to notice that in the development of this theoretical framework, the authors contemplated the possibility for the judge to award super- but also under-compensatory damages. In particular, when the surplus of the social benefit of the invention over the enforcement costs is very low or even non-existant, full restoration of the patentee is considered not adequate, since it would result in a cost for society.[112]

More recently, an alternative model of optimal patent remedies has been elaborated, one derived from a social welfare objective function, that present itself as an application of the theories concerning "*the trade-off between optimizing innovation incentives and minimizing the static monopolization cost of intellectual property*" to patent damages and a development of the literature on general tort theories.[113.]

Patent infringement, as a matter of fact, presents strong peculiarities, that allow distinguishing this context from general torts. The main feature consists in the fact that an IP violation, on the one side, injures the patentee by decreasing his reward to the investment in innovation, while on the other benefits society, by subjecting the patented innovation to competition.[114]

The model then ties the optimal measure of patent damages to the social value of patented invention and presents its result both in a static scenario and in a dynamic one, where the latter includes the investment in innovation.

From the static scenario the authors derive that "*in the absence of substantial post-patent efficiencies, infringement necessarily enhances the innovation surplus*",[115] therefore social welfare may be optimised by avoiding some infringement litigation. On the contrary, it is indeed the presence of post-patent efficiencies that calls for some level of protection and enforcement.

By adding the investment in innovation and, therefore, shifting to a dynamic scenario, it emerges that optimal patent damages are those that balance the marginal cost of patent protection (ie litigation, taking care to avoid infringement costs) to the "dynamic" marginal social benefit of the patented innovation, where the latter is a formula that takes into consideration both the innovation rate of the invention and its social externalities.

As a consequence, an award greater than the sum of loss of profit (or reasonable royalty) damages and litigation costs is justified when the social value of the patent is sufficiently high for the patentee to internalise the expected social gain from the investment.

Being the social value of the patent the sum of the expected profit of the inventor and the expected consumer surplus from the invention, net of litigation and precautionary

---

[112] Thomas F Cotter (n 8) 47.

[113] Keith N Hylton and Mengxi Zhang (n 14) 46.

[114] ibid 45: "*the infringement not only injures the patentee but benefits society by subjecting the patented innovation to competition*".

[115] ibid 48.

costs, this means that super-compensatory damages should be awarded when the patentee expected profit and the consumer surplus are high and enforcement costs are low, so e.g. when the scope of the patent is clear. The sum of the lost profit (or the reasonable royalty) award and a surplus (the "additur"), which comprehends litigation costs and the entire social value of the innovation, constitutes the upper limit of the damage awards.

It is interesting to notice that *a contrario* we can derive that those inventions that do not carry a powerful innovation charge should have a correspondent protection, with a relatively small (even negative) additur.

Finally, the study highlights how efficiency is increased by a combination of injunctive relief and damages, rather than the two measures alone.

5.2 The compensatory logic of torts overcome…

The main findings of the theories – necessarily briefly – described above allow depicting some general considerations that could be then used to discuss the most evident features of the reasonable royalty award.

First of all, they highlight the importance of patent torts in driving innovation: since scope and duration of patents are either fixed or difficult to manage, torts represent a fundamental tool to affect parties' incentives. Among these, incentives to invest in innovation seem directly linked to patent protection, since high level of protection not only increases the trust in the patent enforcement system,[116] but strictly affects investment decision.[117]

Moreover, both models show the limits of the compensatory logic predominant in general torts, where they suggest both to award damages that surpass the actual loss suffered by the patentee when the social value of the invention is sufficiently high and below that measure in the opposite scenario, when it is not better to discourage litigation altogether.

It is interesting to notice that the justifications for allowing super-compensatory damages differ in the two theories: while in Cotter and Blair it depends on the relative efficiency of the two operators, for Hylton and Zhang it is strictly tied to the social value of the invention. This is indeed consistent with the different approach used in the two models, where the first is conceived to incentivise negotiation between the parties and the second one from a social welfare perspective.

In any case, both theories express the necessity to avoid litigation when the social value of the invention is not relevant. To this purpose, a proposed solution is to award low damages, presumably inferior to the loss suffered by the patentee.[118]

---

[116] Nikolaos Papageorgiadisa and Wolfgang Sofka (n 14).
[117] Keith N Hylton and Mengxi Zhang (n 14).
[118] Indeed, ibid 53 suggest the damages award to be 0 when the social value of the invention is low.

5.3 …and the residuality of reasonable royalty awards

Moreover, and for what here is of utmost interest, from the previous analysis emerges that reasonable royalty award, on the one side, does not always encompass for the entire damage suffered by the victim nor, on the other, it is enough when the social value of the invention is sufficiently high.

This seems consistent with the legislative approach described in para 2, where the residuality of the reasonable royalty award emerged quite clearly from all three legislative provisions. The same attitude was found in the relevant case law, notwithstanding the large use of this remedy.

In addition to this - it is expressly stated in the U.S. Code but can also be inferred from the other two systems –, this measure is supposed to constitute only the floor, the minimum compensation for patent infringement awards.

What emerges from practice, nonetheless, is that the other parameters (lost profit and disgorgement of profits) necessarily entice an information disclosure that the economic operators are not willing to subdue. Therefore, when we account for the unwillingness to disclose economic as well as technical information by the litigation parties, the probative difficulties and the uncertainties around an infringement incident, reasonable royalty becomes the primary method to calculate damages in patent infringement.

5.4 The market-based approach

Since it is the most used amongst the parameters at the courts' disposal, it has become increasingly important to follow the case law evolution on the matter, as well as the indications provided by the law and economics doctrine for reasonable royalty calculation.

What emerges from the studies and finds a validation in case law is that being patent value relative, the measure of the reasonable royalty should reflect this characteristic, by taking as basis for calculation the difference between the return on sales from the patented item and the one deriving from the next-best alternative, actualised at the moment at which negotiation is imagined. In this way, the reasonable royalty award functions as a proxy of the social value of the invention, that is to say – incidentally - an amount which is not directly relatable to the damage suffered by the patentee.

The results obtained from this operation represent the incremental profit attributable to the invention, therefore apportionable between the patentee and the willing user though a reasonably royalty.[119]

---

[119] The law and economics analysis has proposed different methods for this apportionment (eg Nash and Rubinstein bargaining solutions), to provide an economic rationale for this process and try to account for the circumstances of the negotiations. Courts in the U.S. have required, though, that these methodologies, in principle admissible, find a justification in the facts of the dispute (*see Uniloc USA, Inc.* v. *Microsoft Corp.* (n 60).

Both German and U.S. case law shows that courts try to reconstruct in a similar way the value of the right of use of the invention, by taking market-based parameters into consideration, but they differ in the moment at which the negotiation is placed in time.

In this sense, the literature examined has shown the advantages of taking an *ex post* perspective, since it allows to determine more accurately the value of the invention, that can be very uncertain at the moment when the infringement began and, therefore, through an *ex ante* evaluation.[120]

5.5 The Chinese experience: is there a sound economic rationale?

More than one international organisation has expressed concerns for the Chinese courts predominant tendency to award statutory damages for infringement dispute, because these seem too low to deter efficiently the wrongful conduct. Something that is particularly condemnable in a word in which piracy and counterfeiting are considered as plagues.

In both models briefly described above *sub* 5.1, some space is dedicated to the analysis of a combination of remedies, injunctive relief and damages. In particular, Hylton and Zhang theory of optimal patent damages suggest that a combination of the two remedies is a more efficient solution, than damages or injunction only,[121] while Cotter affirms that the measure of damages should be reduced, when coupled with injunction.[122]

It is not uncommon nor unreasonable, in fact, that these two are awarded together by courts, since the first aims at eliminating the source of damage for the future and the second one at giving adequate restore for the past harm.

It was shown above that the CIELA statistics seem to suggest that an order to desist is grant by Chinese courts in 75% of the cases in which statutory damages are also awarded.

The interrogative is, therefore, if it is not possible to find an economic rationale in the Chinese experience.

To this end, it should be necessary to analyse a wide range of judgements, to try to understand if a predictability of the measure of damages can be found and if this can in some measure overcome the low amounts awarded; moreover, if the apparent underdeterrence is not – in fact – compensated by the conjoint action by means of an injunctive relief.

6 Conclusions

From the case law comparison proposed, it seems quite clear that, apart from the *ex ante* or *ex post* perspective, the methodology used by German and U.S. court does not

---

[120] Thomas F Cotter (n 8).
[121] Keith N Hylton and Mengxi Zhang (n 14) 50.
[122] Thomas F Cotter (n 8) 63.

dramatically differ. They also seem to be in line with the recent law and economics doctrine, which suggests courts to assess via reasonable royalty by referring to the incremental value attributable to the invention and to apportion this figure between the parties.

It, then, comes as a surprise to ascertain that the concrete figures differ so drastically in these two countries and that Chinese courts, on the other hand, seem not to make use of the reasonable royalty award in the slightest.

Also, considering that these three countries have shared the primacy in terms of number of patents issued and that this numbers have been constant for a while, the great divergence seems to suggest that enforcement does not severely impact investments in innovation.

This, in turn, seems to contradict the results reached by the recent studies in law and economics, which have contributed to outline a direct link between torts and investment decisions of the patentee, being the measure of damages award the primary tool for an adequate patent protection.

The apparent contradictions depicted above can be probably explained by considering the high influence of exogenous factors (such as eg the possibility to award punitive damages, juries involvement in the decision, judicial training and burden of proof) on the level of damages award, so that the a great part of the differences found in the judgements can be largely explained by taking these factors into consideration.

On the other hand, many studies have shown how the high number of patents does not always mean a high level of innovation, since the expected average value of a single patent is very low.[123]

All in all, the differences in the judicial systems and the difficulty in assess innovation seem to explain in great part of the variance in the level of damages registered in the three countries, notwithstanding the analogies between their legislative provisions and, for Germany and the U.S., in the case law analysed.

For torts to gain a primary role in innovation governance, as the law and economics analysis seems to suggest so clearly, some obstacles need to be cleared.

First of all, then, it seems safe to say that the large space accorded to reasonable royalty in case law should be partially downsized or at least adjusted to a sounder economic rational.

The large use of reasonable royalty vouches for the elasticity of this parameter that seems to be able to incorporate a variety of data and circumstances.[124] Courts have then exploited this potentiality to respond to the need for property rights protection when no

---

[123] *See* G Colangelo (n 9) 21 ss.

[124] Interestingly, in the context of 2010 China-Tawain Agreement for property rights protection, some literature suggested courts to work on reasonable royalty and its basis, since it *"provides a flexible concept for accommodating damages"*, *see* Chung-Lun Shen (n 102) 156.

other parameters could satisfy the burden of proof, with the result of the large variance awards that has been illustrated above.

The latter, therefore, appears to be, at least partially, explainable by observing that courts the market-based data merely as a base, to be adjusted in light of the specificities of the case at law. In addition to this, both in the U.S. and in Chinese legal provisions, as well as for exceptional circumstances in German case law, the measure of damages can be enhanced by a multiplier, at the courts' partial discretion.

Contrarily, a reasonable royalty award appears economically suitable in those cases in which the value of the invention is relatively low, so when damages are not supposed to necessarily encompass the whole loss of the patentee.

In this sense, the strict burden of proof required by Chinese courts does not seem totally unjustifiable and some economic rationale might be found in coupling statutory damages (limited in amount by law) with injunctions, when adequate. Contrary to the other two countries, the risk that seems to emerge in Chinese enforcement system through statutory damages (plus injunction) is to excessively reduce the variance, so that damages become almost as a fixed measure (such as patent length), thus failing to have a significant impact on innovation incentives.

What appears then crucial is, then, a more tailored response in terms of damages award, that should not only stem from the impossibility to prove damages in other ways, but one that takes into account a comprehensive estimate of the incremental social value of the invention as well as the relative efficiencies of the parties.

To this end, some probative difficulties should be overcome, in particular those regarding the economic information of the parties, that can provide the interpreter with the tools necessary to correctly assess damages: very welcome, in this direction, the German reform briefly mentioned in note 3, that prescribe to maintain confidentiality on the firm's economic accounts produce in the proceeding.

In addition to this, a propensity for an *ex post* approach should be suggested to interpreters, since this methodology allows relying on a more precise account of the invention value, not only from a private, but also from a social perspective. In this way, the large uncertainties that have been registered in U.S. case law could be reduced, as well as the discretionary power of courts.

By relying on a larger and more precise set of information, it would also be easier to bespeak torts to a sound economic rationale, leaving little space for exogenous factors to influence the measure of damages and in such a way that patent protection can actively contribute to innovation incentives and to the social welfare.

Andrea Monesiglio [*]

PARADIGM SHIFT IN CIVIL JUSTICE IN LIGHT OF TECHNOLOGICAL INNOVATION BETWEEN THE EUROPEAN UNION, THE USA, AND CHINA

Abstract

This article will address the transformation of the paradigm of civil justice that is happening both in the European Union and in other international legal systems in light of technological innovation. I refer to the phenomenon of so-called predictive justice.

The objectives of this paper will be twofold.

The first is to check how and to what extent this paradigm shift is taking place. To this end, I will start by showing some studies carried out by the European Commission and then move on to analyse the use of predictive justice tools in select European countries (Italy, France, Estonia, the Netherlands, and Great Britain). Finally, from an intercontinental comparative point of view, I will analyse the use of those digital tools in other legal systems: China and the US.

The second objective, in light of the Italian legal system, is, on the one hand, to propose some possible practical uses of predictive justice tools and, on the other hand, to analyse their limits.

JEL CLASSIFICATION: K13, K41

1 Social phenomena and changes in legal categories

It is not a new thing to say that the digital transition is revolutionising, among the many sectors, even the legal one.

---

[*] PhD student in civil procedure law in the Faculty of Law of the University of Milano – Bicocca in the field of "PON Ricerca e Innovazione" 2014-2020.

The disruptiveness with which technological systems promise to change the field of law suggests a paradigm shift[1], especially when we talk about predictive justice systems. I'm referring to those devices which, using an appropriately representative dataset, are able to formulate the same response as a human decision-maker would give in the same case.[2]

Predictive justice is a phenomenon that reveals a distortion of the current legal procedural categories and the digital phenomena that are emerging.

In fact, it is not the first time that faced with changes in the economic and social context, the concept of the legal category has changed too. For example, similar transformations happened in the transition between Kant's agricultural Germany[3] and Hegel's industrialised Germany[4] following the first industrial revolution, which deeply affected the elaboration of the legal category of property rights. Indeed, the comparison between the Kantian[5] and Hegelian[6] conceptions of the legal category shows a paradigm

---

[1] See Antonio Carratta, '*Decisione robotica e valori del processo*' (2021) 22(2) Revista Eletrônica de Direito Processual <www.e-publicacoes.uerj.br/index.php/redp/article/view/59558/37720> accessed 18 June 2023; Antonio Punzi, *L'ordine giuridico delle macchine* (Giappichelli 2003).

[2] The Compass tool (Correctional offender management profiling for alternative sanctions) is famous in the United States. Compass is an artificial intelligence tool—mostly used in the states of Wisconsin, Michigan and Florida—used to calculate the risk of recidivism for a subject on the basis of certain variables, such as criminal records, personal attitudes, family structure, social exclusion, lifestyle and the personality of the subject.

[3] 18th century Germany was described by historians as a country that tended to lag behind other European states, for both geographical and sociopolitical reasons. For these considerations see Heinz Schilling, *Ascesa e crisi. La Germania dal 1517 al 1648* (Il Mulino 1997); Heinz Schilling, *Corti e alleanze. La Germania dal 1648 al 1763* (Il Mulino 1999); Tom Kemp, *L'industrializzazione in Europa nell'800* (Il Mulino 1997).

[4] From the second half of the eighteenth century, when Hegel was born and raised, there was a change in judicial trends. These are the years in which the concept of work changed radically thanks to exponential growth in industrial development. The ways of working were radically transformed and the work became more repetitive, specialised and mechanised by economic sector. In this context, see the studies carry out by Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (first published 1776, The University of Chicago Press 1977); David Ricardo, *Principi di economia politica e dell'imposta* (Anna Bagiotti tr, 1st edn, UTET 2006); Karl Marx, *Il Capitale* (first published 1867, Bruno Maffi tr, UTET 2017). On the influence of these authors in Hegelian thought see Remo Bodei, 'Studi sul pensiero politico ed economico di Hegel nell'ultimo trentennio' (1972) 27(4) Rivista Critica di Storia della Filosofia 435.

[5] In the work "Metafisica dei costumi" and in the "Introduzione alla dottrina del diritto §B", Kant writes that rights are the set of conditions under which the will of each agrees with the will of others according to a universal law of freedom. See Immanuel Kant, *La metafisica dei costumi* (first published 1797, Giovanni Vidari tr, Laterza 2009). Suffice it to very briefly mention here, since this is not the place for a comprehensive discussion that the construction of the Kantian juridical category seems to be elaborated in a purely individual key. In fact, in the work just cited, the arbitrary desires of the single individual are elevated to where the system and the juridical category constructed on it both hinge on it. This arbitrary right, however, must not be exercised indiscriminately but is limited by what Kant defines as a categorical imperative, a system of universally recognised moral values.
Also, this system of values (the categorical imperative), which should represent the universal law, finds its foundation, like the will of the individual, in human reason: it should derive from an individual consciousness, from a single subjective determination, an abstract universal law valid for all.

[6] In the work "Lineamenti di filosofia del diritto" the legal category that Hegel proposes reflects some aspects that the author contains within the moment of "morality" (§§ 105-141) and "ethics" (§§ 142-360). See G W F Hegel, *Lineamenti di filosofia del diritto* (first published 1820, Barbara Henry tr, Laterza 2006).
The mutual interdependence between the individual and civil society that emerges from the Hegelian work highlights the change in the economic and social substrate of the time. It is in fact interesting to note that the change in the legal category went hand in hand with the geopolitical and industrial evolution of the country and was also influenced by the classical economic school to which Hegel was drawn. See Remo Bodei, 'Studi sul pensiero politico ed economico di Hegel nell'ultimo trentennio' (1972) 27(4) Rivista Critica di Storia della Filosofia 435; Fulvia De Luise, Giuseppe Farinetti, *Lezioni di Storia della Filosofia* (Zanichelli 2012) 36.

shift that was influenced by the different historical contexts in which Kantian Germany, on the one hand, and Hegelian Germany, on the other, were placed.

A similar phenomenon is occurring in the European context with the continuous development of artificial intelligence tools. In fact, technological developments are delineating a change in the paradigm of civil justice, just as industrialisation allowed a different definition of legal categories in Hegelian Germany.

The example of the Kantian and Hegelian legal categories is only one of the many examples that shows how the social structure and the values embedded in it are able to change existing legal categories and generate new ones.[7]

A similar phenomenon has also occurred in China. In recent years, civil procedural law has undergone significant development following the transition from the Confucian tradition (in which litigation was perceived as the failure of an obligation to maintain social harmony) to the adoption of Western-style jurisdictional models.[8]

In more general terms, economic, productive, and social changes require the adaptation of the legal system through its reformulation to ensure that it is able to faithfully reflect the new social dynamics and, above all, to protect the new emerging (legal) categories that demand recognition.[9]

One recurring aspect that shows the concerns related to the use of artificial intelligence systems is the potential prejudice against fundamental human rights (in this regard, the development of the legal category of neurorights is exemplary).[10]

---

[7] I'm referring to neurorights. On the point see note 10.

[8] On these terms see: Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti, 'Intelligenza artificiale e giurisdizione ordinaria. Una ricognizione delle esperienze in corso' in Alessandro Pajno, Filippo Donati e Antonio Perrucci (eds), *Intelligenza artificiale e diritto: una rivoluzione?,* vol 2 (Il Mulino 2022).

[9] Some examples are the regulation in a single body of consumer rights (d. lgs. n.206/2005, consumer code) in order to ensure adequate protection of consumers against legal entities that are in a dominant position; the regulation of civil unions (d. lgs. n.76/2016); the regulation of workers' rights (l. n.300/1970) following requests for protection in the workplace; the will to enclose in a legislative text (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' COM (2021) 206 final (hereinafter, 'AI Act') the rules governing the placing on the market and use of artificial intelligence.

[10] For a related argument, see Marcello Ienca, 'Tra cervelli e macchine: riflessioni su neurotecnologie e su neurodiritti' (2019) 133 Notizie di Politeia 52; Gaetana Natale, 'Intelligenza artificiale, neuroscienze, algoritmi: le sfide future per il giurista' (2021) 4 Rassegna Avvocatura dello Stato 116; Oreste Pollicino, 'Costituzionalismo, privacy e neurodiritti' (2021) 2 MediaLaw 9 <www.medialaws.eu/rivista/costituzionalismo-privacy-e-neurodiritti/> accessed 22 June 2023. With regard to this category of emerging rights, in 2021 Chile proposed to become the first state in the world to include the protection of neurorights in the constitution. A law of constitutional reform, among various and additional objectives, also set out to give recognition and offer protection for neurorights. An amendment of Art. 9 of the Constitution was proposed with the addition of the following statement: "Scientific and technological development will be at the service of people and will be carried out with respect for life and physical and mental integrity. The law will regulate the requirements, conditions and restrictions for its use by people, and must especially protect brain activity, as well as the information from it". However, in 2022 the constitutional reform text was not approved. More generally on the Chilean path toward constitutional reform see Tania Groppi, Elena Bindi, Andrea Pisaneschi, 'Il Cile verso la Convenzione costituzionale' (2021) 1 DPCE online <www.dpceonline.it/index.php/dpceonline/article/view/1319> accessed 22 June 2023; Tania Groppi, 'Il Cile da un "plebiscito" all'altro. Il rechazo del nuovo testo costituzionale nel referendum del 4 settembre 2022, visto dall'Italia' (2022) 23 Federalismi <https://www.federalismi.it/nv14/editoriale.cfm?eid=645&content=&content_auth=> accessed 22 June 2023.

In fact, from a strictly legal point of view, what clearly differentiates the elaboration of past theories with modern approaches to artificial intelligence tools is the increasing attention that institutions and researchers place on fundamental rights. This has also happened in the study of justice.[11]

In general, the use of digital tools in jurisdictional activity leads to a different connotation of jurisdiction, so much so that the phenomenon has taken the name "digital justice".

In the Italian legal system, this term is also reflected in positive law. The d.lgs n.149/2022 (Cartabia reform) introduced in the implementing provisions of the Code of Civil Procedural Law the Title V ter, entitled "Provisions relating to digital justice".

However, this title contains provisions that refer to a phenomenon with a much more limited scope than the phenomenon of predictive justice. In fact, those rules concern aspects relating to the digitalisation of acts, and they have nothing to do with decision-making artificial intelligence systems.

In any case, the reference to "digital justice" seems to be a timid hint of the Italian legislature's interest in the digital phenomenon. Indeed, the legislature may not be too shy given that the term "digital justice" evokes scenarios that suggest a significant technological advance in the field of justice. Our imagined applications, however, are reframed after we read the provisions, which refer to the simplest issues related to the drafting of acts in a digital format and the transformation of analogic acts into a digital format.[12]

However, the idea of applying artificial intelligence tools in the context of jurisdictional activity has raised heated debates[13] around ethical-legal issues. In fact, on the one hand, there are some extreme theses[14] that predict an apocalyptic technological scenario, and on the other hand, there are more moderate ones[15] that exclude a dystopian future in which knowledge of the law and its application will fall under the monopoly of machines. In the middle, there is the "third way"[16] that is adopted by the European Union in the

---

[11] In this regard, the precursor document is the "European Ethical Charter on the use of artificial intelligence in judicial systems and their environment" adopted by CEPEJ on 3-4 December 2018.

[12] See da Antonio Carratta (n 1).

[13] See Amedeo Santosuosso, Giovanni Sartor 'La giustizia predittiva: una visione realistica' (2022) 7 Giurisprudenza italiana 1760. The first pages of the text, making an analysis around the "*illusions, hopes and fears*" of prediction and algorithmic decisions, offer an effective reconstruction of the feelings generated by the idea of artificial intelligence applied to the field of justice.

[14] See Roberto Bichi, 'Intelligenza digitale, giurmetria, giustizia predittiva e algoritmo decisorio. Machina sapiens e il controllo sulla giurisdizione' in Ugo Ruffolo (ed), *Intelligenza artificiale. Il diritto, i diritti, l'etica* (Giuffrè Francis Lefebvre 2020) 423; Massimo Luciani, 'La decisione giudiziaria robotica' (2018) 3 AIC <www.rivistaaic.it/it/rivista/ultimi-contributi-pubblicati/massimo-luciani/la-decisione-giudiziaria-robotica> accessed 22 June 2023.

[15] See Amedeo Santosuosso, Giovanni Sartor (n. 13); Floris Bex, Henry Prakken, 'On the relevance of algorithmic decision predictors for judicial decision-making', [2021] ICAIL <https://dl.acm.org/doi/abs/10.1145/3462757.3466069> accessed 28 June 2023; Max Tegmark, *Vita 3.0. Essere umani nell'era dell'intelligenza artificiale* (Virginia B. Sala tr, Rafaello Cortina Editore 2018).

[16] See Filomena Santagada, 'Intelligenza artificiale e processo civile' in Rosaria Giordano and others (eds), *Il diritto nell'era digitale. Persona, Mercato, Amministrazione, Giustizia* (Giuffrè Francis Lefebvre 2022); Antonio Punzi, 'Judge

proposal for a European regulation on artificial intelligence (AI Act) and that I will discuss in the next section.

2 Studies carried out by the European Commission

The theme of artificial intelligence has been present since 2004 in the research and development programmes of the European Union,[17] which has allocated a large amount of capital for technological development.[18]

The European Union has in fact intensified its efforts to remain competitive in the field of research and technological development: the stated goal[19] is to become an attractive centre for the global market and to avoid falling behind the US and China.[20]

The first strategic guidelines to concretely outline this objective were defined by the European Commission in the communication of 25 April 2018 COM (2018) 237 final, "*Artificial intelligence for Europe*".[21] The guidelines consist of several profiles: i) a properly strategic profile, relating to the concrete actions to establish itself as a market of excellence, ii) a profile relating to cooperation between Member States, iii) a legal profile and, not by chance, iv) a socioeconomic profile.

The profile dedicated to the legal aspects opens with an affirmation that shows the direction that the European Union would then take: "An environment of trust and accountability around the development and use of AI is needed".[22]

Following this objective, in 2018, the European Commission appointed a group of experts—High-Level Expert Group on Artificial Intelligence (HLEG AI)—with the aim of defining ethical guidelines for the design of artificial intelligence tools to ensure an

---

in the machine. E se fossero le macchine a restituirci l'umanità nel giudicare?' in Alessandra Carleo (ed), *Decisione robotica* (Il Mulino 2019); Antonio Punzi, 'Difettività e giustizia aumentata. L'esperienza giuridica e la sfida per l'umanesimo digitale' (2021) 1 Ars Interpretandi <www.rivisteweb.it/doi/10.7382/100796> accessed 30 June 2023; Edoardo Rulli, 'Giustizia predittiva, intelligenza artificiale e modelli probabilistici. Chi ha paura degli algoritmi?' (2018) 2 Ars Interpretandi <www.rivisteweb.it/doi/10.1433/92116> accessed 30 June 2023. See also Nicola Lettieri, 'Contro la previsione. Tre argomenti per una critica del calcolo predittivo e del suo uso in ambito giuridico' (2021) 1 Ars Interpretandi <www.rivisteweb.it/doi/10.7382/100794> accessed 30 June 2023.

[17] The focus was initially on the robotics sector (2004). Investments in this sector have increased exponentially over the years, reaching up to 700 million euros in the period 2014-2020, in addition to the 2.1 billion euro of private investments in a public-private partnership on robotics. See Commission, Artificial intelligence for Europe (Communication) COM (2018) 237 final.

[18] 95.5 billion euro have been allocated under the "Horizon Europe" programme for the period 2021-2027 and approximately 80 billion euro under the previous "Horizon Europe 2020" programme for the period 2014-2020).

[19] COM (2018) 237 final, 3.

[20] COM (2018) 237 final, 4.

[21] On this communication, the European Economic and Social Committee (EESC) issued its opinion (Comitato Economico e Sociale Europeo, 'Parere del Comitato economico e sociale europeo sulla "Communicazione della Commissione al Parlamento europeo, al Consiglio europeo, al Comitato economico e sociale europeo e al Comitato delle regioni – L'intelligenza artificiale per l'Europa" (2018/C 440/08) <https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52018AE2369&from=IT> accessed 30 June 2023).

[22] COM (2018) 237 final, 13.

adequate legal and ethical framework based on the values of the European Union and consistent with the EU Charter of Fundamental Rights.[23]

To fulfil the strategies contained in COM (2018) 237 final, the European Commission subsequently adopted the "*Coordinated plan on artificial intelligence*",[24] accompanied by the annex "*Coordinated plan on the development and use of artificial intelligence 'Made in Europe' – 2018*"[25]: both documents have been approved by the Council of the European Union. While the first provides a general overview of the objectives that the European Union has set itself, the second defines several concrete actions to be implemented at both the European and national levels in the 2019-2020 period.[26]

Another important initiative of the European Commission was the adoption of the white paper on artificial intelligence on 19 February 2020.[27] In this document, the European Commission defined further strategic guidelines for the safe and reliable development of artificial intelligence tools with the aim of defining an "ecosystem of excellence" and an "ecosystem of trust".

The white paper laid the foundations for the future definition of the regulatory framework for artificial intelligence. The European Commission argued that users' full confidence in AI systems was necessary. The risks associated with the use of artificial intelligence equipment had to be eliminated or reduced to a minimum.

With regard to risk management, the white paper shows that the multilevel regulatory system (cd risk-based approach) adopted by the European Commission in the AI Act had already been envisaged by the German Commission on Data Ethics. This German commission had advanced the idea of "a five-level risk-based system of regulation that would go from no regulation for the most innocuous AI systems to a complete ban for the most dangerous ones".[28]

This approach—as mentioned—was adopted in the proposal to regulate artificial intelligence in AI Act.

The AI Act has been subject of a lengthy confrontation between European political forces, which has led to the formulation of thousands of amendments, so much so that the tripart dialogue between European institutions is still in progress.

---

[23] The High-Level Expert Group on Artificial Intelligence (HLEG AI) developed several documents with the aim of tracing an ethical perimeter within which the design and use of artificial intelligence systems should take place. It thus identified a number of principles (defined ethical imperatives) which should be taken into account: i) respect for human autonomy; ii) prevention of harm; iii) equity; iv) explicability. See High-Level Expert Group on Artificial Intelligence, 'Ethics guidelines for trustworthy AI' (2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> accessed 30 June 2023.
[24] Commission, 'Coordinated Plan on artificial intelligence' (Communication) COM (2018) 795 final.
[25] Commission, 'Coordinated plan on the development and use of artificial intelligence "Made in Europe" – 2018' (Communication) COM (2018) 795 final, Annex.
[26] In addition, the coordinated action plan also provided for continuous monitoring of the plan. In 2018, the "AI Watch" project was launched within the Joint Research Centre (JRC) to provide the necessary analyses for the implementation of the European Artificial Intelligence Initiative.
[27] Commission, 'White paper on Artificial Intelligence - A European approach to excellence and trust' (Communication) COM (2020) 65 final.
[28] COM (2020) 65 final, 10.

A compromise position on the proposal presented by the European Commission was not reached until 9 May 2023, when the LIBE committee and the IMCO committee voted on the amendments previously tabled by the various political forces. The new draft was approved by 84 votes to 7 with 12 abstentions.

Finally, on 14 June 2023, the European Parliament voted in favour of the compromise draft, thus devolving the act to the Council of the European Union in an attempt to achieve the adoption of the regulation by the end of 2023.

The compromise draft applies the "risk-based approach": the regulation of AI systems depends on the degree of risk they pose to human rights, and each risk category is subject to a different legal regime. In particular, the compromise draft approved by the European Parliament subdivides the AI systems into five risk categories.

For the purposes of our investigation, the "high-risk AI systems", which include AI systems "intended to be used by a judicial authority or administrative body or on their behalf to assist a judicial authority or administrative body in researching and interpreting facts and the law and in applying the law to a concrete set of facts or used in a similar way in alternative dispute resolution", are relevant (Annex III, paragraph 8, lett. a).

The formulation of this provision inevitably implies an awareness on the part of the European institutions, which is perhaps still lacking in European national legal systems, although not in all.

3 The predictive justice systems in the European landscape

3.1 The state of the art in the Italian legal system

Applying the provision of Annex III, paragraph 8, lett. a) in the Italian context, we observe that it represents the current experience of both the public administration and the judicial offices.

Regarding the experiences of public administration, research carried out in the framework of the PRIN 2017 project "Governance of/through Big Data: Challenge for European law" has shown that in the Italian legal system, AI systems are adopted in a different way[29]. In particular, three different trends have emerged: i) independent authorities are inclined to resort to independent handling and interfunctional collaborations: the programming of AI systems is tailored to their specific needs; ii) smart cities prefer the use of two alternative models: public tender or self-production; and iii) in the central government, however, there is no single trend: the evaluation of the experiences of the INPS, the Revenue Agency and the Ministry of Justice has led to the

---

[29] See Edoardo Chiti, Barbara Marchetti e Nicoletta Rangone, 'Impiego di sistemi di intelligenza artificiale nelle pubbliche amministrazioni italiane: prove generali' (2022) 2 BioLaw Journal <https://teseo.unitn.it/biolaw/article/view/2351/2296> accessed 2nd July 2023.

affirmation that there are misalignments at the central level regarding the modalities of action.

With reference to the Ministry of Justice, the abovementioned research has shown that the process of digitisation is slow and uneven in the various sectors (civil, criminal and administrative). As a result, this has a negative impact on the possibilities of developing AI systems of predictive justice and putting them into service in the judicial field.

In other public sectors, in recent years, public administration has instead started to use artificial intelligence systems with decision-making functions.[30] One of the sectors that has seen a wide use of this type of tool is the education system.[31] In this field, numerous disputes have emerged because the public offices, using algorithmic decision-making systems, had violated the obligation to always motivate its decisions stated under Art. 3 L. n.241/1990. The disputes thus settled have created an extensive jurisprudence,[32] which has established different principles such as i) the principle of algorithmic nondiscrimination, ii) the principle of nonexclusivity (or the reserve of humanity in the decision) and iii) the principle of comprehensibility-knowability of the decision (that is, the principle that tries to cope with the opacity of the algorithm (so-called "black box") that prevents explaining the logical-argumentative path followed by the software in making the decision.)

With reference to the judicial offices, at the current state of the art, the use of artificial intelligence systems with decision-making functions is not yet recorded. AI systems of this magnitude have not yet been developed. There are, however, numerous projects resulting from agreements concluded between courts on the one hand and universities on the other.

Some examples are presented here.

On 6 April 2018, the University of Brescia, the Ordinary Court of Brescia and the Court of Appeal of Brescia signed the convention "Predictive Justice". The aim of the project is to provide users with information on the foreseeable duration of the trial (regarding the

---

[30] On this point, see Filippo Donati, 'Intelligenza artificiale e giustizia' in Antonio D'Aloia (ed), *Intelligenza artificiale e diritto. Come regolare un mondo nuovo* (Franco Angeli 2020) 248; Silvia Sassi, 'Gli algoritmi nelle decisioni pubbliche tra trasparenza e responsabilità' (2019) 1 Analisi giuridica dell'economia <https://www.rivisteweb.it/issn/1720-951X/issue/7760> accessed 2 July 2023; Giuliano Avanzini, *Decisioni amministrative e algoritmi informatici. Predeterminazione analisi predittiva e nuove forme di intellegibilità* (Editoriale Scientifica 2020).

[31] The issues concerned the use of software by the Ministry of Education in the context of the procedures for the allocation of seats and the mobility procedures of school staff following the law's entry into force c.d. "buona scuola" (L. n. 107/2015).
We also note the recent pronouncement of T.A.R. Campania, 14 November 2022, n. 7003 relating, instead, to the disputes concerning the use by AGEA (Agency for Agricultural Payments) of an algorithmic instrument for redefining the amounts of certain allowances for which the applicant was a beneficiary.

[32] See T.A.R. Lazio, Sez. III-*bis*, 22 March 2017, n. 3769; T.A.R. Lazio, Sez. III-*bis*, 10 September 2018, nn. 9224-9230; Cons. St.., Sez. VI, 8 April 2019, n. 2270; di T.A.R. Lazio, Sez. III-*bis*, 27 May 2019, n. 6606; Cons. St., Sez. VI, 13 December 2019, nn. 8472-8474; Cons. St., Sez. VI, 4 February 2020, n. 881; T.A.R. Lazio, Sez. III-*bis*, 24 June 2021, n. 7589.

specific subjects covered by the project) and the relevant guidelines of the jurisprudence of the two judicial offices that are parties to the convention.[33]

On 29 September 2021, the Electronic Documentation Centre (CED) at the Court of Cassation and the University Higher School IUSS Pavia signed an agreement with the aim of developing technical tools for the collection and organisation of digital legal material through legal analysis. The stated objectives are i) to predict the outcome of judicial, administrative and political proceedings; ii) to extract the argumentative structures from the body of judicial decisions; iii) to automatically create summaries of the judgments; and iv) to design tools that automatically create documents.

A technologically ambitious project is in place at the Ca' Foscari University of Venice in conjunction with the Court of Appeals of Venice and Unionecamere of Veneto in collaboration with Deloitte, which provides technical support.[34] The aim is to create the first prototype of predictive justice software applied in jurisprudence. In particular, 800 decisions issued in the three-year period 2019-2021 by the Territorial Courts in the district of the Court of Appeals of Venice were analysed. These 800 decisions, issued in the field of dismissal for justified objective reasons, have been labelled to train the algorithm that will be called upon to predict the outcome of a dispute in that matter.

A project of wider territorial scope is "NextGeneration UPP",[35] which involves the Judicial Offices of Macro Area A1[36] and numerous universities.[37]

NextGeneration UPP aims to improve the efficiency and performance of justice in north-western Italy[38]. In particular, the objective is to provide the judicial offices with an efficient method of analysis and management of incoming and outgoing litigation, on the one hand, and to provide the "Ufficio per il Processo" attendants (UPP) with cross-cutting skills, on the other.

The attempt is therefore to strengthen the system of digitisation of the judicial offices through tools for legal analytics, thus creating an interdisciplinary environment in which data science, machine learning and natural language processing activities converge.

---

[33] The agreement is available at the following link <www.giustizia.brescia.it/allegatinews/A_18592.pdf>. The project has a dedicated, specific area available at the following link https://giustiziapredittiva.unibs.it/

[34] On 14 November 2014, a conference was held at Ca' Foscari University on this project. The interventions are available on the website of the Court of Appeals of Venice at the following link <https://ca-venezia.giustizia.it/giurisprudenza-predittiva_466.html>.

[35] It is a project that is a part of the Piano Operativo Nazionale – Governance e Capacità Istituzionale 2014-2020 (PON) and it is implemented in synergy with the interventions provided by the National Recovery and Resilience Plan (PNRR).

[36] The Courts of Appeals of Brescia, Genoa, Milan and Turin, the ordinary courts of the respective districts and the courts for minors.

[37] The Next Generation UPP, coordinated by the University of Turin, is realised in partnership with the University of Milan Bicocca, the University of Bergamo, the University of Brescia, the University of Genoa, the University of Insubria, the University of Milan, the University of Eastern Piedmont Amedeo Avogadro, the University of Pavia, the University of Advanced Studies of Pavia, the Polytechnic of Milan and the Polytechnic of Turin.

[38] For an in-depth analysis of the lines of intervention, specific actions and objectives and expected results, please refer to the dedicated website of the University of Milan Bicocca, among others, at the following link <https://giurisprudenza.unimib.it/it/ricerca/next-generation-upp> accessed 5 July 2023

Similar to the latter project are those operating in the other macro areas in which the judicial offices are territorially divided.

In particular, in Macro Area 02, we find the Uni4Justice[39] project that has the University of Bologna as its leader; in Macro Area 03, there is the project Giustizia Agile,[40] coordinated by the University of Tuscia; in Macro Area 04, the project MOD-UPP[41] is coordinated by the University of Naples Federico II; in Macro Area 05, the project StartUPP[42] is coordinated by the University of Bari; and, finally, in Macro Area 06, the project JustSmart[43] is coordinated by the University of Palermo.

Like the NextGenerationUPP, these latest projects are also part of the PON in synergy with the activities of the PNRR and have as an objective the efficiency of the justice system through better management of the operating models and workflows.

However, unlike what happened in some other public offices, to date, no judicial office has yet used predictive justice software with decision-making functions, but the existing projects in the various courts and universities suggest that in the near future, we may reach these developments, as happened in other European cases.

We will now see that, from a comparative point of view, the Italian context regarding the use of predictive justice software is lagging behind that of other European states, in which, instead, initiatives of a publicist matrix have been advanced.

Taking a look at the "Artificial Intelligence Strategic Program 2022-2024" developed by the Italian government, we note that the lines of action are still quite general, and although there is some concern over the enhancement of digital structures in the various public administrations, no targeted policy lines involving the use of predictive justice tools have been specifically identified. This, however, is not much of a surprise, since a proper design of such software would necessitate—or at least, would be facilitated by—the preparation of a public database containing (anonymised) measures issued by the judicial authorities, directly fed by the same Ministry of Justice from which you can draw datasets to train software.

3.2 The state of the art in other European countries

In France, pursuing the goal of encouraging innovation and the digital economy, has already promoted the circulation of data with Loi 2016/1321 by establishing the obligation for public administrations to publish online the main documents in their possession of

---

[39] Please refer to the dedicated site <https://dsg.unibo.it/it/ricerca/progetti-di-ricerca/progetti-nazionali-e-di-ateneo/uni-4-justice> accessed 5 July 2023

[40] Please refer to the dedicated site <https://www.unitus.it/it/unitus/mappatura-della-ricerca/articolo/giustizia-agile> accessed 5 July 2023

[41] Please refer to the dedicated site <http://www.unina.it/-/30852250-nuovo-ufficio-per-il-processo-modelli-organizzativi-e-innovazione-digitale> accessed 5 July 2023

[42] Please refer to the dedicated site <https://www.uniba.it/it/ateneo/rettorato/ufficio-stampa/comunicati-stampa/anno-2022/giustizia-agile> accessed 5 July 2023

[43] Please refer to the dedicated site <https://sites.unica.it/ict4lawforensics/justsmart/> accessed 5 July 2023

economic, social, health and environmental interest. In the area of justice, all judicial decisions have been made public.

While this is certainly favourable for the development of forward-looking justice software, it has raised some concerns.

For the positive aspects, the realisation of some software, such as Predictive, JuriPredis, and Case Law Analytics, must be noted.

Special mention should be made of the DataJust software, which was created following the public initiative taken by the Ministry of Justice, authorised by Decree 2020/356 of 27 March 2020. The goal of DataJust is to carry out an automated processing of data related to settlements in cases with personal damages to realise predictive software for the resolution of new cases.[44]

For the critical aspects, some concerns have rightly arisen in the French legislature regarding both the possible profiling of judges and the profiling of individuals according to their behaviour. Thus, Loi 2019/222 sanctioned a ban on profiling judges, whose violation is subject to a penalty of imprisonment for up to 5 years, and Loi 2018/493 imposed a ban on judicial decisions based on assessments of the behaviour of people obtained through automated processing of data that concern individuals' personalities.[45]

Estonia is a country that is often mentioned when we talk about these issues: a brief look at its context can explain why.

Estonia is a country that has approximately 1.3 million inhabitants, and digitalisation significantly permeates public services. For example, citizens can rely on a digital identity system that allows them to enjoy services such as electronic voting.

More generally, government databases are linked by a system called the X-road that facilitates the exchange of information between public administrations.[46] The latter, therefore, have the possibility of communicating effectively because there is an adequate communication channel.

The sector that interests us is known as an ambitious initiative of the Ministry of Justice, with which it wants to implement artificial intelligence for small claims (the limit set is € 7,000). It is software that is invested in the entire conflict, whose "robotic decision" can be challenged before a human decision-maker.

This is perhaps the most significant project in the field of predictive justice. The ambitious goal is to create a software substitute for the judge, without prejudice to the indispensable right to appeal to a human decision-maker for a possible reform of the digital decision.[47]

---

[44] On the point, see Giada De Pasquale, 'La giustizia predittiva in Francia: il trattamento di DataJust' [2021] Judicium <www.judicium.it/la-giustizia-predittiva-francia-trattamento-datajust/> accessed 5 July 2023. For some reflections on the different reactions of the forensic profession to private and public initiatives see Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

[45] See Edoardo Chiti, Barbara Marchetti e Nicoletta Rangone (n 29).

[46] Antonio A Martino, 'Chi teme i giudici robot' (2020) 2 Rivista italiana di informatica e diritto <www.rivistaitalianadiinformaticaediritto.it/index.php/RIID/article/view/58> accessed 27 June 2023.

[47] Antonio A Martino (n 46); Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

In the Netherlands, as far as we know, there are no predictive justice systems as previously outlined. However, since 2014, an online dispute resolution software (ODR) called Rechtwijzeruit and developed by the University of Twente and Hill (Hague Institute on the Innovation of Law), has been tested, at first limited exclusively to divorce proceedings. In 2017, the platform was redefined and renamed Uitelkaar.nl.

This tool is proposed as an online advice system for private parties who, after having made contact with mediators and lawyers, are followed step by step in the path of their interest (preparation of the divorce plan, the parental plan or both) to see the documents submitted to the court for validation.[48]

Even in the United Kingdom, as far as we know, there are no instruments of predictive justice in the civil sphere. However, mention should be made of the use of risk assessment tools, as is the case in the United States.

The most known application is perhaps the HART software, developed by the University of Cambridge and trained using the data contained in the Durham police archives for the period 2008-2012.[49]

Its operation is based on a particular machine learning model called random forest.[50] Through the analysis of 34 variables—among which there are data on the subject's criminal history, age and gender—we are able to calculate the risk that the arrested person commits further crimes in the following two years.[51]

As with other risk assessment tools used in different jurisdictions (such as the US), the use of HART was not without its critics.[52]

---

[48] On the point, see Laura Kistemaker, '*Rechtwijzer and Uitelkaar.nl. dutch experiences with ODR for divorce*' (2021) 59 (2) Family Court Review <https://search.informit.org/doi/abs/10.3316/agispt.20210506046114> accessed 2 July 2023.

[49] See Mitja Giualuz, 'Quando la giustizia penale incontra l'intelligenza artificiale: luci e ombre dei risk assessment tools tra stati uniti ed europa' [2019] Diritto Penale Contemporaneo <https://archiviodpc.dirittopenaleuomo.org/d/6702-quando-la-giustizia-penale-incontra-l-intelligenza-artificiale-luci-e-ombre-dei-risk-assessment-too> accessed 2 July 2023; Marion Oswald, Jamie Grace, Sheena Urwin and Geffrey C Barnes, 'Algorithmic risk assessment policing models: lessons from the Durham HART model and "Experimental" proportionality' (2018) 27 (2) Information and Communications Technology Law <www.tandfonline.com/doi/full/10.1080/13600834.2018.1458455> accessed 2 July 2023.

[50] Mitja Giualuz (n 49).

[51] For a more in-depth overview see Mitja Giualuz, (n 49); Marion Oswald, Jamie Grace, Sheena Urwin and Geffrey C. Barnes (n 49).

[52] See Big Brother Watch, 'Big Brother Watch's Written Evidence on algorithms in the Justice System for the Law Society's Technology and the Law Policy Commission' (February 2019) <https://bigbrotherwatch.org.uk/wp-content/uploads/2019/02/Big-Brother-Watch-written-evidence-on-algorithms-in-the-justice-system-for-the-Law-Societys-Technology-and-the-Law-Policy-Commission-Feb-2019.pdf> accessed 10 July 2023. With reference to the criticisms about privacy, see Hannah Couchman, 'Policing by machine. Predictive policing and the Threat to Our Rights' (Liberty, January 2019) <www.libertyhumanrights.org.uk/issue/policing-by-machine/> accessed 10 July 2023.

4 The state of the art in China and in the US

In this section, we will consider the Chinese and US legal systems and see how public initiatives relating to the development of digital justice tools in the civil field are placed at opposite ends.

4.1 Digital justice in China

As mentioned in section 1, in recent years in China, there has been a shift from a Confucian tradition to a more Western-style jurisdictional model. We will argue that these changes have led, on the one hand, to an exponential increase in the number of trials and, on the other hand, to the recognition of significant inadequacies in the Chinese judicial system.[53]

These aspects have emerged in a time when the Chinese territory witnessed remarkable growth in technological development, which contributed to the rise of the so-called 'internet court phenomenon'.

China has embarked on an articulated path to exploit the possible advantages deriving from the use of digital systems since 1990. This path can be mainly divided into three phases.[54] The first (1996-2003) saw a progressive development of the digitalisation of civil trials: it began in 1996 following the National Conference on Matters of Court Communication and Computer, and it saw its conclusion in 2003 when the courts finished the process of digitising the files and their websites.[55]

The second phase (2004-2013) is characterised by the use of devices with internet access to allow the conduct of hearings remotely and to facilitate the exchange of documents between the parties and the court. This development may seem trivial, but given the territorial dimensions of China, the establishment of this practice has made it much easier to carry out trials between parties that are at a significant distance from each other. In addition, to monitor the progress of the trials, in 2009, the Beijing High People's Court built a website to allow online access to the ongoing hearings.[56] In this very context, the creation of a database containing both the legal references and the decisions issued by the courts was also envisaged.[57]

---

[53] Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

[54] This subdivision belongs to Shi C, Sourdin T and Li B, 'The Smart Court – A New Pathway to Justice in China?' (2021) 12 International Journal for Court Administration <https://doi.org/10.36745/ijca.367> accessed 5 July 2023. On the same matter, please see: Benjamin Minhao Chen, Zhyu Li, 'How will technology change the face of Chinese justice' (2020) 34 Colum J Asian L <https://heinonline.org/HOL/Page?handle=hein.journals/colas34&div=3&g_sent=1&casa_token=&collection=journals> accessed 5 July 2023.

[55] C Shi, T Sourdin, B Li. (n 54).

[56] Gao Jian, Yao Xueqian, Zhao Yan, 'Beijing High People's Court Now Livestreaming Court Hearings' (Sina News, 17 September 2009) <http://news.sina.com.cn/c/2009-09-17/065816311211s.shtml> accessed 5 July 2023.

[57] Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

The third phase began in 2014, when the Chinese Supreme People's Court expressed its opinion on the 4th Five-Year Reform of the People's Courts (which covered a timespan from 2014 to 2018). The Chinese Supreme People's Court stressed the importance of relying on technology to build judicial mechanisms that are open, dynamic, transparent, and convenient to improve public understanding, trust, and supervision of the judiciary.[58] More recently, in the 5th Five-Year Reform of People's Courts (2019-2023), the Chinese Supreme People's Court confirmed that the creation of smart courts is one of the ten objectives of justice reform.

Thus, in 2017, the first internet court was unveiled in Hangzhou, in the province of Zhejiang.[59]

This is not strictly a system of predictive justice as outlined in the previous sections, as processing final decisions on the basis of a dataset is not the exclusive task of the software. It is indeed software that not only provides a very high level of digitalisation of the trial but also involves digital support from beginning to end. In particular, the parties, through a platform that uses a blockchain system, start the trial, file the motions and proceed with the notification of the acts. The platform can host the hearings in a videoconference and collect evidence online (ie, hearing the testimony of a witness). The Hangzhou internet court exercises its jurisdiction in disputes related to internet use (such as, for example, online purchases), and not coincidentally, it arose at the place where Alibaba[60] has its registered office.

In any case, this is a method of carrying out a trial that is not binding on the parties, who may also opt for the ordinary modalities of carrying out the trial.

It is also interesting to note the role of the central government. In this context, it promotes and encourages decentralised initiatives: Chinese courts, through public funding from local governments, are called to develop autonomous initiatives to improve justice through legaltech systems in synergy with private companies.[61] In fact, following the creation of the first Hangzhou internet court in 2017, others were built in Beijing and Guangzhou.

The main purpose of the central government, however, is not to replace the human judge with internet courts; on the contrary, the internet courts are part of the

---

[58] For the English version of the 4th Five-Year Reform of the People's Courts please refer to the following link www.hshfy.sh.cn/shfy/web/xxnr_yshj.jsp?pa=aaWQ9MjAyMTUxMTQmeGg9MSZsbWRtPUxNMTIxMwPdcssPdcssz&zd= accessed 5 July 2023.

[59] An author's comment following a visit to the Hangzhou Court in 2017: "I was impressed with what I saw: a static robot in the reception area that offered online legal help for court users; on-site facilities for the e-filing of documents; dedicated virtual courtrooms; speaker-independent voice recognition (they no longer need stenographers); and a demonstration of China's first 'internet-court', which resolves internet-related disputes concerning, for example, online loans, e-commerce (contractual and product liability issues), domain name disputes, and online copyright issues. With 800 million users in China, the volume of related disputes has called for new methods. I am told that the court in Hangzhou has now handled more than 10,000 disputes, in roughly half of the time of traditional hearings" in Richard Susskind, *Online Courts and The Future of Justice* (OUP 2019).

[60] Alibaba is a private Chinese multinational enterprise composed of a several companies active in the field of e-commerce.

[61] Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

administration of justice as software to support the human decision-maker. In addition, the positive results that have been recorded support hope for such an outcome. In fact, the Hangzhou internet court reported an increase in the efficiency of the trials and noted that the online collection of evidence allowed the parties to avoid moving, therefore making the litigation more accessible.[62]

According to the data at hand, in two years of operation (2017-2019) 20,000 judgments were issued and there was a reduction of 65% of the time spent in hearings.[63] Given the success of these methods of litigation, the Hangzhou internet court in 2019 published a white paper on the application of internet technology in judicial practice,[64] in which the technologies used by the court were indicated (of particular concern is the use of facial recognition to confirm the identity of the parties in the trial).

4.2 Digital justice in the US

The issue of civil digital justice in the United States seems to stand in contrast to other legal systems examined. In fact, it appears that at the moment, there are no public initiatives aimed at developing predictive justice software in the civil field, not that the courts would make some use of these kinds of software. [65]

However, there are some kinds of software used in law firms that were created as a result of private initiatives. Some examples are Coin (used by J. P. Morgan), Kira (used by DLA Piper), LexMachina (purchased by LexisNexis in 2015) and LinkRFI (used by Linklaters).[66]

It should be noted that, even if there are no civil initiatives, artificial intelligence supports (in particular, risk assessment tools) have been used for a long time in the US in the criminal justice field.

The most popular case on the matter is *Loomies vs. Wisconsin*, in which the COMPASS software was applied. It is a predictive sentencing tool used for the calculation of the risk of recidivism of accused persons in criminal proceedings and whose output is based on a number of variables: criminal records, personal attitudes, family structure, lifestyle, personality and social exclusion. It is an application mostly used in the states of Wisconsin, Michigan and Florida, where its use, such as with the use of the HART software, has not gone without criticism.[67]

---

[62] C Shi, T Sourdin, B Li (n 54).

[63] D Chen, C Wang, 'What Hangzhou Internet Court has brought to us in the past two years' *(*Xinhua Net Legal Daily (, 15 August 2019) <www.zj.xinhuanet.com/2019-08/15/c_1124877777.htm> accessed 15 June 2023.

[64] Beijing Internet Court, 'White Paper on the Application of Internet Technology in Judicial Practice' (17 August 2019) <www.chinadaily.com.cn/specials/WhitePaperontheApplicationofInternetTechnologyinJudicialPractice.pdf> accessed 15 June 2023.

[65] In this term also Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

[66] For a more in-depth overview of these software programmes, see Mario Libertini, Maria Rosaria Maugeri, Enzo Vincenti (n 8).

[67] See the analysis conducted by ProPublica <www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> accessed 15 June 2023 and the subsequent analysis conducted by Northpointe Inc. (now Equivant)

In the field of criminal law, in addition to predictive sentencing tools, there is also the use of law enforcement software. These tools are used by the police to identify areas most at risk: they are therefore aimed at preventing the commission of crimes.[68]

These are software programmes that reflect the debate still alive in the United States (but spread across Europe) on the possibility of predicting the criminal behaviour of individuals.[69]

The strong and constant interest of the United States in these applications could perhaps suggest why no targeted action has yet been taken in relation to predictive justice tools in the civil field.

5 Some benefits and issues of predictive justice tools in light of Italian legal systems

In this section, I will analyse both plausible and critical issues related to the use of predictive justice software within the civil trial. Since it is not possible here to deal exhaustively and comprehensively with all the possible aspects rising around the matter, I have chosen to analyse those of more immediate evidence.

5.1 Potential uses

5.1.1 Immediacy of decisions

Perhaps the first benefit that comes to mind when thinking about opportunities related to the use of predictive justice systems is the potential reduction in the time spent in trials.

One of the features of Italian civil trials is the system of procedural bars. This regime, among other functions, ensures a precise scan of the timing within which to define the perimeter of the *thema decidendum* and the *thema probandum*.

Although the recent reform of D. Lgs. n.149/2022 has reshaped the procedural phase relating to the definition of both the *thema decidendum* and *the thema probandum* with a view to accelerating the trial (Art. 163 and 171-ter Code of Civil Procedure), it will take at least 120 days for these terms to be fully defined.

<http://go.volarisgroup.com/rs/430-MBX-989/images/ProPublica_Commentary_Final_070616.pdf> accessed 15 June 2023.

[68] On the point see Fabio Basile, 'Intelligenza artificiale e diritto penale: quattro possibili percorsi di indagine' (2019) 10 Diritto Penale e Uomo <https://archiviodpc.dirittopenaleuomo.org/d/6821-intelligenza-artificiale-e-diritto-penale-quattro-possibili-percorsi-di-indagine> accessed 15 June 2023.

[69] Clementina Barbaro, 'Uso dell'intelligenza artificiale nei sistemi giudiziari: verso la definizione di principi etici condivisi a livello europeo?' (2018) 4 Questione Giustizia <www.questionegiustizia.it/rivista/articolo/uso-dell-intelligenza-artificiale-nei-sistemi-giud_591.php> accessed 17 June 2023.

To this term, the time necessary for evidentiary instruction, which strongly depends on the written load of the individual judge, and time relating to the decision-making phase must be added.

A significant reduction in the timelines just exposed could be achieved through the use of software similar to the one being designed in Estonia. We could in fact think—at least for a first phase of experimentation—about the use of such software for the resolution of small claims. However, the critical issues that will be further analysed in the following section suggest that such a system should preferably be placed outside the jurisdictional field, at least until the main difficulties can be overcome. It could be better framed as online dispute resolution software (ODR) or as software that can be used by the lawyer to provide to his or her client a preliminary view of the hypothetical outcome of the dispute to better consider the choice of whether to take legal action.

5.1.2 Disincentive of litigation

The use of predictive justice software as outlined in the previous section would have two types of consequences.

We've already spotted one. If the programme ware able to predict the hypothetical outcome of the dispute, this may lead to, on the one hand, economic savings for the customer (who would otherwise be exposed to a vain economic outlay in the event of failure, after considering, for example, the costs of litigation and professional fees) and, on the other hand, a disincentive of litigation[70] with the consequent reduction of pending loads.

The second consequence relates to the knowledge of the probable conclusion of the trial brought by the possible plaintiff, which could be used by the judge as an index to evaluate the abuse of the right of action or the exercise of this merely for the purposes of delay. This is reflected in the topic of *vexatious* litigation.

The judge, in fact, would have at his or her disposal an additional criterion in light of which to assess the diligence of the parties involved and, therefore, to assess whether the plaintiff was actually aware of the manifest groundlessness of his accusations. From this perspective, the way is thus opened to a different evaluation of the judgment that the judge is called to make in relation to "*mala fede o colpa grave*" (Art. 96, co. 1 Code of Civil Procedure).

---

[70] In further discussion of these terms, see also Elena Gabellini, 'La comodità nel giudicare: la decisione robotica' (2019) 4 Riv. trim. dir. proc. civ 1305; Elena Gabellini, 'Algoritmi decisionali e processo civile: limiti e prospettive' (2022) 1 Riv. trim. dir. proc. civ. 59.

5.1.3 Legal certainty and legal calculability

Another positive aspect of predictive judicial systems is the valorisation of precedents.[71] An automated decision would certainly reward the impartiality and objectivity of the previous decision. In fact, in the same cases, unreasonable differential treatment would be avoided, thus achieving better standards of substantial justice[72] (Art. 3 of the Italian Constitution).

It has been pointed out, however, that in a multiplicity of cases, there are factual and legislative elements that may complicate, if not prevent, a correct calculation.[73] Some of them are, for example, i) the crisis of juridical positivism[74] and an increasing stratification of the sources of law[75]; ii) the physiological profiles of juridical incalculability as well as any antinomies and the use by the legislature of flexible terms and general clauses.

5.2 Criticality of predictive justice tools

5.2.1 Opacity of the algorithm

Given the fundamental role played by the motivation of the decision, the issue of the opacity of the algorithm (so-called black box)[76] becomes extremely problematic. The result presented by the programme is in fact not explainable in the same way and extent to which the judge has the obligation to justify his or her conviction (Art. 111 Italian Constitution, 132, n.4 Code of Civil Procedure, 118 provisions implementing the Code of Civil Procedure). In fact, following the implementation of l. n. 69/2009, to respect the

---

[71] On the importance of predictability of decisions for the efficiency of justice see Giorgio Costantino, 'La prevedibilità delle decisioni tra uguaglianza e apparenza (Relazione alla XXIX Conferenza dell'Osservatorio Giordano dell'Amore sui rapporti tra diritto e economia, Milano, 5 febbraio 2015 e al Primo Congresso Giuridico di Monza, Como e Lecco, Monza, 19 febbraio 2015)' (2015) 3 Rivista di diritto processuale 646.

[72] For more discussion of these terms, see also Elena Gabellini, 'La comodità nel giudicare: la decisione robotica' (n 70); Elena Gabellini, 'Algoritmi decisionali e processo civile: limiti e prospettive' (n 70).

[73] Antonio Carratta (n 1).

[74] Vittorio Villa, 'Il problema della scienza giuridica' in Giorgio Pino, Aldo Schiavello, Vittorio Villa (ed), *Filosofia del diritto. Introduzione critica al pensiero giuridico e al diritto positive* (Giappichelli 2013) 387; Pierluigi Chiassoni, *Positivismo giuridico* (Mucchi 2013) 56; E Pattaro, 'Il positivismo giuridico italiano dalla rinascita alla crisi' in Uberto Scarpelli (ed), *Diritto e analisi del linguaggio* (Comunità 1976).

[75] On this point, see G. Pino, 'La gerarchia delle fonti del diritto. Costruzione, decostruzione, ricostruzione' (2011) 1 Ars Interpretandi <www.arsinterpretandi.it/2011-stato-contemporaneo-crisi/> accessed 7 June 2023; Nicolò Lipari, 'I civilisti e la certezza del diritto' (2015) 2 Ars Interpretandi <www.rivisteweb.it/doi/10.7382/82125> accessed 10 June 2023; Natalio Irti, *Un diritto incalcolabile* (Giappichelli 2017); Natalino Irti, 'La crisi della fattispecie' (2014) 1 ivista di diritto processuale.

[76] On this point, see Germana Lo Sapio, 'La black box: l'esplicabilità delle scelte algoritmiche quale garanzia di buona amministrazione' (2021) 16 Federalismi <www.federalismi.it/nv14/articolo-documento.cfm?artid=45610> accessed 26 June 2023; Giorgio Resta, 'Algoritmi, diritto e democrazia' (2019) 4 Giustiziacivile.com <https://giustiziacivile.com/soggetti-e-nuove-tecnologie/editoriali/algoritmi-diritto-democrazia> accessed 26 June 2023; Gherardo Carullo, 'Decisione amministrativa e intelligenza artificiale' (2021) 3 Diritto dell'informazione e dell'informatica 431, 461.

"*constitutional minimum of the motivation*"[77], the decision must contain "*la concisa esposizione delle ragioni di fatto e di diritto*" (Art. 132, n. 4 Code of Civil Procedure), consisting of "*nella succinta esposizione dei fatti rilevanti della causa e delle ragioni giuridiche della decisione, anche con riferimento a precedenti conformi*" (Art. 118, 1 provisions implementing the Code of Civil Procedure). In addition, "*debbono essere esposte concisamente e in ordine le questioni discusse e decise dal collegio ed indicati le norme di legge e i principi di diritto applicati*" (Art. 118, co. 2 disp. att. Code of Civil Procedure).

The output of the algorithm, therefore, cannot be adequately supported by logical legal reasoning. This alone would make the use of predictive justice systems in decision-making incompatible with the Italian framework.

The opacity of the algorithm also creates problems for the impugnation of the judgment. A decision that does not state the logical and legal reasons behind it empties of content the instruments made available by the legislature to ask for the reform of a judgment that is considered unjust or illegitimate.

On this point, Art. 342 Code of Civil Procedure states that for each of the grounds of appeal, we must indicate "*a pena di inammissibilità, in modo chiaro, sintetico e specifico*" and also "*le censure proposte alla ricostruzione dei fatti compiuta dal giudice di primo grado*" and "*le violazioni di legge denunciate e la loro rilevanza ai fini della decisione impugnata*". It is therefore self-evident to say that if the machine does not provide any motivation, the losing party will be at root precluded from denouncing violations of the law and errors to the reconstruction of the facts.

However, these issues could be overcome once again by placing these instruments outside the jurisdictional field by framing them as systems to support lawyers or as systems of ODR.

5.2.2 Independence of the judge

The predictive justice system in providing a solution to the judge influences the decision-making process[78]: this is the c.d. performative effect or self-realisation of the algorithm.[79] Precisely, this effect risks transforming the response of the algorithm from mere indication to prescription.[80]

---

[77] On these terms, see Giovanni Canzio, 'L'art. 111 della Costituzione, commi 6, 7 e 8' [2021] La Magistratura <https://lamagistratura.it/commentario/lart-111-sesto-settimo-e-ottavo-comma-della-costituzione/#:~:text=111%2C%20co.,Cassazione%20per%20violazione%20di%20legge> accessed 5 June 2023.

[78] Filomena Santagada(n 16).

[79] *See note*; Domenico Dalfino, 'Creatività e creazionismo, prevedibilità e predittività' (2018) 12 Il Foro italiano 385. See also Giuseppina Fanelli, 'L'impiego dell'intelligenza artificiale nei processi decisori del giudice, tra la disciplina europea e quella del processo civile' in Rosaria Giordano and others (eds), *Il Diritto nell'Era Digitale. Persona, Mercato, Amministrazione, Giustizia* (Giuffrè Francis Lefebvre 2022) 993.

[80] Filomena Santagada (n 16).

Further examining the question, we can observe friction with the principle of the independence of the judge (Art. 104, paragraph 1 Italian Constitution).

In the face of the risk of generating self-fulfilling prophecy[81] (the prediction that becomes the decision), which derives from the performative effect of the algorithm, lies the risk that the activities of software companies that implement predictive justice systems end up directing judicial activity, thereby undermining its independence.[82] Therefore, with AI decision-making systems, there is a risk of the subjugation of the judge to technology and no longer only to the law.

In this way, among other things, the principle of independence takes on a different connotation that goes beyond its function of safeguarding the judiciary from the other powers of the state, particularly from the government: it becomes a principle of protection against legaltech.[83]

5.2.3 Imputability of the decision

The imputability of the decision depends on the type of mathematical model underlying the predictive justice software.[84] A distinction can be made between expert systems[85] on the one hand and machine learning models on the other.

The first ones are those that rely on inferential mechanisms and follow the rule "if X then Y": upon the occurrence of a specific Condition X, then the machine will provide the answer Y. The software in these hypotheses uses reasoning of deductive type.[86] In these cases, the programmer must provide the machine with the logical rule.

Let us take an example: Art. 10 and ss. Code of Civil Procedure identifies the criteria for determining competence by value, matter and territory. To enable the algorithm to verify whether the trial has been instituted before a competent court, the rules for determining jurisdiction in value, matter and territory should be established for the programme by the programmer.

This is not a difficult task. It is sufficient that the programmer gives the machine the logical inference as a rule "if X then Y", which, after applying it, for example, to Art. 18 Code of Civil Procedure will be a rule of the type "if the defendant has residence in Rome, then the court of Rome will be competent". For the imputability of the decision in these

---

[81] Domenico Dalfino (n 79).

[82] Filomena Santagada (n 16).

[83] Domenico Dalfino, 'Stupidità (non solo) artificiale, predittività e processo' [2019] Questione Giustizia <www.questionegiustizia.it/articolo/stupidita-non-solo-artificiale-predittivita-e-processo_03-07-2019.php> accessed 7 June 2023.

[84] See Gherardo Carullo (n 76).

[85] Filomena Santagada (n 16). See also Piergiuseppe Otranto, 'Riflessioni in tema di decisione amministrativa, intelligenza artificiale e legalità' (2021) 7 Federalismi.it <www.federalismi.it/nv14/articolo-documento.cfm?Artid=45026&content=&content_author=> accessed 13 June 2023.

[86] On the deductive argument, recently, see Roberto Ciuni, Aldo Frigerio, 'Gli argomenti deduttivi' in Damiano Canale, Roberto Ciuni, Aldo Frigerio (eds), *Critical Thinking: An Introduction* (EGEA, 2021). See also Jonathan Bennet, *A philosophical guide to conditional* (Claredon Press 2003); Paolo Legrenzi, Armando Massarenti, *La buona logica* (Raffaello Cortina 2016).

cases, there is no doubt that the algorithm has received the logical rules for reasoning from humans, and therefore, the imputability of the decision can be referred to humans. The problem that is feared here is how the decision can be imputed to the judge who uses the AI system and not to the programmer who designed the software.

The second approach (machine learning models) poses major problems. These can be applied whenever the system has to decide according to the previous case law.[87]

These are models that do not decide through logical rules of inference imparted to the programme by a human programmer. Instead, these programmes, on the basis of a sufficiently representative dataset, generate a mathematical model. They extrapolate the rule by which, given that dataset, that specific response can be provided.

The machine basically replaces the programmer who, in expert systems, provides the rules for the software. In machine learning systems, in fact, it is the same software that, starting from the training set, extrapolates the rule through which it will then provide the subsequent outputs.

Therefore, since there is no human agent here to give the rules for decisions to the software, the problemme of the imputability of the output of the machine is much more evident. Ultimately, this aspect is connected to the black box problem.

6 Conclusions

From the legal panorama described above, we can identify some trends.

The most innovative trend, in terms of the legal regulations of artificial intelligence systems, is certainly the one adopted by the European Union with the AI Act. In fact, Europe is preparing to become the first legal system in the world to regulate the instruments of artificial intelligence, including those of predictive justice (Annex III, paragraph 8, lit. a).

In this context, the question that needs to be addressed is who will be in charge of developing (private or public) predictive justice software.

One initiative that does not seem to be within the scope of Italian institutions at the moment is the creation of a public database to be set up at the Ministry of Justice. We have seen that in those contexts where digital justice is most advanced (Estonia, France and China); a fundamental step has been to establish an open access database that contains all the judgments issued by the judicial authorities. It is therefore a necessary intermediate step, without which it will be very difficult to develop predictive justice tools.

In China, however, the situation seems to be the opposite. This is a context in which the development of digital justice software began years ago (it dates back to the second

---

[87] For this term, see Gherardo Carullo (n 76). See also Roberto Cavallo Perin, Isabella Alberti, 'Atti e procedimenti amministrativi digitali' in Roberto Cavallo Perin, Diana-Urania Galetta (eds), *Il diritto dell'amministrazione pubblica digitale* (Giappichelli 2020).

phase—2003-2014—when the first database containing judgments was created). What is lacking, however, is a legal regulation of these instruments, the use of which is encouraged by the central government, but the implementation of which is left to local courts.

Finally, there seems to be no such initiative in the US legal system in the civil field.

For Italy, from the projects in place at some courts of appeal, it seems that the intention is to approach other legal realities in which projects related to digital civil justice are already in an advanced stage.

In my opinion, the use of artificial intelligence systems in the jurisdictional activity needs to be gradually implemented. A first step could be introducing tools for deciding preliminary ritual issues, such as jurisdiction and competence: in this regard so-called "expert systems" could be use.

However, when it comes to the resolution of prejudicial and dependence issues the implementation of artificial intelligence systems becomes even more complex. Some issues arise from the difficulty of making the machine understand the meaning of technical words such as "prejudicial" and "dependent". Equally if not more challenging is programming of predictive justice tools used for the entire dispute decision.

In any case, a preliminary fundamental step is the creation of open access database that contains all the judgments issued by the judicial authorities. In this regard, it is necessary that its construction will be delegated to the public power in order to prevent private interests from influencing the decision of the jurisdictional authorities.