

The Cambridge Handbook of

# GENERATIVE AI AND THE LAW

EDITED BY

Mimi Zou, Cristina Poncibò,  
Martin Ebers and Ryan Calo

CAMBRIDGE

An abstract watercolor illustration on a yellow background. It features a dense pattern of small black dots. Overlaid on this are several grey washes of varying opacity, some forming circular shapes. There are also several bright blue splatters and three yellow triangles pointing downwards, arranged vertically on the right side of the image.



Shaftesbury Road, Cambridge CB2 8EA, United Kingdom  
 One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
 477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
 314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi - 110025, India  
 103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of Cambridge University Press & Assessment, a department of the University of Cambridge.

We share the University's mission to contribute to society through the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org  
 Information on this title: www.cambridge.org/9781009492584

DOI: 10.1017/9781009492553

© Cambridge University Press & Assessment 2025

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press & Assessment.

When citing this work, please include a reference to the DOI 10.1017/9781009492553

First published 2025

Cover image: "Codes" Gray and Black abstract watercolor painting by Bruce Black, 2020

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging-in-Publication Data

NAMES: Zou, Mimi, 1985- editor | Poncibò, Cristina editor | Ebers, Martin, 1970- editor | Calo, M. Ryan editor  
 TITLE: The Cambridge handbook of generative AI and the law / edited by Mimi Zou, University of New South Wales  
 Sydney; Cristina Poncibò, University of Turin; Martin Ebers, University of Tartu, Estonia; Ryan Calo, University of Washington.

OTHER TITLES: Cambridge handbook of generative Artificial intelligence and the law

DESCRIPTION: Cambridge, United Kingdom ; New York, NY : Cambridge University Press, 2025. | Series: Cambridge law handbooks | Includes bibliographical references and index. | Summary: "This volume is for lawmakers, lawyers, business leaders, and curious minds interested in the legal and regulatory implications of generative AI as the fast-evolving technology is transforming our world. It offers clear and nuanced insights into the challenges of AI for the law, helping readers navigate this complex landscape" - Provided by publisher.

IDENTIFIERS: LCCN 2025007716 (print) | LCCN 2025007717 (ebook) | ISBN 9781009492584 hardback | ISBN 9781009492553 epub

SUBJECTS: LCSH: Artificial intelligence--Law and legislation--United States | Artificial intelligence--Law and legislation  
 CLASSIFICATION: LCC KF390.5.C6 C36 2025 (print) | LCC KF390.5.C6 (ebook) | DDC 343.7309/998--dc23/eng/2  
 0250306

LC record available at <https://lcn.loc.gov/2025007716>

LC ebook record available at <https://lcn.loc.gov/2025007717>

ISBN 978-1-009-49258-4 Hardback

Cambridge University Press & Assessment has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

For EU product safety concerns, contact us at Calle de Jos'e Abascal, 56, 1º, 28003 Madrid, Spain, or email [eugpsr@cambridge.org](mailto:eugpsr@cambridge.org)

## Contents

List of Figures	page ix
List of Tables	xi
List of Contributors	xiii
Foreword	xv
Michael Kirby	xvii
Preface	
<b>PART I UNDERSTANDING GENERATIVE AI FROM MULTIDISCIPLINARY PERSPECTIVES</b>	
1 Generative AI: An Introduction	3
Tom Melham	
2 Unleashing Creative Potential: Nurturing Trustworthy Generative AI	11
Zijie Huang	
3 Normative and Ethical Dimensions of Generative AI: From Epistemological Considerations to Societal Implications	27
Ludovica Paseri and Massimo Durante	
4 Why Generative AI Is Not Cyrano de Bergerac: A Computational Manipulation Perspective on Generative AI	43
Stefano Faraoni	
5 Unnatural Selection? A Darwinian Reading of the Economic Consequences of Generative AI on the Art Market	65
Jerome De Cooman	
<b>PART II EVOLVING REGULATORY AND GOVERNANCE FRAMEWORKS</b>	
6 LLMs Meet the AI Act: Who's the Sorcerer's Apprentice?	85
Ugo Pagallo	

7	Mapping Generative AI Liability Cases in the EU Legal Framework Teresa Rodriguez de las Heras Ballell	98
8	Challenges for Foundation Model Liability and Regulatory Regimes: An Analysis of US Law Peter Henderson	122
9	Navigating China's Regulatory Approach to Generative AI Lu Zhang and Mimi Zou	134
10	Singapore's Evolving AI Governance Framework Jason Allen Grant and Jane Loo	151
11	Shaping Global AI Governance: A Path for the G7 to Foster Rule of Law in a World of Uncertainty Hiroki Habuka and David U. Socol de la Osa	169
12	Generative AI and International Standardisation Sebastian Hallensleben	199
13	Private Ordering and Generative AI: What Can We Learn from Foundation Model Terms and Conditions? Lilian Edwards, Igor Szpotakowski, Gabriele Cifrodelli, Joséphine Sangaré, and James Stewart	201
PART III GENERATIVE AI: SPECIFIC LEGAL QUESTIONS		
14	Generative AI and Non-discrimination Law in the EU Riccardo de Caria	221
15	Generative AI and Data Protection Hannah Ruschemeier	237
16	Generative AI and the Fundamental Limitations of US Privacy Law Elana Zaide	255
17	Generative AI and IP under US Law Bruce E. Boyden	270
18	Copyright and Generative AI in Japan and China Christoph Rademacher and Wanru Cai	300
19	Redefining Rivalry: Generative AI and the Evolving Landscape of Competition Law Sylvia Papadopoulos	317
20	Regulating Hypersuasion Cristina Poncibò	347

21	Generative AI Systems and Corporate Governance, Compliance, and Liability: Rethinking Director and Officer Roles in Light of a New World of Technological, Legal, and Ethical Challenges Patrick O'Malley	367
22	Generative AI and Criminal Guilt: When No One Meant to Harm Beatrice Panattoni	392
PART IV THE USE OF GENERATIVE AI IN LEGAL AND RELATED SECTORS		
23	Automating Legal Tasks: LLMs, Legal Documents, and the AI Act Rūta Liepiņa, Francesca Lagioia, Marco Lippi, Przemysław Pałka, Hans-Wolfgang Micklitz, and Giovanni Sartor	407
24	LawGPT: Generative AI and Legal Services Regulation Martin Ebers	425
25	Generative AI and Article 6 of the European Convention on Human Rights: The Right to a <i>Human Judge</i> ? Mimi Zou and Ellen Lefley	451
26	Generative AI in Public Administration Sophie Weerts	472

Moreover, one of the fundamental principles underpinning the EU AI Act is transparency and explainability,<sup>91</sup> where generative AI systems must be transparent about their capabilities, limitations, and potential impacts on users and society. Users should be informed when they are interacting with AI-generated content and provided with explanations of how the content was generated.

Furthermore, the EU AI Act emphasises the importance of human oversight and control in AI decision-making processes.<sup>92</sup> For generative AI systems, human oversight is essential to ensure that the outputs align with ethical and legal standards, particularly in sensitive domains such as content creation, healthcare, and finance. Users should have the ability to intervene, correct, or override AI-generated outputs when necessary, reducing the risk of unintended consequences or misuse.

Additionally, the EU AI Act introduces requirements for accountability and compliance with regulatory obligations. Developers and deployers of generative AI systems are responsible for ensuring that their systems comply with legal requirements, including those related to data protection, consumer rights, and safety. Overall, the EU AI Act is promoting ethical, transparent, and trustworthy AI practices in a broader area of AI in order to foster the craft of trustworthy AI.

In addition to the EU AI Act, the General Data Protection Regulation (GDPR) has broader implications for the development, deployment, and use of generative AI. For instance, the GDPR sets out fundamental data protection principles<sup>93</sup> that apply to the processing of personal data. Generative AI systems that utilise personal data for training or operation must comply with these principles to ensure that data subjects' rights are respected and protected. Moreover, generative AI systems must have a lawful basis for processing personal data under the GDPR. Developers and deployers of generative AI systems must ensure that they have a valid legal basis for processing personal data in accordance with the GDPR requirements. Furthermore, the GDPR mandates that personal data be processed in a manner that ensures appropriate security and confidentiality, including protection against unauthorised or unlawful processing and accidental loss, destruction, or damage. Generative AI systems must implement technical and organisational measures to safeguard personal data from unauthorised access, disclosure, alteration, or destruction, in line with the GDPR's security requirements. Compliance with the GDPR is essential for building trust in generative AI and promoting responsible data-handling practices in the AI ecosystem.

## 2.6 CONCLUSION

We need to prioritise the development of trustworthy generative AI to ensure that the powerful technologies are designed and used in ways that promote fairness, transparency, and accountability. The future of generative AI holds great promise, but it also poses significant risks if not handled responsibly. Some of the potential future challenges we may face include the need to address increasing concerns around bias and discrimination, the rise of AI-generated deepfakes and disinformation, and the potential for generative AI to disrupt traditional employment and social structures. Regulatory measures that are underlined by these key principles serve as crucial safeguards for building trust in generative AI and AI systems generally.

<sup>91</sup> Transparency Obligations. [www.euaiact.com/key-issue/5#:~:text=Article%2013%20of%20the%20EU%20AI%20Act%20provides%20the%20requirement,providers%20and%20users%20to%20reasonably%20aim%20at%20persist%20notwithstanding%20the%20application](https://www.euaiact.com/key-issue/5#:~:text=Article%2013%20of%20the%20EU%20AI%20Act%20provides%20the%20requirement,providers%20and%20users%20to%20reasonably%20aim%20at%20persist%20notwithstanding%20the%20application)

<sup>92</sup> Article 14: Human Oversight. <https://artificialintelligenceact.eu/article/14#:~:text=Human%20oversight%20shall>

<sup>93</sup> 'A guide to the data protection principles.' <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/>

## Normative and Ethical Dimensions of Generative AI From Epistemological Considerations to Societal Implications

Ludovica Paseri and Massimo Durante

### 3.1 INTRODUCTION

Since the summer of 1956, when John McCarthy, Marvin Minsky, Nathaniel Rochester, and Claude Shannon presented their project on artificial intelligence (AI) at a conference at Dartmouth College in Hanover,<sup>1</sup> much has happened. Starting 'on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it',<sup>2</sup> these pioneers paved the way for the following decades of research in this field. While AI originally referred to a new discipline set out to investigate the ways in which machines could simulate human intelligence, today, after numerous winters and rebirths of AI, this field of research has been profoundly changed, definitions have multiplied, and the evolution of AI has called for new classifications and more accurate taxonomies.<sup>3</sup>

From the legal perspective, the European Regulation on artificial intelligence (AI Act)<sup>4</sup> defines an AI system as:

a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.<sup>5</sup>

This definition encompasses two main *characteristics*, namely autonomy and adaptability, which abstract from direct reference to intelligence but characterise a process that allows the AI system to map the relationship between input and output and thereby learn to generate results that can

Ludovica Paseri wrote Sections 3.1, 3.2, and 3.4; Massimo Durante wrote Sections 3.3 and 3.5.

<sup>1</sup> J. McCarthy, M. L. Minsky, N. Rochester, C. E. Shannon, 'A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955', *AI Magazine* 27.4, 2006, p. 12.

<sup>2</sup> *Ibid.*

<sup>3</sup> N. Cristianini, *The shortcut: why intelligent machines do not think like us*, CRC Press, Boca Raton, FL, 2023; L. Floridi, *The ethics of artificial intelligence: principles, challenges, and opportunities*, Oxford University Press, Oxford, 2023.

<sup>4</sup> Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No. 300/2008, (EU) No. 167/2013, (EU) No. 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), 12.7.2024, ELI: <http://data.europa.eu/eli/reg/2024/1689/oj>.

<sup>5</sup> Art. 3(1) AI Act.

have an influence on the environment in physical or virtual terms. Outputs are exemplified in (i) predictions, (ii) content, (iii) recommendations, or (iv) decisions.

Recent developments in the field of AI have also prompted European lawmakers to include a further definition of so-called general-purpose AI models, defined as follows:

an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market.<sup>6</sup>

In Executive Order No. 14110 on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, signed by US President Joe Biden on 30 October 2023,<sup>7</sup> the United States takes a distinct approach, proposing three different definitions of AI,<sup>8</sup> AI model,<sup>9</sup> and AI system.<sup>10</sup> In addition to these general definitions, the Executive Order also includes a specific mention of generative AI, defined as follows: 'the class of AI models that emulate the structure and characteristics of input data in order to generate derived synthetic content. This can include images, videos, audio, text, and other digital content'.<sup>11</sup>

The approaches taken by the United States and the European Union are different. In the first case, generative AI (Gen AI) is explicitly mentioned, of which large language models (LLM) are the best-known and most widely used example. In the latter case, the EU legislator ultimately decided not to explicitly reference Gen AI but, instead, to include general-purpose AI models as an example of which are the so-called large generative AI models, as specified in Recital 99 of the AI Act. Thus, our study focuses on Gen AI systems, examining their ability to 'automatically generate content in response to prompts written in natural-language conversational interfaces'.<sup>12</sup> In particular, the intention is to shed light on the ethical challenges and social implications raised by Gen AI. The issue is often addressed through a normative approach (typical, for instance, of European institutions), which aims to identify the advantages and disadvantages of applying a specific technology. However, the ethical and social challenges of Gen AI also require adopting an epistemological approach, which calls for examining the differences between the representation of reality of human beings and that of Gen AI.

Therefore, drawing attention to the normative approach (Section 3.2) and, in particular, to the uses (Section 3.2.1), abuses (Section 3.2.2), and under-uses (Section 3.2.3) of Gen AI, the

<sup>6</sup> Art. 3(63) AI Act.

<sup>7</sup> The White House, *Executive order on the safe, secure, and trustworthy development and use of artificial intelligence* no. 14110, 30 October 2023, available at [www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/](http://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/). Consider that the executive order has recently been overruled, see: Initial Rescissions Of Harmful Executive Orders And Actions, The White House, January 20, Washington D.C., USA, <https://www.whitehouse.gov/presidential-actions/2025/01/initial-rescissions-of-harmful-executive-orders-and-actions/>.

<sup>8</sup> In §3(b) Executive Ord., AI was defined as 'a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. Artificial intelligence systems use machine- and human-based inputs to perceive real and virtual environments; abstract such perceptions into models through analysis in an automated manner; and use model inference to formulate options for information or action'.

<sup>9</sup> In §3(c) Executive Ord., the AI model was defined as 'a component of an information system that implements AI technology and uses computational, statistical, or machine-learning techniques to produce outputs from a given set of inputs'.

<sup>10</sup> In §3(e) Executive Ord., the term AI system was defined as 'any data system, software, hardware, application, tool, or utility that operates in whole or in part using AI'.

<sup>11</sup> §3(p) Executive Ord.

<sup>12</sup> UNESCO, *Guidance for generative AI in education and research*, UNESCO Publishing, 2023, <https://unesco.org/ark:/48223/pf0000386693>, p. 8.

chapter focuses on the epistemological approach (Section 3.3), wondering what representation of reality Gen AI produces and how we interact with this representation. Four epistemological challenges of Gen AI are identified in terms of qualification (Section 3.4.1); reliability (Section 3.4.2); pluralism and novelty (Section 3.4.3); and technological dependence (Section 3.4.4). In light of these four epistemological challenges, the social implications of Gen AI are framed, proposing some final remarks (Section 3.5).

### 3.2 NORMATIVE APPROACH

The question of normativity, applied to the impact of digital technologies on society, investigates how this impact shapes or redesigns our society,<sup>13</sup> identifying advantages and disadvantages in facing the challenges posed by new technologies. This approach, typically adopted by European institutions, is also generally applied to the analysis of the impact of AI and, in particular, Gen AI on our society. This investigation allows us to highlight the knots and potential risks of harm that arise in this context, further distinguishing the harms resulting from uses (Section 3.2.1), misuses or abuses (Section 3.2.2) and underuses (Section 3.2.3) of Gen AI.

#### 3.2.1 Uses

Against some alarmist and sensationalist representations, it must be recognised that 'while generative AI may be new, its harms are not'.<sup>14</sup> By this we do not intend to underestimate the dangers of using Gen AI but to put them into proper perspective. For instance, what is groundbreaking is that 'today, artificial intelligence (AI) manages the properties of electromagnetism to process texts with extraordinary success and often with outcomes that are indistinguishable from those that human beings could produce'.<sup>15</sup> In addition to the case where there is a specific intention to generate harm through Gen AI, the ordinary use of such systems can entail three main risks: (i) concentration of power; (ii) perpetuation of bias and discrimination over time; and (iii) technological solutionism.

As regards the concentration of power, the recent and increasing focus on Gen AI has shifted the attention from research, transparency, sharing, and experimentation in the field of AI<sup>16</sup> to

<sup>13</sup> M. Durante, 'Normativity, constructionism, and constraining affordances', *Etica & Politica* 13, 2011, p. 180. See also M. Durante, 'How does the evolution of ICTs change the law? An approach to law through the philosophy of information of Luciano Floridi', in M. Arias-Oliva, T. Ward Bynum, S. Rogerson, T. Torres-Coronas (eds.), *The 'backwards, forwards and sideways' changes of ICT*, University Rovira I Virgili Press, Tarragona, 2010, pp. 136-144.

<sup>14</sup> A report by the Electronic Privacy Information Centre (EPIC) provides a classification of the types of harm caused by Gen AI, including physical harms; economic harms; reputational harms; psychological harms; autonomy harms; discrimination harms; relationship harms; loss of opportunity; social stigmatisation; and dignitary harm. See G. Fergusson et al., 'Generating harms: generative AI's impact & paths forward', *Electronic Privacy Information Center*, 2023, <https://epic.org/documents/generating-harms-generative-ais-impact-paths-forward/>, pp. 65-66.

<sup>15</sup> L. Floridi, 'AI as agency without intelligence: on ChatGPT, large language models, and other generative models', *Philosophy & Technology* 36, 2023, p. 1.

<sup>16</sup> Consider the scenario described in M. Hutson, 'Artificial intelligence faces reproducibility crisis', *Science*, 359.6377, 2018, p. 725: 'The most basic problem is that researchers often don't share their source code. At the AAAI [Association for the Advancement of Artificial Intelligence] meeting, Odd Erik Gundersen, a computer scientist at the Norwegian University of Science and Technology in Trondheim, reported the results of a survey of 400 algorithms presented in papers at two top AI conferences in the past few years. He found that only 6% of the presenters shared the algorithm's code. Only a third shared the data they tested their algorithms on, and just half shared "pseudocode" - a limited summary of an algorithm. (In many cases, code is also absent from AI papers published in journals, including *Science* and *Nature*).' Although 'there are many reasons for the missing details', the percentages highlight an environment that is not prone to sharing.

profit, opacity, and, consequently, the concentration of power.<sup>17</sup> Consider, for example, that in the original AI Act proposal,<sup>18</sup> there was no exception for research purposes, with effects described in the literature as 'detrimental to the entire scientific research ecosystem'.<sup>19</sup> Secondly, the risk of perpetuating bias and discrimination over time can also be highlighted here. Gen AI relies on its ability to collect (often, scrape) and process huge amounts of data from multiple sources.<sup>20</sup> Personal data is also highly involved: recently, many leading digital companies have changed their privacy policies to expand the amount of data processed for training, as recently warned by the US consumer protection agency, the Federal Trade Commission (FTC).<sup>21</sup> Once again, the processing of big data, including personal data, is reinforcing forms of discrimination perpetuated over time, crystallising vulnerability and inequality. Even if there were no specific intention to discriminate against or harm a particular category, imprisoning this vulnerable category in a representation that draws only from the past prevents it from evolving and thriving.<sup>22</sup>

Third, the strong advancement of Gen AI raises a risk of technological solutionism, which represents a revival of forms of techno-determinism that marked the early years of the digital revolution.<sup>23</sup> Recently, the Stockholm Declaration on AI for Science has been released, stating that 'by harnessing the power of AI we can propel humanity toward a future where groundbreaking achievements in science, even achievements worthy of a Nobel Prize, can be fully automated. We believe that this is achievable by the year 2050'.<sup>24</sup> This statement seems to echo the

<sup>17</sup> P. Verdegem, 'Dismantling AI capitalism: the commons as an alternative to the power concentration of Big Tech', *AI & Society*, 39, 2024, pp. 727–737. See also P. Nemitz, 'Constitutional democracy and technology in the age of artificial intelligence', *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376.2133, 2018, p. 3, claiming that 'The accumulation of digital power, which shapes the development and deployment of AI as well as the debate on its regulation, is based on four sources of power'.

<sup>18</sup> 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. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>.

<sup>19</sup> M. Ebers et al., 'The European Commission's proposal for an Artificial Intelligence Act – a critical assessment by members of the Robotics and AI Law Society (RAILS)', *Multidisciplinary Scientific Journal* 4.4, 2021, p. 591. The latest version of the AI Act appears to be different. Article 2(6) states: 'This Regulation does not apply to AI systems or AI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development'.

<sup>20</sup> On the phenomenon of 'data scraping' or 'web scraping' see J. Ciani Sciolla, 'The normative challenges of data scraping: legal hurdles and steps forward', *i-lex* 16.2, 2023, p. ii; M. A. Khder, 'Web scraping or web crawling: state-of-the-art, techniques, approaches and application', *International Journal of Advances in Soft Computing & Its Applications* 13.3, 2021, pp. 144–168.

<sup>21</sup> 'It may be unfair or deceptive for a company to adopt more permissive data practices – for example, to start sharing consumers' data with third parties or using that data for AI training – and to only inform consumers of this change through a surreptitious, retroactive amendment to its terms of service or privacy policy': Federal Trade Commission Office of technology and the division of privacy and identity protection, 'AI (and other) companies: Quietly changing your terms of service could be unfair or deceptive', *Technology Blog*, 13 February 2024, [www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/02/ai-other-companies-quietly-changing-your-terms-service-could-be-unfair-or-deceptive](http://www.ftc.gov/policy/advocacy-research/tech-at-ftc/2024/02/ai-other-companies-quietly-changing-your-terms-service-could-be-unfair-or-deceptive).

<sup>22</sup> On the ability to amplify disadvantages and bias in AI applications in forensic psychiatry and criminal justice see L. Testa, 'Beyond discrimination: generative AI applications and ethical challenges in forensic psychiatry', *Frontiers in Psychiatry* 14, 2024, p. 7: 'the application of generative AI models in critical tasks that encompass life-altering outcomes, such as risk assessment, sentencing recommendation and treatment and rehabilitation planning, will not only reiterate but significantly magnify existing biases, exacerbating discrimination against forensic psychiatry patients, particularly those from historically marginalised groups, and reinforcing the stigma they experience across multiple levels of society'.

<sup>23</sup> U. Pagallo, *Il diritto nell'età dell'informazione: il riposizionamento tecnologico degli ordinamenti giuridici in complessità sociale, lotta per il potere e tutela dei diritti*, Giappichelli Editore, Torino, 2014, pp. 20–21. See also M. Dafoe, 'On technological determinism: a typology, scope conditions, and a mechanism', *Science, Technology & Human Values* 40.6, 2015, pp. 1047–1076.

<sup>24</sup> The Stockholm Declaration on AI for Science, <https://sites.google.com/view/stockholm-declaration/#h.t64rqc0y0p626.8000>, 2024, pp. 716–716.

Anderson's controversial 2008 article arguing that big data would make the scientific method obsolete.<sup>25</sup> As critically underlined in discussing Anderson's position, 'knowledge is more than information, because it requires explanations and understanding, not just truths or correlations'.<sup>26</sup> Similarly, as far as the Stockholm Declaration is concerned, it is fair to say that automation is not a value or an achievement in itself,<sup>27</sup> and, above all, to believe that scientific discoveries that address the major challenges of our time can be fully automated reveals a profound misunderstanding of the epistemic dimension of AI.

The context further thickens when the spotlight is on the abuses and misuses of Gen AI, which are explored in the following section.

### 3.2.2 Misuses and Abuses

Primarily because of its ability to blur the boundaries between true and false, Gen AI can become an exceptionally effective tool for generating harm. Specifically, three forms of harm can result from abuses and misuses of Gen AI: (i) disinformation; (ii) fraud; and (iii) manipulation.

Among the various forms of information distortion (from misinformation to malinformation, etc.), disinformation, broadly understood, is specifically driven by the malice of an actor who intends to make false information go viral and take advantage of it.<sup>28</sup> The role of Gen AI in this scenario is highly relevant because it increasingly challenges our ability to distinguish historical from synthetic data and, therefore, true from false. The harm that Gen AI abuses can perpetrate varies and involves both the private and public dimensions. Concerning the private dimension, an example is provided by the case of the high school students in New Jersey who shared AI-generated fake pornographic images of classmates.<sup>29</sup> Regarding the public sphere, consider the case of the fake recording, gone viral, of a candidate in the Slovakia elections in which he appeared to be claiming to have rigged the elections.<sup>30</sup> While these forms or attempts to distort reality have always existed, what makes the difference at present is the *qualitative* aspect of the distorted representation enacted by AI.

Besides a qualitative increase, there is also a quantitative increase in the harms occurring from the misuses of Gen AI. Recently, the FTC released some particularly disconcerting data: 'consumers reported losing more than \$10 billion to fraud in 2023, marking the first time that fraud losses have reached that benchmark. This marks a 14% increase over reported losses in

<sup>25</sup> C. Anderson, 'The end of theory: data deluge makes the scientific method obsolete', *Wired Magazine*, 23 June 2008.

<sup>26</sup> L. Floridi, *The fourth revolution: how the infosphere is reshaping human reality*, Oxford University Press, Oxford, 2014, p. 130.

<sup>27</sup> On the limits of automation see P. Nokelainen, T. Nevalainen, K. Niemi, 'Mind or machine? Opportunities and limits of automation', in C. Harteis (ed.), *The impact of digitalization in the workplace: an educational view*, Springer, Dordrecht, 2018, pp. 13–24. With specific reference to the health sector, see G. Lippi, G. Da Rin, 'Advantages and limitations of total laboratory automation: a personal overview', *Clinical Chemistry and Laboratory Medicine* 57.6, 2019, pp. 802–811.

<sup>28</sup> On disinformation and other forms of distortion of the information communication see O. Pollicino, 'The European approach to disinformation: comparing supranational and national measures', *Annuario di diritto comparato e di studi legislative*, 2020, pp. 175–212; O. Pollicino (ed.), *Freedom of speech and the regulation of fake news*, Intersentia, Cambridge, 2023.

<sup>29</sup> J. Jargon, 'Fake nudes of real students cause an uproar at a New Jersey high school', *The Wall Street Journal*, 2 November 2023, [www.wsj.com/tech/fake-nudes-of-real-students-cause-an-uproar-at-a-new-jersey-high-school-df10f1bb](http://www.wsj.com/tech/fake-nudes-of-real-students-cause-an-uproar-at-a-new-jersey-high-school-df10f1bb).

<sup>30</sup> C. Devine, D. O'Sullivan, S. Lyngaas, 'A fake recording of a candidate saying he'd rigged the election went viral. Experts say it's only the beginning', *CNN*, 1 February 2024, <https://edition.cnn.com/2024/02/01/politics/election-deepfake-threats-invs/index.html>.

2022'.<sup>31</sup> The part played by AI in this scenario is remarkable, so much so that the Norwegian Authority in Consumer Protection, in a report on the impact of Gen AI, stated that 'Generative AI models can be abused by malicious actors to augment or supercharge criminal activities'.<sup>32</sup> In response to these challenges, the US FTC announced a 'Voice Cloning Challenge to help promote the development of ideas to protect consumers from the misuse of artificial intelligence-enabled voice cloning for fraud and other harms'.<sup>33</sup>

Furthermore, an improper use of Gen AI can result in 'automated, effective manipulation at scale'.<sup>34</sup> In addition to the harms caused by disinformation, it is also necessary to consider the harms of manipulation since, although the concepts are often considered together and sometimes even represented as overlapping, they have dissimilar meanings and implications.

Whereas in the case of disinformation, a plausible but false scenario is provided (which may trigger different reactions), in the case of manipulation, the misrepresentation of a scenario is intended to induce people to act differently from how they would have acted in that specific circumstance, in order to achieve a certain outcome, thereby altering the individual's self-determination. While the primary intent of disinformation is the dissemination of false information, the primary intent of manipulation is to exert a form of control over the manipulated individual that leads to a desirable outcome. This is why Gen AI can be nimbly used to manipulate consumers<sup>35</sup> or, in society at large, conspiracy theories can be sharpened by providing plausible findings to support any position.<sup>37</sup>

With regard to the harms resulting from Gen AI abuse and misuse, it is crucial to consider the effects. There are three notable consequences that are crucial to grasping and measuring the extent of these forms of harm: (a) a general lack of trust; (b) threats to democracy and human rights; and (c) security dangers.

Distorted uses of Gen AI have been described as causing a 'disintegration of trust'.<sup>38</sup> Not surprisingly, the AI Act speaks of a 'need to build trust',<sup>39</sup> and the US executive order on AI, with specific reference to the education sector, advocated the need to develop an "AI toolkit" for

<sup>31</sup> Federal Trade Commission, 'As nationwide fraud losses top \$10 billion in 2023, FTC steps up efforts to protect the public', *FTC Press Release*, 9 February 2024, [www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public](http://www.ftc.gov/news-events/news/press-releases/2024/02/nationwide-fraud-losses-top-10-billion-2023-ftc-steps-efforts-protect-public).

<sup>32</sup> Forbrukerrådet, *Ghost in the machine. Addressing the consumer harms of generative AI*, June 2023, <https://storage.googleapis.com/forbrukerradet.no/media/2023/06/generative-ai-rapport-2023.pdf>, p. 32. The quantitative increase is emphasised in the report: 'Large language models can be used by scammers to generate a large amount of convincing-looking text to deceive victims. Similarly, catfishing scams, where the scammer builds trust with the victim over time through regular contact, can potentially be automatized convincingly by the use of advanced chatbots. This means that the criminal can effectively scam more victims using less time and resources' (p. 32).

<sup>33</sup> Federal Trade Commission, 'FTC announces exploratory challenge to prevent the harms of AI-enabled voice cloning', *FTC Press Release*, 16 November 2023, [www.ftc.gov/news-events/news/press-releases/2023/11/ftc-announces-exploratory-challenge-prevent-harms-ai-enabled-voice-cloning](http://www.ftc.gov/news-events/news/press-releases/2023/11/ftc-announces-exploratory-challenge-prevent-harms-ai-enabled-voice-cloning).

<sup>34</sup> M. Klenk, 'Ethics of generative AI and manipulation: a design-oriented research agenda', *Ethics and Information Technology*, 26.9, 2024, p. 1.

<sup>35</sup> For instance, J. E. Fossum, 'The context of fake news, disinformation, and manipulation', in M. Conrad et al. (eds.), *Europe in the age of post-truth politics: populism, disinformation and the public sphere*, Springer, Cham, 2022, pp. 31–52; E. Culloty, J. Suiter, *Disinformation and manipulation in digital media: information pathologies*, Routledge, London, 2021.

<sup>36</sup> Forbrukerrådet, *Ghost in the machine*, p. 22.

<sup>37</sup> S. Gabriel et al., 'Generative AI in the era of "alternative facts"', in D. Huttenlocher, A. Ozdaglar (eds.), *An MIT exploration of generative AI: from novel chemicals to opera*, MIT Press, Cambridge, MA, pp. 1–24.

<sup>38</sup> Fergusson et al., 'Generating harms', p. 64.

<sup>39</sup> Recital 6 AI Act.

education leaders implementing recommendations from the Department of Education's AI ... designing AI systems to enhance trust'.<sup>40</sup>

Furthermore, the consequences of such harms can pose threats to both democracy (consider, for instance, manipulation of the election process) and the protection of human rights. A recent report by the United Nations Human Rights Office of the High Commissioner, in addition to presenting a taxonomy of human rights risks connected to Gen AI,<sup>41</sup> also points out that 'we can expect that human rights will be threatened in new ways',<sup>42</sup> thereby opening future discussions.

Finally, the harms identified above can have consequences for the security and safety of both individuals and society. The topic is highly researched,<sup>43</sup> and it is significant that not only institutions but also companies are paying close attention to the issue.<sup>44</sup> For instance, the Google report 'Cybersecurity Forecast 2024. Insights for future planning' says that 'In 2024, the rapidly evolving world of gen AI will provide attackers with new ways to conduct convincing phishing campaigns and information operations at scale'.<sup>45</sup> However, it is worth noting that it also states that 'defenders will use the same technologies to strengthen detection, response, and attribution of adversaries – and more broadly reduce toil, address threat overload, and close the widening skills gap'.<sup>46</sup> This consideration, therefore, reminds us that in providing a comprehensive analysis of the socio-ethical impact of Gen AI adopting the normative approach, it is necessary not only to look at harms from use and misuse but also to draw attention to cases of underuse, which are analysed in the following section.

### 3.2.3 Underuses

The term underuse refers to the lack of exploitation of the potential of AI in a specific domain, leading to missed opportunities, inefficiencies, and related costs. In the literature, beginning with studies on AI underuse in the health sector,<sup>47</sup> the issue has also been analysed with regard

<sup>40</sup> §8c(iii)(d) Executive Ord.

<sup>41</sup> United Nations Human Rights Office of the High Commissioner, 'Taxonomy of human rights risks connected to generative AI', Supplement to B-Tech's Foundational Paper on the Responsible Development and Deployment of Generative AI, 2023, pp. 1–23, [www.ohchr.org/sites/default/files/documents/issues/business/b-tech/taxonomy-GenAI-Human-Rights-Harms.pdf](http://www.ohchr.org/sites/default/files/documents/issues/business/b-tech/taxonomy-GenAI-Human-Rights-Harms.pdf).

<sup>42</sup> *Ibid.*, p. 17.

<sup>43</sup> Consider, ex multis, K. Huang, *Generative AI security: theories and practices*, Springer Nature, Cham, 2024; M. A. Ferrag, M. Amine, M. Debbah, M. Al-Hawawreh, 'Generative AI for cyber threat-hunting in 6G-enabled IOT networks', 2023 *IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)*, IEEE, 2023, pp. 16–25; S. Weedon, 'Generative AI: security implications for business automation', *Network Security* 9, 2023, pp. 1353–4858.

<sup>44</sup> On different regulatory approaches to cybersecurity and Gen AI: C. Novelli, F. Casolari, P. Hacker, G. Spedicato, L. Floridi, 'Generative AI in EU law: liability, privacy, intellectual property, and cybersecurity', 2024, pp. 22–27, <https://arxiv.org/pdf/2401.07348>.

<sup>45</sup> Google Cloud, *Cybersecurity Forecast 2024. Insights for future planning*, 2024, p. 17, <https://cloud.google.com/resources/security/cybersecurity-forecast>.

<sup>46</sup> *Ibid.*

<sup>47</sup> See U. Pagallo et al., 'The underuse of AI in the health sector: opportunity costs, success stories, risks and recommendations', *Health and Technology* 14.1, 2024, pp. 1–14; U. Pagallo, 'The politics of data in EU law: will it succeed?', *Digital Society* 1.3, 2022, pp. 1–20; U. Pagallo, *Il dovere alla salute: sul rischio di sottoutilizzo dell'intelligenza artificiale in ambito sanitario*, Mimesis, Milano-Udine, 2022, p. 64; U. Pagallo, 'La grande trasformazione. Datificazione della società, tutela dell'ambiente e rischi e opportunità dell'innovazione tecnologica', in M. Durante, U. Pagallo, *La politica dei dati: il governo delle nuove tecnologie tra diritto, economia e società*, Mimesis, Milano-Udine, 2022, pp. 123–140; and for a comment see M. Orofino, 'La questione del sotto utilizzo dell'intelligenza artificiale in campo sanitario: spunti di rilievo costituzionale', *Queste istituzioni*, 4, 2022, pp. 158–171.

to the underuse of research data in the scientific research sector,<sup>48</sup> the environmental challenges of AI,<sup>49</sup> and public-sector data.<sup>50</sup>

Like the underuse of AI in the healthcare system, which gives rise to severe opportunity costs for the public, cases of harm resulting from the underuse of Gen AI can also be outlined. These forms of harm can be represented by loss of opportunity cost, loss of innovation, and inequitable distribution of benefits. First, it should be recalled that there are multiple benefits to using Gen AI.<sup>51</sup> In terms of opportunity costs, consider the effective potential strained to process efficiencies, optimisation, and efficiency gains. Regarding the research and education sector, while recognising that Gen AI 'harbours risks, such as the large-scale generation of disinformation and other unethical uses with significant societal consequences',<sup>52</sup> it also 'provides many opportunities'.<sup>53</sup> Second, the underuse of Gen AI can lead to a loss or delay in innovation as well as form of inefficiency in generating new ideas.<sup>54</sup> Third, cases of Gen AI underutilisation may exacerbate already disadvantageous positions. The problems related to the digital divide are not amplified by generating unequal situations between those who have access to Gen AI and those who do not. In addition, 'The rapid pervasion of Gen AI in technologically advanced countries and regions has exponentially accelerated the generation and processing of data and has simultaneously intensified the concentration of AI wealth in the Global North',<sup>55</sup> exacerbating forms of inequity.

Having made this analysis from a normative point of view, it should be noted that a further identification and evaluation of the risks and opportunities of Gen AI depends both on how the form of AI works from an epistemological point of view and on our ability to interact with it. For this reason, the following section focuses on the epistemological approach.

### 3.3 EPISTEMOLOGICAL APPROACH

Gen AI systems produce what can be described as pieces of knowledge<sup>56</sup> (text, audio, image, and video), offering a specific representation of reality. From a qualitative point of view, Ge

<sup>48</sup> L. Paseri, 'Sharing and reuse of research data in EU: the risk of over regulation and its implications', in M. Palmieri (ed.), *Artificial Intelligence and the Complexity of Legal Systems (AICOL) volume*, Springer, Cham, forthcoming.

<sup>49</sup> U. Pagallo, J. Ciani Sciolla, M. Durante, 'The environmental challenges of AI in EU law: lessons learned from the Artificial Intelligence Act (AIA) with its drawbacks', *Transforming Government: People, Process and Policy* 16.3, 2022, pp. 359–376.

<sup>50</sup> C. V. Buttow, S. Weerts, 'Managing public sector data: national challenges in the context of European Union's data governance models', *Information Polity* 29.3, 2023, pp. 1–16.

<sup>51</sup> Although in this chapter we focus specifically on Gen AI issues, we are aware of the benefits and opportunities that can arise from both the use of Gen AI and, more broadly, of synthetic data, in different areas that include but are not limited to law. Consider, for instance, the scientific research sector: S. Kurapati, L. Gilli, 'Synthetic data: convergence between Innovation and GDPR', in L. Paseri (ed.), 'Special issue on Open Science and Data Protection', *J. Open Access L.* 11(1), 2023, pp. 1–12.

<sup>52</sup> European Commission, DG for Research and Innovation, *Living guidelines on the responsible use of generative AI in research*, Publications Office of the European Union, Luxembourg, 2024, p. 3.

<sup>53</sup> *Ibid.*

<sup>54</sup> U. Pagallo, M. Durante, 'The good, the bad, and the invisible with its opportunity costs: introduction to the "J" special issue on "The impact of artificial intelligence on law"', *Multidisciplinary Scientific Journal* 5.1, 2022, p. 147.

<sup>55</sup> U. Pagallo, *The laws of robots: crimes, contracts, and torts*, Springer, Dordrecht, 2013.

<sup>56</sup> UNESCO, *Guidance for generative AI in education and research*, p. 14.

The expression 'pieces of knowledge' has often been used in computer science studies of knowledge representation. See for example D. B. Lenat, E. A. Feigenbaum, 'On the thresholds of knowledge', *Proceedings of the International Workshop on Artificial Intelligence for Industrial Applications*, 25–27 May 1988, 87, 1988, pp. 1172–1176, [www.iaai.org/Proceedings/87-2/Papers/122.pdf](http://www.iaai.org/Proceedings/87-2/Papers/122.pdf); J. Kelemen, J. Hvorecký, 'On knowledge, knowledge systems, and knowledge management', *Proceedings of the 9th International Conference on Computational Intelligence and Information*

AI's representation of reality reaches high standards and, as a result, discerning between synthetic and non-synthetic outputs becomes challenging for humans. Accordingly, apart from normative considerations, it becomes essential to scrutinise the representation of reality offered by Gen AI and how human beings become subjects of interaction, since they create machines that make decisions, artificial agents that act autonomously, and programs that are endowed with ever-increasing computational power that produce significant consequences for human life.<sup>57</sup>

From an epistemological perspective, Gen AI and humans are indeed different. While human beings are extraordinary 'semantic engines' (their understanding and representation of reality is rooted in semantic skills – i.e., the ability to give meaning and sense to the world), Gen AI is an extraordinary 'syntactic engine'<sup>58</sup> (operating on the basis of the ability to process vast amounts of data and parameters through structures of great complexity, such as neural networks, without possessing knowledge of what is being processed). Unlike the scientific method, where knowledge entails verifiable predictions based on models and explanations, AI-generated outputs lack such theoretical grounding. Thus, questioning the underlying functioning of Gen AI is pivotal due to its profound impact on human actions and the criteria for justifying them.

In doing so, we necessarily limit ourselves to considering the possible comparison with a specific epistemology that is the focus of our attention, legal epistemology. Without going into the age-old and thorny question of the definition of law and its role, we can say that, from an epistemological standpoint, the law acts as a normative system. It interprets and applies legal statements within a shared legal framework, verifying the reliability of this interpretation and application through forms of reasoning and specific tools (principles, categories, evidence, review systems, and so on) that the law itself provides and regulates. These have a fundamental common feature: every piece of knowledge produced by the interpretation and application of the law is offered to the scrutiny, control, and discussion of a counterparty and a third party (as well as more widely to the community of legal scholars over time).<sup>59</sup> Legal science is progressive and incremental and uses the contribution of all legal experts. This method, and the resulting representation of reality that it conveys, is in contrast with Gen AI's representation of reality for three main reasons.

First, the representation produced by Gen AI may happen to be accurate (where there is a validated standard or metric to measure accuracy), but it is not based on the ability to attribute meaning and intersubjectively verify its epistemic reliability. In more analytical terms, LLMs (and more widely Gen AI) do not have the ability to operate a semantic ascent<sup>60</sup> or descent in relation to artificially produced sentences; that is, they are not able to move from the linguistic to

Budapest Tech., Budapest, 2008, p. 28; D. Lenat, G. Marcus, 'Getting from generative AI to trustworthy AI: what LLMs might learn from CYC', 2023, pp. 1–21, <https://arxiv.org/pdf/2308.04445>.

<sup>57</sup> M. Durante, *Computational power: the impact of ICT on law, society and knowledge*, Routledge, London, 2021, p. 13.

<sup>58</sup> Floridi, *The ethics of artificial intelligence*, p. 44: 'with the staggering growth of available data, quantity and speed of calculation, and ever-better algorithms, they can do statistically . . . what we do semantically'.

<sup>59</sup> According to a neo-positivist conception of legal science that has also been called logical positivism, 'the paradigmatic instance of knowledge is represented by the empirical sciences, which adopt the principle of verification (or, from a certain point on, falsification). The latter thus becomes the fundamental methodological principle of every sphere of knowledge. Just as the proposition "this is chalk" implies that if one observes a piece of chalk at the microscope certain structural qualities will become apparent, similarly the proposition "Section 62 of the Uniform Negotiable Instruments Act is valid Illinois law" implies that Illinois courts, given certain conditions, will behave in a certain way': A. Schiavello, 'La scienza giuridica analitica dalla nascita alla crisi', *Ragion pratica* 1, 2020, p. 148 [our translation].

<sup>60</sup> This is an expression from W. V. Quine, *From a logical point of view*, Harvard University Press, Cambridge, MA, 1953, that we refer to only by analogy in this context. This reference to Quine can be found in L. Floridi, *The philosophy of information*, Oxford University Press, Oxford, 2013, p. 202.

Article 5 (1) (f). The Article does not apply to CM that causes non-significant harm, and in case, not to a scenario such as that of the e-cigarette.

Furthermore, the legislation referred to in the AIA, which, in the EU's opinion, should protect individuals when formal consent to personalisation and profiling is present. The DMA and the DMA leave websites, small platforms, or small businesses that use computational manipulation out of their scope. Moreover, Article 25 of the DSA does not apply to what is already protected by the GDPR and the UCPD. The UCPD and its Guidance leave consumers and CM<sub>2</sub> out of their scope.

Therefore, the existing EU legislation should be reconsidered to render it coherent, include non-consumers and the kind of CM able to covertly identify and exploit routes via LGAIM that is not subliminal and does not use deception or false information, undue influence, and that does not cause relevant harm but still impacts the right to mental determination of individuals.

## Unnatural Selection?

### *A Darwinian Reading of the Economic Consequences of Generative AI on the Art Market*

Jerome De Cooman

#### 5.1 INTRODUCTION

March 1884. The Supreme Court of the United States held in *Burrow-Giles* that a portrait photograph is copyrightable.<sup>1</sup> The Court explained that even if photography is by definition a mechanical process that takes a snapshot of reality, the author can still exercise control during the creative process.

December 2011. The European Court of Justice (ECJ) reached the same conclusion in *Painer*. A portrait photograph is protected by copyright if the photographer was able to make 'free and creative choices' during its production.<sup>2</sup>

February 2023. The US Copyright Office (USCO) held that the images contained within Kristina Kashtanova's comic book *Zarya of the Dawn* and drawn by a generative artificial intelligence (GenAI) were not copyrightable.<sup>3</sup> Referring to *Burrow-Giles*, the Office concluded that the user of a GenAI lacks 'sufficient control over the generated images.'<sup>4</sup>

April 2023. German photographer Boris Eldagsen declined the award for the Creative Open Competition Category of the 2023 Sony World Photography Award after having revealed it was an AI creation.<sup>5</sup> He explained that he willingly misled the jury in order to publicly engage in an open discussion on the future of photography.

What can be learnt from that background? First, *Burrow-Giles* and *Painer* teach us that the use of a machine in the creation of a work does not prevent it from being copyrighted as long as the human user is able to exercise a certain creative control over the creation. Second, the USCO decision related to *Zarya of the Dawn* suggests that users of GenAI do not reach the required threshold. Third, the declined photography award suggests that human-made and machine-

<sup>1</sup> *Burrow-Giles Lithographic Co. v. Sarony* [1884] 111 US 53, 60.

<sup>2</sup> Case C-145/10, *Eva-Maria Painer v. Standard Verlag GmbH and others*, ECLI:EU:C:2011:798, §94.

<sup>3</sup> United States Copyright Office, 'Zarya of the Dawn (Registration # VAu001480196)', 21 February 2023, <https://copyright.gov/docs/zarya-of-the-dawn.pdf>.

<sup>4</sup> *Ibid.*, 9.

<sup>5</sup> Matt Novak, 'Artist Reveals His Award-Winning "Photo" Was Created Using AI', *Forbes*, 17 April 2023, [www.forbes.com/sites/mattnovak/2023/04/17/artist-reveals-his-award-winning-photo-was-created-using-ai/?sh=3aa5e812f0fb](https://www.forbes.com/sites/mattnovak/2023/04/17/artist-reveals-his-award-winning-photo-was-created-using-ai/?sh=3aa5e812f0fb).

enabled artworks are indistinguishable.<sup>6</sup> This indistinguishability implies asymmetric information: artists do know whether the artwork is machine-enabled or human-made while buyers do not. George Akerlof seminally described asymmetric information as a 'lemons problem'. When buyers ('beholders') are unable to distinguish good-quality (known as 'peaches') from poor-quality (known as 'lemons') products (artworks), they assess quality on average. If sellers have an incentive to cheat, buyers internalise the risk of buying a poor-quality product at the price of a good-quality one. Therefore, the sale price of peaches is systematically lower than the value while the sale price of lemons is higher than what they are worth. Consequently, sellers of peaches withdraw from the market. This further erodes the average quality of available products and, eventually, the market collapses.

The art market constitutes fertile ground for the lemons problem, a fortiori with the advent of GenAI: since copyrightable human-made artworks and uncopyrightable machine-enabled artworks are indistinguishable, sellers of the latter are incentivised not to disclose whether AI was involved in the creation process.<sup>8</sup> This has economic consequences given that copyright is an exclusive property right that grants its owner a temporary monopoly.<sup>9</sup> Assuming therefore that the market value of a machine-enabled artwork (disclosed as such) is lower than the market value of a human-made one, art market buyers will internalise the average risk of being sold a machine-enabled artwork at the price of a human-made one. This, in turn, suggests that human-made artists cannot receive the true value associated with their artistic production.

This chapter proposes a Darwinian reading of this market inefficiency. If Charles Darwin had been faced with the lemons problem, he would probably have reached the same conclusion as Akerlof: the exclusion of one type of product from the market. He explained that 'if one species has any advantage whatever over another, it will in a very brief time wholly or in part supplant it'.<sup>10</sup> Asymmetric information constitutes such advantages that allows dishonestly sold machine-enabled products (the advantaged species) to drive human-made products (the other species) out of the market (one species supplants the other). Yet, for Darwin, this 'struggle for existence' refers to the process by which only the fittest species survive.<sup>11</sup> The survival of machine-enabled artworks at the expense of human-made art seems not in line with natural selection that is solely through the preservation of variations in some way advantageous, which consequently 'endure'.<sup>12</sup> In economic terms, natural selection is akin to competition on the merits. This means that 'firms must compete through the intrinsic qualities of their products rather than through extraneous conditions'.<sup>13</sup> The ECJ has confirmed that 'competition on the merits means, in definition, lead to the departure from the market or the marginalisation of competitors that are less efficient and so less attractive to consumers from the point of view of, among other things,

<sup>6</sup> This chapter endorses the terminology suggested by Jane C. Ginsburg and Luke Ali Budiardjo, 'Authenticity and Machines' (2020) 34(1) *Berkeley Technology Law Journal* 343–448 (who explained at 348n17 that it is more relevant to use 'enabled' rather than 'generated' to emphasise that a machine does not produce the output of its own volition but that it is a human who creates the output with the assistance of a machine).

<sup>7</sup> George Akerlof, 'The Market for "Lemons": Quality Uncertainty and the Market Mechanism' (1970) 84(3) *Quarterly Journal of Economics* 488–500.

<sup>8</sup> Jerome De Cooman, 'When Art Becomes a Lemon: The Economics of Machine-Enabled Artworks and the Need for a Rule of Origin' (2023) 5(1) *Law, Technology and Humans* 24–39.

<sup>9</sup> Justine Pila and Paul Torremans, *European Intellectual Property Law* (2nd ed., Oxford: Oxford University Press, 2018).

<sup>10</sup> Charles Darwin, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (London: John Murray, 1859) 402 (emphasis added).

<sup>11</sup> *Ibid.*, 60–79.

<sup>12</sup> *Ibid.*, 109.

<sup>13</sup> Harry S. Gerla, 'Competition on the Merits – A Sound Industrial Policy for Antitrust Law' (1984) 36(4) *University of Florida Law Review* 553–590, 558.

price, choice, quality or innovation'.<sup>14</sup> In other words, natural selection and competition on the merits both lead the more efficient species or products (respectively) to prevail while others vanish. This is not the outcome of the lemons problem. In the latter case, the selection of the prevailing species is more artificial than natural. There is no competition on the merits between human-made and machine-enabled artworks. The latter will only prevail because of asymmetric information (extraneous condition) and not due to its intrinsic qualities (one species did not prevail because it was most fitted to its environment but because of farm breeding).<sup>15</sup> The lemons problem provokes unnatural selection.

This is not inevitable, however. Darwin went on to explain that 'if both [species] are equally well fitted for their own places in nature, both probably will hold their own places and keep separate for almost any length of time'.<sup>16</sup> This chapter explores how to ensure both human-made and machine-enabled artworks are well-fitted for the art market. It argues that a rule of origin inspired by art authentication and EU customs and tax law is a way of ensuring that both types of artworks hold their own place and keep separate in their ecosystem. To build the argument, this chapter proposes a fivefold structure. Section 5.1 sets the scene by defining the lemons problem and transposing it on the art market. Section 5.2 suggests the lemons problem is an appropriate framework to evaluate the economic consequences of GenAI on that market. Section 5.3 recommends adopting a rule of origins to erase asymmetric information and suggests workable enforcement solutions. Section 5.4 summarises the argument and concludes that the proposed solution is not a Luddite, but a proportionate response to the (potential) economic consequences of AI on the art market.

## 5.2 THE LEMONS PROBLEM IN THE ART MARKET

Imagine that in the market for used cars there are only two different types of products: the good-quality used cars (in American slang, the peaches) and the bad-quality ones (the lemons).<sup>17</sup> Being in the used car market, sellers know whether the car being sold is a lemon or not. Buyers do not.<sup>18</sup> This asymmetric information incentivises dishonest sellers to sell a lemon at the price of a peach.<sup>19</sup> Buyers, in turn, will value a particular car based on the average probability that that car is a peach (assuming they are risk neutral) and adapt their willingness-to-pay accordingly. Therefore, the market price equilibrium will systematically be lower than what a peach is worth, but higher than what a lemon is worth.

This is market inefficiency par excellence: asymmetric information forces buyers to misallocate their resources, resulting in a systemic inefficiency that harms both buyers' and other sellers' welfare.<sup>20</sup> Under these conditions, the rational behaviour of sellers of peaches is either to leave the market, or to voluntarily decrease the quality of their products.<sup>21</sup> Perhaps more critical, the more sellers of peaches leave the market, the greater the probability of selling a lemon at the

<sup>14</sup> Case C-413/14 P, *Intel v. Commission*, ECLI:EU:C:2017:632, §134. See Nicolas Petit, 'The Judgement of the EU Court of Justice in Intel and the Rule of Reason in Abuse of Dominance Cases' (2018) 43(5) *European Law Review* 728–750.

<sup>15</sup> Darwin, *Origin of Species*, 446.

<sup>16</sup> *Ibid.*, 402.

<sup>17</sup> Akerlof, 'Lemons'.

<sup>18</sup> Jonathan Levin, 'Information and the Market for Lemons' (2001) 32(4) *The RAND Journal of Economics* 657–666.

<sup>19</sup> Charles Wilson, 'Adverse Selection' in John Eatwell, Murray Milgate and Peter Newman (eds.), *Allocation, Information and Markets* (Cham: Springer, 1989) 31–34.

<sup>20</sup> *Ibid.*, 12.

<sup>21</sup> Jack Hirschleifer, 'Where Are We in the Theory of Information' (1973) 63(2) *American Economic Review* 31–39, 37.

price of a peach and, therefore, the lower the market price equilibrium. This leads the seller to moderate-quality products to leave the market, hence further increasing the risk of dishonest deals and decreasing the market price equilibrium accordingly, until the market collapses. The ever-decreasing product quality and price suggest that, ultimately, the only market equilibrium price is null and therefore that no transaction occurs.<sup>23</sup>

Leaving aside the market for used cars, the lemons problem (if slightly adjusted) is also of interest for artworks. Identifying peaches and lemons among artworks is a tricky exercise. As the example par excellence of a market where quality is multidimensional, unquantifiable, and irreducibly subjective.<sup>24</sup> This explains why some people are willing to pay an astronomical price 'for what appears to be a watercolour sketch [that] was likely painted by Jackson Pollock [which] many more aesthetically pleasing works have sold for significantly less'.<sup>25</sup>

For the sake of this chapter, a line has to be drawn between *aesthetic value* and *market value*. This is not alien to the market for artworks: price reacts negatively when a work is discovered to be a fake or a forgery.<sup>27</sup> Leonardo da Vinci's *Salvator Mundi* is the most striking example. In the 1950s, it was sold as a copy of the original, completed by one of da Vinci's students for approximately US\$60 (approximately US\$755 today).<sup>29</sup> After its authentication as a genuine Da Vinci, it was auctioned for US\$450.3 million in 2017 (approximately US\$557.3 million today). However, experts expressed doubts afterwards and the painting has not been exhibited since.<sup>30</sup> Authentication did not change the aesthetic value of *Salvator Mundi*, but did multiply its market value by more than 700,000 – before disappearing from the museum's walls.<sup>31</sup> It is simply a fact that, once discovered, forgeries become monstrous doppelgängers. Yet this devaluation does not suggest the aesthetic value has decreased: only the market did.<sup>32</sup> Therefore, aesthetic value is not the prime factor of market value,<sup>33</sup> authentication is.<sup>34</sup> This makes sense given that the main reason to acquire an artwork today is its resale potential: authentication ensures a proper return on investment.<sup>35</sup>

<sup>22</sup> Akerlof, 'Lemons', 491 and 495.

<sup>23</sup> Wilson, 'Adverse Selection', 31.

<sup>24</sup> Hirschleifer, 'Theory of Information', 37.

<sup>25</sup> Gregory Day, 'Explaining the Art Market's Thefts, Frauds, and Forgeries (and Why the Art Market Does Not See Care)' (2014) 16(3) *Vanderbilt Journal of Entertainment & Technology Law* 457–495, 467.

<sup>26</sup> Malcolm Budd, 'Artistic Value' in Peter Lamarque and Stein Haugom Olsen (eds.), *Aesthetics and the Philosophy of Art: The Analytic Tradition, An Anthology* (2nd ed., Hoboken, NJ: Wiley, 2018) 236–246, 236.

<sup>27</sup> Fabian Bocart and Kim Oosterlinck, 'Discoveries of Fakes: Their Impact on the Art Market' (2011) 113(2) *Economic Letters* 124–126; Jonathon Keats, *Forged: Why Fakes Are the Great Art of Our Age* (New York: Oxford University Press, 2013) 8–11.

<sup>28</sup> But there are others. See Carolyn Embree and David A. Scott, 'The Multifarious Nature of Art Forgery in France: Four Case Studies of Belle Époque Fakes and Forgeries' (2015) 13 *Journal of Art Crime* 3–16, 11.

<sup>29</sup> All 2023 amounts were computed using [www.officialdata.org/](http://www.officialdata.org/).

<sup>30</sup> Katie Dixon and Zachary Schufro, 'Risky Business: Fraud, Authenticity, and Limited Legal Protections in the Art Market' (2021) 10(2) *Journal of Intellectual Property and Entertainment Law* 246–296, 247.

<sup>31</sup> Similarly, when the Musée du Louvre discovered that the *Tiara of Saitapharnès* was a forgery, it chose to hide it from the public for almost a century (Embree and Scott, 'Multifarious Nature', 7).

<sup>32</sup> Jack W. Meiland, 'Originals, Copies, and Aesthetic Values' in Peter Lamarque and Stein Haugom Olsen (eds.), *Aesthetics and the Philosophy of Art: The Analytic Tradition, An Anthology* (2nd ed., Hoboken, NJ: Wiley, 2018) 229–235, 230.

<sup>33</sup> Peter Martin Jaworski, 'In Defense of Fakes and Artistic Treason: Why Visually-Indistinguishable Duplicate Paintings Are Just as Good as the Originals' (2013) 47(4) *Journal of Value Inquiry* 391–405 at 393.

<sup>34</sup> Dixon and Schufro, 'Risky Business', 247.

<sup>35</sup> Leila Amineddoleh, 'Purchasing Art in a Market Full of Forgeries: Risks and Legal Remedies for Buyers' (2015) 3 *International Journal of Cultural Property* 419–435; Justine M. Bonner, 'Let Them Authenticate: Determining Fraud' (2017) 24(1) *UCLA Entertainment Law Review* 19–50, 24.

TABLE 5.1 *The lemons problem in the markets for (1) used car and (2) artworks*

	Market for used cars	Market for artworks
Criterion	Worth of a used car → quality of that car	Worth of an artwork → origin of that artwork
Lemons	Cars that often need repairs	Fakes and forgeries notwithstanding their aesthetic value
Peaches	Reliable cars	Genuine artworks notwithstanding their aesthetic value

For this exact reason, this chapter argues that the art market faces an idiosyncratic lemons problem: on that market, poor-quality products are not necessarily those with low aesthetic value but those with little-valued origin (if valued at all): fakes or forgeries. Genuine artworks are peaches merely because they are authentic (see Table 5.1).

To sum up the argument so far, without authentication it is impossible for the buyer to know whether they are acquiring an original artwork or a forgery.<sup>36</sup> On the contrary, the seller possesses this knowledge – assuming they are the forger or know from the forger that the artwork is forged.<sup>37</sup> This is asymmetric information. Therefore, buyers will revise their willingness-to-pay to take into account the average probability the artwork is a forgery. Yet the art market has not collapsed: consumers continue to buy art en masse while 'most dealers and patrons adamantly defend the traditional ways in which they have sold and traded art'.<sup>38</sup> The fact that art is an investment explains why neither buyers nor sellers have asked for more market transparency. Their interests are aligned: buyers do not want symmetric information because they know one day they will be in the sellers' shoes.<sup>39</sup> This does not change the fact that information asymmetry 'has created substantial societal harms'.<sup>40</sup> Preventing their occurrence is precisely the *raison d'être* of authentication.<sup>41</sup>

### 5.3 GENAI AND THE MARKET FOR LEMONS

For a lemons problem to exist, good-quality and poor-quality products must have a different value and be indistinguishable, so that there is an incentive to sell a lemon for the price of a peach. This section discusses whether GenAI is concerned by information asymmetry and incentive to cheat.

#### 5.3.1 Indistinguishability

Machine-enabled artworks and human-made ones are indistinguishable. Anyone willing to buy an artwork would be incapable of identifying its human or machine origin by simply looking at it.<sup>42</sup> This is easily explained technologically. Suffice it to analyse the functioning of generative

<sup>36</sup> This indistinguishability is reinforced by the fact that (a) the primary art market is relatively small while the secondary (resale) market is much larger and (b) that art galleries conduct their business privately and confidentially – which reduces price transparency (Day, 'Art Market's Thefts', 466).

<sup>37</sup> David Crystal-Kirk, 'Forgery Reforged: Art-Faking and Commercial Passing-Off since 1981' (1986) 49(5) *Modern Law Review* 608–616, 609.

<sup>38</sup> Day, 'Art Market's Thefts', 486.

<sup>39</sup> *Ibid.*, 486–488.

<sup>40</sup> *Ibid.*, 488.

<sup>41</sup> Bonner, 'Let Them Authenticate'.

<sup>42</sup> Hossein Aboutalebi et al., 'DeepFakeArt Challenge: A Benchmark Dataset for Generative AI Art Forgery and Data Poisoning Detection', arXiv, 2 June 2023, <https://doi.org/10.48550/arXiv.2306.01272>.