BRUSSELS PRIVACY HUB

From Transparency to Justification: Toward Ex Ante Accountability for AI

Gianclaudio Malgieri and Frank Pasquale[1]

Abstract

At present, policymakers tend to presume that AI used by firms is legal, and only investigate and regulate when there is suspicion of wrongdoing. What if the presumption were flipped? That is, what if a firm had to demonstrate that its AI met clear requirements for security, non-discrimination, accuracy, appropriateness, and correctability, before it was eployed? his paper proposes a system of "unlawfulness by default" for AI systems, an ex-ante model where some AI developers have the burden of proof to demonstrate that their technology is not discriminatory, not manipulative, not unfair, not inaccurate, and not illegitimate in its legal bases and purposes. The EU's GDPR and proposed AI Act tend toward a sustainable environment of AI systems. However, they are still too lenient and the sanction in case of non-conformity with the Regulation is a monetary sanction, not a prohibition. This paper proposes a pre-approval model in which some AI developers, before launching their systems into the market, must perform a preliminary risk assessment of their technology followed by a self-certification. If the risk assessment proves that these systems are at high-risk, an approval request (to a strict regulatory authority, like a Data Protection Agency) should follow. In other terms, we propose a presumption of unlawfulness for high-risk models, while the AI developers should have the burden of proof to justify why the AI is not illegitimate (and thus not unfair, not discriminatory, and not inaccurate). Such a standard may not seem administrable now, given the widespread and rapid use of AI at firms of all sizes. But such requirements could be applied, at first, to the largest firms' most troubling practices, and only gradually (if at all) to smaller firms and less menacing practices.

Keywords: AI, Accountability, Justification, GDPR

[1] *Gianclaudio Malgieri is Associate Professor of Law, Augmented Law Institute, EDHEC Business School; Co-Director of the Brussels Privacy Hub. Frank Pasquale is Professor of Law, Brooklyn Law School (US). The authors, listed above in alphabetical order, contributed equally to this paper.*

Contents

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. Available at www.brusselsprivacyhub.org/publications.html ISSN N° 2565-9979. This version is for academic use only. This is a first draft working paper; the final version to be published in Computer Law & Security Review, for the forthcoming special issue on «EU Data legislative revolution: the ethical issues that still remain».

1 Introduction

Those regulating artificial intelligence (AI) face a crisis of overwork and under-resourcing. Enforcement of relevant laws is too often belated if it comes at all. Massive firms face fines for AI misuse that are the economic equivalent of a parking ticket. Moreover, thanks to the well-recognized "black box" problem, identifiable AI abuses are only the tip of an iceberg of problems.[2] Even the most diligent regulators and civil society groups have little idea of the full scope and intensity of AI use at leading firms, given the triple barriers of trade secrecy, nondisclosure agreements, and technical complexity now effectively hiding their actions from public scrutiny. This crisis is likely to continue unless there is a fundamental shift in the way we regulate AI.

At present, policymakers tend to presume AI use at firms is legal, and only investigate and regulate when there is suspicion of wrongdoing. What if the presumption were flipped? That is, what if a firm had to certify that its AI met clear requirements for security, nondiscrimination, accuracy, appropriateness, and correctability, before it collected, analyzed, or used data?[3] Such a standard may not seem administrable now, given the widespread and rapid use of AI at firms of all sizes. But such requirements could be applied, at first, to the largest firms' most troubling practices, and only gradually (if at all) to smaller ones and less menacing applications of AI. For example, would it really be troubling to require firms to demonstrate basic practices of fairness, accuracy, and validity, once they have used an AI system in use by over 1 million people?[4] Scholars have argued that certain data practices should not be permitted at all.[5] Rather than expecting underfunded, understaffed regulators to overcome the monumental administrative and black box problems mentioned above, responsibility could be built into the structure of data-driven industries via licensure schemes that require certain standards to be met before large scale data practices expand even further.[6]

---

[2] *Frank Pasquale, The Black Box Society: The Secret Algorithms That Control Money and Information (Harvard Univ Pr 2015).*

[3] *For earlier examples of this kind of move to supplement ex post regulation with ex ante licensure, see Saule Omarova, 'License to Deal: Mandatory Approval of Complex Financial Products' (2012) 90 Washington University Law Review 064; Andrew Tutt, 'An FDA for Algorithms' (2017) 69 Administrative Law Review 83; Pasquale, The Black Box Society (n 2) 181; The Federal Communications Commission's power to license spectrum and devices is also a useful precedent here as well. Data may usefully be considered as a public resource. Salomé Viljoen, 'A Relational Theory of Data Governance' [2021] The Yale Law Journal 82.*

[4] *For an overview of what such practices may entail Timnit Gebru and others, 'Datasheets for Datasets' (2021) 64 Communications of the ACM 86; Matthew Zook and others, 'Ten Simple Rules for Responsible Big Data Research' (2017) 13 PLOS Computational Biology e1005399.*

[5] *Siddharth Venkataramakrishnan, 'Top Researchers Condemn "Racially Biased" Face-Based Crime Prediction' Financial Times (24 June 2020) <https://www.ft.com/content/aaa9e654-c962-46c7-8dd0-c2b4af932220> accessed 21 January 2022 ("More than 2,000 leading academics and researchers from institutions including Google, MIT, Microsoft and Yale have called on academic journals to halt the publication of studies claiming to have used algorithms to predict criminality. The nascent field of AI-powered 'criminal recognition' trains algorithms to recognise complex patterns in the facial features of people categorised by whether or not they have previously committed crimes."). ; For more on the problems of face-focused prediction of criminality by AI, see Frank Pasquale, 'When Machine Learning Is Facially Invalid' (2018) 61 Communications of the ACM 25, 25.*

[6] *See also, Data Ethics Commission of the Federal Government of Germany, 'Opinion of the Data Ethics Commission' <https://www.bmj.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN_lang.html> accessed 21 January 2022 (calling for "Preventive official licensing procedures for high-risk algorithmic systems"). The DEC observes that, "[I]n the case of algorithmic systems with regular or appreciable (Level 3) or even significant potential for harm (Level 4), in addition to existing regulations, it would make sense to establish licensing procedures or preliminary checks carried out by supervisory institutions in order to prevent harm to data subjects, certain sections of the population or society as a whole." Id. Such licensing could also be promulgated by national authorities to enforce the European Union's proposed AI Act. ; Frank Pasquale and Gianclaudio Malgieri, 'Opinion | If You Don't Trust A.I. Yet, You're Not Wrong' The New York Times (30 July 2021) <https://www.nytimes.com/2021/07/30/opinion/artificial-intelligence-european-union.html> accessed 21 January 2022.*

This article addresses the potential for licensure in the realm of product-based and services-based AI, including automobiles, aircraft, logistics, smart infrastructure, and medical equipment, as well as in AI services.[7] There is increasing concern about the validity of the data used in AI, and the algorithms it is based on. Rather than addressing all these concerns in a post hoc way, via tort-based liability, the *ex ante* approach of licensure must be part of the regulatory armamentarium. There are some wrongs that can arise out of AI that are too serious to be recompensed ex post. [8]

To give a concrete example motivating this flipped presumption about the use of AI, consider the growing prevalence of AI diagnostics and safety tools. For example, in the case of computerized physician order entry (CPOE) for prescriptions, a "drug-drug interaction" alert (DDI) could simply warn a physician about possible side effects from simultaneous ingestion of two pills, or troubling side effects for a given medication for persons with a specific clinical background.[9] There are new reports about such potential adverse events daily. How can patients be assured that they are being treated with AI that is up to date? Licensure is one way to ensure that ongoing duties to maintain and update AI are respected. Other mission-critical applications include automobiles and aircraft, where proper operation can mean the difference between life and death for passengers.

A licensure regime for AI would enable citizens to democratically shape data's scope and proper use, rather than resigning ourselves to being increasingly influenced and shaped by forces beyond our control. To ground the case for more *ex-ante* regulation, Part I describes the expanding scope of AI, and the threats that scope poses. Part II describes the dominant current modes of AI regulation, while Part III examines the substantive foundation of licensure models by elaborating a jurisprudential conception of justification. Part IV addresses the institutional dimensions of our licensure proposal, and addresses objections. Part V concludes with reflections on the opportunities created by AI licensure frameworks and potential limitations upon them.

This paper will make reference to both the EU and the US legal framework. As regards the EU legal framework, this paper will build mostly on the approved laws, in particular the GDPR, while the proposed EU AI Act (whose initial text is still under discussion and during the preparation of this paper) will not be analyzed in detail here, but just considered as an additional reference. An analysis of that specific proposal is beyond the scope of this article.

2 The Dominant Current Frameworks for AI Regulation

There are good reasons to be skeptical of artificial intelligence. Tesla crashes have dented the dream of self-driving cars.[10] Even in areas where A.I. seems to be an unqualified good, like machine learning to better spot melanoma, researchers are worried that current data sets do not adequately represent all patients' racial backgrounds.[11] While

---

[7] *We set aside, for now, the types of evaluative AI that are being deployed to rank and rate persons for job aptitude, educational admissions, credit, and similar opportunities. These may well be optimally subject to a licensing regime, but they raise enough distinctive issues that a focus on product-based AI is necessary to delimit the scope of this paper.*

[8] *Our model is meant to complement ex post approaches of tort and audit, with ex ante licensure. For more on the importance of audits (whose results could indeed feed into the information necessary for a valid licensing scheme, see Gregory Falco and others, 'Governing AI Safety through Independent Audits' [2021] 3 Nature Machine Intelligence <https://uwe-repository.worktribe.com/output/7562797/governing-ai-safety-through-independent-audits> accessed 21 January 2022.*

[9] *For a good typology of potential scenarios arising in the context of assistive AI, see generally W Nicholson Price, Sara Gerke and I Glenn Cohen, 'Potential Liability for Physicians Using Artificial Intelligence' (2019) 322 JAMA 1765.*

[10] *Prescient commentators warned of this possibility. See Meredith Broussard, Artificial Unintelligence: How Computers Misunderstand the World (MIT Press 2018).*

[11] *Angela Lashbrook, 'AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind' The Atlantic (16 August 2018) <https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/> accessed 3 May 2022.*

machines are proving "better than human" at some narrow tests, that superiority is fragile, given the dependence of many forms of AI on data sets that change over time.[12]

As AI becomes more prevalent, massive firms are privy to exceptionally comprehensive and intimate details about individuals. Mysterious algorithms predict job applicants' performance based on little more than video interviews. [13] Similar technologies may soon be headed to the classroom, as administrators use "learning analytics platforms" to scrutinize students' written work and emotional states.[14] Financial technology companies are using social media and other sensitive data to set interest rates and repayment terms.[15]

In short, sectors ranging from transport, financial, retail, health, leisure, and entertainment are all being increasingly affected by AI. Once large enough stores of data are created, there are increasing opportunities to create AI-driven inferences about persons based on extrapolations from both humanly recognizable and ad hoc, machine learning-recognizable groups. Machines as well are increasingly directed by AI.

This Part surveys existing efforts to address the challenges posed by AI. In Section 2.1, self-help, disclosure, and notice and consent approaches are analyzed. Section 2.2 drills down on the promise and limits of explanatory AI (XAI). Whatever the merits of extant approaches, they should be complemented by ex-ante, regulatory approaches based on licensing, at least with respect to some AI applications.

2.1The Limits of Self-Help, Notice, and Consent

Another simple way to regulate AI in reputational and evaluative contexts is to set a rule that persons must consent to its application before it may be applied. With respect to products, this would likely amount to a mere notification rule. Consumers would be notified if the product they were buying had significant AI in it and could then decide whether or not to purchase it. Similarly, employers might be required to disclose if they use AI tools in hiring. And litigants could be required to publicly acknowledge their utilization of such tools, as Pasquale & Cashwell have recommended.[16]

How can a person with a job and family to take care of, try to figure out which of thousands of AI controllers has information about them, has correct information, and has used it in a fair and rigorous manner? In the U.S., even the diligent will all too often run into the brick walls of trade secrecy, proprietary business methods, and malign neglect if they do so much as ask about how their AI has been used, with whom it has been shared, and how it has been analyzed.[17]

---

[12] *Eric Topol, Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again (Illustrated edition, Basic Books 2019); Gary Marcus and Ernest Davis, Rebooting AI: Building Artificial Intelligence We Can Trust (Vintage 2019).*

[13] *Drew Harwell, 'A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job' Washington Post <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/> accessed 3 May 2022; See also Zoë Corbyn, '"Bossware Is Coming for Almost Every Worker": The Software You Might Not Realize Is Watching You' The Guardian (27 April 2022) <https://www.theguardian.com/technology/2022/apr/27/remote-work-software-home-surveillance-computer-monitoring-pandemic> accessed 3 May 2022.*

[14] *Deborah Lupton and Ben Williamson, 'The Datafied Child: The Dataveillance of Children and Implications for Their Rights' (2017) 19 New Media & Society 780.*

[15] *Kristin Johnson, Frank Pasquale and Jennifer Chapman, 'Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation' (2019) 88 Fordham Law Review 31.*

[16] *Frank Pasquale and Glyn Cashwell, 'Prediction, Persuasion, and the Jurisprudence of Behaviorism' [2018] Faculty Scholarship <https://digitalcommons.law.umaryland.edu/fac_pubs/1604>.*

[17] *Even in the health care system, where access to such information is supposed to be guaranteed by federal health privacy laws, patients find considerable barriers to the exercise of their rights.*

Europeans may make Subject Access Requests, but there are far too many AI-gathering and AI-processing firms for the average person to conduct review of their results in a comprehensive way.

The list of potential targets of disclosure is endless. However, the benefits of disclosure are not nearly as extensive. First, the growing prevalence of AI may make the "right" to avoid its use nugatory. Eventually, every automobile may include it, rendering the disclosure a mere notice without opportunity to act upon it, much as HIPAA notices operate in the U.S. medical context. In other words: if it is a near-inevitability that such technologies will be an increasingly important part of the products surrounding us, the question is less how to give individuals a chance to "opt out," than how to ensure the inevitable accoutrements of their daily lives are functioning in a responsible and accountable manner.

This consent-based approach has multiple infirmities.[18] Much AI arises out of observation unrestricted by even theoretical contracts. To give an example: a person may be put in situations where it is impractical to "consent" to AI use—for example, when entering another person's car, home, or office.

There are also practices that it may be unwise to permit persons to consent to. For example, a driver may freely choose an autonomous vehicle programmed to save the driver in cases of unavoidable tragedy, even if that means taking the lives of many others. (Imagine, for instance, a car facing an oncoming truck which an only avoid a head-on collision by colliding with a crowd on a sidewalk.) Such a selfish action should not be permitted. Moreover, by ruling it out of hand on a regulatory level, regulators can nip in the bud potential arms races of AVs designed to protect occupant safety above all other concerns.

2.2 The limits of AI "explanation"

A deeper version of a disclosure approach involves AI explanation. Such a rule would require that vendors not only disclose the presence of AI in a product or service, but also explain how it works. Legal scholars and computer scientists have discussed widely *how* to reach a good level of AI explainability and a good level of algorithmic accountability and fairness.

In general terms, *explaining* decision-making is a complex task.[19] Many commentators have interrogated the notion of explanation in AI in particular.[20] In general terms, explaining means making (an idea or a situation) clear to someone by describing it in more detail or revealing relevant facts.[21] In other terms, the explanation is an act of spotting the main reasons or *factors* that led to a particular consequence, situation or decision.[22]

---

[18] *Julie E Cohen, 'Turning Privacy Inside Out' (2019) 20 Theoretical Inquiries in Law <http://www7.tau.ac.il/ojs/index.php/til/article/view/1607> accessed 23 January 2019; Gabriela Fortuna-Zanfir, 'Forgetting about Consent. Why the Focus Should Be on "Suitable Safeguards" in Data Protection Law' in Serge Gutwirth, Ronald Leenes, Paul De Hert (ed), Reloading Data Protection (Springer 2014); Bart W Schermer, Bart Custers and Simone van der Hof, 'The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection' (2014) 16 Ethics and Information Technology 171.*

[19] *Charles Tilly, Why? (2008) <https://press.princeton.edu/books/paperback/9780691136486/why> accessed 21 January 2022.*

[20] *Tim Miller, 'Explanation in Artificial Intelligence: Insights from the Social Sciences' (2019) 267 Artificial Intelligence 1.*

[21] *'EXPLAIN | Meaning & Definition for UK English | Lexico.Com' (Lexico Dictionaries | English) <https://www.lexico.com/definition/explain> accessed 21 January 2022.*

[22] *Andrew D Selbst and Julia Powles, 'Meaningful Information and the Right to Explanation' (2017) 7 International Data Privacy Law 233.*

In the field of Computer Science, explanation (of AI) has been referred to as making it possible for a human being (designer, user, affected person, etc.) to understand a result or the whole system.[23] Miller, analyzing the structure and expectations of explanations, identified four characteristics of explanations [24] they are a) *contrastive*, i.e. mostly in response to some counterfactuals;[25] b) *selected*, i.e. not comprehensive, but based only on the few main factors that influenced the final decision; c) *causal* rather than correlational/statistical; d) social and contextual, i.e. depending on the specific social relations and contexts at stake.[26]

As affirmed in legal theory, an explanation attempts to render a situation or a process understandable under a *causal*, *intentional, or narrative* perspective.[27] The causal nature of explanation is based on the link between cause and effect ("what are the causes behind this decision?"); while its intentional nature is based on the motives of the actor and her beliefs regarding reality ("what are purposes or intentions behind this decision?"). Considering these two sides of the coins, the explanation is the "answer to the question of *why* something happened or why someone acted as he did." Said in other terms, an explanation is a framework for understanding the action that has happened. [28]

The GDPR (and in particular the provisions in Article 22 and recital 71) are often interpreted as referring to only "one" kind of explanation. Actually, there is no unique explanation in practice:[29] each form of explanation highly depends on the context at issue.[30] More importantly, the capability to give a fair and satisfactory explanation depends also on the possibility to show *causal* links between the input data (and in particular some crucial factors within the input information) and the final decision. However, this is not always possible: while for traditionally data-based decision-making it might be easier to give adequate explanations, addressing the causes, the determining factors and the counterfactuals; in more complex AI-based decisions it might be hard to reach this high level of explainability. Indeed, looking at the quick development of deep learning in different forms of automated decisions (even COVID-19 automated diagnosis based on, e.g., lung images), explaining the specific reasons and factors of an individual decision might be nearly impossible.[31] An explanation which is neither causal, nor contextual is perhaps inadequate to show to the data subject eventual grounds for challenging the decision and then unsuitable under Article 22(3) of the GDPR. Even the proposed EU AI Act requires transparency measures to make the AI more interpretable by users (Article 13), with through some human oversight duties too (Article 14).

---

[23] *Clement Henin and Daniel Le Métayer, 'A Multi-Layered Approach for Interactive Black-Box Explanations' 38.*

[24] *Miller (n 24).*

[25] *Sandra Wachter, Brent Mittelstadt and Chris Russell, 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR' [2018] Harvard Journal of Law & Technology <http://arxiv.org/abs/1711.00399> accessed 16 September 2019.*

[26] *Miller (n 24).*

[27] *Aulis Aarnio, The Rational as Reasonable: A Treatise on Legal Justification (Springer Science & Business Media 1986).*

[28] *ibid.*

[29] *Miller (n 24); Margot E Kaminski, 'Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability' (2019) 92 Southern California Law Review <https://papers.ssrn.com/abstract=3351404> accessed 23 April 2019.*

[30] *Clement Henin and Daniel Le Métayer, 'A Framework to Contest and Justify Algorithmic Decisions' [2021] AI and Ethics <https://hal.inria.fr/hal-03127932> accessed 21 January 2022.*

[31] *Ronan Hamon and others, 'Impossible Explanations? Beyond Explainable AI in the GDPR from a COVID-19 Use Case Scenario', Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (Association for Computing Machinery 2021) <https://doi.org/10.1145/3442188.3445917> accessed 27 May 2021.*

3 Toward Justification of High-Risk AI

To overcome the abovementioned limits of disclosure, notice and consent, and explanation-driven approaches to AI regulation, a possible solution might be inspired by elements of the GDPR that focus on the legitimacy and value of data use. Article 22(3) and recital 71, when mentioning the possible measures to make automated decisions more accountable, do not address only the right to an individual explanation, but several other complementary tools (e.g., a right of contestation and rights to human involvement and algorithmic auditing). In particular, there are several principles and concepts that might influence the interpretation of accountability duties also in case of algorithmic decision-making: the fairness principle (Article 5(1)(a)), the lawfulness principle (Article 5(1)(a)), the accuracy principle (Article 5(1)(d)), the risk-based approach (Articles 24, 25, 35), and the data protection impact assessment model (Article 35).

As these provisions suggest, *justification* of automated decisions taken is not only more *feasible* but also more *useful* and *desirable* than alternative approaches discussed so far.[32] Justifying AI means not merely explaining the logic and the reasoning behind it, but also explaining why it operates in a legally acceptable (correct, lawful and fair) way (e.g., why decisions made by the AI comply with the core of the GDPR and are based on proportional and necessary data processing, using pertinent categories of data and relevant profiling mechanisms).

This justification process will be addressed in the next section. However, at this moment we can already affirm that justification and explanation complement one other: when explanations are not satisfactory or feasible, the data controller should anyway implement some alternative accountability tools.[33] In a previous paper, Kaminski and Malgieri proposed to disclose meaningful information about a Data Protection Impact Assessment (DPIA) on the algorithmic decision-making system. The DPIA, as mentioned in Article 35 of the GDPR, is a process to assess and mitigate the impact of data processing operations on fundamental rights and freedoms of data subjects.[34] This paper, in addition to that proposal, introduces a practical description of a possible *justification test* on the algorithm, where the data controller explains why the algorithm (analysed on the aggregated final effects on different data subjects, but also analysed in its purposes, intentions, etc.) is not unfair, unlawful, inaccurate, and beyond the purpose limitation of relevant data.

This Part proposes a shift from disclosure/explanation to justification of AI. Section 3.1 describes the concept of justification generally, while Section 3.2 focuses on justification in legal contexts. In Section 3.3, the focus is even narrower, on the nature of justification in the General Data Protection Regulation (GDPR). Section 3.4 applies the lessons of Section C to AI in particular. In the next Part, we propose a licensure regime to provide an institutional framework for ensuring accountability *ex ante*, rather than merely chasing after it *ex post*.

3.1 The Nature of Justification

Before describing the practicalities of a possible justification model and before exploring the advantages of this approach, it is useful to understand *what* justification *means* in general and more specifically in the legal field (and in the data protection field in particular).

---

[32] *Kaminski (n 33).*

[33] *Lilian Edwards and Michael Veale, 'Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?' (2018) 16 IEEE Security & Privacy 46.*

[34] *Margot Kaminski and Gianclaudio Malgieri, 'Multi-Layered Explanation from Algorithmic Impact Assessments in the GDPR', FAT 2020 Proceedings (ACM publishing 2020).*

In general terms, a justification is an action to prove or show (a person, an action, opinion etc.) to be *just*, right, desirable or reasonable.[35] The meaning of justification changes in different fields and contexts.[36] For example, in theology the justification is the action of declaring or making "righteous" in the sight of God.[37] Similarly, in philosophical terms, the justification of decision-making that affects human agents and human societies means proving (usually with appeals to a meta-ethical framework like utilitarianism or deontology) that a theory, opinion, or approach to a problem is validated by relevant meta-ethical criteria. [38]. In scientific terms, justifying means proving that a theory or a statement is correct and verified through the scientific method.[39]

While explanation aims to improve understanding about why a decision was taken, a justification aims at convincing an observer that the decision is "just" or "right" (following different benchmarks of correctness or validity in different fields).[40] In different terms, while explanations are *descriptive* and *intrinsic* because they only depend on the system itself, justifications are *normative* and *extrinsic* because they are grounded on external references, namely a "norm" according to which we can assess the validity of the decision. [41] This means that a justification requires two elements: a) the reference *norm* and b) the *proof* that that case or decision applies to that norm.

Whether the proof can follow logical reasoning standards, the "norm" depends on the specific context at issue. As shown above, the norm can be based on theological grounds, philosophical grounds (utilitarian norm, deontological norm, etc.), scientific grounds (scientific method) and, of course, legal grounds. Indeed, in legal terms, justification means proving that a certain action or act respects the current law and, more in general, the *legality* principle.[42]

Actually, as Loi et al. argue,[43] the two-dimensional justification that we mention above (norm and proof) should have a hybrid nature. In particular, the norms can be also from different sources (e.g., utilitarian and legal): a decision-maker can justify a decision on her "primary goals" (based on utilitarian norms, i.e. business objectives), but she is also asked to justify her decision on "constraining goals" imposed by law and, thus, based on legal norms (or other ethical values), such as privacy, fairness, etc.[44] Justifying a decision on the primary goals aims to show that the decision is not morally arbitrary, while justifying it on the constraining goals aims to prove the legality of that decision.

---

[35] *'JUSTIFICATION | Meaning & Definition for UK English | Lexico.Com' (Lexico Dictionaries | English) <https://www.lexico.com/definition/justification> accessed 21 January 2022.*

[36] *Luc Boltanski and Laurent Thévenot, On Justification (2006) <https://press.princeton.edu/books/paperback/9780691125169/on-justification> accessed 21 January 2022.*

[37] *'JUSTIFICATION | Meaning & Definition for UK English | Lexico.Com' (n 39).*

[38] *See, in general, Larry Alexander and Michael Moore, 'Deontological Ethics' in Edward N Zalta (ed), The Stanford Encyclopedia of Philosophy (Winter 2020, Metaphysics Research Lab, Stanford University 2020) <https://plato.stanford.edu/archives/win2020/entries/ethics-deontological/> accessed 1 December 2020.*

[39] *Paul K Moser, 'Justification in the Natural Sciences' (1991) 42 The British Journal for the Philosophy of Science 557; Mario Bunge, Philosophy of Science: From Problem to Theory (Transaction Publishers 1998).*

[40] *Or Biran and Courtenay V Cotton, 'Explanation and Justification in Machine Learning : A Survey' </paper/Explanation-and-Justification-in-Machine-Learning-%3A-Biran-Cotton/02e2e79a77d8aabc1af1900ac80ceebac20abde4> accessed 26 November 2020.*

[41] *Henin and Métayer (n 34).*

[42] *Aarnio (n 31); Mireille Hildebrandt, Law for Computer Scientists and Other Folk (Oxford University Press 2020).*

[43] *Michele Loi, Andrea Ferrario and Eleonora Viganò, 'Transparency as Design Publicity: Explaining and Justifying Inscrutable Algorithms' [2020] Ethics and Information Technology <https://doi.org/10.1007/s10676-020-09564-w> accessed 30 November 2020.*

[44] *ibid.*

3.2 Legal Justification

Coming back to the notion of *legal* justification, scholars have proposed different approaches to it,[45] in particular observing judgements and the reasoning behind judicial acts, which have a function for appeal.[46] In regulatory contexts, an agency must justify the rules it promulgates. As large firms deploying AI increasingly govern aspects of common life,[47] they should expect to see more societal demands that their products provide similar justifications.[48] Before explaining how such reason-giving may be institutionalized, it is helpful to review the special case of legal justification.

In general terms, there are strictly legal positivist approaches (i.e., a valid law in itself is a sufficient justification) and more balanced approaches that concede the dependence of some degree of legal validity on normative legitimacy (i.e., a justification lies on a balance between the letter of the law and other grounds having significance in the decision-making).[49] A more balanced approach might better solve different issues related to the law's open nature and the defeasible nature of legal justification (if additional information is taken into account, the status of a conclusion can change).[50] These considerations are also evident in criminal law, where the "justification" is an exception to the prohibition of committing certain offenses that renders a nominal violation of the criminal law lawful and therefore exempt from criminal sanctions. In doing so, such a justification balances a general legal norm with other contextual interests at issue.[51]

In sum, while an explanation tends to clarify only why a decision was taken (on which "primary goals", on which practical interests and needs it was taken), [52] a "legalistic" justification usually tends just to focus on the mere written law, without a contextual consideration of the balance of interests.

Both these approaches appear incomplete to our purposes (justification of algorithmic decisions). A desirable justification should not merely show the compliance with the "law", but with the *core* or essence of the legal principles, i.e., with the *legality* principle.[53] As we will argue below, the core of data protection in the GDPR is summarized in the data protection principles in Article 5. Accordingly, justifying an automated decision-making under the data protection *goals* and *norms* means – at least – showing the respect with the principles of data protection in Article 5.

---

[45] *Aarnio (n 31); Arno R Lodder, Dialaw: On Legal Justification and Dialogical Models of Argumentation (1999 ed, Kluwer Academic Pub 1999).*

[46] *Aarnio (n 31).*

[47] *Elizabeth Anderson, Private Government: How Employers Rule Our Lives (Princeton University Press 2017); Frank Pasquale, New Laws of Robotics: Defending Human Expertise in the Age of Ai (Belknap Pr 2020).*

[48] *Anderson (n 51); Frank Pasquale, 'Licensure as Data Governance' [2021] Knight First Amendment Institute at Columbia University <https://knightcolumbia.org/content/licensure-as-data-governance> accessed 21 January 2022.*

[49] *Aarnio (n 31).*

[50] *Lodder (n 49).*

[51] *JC Smith, Justification and Excuse in the Criminal Law (Stevens 1989); Donald L Horowitz, 'Justification and Excuse in the Program of the Criminal Law' (1986) 49 Law and Contemporary Problems 109.*

[52] *Kiel Brennan-Marquez, '"Plausible Cause": Explanatory Standards in the Age of Powerful Machines' 70 Vanderlibt Law Review 53; Kaminski (n 33).*

[53] *Hildebrandt (n 46).*

3.3 Justification of Data Processing in the GDPR

Although in the proposed EU AI Act there is no explicit justification requirement, but only specific design requirements and –inter alia – risk assessment duties for high-risk AI systems (Articles 6-15), in the GDPR we observe several references to justification of data processing in general, and of automated decision-making in particular. In different parts of the GDPR, when there is a prohibition (e.g., the prohibition to repurpose the data processing as stated in Article 5(1)(b); the prohibition to process sensitive data as stated in Article 9(1); the prohibition to conduct automated decision-making as stated in Article 22(1); the prohibition of transferring data outside the EU as mentioned in Article 44, etc.) there is always a list of exceptions, often accompanied by some safeguards to protect fundamental rights and freedoms of the data subject. This combination of exception and safeguards is the basis of what we can consider a *justification*. In addition, in these cases the GDPR often refers to the "principles of data processing" as the overarching norm or goal that the data controller needs to comply with in order to *justify* the legality of some nominally illegal acts (see, e.g., recital 72 about profiling or recital 108 about data transfer).

We might observe another strong example of justification in the GDPR: it is the case of high-risk data processing (Article 35). Under the Data Protection Impact Assessment (DPIA) model, data controllers must prove the legal proportionality and necessity of the data processing, and thus the legal necessity and proportionality of eventual automated decisions taken (Art. 35(7)(d)). This may constitute a form of *justification of data processing* on the basis of legality and legitimacy, aiming at the "essence" of data protection.[54]

In addition, the Article 29 Working Party Guidelines on profiling recommend that data controllers (in order to comply with Articles 13-15) explain the pertinence of categories of data used and the relevance of the profiling mechanism. [55] Assessing whether the data used are pertinent and the profile is relevant for a decision, as well as assessing the necessity and proportionality of the data processing in an automated decision-making system, seems to constitute a call for justification. The purpose of such assessment is not just transparency about the technology and its processes, but a justification about the lawfulness, fairness, necessity, accuracy and legitimacy of certain automated decisions.[56]

Interestingly, empirical research revealed that justification of algorithms (defined as showing the fairness of goals and rationales behind each step in the decision) is the most effective type of explanation in changing users' attitudes towards the system.[57]

3.4 Specific Grounds for Algorithmic Justifications in the GDPR

While some scholars have already addressed the need for justification of automated decision-making (rather than a mere need for explanation), very few authors tried to clarify *what* this ADM justification should be and *how* it should be conducted under the GDPR rules. This article argues that, considering the meaning of "legal justification" as mentioned in

---

[54] *Dariusz Kloza and others, 'Data Protection Impact Assessment in the European Union: Developing a Template for a Report from the Assessment Process' (LawArXiv 2020) DPiaLab Policy Brief <https://osf.io/7qrfp> accessed 1 December 2020.*

[55] *Article 29 Working Party, 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' (2017).*

[56] *Kaminski and Malgieri (n 38).*

[57] *Biran and Cotton (n 44); Kaminski (n 33); Tom R Tyler, 'Procedural Justice, Legitimacy, and the Effective Rule of Law' (2003) 30 Crime and Justice 283, 283.*

the previous sections, justifying an algorithmic decision should lead to *prove the legality of that decision*. For "legality", we mean not just lawfulness, but also accountability, fairness, transparency, accuracy, integrity, necessity.

In the last years, scholars have called for fair algorithms,[58] or for accountable algorithms[59] or for transparent algorithmic decisions[60] or, again, for lawful, accurate and integrous automated decisions. Justifying ADM means calling for algorithmic decision processes that prove to have *all* the aforementioned characteristics and respect the *essence* or the core of data protection.[61] The author argues that the essence of data protection in the GDPR consists in the data protection *principles* in Article 5. Accordingly, *justifying* automated decisions means proving that they comply (or adjusting them in order to comply) with data protection principles in Article 5.

Interestingly, the principles of data protection seem to lead to the desirable characteristics of automated decision-making as mentioned above. We will now analyze them one-by-one, contextualizing them to the case of algorithmic decision-making.

Article 5(1)(a) refers to lawfulness, transparency and fairness. As regards *lawfulness*, automated decision-making should be lawful, i.e. having a legal ground and respect fundamental rights and freedom. Such a legal basis should be found not only in Article 6(1) (or in Article 9(2) in case of special categories of personal data), but also in Article 22. Since Article 22(1) is interpreted as a *prohibition* of automated decision-making,[62] in order to make it lawful it is necessary to prove that one of the exceptions in Article 22(2) (consent, contract, Union or national law) applies, with the related requirements in Article 22(3) (suitable measures to safeguard the data subject's rights, including at least the right to human intervention, to express his or her point of view and to contest the decision). This part of "justification" is the most formal one: the controller needs to *justify* why an activity which is apparently unlawful (profiling individuals or taking significant decisions on automated bases) is instead lawful. In this sense, this part of justification reminds the legal justification in criminal law as mentioned above.[63]

As regards *fairness* justification, the data controller should prove that the decision-making processing is fair, i.e. non-discriminatory, unbiased, non-manipulative and that in general it does not exploit a significant imbalance between the controller and the subject in particular contexts (vulnerable individuals).[64] In general, the algorithmic processing should

---

[58] *Future of Privacy Forum, 'Unfairness By Algorithm: Distilling the Harms of Automated Decision-Making' (2017) <https://fpf.org/2017/12/11/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making/> accessed 8 February 2020; Sainyam Galhotra, Yuriy Brun and Alexandra Meliou, 'Fairness Testing: Testing Software for Discrimination', Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering - ESEC/FSE 2017 (ACM Press 2017) <http://dl.acm.org/citation.cfm?doid=3106237.3106277> accessed 31 May 2019; Andrew D Selbst, 'Disparate Impact in Big Data Policing' (2018) 52 Georgia Law Review 109.*

[59] *Joshua Kroll and others, 'Accountable Algorithms' (2017) 165 University of Pennsylvania Law Review 633.*

[60] *Bruno Lepri and others, 'Fair, Transparent, and Accountable Algorithmic Decision-Making Processes' (2018) 31 Philosophy & Technology 611; Bilyana Petkova and Philipp Hacker, 'Reining in the Big Promise of Big Data: Transparency, Inequality, and New Regulatory Frontiers' [2016] Lecturer and Other Affiliate Scholarship Series <https://digitalcommons.law.yale.edu/ylas/13>; Mireille Hildebrandt, 'Profile Transparency by Design? Re-Enabling Double Contingency' <https://works.bepress.com/mireille_hildebrandt/63/> accessed 3 January 2019.*

[61] *Maja Brkan, 'The Essence of the Fundamental Rights to Privacy and Data Protection: Finding the Way Through the Maze of the CJEU's Constitutional Reasoning' (2019) 20 German Law Journal 864.*

[62] *Article 29 Working Party (n 59); Michael Veale and Lilian Edwards, 'Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling' (2018) 34 Computer Law & Security Review 398.*

[63] *Smith (n 55).*

[64] *Damian Clifford and Jef Ausloos, 'Data Protection and the Role of Fairness' (2018) 37 Yearbook of European Law 130; Gianclaudio Malgieri, 'The Concept of Fairness in the GDPR: A Linguistic and Contextual Interpretation', Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (Association for Computing Machinery 2020) <https://doi.org/10.1145/3351095.3372868> accessed 29 January 2020.*

not violate the expectations of the subjects,[65] and its effects should not impair human dignity, autonomy, safety and other fundamental rights set out in the EU Charter of fundamental rights.[66]

As regards *transparency* justification, the data controller should prove that the algorithmic processing is legible[67] in the sense that, at least, meaningful information about the logic, the significance and envisaged consequences of the decision-making are communicated to the subject at the beginning of the data processing (Articles 13(2)(f) and 14(2)(g)) and, upon request, after the processing has started (Articles 15(1)(h)). As argued in a different article, [68] there are at least three levels of possible transparency: general (or "global") information, group-based explanation or individual (or "local") explanation (implementing recital 71). Each level of transparency should depend on the level of risk of that algorithmic decision-making process.[69] This multi-layered approach has been already discussed and endorsed also in the field of computer science.[70] Adding the transparency requirement in our justificatory models is not a contradiction of our shift from transparency to justification: explanations and justifications are not alternative elements, but they should read in conjunction.

Article 5(1)(b) refers, then, to *purpose limitation*. According to this principle, the justification should also prove that the ADM system is based just on data collected for the specific (licit and declared) purpose of obtaining an automated decision affecting the data subject. Under a broader perspective, the purpose limitation justification should also clarify that the algorithm was not originally developed for other purposes (military, commercial, etc.) and then eventually re-purposed for the processing at stake.[71] This would help to prevent algorithmic biases based on a decontextualization of algorithms.[72]

Article 5(1)(c) mentions the principle of *data minimization*. Under this principle, the justification of the data controller should prove that the ADM is based on the processing of only data that are adequate, relevant and limited to what is necessary for the purpose of taking that automated decision. To make an example, if the controller is an employer that needs to hire a new employee and she declares that the automated decision-making processing has the purpose of selecting the worthiest candidate, any information about, e.g., the sexual orientation, the ethnic origin, the religion or the possibility to take maternity leave (fertility, marital status, etc.), are unnecessary and should not be collected. This might be a way to prevent also intentional discrimination[73] hidden through "masking"[74]: when the data controller tries to cover intentional discrimination behind the shield of data analytics. In those cases, the data minimization justification could be

---

[65] *Michael Butterworth, 'The ICO and Artificial Intelligence: The Role of Fairness in the GDPR Framework' (2018) 34 Computer Law & Security Review 257.*

[66] *European Parliament Resolution, 'Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies - Tuesday, 20 October 2020' <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html> accessed 21 January 2022.*

[67] *Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 243.*

[68] *Kaminski and Malgieri (n 38).*

[69] *Margot E Kaminski and Gianclaudio Malgieri, 'Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations' (2019) 19–28 U of Colorado Law Legal Studies Research Paper <https://papers.ssrn.com/abstract=3456224> accessed 28 October 2019.*

[70] *Karthikeyan Natesan Ramamurthy and others, 'Model Agnostic Multilevel Explanations' <https://arxiv.org/abs/2003.06005v1> accessed 25 March 2020; Henin and Métayer (n 27).*

[71] *Malgieri and Comandé (n 71).*

[72] *Jonida Milaj, 'Privacy, Surveillance, and the Proportionality Principle: The Need for a Method of Assessing Privacy Implications of Technologies Used for Surveillance' (2016) 30 International Review of Law, Computers & Technology 115; Miller (n 24).*

[73] *Pauline T Kim, 'Data-Driven Discrimination at Work' 58 81.*

[74] *Solon Barocas and Andrew D Selbst, 'Big Data's Disparate Impact' (2016) 104 California Law Review 671; Cynthia Dwork and Deirdre K Mulligan, 'It's Not Privacy, and It's Not Fair' (2013) 66 Stanford Law Review 6; Kroll and others (n 63).*

helpful. At the same time, it is helpful when the processed data are not explicitly about protected categories of information but could also reveal information that might potentially lead to discrimination.[75]

Article 5(1)(d) refers to data *accuracy*. When justifying ADM, accuracy is also fundamental. The data controller should prove that the algorithmic decision is correct and accurate. Recital 71 (addressing ADM) requires the data controllers to make sure "that factors which result in *inaccuracies* in personal data are corrected and the risk of errors is minimized" (italics added). Indeed, accuracy (of input data and of the final product-decision) has generally been considered one of the main elements to justify the use of certain algorithms.[76] WP29 has referred to inaccuracy as one of the main issues of automated decision making, since these errors in data or in the ADM process might result in: "incorrect classifications; and assessments based on imprecise projections that impact negatively on individuals".[77] To make a practical example, the European Bank Authority, in its report on advanced analytics, has given great importance to data accuracy for justifying algorithms in the bank sector and has developed that concept through different sub-concepts: accuracy and integrity, timeliness, consistency and completeness of data.[78] The accuracy justification should result not only in proving the accuracy of input data, but also to prove that the chosen algorithm is fit-for-purpose, i.e. produces accurate results. Indeed, often discriminatory decisions are also inaccurate and incorrect.[79] Empirical studies also confirm that the "usefulness" of an algorithmic decision is a key component in their social acceptance.[80]

Article 5(1)(e) mentions the principle of *storage limitation*. Although in the field of ADM this principle seems not so pertinent, its function is also important. This principle requires that data should be stored for no longer than necessary for the purpose of the processing: this time limitation should apply also to algorithmic decision making. In other words, ADM should not be based on data that are no longer necessary (e.g., outdated data) for the purpose and the context of the decision. At the same time, controllers should not use algorithms that are no longer necessary for the declared purposes.

Article 5(1)(f) mentions the principle of *integrity and confidentiality*. In the context of ADM, it is central that algorithmic decisions are integrous and do not lead to cybersecurity risks that could adversely affect the safety (or any other fundamental right or freedom) of the data subject. Recital 71 also indirectly refers to these "risks" when mentioning automated decisions. However, cybersecurity, safety and integrity are central elements to consider when justifying algorithms. A "just" algorithm is based on and produce integrous data, is based on integrous steps and does not endanger the (digital or physical) safety of the data subject.[81]

The last principle in Article 5 is *accountability* (Article 5(2)). Accountability of ADM is an overarching goal that is considered the final objective of legally desirable AI, in particular in the data protection framework.[82] This is a "meta-principle", i.e. a methodology to apply and implement all the other data protection principles in Article 5. We can identify two

---

[75] *Sandra Wachter, 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising' (2019) 35 Berkeley Technology Law Journal <https://papers.ssrn.com/abstract=3388639> accessed 2 June 2019.*

[76] *Kroll and others (n 63); Cynthia Rudin, 'Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead' (2019) 1 Nature Machine Intelligence 206; Zachary C Lipton, 'The Mythos of Model Interpretability' (2018) 61 Communications of the ACM 36.*

[77] *Article 29 Working Party (n 59).*

[78] *Benjamin T Hazen and others, 'Data Quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An Introduction to the Problem and Suggestions for Research and Applications' (2014) 154 International Journal of Production Economics 72.*

[79] *Julia Dressel and Hany Farid, 'The Accuracy, Fairness, and Limits of Predicting Recidivism' (2018) 4 Science Advances eaao5580.*

[80] *Theo Araujo and others, 'In AI We Trust? Perceptions about Automated Decision-Making by Artificial Intelligence' (2020) 35 AI & SOCIETY 611.*

[81] *European Parliament Resolution (n 70).*

[82] *Sonia K Katyal, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66 UCLA Law Review 88; Kroll and others (n 63); Kaminski (n 33); Lepri and others (n 64).*

perspectives of accountability justification in the GDPR: a practical perspective and a methodological one. The practical accountability justification should lead to demonstrate that the data controller has proactively implemented some suitable ADM measures under Article 22(3) and recital 71,[83] that she is ready to make data subjects exercise their ADM-related rights (within and beyond Article 22) and that those rights are effective: the right to contest the algorithm, e.g., should be made effective through clear information about the system[84] and the decision and there should be concrete technical or organizational steps to take into account the eventual data subjects' contestation, to comply with it or to explain why such a request is unreasonable.[85]

On the other hand, the methodological perspective of accountability indicates *how* the justification should be conducted, i.e. how the justificatory auditing should be carried out (see section below) and what the *legal approach* to justification should be. In particular, the accountability principle – as Article 5(2) indicates – put the burden of proving data processing compliance on the data controller.[86] This means that there is a rebuttable presumption (*praesumptio iuris tantum*) that the data processing activity at stake – and, thus, any ADM processing too – is not compliant with the data protection principles. The burden of proof about legality is on the data controller.[87] In other terms, we should consider that algorithmic decisions are illegal by-default, unless the data controller *justifies* them through a valid justification, meant both as a process of justificatory auditing and an eventual final justification statement.

4 Institutionalizing Justification Via Licensure

Of course, all these values and goals as expressed in law are mere dead letters if they are not realized in an institutional framework for their effective realization (or progressive realization, to borrow terminology from the discourse of cultural and social rights).[88] One way to ensure proper justification of AI along the lines developed above is to create mechanists that promote proper scrutiny occurs before the collection, analysis, and use of the data and algorithms fueling AI, to be followed by ongoing monitoring of AI's effects and results. If enacted via a licensure regime, this scrutiny would enable a true industrial policy for AI, deterring misuses and thereby helping to channel AI development in more socially useful directions. As AI becomes more invasive and contested, there will be increasing calls for licensure regimes. To be legislatively viable, proposals for licensure need theoretical rigor and practical specificity.

Cognizant of these queries, some legislators and regulators have begun to develop an explicitly justification-driven approach to AI.[89] While not embracing licensure, U.S. Sen. Sherrod Brown has demonstrated how substantive limits may

---

[83] *See, e.g., Antoni Roig, 'Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)' (2018) 8 European Journal of Law and Technology <http://ejlt.org/article/view/570> accessed 15 January 2019.*

[84] *Kaminski and Malgieri (n 38).*

[85] *Gianclaudio Malgieri, 'Automated Decision-Making in the EU Member States: The Right to Explanation and Other "Suitable Safeguards" in the National Legislations' (2019) 35 Computer Law & Security Review 105327.*

[86] *Information Commissioner's Officer, 'Accountability and Governance' (1 October 2020) <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/> accessed 29 November 2020.*

[87] *Raluca Oprișiu, 'Reversal of "the Burden of Proof" in Data Protection | Lexology' <https://www.lexology.com/library/detail.aspx?g=e9e8c734-23d9-41bb-a723-5d664b3c86cc> accessed 29 November 2020.*

[88] *Eric Lander, 'Americans Need a Bill of Rights for an AI-Powered World' Wired <https://www.wired.com/story/opinion-bill-of-rights-artificial-intelligence/> accessed 21 January 2022.*

[89] *European Data Protection Board, 'Guidelines 1/2018 on Certification and Identifying Certification Criteria in Accordance with Articles 42 and 43 of the Regulation - Version Adopted after Public Consultation | European Data Protection Board' (2018) 20 <https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-12018-certification-and-identifying_en> accessed 21 January 2022.*

be enforced via licensure restrictions for the large-scale data collection, analysis, and use at the heart of so much AI. His proposed Data Accountability and Transparency Act would amount to a Copernican shift in U.S. governance of data, putting civil rights protection at the core of public concern.[90] This reflects a deep concern about the dangers of discrimination against minoritized or disadvantaged groups, as well as against the "invisible minorities" previously described in *The Black Box Society*.[91]

4.1 A. Case Study: Facial Recognition

Consider a concrete example of an AI diagnostic technology that could have dual uses, some to be licensed, and some not to be (and thus forbidden). Researchers have analyzed certain activities of people who extensively searched for information about Parkinson's disease on Bing, including their mouse movements six months before they entered those search terms.[92] Most users of the internet are probably unaware that not just what they click on, but how fast and smoothly they move their mouse to do so, can be recorded and traced by the sites they are using. The group of Bing users who searched for Parkinson's—which it is probably safe to assume is far more likely to have Parkinson's than the population as a whole—tended to have certain tremors in their mouse movements distinct from other searchers. These tremor patterns were undetectable by humans—only machine learning could distinguish the group identified to have a higher propensity to have Parkinson's, based in part on microsecond-by-microsecond differences in speed and motion of hand movement.

A licensure regime would likely forbid the calculation of the inference itself by entities that intend to discriminate based on it (or, more broadly, entities that have not demonstrated a personal or public health rationale for creating, disseminating, or using the inference).[93] But licenses could be granted to physicians to use these inferences to give early diagnosis and support to the person whose data was analyzed in this way. General inferences that enable other diagnostic programs may be permissible as a way of conducting "public or peer-reviewed scientific, historical, or statistical research in the public interest."[94] Thus, the generalizable finding may be made public, but its harmful use against an individual would be precluded by preventing a firm with no reasonable method of improving the person's health from making the inference. This avoids the "runaway AI" problem described in Pasquale's *Black Box Society*, where predictive analytics initially deemed promising and helpful becomes a bane for individuals stigmatized by them.

---

[90] *Press release, 'Brown Releases New Proposal That Would Protect Consumers' Privacy from Bad Actors | U.S. Senator Sherrod Brown of Ohio' <https://www.brown.senate.gov/newsroom/press/release/brown-proposal-protect-consumers-privacy> accessed 21 January 2022.*

[91] *Pasquale, The Black Box Society (n 2) 2; Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) 2 Columbia Business Law Review <https://papers.ssrn.com/abstract=3248829> accessed 18 December 2018; Gianclaudio Malgieri and Jedrzej Niklas, 'The Vulnerable Data Subject' (2020) 37 Computer Law & Security Review.*

[92] *Ryen W White, P Murali Doraiswamy and Eric Horvitz, 'Detecting Neurodegenerative Disorders from Web Search Signals' (2018) 1 npj Digital Medicine 1; In this case, the source of the information was clear: Microsoft itself, which operates Bing, permitted the researchers to study anonymized databases. In the U.S., such data is now well beyond the scope of the privacy and security protections guaranteed pursuant to the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act], see Bill Stead, NCVHS Chair and Linda Kloss, 'Health Information Privacy Beyond HIPAA: A Framework for Use and Protection' 21.*

[93] *Data Accountability and Transparency Act (DATA Act), S. 20719, 116th Cong. § 102(b)(4) (as proposed to the Senate, 2020) [hereinafter AI Act]. The proposed act states that data aggregators "shall not collect, use, or share, or cause to be collected, used, or shared, any personal data unless the aggregator can demonstrate that such personal data is strictly necessary to carry out a permissible purpose under section 102." Id. at § 101.*

[94] *European Data Protection Supervisor, 'Preliminary Opinion on Data Protection and Scientific Research | European Data Protection Supervisor' (2020) <https://edps.europa.eu/data-protection/our-work/publications/opinions/preliminary-opinion-data-protection-and-scientific_en> accessed 21 January 2022.*

Sensitive to misuses of AI, ethicists have called for restrictions on certain types of AI, with a presumption that it be banned. For example, facial recognition is widely regarded as particularly dangerous and deserving of a ban.[95] The proposed EU AI Act already provides a black-list of AI practices that should be banned (Article 5), but for the large majority of risky AI (the so-called high-risk AI), there is neither a ban nor a justificatory requirement, but only some specific design and organizational duties (Articles 6-15). But licensure allows for society to permit some of the highest value cases of facial recognition, while preventing all others. For example, it may be reasonable to develop highly specialized databases of the faces of terrorists. But to deploy such powerful technology to ticket speeders or ferret out benefits fraud is inappropriate, like using a sledgehammer to kill a fly.[96] A rational government would not license the technology for such purposes, even if it would be entirely reasonable to do so for other purposes (for example, to prevent pandemics via early detection of infection clusters). Nor would it enable many of the forms of discrimination and mischaracterization now enabled by light-to-nonexistent regulation of large-scale AI.

A licensure regime would help ensure that inaccurate, irresponsible, and damaging AI is limited. Rather than assuming that AI use is in general permitted, and that regulators must struggle to catch up and outlaw particular bad acts, a licensure regime flips the presumption. Under it, firms would need to apply for permission for their AI to be deployed in mission-critical and sensitive contexts (at the very least for new AI applications, if older ones are "grandfathered" and thus assumed to be licensed).

4.2 The Finance Precedent

The shift to thinking of AI use as a privilege, instead of as a right, may seem jarring to American ears, given the expansion of First Amendment coverage over the past century. However, even in the U.S. it is roundly conceded that there are certain particularly sensitive pieces of "information" that cannot simply be collected and disseminated. A die-hard cyberlibertarian or anarchist may want to copy and paste bank account numbers or government identification numbers onto anonymous websites, but that is illegal because complex sociotechnical systems like banks and the Social Security Administration can only function on a predicate of privacy and informational control.[97] AI that enables, say, the automation of constant attempts to break into websites, or massive misuse and wasting of computational powers, should be similarly suspect and restricted.

Just as there is regulation of federally funded human subjects research, similar patterns of review and limitation must apply to the new forms of human classification and manipulation now enabled by AI.[98] A licensure regime for AI also puts some controls on the speed and ubiquity of the correlations such systems can make. Just as policymakers may want to prevent automated bots from dominating forums like Twitter (while permitting their development in other settings), we

---

[95] *Woodrow Hartzog and Evan Selinger, 'Why You Can No Longer Get Lost in the Crowd' The New York Times (17 April 2019) <https://www.nytimes.com/2019/04/17/opinion/data-privacy.html> accessed 21 January 2022.*

[96] *For an example of other such potential excessive uses, see Robert Pear, 'On Disability and on Facebook? Uncle Sam Wants to Watch What You Post' The New York Times (10 March 2019) <https://www.nytimes.com/2019/03/10/us/politics/social-security-disability-trump-facebook.html> accessed 21 January 2022.*

[97] *For a broader argument on the limits of First Amendment protection for operational code, see David Golumbia, 'Code Is Not Speech' (Social Science Research Network 2016) SSRN Scholarly Paper ID 2764214 <https://papers.ssrn.com/abstract=2764214> accessed 21 January 2022.*

[98] *For an analysis of the potential and limits of this analogy, see James Grimmelmann, 'Law and Ethics of Experiments on Social Media Users' [2015] Cornell Law Faculty Publications <https://scholarship.law.cornell.edu/facpub/1487>.*

can and should develop a societal consensus toward limiting the degree to which automated correlations of often biased, partial, and secret AI influence our reputations and opportunities.[99]

This commitment is already a robust part of finance regulation. For example, when credit scores are calculated, the Fair Credit Reporting Act imposes restrictions on the AI that can affect them.[100] Far from being a forbidden content-based restriction on the "speech" of scoring, such restrictions are vital to a fair credit system.[101] The Equal Credit Opportunity Act takes the restrictions further regarding a creditor's scoring system.[102] Such scoring systems may not use certain characteristics—such as race, sex, gender, marital status, national origin, religion, or receipt of public assistance—as a factor regarding a customer's credit worthiness.[103]. Far from being a relic of the activist 1970s, restrictions like this are part of contemporary efforts to ensure a fairer credit system.[104]

European examples abound as well. In Germany, the United Kingdom, and France, agencies cannot use ethnic origin, political opinion, trade union membership, or religious beliefs when calculating credit scores.[105] Germany and the United Kingdom also prohibit the use of health AI, while France allows the use of health AI in credit score calculations.[106] Such restrictions might be implemented as part of a licensure regime for use of AI-driven propensity scoring in many fields. For example, authorities may license systems that credibly demonstrate to authorized testing and certification bodies that they do not process AI on forbidden grounds, while denying a license to those that do.

Moreover, credit scores themselves feature as forbidden AI in some other determinations. For example, many U.S. states prevent them from being used by employers. [107] California, Hawaii, and Massachusetts ban the use of credit scoring for automobile insurance.[108] A broad coalition of civil rights and workers' rights groups reject these algorithmic assessments

---

[99] *On policy rationales for limiting automated bot speech, see Frank Pasquale, 'Preventing a Posthuman Law of Freedom of Expression' in David E Pozen (ed), The Perilous Public Square: Structural Threats to Free Expression Today (Columbia University Press 2020).*

[100] *U.S. Fair Credit Reporting Act (FCRA) § 609, 15 U.S.C. § 1681(g) (2011)*

[101] *The FCRA provides further language limiting what information may by contained in a consumer report. 15 U.S.C. 1681(c) (2011). Consumer reports cannot contain: Title 11 cases over ten years old; civil suits, judgments, or arrest records over seven years old; paid tax liens over seven years old; accounts placed for collection or charged to profit and loss over seven years old; or any other adverse information, other than criminal convictions, over seven years old. These restrictions have not been successfully challenged as content-based restrictions under the First Amendment.*

[102] *A creditor is defined by the Equal Credit Opportunity Act as those who "extend, renew, or continue credit." 15 U.S.C. § 1691(a)(e) (2010).*

[103] *15 U.S.C. § 1691(a).*

[104] *Keshia Clukey, 'Social Networks Can't Go Into Credit Decisions Under N.Y. Ban (1)' (News Bloomberg Law) <https://news.bloomberglaw.com/banking-law/social-networks-cant-go-into-credit-decisions-under-n-y-ban> accessed 21 January 2022.*

[105] *Nicola Jentzsch, Financial Privacy: An International Comparison of Credit Reporting Systems (Springer Science & Business Media 2007).*

[106] *Id. The same restriction applies in the U.S. "A consumer reporting agency shall not furnish … a consumer report that contains medical information (other than medical contact information treated in the manner required under section 1681(c)(a)(6) of this title) about a consumer, unless—the consumer affirmatively consents, … if furnished for employment purposes, … the information is relevant to the process or effect the employment or credit transaction, … the information to be furnished pertains solely to transactions, accounts, or balances relating to debts arising from the receipt of medical services, products, or devises, … a creditor shall not obtain or use medical information … in connection with any determination of the consumer's eligibility, or continued eligibility, for credit." Fair Credit Reporting Act, 15 U.S.C. § 1681(b)(g) (2020).*

[107] *Microbilt, 'State Laws Limiting Use of Credit Information For Employment' <https://www.microbilt.com/Cms_Data/Contents/Microbilt/Media/Docs/MicroBilt-State-Laws-Limiting-Use-of-Credit-Information-For-Employment-Version-1-1-03-01-17-.pdf>.*

[108] *ibid.*

of personal worth and trustworthiness.[109] The logical next step for such activism is to develop systems of evaluation that better respect human dignity and social values in the construction of actionable reputations—those with direct and immediate impact on how we are classified, treated, and evaluated. For example, many have called for the nationalization of at least some credit scores.[110] Compared with that proposal, a licensure regime for such algorithmic assessments of propensity to repay is moderate.

To be sure, there will be some difficult judgment calls to be made, as in the case with any licensure regime. But size-based triggers can blunt the impact of licensure regimes on innovation by small and medium sized entities, focusing restrictions on firms with the most potential to cause harm. These firms are so powerful that they are almost governmental in their own right.[111] The EU's Digital Services Act proposal, for example, includes obligations that would only apply to platforms that reach 10 percent of the EU population (about 45 million people).[112] The Digital Markets Act proposal includes obligations that would only apply to firms that provide "a core platform service that has more than 45 million monthly active end users established or located in the Union and more than 10,000 yearly active business users established in the Union in the last financial year." [113] In the U.S., the California Consumer Privacy Act applies to companies that have AI on 50,000 California residents. [114] Many U.S. laws requiring security breach notifications generally trigger at around 500-1,000 records breached. [115] In short, a nuanced licensing regime can be developed that is primarily aimed at the riskiest collections of AI, and only imposes such obligations (or less rigorous ones) on smaller entities as the value and administrability of requirements for larger firms is demonstrated.

---

[109] *NYC Commission on Human Rights Legal Enforcement Guidance on the Stop Credit Discrimination in Employment Act, N.Y.C. Admin. Code §§ 8-102(29), 8-107(9)(d), (24); Local Law No. 37 (2015), 'Stop Credit Discrimination in Employment Act: Legal Enforcement Guidance' <https://www1.nyc.gov/site/cchr/law/stop-credit-discrimination-employment-act.page> accessed 21 January 2022.*

[110] *McKenna Moore, 'Biden Wants to Change How Credit Scores Work in America' Fortune <https://fortune.com/2020/12/18/biden-public-credit-agency-economic-justice-personal-finance-racism-credit-scores-equifax-transuion-experian-cfpb/> accessed 21 January 2022; Amy Traub, 'Establish a Public Credit Registry' Demos <https://www.demos.org/policy-briefs/establish-public-credit-registry> accessed 21 January 2022; 'The Biden Plan for Investing in Our Communities through Housing' (Joe Biden for President: Official Campaign Website) <https://joebiden.com/housing/> accessed 21 January 2022.*

[111] *Frank Pasquale, 'From Territorial to Functional Sovereignty: The Case of Amazon' [2017] LPE Project <https://lpeproject.org/blog/from-territorial-to-functional-sovereignty-the-case-of-amazon/> accessed 21 January 2022.*

[112] *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act), at 3, COM (2020) 825 final (Dec. 15, 2020) ("The operational threshold for service providers in scope of these obligations includes those online platforms with a significant reach in the Union, currently estimated to be amounting to more than 45 million recipients of the service. This threshold is proportionate to the risks brought by the reach of the platforms in the Union; where the Union's population changes by a certain percentage, the Commission will adjust the number of recipients considered for the threshold, so that it consistently corresponds to 10% of the Union's population."); Id. at 31 ("Such significant reach should be considered to exist where the number of recipients exceeds an operational threshold set at 45 million, that is, a number equivalent to 10% of the Union population. The operational threshold should be kept up to date through amendments enacted by delegated acts, where necessary."). Such thresholds reflect a risk-focused model of regulation commended by the German AI Ethics Commission. AI Ethics Comm'n Fed. Gov't Ger., Opinion of the AI Ethics Commission (2019), 177.*

[113] *Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act), at 36–37, COM (2020) 842 final (Dec. 15, 2020) ("A provider of core platform services shall be presumed [an important gateway for business users to reach end users] where it provides a core platform service that has more than 45 million monthly active end users established or located in the Union and more than 10,000 yearly active business users established in the Union in the last financial year.").*

[114] *CAL. CIV. CODE § 1798.140(c)(1)(B) (West 2020) (covering businesses that "[a]lone or in combination, annually buys, receives for the business's commercial purposes, sells, or shares for commercial purposes, alone or in combination, the personal information of 50,000 or more consumers, households, or devices").*

[115] *See, e.g., 16 C.F.R. § 318.5(b)–(c) ("A vendor of personal health records or PHR related entity shall provide notice to prominent media outlets serving a State or jurisdiction, following the discovery of a breach of security, if the unsecured PHR identifiable health information of 500 or more residents of such State or jurisdiction is, or is reasonably believed to have been, acquired during such breach."); SECURITY BREACH NOTIFICATION LAWS, https://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx [https://perma.cc/BS39-J2RE] (last visited May 13, 2021) (36 states set notification thresholds at 500 or 1,000).*

4.3 Anticipating Objections

There will, of course, be many objections to our proposal. The division of responsibilities among the European Commission and member states can become dizzyingly complicated, as evidenced in recent concerns about the EU AI Act's apparent delegation of important functions to standardization bodies. Veale and Borgesius have complained that the standardization bodies that are slated to play an important role in EU AI regulation are not, at present, constituted to fully grasp (let alone regulate) the full panoply of civil rights, safety, and other normative issues raised by AI.[116] We agree that it would take some investment and empowerment of such institutions to address the full array of concerns raised. However, until more apt regulatory bodies are proposed, it may well be necessary to institutionally house licensure and justification regimes in institutions that will need to adapt to the role.

Given their regulation of information and information flows, licensure regimes will face challenges in some jurisdictions based on free expression rights.[117] For some commentators, AI and robots are tantamount to persons, and thus deserve free speech rights.[118] While understandable as a futuristic possibility, the problems of such "rights for machines" become clear upon further reflection. As Birhane and van Dijk argue, so-called "intelligent machines" are "increasingly used in sustaining forms of oppression."[119] Consider the case of facial recognition. It is one thing to go to a protest when security personnel watch from afar. It is quite another when the police can immediately access your name, address, and job from a quick face scan purchased from an unaccountable private firm using machine vision.

This may be one reason why the American Civil Liberties Union decisively supported the regulation of Clearview AI (a firm providing facial recognition services) under the Illinois Biometric Information Privacy Act (BIPA), despite Clearview's insistence (to courts and the public at large) that it has a First Amendment right to gather and analyze AI unimpeded by BIPA. If unregulated, the firm's activities seem far more likely to undermine a robust public sphere than to promote it. Moreover, even if its AI applications were granted free expression protections, such protections may be limited by "time, place, and manner" restrictions. In that way, the licensure regime proposed here is much like permit requirements for parades, which recognize the need to balance the parade organizers' and marchers' free expression rights against the public need for safe and orderly streets. Given the privacy, security, and safety concerns raised by many forms of AI, a tailored licensing regime may be subject to only intermediate scrutiny in the U.S (ACLU v. Clearview AI, Case 20 CH 4353, Aug. 27, 2021: "BIPA's speaker-based exemptions do not appear to favor any particular viewpoint. As BIPA's restrictions are content neutral, the Court finds that intermediate scrutiny is the proper standard."). Far less free expression protection

---

[116] *Michael Veale and Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach' (2021) 22 Computer Law Review International 97.*

[117] *These rights claims will be particularly salient in the U.S., whose courts have expanded the scope of the First Amendment to cover many types of activity that would not merit free expression elsewhere, or would merit much less intense free expression protection, given the importance of competing rights to privacy, security, and AI protection. On the general issue of information processing being categorized as speech, see Jack M Balkin, 'Information Fiduciaries and the First Amendment' (2016) 49 UC Davis Law Review 52; Jane Bambauer, 'Is Data Speech?' (2014) 66 Stanford Law Review 57; Paul M Schwartz, 'Free Speech vs. Information Privacy: Eugene Volokh's First Amendment Jurisprudence' (2000) 52 Stanford Law Review 1559; James Hilmert, 'The Supreme Court Takes on the First Amendment Privacy Conflict and Stumbles: Bartnicki v. Vopper, the Wiretapping Act, and the Notion of Unlawfully Obtained Information' (2002) 77 77 Indiana Law Journal 639 (2002) <https://www.repository.law.indiana.edu/ilj/vol77/iss3/5>; Eric Easton, 'Ten Years After: Bartnicki v. Vopper as a Laboratory for First Amendment Advocacy and Analysis' [2011] SSRN Electronic Journal <http://www.ssrn.com/abstract=1986895> accessed 21 January 2022.Bambauer.*

[118] *John Frank Weaver, 'Why Robots Deserve Free Speech Rights' [2018] Slate <https://slate.com/technology/2018/01/robots-deserve-a-first-amendment-right-to-free-speech.html> accessed 21 January 2022.*

[119] *Abeba Birhane and Jelle van Dijk, 'Robot Rights? Let's Talk about Human Welfare Instead' [2020] Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 207.*

would be due in the EU, Canada, and Australia.[120] And the Chinese government, a leader in this space, has even more freedom of maneuver.[121]

5 Conclusion

Without proper assurances that the abuse of AI has been foreclosed, citizens should not accede to the large-scale application of AI now underway. Not only *ex post* enforcement, but also *ex ante* licensure are necessary to ensure that AI are only used for permissible purposes. This article has sketched the first steps toward translating the general normative construct of a "social license" for AI use into a specific licensure framework, building on the existing legal framework (e.g., the GDPR) and considering also the new policy proposals.

Of course, more conceptual work remains to be done, both substantively (elaborating grounds for denying a license) and practically (to estimate the resources needed to develop the first iteration of the licensing proposal).[122] The notice and consent model has enjoyed the benefits of such conceptual work for decades; now it is time to devote similar intellectual energy to a licensing model.

Ex ante licensure of large-scale AI use should become common in jurisdictions committed to enabling democratic governance of AI. Defining permissible purposes for the licensure of AI will take up an increasing amount of time for regulators, and law enforcers will need new tools to ensure that regulations are actually being followed. The articulation and enforcement of these specifications will prove an essential foundation of an emancipatory industrial policy for AI.

---

[120] *Office of the Privacy Commissioner of Canada, 'Joint Investigation of Clearview AI, Inc. by the Office of the Privacy Commissioner of Canada, the Commission d'accès à l'information Du Québec, the Information and Privacy Commissioner for British Columbia, and the Information Privacy Commissioner of Alberta' (3 February 2021) <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2021/pipeda-2021-001/> accessed 21 January 2022.*

[121] *'Translation: Internet Information Service Algorithmic Recommendation Management Provisions – Effective March 1, 2022' (DigiChina) <https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/> accessed 21 January 2022.*

[122] *To provide the proper level of resources, the "self-funding agency" model is useful. Certain financial and medical regulators are funded in part via fees paid by regulated entities which must apply to engage in certain activities. For example, fees paid pursuant to the Prescription Drug User Fee Act (PDUFA) fund the Food and Drug Administration (which essentially licenses drugs for sale in the U.S.). For background on this Act and its Amendments, see U.S. Food and Drug Administration, Prescription Drug User Fee Amendments, at https://www.fda.gov/industry/fda-user-fee-programs/prescription-drug-user-fee-amendments (last updated Aug. 25, 2021).*

Bibliography

Aarnio A, The Rational as Reasonable: A Treatise on Legal Justification (Springer Science & Business Media 1986)

Alexander L and Moore M, 'Deontological Ethics' in Edward N Zalta (ed), The Stanford Encyclopedia of Philosophy (Winter 2020, Metaphysics Research Lab, Stanford University 2020) <https://plato.stanford.edu/archives/win2020/entries/ethics-deontological/> accessed 1 December 2020

Anderson E, Private Government: How Employers Rule Our Lives (Princeton University Press 2017)

Araujo T and others, 'In AI We Trust? Perceptions about Automated Decision-Making by Artificial Intelligence' (2020) 35 AI & SOCIETY 611

Article 29 Working Party, 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' (2017)

Balkin JM, 'Information Fiduciaries and the First Amendment' (2016) 49 UC Davis Law Review 52

Bambauer J, 'Is Data Speech?' (2014) 66 Stanford Law Review 57

Barocas S and Selbst AD, 'Big Data's Disparate Impact' (2016) 104 California Law Review 671

Biran O and Cotton CV, 'Explanation and Justification in Machine Learning : A Survey' </paper/Explanation-and-Justification-in-Machine-Learning-%3A-Biran-Cotton/02e2e79a77d8aabc1af1900ac80ceebac20abde4> accessed 26 November 2020

Birhane A and van Dijk J, 'Robot Rights? Let's Talk about Human Welfare Instead' [2020] Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society 207

Boltanski L and Thévenot L, On Justification (2006) <https://press.princeton.edu/books/paperback/9780691125169/on-justification> accessed 21 January 2022

Brennan-Marquez K, '"Plausible Cause": Explanatory Standards in the Age of Powerful Machines' 70 Vanderlibt Law Review 53

Brkan M, 'The Essence of the Fundamental Rights to Privacy and Data Protection: Finding the Way Through the Maze of the CJEU's Constitutional Reasoning' (2019) 20 German Law Journal 864

Broussard M, Artificial Unintelligence: How Computers Misunderstand the World (MIT Press 2018)

Bunge M, Philosophy of Science: From Problem to Theory (Transaction Publishers 1998)

Butterworth M, 'The ICO and Artificial Intelligence: The Role of Fairness in the GDPR Framework' (2018) 34 Computer Law & Security Review 257

Canada O of the PC of, 'Joint Investigation of Clearview AI, Inc. by the Office of the Privacy Commissioner of Canada, the Commission d'accès à l'information Du Québec, the Information and Privacy Commissioner for British Columbia, and the Information Privacy Commissioner of Alberta' (3 February 2021) <https://www.priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2021/pipeda-2021-001/> accessed 21 January 2022

Clifford D and Ausloos J, 'Data Protection and the Role of Fairness' (2018) 37 Yearbook of European Law 130

Clukey K, 'Social Networks Can't Go Into Credit Decisions Under N.Y. Ban (1)' (News Bloomberg Law) <https://news.bloomberglaw.com/banking-law/social-networks-cant-go-into-credit-decisions-under-n-y-ban> accessed 21 January 2022

Cohen JE, 'Turning Privacy Inside Out' (2019) 20 Theoretical Inquiries in Law <http://www7.tau.ac.il/ojs/index.php/til/article/view/1607> accessed 23 January 2019

Corbyn Z, '"Bossware Is Coming for Almost Every Worker": The Software You Might Not Realize Is Watching You' The Guardian (27 April 2022) <https://www.theguardian.com/technology/2022/apr/27/remote-work-software-home-surveillance-computer-monitoring-pandemic> accessed 3 May 2022

Data Ethics Commission of the Federal Government of Germany, 'Opinion of the Data Ethics Commission' <https://www.bmj.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN_lang.html> accessed 21 January 2022

Dressel J and Farid H, 'The Accuracy, Fairness, and Limits of Predicting Recidivism' (2018) 4 Science Advances eaao5580

Dwork C and Mulligan DK, 'It's Not Privacy, and It's Not Fair' (2013) 66 Stanford Law Review 6

Easton E, 'Ten Years After: Bartnicki v. Vopper as a Laboratory for First Amendment Advocacy and Analysis' [2011] SSRN Electronic Journal <http://www.ssrn.com/abstract=1986895> accessed 21 January 2022

Edwards L and Veale M, 'Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?' (2018) 16 IEEE Security & Privacy 46

European Data Protection Board, 'Guidelines 1/2018 on Certification and Identifying Certification Criteria in Accordance with Articles 42 and 43 of the Regulation - Version Adopted after Public Consultation | European Data Protection Board' (2018) <https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-12018-certification-and-identifying_en> accessed 21 January 2022

European Data Protection Supervisor, 'Preliminary Opinion on Data Protection and Scientific Research | European Data Protection Supervisor' (2020) <https://edps.europa.eu/data-protection/our-work/publications/opinions/preliminary-opinion-data-protection-and-scientific_en> accessed 21 January 2022

European Parliament Resolution, 'Framework of Ethical Aspects of Artificial Intelligence, Robotics and Related Technologies - Tuesday, 20 October 2020' <https://www.europarl.europa.eu/doceo/document/TA-9-2020-0275_EN.html> accessed 21 January 2022

'EXPLAIN | Meaning & Definition for UK English | Lexico.Com' (Lexico Dictionaries | English) <https://www.lexico.com/definition/explain> accessed 21 January 2022

Falco G and others, 'Governing AI Safety through Independent Audits' (2021) 3 Nature Machine Intelligence <https://uwe-repository.worktribe.com/output/7562797/governing-ai-safety-through-independent-audits> accessed 21 January 2022

Fortuna-Zanfir G, 'Forgetting about Consent. Why the Focus Should Be on "Suitable Safeguards" in Data Protection Law' in Serge Gutwirth, Ronald Leenes, Paul De Hert (ed), Reloading Data Protection (Springer 2014)

Future of Privacy Forum, 'Unfairness By Algorithm: Distilling the Harms of Automated Decision-Making' (2017) <https://fpf.org/2017/12/11/unfairness-by-algorithm-distilling-the-harms-of-automated-decision-making/> accessed 8 February 2020

Galhotra S, Brun Y and Meliou A, 'Fairness Testing: Testing Software for Discrimination', Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering - ESEC/FSE 2017 (ACM Press 2017) <http://dl.acm.org/citation.cfm?doid=3106237.3106277> accessed 31 May 2019

Gebru T and others, 'Datasheets for Datasets' (2021) 64 Communications of the ACM 86

Golumbia D, 'Code Is Not Speech' (Social Science Research Network 2016) SSRN Scholarly Paper ID 2764214 <https://papers.ssrn.com/abstract=2764214> accessed 21 January 2022

Grimmelmann J, 'Law and Ethics of Experiments on Social Media Users' [2015] Cornell Law Faculty Publications <https://scholarship.law.cornell.edu/facpub/1487>

Hamon R and others, 'Impossible Explanations? Beyond Explainable AI in the GDPR from a COVID-19 Use Case Scenario', Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (Association for Computing Machinery 2021) <https://doi.org/10.1145/3442188.3445917> accessed 27 May 2021

Hartzog W and Selinger E, 'Why You Can No Longer Get Lost in the Crowd' The New York Times (17 April 2019) <https://www.nytimes.com/2019/04/17/opinion/data-privacy.html> accessed 21 January 2022

Harwell D, 'A Face-Scanning Algorithm Increasingly Decides Whether You Deserve the Job' Washington Post <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/> accessed 3 May 2022

Hazen BT and others, 'Data Quality for Data Science, Predictive Analytics, and Big Data in Supply Chain Management: An Introduction to the Problem and Suggestions for Research and Applications' (2014) 154 International Journal of Production Economics 72

Henin C and Métayer DL, 'A Framework to Contest and Justify Algorithmic Decisions' [2021] AI and Ethics <https://hal.inria.fr/hal-03127932> accessed 21 January 2022

——, 'A Multi-Layered Approach for Interactive Black-Box Explanations' 38

Hildebrandt M, 'Profile Transparency by Design? Re-Enabling Double Contingency' <https://works.bepress.com/mireille_hildebrandt/63/> accessed 3 January 2019

——, Law for Computer Scientists and Other Folk (Oxford University Press 2020)

Hilmert J, 'The Supreme Court Takes on the First Amendment Privacy Conflict and Stumbles: Bartnicki v. Vopper, the Wiretapping Act, and the Notion of Unlawfully Obtained Information' (2002) 77 77 Indiana Law Journal 639 (2002) <https://www.repository.law.indiana.edu/ilj/vol77/iss3/5>

Horowitz DL, 'Justification and Excuse in the Program of the Criminal Law' (1986) 49 Law and Contemporary Problems 109

Information Commissioner's Officer, 'Accountability and Governance' (1 October 2020) <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/> accessed 29 November 2020

Jentzsch N, Financial Privacy: An International Comparison of Credit Reporting Systems (Springer Science & Business Media 2007)

Johnson K, Pasquale F and Chapman J, 'Artificial Intelligence, Machine Learning, and Bias in Finance: Toward Responsible Innovation' (2019) 88 Fordham Law Review 31

'JUSTIFICATION | Meaning & Definition for UK English | Lexico.Com' (Lexico Dictionaries | English) <https://www.lexico.com/definition/justification> accessed 21 January 2022

Kaminski M and Malgieri G, 'Multi-Layered Explanation from Algorithmic Impact Assessments in the GDPR', FAT 2020 Proceedings (ACM publishing 2020)

Kaminski ME, 'Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability' (2019) 92 Southern California Law Review <https://papers.ssrn.com/abstract=3351404> accessed 23 April 2019

Kaminski ME and Malgieri G, 'Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations' (2019) 19–28 U of Colorado Law Legal Studies Research Paper <https://papers.ssrn.com/abstract=3456224> accessed 28 October 2019

Katyal SK, 'Private Accountability in the Age of Artificial Intelligence' (2019) 66 UCLA Law Review 88

Kim PT, 'Data-Driven Discrimination at Work' 58 81

Kloza D and others, 'Data Protection Impact Assessment in the European Union: Developing a Template for a Report from the Assessment Process' (LawArXiv 2020) DPIaLab Policy Brief <https://osf.io/7qrfp> accessed 1 December 2020

Kroll J and others, 'Accountable Algorithms' (2017) 165 University of Pennsylvania Law Review 633

Lander E, 'Americans Need a Bill of Rights for an AI-Powered World' Wired <https://www.wired.com/story/opinion-bill-of-rights-artificial-intelligence/> accessed 21 January 2022

Lashbrook A, 'AI-Driven Dermatology Could Leave Dark-Skinned Patients Behind' The Atlantic (16 August 2018) <https://www.theatlantic.com/health/archive/2018/08/machine-learning-dermatology-skin-color/567619/> accessed 3 May 2022

Lepri B and others, 'Fair, Transparent, and Accountable Algorithmic Decision-Making Processes' (2018) 31 Philosophy & Technology 611

Lipton ZC, 'The Mythos of Model Interpretability' (2018) 61 Communications of the ACM 36

Lodder AR, Dialaw: On Legal Justification and Dialogical Models of Argumentation (1999 ed, Kluwer Academic Pub 1999)

Loi M, Ferrario A and Viganò E, 'Transparency as Design Publicity: Explaining and Justifying Inscrutable Algorithms' [2020] Ethics and Information Technology <https://doi.org/10.1007/s10676-020-09564-w> accessed 30 November 2020

Lupton D and Williamson B, 'The Datafied Child: The Dataveillance of Children and Implications for Their Rights' (2017) 19 New Media & Society 780

Malgieri G, 'Automated Decision-Making in the EU Member States: The Right to Explanation and Other "Suitable Safeguards" in the National Legislations' (2019) 35 Computer Law & Security Review 105327

——, 'The Concept of Fairness in the GDPR: A Linguistic and Contextual Interpretation', Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (Association for Computing Machinery 2020) <https://doi.org/10.1145/3351095.3372868> accessed 29 January 2020

Malgieri G and Comandé G, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 243

Malgieri G and Niklas J, 'The Vulnerable Data Subject' (2020) 37 Computer Law & Security Review

Marcus G and Davis E, Rebooting AI: Building Artificial Intelligence We Can Trust (Vintage 2019)

Microbilt, 'State Laws Limiting Use of Credit Information For Employment' <https://www.microbilt.com/Cms_Data/Contents/Microbilt/Media/Docs/MicroBilt-State-Laws-Limiting-Use-of-Credit-Information-For-Employment-Version-1-1-03-01-17-.pdf>

Milaj J, 'Privacy, Surveillance, and the Proportionality Principle: The Need for a Method of Assessing Privacy Implications of Technologies Used for Surveillance' (2016) 30 International Review of Law, Computers & Technology 115

Miller T, 'Explanation in Artificial Intelligence: Insights from the Social Sciences' (2019) 267 Artificial Intelligence 1

Moore M, 'Biden Wants to Change How Credit Scores Work in America' Fortune <https://fortune.com/2020/12/18/biden-public-credit-agency-economic-justice-personal-finance-racism-credit-scores-equifax-transuion-experian-cfpb/> accessed 21 January 2022

Moser PK, 'Justification in the Natural Sciences' (1991) 42 The British Journal for the Philosophy of Science 557

NYC Commission on Human Rights Legal Enforcement Guidance on the Stop Credit Discrimination in Employment Act, N.Y.C. Admin. Code §§ 8-102(29), 8-107(9)(d), (24); Local Law No. 37 (2015), 'Stop Credit Discrimination in Employment Act: Legal Enforcement Guidance' <https://www1.nyc.gov/site/cchr/law/stop-credit-discrimination-employment-act.page> accessed 21 January 2022

Omarova S, 'License to Deal: Mandatory Approval of Complex Financial Products' (2012) 90 Washington University Law Review 064

Oprișiu R, 'Reversal of "the Burden of Proof" in Data Protection | Lexology' <https://www.lexology.com/library/detail.aspx?g=e9e8c734-23d9-41bb-a723-5d664b3c86cc> accessed 29 November 2020

Pasquale F, The Black Box Society: The Secret Algorithms That Control Money and Information (Harvard Univ Pr 2015)

——, 'From Territorial to Functional Sovereignty: The Case of Amazon' [2017] LPE Project <https://lpeproject.org/blog/from-territorial-to-functional-sovereignty-the-case-of-amazon/> accessed 21 January 2022

——, 'When Machine Learning Is Facially Invalid' (2018) 61 Communications of the ACM 25

——, New Laws of Robotics: Defending Human Expertise in the Age of Ai (Belknap Pr 2020)

——, 'Preventing a Posthuman Law of Freedom of Expression' in David E Pozen (ed), The Perilous Public Square: Structural Threats to Free Expression Today (Columbia University Press 2020)

——, 'Licensure as Data Governance' [2021] Knight First Amendment Institute at Columbia University <https://knightcolumbia.org/content/licensure-as-data-governance> accessed 21 January 2022

Pasquale F and Cashwell G, 'Prediction, Persuasion, and the Jurisprudence of Behaviorism' [2018] Faculty Scholarship <https://digitalcommons.law.umaryland.edu/fac_pubs/1604>

Pasquale F and Malgieri G, 'Opinion | If You Don't Trust A.I. Yet, You're Not Wrong' The New York Times (30 July 2021) <https://www.nytimes.com/2021/07/30/opinion/artificial-intelligence-european-union.html> accessed 21 January 2022

Pear R, 'On Disability and on Facebook? Uncle Sam Wants to Watch What You Post' The New York Times (10 March 2019) <https://www.nytimes.com/2019/03/10/us/politics/social-security-disability-trump-facebook.html> accessed 21 January 2022

Petkova B and Hacker P, 'Reining in the Big Promise of Big Data: Transparency, Inequality, and New Regulatory Frontiers' [2016] Lecturer and Other Affiliate Scholarship Series <https://digitalcommons.law.yale.edu/ylas/13>

Press release, 'Brown Releases New Proposal That Would Protect Consumers' Privacy from Bad Actors | U.S. Senator Sherrod Brown of Ohio' <https://www.brown.senate.gov/newsroom/press/release/brown-proposal-protect-consumers-privacy> accessed 21 January 2022

Price WN, Gerke S and Cohen IG, 'Potential Liability for Physicians Using Artificial Intelligence' (2019) 322 JAMA 1765

Ramamurthy KN and others, 'Model Agnostic Multilevel Explanations' <https://arxiv.org/abs/2003.06005v1> accessed 25 March 2020

Roig A, 'Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)' (2018) 8 European Journal of Law and Technology <http://ejlt.org/article/view/570> accessed 15 January 2019

Rudin C, 'Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead' (2019) 1 Nature Machine Intelligence 206

Schermer BW, Custers B and van der Hof S, 'The Crisis of Consent: How Stronger Legal Protection May Lead to Weaker Consent in Data Protection' (2014) 16 Ethics and Information Technology 171

Schwartz PM, 'Free Speech vs. Information Privacy: Eugene Volokh's First Amendment Jurisprudence' (2000) 52 Stanford Law Review 1559

Selbst AD, 'Disparate Impact in Big Data Policing' (2018) 52 Georgia Law Review 109

Selbst AD and Powles J, 'Meaningful Information and the Right to Explanation' (2017) 7 International Data Privacy Law 233

Smith JC, Justification and Excuse in the Criminal Law (Stevens 1989)

Stead B, Chair N and Kloss L, 'Health Information Privacy Beyond HIPAA: A Framework for Use and Protection' 21

'The Biden Plan for Investing in Our Communities through Housing' (Joe Biden for President: Official Campaign Website) <https://joebiden.com/housing/> accessed 21 January 2022

Tilly C, Why? (2008) <https://press.princeton.edu/books/paperback/9780691136486/why> accessed 21 January 2022

Topol E, Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again (Illustrated edition, Basic Books 2019)

'Translation: Internet Information Service Algorithmic Recommendation Management Provisions – Effective March 1, 2022' (DigiChina) <https://digichina.stanford.edu/work/translation-internet-information-service-algorithmic-recommendation-management-provisions-effective-march-1-2022/> accessed 21 January 2022

Traub A, 'Establish a Public Credit Registry' Demos <https://www.demos.org/policy-briefs/establish-public-credit-registry> accessed 21 January 2022

Tutt A, 'An FDA for Algorithms' (2017) 69 Administrative Law Review 83

Tyler TR, 'Procedural Justice, Legitimacy, and the Effective Rule of Law' (2003) 30 Crime and Justice 283

Veale M and Borgesius FZ, 'Demystifying the Draft EU Artificial Intelligence Act — Analysing the Good, the Bad, and the Unclear Elements of the Proposed Approach' (2021) 22 Computer Law Review International 97

Veale M and Edwards L, 'Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling' (2018) 34 Computer Law & Security Review 398

Venkataramakrishnan S, 'Top Researchers Condemn "Racially Biased" Face-Based Crime Prediction' Financial Times (24 June 2020) <https://www.ft.com/content/aaa9e654-c962-46c7-8dd0-c2b4af932220> accessed 21 January 2022

Viljoen S, 'A Relational Theory of Data Governance' [2021] The Yale Law Journal 82

Wachter S, 'Affinity Profiling and Discrimination by Association in Online Behavioural Advertising' (2019) 35 Berkeley Technology Law Journal <https://papers.ssrn.com/abstract=3388639> accessed 2 June 2019

Wachter S and Mittelstadt B, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) 2 Columbia Business Law Review <https://papers.ssrn.com/abstract=3248829> accessed 18 December 2018

Wachter S, Mittelstadt B and Russell C, 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR' [2018] Harvard Journal of Law & Technology <http://arxiv.org/abs/1711.00399> accessed 16 September 2019

Weaver JF, 'Why Robots Deserve Free Speech Rights' [2018] Slate <https://slate.com/technology/2018/01/robots-deserve-a-first-amendment-right-to-free-speech.html> accessed 21 January 2022

White RW, Doraiswamy PM and Horvitz E, 'Detecting Neurodegenerative Disorders from Web Search Signals' (2018) 1 npj Digital Medicine 1

Zook M and others, 'Ten Simple Rules for Responsible Big Data Research' (2017) 13 PLOS Computational Biology e1005399

The Brussels Privacy Hub Working Papers series

The Brussels Privacy Hub Working Papers are intended to circulate research in progress for comment and discussion. The Working Papers focus on all areas of data protection and privacy research and can contain empirical research on privacy issues and analytical work on privacy governance and regulation in the EU; global flows of data; reconciling law enforcement and privacy interests; privacy challenges posed by new technologies; comparative research on privacy in different regions; jurisprudential issues of privacy protection; and many others.

Available at https://brusselsprivacyhub.com/working-papers/

Editorial Board: Paul De Hert and Christopher Kuner

Contact: info@brusselsprivacyhub.eu

N°1 "The data protection regime applying to the inter-agency cooperation and future architecture of the EU criminal justice and law enforcement area" (November 2014) by Paul De Hert and Vagelis Papakonstantinou (35 pages)

N°2 "The new cloud computing ISO/IEC 27018 standard through the lens of the EU legislation on data protection" (November 2014) by Paul de Hert, Vagelis Papakonstantinou, Irene Kamara (25 pages)

N°3 "Towards efficient cooperation between supervisory authorities in the area of data privacy law" (October 2015) by Dariusz Kloza, Antonella Galetta (24 pages)

N°4 "The data protection regime in China" (November 2015) by Paul De Hert and Vagelis Papakonstantinou (30 pages)

N°5 "The right to privacy and personal data protection in Brazil: time for internet privacy rights?" (February 2016) by Vinícius Borges Fortes (23 pages)

N°6 "Permissions and Prohibitions in Data Protection Jurisdiction" (May 2016) by Mistale Taylor (25 pages)

N°7 "Structure and Enforcement of Data Privacy Law in South Korea" (October 2016) by Haksoo Ko, John Leitner, Eunsoo Kim and Jong-Gu Jung (20 pages)

N°8 "The "Right to be Forgotten" and Search Engine Liability" (December 2016) by Hiroshi Miyashita (15 pages)

N°9 "European Human Rights, Criminal Surveillance, and Intelligence Surveillance: Towards "Good Enough" Oversight, Preferably but Not Necessarily by Judges" (March 2017) by Gianclaudio Malgieri and Paul De Hert (25 pages)

N°10 "Does Technology Drive Law? The Dilemma of Technological Exceptionalism in Cyberlaw" (July 2017) by Meg Leta Jones, JD, PhD (31 pages)

N°11 "The Microsoft Ireland case and the cyberspace sovereignty trilemma. Post-territorial technologies and companies question territorial state sovereignty and regulatory state monopolies" (July 2018) by Paul De Hert and Johannes Thumfart (27 pages)

N°12 "Understanding the balancing act behind the legitimate interest of the controller ground: a pragmatic approach" (August 2018) by Irene Kamara and Paul De Hert (35 pages)

N°13 "Big data analytics by telecommunications operators and the draft ePrivacy Regulation" (September 2018) by Vagelis Papakonstantinou and Paul de Hert (13 pages)

N°14 "Enforcement in Indonesia Data Privacy Laws: The 2018 Facebook-Cambridge Analytica scandal as a case study" (October 2018) by Anbar Jayadi (21 pages)

N°15 "Belgium, Courts, Privacy and Data Protection. An inventory of Belgian case law from the pre-GDPR regime (1995-2015)." (January 2019) by Paul De Hert (34 pages)

N°16 Big data analytics in electronic communications: A reality in need of granular regulation (even if this includes an interim period of no regulation at all) (June 2019) by Vagelis Papakonstantinou and Paul de Hert (25 pages)

N°17 Data Localisation: Deconstructing myths and suggesting a workable model for the future. The cases of China and the EU (September 2019) by Author: Yanqing Hong, Senior Fellow, Law and Development Institute, Peking University of China, Edited by Vagelis Papakonstantinou, Brussels Privacy Hub (31 pages)

N°18 Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination (January 2020) by Dr Monique Mann and Professor Tobias Matzner (18 pages)

N°19 Access to the Internet in the EU: a policy priority, a fundamental, a human right or a concern for eGovernment? (February 2020) by Lina Jasmontaite and Paul de Hert (23 pages)

N°20 The Proposed ePrivacy Regulation: The Commission's and the Parliament's Drafts at a Crossroads? (March 2020) by Elena Gil González, Paul De Hert & Vagelis Papakonstantinou (31 pages)

N°21 Article 8 ECHR compliant and foreseeable surveillance: the ECtHR's expanded legality requirement copied by the CJEU. A discussion of European surveillance case law (April 2020) by Paul De Hert & Gianclaudio Malgieri (42 pages)

N°22 The "Ethification" of Privacy and Data Protection Law in the European Union. The Case of Artificial Intelligence (May 2020) by Niels van Dijk and Simone Casiraghi (23 pages)

N°23 Logic and Key Points of China's Cybersecurity Review Measures (June 2020) by Yanqing Hong, Senior Fellow, Law and Development Institute, Peking University of China, Edited by Vagelis Papakonstantinou, Brussels Privacy Hub (9 pages)

N°24 Individuation: re-imagining data privacy laws to protect against digital harms (July 2020) by Anna Johnston (22 pages)

N°25 The Dark Side of the MOOC? The Rise of EdTech in Times of COVID-19: A Data Protection Challenge for Universities (August 2020) by Jonas Botta (16 pages)

N°26 European Law Enforcement and US Data Companies: A Decade of Cooperation Free from Law (September 2020) by Angela Aguinaldo and Paul De Hert (16 pages)

N°27 Fashion ID and Decisively Influencing Facebook Plugins: A Fair Approach to Single and Joint Controllership (June 2021) by Paul De Hert and Georgios Bouchagiar (24 pages)

N°28 Adding and removing elements of the proportionality and necessity test to achieve desired outcomes. Breyer and the necessity to end anonymity of cell phone users (September 2021) by Paul De Hert and Georgios Bouchagiar (26 pages)

N°29 Facial recognition, visual and biometric data in the US. Recent, promising developments to regulate intrusive technologies (October 2021) by Paul De Hert and Georgios Bouchagiar (46 pages)

N°30 Necessity knows no law in contaminated times: the rule of law under pandemic police and pandemic legislation' ('Nood breekt wet in besmette tijden: de rechtsstatelijkheid van de pandemiepolitie en pandemiewetgeving') (November 2021) by Paul De Hert (33 pages)

N°31 The fundamental right to personal data protection in criminal investigations and proceedings: framing big data policing through the purpose limitation and data minimisation principles of the Directive (EU) 2016/680 (December 2021) by Paul De Hert and Juraj Sajfert (17 pages)

N°32 Humans in the GDPR and AIA governance of automated and algorithmic systems. Essential pre-requisites against abdicating responsibilities (January 2022) by Guillermo Lazcoz andPaul de Hert (31 pages)

N°33 From Transparency to Justification: Toward Ex Ante Accountability for AI (May 2022) by Gianclaudio Malgieri and Frank Pasquale (30 pages)