RETHINKING EXPLAINABLE MACHINES: THE GDPR'S "RIGHT TO EXPLANATION" DEBATE AND THE RISE OF ALGORITHMIC AUDITS IN ENTERPRISE

Bryan Casey,[*] *Ashkon Farhangi,*[**] *& Roland Vogl*[***]

ABSTRACT. The public debate surrounding the General Data Protection Regulation's ("GDPR") "right to explanation" has sparked a global conversation of profound social and economic significance. But from a practical perspective, the debate's participants have gotten way ahead of themselves. In their search for a revolutionary new data protection within the provisions of a single chapter of the GDPR, many prominent contributors to the debate have lost sight of the most revolutionary change ushered in by the Regulation: the sweeping new enforcement powers given to European data protection authorities ("DPAs") by Chapters 6 and 8 of the Regulation. Unlike the 1995 Data Protection Directive that it will replace, the GDPR's potent new investigatory, advisory, corrective, and punitive powers granted by Chapters 6 and 8 will render DPAs de facto interpretive authorities of the Regulation's controversial "right to explanation." Now that the DPAs responsible for enforcing the right have officially weighed in, this Article argues that at least one matter of fierce public debate can be laid to rest. The GDPR provides an unambiguous "right to explanation" with sweeping legal implications for the design, prototyping, field testing, and deployment of automated data processing systems. While the protections enshrined within the right may not mandate transparency in the form of a complete individualized explanation, a holistic understanding of the Regulation's interpretation by DPAs reveals that the right's true power derives from its synergistic effects when combined with the algorithmic auditing and "data protection by design" methodologies codified by the Regulation's subsequent chapters. Accordingly, this Article predicts that algorithmic auditing and "data protection by design" practices will likely become the new gold standard for enterprises deploying machine learning systems both inside and outside of the EU bloc.

---

[*] *CodeX: The Stanford Center for Legal Informatics, Stanford, CA.*

[**] *Google, Mountain View, CA; CodeX: The Stanford Center for Legal Informatics, Stanford, CA.*

[***] *Lecturer in Law at Stanford Law School; Executive Director of CodeX: The Stanford Center for Legal Informatics, Executive Director of the Stanford Program in Law, Science, and Technology, Stanford, CA.*

INTRODUCTION

The year is 1995 and a spate of pioneering companies, including the upstarts Amazon.com and eBay, are staking their financial futures on an emerging technology that appears poised to forever transform the computing and communications worlds.[1] The technology, known among its proselytizers as the "Net," represents a new form of digital infrastructure that facilitates the worldwide sharing of data and communications without

[1] *See* Harry McCracken, *1995: The Year Everything Changed*, FAST COMP. (December 30 2015), https://www.fastcompany.com/3053055/1995-the-year-everything-changed. Ebay launched under the name of AuctionWeb at the time. *See id.*

regard for geographic location.[2] Though adoption rates of this mysterious new technology remain relatively low, European anxieties surrounding its increasingly widespread use are in full swing—precipitating the passage of legislation known as the Data Protection Directive ("DPD") designed to grapple with the societal and technical complexities of a world on the cusp of a new digital era.[3]

Fast forward twenty years to the present and the internet is, decidedly, old hat. But a technology of seemingly equal allure to the "Net" circa 1995 is enjoying a period of similarly rapid ascendance. The technology is known as "machine learning"[4]—or, for those of a more poetic bent, "artificial intelligence."[5] And the level of optimism surrounding its potential to transform the world by turning machines into "intelligent"[6]

---

[2] *See id.*

[3] *See* Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data, 1995 O.J. (L 281) 31 [hereinafter "DPD"]; EUROPEAN COMMISSION, COMMISSION PROPOSES A COMPREHENSIVE REFORM OF DATA PROTECTION RULES TO INCREASE USERS' CONTROL OF THEIR DATA AND TO CUT COSTS FOR BUSINESSES (2012), http://europa.eu/rapid/press-release_IP-12-46_en.htm [hereinafter 'GDPR Proposal'].

[4] Machine learning can be described as a field of computer science that gives computers the ability to solve problems without being explicitly programmed to do so—*i.e.* the ability to "learn" by progressively improving performance on specific tasks. For references to definitions proffered by EU data authorities, *see also e.g.,* DATATILSYNET: THE NORWEGIAN DATA PROTECTION AUTHORITY, ARTIFICIAL INTELLIGENCE AND PRIVACY (January 2018), https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf [hereinafter "Norwegian Data Protection Authority"]; COMM. ON THE ANALYSIS OF MASSIVE DATA ET AL., FRONTIERS IN MASSIVE DATA ANALYSIS 101 (2013), http://www.nap.edu/catalog.php?record_id=18374.

[5] While it is not wholly accurate to define "machine learning" and "artificial intelligence" as coextensive, for practical purposes this Article adopts to convention of treating the two terms as synonymous. *See* Norwegian Data Protection Authority, *supra* note __ (defining artificial intelligence as "the concept used to describe computer systems that are able to learn from their own experiences and solve complex problems in different situations – abilities we previously thought were unique to mankind"). "Artificial intelligence is an umbrella term that embraces many different types of machine learning." *See id.*

[6] The word intelligent, here, is used in quotes because of the fraught definitional issues associated with the term. As the scholar, Ryan Calo, notes, "Few complex technologies have a single, stable, uncontested definition [and] [r]obots

decision-makers is matched only by the level of anxiety felt by those who fear the potential for bias to infiltrate machine decision-making systems once humans are removed from the equation.[7]

As recently as a decade ago, concerns surrounding bias within these types of complex automated systems would likely have struck many observers as far-fetched. Ever since the birth of computation with Turing, humans have ascribed a kind of perfect "objectivity" to the mechanistic processes underlying algorithmic decision-making—a propensity now known as "automation bias."[8] Study after study has documented an innate human tendency to assume the validity of decisions made by algorithms,[9] even when presented with information that directly contradicts the decision's apparent validity.[10] The drafters of Europe's DPD, themselves,

---

are no exception." Ryan Calo, *Robotics and the Lessons of Cyberlaw*, 103 CALIF. L. REV. 513 (2015). For stylistic purposes, this Article uses "machine learning" and "artificial intelligence" interchangeably. Both terms lack a universally accepted definition but, in this Article, refer broadly to any "computerized system that exhibits behavior that is commonly thought of as requiring intelligence." EXECUTIVE OFFICE OF THE PRESIDENT NAT'L SCI. AND TECH. COUNCIL COMM. ON TECH., PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE 6 (Oct. 2016).

[7] *See infra* notes __ – __ and accompanying text.

[8] *See, e.g.,* Linda J Skitka et al., Accountability and Automation Bias, 52 INT'L J. HUMAN COMPUTER STUD. 4, 4 (2000); Christian Sandvig, Seeing the Sort: The Aesthetic and Industrial Defence of "the Algorithm," 11 MEDIA-N 1 (2015); A HISTORY OF ALGORITHMS: FROM THE PEBBLE TO THE MICROCHIP (Jean-Luc Chabert et al. eds., 1999) ; Kate Goddard et al., *Automation Bias: A Systematic Review of Frequency, Effect Mediators, and Mitigators*, 19 J. OF THE AMER. MED. INFORMATICS ASS'N 121, 121–127 (2012) (doi:10.1136/amiajnl-2011-000089); Mary Cummings, *Automation Bias in Intelligent Time Critical Decision Support Systems,* AIAA 1ST INT. SYS. TECHNICAL CONFERENCE (2004) (doi:10.2514/6.2004-6313).

[9] An "algorithm" can be defined as "a formally specified sequence of logical operations that provides step-by-step instructions for computers to act on data and thus automate decisions." Solon Baracas & Andrew Selbst, Big Data's Disparate Impact, 104 CAL. L. REV. 671, 673 fn. 10 (quoting SOLON BAROCAS ET AL., DATA & CIVIL RIGHTS: TECH. PRIMER (2014), http://www.datacivilrights.org/pubs/2014- 1030/Technology.pdf). *See also,* A HISTORY OF ALGORITHMS: FROM THE PEBBLE TO THE MICROCHIP (Jean-Luc Chabert et al. eds., 1999) at 2 (defining "algorithm even more broadly as "any process that can be carried out automatically").

[10] *See* Goddard et al., *supra* note __; Mary Cummings, *Automation Bias in Intelligent Time Critical Decision Support Systems*, AIAA 1ST INT. SYS.

explicitly acknowledged this phenomenon in 1992. So worried were they that "machine[s] using more and more sophisticated software" might be perceived as having "an apparently objective and incontrovertible character," they felt it necessary to legislate specific measures guarding against it.[11]

In recent years, however, society's deferential attitude toward algorithmic objectivity has begun to wane—thanks, in no small part, to a flurry of influential publications examining bias within complex computational systems.[12] Particularly in the last five years, numerous

---

TECHNICAL CONFERENCE (2004) (doi:10.2514/6.2004-6313); Kathleen Mosier et al., *Automation Bias: Decision Making and Performance in High-Tech Cockpits*, 8 INT'L J. OF AVIATION PSYCH. 1, 47-63 (Feb. 1997).

[11] *See Amended Proposal for a Council Directive on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data*, at 26, COM(92) 422 final—SYN 297 (Oct. 15, 1992).

[12] Jonas Lerman, *Big Data and Its Exclusions*, 66 STAN. L. REV. ONLINE 55, 57 (2013); Brent Mittelstadt et al., T*he Ethics of Algorithms: Mapping the Debate*, 3 BIG DATA & SOC'Y 2 (2017); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV.1,16 (2014); Solon Barocas, *Data Mining and the Discourse on Discrimination*, PROC. DATA ETHICS WORKSHOP (2014), https://dataethics.github.io/proceedings/DataMiningandtheDiscourseOnDiscrimination.pdf; Bart Custers, *Data Dilemmas in the Information Society: Introduction and Overview*, in DISCRIMINATION AND PRIVACY IN THE INFO. SOC. 3, 20 (Bart Custers et al. eds., 2013); Latanya Sweeney, *Discrimination in Online Ad Delivery*, COMM. ACM, May 2013, at 44, 47 (2013); Shoshana Zuboff, *Big Other: Surveillance Capitalism and the Prospects of an Information Civilization*, 30 J. INFO. TECH. 1, 75–89 (2015); Kate Crawford & Jason Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, 55 B.C. L. REV. 93, 101 (2014) (noting "housing providers could design an algorithm to predict the [race, gender, or religion] of potential buyers or renters and advertise the properties only to those who [meet certain] profiles." *See also, e.g.,* Julia Angwin & Terry Parris Jr., *Facebook Lets Advertisers Exclude Users by Race*, PROPUBLICA (Oct. 28, 2016), https://www.propublica.org/article/facebook-lets-advertisersexclude-users-by-race; Julia Angwin et al., *Facebook (Still) Letting Housing Advertisers Exclude Users by Race*, PROPUBLICA (Nov. 21, 2017), https://www.propublica.org/article/facebook-advertising-discrimination-housing-racesex-national-origin; Devah Pager & Hana Shepherd, *The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets*, 34 ANN. REV. SOC. 181, 182 (2008); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 4 (2014) ("Because human beings program predictive

studies across multiple industry sectors and social domains have revealed the potential for algorithmic systems to produce disparate real world impacts on vulnerable groups.[13] These revelations, in turn, have had a pronounced effect on scholars, policymakers, industry leaders, and society *writ large*—often serving as a rallying cry for greater efforts to promote fairness, accountability, and transparency in the design and deployment of highly automated systems.[14]

---

algorithms, their biases and values are embedded into the software's instructions. . . ."); Danielle Keats Citron, *Technological Due Process*, 85 WASH. U. L. REV. 1249, 1254 (2008); Nell Greenfieldboyce, *Big Data Peeps at Your Medical Records to Find Drug Problems*, NPR (July 21, 2014, 5:15 AM), http://www.npr.org/blogs/health/2014/07/21/332290342/big-datapeeps-at-your-medical-records-to-find-drug-problems; Tanzina Vega, *New Ways Marketers Are Manipulating Data to Influence You,* N.Y. TIMES: BITS (June 19, 2013, 9:49 PM), http://bits.blogs.nytimes.com/2013/06/19/new-waysmarketers-are-manipulating-data-to-influence-you; Nadya Labi, *Misfortune Teller*, ATLANTIC (Jan.–Feb. 2012), http://www.theatlantic.com/ magazine/archive/2012/01/misfortune-teller/308846; Anders Sandberg, *Asking the Right Questions: Big Data and Civil Rights*, PRAC. ETHICS (Aug. 16, 2012), http://blog.practicalethics.ox.ac.uk/2012/08/asking-the-right-questions-big-data-and-civilrights; Alistair Croll, *Big Data Is Our Generation's Civil Rights Issue, and We Don't Know It*, SOLVE FOR INTERESTING (July 31, 2012), http://solveforinteresting.com/big-datais-our-generations-civil-rights-issue-and-we-dont-know-it; Kate Crawford, *The Hidden Biases in Big Data*, HARV. BUS. REV. (Apr. 1, 2013), https://hbr.org/2013/04/the-hidden-biases-in-big-data; Moritz Hardt, *How Big Data Is Unfair*, MEDIUM (Sept. 26, 2014), https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de.

[13] *See infra* notes __ and accompanying text.

[14] *See, e.g.,* EUROPEAN COMMISSION, DATA PROTECTION EUROOBAROMETER, (June 2015), CABINET OFFICE, DATA SCIENCE ETHICAL FRAMEWORK (HM Government, May 2016), https://www.gov.uk/government/publications/data-science-ethical-framework; INFORMATION COMMISSIONERS OFFICE (ICO), BIG DATA, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DATA PROTECTION (2017); EUROPEAN DATA PROTECTION SUPERVISOR (EDPS), MEETING THE CHALLENGES OF BIG DATA: A CALL FOR TRANSPARENCY, USER CONTROL, DATA PROTECTION BY DESIGN AND ACCOUNTABILITY [OPINION 7/2015] (2015); ROYAL SOCIETY, MACHINE LEARNING: THE POWER AND PROMISE OF COMPUTERS THAT LEARN BY EXAMPLE (2017); Wetenschappelijke Raad voor het Regeringsbeleid [Dutch Scientific Council for Government Policy (WRR)], *Big data in een vrije en veilige samenleving* [Big data in a free and safe society], WRR-Rapport 95 (2016); Commons Science and Technology Select Committee, *Algorithms in Decision-Making Inquiry Launched*, UK PARLIAMENT (Feb. 28, 2017); NATIONAL SCIENCE

Yet, in spite of society's recent shift in attitude toward these types of algorithmic systems, the inexorable march of "machine learning eating the world" is only accelerating.[15] Across a diverse array of industries—from private social networks to public sector courtrooms[16]—machine learning applications are witnessing unprecedented rates of adoption due to their ability to radically improve data-driven decision-making at a cost and scale incomparable to that of humans.[17] Today, virtually all experts agree that machine learning algorithms processing vast troves of data will only continue to play an increasingly large role in regulating our lives. The question, thus, becomes: How are we to regulate these algorithms?

In 2016, the EU sought to become a global pioneer in answering this question by replacing its 1990s-era DPD with comprehensive reform legislation known as the General Data Protection Regulation ("GDPR").[18] Among the numerous protections introduced by the GDPR was an update

---

AND TECHNOLOGY COUNCIL, PREPARING FOR THE FUTURE OF AI (2016), INFORMATION COMMISSIONER'S OFFICE, OVERVIEW OF THE GENERAL DATA PROTECTION REGULATION (2016) 1.1.1; HOUSE OF COMMONS SCIENCE AND TECHNOLOGY COMMITTEE, ROBOTICS AND ARTIFICIAL INTELLIGENCE (2016) HC 145; EUROPEAN PARLIAMENT COMMITTEE ON LEGAL AFFAIRS, REPORT WITH RECOMMENDATIONS TO THE COMMISSION ON CIVIL LAW RULES ON ROBOTICS (2017) 2015/2103(INL); Sophie Curtis, *Google Photos Labels Black People as 'Gorillas*, THE TELEGRAPH (May 4, 2017), http://www.telegraph.co.uk/technology/google/11710136/ Google-Photos-assigns-gorilla-tag-to-photos-of-black-people.html

[15] *See* Tom Simonite, *NVidia CEO: Software is Eating the World, but AI is Going to Eat Software*, MIT TECH. REV. (May 12 2017), https://www.technologyreview.com/s/607831/nvidia-ceo-software-is-eating-the-world-but-ai-is-going-to-eat-software/.

[16] *See, e.g.,* Corbett-Davies et al., *Algorithmic Decision Making and the Cost of Fairness*, arXiv:1701.08230v4 (Jun. 2017), https://arxiv.org/pdf/1701.08230.pdf; Nikolaj Tollenaar et al., *StatRec — Performance, Validation and Preservability of a Static Risk Prediction Instrument*, 129 BULL. SOC. METHODOLOGY 25 (2016) (detailing published UK and Dutch predictive models involving recidivism).

[17] *See id.*

[18] *See* Regulation (EU) 2016/679, of the European Parliament and the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1 [hereinafter "GDPR"].

to the DPD's rights surrounding automated decision-making.[19] The update formally enshrined what has since come to be referred to as the "right to explanation."[20] The right mandates that entities handling the personal data of EU citizens "ensure fair and transparent processing" by providing them with access to "meaningful information about the logic involved" in certain automated decision-making systems.[21]

Many view the GDPR's "right to explanation" as a promising new mechanism for promoting fairness, accountability, and transparency in the types of machine learning systems being deployed with increasing regularity by companies across the globe.[22] But as is true of numerous other rights enshrined within the GDPR, the precise contours of its protections are less than clear—leading observers to wonder exactly how the "right to explanation" will impact the use of machine learning algorithms in enterprise.

In the two years since the GDPR's official publication, this uncertainty has ignited a heated global debate surrounding the right's actual substantive protections.[23] The debate has centered on a cluster of four provisions found in Chapter 3 of the Regulation that circumscribe the specific text giving rise to the right. Scholars, industry leaders, and media sources across the globe have scoured the language of these provisions, proffering all variety of competing interpretations of what the GDPR's new, and potentially revolutionary, "right to explanation" entails.[24] But lost in the

---

[19] *See id.*

[20] *See infra* Part I and accompanying notes.

[21] *See* GDPR, *supra* note __.

[22] *See infra* Part III(C) and accompanying notes. *See also, e.g.,* EXECUTIVE OFFICE OF THE PRESIDENT NAT'L SCI. AND TECH. COUNCIL, PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE (Oct. 2016), https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf; Catherine Stupp, *Commission to Open Probe into Tech Companies' Algorithms Next Year*, EURACTIV (Nov. 7 2016), https://www.euractiv.com/section/digital/news/commission-to-open-probe-into-tech-companies-algorithms-next-year/; GOV'T OFF. FOR SCI., ARTIFICIAL INTELLIGENCE: AN OVERVIEW FOR POLICY-MAKERS (2016), https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf.

[23] *See infra* Part I and accompanying notes.

[24] *See infra* Part II and accompanying notes. *See also, e.g.,* FRANCESCA ROSSI, ARTIFICIAL INTELLIGENCE: POTENTIAL BENEFITS AND ETHICAL

debate's singular focus on Chapter 3 has been a recognition of the most revolutionary change of all ushered in by the Regulation: the sweeping new enforcement powers given to Europe's data protection authorities by Chapters 6 and 8.[25]

Unlike the DPD that it will replace, Chapters 6 and 8 of the GDPR grant EU data authorities vastly enhanced investigatory powers, a broad corrective "tool kit," and the capacity to levy fines several thousand times larger than the current maximum available under EU law.[26] The practical importance of this paradigm shift is difficult to overstate. Thanks to the GDPR's introduction of truly threatening administrative powers, EU data authorities will no longer be rendered the toothless data watchdogs many companies have long viewed them to be.[27] Rather, these newly empowered authorities will play a weighty role in enforcing and, therefore, *interpreting* the GDPR's numerous protective mandates.[28]

Viewed through this lens, it becomes apparent that many participants in the public debate surrounding the "right to explanation" have simply gotten ahead of themselves. While commentator after commentator has raced to announce their own rarified linguistic interpretation of Chapter 3, those tasked with actually enforcing the "right to explanation" have

---

CONSIDERATIONS, EUROPEAN PARLIAMENT: POLICY DEPARTMENT C: CITIZENS' RIGHTS AND CONSTITUTIONAL AFFAIRS BRIEFING PE 571.380 (2016). For media perspectives, *see also , e.g.,* Cade Metz, *Artificial Intelligence Is Setting Up the Internet for a Huge Clash With Europe*, WIRED (Jul. 11, 2016), https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/; Bernard Marr, *New Report: Revealing The Secrets Of AI Or Killing Machine Learning?*, FORBES (Jan. 12, 2017), https://www.forbes.com/sites/bernardmarr/2017/01/12/new-report-revealing-the-secrets-of-ai-or-killing-machine-learning/#35a503e543ef; Liisa Jaakonsaari, *Who Sets the Agenda on Algorithmic Accountability?*, EURACTIV (Oct. 29, 2016), https://www.euractiv.com/section/digital/opinion/who-sets-the-agenda-on-algorithmic-accountability/; Nick Wallace, *EU's Right to Explanation: A Harmful Restriction on Artificial Intelligence*, TECHZONE360 (Jan. 25, 2017), http://www.techzone360.com/topics/techzone/articles/2017/01/25/429101-eus-right-explanation-harmful-restriction-artificial-intelligence.htm.

[25] *See* GDPR, *supra* note __, at chs. 6 and 8.

[26] *See id.* The exact multiple can vary depending on the company's annual turnover. *See infra* Part II.

[27] *See infra* Part II and accompanying notes.

[28] *See id.*

quietly gone to work.[29] In the last six months, these authorities have produced an extensive corpus of guidance offering a richly detailed framework for companies seeking to promote compliance with the GDPR's "right to explanation."[30]

Now that the dust from this recent burst of activity by data authorities has begun to settle, this Article attempts to take stock of the new developments—just in time for the Regulation's effectuation in May 2018. In doing so, it seeks to turn the page within the GDPR's fraught "right to explanation" debate by answering a question that has, thus far, gone almost entirely overlooked: What do those actually tasked with enforcing the right think it entails?

Stepping outside of the debate's narrow focus on Chapter 3, this Article adopts a holistic approach to understanding the GDPR's somewhat loosely-worded mandate. It contextualizes the specific provisions pertaining to the "right to explanation" by setting them against the backdrop of the potent range of new administrative capabilities prescribed by Chapters 6 and 8 that, effectively, render Europe's data protection agencies *de facto* interpretive authorities.[31] In adopting this approach, it takes particular pains to let the words of the Regulation and its downstream interpreters speak for themselves—making use of direct quotes or passages whenever possible, so as to minimize the likelihood of editorializing.[32]

Through the words of the authorities that will soon take the lead in enforcing the GDPR, this Article finds a muscular "right to explanation" enshrined within the Regulation—albeit one that is subtly different from the competing visions contemplated by many scholars and industry experts. Europe's data protection authorities consistently reveal that they envisage the "right to explanation" not only as an individual remedial mechanism, but also as a general form of oversight with broad implications for the design and deployment of automated systems that process personal data.[33]

---

[29] *See supra* note __ and accompanying text; *infra* Part III.

[30] *See infra* Part III and accompanying notes.

[31] *See infra* Part II and accompanying notes.

[32] The use of the verb "minimize," as opposed to "eliminate," is intentional—as even the choice of quoted text involves a process of editorial selection.

[33] *See infra* Part II and accompanying notes.

This Article seeks to better understand this newly articulated "right to explanation" and, in doing so, sheds light on how enterprises can prepare for, react to, and promote compliance with what will doubtless be one of the most influential data protection frameworks of the coming decades. It proceeds in five parts. Part I traces the history of the public debate surrounding the "right to explanation." It begins with the right's origins in the specific text of Chapter 3, and proceeds to overview several of the most prominent contributions to the public debate thus far. In highlighting the debate's relative merits and demerits, it argues that the participants' general failure to countenance the substantive changes to enforcement introduced by Chapters 6 and 8 of the Regulation represents a fundamental oversight—one that has hindered a genuine understanding of the right's substantive protections.

Part II turns the page in the debate by broadening its focus to include Chapters 6 and 8 of the GDPR. It contextualizes the newfound role that enforcement agencies will play by detailing their current limitations under the Data Protection Directive and outlining their vastly enhanced administrative powers granted by Chapters 6 and 8. It argues that these newly empowered data watchdogs will serve as *functional* interpretive authorities of the GDPR's "right to explanation," even if other legislative or judicial authorities may, theoretically, have the final say. Because these agencies will be on the front lines of enforcement, their interpretations will, of necessity, be the most relevant for enterprises seeking to promote GDPR compliance. Fortunately, these very agencies have recently produced extensive guidance describing their interpretations of the "right to explanation" that offers powerful insights into the substantive protections afforded by the GDPR's vaguely-worded mandate.

Part III details this newly issued guidance and summarizes its implications for companies seeking to better understand what compliance with the GDPR's "right to explanation" actually entails. It reveals that Europe's data authorities have repeatedly envisioned the "right to explanation" as a robust data protection whose true power lies in its synergistic combination with the "data protection by design" principles codified in the Regulation's subsequent chapters. In examining this new interpretation of the "right to explanation," it argues that data auditing methodologies designed to safeguard against algorithmic bias throughout the entire product life cycle will likely become the new norm for promoting compliance in automated systems. It further argues that this more general version of a "right to explanation" offers greater hope of promoting genuine

"algorithmic accountability" than the individualized remedial mechanism many commentators have presumed it to be.

Part IV concludes the piece by examining the global implications of the GDPR for companies and countries—both inside and outside of Europe—grappling with compliance. It argues that the Regulation will likely have an outsized extraterritorial impact due, not only to the well-documented "Brussels Effect," but also due to the introduction of several legal mechanisms with implications for entities operating outside of the EU. Thanks to the far-flung legal reach of the Regulation, it argues that the "right to explanation"—as envisioned by the GDPR's enforcement authorities—appears destined to become part of a new global data protection standard for companies handling personal information. And while this new standard will certainly pose challenges for enterprises seeking to deploy sophisticated algorithms, it argues that the speed and scale of the global response so far is cause for genuine optimism among those who hope for a future with more fair, accountable, and transparent automated decision-making systems.

I. DOES THE GDPR ENVISAGE A RIGHT TO EXPLANATION?

In January 2012, the European Commission made global headlines by submitting a proposal to "update and modernise the principles enshrined in the 1995 Data Protection Directive."[34] For seventeen years, the DPD had reigned as Europe's preeminent legislation governing the processing of digital data. But after nearly two decades, the longstanding Directive was beginning to show signs of age. As the Commission noted, "technological progress [had] profoundly changed the way [] data is collected, accessed and used" since the DPD's original passage, at a time when "less than 1% of Europeans used the internet."[35]

The press release accompanying the Commission's announcement set the stage for "a comprehensive reform of [the DPD's] data protection rules" designed "to increase users' control of their data," to "provide[] for increased responsibility and accountability for those processing personal data," and to create a "single set of rules" that would be "valid across the

[34] *See* GDPR Proposal, *supra* note __.
[35] *See* GDPR Proposal, *supra* note __.

EU."[36] More than three years of negotiations followed the preliminary proposal, eventually culminating in the formal adoption of the General Data Protections Regulation ("GDPR") in April of 2016.[37] The finalized Regulation constituted a major overhaul of European data processing standards, enumerating a litany of powerful protections intended to make the EU bloc "fit for the digital age."[38]

One such protection—stemming from a cluster of provisions located within Chapter 3 of the GDPR—sets forth what the Regulation describes as the "right not to be subject to a decision based solely on automated processing."[39] The protection establishes a number of safeguards designed to ensure the "fair and transparent processing" of personal data, including an obligation that entities provide "meaningful information about the logic involved" in certain types of highly automated decision-making systems.[40] The protection's mandated disclosure of "meaningful information" has led it to be variously characterized as enshrining a "right to information," a "right to be informed," or, most commonly, a "right to explanation."[41]

---

[36] *See id.*; CONSOLIDATED VERSION OF THE TREATY ON THE FUNCTIONING OF THE EUROPEAN UNION art. 288, 2008 O.J. C 115/47 (emphasis removed).

[37] *See* GDPR, *supra* note __.

[38] EUROPEAN COMMISSION, REFORM OF EU DATA PROTECTION RULES (2017), http://ec.europa.eu/justice/data-protection/reform/ index_en.htm.

[39] *See* GDPR, *supra* note __, at arts. 13(2)(f), 14(2)(g), 15(1)(h), 22.

[40] *See id.*

[41] *See, e.g.,* Bryce Goodman & Seth Flaxman, *EU Regulations on Algorithmic Decision Making and "a Right to an Explanation*," ICML WORKSHOP ON HUMAN INTERPRETABILITY IN ML (2016), arXiv:1606.08813 (v1); Sandra Wachter et al., *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, 7 INT'L DATA PRIV. L. 2, 76–99 (2017), doi:10.1093/idpl/ipx005; Andrew Selbst & Julia Powles, *Meaningful Information and the Right to Explanation*, 7 International Data Privacy Law 3; ARTICLE 29 WORKING PARTY, GUIDELINES ON AUTOMATED INDIVIDUAL DECISION-MAKING AND PROFILING FOR THE PURPOSES OF REGULATION 2016/679 (2017), http://ec.europa.eu/newsroom/just/document.cfm?doc_id=47963 [hereinafter "A29WP Automated Decision-Making Guidelines"].

As the first[42] piece of legislation to explicitly gesture toward such a right, the substantive protections that eventually flow from it will set a precedent with ramifications extending far beyond the technology sector. While the usual suspects, such as Facebook, may have recently grabbed global headlines by announcing millions of dollars spent toward promoting GDPR compliance, the rapid proliferation of machine learning technology across diverse industries indicates that the right will soon be felt across vast swaths of the private sector. Depending on how the protection is eventually applied in practice, it could have profound implications for the use of some of the most powerful computational techniques available to modern enterprises. But as is true of many protections enshrined within the legislative text of the GDPR, the precise reach of the right is far from certain. A careful examination of the language giving rise to the protection provides a useful starting point for understanding and contextualizing it.

A. Specific Text Giving Rise to the "Right to Explanation"

Article 22 of the GDPR grants all data subjects[43] a rebuttable[44] "right not to be subject to a decision based solely[45] on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her."[46] The GDPR defines "processing" as:[47]

> any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated

---

[42] This could be better phrased as the first piece of legislation *with meaningful threat of enforcement* to gesture toward such a right—a nuance that is covered in greater detail in Part II *infra*.

[43] *See* GDPR, *supra* note __, at art. 4. The GDPR defines "data subjects" as "any identified or identifiable natural person" and "an identifiable natural person" as "one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person." *See id*.

[44] *See* GDPR, *supra* note __, at art. 22(2)–(4) specifying limited circumstances where automated decision-making is permitted, and providing for different data safeguards.

[45] This term has recently been subject to clarification. *See infra* notes __–__ and accompanying text.

[46] *See* GDPR, *supra* note __, at art. 22.

[47] *See* GDPR, *supra* note __, at art. 4.

> means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction[.]

The GDPR's use of the term "profiling" introduces a relatively novel concept under EU data protection law.[48] The regulation defines the term as:[49]

> any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements[.]

Article 22(2) enumerates a limited number of circumstances in which companies[50] processing personal data are exempt from its prohibitions—including when automated decision-making is done consensually or is necessary for contracting.[51] But even in such instances, Article 22 requires that companies nevertheless "implement suitable

---

[48] *See* Frederike Kaltheuner and Elettra Bietti, *Data is Power: Towards Additional Guidance on Profiling and Automated Decision-Making in the GDPR* (2018), 2 J. OF INFO. RIGHTS POL'Y AND PRACTICE 1, 2-5.

[49] *See* GDPR, *supra* note __, at art. 4. Recital 71 of the GDPR adds: "Such processing includes 'profiling' that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyse or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behaviour, location or movements, where it produces legal effects concerning him or her or similarly significantly affects him or her." *See* GDPR, *supra* note __, at recital 71. *See also* Mireille Hildebrandt, *Defining Profiling: A New Type of Knowledge?*, in PROFILING THE EUROPEAN CITIZEN 17 (Mireille Hildebrandt & Serge Gutwirth eds., Springer 2008).

[50] *See* GDPR, *supra* note __, at art. 4. The GDPR does not single out companies, but instead uses the term "controller" which "means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data; where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law." *See id.*

[51] *See* GDPR, *supra* note __, at art. 22(2)–(4).

measures to safeguard the data subject's rights and freedoms and legitimate interests" which, at a minimum, includes the subject's "the right to obtain human intervention on the part of the [company]" or "to express his or her point of view and to contest the decision."[52]

Article 22's protections are buttressed by those located within Articles 13–15, pertaining to the rights of data subjects whose personal information is directly or indirectly implicated by automated processing techniques. According to the GDPR's text, these Articles are intended to "provide the data subject with the . . . information necessary to ensure fair and transparent processing."[53] In fulfilling this goal, Articles 13(2)(f), 14(2)(g), and 15(1)(h) mandate that companies provide subjects with information regarding "the existence of automated decision-making, including profiling, referred to in Article 22 . . . and, at least in those cases, *meaningful information about the logic involved*, as well as the significance and the envisaged consequences of such processing for the data subject."[54]

In additional to the text of the GDPR itself, the accompanying nonbinding Recital 71 offers further clarification regarding the regulation's protections pertaining to automated decision-making.[55] The Recital states

---

[52] *See* GDPR, *supra* note __, at art. 22.

[53] *See* GDPR, *supra* note __, at art. 13–15.

[54] *See id.* (emphasis added). This disclosure requirement extends even to data subjects whose personal information has not been directly obtained by a company.

[55] *See* Tadas Klimas & Jurate Vaiciukaite, *The Law of Recitals in European Community Legislation*, 15 ILSA J. OF INT'L AND COMP. LAW 1 61, 92 (2008). Recitals in EU law lack "independent legal value, but they can expand an ambiguous provision's scope. . . . They cannot, however, restrict an unambiguous provision's scope, but they can be used to determine the nature of a provision, and this can have a restrictive effect." *See id.* "Recitals explain the background to the legislation and the aims and objectives of the legislation. They are, therefore, important to an understanding of the legislation which follows." EUROPA, 'Guide to the Approximation of EU Environmental Legislation ANNEX I' (Environment, 2015) accessed 3 March 2017. See also Judgement of 15 5 1997 - Case C-355/95 P Textilwerke Deggendorf GmbH (TWD) v Commission of the European Communities and Federal Republic of Germany [1997] European Court of Justice C-355/95 P [21]: "In that regard, it should be stated that the operative part of an act is indissociably linked to the statement of reasons for it, so that, when it has to be interpreted, account must be taken of the reasons which led to its adoption."

that the data processing techniques implicating personal data "should be subject to suitable safeguards, which should include [the provision of] specific information to the data subject," as well as the rights "to obtain human intervention," "to express his or her point of view," "to *obtain an explanation of the decision reached* after such assessment," and "to challenge the decision."[56] The Recital further stipulates:[57]

> In order to ensure fair and transparent processing . . . [companies] should use appropriate mathematical or statistical procedures for the profiling, implement technical and organisational measures appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimised, secure personal data in a manner that takes account of the potential risks involved for the interests and rights of the data subject and that prevents, inter alia, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, or that result in measures having such an effect.

While the authority of the Recital is nonbinding under EU law, it nonetheless provides a critical reference point for future interpretations by data protection agencies, as well as for co-determinations of positive law that may be made by legislators, courts, and other authorities.[58]

European Court of Justice (ECJ) jurisprudence reveals that the role of Recitals is "to dissolve ambiguity in the operative text of a framework." *See* Watcher et. al, *supra* note __. According to the ECJ:

> "Whilst a recital in the preamble to a regulation may cast light on the interpretation to be given to a legal rule, it cannot in itself constitute such a rule." Case 215/88 Casa Fleischhandels [1989] European Court of Justice ECR 2789 [31].

See also Roberto Baratta, *Complexity of EU Law in the Domestic Implementing Process*, 2 THE THEORY AND PRACTICE OF LEGISLATION 293 (2014); Tadas Klimas and Jurate Vaiciukaite, *The Law of Recitals in European Community Legislation* 15 ILSA Journal of International & Comparative Law 32–33 (2008).

[56] *See* GDPR, *supra* note __, at recital 71 (emphasis added).

[57] *See id.*

[58] These authorities, among other, include the GDPR's designated "Supervisory Authorities," the Article 29 Working Party, the European Data Protection Board, the European Data Protection Supervisor, and the European Data Protection Supervisor's Ethics Advisory Group.

B. The "Right to Explanation" Debate

Despite the GDPR's concerted efforts to detail the protections enshrined under Articles 13, 14, 15, and 22, much uncertainty continues to shroud the Regulation's so-called "right to explanation." This phenomenon owes, in large part, to the GDPR's somewhat fuzzy mandate that entities "ensure fair and transparent processing" by providing "meaningful information about the logic involved" in automated decision-making systems. At a minimum, the protection appears to envisage a limited right for data subjects to understand and verify the basic functionality of certain automated decision-making systems. But beyond that minimum threshold, the precise scope and force of the "right to explanation" has been the subject of much speculation—giving rise to an "explosive" public debate that has swept across the global community.[59]

Among the most prominent contributions to the debate, thus far, have been three distinct perspectives originating from scholars within the U.K. and the U.S.[60] Their claims and critiques are set forth below.

1. The Original Claim

Goodman's and Flaxman's conference paper—'European Union Regulations on Algorithmic Decision-making and a "Right to Explanation"'—first popularized the knotty, sometimes vexing, issues at the heart of the GDPR's "right to explanation."[61] Published just two months after the Regulation's official release, the piece drew widespread attention to the technical and societal challenges inherent to "explain[ing] an algorithm's decision" when made by a so-called "black box" system.[62]

---

[59] *See* Selbst & Powles, *supra* note __.

[60] Many other contributors beyond these three have also thrown their hat in the ring.

[61] *See* Goodman & Flaxman, *supra* note __. It should be noted that this paper was subsequently revised.

[62] *See* FRANK PASQUALE, THE BLACK BOX SOCIETY 3–4 (2015); Brenda Reddix-Smalls, *Credit Scoring and Trade Secrecy: An Algorithmic Quagmire or How the Lack of Transparency in Complex Financial Models Scuttled the Finance Market*, 12 U.C. DAVIS BUS. L. J. 87 (2011); Frank Pasquale, *Restoring Transparency to Automated Authority*, 9 J. ON TELECOMM. & HIGH TECH. L. 235, 237 (2011); Maayan Perel & Nova Elkin-Koren, *Accountability in Algorithmic Copyright Enforcement*, 19 STAN. TECH. L. REV. 473, 478 (2016); NICHOLAS

Goodman and Flaxman observed that the algorithms of past decades tended to rely on explicit, rules-based logic for processing information—an architecture that typically made explaining the system's underlying decision-making relatively straightforward. But, crucially, the scholars noted that many of the most powerful contemporary algorithms instead relied on "models exhibiting implicit, rather than explicit, logic[] usually not optimised for human-understanding"—thereby rendering the logic underlying their decision-making an uninterpretable "black box."[63]

While these types of "black box" algorithms had existed in research labs since the 1970s, Goodman and Flaxman made the prescient observations that their recent proliferation throughout industry presented many challenges for companies and governments seeking to comply with the GDPR. The scholars demonstrated how numerous factors—including potentially biased training sets, "data quality,"[64] the complexity of the most powerful predictive models, and the steep barriers to technical fluency—could pose significant challenges for modern enterprises seeking to promote compliance with the GDPR's mandate of algorithmic explicability.[65]

Although the scholars' work was widely crediting with sparking the "right to explanation" debate, their piece was actually less a legal treatise than a technical primer. Their analysis offered relatively little commentary regarding the right's substantive protections and made only a passing reference to the GDPR's newly introduced enforcement provisions. When the piece did, in fact, discuss the "right to explanation" directly, Goodman and Flaxman tended to construe the protection as relatively narrow. Aside

---

DIAKOPOULOS, ALGORITHMIC ACCOUNTABILITY REPORTING: ON THE INVESTIGATION OF BLACK BOXES (Tow Centre for Digital Journalism, 2013); Goodman & Flaxman, *supra* note __.

[63] *See id.* Machine learning techniques that explicitly encode logic do exist—particularly in the natural language processing and in bioinformatics realms—but are not focused on for purposes of concision.

[64] "Data quality" is a broadly construed term whose components include "accuracy, precision, completeness, consistency, validity, and timeliness, though this catalog of features is far from settled." *See* Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For*, 16 Duke L. & Tech. Rev. 21. *See also generally, e.g.,* Luciano Floridi, I*nformation Quality*, 26 PHIL. & TECH. 1 (2013); Richard Y. Wang & Diane M. Strong, *Beyond Accuracy: What Data Quality Means to Data Consumers*, 12 J. MGMT. INFO. SYS. 5 (1996); LARRY P. ENGLISH, INFORMATION QUALITY APPLIED (2009).

[65] *See id.*

from a single loosely-worded sentence in the paper's abstract that received outsized attention, the scholars suggested that the "right to explanation" could be satisfied with a minimalist explanation of input features and predictions sufficient to answer questions such as: "*Is the model more or less likely to recommend a loan if the applicant is a minority?*" or "*Which features play the largest role in prediction?*"[66]

2. The Response

In response to the widespread attention garnered by Goodman's and Flaxman's conference paper, Watcher et al. entered into the public arena with the provocatively titled piece, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation.'[67] The scholars wasted no time taking to the offensive, immediately calling into doubt both the legal existence and the technical feasibility of what Goodman and Flaxman referred to as the GDPR's "right to explanation." Watcher et al.'s contribution offered a richly detailed tour of the Regulation's relevant text and associated Recital—one that reached far greater analytic depths than the technically-oriented conference paper it criticized. Among the scholars' most important contributions was a powerful tripartite framework for distinguishing questions of algorithmic explicability along chronological and functional dimensions—one that has been replicated by numerous researchers since.[68]

But thoroughgoing as Wachter's et. al's analysis may be have been, their focus was also highly selective. Several of their arguments largely ignored key terms within Articles 13, 14, 15 and 22—the most egregious of which involved their disregard of the word "meaningful" as applied to a substantive analysis of the phrase "meaningful information about the logic involved" in automated decision-making.[69]As importantly, their piece barely mentioned the powerful new administrative capabilities introduced by Chapters 6 and 8 of the Regulation. Instead, their discussion of the

---

[66] *See id.* (emphasis added). The scholars offered virtually no substantive support for their argument that the right could be satisfied with these types of explanations.

[67] *See* Wachter et a., *supra* note __.

[68] *See, e.g.*, Edwards & Veale, *supra* note __.

[69] *See* Wachter et a., *supra* note __. The scholars also made a few claims of astonishing scope, include one assertion that, "There are no ambiguities in the language [of the GDPR] that would require further interpretation with regard to the minimum requirements that must be met by data controllers."

GDPR's new enforcement capabilities was limited single footnote.[70] Most strikingly of all, the central thesis advanced by the scholar's piece was outright contradicted by their own analysis. After electing to title their work 'Why a Right to Explanation of Automated Decision-Making *Does Not Exist* in the General Data Protection Regulation'[71] the scholars went on to repeatedly acknowledge that just such a right existed. Rather than calling it a "right to explanation," however, the scholars instead sought to replace it with a phrase of seemingly narrower implications. They insisted that "the GDPR does not . . . implement a right to explanation, but rather [a] 'right to be informed'"—a distinction that their own subsequent analysis revealed to be of little more than semantic significance.[72]

3. The Rebuttal

In November 2017—with the GDPR just 6 months away and the "right to explanation" debate rapidly rising to a fevered pitch—Selbst and Powles entered into the fray with a point-by-point takedown of Watcher et al. Their contribution sought to address what they described as the numerous "unfounded assumptions and unsettling implications of [Watcher et al.'s] analytical frame."[73] In doing so, Selbst and Powles "offer[ed] a positive conception of the right [to explanation] located in the text and purpose of the GDPR" that they convincingly argued "should be interpreted functionally, flexibly, and should, at a minimum, enable a data subject to exercise his or her rights under the GDPR and human rights law."[74]

Selbst's and Powles's piece represented a vital course correction in a public debate that had begun to more closely resemble a rebranding effort than an actual refutation of the substantive right itself.[75] But much like the scholars they criticized, Selbst and Powles all but disregarded the profound

---

[70] *See* Wachter et a., *supra* note __ at 47.

[71] *See id.* (emphasis added). The scholars Selbst and Powles correctly noted that this tactic was "not only disingenuous but dangerous, as it invites less scrupulous or more time pressed advocates to cite the paper for the proposition that there is no right to explanation, which is not even what the paper argues in substance." *See* Selbst and Powles *supra* note __.

[72] *See* Selbst & Powles *supra* note __.

[73] Numerous of these criticisms are outlined in the section above.

[74] *See* Selbst & Powles *supra* note __.

[75] Watcher et al.'s piece continues to enjoy widespread popularity among more casual observers—with many remaining unaware of the important counterweight provided by Selbst & Powles.

changes to data enforcement ushered in by Chapters 6 and 8 of the GDPR. As a result, the scholars paid undeservedly little attention to the perspectives of the agencies that would soon be tasked with enforcing the Regulation, despite the fact that their piece was written after the release of extensive GDPR guidance by Europe's most influential data protection authority.[76]

C. Lost in the Fog of Battle

Since its origins with Goodman and Flaxman, the GDPR's "right to explanation" debate has fostered a conversation of profound global significance—exploring the economic benefits, technical feasibility, and social tradeoffs of applying "algorithmic accountability" practices in enterprise and government.[77] The contributions of Goodman, Flaxman, Selbst, Powles, and Watcher et. al constitute just a tiny sample of the vast, and impressively diverse, array of perspectives on offer. Over a period of just eighteen months, countless industry leaders, media sources, and researchers of various backgrounds have also contributed their unique perspectives.[78] But mystifyingly, many of the most distinguished

---

[76] *See* Selbst & Powles *supra* note __.

[77] *See infra* notes __ and accompanying text.

[78] *See e.g.*, Rich Caruana et al., *Intelligible Models for Healthcare: Predicting Pneumonia Risk and Hospital 30-Day Readmission* (2015), PROCEEDINGS OF THE 21ST ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING (KDD '15) 1721; David Bamman, *Interpretability in Human-Centered Data Science* (2016), CSCW WORKSHOP ON HUMAN-CENTERED DATA SCIENCE; Michael Gleicher, *A Framework for Considering Comprehensibility in Modeling* (2016) 4 Big Data 4, 75; Finale Doshi-Valez & Been Kim, *A Roadmap for a Rigorous Science of Interpretability* (2017), arXiv:1702.08608; Eric Horvitz, *On the Meaningful Understanding of the Logic of Automated Decision Making,* BCLT PRIVACY LAW FORUM (Mar. 24 2017), https://www.law.berkeley.edu/wpcontent/uploads/2017/03/BCLT_Eric_Horvitz_March_ 2017.pdf; Ethan Chiel, *EU Citizens Might Get a "Right to Explanation" About the Decisions Algorithms Make*, FUSION (Jul. 5, 2016), http://fusion.kinja.com/eu-citizens-might-get-aright-to-explanation-about-the-1793859992; Cade Metz, *Artificial Intelligence Is Setting Up the Internet for a Huge Clash With Europe*, WIRED (Jul. 11, 2016), https://www.wired.com/2016/07/artificial-intelligence-setting-internet-huge-clash-europe/; Ian Sample, *AI Watchdog Needed to Regulate Automated Decision-making, Say Experts,* THE GUARDIAN, (Jan. 27, 2017), https://www.theguardian.com/technology/ 2017/jan/27/ai-artificial-intelligence-watchdog-needed-to-prevent-discriminatoryautomated-decisions; Matt Burgess,

contributions to this multifaceted debate have essentially overlooked what is likely the most profound change of all heralded by the GDPR: the sweeping new enforcement powers granted to EU data protection authorities under the new Regulation.

Beyond the "right to explanation" debate's narrow focus on Articles 13, 14, 15, and 22 lay a series of articles that appear destined to forever change the practical reality of enforcement by data protection authorities. These articles—contained in Chapters 6 and 8 of the Regulation—grant vast new administrative powers to EU watchdog agencies that have long been viewed as toothless under the DPD.[79] Failing to acknowledge the role that these newly empowered agencies will play in enforcing and, therefore, *interpreting* the GDPR's "right to explanation" currently represents a major blind-spot within the public debate—one that, if left unaddressed, risks allowing it to move in an unproductive, and unnecessarily adversarial, direction.

II. TURNING THE PAGE IN THE "RIGHT TO EXPLANATION" DEBATE

Although the introduction of the GDPR will represent the largest overhaul of EU data protections in twenty years, many experts agree that the most revolutionary change ushered in by the Regulation actually involves the addition of a host of new legal mechanisms for promoting enforcement.[80] After all, the EU has long boasted an extensive list of rules[81]

Watching Them, Watching Us: Can We Trust Big Tech to Regulate Itself?, CREATIVE REVIEW, (Apr. 2017), https://www.creativereview.co.uk/watching-watching-us/; ACM US Public Policy Council, Statement on Algorithmic Transparency and Accountability, ACM US PUB. POL'Y COUNCIL & ACM EUROPE COUNCIL (May 25 2017), http://www.acm.org/binaries/content/assets/public-policy/2017_joint_statement_algorithms.pdf.

[79] *See* GDPR, *supra* note __, at chs. 6 and 8.

[80] *See* Natasha Lomas, *WTF is GDPR*, TECHCRUNCH, January 2018, https://techcrunch.com/2018/01/20/wtf-is-gdpr/.

[81] In addition to the DPD, there are numerous other regulations that allude to rights involving automated decision-making explicability. "For example, the public sector is subject to the Public Administration Act that requires, inter alia, individual decisions to be substantiated. The person concerned has the right to be informed of the regulations and the actual circumstances underpinning a decision, as well as the main considerations that have been decisive." *See* Wachter et at., *supra* note __ (quoting Public Administration Act: Sections 24 and 25). The EU also explicitly treats privacy protection as a fundamental right.

that set a high bar for data protection—including rights that specifically address automated decision-making.[82] What these rules have lacked, however, is a meaningful threat of enforcement.[83]

Under the current Data Protection Directive ("DPD"), EU agencies tasked with carrying out its mandate are highly limited in their capacity to levy financial penalties against entities in breach.[84] The UK's Information Commissioner's Office ("ICO"), for example, is capped at a maximum fine of just £500,000 for violations. Facebook's annual revenue for the 2017 fiscal year, by comparison, topped £28.3B—meaning that, at most, the ICO could hope to impose a fine representing a paltry 0.000018% of the company's annual turnover.[85]

---

[82] *See* DPD, *supra* note __. *See also, e.g.,* Izak Mendoza & Lee A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, in EU INTERNET LAW: REGULATION AND ENFORCEMENT 2 (Tatiani Synodinou et al. eds., Springer, forthcoming 2017); Lee A. Bygrave, *Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling*, 17 COMPUTER L. & SECURITY REP., 17 (2001); Alfred Kobsa, T*ailoring Privacy to Users' Needs*, in USER MODELING, (M. Bauer et al. eds., Springer 2001), doi:10.1007/3-540-44566-8_52; Mireille Hildebrandt, *Profiling and the rule of law*, 1 IDENTITY IN THE INFORMATION SOCIETY 1, 55 (2008), doi:10.1007/s12394-008-0003-1.

Wachter, et al. actually discuss this phenomenon, noting:

> Interestingly, despite years of negotiations, the final wording of the GDPR concerning protections against profiling and automated decision-making hardly changed from the relevant Articles and Recitals of the Data Protection Directive.

But their failure to address the enhanced enforcement powers introduced by the GDPR renders moot their underlying argument that the new provisions will do little to change the current regulatory landscape.

[83] *See* Izak Mendoza & Lee A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, in EU INTERNET LAW: REGULATION AND ENFORCEMENT 2 (Tatiani Synodinou et al. eds., Springer, forthcoming 2017) (describing art. 15 as "a second class data protection right: rarely enforced, poorly understood and easily circumvented").

[84] *See id.*

[85] *See* FACEBOOK INVESTOR RELATIONS, FACEBOOK REPORTS FOURTH QUARTER AND FULL YEAR 2017 RESULTS', FACEBOOK (2018), https://investor.fb.com/investor-news/press-release-details/2018/Facebook-Reports-Fourth-Quarter-and-Full-Year-2017-Results/default.aspx.

Moreover, the replacement of the DPD with the GDPR represents an instance of an EU Regulation replacing a Directive. While directives "set out general rules to be transferred into national law by each country as they deem appropriate," regulations constitute a single, uniform law that is "directly applicable" to all EU Member States.[86] The differences between these two paths to legislative implementation may seem trivial to outsiders looking in, but their practical effects are not. Unlike the GDPR, the DPD is currently subject to twenty-eight different interpretations and enforcement regimes—leading to differences that can foment confusion and inconsistency among industry leaders and data protection authorities alike.

The combined effect of these DPD enforcement limitations has produced a pack of EU data watchdogs tethered to a markedly short regulatory leash. For over two decades, the Directive has set a high standard for data protection for companies handling the personal information of EU citizens. But the watchdogs responsible for actually enforcing these protections have long been perceived as lacking real bite.

A. New Watchdogs on the Bloc

Viewed through this lens, it is easy to understand why the debate surrounding the "right to explanation" has seen comparatively little attention paid to the authorities that will actually be tasked with enforcing it. For if the past were prologue, they could be expected to play a decidedly peripheral role in carrying out the right's protective mandate.

With the passage of the GDPR, however, all of that is set to change, thanks to Chapters 6 and 8 of the Regulation which grant data authorities vastly increased investigatory powers, an enhanced "enforcement tool kit," and the capacity to levy far greater financial penalties against entities in breach.[87] No longer will EU data authorities be constrained by the limited range of enforcement options available under the DPD. Instead, these authorities will have far-reaching investigatory and corrective powers that allow them to issue sanctions against data protection violations that are

---

[86] Art. 288 of the Treaty on the Functioning of the European Union provides that: "A directive shall be binding, as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods." The Article states that a regulation, on the other hand, "shall be binding in its entirety and directly applicable in all Member States."

[87] *See infra* Part II(B) and accompanying notes.

"effective, proportionate," and, most importantly, "dissuasive." The significance of this paradigm shift is difficult to overstate. Indeed, it has led some commentators to assert that the transition from the DPD to the GDPR should be understood as less about "individual EU Member States . . . getting stronger privacy laws" and more about EU data authorities finally starting "to bark and bite like proper watchdogs."[88]

Lamentably, the practical implications of this shift in power have been largely drowned out by the sound and fury generated over rarified textual interpretations of Chapter 3 of the GDPR. If only the experts would turn the page to Chapters 6 and 8, which have quietly announced the dawn of a new era of data enforcement—one with profound implications for the ultimate resolution of the "right to explanation" debate.

B. The Ascent of Enforcement

Chapter 6 of the GDPR provides for the appointment, by each Member State, of "one or more independent public authorities to be responsible for monitoring [its] application."[89] The legislation endows these agencies—which it terms "supervisory authorities" ("SAs")—with broad "investigatory," "advisory," and "corrective" powers of far greater scope than those currently available under the DPD. According to Chapter 6, these powers are designed to ensure the "consistent application" of the GDPR throughout the EU and include, among many other provisions, the ability: (1) "to obtain . . . access to all personal data [belonging to a company] and to all information necessary for the performance of [investigatory] tasks," (2) "to carry out investigations in the form of data protection audits,"[90] (3) "to issue warnings [or] reprimands to a [company]," (4) "to impose a temporary or definitive limitation [against companies] including a ban on processing," and (5) "to order the suspension of data flows to a recipient in a third country[91] or to an international organization."[92]

Chapter 6's expansive set of investigatory and corrective powers are buttressed by an equally expansive set of remedial powers laid out in Chapter 8—intended to provide supervisory agencies with the authority to

---

[88] *See* Natasha Lomas, *WTF is GDPR*, TECHCRUNCH, January 2018, https://techcrunch.com/2018/01/20/wtf-is-gdpr/.

[89] *See* GDPR, *supra* note __, at ch. 6.

[90] Data protection audits are discussed in greater detail *infra*.

[91] This term is discussed in detail in Part IV *infra*.

[92] *See* GDPR, *supra* note __, at art. 58.

impose administrative fines that are "effective, proportionate, and dissuasive."[93] Under Chapter 8, SAs can fine companies that violate the GDPR's basic administrative or technical requirements up to €10 million, or up to 2% of the companies' total annual revenue for the proceeding financial year, "which[ever] is higher."[94] For violations of provisions more fundamental to the GDPR's data protection mandate[95]—including Articles 13, 14, 15, and 22—the maximum allowable fine increases precipitously. SAs can punish infringers of these provisions with fines of up to €20 million, or up to 4% of the companies' total annual revenue for the proceeding financial year—again, "which[ever] is higher."[96]

The operative adjective, in both such instances, is the word "higher." To return to the example of the tech giant Facebook, whose annual revenues approximate €32.1 billion, a fine of 4% of annual turnover could total €1.28 billion—more than 2500 times larger than the maximum fine currently afforded under the DPD. This switch from proportional, as opposed to fixed, financial penalties under the GDPR ensures that even the titans of industry will not be immune from enforcement.

But for any in-house practitioners whose pulse doubled at the sight of such a multiple, the regulation also provides cause for relief. First, the GDPR makes clear that punishment for breaches should be individualized and proportionate—and it does not mandate the use of fines for all enforcement actions.[97] Article 83 outlines an extensive list of considerations for SAs seeking to ensure that their punishments are commensurate with the alleged violation.[98] These factors shift the administrative focus to the actual impacts of the violation, including the number of individuals affected, the actual damages suffered, and the sensitivity of the personal data at root.[99] Second, the GDPR also stipulates that good faith efforts to proactively implement protective policies, ensure transparency, notify enforcement

---

[93] *See* GDPR, *supra* note __, at ch. 8.

[94] *See* GDPR, *supra* note __, at art. 84.

[95] "Examples that fall under this category are non-adherence to the core principles of processing personal data, infringement of the rights of data subjects and the transfer of personal data to third countries or international organizations that do not ensure an adequate level of data protection." *See* GDPR, *supra* note __, at art. 84.

[96] *See* GDPR, *supra* note __, at art. 84.

[97] *See* GDPR, *supra* note __, at art. 83.

[98] *See id.*

[99] *See id.*

agencies, and cooperate with SA oversight will further reduce the likelihood of companies facing serious sanction.[100]

C. The Importance of Understanding When the Watchdogs Might Bite

With great power, of course, comes great interpretive responsibility. After all, what better source of guidance could there be for companies seeking to ensure compliance with the GDPR's "right to explanation" than the data authorities likeliest to bring enforcement action against them? Any agency action will, admittedly, be subject to the slower-burning process of judicial clarification through national and international litigation. But while any such activity percolates through the EU's multi-layered legal system, the *de facto* interpretive authorities of the "right to explanation" will be those whose primary responsibility it is to investigate and punish companies in breach.

Already, data protection authorities have begun to signal their anticipated ascendance by flexing additional regulatory muscle in the lead up to the GDPR's effectuation.[101] According to a recent report, the total

---

[100] *See id.*

[101] *See* Max Metzger, *Sharp Rise in ICO Fines and Enforcement Notices as GDPR Races Closer*, SC MEDIA (June 01 2017), https://www.scmagazineuk.com/sharp-rise-in-ico-fines-and-enforcement-notices-as-gdpr-races-closer/article/665466/. *See also, e.g., See* INFORMATION COMMISSIONERS OFFICE (ICO), BIG DATA, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DATA PROTECTION (2017); Elizabeth Denham, the residing commissioner, remarked:

> In this world of big data, AI and machine learning, my office is more relevant than ever. I oversee legislation that demands fair, accurate and non-discriminatory use of personal data; legislation that also gives me the power to conduct audits, order corrective action and issue monetary penalties. Furthermore, under the GDPR my office will be working hard to improve standards in the use of personal data through the implementation of privacy seals and certification schemes. We're uniquely placed to provide the right framework for the regulation of big data, AI and machine learning, and I strongly believe that our efficient, joined-up and co-regulatory approach is exactly what is needed to pull back the curtain in this space.

See also, Jamie Doward, *Watchdog to Launch Inquiry into Misuse of Data in Politics*, THE GUARDIAN (Mar. 4, 2017),

monetary value of fines levied by the UK's ICO doubled in 2016—coinciding with a steep uptick in the number of enforcement notices issued by the agency and a nearly 100% increase in the size of its fines.[102] This increased enforcement activity also came amid calls by the agency to increase its staff size in advance of the GDPR's May 2018 effectuation.[103]

III. THE NEXT CHAPTER IN THE DEBATE: SA ENFORCEMENT AND THE RISE OF DATA AUDITS

Viewed against the backdrop of Chapter 6's and 8's vastly enhanced enforcement powers, it becomes immediately apparent that the public debate over the "right to explanation" can no longer be confined exclusively to Chapter 3. Instead, the right must be understood holistically, with a newfound deference owed to the downstream interpretations by EU data watchdogs whose regulatory bark and bite will soon become far costlier for companies to ignore. Fortunately, a recent burst of activity by these very data authorities has provided extensive guidance for enterprises seeking to better understand what meaningful compliance with the GDPR's controversial "right to explanation" will actually entail in practice.

A. The Interpretation of the Article 29 Data Protection Working Party

In October 2017, the Article 29 Data Protection Working Party ("A29WP") published its official "Guidelines on Automated Individual Decision-Making and Profiling" for the GDPR.[104] The A29WP "is the European Commission's most senior advisory body on data protection and information security matters" and serves as a central authority for all EU data protection agencies.[105] Although its guidelines are nonbinding, they

---

https://www.theguardian.com/technology/2017/mar/04/cambridge-analytics-data-brexittrump.

[102] *See* Metzger *supra* note __.

[103] *See id.*

[104] *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[105] *See* DATATILSYNET: THE NORWEGIAN DATA PROTECTION AUTHORITY, ARTIFICIAL INTELLIGENCE AND PRIVACY (January 2018), https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf. The A29WP, which launched in 1996, derives its name from Article 29 of the DPD setting out its composition and purpose. *See* DPD, *supra* note __ at art. 29. It is a a representative body composed of data protection authorities from each EU Member State and also includes the European Data Protection Supervisor and the

constitute a vital reference point for the individual SAs appointed by EU Member States and are, therefore, critical to understanding how those authorities will eventually interpret the GDPR.

The A29WP's guidance on automated decision-making included numerous provisions intended to clarify the elusive "right to explanation"—stemming from a collection of rights that the A29WP referred to as the rights "to be informed," "to obtain human intervention," and "to challenge [a] decision" made by certain automated systems.[106] According to the A29WP, the "complexity of machine-learning" algorithms used in such systems "can make it challenging to understand how an automated decision-making process or profiling works." But such complexity, it insisted, "is no excuse for failing to provide information" to data subjects.[107] The A29WP instructed that companies making automated decisions which fall under Article 22(1) "should find simple ways to tell the data subject about the rationale behind, or the criteria relied on in reaching the decision"—albeit "without necessarily always attempting a complex explanation of the algorithms used or [a] disclosure of the full algorithm."[108] In doing so, the A29WP stipulated that companies must:[109]

- tell the data subject that they are engaging in this type of activity;
- provide meaningful information about the logic involved; and
- explain the significance and envisaged consequences of the processing.

The A29WP further clarified that the phrase "[m]eaningful information about the logic involved will in most cases require controllers to provide details such as:"[110]

- the information used in the automated decision-making process, including the categories of data used in a profile;
- the source of that information;

European Commission. Once the GDPR takes effect, it will be replaced by the "European Data Protection Board." *See* GDPR *supra* note __.

[106] *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[107] *See id.*

[108] *See id.*

[109] *See id.*

[110] *See id.*

- how any profile used in the automated decision-making process is built, including any statistics used in the analysis;
- why this profile is relevant to the automated decision-making process; and
- how it is used for a decision concerning the data subject.

The A29WP added that it was "good practice [for companies] to provide the above information *whether or not* the processing falls within the narrow Article 22(1) definition."[111] The agency also insisted that companies could not avoid Article 22 by simply "fabricating" *de minimus* human involvement in decision-making.[112] According to the A29WP, companies must ensure that any human "oversight of [a] decision is meaningful, rather than just a token gesture" if they intend for their systems to fall outside the scope of Article 22's provisions pertaining to decisions "based *solely* on automated processing."[113]

In addition to the specific explanatory measures outlined above, the A29WP also recommended that companies introduce more general "procedures and measures to prevent errors, inaccuracies or discrimination" in data processing.[114] The guidelines suggested that companies "carry out

---

[111] *See id.* (emphasis added). This justification stemmed, in part, from GDPR Recital 60 stating:

> The controller should provide the data subject with any further information necessary to ensure fair and transparent processing taking into account the specific circumstances and context in which the personal data are processed. Furthermore the data subject should be informed of the existence of profiling and the consequences of such profiling.

[112] *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[113] *See id.* (emphasis added). This question, too, has been the subject of heated debate due to Article 22's use of the phrase "solely" in its provisions related to automated decision-making. *See, e.g.,* Watchter et al. *supra* note __; Selbst & Powles *supra* note __. The A29WP further clarified that:

> To qualify as human intervention, the [data] controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the available input and output data.

[114] *See id.*

frequent assessments on the data sets they process to check for any bias, and develop ways to address any prejudicial elements, including any over-reliance on correlations."[115] According to the A29WP, these assessments should be conducted "on a cyclical basis; not only at the design stage, but also continuously, as the profiling is applied to individuals," so that the "outcome of such testing [can] feed back into the system design."[116]

One such safeguard repeatedly invoked by the A29WP involves the use of the "Data Protection Impact Assessment" ("DPIA"), originating under Article 35 of the GDPR.[117] Although the GDPR does not formally define the concept of the DPIA, the A29WP described it as "a process for building and demonstrating" compliance by systematically examining automated processing techniques to determine the measures necessary to "manage the risks to the rights and freedoms of natural persons resulting from the processing of personal data."[118] Article 35(7) of the GDPR enumerates four basic features that all DPIAs must, at a minimum, contain:[119]

> 1. a systematic description of the envisaged processing operations and the purposes of the processing, including, where applicable, the legitimate interest pursued by the controller;
> 2. an assessment of the necessity and proportionality of the processing operations in relation to the purposes;
> 3. an assessment of the risks to the rights and freedoms of data subjects[; and]
> 4. the measures envisaged to address the risks, including safeguards, security measures and mechanisms to ensure the protection of personal data and to demonstrate compliance with this Regulation taking into account the rights and legitimate interests of data subjects and other persons concerned.

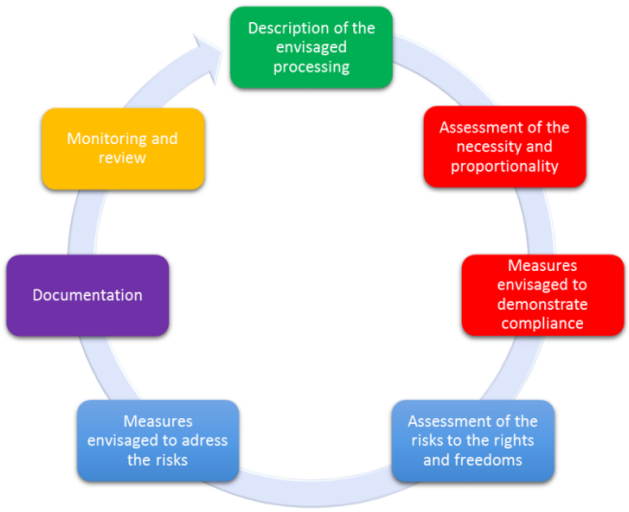
While noting that the GDPR provides companies with considerable "flexibility to determine the precise structure and form of the DPIA," the A29WP stipulated that the DPIA represented a fundamentally "iterative

---

[115] *See id.*
[116] *See id.*
[117] *See id.*
[118] *See id.*
[119] *See id.*

process" with "common criteria" for carrying it out.[120] According to the A29WP, these criteria were best understood as falling within the GDPR's broader "data protection by design" principles which apply at all stages of a system's life cycle.[121]

FIGURE 3.1: THE ITERATIVE DPIA PROCESS



Under the GDPR's "data protection by design" mandate, companies must "tak[e] into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by [] processing."[122] The A29WP recommends DPIAs as a means of proactively identifying and addressing these considerations, so that companies can effectively "implement appropriate technical and organisational . . . safeguards into the[ir] processing [operations]."[123]

1. When Are DPIAs More Than Mere Recommendations?

---

[120] The A29WP's Annex's 1 and 2 provide additional details regarding these requirements. *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[121] *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[122] *See id.*

[123] *See id.* The A29WP Guidelines explicitly recommend "measures, such as pseudonymisation, which are designed to implement data-protection principles, [and] data minimization." *See id.*

The A29WP's guidance stresses that, in many circumstances, DPIAs are not merely recommended as a matter of best practices but are compulsory. In determining whether a DPIA is or is not compulsory, Article 35(1) of the GDPR relies, primarily, on the heuristic of so-called "high risk" data processing operations.[124] According to the Regulation, DPIAs are mandatory "[w]here a type of processing[,] taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons."[125] Article 35 establishes a non-exhaustive list of scenarios likely to be deemed high risk, including when operations involve:[126]

> a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;
> b) processing on a large scale of special categories of data referred to in Article 9(1),[127] or of personal data relating to criminal convictions and offences referred to in Article 10;[128] or

---

[124] *See* GDPR, *supra* note __, at art. 35.

[125] *See id.*

[126] *See id.* The GDPR notes that the use of "new technologies" is "particularly" likely to produce high risks.

[127] Article 9(1) states:

> Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited.

[128] Article 10 states:

> Processing of personal data relating to criminal convictions and offences or related security measures based on Article 6(1) shall be carried out only under the control of official authority or when the processing is authorised by Union or Member State law providing for appropriate safeguards for the rights and freedoms of data subjects. Any comprehensive register of criminal convictions shall be kept only under the control of official authority.

c) a systematic monitoring of a publicly accessible area on a large scale.

The A29WP's guidance elaborates on this list by enumerating ten specific scenarios that "provide a more concrete" set of criteria for determining whether operations are "high risk." These include instances where processing involves: (1) evaluating or scoring, (2) automated decision-making with legal or similarly significant effects, (3) systematic monitoring, (4) sensitive data, (5) data processed on a large scale, (6) datasets that have been matched or combined, (7) data concerning vulnerable data subjects, (8) innovative use or applying technological or organizational solutions, (9) data transfer across borders outside the European Union, and (10) processing that inherently "prevents data subjects from exercising a right or using a service or a contract."[129]

Although the A29WP was careful to emphasize that DPIAs are not obligatory "for every processing operation which may result in risks," the GDPR's requirement that an *ex ante* assessment be conducted for all processing operations produces a distinctly circular effect. In cases where it is unclear whether a given operation requires a DPIA, carrying out a preliminary DPIA to assess the risks may be the best means of ensuring compliance. In other words, demonstrating that a DPIA is not necessary will, in many instances, itself require a DPIA.[130]

Crucially, these *ex ante* assessments are required even when the GDPR's provisions pertaining to decision-making "based *solely* on automated processes" are not directly implicated.[131] The A29WP repeatedly highlighted that Article 35(3)(a)'s deliberate exclusion of the word "solely" meant that the Article "appl[ied] in the case of decision-making including profiling with legal or similarly significant effects that is *not wholly*

Article 6(1) includes a list of criteria for establishing the lawfulness of processing.

[129] *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[130] The A29WP stressed that: "In order to enhance compliance with this Regulation where processing operations are likely to result in a high risk to the rights and freedoms of natural persons, the controller should be responsible for the carrying-out of a data protection impact assessment to evaluate, in particular, the origin, nature, particularity and severity of that risk." *See* A29WP Automated Decision-Making Guidelines, *supra* note __.

[131] *See id.* (emphasis added).

automated, as well as solely automated decision-making defined in Article 22(1)."[132]

2. What Kinds of Documented "Explanations" Do DPIAs Require?

As a means of promoting additional transparency through DPIAs, the A29WP instructed that when data "processing is wholly or partly performed by a [company]," the company should assist SAs "in carrying out [a] DPIA and provide any necessary information" to them.[133] Moreover, the A29WP emphasized that, under Article 35(9), companies are required, "where appropriate," to actively "seek the views of data subjects or their representatives" during the DPIA process.[134] In fulfilling this obligation, the A29WP stated that the views of data subjects could be solicited by a variety of means "depending on the context," including "an internal or external study related to the purpose and means of the processing operation," "a formal question" directed to the relevant stakeholders," or "a survey sent to the data controller's future customers."[135] The A29WP also noted that when a company's "final decision" to proceed with a particular process operation "differ[ed] from the views of the data subjects, its reasons for going ahead or not should be [also] documented."[136] And even in instances where a company has decided that soliciting the views of data subjects is not appropriate, the A29WP insisted that the company should nonetheless document "its justification for not seeking the views of data subjects."[137]

Finally, the A29WP added that, while publicly releasing "a DPIA is not a legal requirement under the GDPR," companies "should consider publishing their DPIA[s]" either in full or in part.[138] The A29WP stated that the "purpose of such a process would be to help foster trust in the controller's processing operations, and demonstrate accountability and transparency"—particularly "where members of the public are affected by the processing operation."[139] According to the institution, the "published DPIA does not need to contain the whole assessment, especially when the

---

[132] *See* A29WP Automated Decision-Making Guidelines, *supra* note __ (emphasis added).

[133] *See id.*

[134] *See id.*

[135] *See id.*

[136] *See id.*

[137] *See id.*

[138] *See id.*

[139] *See id.*

DPIA could present specific information concerning security risks for the data controller or give away trade secrets or commercially sensitive information" and "could even consist of just a summary of the DPIA's main findings."[140]

B. From the A29WP to Supervisory Authorities

From the central guidance provided by the A29WP comes the specific downstream interpretations of data authorities appointed by individual Member States. Although the individual interpretations of these SAs are, by design, the furthest from the textual wellspring of the GDPR, they are far and away the most relevant for companies seeking to promote compliance. As the agencies on the front lines of overseeing investigations and issuing sanctions, the interpretations they provide will constitute the clearest signals for companies attempting to understand the substantive protections afforded by the GDPR's "right to explanation."

1. Why the ICO?

The analysis that follows focuses on one such authority—the UK's Information Commissioner's Office ("ICO"). The reasons for its focus on the ICO are at least twofold. First, surveying all twenty-eight agencies would be needlessly exhaustive, as each agency's interpretation draws directly from the GDPR as opposed to drawing indirectly from 28 individual legislative enactments, as is currently the case under the DPD. Second, and most importantly, the UK's imminent exit from the EU makes the ICO a particularly informative example—as the country is seeking to ensure the continuing free flow of data between itself and Continental Europe by promoting domestic compliance with the GDPR, despite the fact that it will soon separate from the European bloc. Thus, the fact that the ICO is, in one sense, a bad example makes it an especially good one.

2. The ICO's Guidance

Since the A29WP's release of its GDPR guidance in October 2017, the ICO—along with every other EU data authority—has published extensive guidelines for organizations seeking to comply with the GDPR's

[140] *See id.*

requirements.[141] The agency describes these guidelines as a "living document" subject to elaboration or alteration on an ongoing basis.[142] Among the ICO's many provisions interpreting the GDPR are those pertaining to the data subjects' "rights related to automated decision-making including profiling."[143] According to the ICO, companies processing data "must identify whether any of [their] processing falls under Article 22 and, if so, make sure that" they:[144]

- give individuals information about the processing;
- introduce simple ways for them to request human intervention or challenge a decision;
- carry out regular checks to make sure that your systems are working as intended.

When processing operations fall under Article 22's specific purview,[145] the ICO also requires that companies carry out a DPIA "to identify the risks to individuals," to "show how [they] are going to deal with them," and to demonstrate the "measures [they] have in place to meet GDPR requirements."[146]

---

[141] *See* INFORMATION COMMISSIONER'S OFFICE, GUIDE TO THE GENERAL DATA PROTECTION REGULATION (2018) https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/. The UK Government has also issued new data protection legislation that will implement the standards set forth by the GDPR. *See* GDPR FACT SHEET, BENEFACTO (2018), https://benefacto.org/gdpr-fact-sheet/. These laws include a number additional protections going above and beyond the baseline set by the GDPR which extend to "journalists, scientific and historical researchers, and anti-doping agencies who handle people's personal information." *See id.*

[142] *See id.*

[143] *Seee* INFORMATION COMMISSIONER'S OFFICE, RIGHTS RELATED TO AUTOMATED DECISION MAKING INCLUDING PROFILING (2018) https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/ [hereinafter "ICO Automated Decision Making Guidelines"].

[144] *See id.*

[145] *See id.* Some instances do not apply. See section above.

[146] *See* ICO Automated Decision Making Guidelines, *supra* note __. Even in instances where Article 22's requirements do not apply, the ICO recommends that companies nonetheless "carry out a DPIA to consider and address the risks before [they] start any new automated decision-making or profiling" and "tell [] customers about the profiling and automated decision-making [they] carry out, what information [they] use to create the profiles and where [they] get this information from." *See id.*

Even when processing operations fall outside of Article 22, the ICO's guidelines explicitly endorse the use of a DPIA as part of a broader compliance tool kit based on the same principles of "data protection by design" ("DPbD") identified by the A29WP.[147] In addition to the comprehensive set of recommendations involving DPbD detailed its public discussion paper, 'Big Data, Artificial Intelligence, Machine Learning and Data Protection,'[148] the ICO states that companies "have a general obligation to implement technical and organisational measures to show that [they] have considered and integrated data protection into [their] processing activities."[149]

C. The Rise of the DPIA and Data Protection by Design

From the guidance set forth by the A29WP and the ICO, one fact is made overwhelmingly clear: The GDPR's "right to explanation" is no mere remedial mechanism to be invoked by data subjects on an individual basis, but implies a more general form of oversight with broad implications for the design, prototyping, field testing, and deployment of data processing systems. The "right to explanation" may not require that companies pry open their "black boxes" *per se*, but it does require that they evaluate the interests of relevant stakeholders, understand how their systems process data, and establish policies for documenting and justifying key design features throughout a system's life cycle. Not only must companies convey many of these details directly to downstream data subjects,[150] but they must also document and explain the safeguards in place for managing data processing risks, either through a DPIA as described in Article 35 or through a substantively similar mechanism. Indeed, it is perhaps no coincidence that the formulation of Article 35(1) bears such a striking similarity to that of Article 22(1). Taken together, these two mandates produce a powerful synergistic effect that promotes the kinds of prophylactic "data protection by design" ("DPbD") principles prevalent throughout the GDPR.[151] As a

---

[147] *See* Information Commissioner's Office, DATA PROTECTION BY DESIGN AND DEFAULT (2018), https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-by-design-and-default/.

[148] *See* INFORMATION COMMISSIONERS OFFICE, BIG DATA, ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND DATA PROTECTION (2017).

[149] *See id.*

[150] *See supra* Part III(A).

[151] *See* GDPR, *supra* note __, at art. 25 and Recital 78.

consequence, it now appears that *ex ante* DPIAs—as opposed to *ex poste* invocations of an individual "right to explanation"—are destined to "become the required norm for algorithmic systems, especially where sensitive personal data, such as race or political opinion, is processed on a large scale."[152]

The advantages of shifting the dialogue surrounding the GDPR's "right to explanation" from one involving individual remedies to one involving more general DPbD principles are manifold. First, mere algorithmic explicability is not the panacea it is often presumed to be.[153] As numerous experts of diverse backgrounds have noted, the reliance on transparency as an individualized mechanism often places excessive burdens on resource-constrained users to "seek out information about a system, interpret it, and determine its significance, only then to find out they have little power to change things anyway, being disconnected from

[152] *See* Lilian Edwards & Michael Veale, *Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For*, 16 Duke L. & Tech. Rev. 21, 78 (quoting GDPR, art. 35(3)(b)) (quotations removed). *See also, e.g.,* DATA PROTECTION WORKING PARTY, *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation* 2016/679, Art. 29, WP 248 (Apr. 4, 2017). This prediction involving the rise of data auditing methodologies is also supported by additional legal mechanisms within the GDPR that, for purposes of concision, are not addressed by this Article. *See, e.g.,* GDPR *supra* note __ at art. 42 (requiring "the establishment of data protection certification mechanisms and data protection seals and marks. . . available via a process that is transparent" and subject to regular review); GDPR *supra* note __ at art. 40 (recommending that companies "prepare codes of conduct...such as with regard to fair and transparent processing" and "to carry out the mandatory monitoring of compliance").

[153] *But see* Kasper Lippert-Rasmussen, *"We Are All Different": Statistical Discrimination and the Right to Be Treated as an Individual*, 15 J. ETHICS 47, 54 (2011) ("[O]btaining information is costly, so it is morally justified, all things considered, to treat people on the basis of statistical generalizations even though one knows that, in effect, this will mean that one will treat some people in ways, for better or worse, that they do not deserve to be treated."); Brian Dalessandro et al., *Bigger Is Better, but at What Cost?: Estimating the Economic Value of Incremental Data Assets*, 2 BIG DATA 87 (2014); Cynthia Dwork et al., *Fairness Through Awareness*, 3 PROC. INNOVATIONS THEORETICAL COMPUTER SCI. CONF. 214 app. at 226 (2012); Tal Zarsky, *Transparency in Data Mining: From Theory to Practice*, DISCRIMINATION AND PRIVACY IN THE INFO. SOC'Y 301, 310 (2013); Mireille Hildebrandt, *The Dawn of a Critical Transparency Right for the Profiling Era*, DIGITAL ENLIGHTENMENT YEARBOOK 2012 (Jacques Bus et al. eds., 2012).

power."[154] Though transparency may often feel like a robust solution intuitively, explainable artificial intelligence—or "XAI"[155] as it is increasingly called—is especially unlikely to provide significant remedial utility to individuals in instances where the discrimination involved is only observable at the statistical scale. Moreover, some commentators have convincingly argued that too great a focus on individualized explanations—as opposed to broader, multi-methodological design practices for mitigating unfairness—could "nurture a new kind of transparency fallacy," wherein providing a basic explanation to individual users could provide false cover for companies whose processing operations may be biased for other reasons.[156]

Second, allowing enterprises a broader range of compliance options could allow them greater flexibility to deploy powerful algorithms that may make more conventional forms of explicability impractical or impossible. Under the current state-of-the-art, it is widely acknowledged that many of the highest performing machine learning algorithms pose significant "tradeoff[s] between the representational capacity of a model and its interpretability."[157] Techniques capable of achieving the richest predictive

---

[154] *See* Edwards & Veale, *supra* note __ (quoting Mike Annany & Kate Crawford, *Seeing Without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability*, New Media & Society 1, 5) (quotations removed). *See also, e.g.,* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (Harvard University Press 2015) (arguing that sheer transparency has not created any real effects on or accountability of in many contexts); Joshua Kroll et al., *Accountable Algorithms*, 165 U.PA.L. REV. 633, 638 (2017) (rejecting transparency as a meaningful remedy for promotoing accountability); Brendan Van Alsenoy et al., *Privacy Notices Versus Informational Self-Determination: Minding The Gap*, 28 INT'L REV. OF LAW, COMPUTERS & TECHNOLOGY 2 185–203 (2014).

[155] *See* Tim Miller, *Explanation in Artificial Intelligence: Insights from the Social Sciences* (June 2017), arXiv.1706.07269v1. The initialism stands for "eXplainable artificial intelligence."

[156] *See* Edwards & Veale, *supra* note __. *See also, e.g.,* Toon Calders & Indrė Žliobaitė, *Why Unbiased Computational Processes Can Lead to Discriminative Decision Procedures,* in DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY 43, 46 ("[T]he selection of attributes by which people are described in [a] database may be incomplete."); Annamarie Carusi, *Data as Representation: Beyond Anonymity in E-Research Ethics*, 1 INT'L J. INTERNET RES. ETHICS 37, 48–61 (2008).

[157] *See* Goodman & Flaxman, *supra* note __. *See also, e.g.,* MIREILLE HILDEBRANDT, THE NEW IMBROGLIO - LIVING WITH MACHINE ALGORITHMS,

results tend to do so through the use of aggregation, averaging, or multilayered techniques which, in turn, make it difficult to determine the exact features that play the largest predictive role.[158] Depending on the circumstances, performance losses associated with adopting a more explicable approach could prove far costlier than the social utility of providing individualized explanations.[159] Particularly in instances where the leading techniques far outpace the remedial options available to data subjects, a one-size-fits-all approach to oversight could lead to unnecessary bureaucratic roadblocks for technologies with massively beneficial social impacts.[160]

(The Art of Ethics in the Information Society 2016), https://works.bepress.com/mireille_hildebrandt/75/; IEEE GLOBAL INITIATIVE, ETHICALLY ALIGNED DESIGNED - A VISION FOR PRIORITIZING HUMAN WELLBEING WITH ARTIFICIAL INTELLIGENCE AND AUTONOMOUS SYSTEMS (IEEE 2016 Version 1), http://standards.ieee.org/develop/indconn/ec/ead_v1.pdf; BEN WAGNER, EFFICIENCY VS. ACCOUNTABILITY? – ALGORITHMS, BIG DATA AND PUBLIC ADMINISTRATION.

[158] *See* Wojciech Samek et al., *Evaluating the Visualization of What a Deep Neural Network Has Learned*, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS 1, 7 (2016); Marco Tulio Ribeiro et al. *"Why Should I Trust You?": Explaining the Predictions of Any Classifier*, PROCEEDINGS OF THE 22ND ACM SIGKDD INTERNATIONAL CONFERENCE ON KNOWLEDGE DISCOVERY AND DATA MINING 1135–44 (2016); Jon Kleinberg et al., *Human Decisions and Machine Predictions*, NAT''L BUREAU OF ECON. RES. WORKING PAPER SERIES (2017), http://www.nber.org/papers/w23180.

[159] *See id.* This, however, may eventually prove to be a moving target.

[160] *See, e.g.,* Toon Calders & Sicco Verwer, *Three Naive Bayes Approaches for Discrimination-Free Classification*, 21 DATA MINING AND KNOWLEDGE DISCOVERY 277 (Jul. 2010) (describing trade-off between discrimination removal and classifier performance); Faisal Kamiran & Toon Calders, *Data Preprocessing Techniques for Classification Without Discrimination*, 33 KNOWLEDGE AND INFO. SYS. 1 (Dec. 2012) (describing trade-off between discrimination removal and classifier performance); Jagriti Singh & S.S. Sane, *Preprocessing Technique for Discrimination Prevention in Data Mining*, 3 THE INT'L J. OF ENGINEERING AND SCI. (IJES) 12 (2014) (noting inherent trade-offs in the current state-of-the-art); Corbett-Davies et al., *Algorithmic Decision Making and the Cost of Fairness*, arXiv:1701.08230v4 (Jun. 2017), https://arxiv.org/pdf/1701.08230.pdf. These tradeoffs will likely be a moving target. Indeed, Edwards & Veale, note that the inevitability of these tradeoffs may only be an "an interim conclusion" and are "convinced that recent research in ML explanations shows promise" for reducing or eliminating some of these tradeoffs. *See* Edwards & Veale, *supra* note __.

Finally, and perhaps most importantly, system-wide audits of the type envisioned by DPIAs already have a well-documented track record of detecting and combating algorithmic discrimination in otherwise opaque systems. As Sandvig et al. note, audit studies are "the most prevalent social scientific methods for the detection of discrimination" in complex computational systems.[161] In recent years, these auditing techniques have been used by researchers and journalists to successfully detect and document algorithmic bias across diverse industry sectors and social domains.[162] Further, this approach includes the added benefit of allowing outside entities that may have more resources than individuals to scrutinize the integrity of complex computational systems. Regulators, NGOs, media outlets, and public interest organizations that specialize in this area will be able to invest in the expertise necessary to not only provide data subjects with the right answers, but also to ensure that the right questions are asked.

---

[161] *See* Christian Sandvig et al., *Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms*, DATA AND DISCRIMINATION: CONVERTING CRITICAL CONCERNS INTO PRODUCTIVE INQUIRY (2014), http://social.cs.uiuc.edu/papers/pdfs/ ICA2014-Sandvig.pdf; Andrea Romei & Salvatore Ruggieri, *Discrimination Data Analysis: A Multi-Disciplinary Bibliography*, in DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY 109, 120; Faisal Kamiran, Toon Calders & Mykola Pechenizkiy, *Techniques for Discrimination Free Predictive Models*, in DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY 223–24.

[162] *See e.g.*, James Grimmelmann and Daniel Westreich, *Incomprehensible Discrimination*, 7 CAL. L. REV. ONLINE 164, 173 (2017); FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION (Harvard University Press 2015); MIREILLE HILDEBRANDT, THE NEW IMBROGLIO - LIVING WITH MACHINE ALGORITHMS, (The Art of Ethics in the Information Society 2016), https://works.bepress.com/mireille_hildebrandt/75/; Kiel Brennan-Marquez, *"Plausible Cause": Explanatory Standards in the Age of Powerful Machines*, (2017) 70 VANDERBILT L. REV. 1249; Andrew Selbst, *A Mild Defense of Our New Machine Overlords* (2017) 70 VANDERBILT L. REV. EN BANC 87; Reuben Binns, *Algorithmic Accountability and Public Reason*, PHILOSOPHY & TECHNOLOGY, (forthcoming); Katherine Strandburg, *Decision-Making, Machine Learning and the Value of Explanation*, THE HUMAN USE OF MACHINE LEARNING: AN INTERDISCIPLINARY WORKSHOP, (Dec. 16 2016), http://www.dsi.unive.it/HUML2016/assets/Slides/Talk%202.pdf.

Although data audit and DPbD methodologies come with their own unique set of challenges,[163] the multifaceted advantages[164] offered by these approaches present exciting new possibilities for fostering genuine algorithmic accountability in enterprises without stifling technological and business advances.[165] In contrast to a remedial "right to explanation" invoked on an individual basis by downstream data subjects, properly implemented auditing and DPbD can provide the evidence necessary to inform and vet the design and deployment of more fair, accountable, and transparent algorithmic systems.[166]

IV. EXPORTING THE "RIGHT TO EXPLANATION:" THE BRUSSELS EFFECT AND THE GDPR'S LONG TENTACLES

Although the EU is sometimes maligned as a declining force on the world stage, numerous recent studies have demonstrated that it actually exercises "unprecedented global power . . . through its legal institutions and standards that it successfully exports to the rest of the world."[167] This

---

[163] *See* Bryce Goodman, *A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union General Data Protection*, NIPS 29th Conference on Neural Information Processing Systems 1, 7 (2017) http://www.mlandthelaw.org/papers/goodman1.pdf ("[A] process that passes a safety audit may fail for other reasons (e.g. inefficiency). Passing a safety audit does not mean that all risk is eliminated but, rather, that risk is reduced to an acceptable level. Choosing an acceptable level of risk depends in turn on the process evaluated and, in particular, both the likelihood and severity of a failure."). *See also* Lior Jacob Strahilevitz, *Privacy Versus Antidiscrimination*, 75 U. CHI. L. REV. 363, 364 (2008).

[164] The list enumerated above is, of necessity, far from exhaustive.

[165] *See* Bryce Goodman, *A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union General Data Protection*, NIPS 29th Conference on Neural Information Processing Systems 1, 7 (2017) http://www.mlandthelaw.org/papers/goodman1.pdf.

[166] *See id. See also* Anupam Datta et al., *Algorithmic Transparency via Quantitative Input Influence*, in TRANSPARENT DATA MINING FOR BIG AND SMALL DATA (Tania Cerquitelli et al. eds. Springer 2017).

[167] *See* Anu Bradford, *The Brussels Effect*, 107 Northwestern University Law Review 1 (2012). *See also, e.g.,* Christopher Kuner, *The Internet and the Global Reach of EU Law*, LSE LAW, SOCIETY AND ECONOMY WORKING PAPERS 4/2017 OF THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE LAW DEPARTMENT (2017); David Scheer, *For Your Eyes Only—Europe's New High-Tech Role: Playing Privacy Cop to the World,* WALL ST. J. (Oct. 10, 2003); Brandon Mitchener, *Standard Bearers: Increasingly, Rules of Global Economy*

"export" effect occurs through the process of "unilateral regulatory globalization," wherein "a single state is able to externalize its laws and regulations outside its borders through market mechanisms, resulting in the globalization of standards."[168] Particularly in the last decades, the EU has evinced "a strong and growing ability to promulgate regulations that become entrenched in the legal frameworks of developed and developing markets alike" without relying on international institutions or intergovernmental negotiations. This phenomenon has since come to be described as the "Brussels Effect."[169]

There is, perhaps, no better exemplar of the "Brussels Effect" in action than the DPD itself, which has become a *de facto* standard for data privacy protection across the globe.[170] Since its passage in 1995, more than thirty countries have heeded Brussels' call by "adopt[ing] EU-type privacy

---

Are Set in Brussels, WALL ST. J. (Apr. 23, 2002); Editorial, *Regulatory Imperialism*, WALL ST. J. (Oct. 26, 2007), http://online.wsj.com/article/SB1193347205 39572002.html; see Case COMP/M.5984, Intel/McAfee, EUR-Lex 32011M5984 (Jan. 26, 2011).

[168] *See* Bradford *supra* note __. *See also, e.g.,* Daniel W. Drezner, *Globalization, Harmonization, and Competition: The Different Pathways to Policy Convergence*, 12 J. EUR. PUB. POL'Y 841, 841–59 (2005) ("[A] . . . reasonable conjecture would be to say that the public good benefits from regulatory coordination depend upon the size of the newly opened market."); Beth Simmons, *The International Politics of Harmonization: The Case of Capital Market Regulation*, in DYNAMICS OF REGULATORY CHANGE: HOW GLOBALIZATION AFFECTS NATIONAL REGULATORY POLICIES, at 42, 50–52; David A. Wirth, *The EU's New Impact on U.S. Environmental Regulation*, 31 FLETCHER F. WORLD AFF. 91, 96 (2007) ("If [a] jurisdiction's market share is sufficiently large, [its] regulatory requirements can affect an even larger area, including those under the control of other sovereign authorities.").

"This process can be distinguished from political globalization of regulatory standards where regulatory convergence results from negotiated standards, including international treaties or agreements among states or regulatory authorities." *See* Bradford *supra* note __. "It is also different from unilateral coercion, where one jurisdiction imposes its rules on others through threats or sanctions." *See id.* "Unilateral regulatory globalization is a development where a law of one jurisdiction migrates into another in the absence of the former actively imposing it or the latter willingly adopting it." *See id.*

[169] *See* Bradford *supra* note __.

[170] *See id.*

laws, including most countries participating in the Organization for Economic Cooperation and Development."[171]

According to those who have studied the "Brussels Effect" closely, its underlying mechanics are relatively intuitive. Countries confronted with the stringent standards established by many EU regulations face a stark choice. They can either revise their own domestic policies to reflect those within Europe, or risk breaking economic ties with the world's largest trading bloc.[172] For most, the decision requires little more than a moment's contemplation. Aside from a few notable outliers—such as the United States,[173] Russia, and China—most countries simply make the rational calculation that the costs of exclusion from a market consisting of 500 million of the globe's most affluent inhabitants far outweigh the costs of complying with Europe's higher standards.[174]

---

[171] *See, e.g.,* Debora L. Spar & David B. Yoffie, *A Race to the Bottom or Governance from the Top?*, in COPING WITH GLOBALIZATION 31, 31–51 (Aseem Prakash & Jeffrey A. Hart eds., 2000); David Vogel, *Trading Up and Governing Across: Transnational Governance and Environmental Protection*, 4 J. EUR. PUB. POL'Y 556, 563 (1997); Vogel & Kagan, supra note 15, at 2–8; ELIZABETH R. DESOMBRE, FLAGGING STANDARDS: GLOBALIZATION AND ENVIRONMENTAL, SAFETY, AND LABOR REGULATIONS AT SEA (2006); Daniel W. Drezner, *Globalization, Harmonization, and Competition: The Different Pathways to Policy Convergence*, 12 J. EUR. PUB. POL'Y 841, 841–59 (2005); Beth Simmons, *The International Politics of Harmonization: The Case of Capital Market Regulation*, in DYNAMICS OF REGULATORY CHANGE: HOW GLOBALIZATION AFFECTS NATIONAL REGULATORY POLICIES, at 42, 50–52; Katharina Holzinger & Thomas Sommerer, *'Race to the Bottom' or 'Race to Brussels'?: Environmental Competition in Europe*, 49 J. COMMON MARKET STUD. 315, 329 (2011); David Bach & Abraham Newman, *The European Regulatory State and Global Public Policy: Micro-Institutions, Macro-Influence*, 14 J. OF EUR. PUB. POL'Y 827 (2007); Mark F. Kightlinger, *Twilight of the Idols? EU Internet Privacy and the Post Enlightenment Paradigm*, 14 COLUM. J. EUR. L. 1, 2–3 (2007).

[172] *See* David Bach & Abraham L. Newman, *The European Regulatory State and Global Public Policy: Micro-Institutions, Macro-Influence*, 14 J. EUR. PUB. POL'Y 827, 831 (2007); Bradford *supra* note __, at 11–14. There are, of course, other factors that contribute to this effect. *See* Bradford *supra* note __, at 11–28.

[173] *See* Bradford *supra* note __.

[174] The EU's population exceeds 500 million, and its GDP per capita exceeds $35,000. *See* EUROPEAN COMMISSION, EUROSTAT: GDP PER CAPITA IN PPS, http://ec.europa.eu/eurostat/web/products-datasets/-/tec00114.

And lest those powerful incentives prove to be insufficient, the GDPR itself also includes a number of notable changes intended to promote extraterritorial compliance that look likely to extend its regulatory reach above and beyond the baseline already established by the "Brussels Effect." The most significant changes, in this realm, are those involving the Regulation's "adequacy decision" used to determine whether "third countries"—*i.e.* countries outside of the EU—have sufficient protections in place to warrant the transfer of personal data between themselves and EU Member States.[175] Once a country is deemed "adequate" through an assessment by the European Commission, data can flow freely without the need for additional protective measures. But unlike under the DPD, adequacy decisions made under the GDPR will be subject to a periodic review at least every once every four years, and will also be subject to repeal, amendment, or suspension on an ongoing basis.[176]

Thanks, in no small part, to the introduction of these far-reaching forms of regulatory oversight, the GDPR is already showing signs of its global standard-setting authority. Countries such as Israel, New Zealand, Argentina, and Japan have all recently undergone efforts to receive EU "adequacy" certifications by ensuring that their domestic data protections rise to the level of Europe's.[177] And "[o]ther countries, from Colombia to South Korea to the tiny island nation of Bermuda, are similarly rebooting [their] domestic legislation. . . [which at times] involves adopting European rules almost word for word."[178]

Though the "Europeanization" of global regulatory standards is often most pronounced at the national level, a phenomenon not unlike the one occurring on the global scale due to the "Brussels Effect" is also taking place within individual enterprises. According to a recent headline-grabbing announcement by Facebook, "[d]ozens of people at [the company] are working full time on" GDPR compliance—requiring upwards of a 250%

---

[175] *See* GDPR, *supra* note __, at art. 45.

[176] *See id.*

[177] *See* Mark Scott and Laurens Cerulus, *Europe's New Data Protection Rules Export Privacy Standards Worldwide*, Politico (Jan. 31, 2018), https://www.politico.eu/article/europe-data-protection-privacy-standards-gdpr-general-protection-data-regulation/.

[178] *See id.*

increase in staffing related to EU data protection.[179] A company spokesperson noted that, "It is hard for us to put an exact figure on it, but when you take into account the time spent by our existing teams, the research and legal assessments and the fact that we have had to pull in teams from product and engineering, it is likely to be millions of dollars."[180] Recent reporting by *The Financial Times* provided even further confirmation of this phenomenon. The media outlet—which contracted twenty "of the largest social media, software, financial technology and internet companies with EU operations"—noted that its inquiries "revealed that the sector is scrambling to hire new staff and redesign products as it faces millions of dollars in higher costs and lost revenues."[181] And while not every company has quite the multinational reach of the average tech giant, this extraterritorial affect is made all the more pronounced by the GDPR's applicability to *any* company processing the data of EU citizens, not just those companies actually located within the EU itself.[182]

For some companies operating outside of the GDPR's immediate purview, it may be feasible to fragment their internal processing pipelines by treating data originating in Europe differently from that of other geographies. But doing so could prove administratively onerous—requiring multiple, separate handling processes for data flowing through any given enterprise. Moreover, this type of maneuver may also be perceived as a public relations risk for companies concerned about being "outed as deliberately offering a lower privacy standard to [their] home users [versus] customers abroad."[183] Thus, just as is true at the national level as a result of

---

[179] *See* Aliya Ram, *Tech Sector Struggles to Prepare for New EU Data Protection Laws*, THE FIN. TIMES (Aug. 29 2017), https://www.ft.com/content/5365c1fa-8369-11e7-94e2-c5b903247afd.

[180] *See id.*

[181] *See id.* This phenomenon has led some experts to speculation speculate that the "GDPR could be one of the most expensive pieces of regulation in the [technology] sector's history." *See id.*

[182] *See* GDPR *supra* note __. *See also, e.g.,* Goodman & Flaxman, *supra* note __ (commenting that the GDPR's "requirements do not just apply to companies that are headquartered in the EU but, rather, to any companies processing EU residents' personal data . . . [thus] [f]or the purposes of determining jurisdiction, it is irrelevant whether that data is processed within the EU territory, or abroad"); Natasha Lomas, *WTF is GDPR*, TECHCRUNCH, January 2018, https://techcrunch.com/2018/01/20/wtf-is-gdpr/ (noting "that GDPR does not merely apply to EU businesses; any entities processing the personal data of EU citizens need to comply).

[183] *See* Scott & Cerulus, *supra* note __.

the "Brussels Effect," the path of least resistance for many companies will likely entail treating the GDPR as the new "gold standard"—dictating that they handle all personal data, regardless of geography, in accordance with the baseline articulated by the Regulation's enforcement agencies.[184] And while the precise contours of this new gold standard may continue to be clarified for some time to come, it is now clear that it includes a muscular "right to explanation" with sweeping implications for companies and countries thoroughout the world. As one commentator working to promote GDPR compliance as far away as South Africa recently noted, any entity not currently addressing it will soon realize that the "GDPR has long tentacles."[185]

CONCLUSION

Now that the data protection authorities responsible for enforcing the GDPR's "right to explanation" have weighed in, at least one matter of fierce public debate can be laid to rest. The GDPR provides an unambiguous "right to explanation" with sweeping legal implications for the design, prototyping, field testing, and deployment of automated data processing systems. Failing to countenance this right could subject enterprises to economic sanctions of truly historic magnitudes—a threat that simply did not exist under the GDPR's predecessor.

Although the protections enshrined within the right may not mandate transparency in the form of a complete individualized explanation, a holistic examination of the Regulation reveals that the right's true power derives from its synergistic combination with DPbD practices codified by the Regulation's subsequent chapters. While these new design standards will, doubtless, pose significant challenges for the enterprises that fall within the GDPR's purview, the speed and scale of the global response thus far is cause for genuine optimism. Indeed, there is perhaps no more hopeful bookend to this profoundly important debate than the recent words of Bryce Goodman, one of the authors responsible for first sparking the controversy: "In the past, companies have devoted immense resources to improving

---

[184] *See* EUROPEAN DATA PROTECTION SUPERVISOR, THE HISTORY OF THE GENERAL DATA PROTECTION REGULATION, https://edps.europa.eu/data-protection/data-protection/legislation/history-general-data-protection-regulation_en.

[185] *See* Scott & Cerulus, *supra* note __.

algorithmic performance. . . . Going forward, one hopes to see similar investments in promoting fair and accountable algorithms."[186]

[186] *See* Bryce Goodman, *A Step Towards Accountable Algorithms?: Algorithmic Discrimination and the European Union General Data Protection*, NIPS 29th Conference on Neural Information Processing Systems 1, 7 (2017) http://www.mlandthelaw.org/papers/goodman1.pdf.