Algorithmic Impact Assessments under the GDPR:

Producing Multi-layered Explanations

Margot E. Kaminski & Gianclaudio Malgieri*

ABSTRACT

Policy-makers, scholars, and commentators are increasingly concerned with the risks of using profiling algorithms and automated decision-making. The EU's General Data Protection Regulation (GDPR) has tried to address these concerns through an array of regulatory tools. As one of us has argued, the GDPR combines individual rights with systemic governance, towards algorithmic accountability. The individual tools are largely geared towards individual "legibility": making the decision-making system understandable to an individual invoking her rights. The systemic governance tools, instead, focus on bringing expertise and oversight into the system as a whole, and rely on the tactics of "collaborative governance," that is, use public-private partnerships towards these goals. How these two approaches to transparency and accountability interact remains a largely unexplored question, with much of the legal literature focusing instead on whether there is an individual right to explanation.

The GDPR contains an array of systemic accountability tools. Of these tools, impact assessments (Art. 35) have recently received particular attention on both sides of the Atlantic, as a means of implementing algorithmic accountability at early stages of design, development, and training. The aim of this paper is to address how a Data Protection Impact Assessment (DPIA) links the two faces of the GDPR's approach to algorithmic accountability: individual rights and systemic collaborative governance. We address the relationship between DPIAs and individual transparency rights. We propose, too, that impact assessments link the GDPR's two methods of governing algorithmic decision-making by both providing systemic governance and serving as an important "suitable safeguard" (Art. 22) of individual rights.

After noting the potential shortcomings of DPIAs, this paper closes with a call—and some suggestions—for a Model Algorithmic Impact Assessment in the context of the GDPR. Our examination of DPIAs suggests that the current focus on the right to explanation is too narrow. We call, instead, for data controllers to consciously use the required DPIA process to produce what we call "multi-layered explanations" of algorithmic systems. This concept of multi-layered explanations not only more accurately describes what the GDPR is attempting to do, but also normatively better fills potential gaps between the GDPR's two approaches to algorithmic accountability.

* Associate Professor of Law, Colorado Law School; Doctoral Researcher at LSTS, Vrije Universiteit Brussels. Both authors contributed equally to this paper.

1. Introduction

To date, the discussion of the GDPR's regulation of algorithmic accountability has largely focused on whether there is an individual right to explanation of an algorithmic decision. Only more recently have legal scholars begun to focus on the GDPR's systemic accountability tools.

Impact assessments have received particular attention, on both sides of the Atlantic, as a tool for algorithmic accountability. The aim of this paper is to address how Data Protection Impact Assessments (DPIAs) (Art. 35) link the GDPR's two approaches to algorithmic accountability—individual rights and systemic governance—and potentially lead to more accountable and explainable algorithms. Examining the GDPR's approach to impact assessments suggests that the scholarship has been getting explanations wrong. Algorithmic explanations should not be understood as static statements, but as a circular and multi-layered process. The literature has largely to date focused on what information goes to whom, when; we argue that the impact assessment process plays a crucial role in connecting internal company heuristics and risk mitigation to outward-facing rights, and in forming the substance of several different kinds of explanations.

We begin by introducing the algorithmic accountability tools in the GDPR (§2). In Section 3, we explore the individual rights of data subjects as regards algorithmic decisions. In Section 4, we explain the GDPR's collaborative governance of algorithms. Section 5 discusses the broader literature on Algorithmic Impact Assessments, and Section 6 turns to how the DPIA in the GDPR might function as an Algorithmic Impact Assessment. The final section (§7) explains how our approach to the GDPR's Impact Assessments in fact leads to a better, more complex understanding of the GDPR's explanations of algorithmic decision-making than a focus on Article 22 alone. We close by calling for what we call a multi-layered approach to explanations, stemming from the Impact Assessment process.

2. Algorithmic Accountability in the GDPR

The GDPR has significant implications for algorithmic decision-making. At first, the legal debate focused on whether the GDPR created an individual right to an explanation of an individual algorithmic decision.[1]

Subsequent legal analysis, however, began to focus instead on other accountability tools,[2] either required in the text of the GDPR or recommended in interpretative guidelines from the Article 29 Working Party.[3] These tools include third-party auditing, the appointment of Data Protection Officers (DPOs)(Art.37), and the requirement of Data Protection Impact Assessments (DPIAs)(Art.35).

As one of us has argued, the GDPR combines a series of individual rights (Arts.12-23) with a systemic governance regime overseen by regulators, targeted at more comprehensive oversight over the algorithm and the people around it (Arts.24-43 & throughout). These two systems interact and overlap. An individual right is often also a company's duty. But even if individuals (data subjects) fail to invoke their rights, companies (data controllers) have significant obligations—both procedural and substantive—under the GDPR. [4]

For example, in the algorithmic governance context, a data subject has a right to contest an individual algorithmic decision (Art.22), to receive notice of solely automated decision-making (Art.13), and to request access to "meaningful information about the logic involved" (Art.15). Should she fail to invoke any of these rights, however, the GDPR still puts in place significant obligations on data controllers using automated decision-making, whether that decision-making involves a human or not.[5] The GDPR requires an array of systemic accountability tools, including: third-party auditing, the appointment of Data Protection Officers (DPOs)(Art. 37), and Data Protection Impact Assessments (DPIAs)(Art. 35). These obligations arise both from the text of the law, and in accompanying Recitals, and in the Guidelines on Automated Individual Decision-making and Profiling ("Guidelines on ADM") released in October 2017 and revised in February 2018 by the Article 29 Working Party (now the European Data Protection Board).[6]

---

[1] Sandra Wachter, Brent Mittelstadt, and Luciano Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation', *International Data Privacy Law* 7, no. 2 (1 May 2017): 76–99, https://doi.org/10.1093/idpl/ipx005; Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation', *International Data Privacy Law* 7, no. 4 (1 November 2017): 243–65, https://doi.org/10.1093/idpl/ipx019. B. Goodman and S. Flaxman, 2016 EU Regulations on Algorithmic Decision-Making and a "right to Explanation," http://arxiv.org/abs/1606.08813, accessed 30 June 2018. A. Selbst and J. Powles; Meaningful information and the right to explanation, *International Data Privacy Law*, Volume 7, Issue 4, 1 November 2017, Pages 233–242, https://doi.org/10.1093/idpl/ipx022 . Maja Brkan, 'Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and Beyond', *International Journal of Law and Information Technology*, accessed 24 April 2019, https://doi.org/10.1093/ijlit/eay017; Margot E. Kaminski, 'The Right to Explanation, Explained', *Berkeley Technology Law Journal*, 34, no. 1 (2019), https://papers.ssrn.com/abstract=3196985.

[2] Antoni Roig, 'Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)', *European Journal of Law and Technology* 8, no. 3 (21 January 2018), http://ejlt.org/article/view/570; Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For', *Duke Law & Technology Review* 16, no. 1 (4 December 2017): 18–84; Bryan Casey, Ashkon Farhangi, and Roland Vogl, 'Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise', *Berkeley Technology Law Journal* 34, no. 2019 (19 February 2018), https://papers.ssrn.com/abstract=3143325.

[3] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, Adopted on 3 October 2017, As last Revised and Adopted on 6 February 2018, WP251rev.01, 29.

[4] Margot E. Kaminski, 'Binary Governance: Lessons from the GDPR's Approach to Algorithmic Accountability', *Southern California Law Review* 92, no. 6 (2019), https://papers.ssrn.com/abstract=3351404; Kaminski, 'The Right to Explanation, Explained'.

[5] Edwards and Veale, 'Slave to the Algorithm?', 74–80.

[6] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, Adopted on 3 October 2017, As last Revised and Adopted on 6 February 2018, WP251rev.01, 29. See Michael Veale and Lilian Edwards, 'Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling', *Computer Law & Security Review* 34, no. 2 (1 April 2018): 398–404, https://doi.org/10.1016/j.clsr.2017.12.002.

It is also crucial to understand the mode through which the GDPR governs. The GDPR largely governs—both in the sense of coming up with the substance of data controllers' duties, and in the sense of monitoring compliance—through an approach known in the legal literature as "collaborative governance": the use of public-private partnerships.[7] This form of regulatory design has alternatively been referred to as "new governance," "co-governance," partial delegation to the private sector, and "meta-regulation." Importantly, it is not equivalent to self-regulation; the government still has an important, even central, role to serve.

Because the GDPR often effectively outsources governance decisions to private companies, accountability takes on added significance. Accountability in the GDPR is not just about protecting individual rights. It is about ensuring that this process of co-governing with private parties receives appropriate oversight from the public, from civil society, and from both expert and affected third parties.[8]

With this background in mind, the next two sections of this paper go into more detail on both the individual rights and systemic governance elements of the GDPR's approach to algorithmic accountability, before turning to the role of the Data Protection Impact Assessment (DPIA) in linking the two facets.

3. Individual Rights in the GDPR and the Multi-layered Explanation

The GDPR gives individuals several important rights with respect to algorithmic decision-making. The GDPR contains both general data protection rights and rights specific to profiling that also apply to algorithmic decision-making.[9] On top of this, it establishes rights specific to algorithmic decision-making, which include: a right to be notified of solely automated decision-making (Arts. 13, 14); a right of both notification and access to meaningful information about the logic involved (Arts. 13, 14, 15); a right to be informed of the significance of and envisaged effects of solely automated decision-making (Arts. 13, 14, 15); and a right not to be subject to solely automated decision making (Art. 22), with safeguards and restraints for the limited cases in which automated decision-making is permitted. Those safeguards include, but are not limited to, a right to contest a decision, to express one's point of view, and to human intervention (Art. 22).

We do not intend to revisit the legal debate over these rights in detail here, but an overview may be useful. As mentioned, discussion of these individual rights has largely focused on whether or not—or really, how—solely automated decision-making must be explained to individuals. As Selbst & Powles point out, it is disingenuous to say that there is no right to an explanation in the GDPR; the GDPR's text clearly requires companies to explain at least "meaningful information about the logic involved" in automated decision-making, in addition to its significance and envisioned effects (Art. 13, 14, 15).[10] What this information constitutes in practice, however, has been the subject of hot debate, including whether it is a system-wide (model-wide) explanation or specific to individual decisions, and what depth of explanation is required.[11]

---

[7] See Jody Freeman, 'The Private Role in the Public Governance', *New York University Law Review* 75 (2000): 543; K. Bamberger, 'Regulation as Delegation: Private Firms, Decisionmaking, and Accountability in the Administrative State', *Duke Law Journal*, 1 January 2006, 377.

[8] Kaminski, 'Binary Governance', 28.

[9] Lilian Edwards and Michael Veale, 'Enslaving the Algorithm: From a "Right to an Explanation" to a "Right to Better Decisions"?', *IEEE Security & Privacy* 16, no. 3 (2018): 46–54.

[10] Andrew D. Selbst and Julia Powles, 'Meaningful Information and the Right to Explanation', *International Data Privacy Law* 7, no. 4 (1 November 2017): 233–42, https://doi.org/10.1093/idpl/ipx022.

[11] Wachter, Mittelstadt, and Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation', 78; Malgieri and Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation', 244; Selbst and Powles, 'Meaningful Information and the Right to Explanation', 240–41.

The core debate has primarily focused on whether or not Article 22 creates an *ex post* right to explanation of an *individual* decision made by an automated system.[12] Our view, discussed at length by each of us elsewhere, is that it does.[13] Automated decisions with significant effects must be made "legible" to individuals, in the sense that individuals must be able to understand enough about the decision-making process to be able to invoke their other rights under the GDPR, including the right to contest a decision.[14] And as one of us has noted, several of the Member States implementing Article 22(2)b of the GDPR have outlined the Article 22 explanation duties in greater detail.[15] As we discuss below in Section 7, our view is that the GDPR's transparency rights, too, are best discussed together as a system. The GDPR, that is, is best understood as establishing a system of *multi-layered explanations*. Individuals have a right to both a system-wide but detailed description of the logic of an algorithm (Arts. 13, 14, 15), and more specific insights on individual decisions taken.[16] We discuss possible further layers of explanations in Section 7 below.

It is important to remark that different data controllers may have different accountability duties. Article 24(1) of the GDPR states that taking into account the nature, scope, context and risks of data processing, the controller shall implement appropriate technical and organisational measures to ensure compliance with the GDPR. Accordingly, algorithmic decision making involving bigger risks for data subjects should entail more safeguards.[17] As we will explain in Section 4, data controllers to a certain extent choose their own algorithmic accountability safeguards: some are required in the GDPR both at Article 22(3) and at recital 71 (algorithmic auditing, the rights to contest, to have a new decision, to a human in the loop, and to explanation), but these are not closed lists, and the Guidelines on ADM suggest additional techniques. In the case of more intrusive and riskier automated decision-making processes, the data controller should and likely will implement all possible safeguards, including a right to explanation of individual decision taken.[18]

---

[12] Brkan, 'Do Algorithms Rule the World?'; Stefanie Hänold, 'Profiling and Automated Decision-Making: Legal Implications and Shortcomings', in Marcelo Corrales, Mark Fenwick, and Nikolaus Forgó (eds.), *Robotics, AI and the Future of Law*, Perspectives in Law, Business and Innovation (Singapore: Springer Singapore, 2018), 123–53, https://doi.org/10.1007/978-981-13-2874-9_6, See also Edwards & Veale, Slave to the Algorithms, passim.

[13] Malgieri & Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation'; Kaminski, 'The Right to Explanation, Explained', cit.

[14] Malgieri and Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation', 250.

[15] Gianclaudio Malgieri, 'Automated Decision-Making in the EU Member States: The Right to Explanation and Other "Suitable Safeguards" in the National Legislations', *Computer Law & Security Review*, 9 July 2019, 105327, https://doi.org/10.1016/j.clsr.2019.05.002 See in particular the cases of French and Hungarian laws, that provides more explicit explanation of individual decisions taken (based on criteria and methods used in algorithmic processing).

[16] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation, 25: "The controller should find simple ways to tell the data subject about the rationale behind, or the criteria relied on in reaching the decision. The GDPR requires the controller to provide meaningful information about the logic involved, not necessarily a complex explanation of the algorithms used or disclosure of the full algorithm. The information provided should, however, be sufficiently comprehensive for the data subject to understand the reasons for the decision". See also at page 27: "the controller should provide the data subject with general information (notably, on factors taken into account for the decision-making process, and on their respective 'weight' on an aggregate level) which is also useful for him or her to challenge the decision" and "Recital 71 highlights that *in any case* suitable safeguards should also include: specific information to the data subject and the right (…) to obtain an explanation of the decision reached after such assessment and to challenge the decision".

[17] Raphael Gellert, 'We Have Always Managed Risks in Data Protection Law: Understanding the Similarities and Differences between the Rights-Based and the Risk-Based Approaches to Data Protection', *European Data Protection Law Review (EDPL)* 2 (2016): 481.

[18] See European Commission's High Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Thrustworthy AI*, https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines, 15 ("The degree to which explicability is needed is highly dependent on the context and the severity of the consequences if that output is erroneous or otherwise inaccurate").

However, there have been legitimate concerns voiced in the legal literature both in Europe and in the United States about the capacity of individuals to both invoke their rights and execute oversight over algorithmic decision-making.[19] These range from concerns about access to justice to concerns about individual capacity and expertise. Consequently, most policy proposals call either for a dual regime, like the GDPR, that mixes individual rights with other more systemic forms of accountability;[20] or for foregoing individual accountability in favour of expert and external oversight.[21]

The latter approach—foregoing individual rights entirely—ignores the dignitary and legitimizing value of such rights.[22] Individual rights allow individuals to exhibit autonomy and exert control, and to protest or reject their objectification by profiling or decision-making machines.[23] Individualized explanations also serve to establish the legitimacy, or illegitimacy, of a decision-making system by subjecting its logics and performance to inspection and assessment as to whether they are socially acceptable or even illegal (what we and others call a "justification" of algorithmic decisions).[24] Rejecting individual rights, as we discuss below, also ignores the symbiosis between the GDPR's two regimes. Individual rights can play a crucial role in the GDPR's systemic collaborative governance. The GDPR's dual approach to algorithmic accountability has the potential to answer important questions in the literature about the value, in practice, of individual rights in algorithmic accountability.[25]

4. Collaborative Governance in the GDPR

The other side of algorithmic governance in the GDPR is its systemic governance regime. This regime aims, largely, to address instrumental goals: preventing error, bias, and discrimination.[26] As one of us discusses at length elsewhere, it is largely constituted through collaborative governance, or a cooperative public-private approach to regulation. We here illustrate two examples of how this works in the GDPR.

Article 22's suitable safeguards on automated decision-making are one example of this in practice. The GDPR's text does not comprehensively dictate what companies using automated decision-making must do to protect individual rights (Art.22). Instead, it lists examples of safeguards (contestation, expression, human intervention), but leaves it to both companies and regulators to determine what additional safeguards are necessary. The

[19] Mike Ananny and Kate Crawford, 'Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability', *New Media & Society* 20, no. 3 (1 March 2018): 973–89, https://doi.org/10.1177/1461444816676645; M. Hildebrandt, 'The Dawn of a Critical Transparency Right for the Profiling Era', in *Digital Enlightenment Yearbook*, ed. J. Bus, 2012th ed. (Amsterdam : IOS Press, 2012), 41–56, https://repository.ubn.ru.nl/handle/2066/94126; Edwards and Veale, 'Slave to the Algorithm?', 67; Bryce Goodman, 'A Step Towards Accountable Algorithms ? : Algorithmic Discrimination and the European Union General Data Protection', 2016, 3–4. Joshua Kroll et al., 'Accountable Algorithms', *University of Pennsylvania Law Review* 165, no. 3 (1 January 2017): 633; Deven R. Desai and Joshua A. Kroll, 'Trust But Verify: A Guide to Algorithms and the Law', *Harvard Journal of Law & Technology*, 27 April 2017, https://papers.ssrn.com/abstract=2959472; Edwards and Veale, 'Slave to the Algorithm?', 65–67.

[20] See generally Kaminski, 'Binary Governance'. See also Kate Crawford and Jason Schultz, 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms', *Boston College Law Review* 55, no. 1 (29 January 2014): 93. Danielle Citron, 'Technological Due Process', *Faculty Scholarship*, 30 April 2009, 1310, https://digitalcommons.law.umaryland.edu/fac_pubs/1012; Danielle Citron and Frank Pasquale, 'The Scored Society: Due Process for Automated Predictions', *Faculty Scholarship*, 1 January 2014, 20, 26, https://digitalcommons.law.umaryland.edu/fac_pubs/1431.

[21] Kroll et al., 'Accountable Algorithms', 660-663; Desai and Kroll, 'Trust but Verify', 39; Edwards and Veale, 'Slave to the Algorithm?', 76.

[22] Lee A Bygrave, 'Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling', *Computer Law & Security Review* 17, no. 1 (1 January 2001): 18, https://doi.org/10.1016/S0267-3649(01)00104-2.

[23] Kaminski, 'Binary Governance', 4; Hildebrandt, 'The Dawn of a Critical Transparency Right for the Profiling Era', 47.

[24] Kaminski, 'Binary Governance', 15.

[25] Id., passim.

[26] Id., 27.

accompanying Recital famously adds a right to individual explanation (Rec.71). The Guidelines, too, fill in this gap, proposing a list of best practices.[27] These include, but are not limited to: regular quality assurance checks, algorithmic auditing, independent auditing, establishing data minimisation and clear retention periods, using pseudonymisation techniques, certification mechanisms, ethical review boards, and more.[28]

All of these are attempts at systemic accountability and oversight, in a comprehensive and ongoing manner. But the Guidelines make clear that what counts as adequate safeguards will be established through an ongoing conversation between companies and regulators, involving government guidelines, and potentially involving industry-wide efforts to come up with codes of conduct or other forms of standards (Art.40).[29] The GDPR thus harnesses companies to help come up with both the what and the how of regulation in this space.

Another example of collaborative governance in action is the GDPR's approach to preventing bias and discrimination in algorithmic decision-making. Recital 71 tasks companies with preventing "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" in profiling and algorithmic decision-making. The GDPR does not lay out *how* to do this. Instead, the Guidelines suggest that companies check data sets for bias, regularly review the accuracy and relevance of decisions, deploy systems that audit algorithms, and use "appropriate procedures and measures to prevent errors, inaccuracies or discrimination" on the basis of sensitive data such as race, religion, or health information, deployed on a cyclical basis.[30]

Again, the GDPR does not tell companies precisely what to do. It identifies the problem, provides suggestions of what regulators might consider adequate, but also tasks companies with cooperatively coming up with solutions. Such company-created solutions may then feed back into what regulators ultimately require.[31]

5. Algorithmic Impact Assessments

Within this dual system of algorithmic accountability—individual rights accompanied by extensive but collaborative governance of companies' behaviour—the Data Protection Impact Assessment (DPIA) has a special role. We claim here that the DPIA is best understood as a nexus between the GDPR's two approaches to algorithmic accountability. Understanding it in this way allows us to better understand what is or might be required, and to observe the tool's shortcomings as implemented in the GDPR. Our analysis also provides lessons for discussions of Algorithmic Impact Assessments elsewhere. We turn here to both the general discussion that has arisen recently over Algorithmic Impact Assessments, and to the specific role of the DPIA in the GDPR, including as implemented by one Member State, Slovenia.

[27] Article 29 Working Party, Guidelines on Automated individual decision-making, 31-34.
[28] Article 29 Working Party, Guidelines on Automated individual decision-making, 32.
[29] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 32. On Certifications and algorithms see also Edwards and Veale, 'Enslaving the Algorithm', 50.
[30] Article 29 Working Party, Guidelines on Automated individual decision-making, 28.
[31] Roig, 'Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing', at 2 ("the requirement of data protection impact assessment (DPIA)… could compile all the relevant safeguards for specific technologies and automatic processing and turn into a data generator for policy purposes").

We are not the first to focus on Algorithmic Impact Assessments, or impact assessments in closely related fields.[32] We are, however, the first to discuss Algorithmic Impact Assessments not in isolation, but as a central component, among many components, of the GDPR's two-prong approach to algorithmic accountability. This changes the nature of the conversation. Instead of examining impact assessments in isolation from other accountability tools, it situates them within an overarching system of data governance. Our GDPR-specific analysis, then, may have implications for proposals for algorithmic impact assessments in other legal systems.[33] It suggests that impact assessments best serve a role in conversation with other accountability tools, as part of overarching regulatory design.[34] And it suggests that impact assessments play a central role both as a source of, and mediator between, the multi-layered explanations we believe are indicated in the GDPR.

5.1. Proposals for Algorithmic Impact Assessments

Algorithmic Impact Assessments have received a good deal of attention on both sides of the Atlantic as possible tools to address problems of algorithmic discrimination, bias, and unfairness—including in at least one proposed U.S. federal law.[35] We here briefly discuss several important precursors to the AIA: Environmental Impact Statements (EIS), Human Rights Impact Assessments (HRIA), Privacy Impact Assessments (PIA), Ethical Impact Assessments (EIA), and Surveillance Impact Assessments (SIA). We provide a short overview, too, of recent proposals for Algorithmic Impact Assessments: their origins, their substance, their overlaps, and their differences.

The inspiration for many U.S.-based impact assessment proposals is the Environmental Impact Statement (EIS), established in the United States in 1969 in the National Environmental Policy Act (NEPA).[36] NEPA's impact statement requirement applies when a federal agency proposes to take a "major Federal action significantly affecting the quality of the human environment."[37] As a threshold matter, agencies assess coverage and query whether any

[32] See e.g. Kenneth A. Bamberger & Deirdre K. Mulligan, 'PIA Requirements and Privacy Decision-Making in US Government Agencies', in D. Wright & P. De Hert (eds.), *Privacy Impact Assessment* (2012) at 225; Reuben Binns, 'Data protection impact assessments: a meta-regulatory approach', 7 Int'l. Data Priv. L. 22 (2017); Bryan Casey, Ashkon Farhangi, and Roland Vogl, 'Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise', *Berkeley Technology Law Journal* 34, no. 2019 (19 February 2018), https://papers.ssrn.com/abstract=3143325; Roger Clarke, 'Privacy impact assessment: Its origins and development', 25 *Comp. L. & Sec. Rev*. 123 (2009); A. Michael Froomkin, 'Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements', 2015 *U. Ill. L. Rev.* 1713 (2015); Chris Jay Hoofnagle, 'Assessing the Federal Trade Commission's Privacy Assessments', 14(2) *IEEE Sec. U& Priv.* 58 (2016); Sonia K. Katyal, 'Private Accountability in the Age of Artificial Intelligence', 66 *UCLA L. Rev.* 54, 112 (2019); Alessandro Mantelero, 'AI and Big Data: A blueprint for a human rights, social and ethical impact assessment', 34 *Comp. L. & Sec. Rev.* 754 (2018); David Wright & Charles D. Raab, 'Constructing a surveillance impact assessment', 28 *Comp. L. & Sec. Rev.* 613 (2012); Marc L. Roark, 'Human Impact Statements', 54 *WASHBURN L.J.* 649 (2015); Reisman et al., 'Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability'; Andrew D. Selbst, 'Disparate Impact in Big Data Policing', 52 *GA. L. REV*. 109 (2017), 169.David Wright & Michael Friedewald, 'Integrating Privacy and Ethical Impact Assessments', 40 *Sci. & Pub. Pol.* 755 (2013).

[33] See, e.g., more generally Dillon Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability' (AI Now Institute, n.d.), https://ainowinstitute.org/aiareport2018.pdf.

[34] Katyal, 'Private Accountability in the Age of Artificial Intelligence', cit., 117 suggests this by emphasizing the concurrent need for whistleblower protection. But not an overarching governance system.

[35] Wyden, Clarke, and Booker's Algorithmic Accountability Act. See https://www.wyden.senate.gov/news/press-releases/wyden-booker-clarke-introduce-bill-requiring-companies-to-target-bias-in-corporate-algorithms-.

[36] 42 U.S.C. § 4332(C) (2012). See Reisman et al at 7 ("The EIS process combines a focus on core values with a means for the public, outside experts, and policymakers to consider complex social and technical questions"); Selbst, 'Disparate Impact in Big Data Policing', 168 ("before adopting predictive policing technology, police should be required to create "algorithmic impact statements" (AISs), modeled on the environmental impact statements (EISs) of the National Environmental Policy Act (NEPA)." See also Froomkin, 'Regulating Mass Surveillance', 1749; See Roark, 'Human Impact Statement', 18.

[37] 42 U.S.C. § 4332(2)(C); 40 C.F.R. §1508.18 (defining "Major Federal Action"). This includes "projects and programs entirely or partly financed, assisted, conducted, regulated, or approved by federal agencies." Thus NEPA can apply to the behaviour of private actors, where

"Categorical Exclusions" apply.[38] If no exclusion applies, the agency performs an Environmental Assessment (EA), a public document that must "provide sufficient evidence and analysis for determining whether to prepare an environmental impact statement."[39] Then the agency either issues a Finding of No Significant Impact (FONSI), or goes on to prepare an Environmental Impact Statement. The full EIS must contain a detailed statement on environmental impact, including both any adverse effects which cannot be avoided, and alternatives,[40] before a project can go forward. The EIS is subject to comment by the public and other agencies,[41] and individuals can sue, and thus delay the project, if an EIS is incomplete or inadequate.[42]

A number of U.S. commentators have recently picked up the EIS as a model for impact assessments in other contexts. Froomkin, for example, touts the EIS as an effective alternative to command-and-control regulation, and a model for his proposed Privacy Impact Notice.[43] According to Froomkin, the EIS is a good regulatory model because it (1) pushes agencies to "consider…issues in the early design phase of their projects"[44] and (2) informs the public and solicits public feedback.[45] Selbst, who similarly bases his call for Algorithmic Impact Statements on the EIS model, agrees that it is an "action-forcing" regulation that "push[es] decision-makers to do their homework and engage with the public."[46] Selbst describes the EIS model not as an alternative to substantive regulation, but as a necessary precursor to it.[47] Froomkin, too, notes that public transparency can "ignite a regulatory dynamic by collecting information about the privacy costs of previously unregulated activities that should, in the end, lead to significant results".[48]Thus the EIS potentially has positive consequences both for the particular project at issue, and for forward movement in the larger policy debate.

For some, however, the EIS model fails to go far enough. The EIS process is static in nature, taking place only prior to the commencement of a project.[49] It is procedural, rather than substantive; it does not set substantive requirements, nor prohibit anybody from doing anything.[50] And while the EIS process requires public transparency and input, it does not require ongoing monitoring for compliance.[51]

Other proposals for impact assessments draw on additional sources as models, some of which in turn also trace their origins to the EIS.[52] Mantelero, for example, draws partially on the model of Human Rights Impact Assessments

---

they are funded or permitted by a federal agency. In fact, private actors applying for federal permits often participate in the EIS/EA process. See Froomkin, 'Regulating Mass Surveillance', 1751, n. 205.

[38] See Froomkin, 'Regulating Mass Surveillance', 1750.

[39] 40 CFR §1508.9(1), (3)(b).

[40] 42 U.S.C. § 4332(C) (2012); 40 CFR §1502.14.

[41] See Selbst, 'Disparate Impact in Big Data Policing', 178 (describing the two notice-and-comment periods, one to define the scope of the EIS and the second on the draft).

[42] Froomkin, 'Regulating Mass Surveillance', 1751.

[43] Id., 1755.

[44] Id., 1756

[45] Id., 1746

[46] Selbst, 'Disparate Impact in Big Data Policing', 169.

[47] Id. at 168 ("It is hard to say in the abstract what stronger regulatory solutions may be required, or how big a problem the technology poses in reality, until more information about the technology's implementation is created").

[48] Fromkin, 'Regulating Mass Surveillance', 1747.

[49] Selbst, 'Disparate Impact in Big Data Policing', 172.

[50] See Katyal, 'Private Accountability in the Age of Artificial Intelligence', 115.

[51] Selbst, 'Disparate Impact in Big Data Policing', 188; Clarke, 'Privacy Impact Assessments',125 (describing an EIS as "insufficiently auditable").

[52] See Clarke, 'Privacy Impact Assessment', 125. See also Mantelero, 'AI and Big Data', 757, that describes HRIA's as having their roots in the EIS.

(HRIA).[53] Katyal, too, references the HRIA process.[54] The HRIA process outlined by the United Nations[55] is a comparatively time- and resource-intensive process conducted on a business by third-party assessors, who collect data and interview stakeholders, experts, and management.[56] Wright and Friedewald look to Ethical Impact Assessments (EIA), voluntary assessments that go beyond legal compliance to assess the ethical implications of new technologies, and involve consultation with a wide number of stakeholders, and publication of the assessment.[57]

However, the most direct precursor for the notion of the AIA, especially in the context of the GDPR, is the Privacy Impact Assessment.[58] As Clarke explained as early as 2009, PIAs originated in the 1990s around the world, with multiple regulators issuing guidance in the early 2000s.[59] While Privacy Impact Assessments as conducted in the United States have been widely decried as toothless,[60] elsewhere they are considerably more substantial.[61] Clarke identifies EIS as a "progenitor" of the PIA, but goes on to identify a number of important differences.[62] It is interesting to note that in at least several European countries, the PIA may have originated in part in the system of "prior checking" under earlier data protection regimes, which was effectively a system of government registration or licensing of data processing systems prior to processing.[63] In order to receive a license from a national authority, a company had to assess whether it was in compliance with national data protection law. This differs vastly from the EIS, which has no substantive underpinnings and does not serve as the basis for a licensing regime.

Clarke characterizes the characteristics of the ideal PIA as: being performed on a project rather than an organization; being anticipatory in nature rather than retrospective; being broad in scope with respect to individual, group, community, and other "dimensions" of privacy; taking into account the perspectives of affected segments of the population; being broader than legal compliance; being oriented towards surfacing solutions, not just problems; emphasizing process over product; and requiring engagement from executives and managers.[64] In 2012, Wright and Raab proposed the concept of a Surveillance Impact Assessment (SIA),[65] wider in scope than a PIA but consisting of a similar "process of engaging stakeholders in order to identify the impacts on privacy and other values of a new

[53] Mantelero, 'AI and Big Data', 762. His HRSEIA is a (voluntary) hybrid. Lighter touch than HRIA, but takes into account ethical, social, human rights (grounded in human rights law).
[54] Katyal, 'Private Accountability in the Age of Artificial Intelligence', 112.
[55] United Nations, Human Rights Council, Office Of The High Comm'r, Guiding Principles on Business and Human Rights 23–26 (2011), https://www.ohchr.org/Documents/Publications/GuidingPrinciplesBusinessHR_EN.pdf
[56] Mantelero, 'AI and Big Data', 764.
[57] David Wright and Michael Friedewald, 'Integrating Privacy and Ethical Impact Assessments', *Science and Public Policy* 40, no. 6 (1 December 2013): 755–66, https://doi.org/10.1093/scipol/sct083; See previously on this point David Wright and Emilio Mordini, 'Privacy and Ethical Impact Assessment', in *Privacy Impact Assessment*, ed. David Wright and Paul De Hert, Law, Governance and Technology Series (Dordrecht: Springer Netherlands, 2012), 397–418, https://doi.org/10.1007/978-94-007-2543-0_19.
[58] See Binns, 'Data Protection Impact Assessments', 23; Clarke, 'Privacy Impact Assessment'; Wright & Friedewald, 'Integrating privacy and ethical impact assessments', 757-758. Wright & Raab, 'Constructing a Surveillance Impact Assessment', 755.
[59] Clarke, 'Privacy Impact Assessments'.
[60] See, e.g., Clarke, 'Privacy Impact Assessments', 128; Hoofnagle, 'Assessing the Federal Trade Commission's Privacy Assessments', 64; Bamberger & Mulligan, 'PIA Requirements and Privacy Decision-Making in US Government Agencies', 250.
[61] Clarke, 'Privacy Impact Assessments', passim.
[62] Clarke, 'Privacy Impact Assessments', 125.
[63] Binns, 'Data Protection Impact Assessments', 24; Clarke, 'Privacy Impact Assessments', 125. See also G. Le Grand, & E. Barrau, 'Prior Checking, A Forerunner to Privacy Impact Assessments', in D. Wright & P. De Hert (eds.), *Privacy Impact Assessment* (2012), 97–115.
[64] Clarke, 'Privacy Impact Assessments', at 124-125.
[65] See, previously, from the same authors: Charles Raab and David Wright, 'Surveillance: Extending the Limits of Privacy Impact Assessment'. In: Wright David, De Hert Paul, Editors. Privacy Impact Assessment. Dordrecht', in *Privacy Impact Assessment*, ed. David Wright and Paul De Hert (Dordrecht: Springer, 2012), 363–83.

project, technology, service or other initiative in order to take remedial action to minimise, avoid or overcome the risks."[66]
But this is the ideal. Mantelero observes that in practice, DPIAs in the European context have tended to focus on data quality and data security, leaving out broader social and legal impact despite aspirational language to the contrary.[67]

We now turn to recent proposals for Algorithmic Impact Assessments, which draw to varying degrees on these precursors. We find both common threads and significant differences in the proposals. We find, too, a significant gap in this literature that our perspective on the GDPR helps to fill.

Selbst proposes the use of an Algorithmic Impact Statement, modelled after the EIS with some modifications. His AIS would apply narrowly to police departments looking to acquire and use predictive policing technologies. An AIS would, in Selbst's proposal, be performed prior to using such technology. First, the AIS would, like an EIS, require policy departments to "rigorously explore and objectively evaluate all reasonable alternatives," including by having third-party vendors "(1) explain the various design choices, (2) measure the resulting efficacy using the best available audit methods, and (3) evaluate the resulting disparate impact for the various systems and configurations."[68] Second, a police department would have to "devote substantial treatment to each alternative."[69] It would be required to "include the alternative of no action."[70] It would be required to identify a preferred alternative among the various algorithm design choices disclosed.[71] And finally, police would have to include proposed mitigation measures in the AIS.[72] To address various concerns about the EIS model, Selbst emphasizes the importance of public disclosure and comment, and judicial oversight with not just procedural but substantive bite.[73]

[66] See Wright and Raab, 'Constructing a Surveillance Impact Assessment', 615, describing the sixteen steps as follows:
1. Determine whether a PIA (or SIA) is necessary (threshold analysis).
2. Identify the PIA (or SIA) team and set the team's terms of reference, resources and time frame.
3. Prepare a PIA (or SIA) plan.
4. Determine the budget for the PIA (or SIA).
5. Describe the proposed project to be assessed.
6. Identify stakeholders.
7. Analyse the information flows and other impacts.
8. Consult with stakeholders.
9. Determine whether the project complies with legislation.
10. Identify risks and possible solutions.
11. Formulate recommendations.
12. Prepare and publish the report, e.g., on the organisation's website.
13. Implement the recommendations.
14. Ensure a third-party review and/or audit of the PIA (or SIA).
15. Update the PIA (or SIA) if there are changes in the project.
16. Embed privacy awareness throughout the organisation and ensure accountability.

[67] Mantelero, 'AI and Big Data', 761.
[68] Selbst, 'Disparate Impact in Big Data Policing', 173.
[69] Id.
[70] Id. at 176
[71] Id.
[72] Id. at 177.
[73] Id. at 178.

Katyal incorporates elements of Selbst's proposal into her suggestion of a Human Impact Statement in Algorithmic Decision-making.[74] She recommends as a backstop a substantive, rather than purely procedural, commitment to algorithmic accountability and antidiscrimination.[75] And she adds that companies should also (1) identify potentially impacted populations and determine their status-based categories; (2) identify the effect of uncertainty or error on those groups; and (3) study whether the decision will have an adverse impact on a particular subpopulation.[76] The single biggest difference, however, between Katyal's and Selbst's proposals is that Katyal recommends the HIA in AI as a voluntary measure undertaken by private industry, rather than required by law.

The AI Now Institute, a research institute housed at New York University,[77] issued a report building on Selbst's proposal.[78] The report's authors call for a pre-procurement Algorithmic Impact Assessment before all public agencies (not just police) commit to the use of an automated decision-making system.[79] Like Selbst's proposal, the AI Now proposal is limited to covering the public sector. Like Selbst's proposal, it would be mandatory rather than voluntary. Unlike Selbst's proposal, it goes beyond the policing context.[80]

At first glance, the AI Now proposal looks similar to an EIS in that it must be done prior to implementing a project. Like an EIS, the proposed AIA requires agency disclosure and a public comment period. Unlike an EIS, however, the proposed AIA is envisioned as being renewed every two years.[81] And a substantial portion of AI Now's proposal is dedicated to ongoing processes to be established by the AIA, including both meaningful access for researchers and auditors once systems are deployed,[82] and individual due process for those affected by the system's decisions.[83]

Finally, we return to the context of the GDPR. Mantelero discusses the idea of a Human Rights, Social and Ethical Impact Assessment (HRSEIA) in the AI context.[84] A hybrid between a Human Rights Impact Assessment and a Privacy Impact Assessment, the HRSEIA suggests that businesses voluntarily take into account ethical and social impact, in addition to human rights.[85] Mantelero emphasizes the role of such an impact assessment in addressing the collective dimensions of data harms.

At its core, Mantelero's HRSEIA has three features: it is participatory, it is transparent, and it is circular in nature.[86] Practically, it consists of a self-assessment questionnaire, sometimes leading to evaluation by an ad hoc committee of experts.[87] Stakeholder engagement is encouraged but not required.[88] Similarly, public disclosure is encouraged.[89] Mantelero explains that while this proposal is "[i]n line with the declared intent of the GDPR," he does not

[74] Katyal, 'Private Accountability in the Age of Artificial Intelligence', 115.
[75] Id.
[76] Id. at 116.
[77] Resiman et al., Algorithmic Impact Assessment, cit.
[78] Id.
[79] Id. at 8.
[80] Id. at 6.
[81] Id. at 10.
[82] Id. at 18
[83] Id. at 16.
[84] Mantelero, 'AI and Big Data', passim.
[85] Id. at 762.
[86] Id. at 759.
[87] Id at 758.
[88] Id. at 769.
[89] Id.

understand the HRSEIA to be required by the GDPR.[90] Several other commentators have recently discussed the DPIA and the role it plays in the context of algorithmic accountability more generally.[91]

Notably, most of these proposals for Impact Assessments centrally emphasize release of information to the public.[92] This is necessary both to obtain external input into how a system is developed, trained, or monitored, and to gain public legitimacy and acceptance for the use of a system. The kind of information released to the public can be more in the nature of a summary or an overview; it is not necessarily the source code.[93] Some suggest a tiered release of information, with summaries released to the public and detailed or sensitive information released only to regulators or experts.[94] Thus more recent proposals also call for expert input and oversight as a central component of the impact assessment process—that companies (or government agencies) use Impact Assessments to come up with, and stick to, a plan for third-party expert oversight over a system's development and eventual ongoing use.[95]

6. The Data Protection Impact Assessment (DPIA) as an Algorithmic Impact Assessment

The GDPR's Data Protection Impact Assessment (DPIA) will serve, in the automated decision-making context, as a version of an Algorithmic Impact Assessment. It thus may prove to be an example for governments around the world considering using impact assessments as a tool to achieve algorithmic accountability.

Article 35(3)(a) requires a DPIA in case of "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"(emphasis added). Interpreting this provision, the Guidelines on ADM mandate DPIAs for all automated decision-making, creating a categorical requirement that applies "in the case of decision-making including profiling with legal or similarly significant effects that is not wholly automated, as well as solely automated decision-making defined in Article 22(1)".[96] Moreover, as Casey, Farhangi, & Vogl have noted, "demonstrating that a DPIA is not necessary will, in many instances, itself require a DPIA".[97] We note, too, that at least one Member State, Slovenia, requires algorithmic impact assessments as a specific safeguard in case of automated decision-making under Article 22(1) of the GDPR.[98]

---

[90] Id. at 762

[91] Bryan Casey, Ashkon Farhangi, & Roland Vogl, 'Rethinking Explainable Machines: The GDPR's 'Right to Explanation' Debate and the rise of Algorithmic Audits in Enterprise', 34 *Berk. Tech. L. J.* 143 (2019), 170-184; 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. 18 (2017), 77-80.

[92] Andrew D. Selbst, 'Disparate Impact in Big Data Policing', *Georgia Law Review* 52, no. 1 (19 February 2018): 118; Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', 4.

[93] Council of Europe, 'Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data' (Strasbourg, 23 January 2017), 4; Selbst, 'Disparate Impact in Big Data Policing', 190; Kristian Lum and William Isaac, 'To Predict and Serve?', *Significance* 13, no. 5 (2016): 14–19, https://doi.org/10.1111/j.1740-9713.2016.00960.x.

[94] Mantelero, 'AI and Big Data', 766.

[95] Christian Sandvig et al., 'Auditing Algorithms : Research Methods for Detecting Discrimination on Internet Platforms', 2014; Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', 18–20.

[96] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, Adopted on 3 October 2017, As last Revised and Adopted on 6 February 2018, WP251rev.01, 29.

[97] Casey, Farhangi, and Vogl, 'Rethinking Explainable Machines', 176.

[98] Predlog Zakona o varstvu osebnih podatkov – predlog za obravnavo – nujni postopek – Novo Gradivo ŠT. 2, http://www.mp.gov.si/fileadmin/mp.gov.si/pageuploads/mp.gov.si/novice/2018/ZVOP-2_NG_2_apr.pdf

In this section, we address the DPIA as an Algorithmic Impact Assessment. We identify what the purpose of the DPIA is in the GDPR, and what it must include. Understanding the DPIA's purpose in algorithmic governance both clarifies what the content should be, and points to several shortcomings in the current conception of it.

6.1. What is Required in a DPIA?

The GDPR describes a DPIA as "an assessment of the impact of the envisaged processing operations on the protection of personal data" (Art.35). That assessment, per the text of the GDPR, must include: a description of the "processing operations" (in this case, the algorithm) and the purpose of the processing; an assessment of the necessity of processing in relation to the purpose; an assessment of the risks to individual rights and freedoms; and importantly, the measures a company will use to address these risks and demonstrate GDPR compliance, including security measures (Art.35(7))(Rec. 84, 90).

The DPIA must occur before a company implements a system. That is, a company must assess a system, and propose risk-mitigation measures, before data processing takes place (Art.35(1)). But the GDPR also envisions iteration. For example, if the risk posed by a system changes, a company must assess whether it is complying with its own Impact Assessment (Art.35(11)). It should also under such circumstances review and possibly revise the DPIA itself.

The DPIA Guidelines suggest an even more dynamic view of DPIAs. They suggest that DPIAs should as a matter of good practice actually be continuous, "updated throughout the lifecycle [of the] project," and that they should be re-assessed or revised at least every 3 years. "Carrying out a DPIA is a continual process, not a one-time exercise," per the DPIA Guidelines.[99] This continual process involves assessing risk, deploying risk-mitigation measures, documenting their efficacy through monitoring, and feeding that information back into the risk assessment and ongoing process. The DPIA Guidelines envision this process as running "multiple times."

The GDPR also lays out procedural requirements for the DPIA. Differing from the Impact Assessments imagined in the literature, DPIAs do not involve a period of public comment or input. They do require consultation with an internal but independent Data Protection Officer, if a company has one. Many companies that are required to perform algorithmic impact assessments will likely have Data Protection Officers in place (Art.38).[100]

In lieu of public or formal stakeholder consultation, the GDPR requires consultation "where appropriate" with impacted individuals (Art.35(9)).[101] This puts in place one method for external input, from impacted individuals rather than external experts or the public. The DPIA Guidelines envision that this input could be, for example, in the form of surveys crafted by companies and sent to future customers. This would make external input less meaningful than, say, deep consultation with a board of representative of civil society members or chosen

[99] Article 29 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, 14.

[100] Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA), 15. See Also Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 30: "An additional accountability requirement is the designation of a DPO, where the profiling and/or the automated decision-making is a core activity of the controller and requires regular and systematic monitoring of data subjects on a large scale (Article 37(1)(b)".

[101] In the original proposal of the Commission, consultation with data subjects was mandatory (Article 33[4]). The Parliament's text argued that this 'represents a disproportionate burden on data controllers' (amendment 262). Accordingly, the approved Article 35(9) requires consultation only 'where appropriate' and 'without prejudice to the protection of commercial or public interests or the security of the processing operations''. Reuben Binns, 'Data Protection Impact Assessments: A Meta-Regulatory Approach', *International Data Privacy Law* 7, no. 1 (1 February 2017): 28, https://doi.org/10.1093/idpl/ipw027.

community representatives, as envisioned in the literature on impact assessments reviewed above.[102] The DPIA Guidelines explain that if companies do not seek these external views, they have an obligation to justify this decision.[103] In addition, if companies do seek these views and then disregard them, they must document why they have chosen to disregard external input. [104]

As for other forms of external oversight, the DPIA Guidelines recommend, but do not require, seeking advice from independent experts, ranging from lawyers and sociologists to data security experts.[105] The GDPR does *not* generally require most DPIAs to be overseen by a public authority (the Data Protection Authority). But if a risk assessment indicates that processing would result in *high risk* in the absence of measures taken by the controller to mitigate the risk, then a company must consult with the regulator before processing (Art.36). Thus a company effectively decides itself whether it should be subject to regulatory oversight, as part of the impact assessment process.

In the biggest departure from the Impact Assessment proposals discussed above, DPIAs are not legally required to be released to the public, even when finalized.[106] As the Guidelines explain, "[p]ublishing a DPIA is not a legal requirement of the GDPR…[h]owever, data controllers should consider publishing their DPIA, or perhaps part of their DPIA."[107] The Guidelines caution that it is a good practice to publish DPIAs, especially where members of the public are impacted. But companies need not publish the entire assessment; the published DPIA "could even consist of just a summary of the DPIA's main findings."[108] Additionally, as some scholars have remarked, there are cases in which full disclosure of the assessment results may be limited by the legitimate interests of the data controller, such as interests in the confidentiality of information, in security, and in competition.[109]

The GDPR text and DPIA Guidelines thus give an overview, but little specific guidance on what exactly a company must put in a DPIA report in the context of algorithmic impact assessments. Unlike the Impact Assessments proposed in the legal literature, neither the text nor Guidelines require public input or public disclosure, though they each suggest both as best practices. This has led one proposal to dismiss the GDPR's DPIAs as "not shared with the public, and hav[ing] no built-in external researcher review or other individualized due process mechanisms".[110] As we discuss below, this is not entirely correct, at least in the context of automated decision-making.

6.2. What is the Purpose of a DPIA, in the context of the GDPR's Algorithmic Governance?

[102] Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA), 15. See also Dariusz Kloza et al., 'Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework towards a More Robust Protection of Individuals', D.Pia.Lab Policy Brief No. 1/2017, n.d., 4, https://cris.vub.be/files/32009890/dpialab_pb2017_1_final.pdf.
[103] Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA), 15.
[104] Id., 15.
[105] Id., 15.
[106] Id.,18.
[107] Id., 17.
[108] Id.
[109] Alessandro Mantelero, 'AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment', *Computer Law & Security Review* 34, no. 4 (1 August 2018): 766, https://doi.org/10.1016/j.clsr.2018.05.017; Frank Vanclay et al., 'Social Impact Assessment: Guidance for Assessing and Managing the Social Impacts of Projects' (International Association for Impact Assessment, April 2015), https://www.iaia.org/uploads/pdf/SIA_Guidance_Document_IAIA.pdf; Simon Walker, *The Future of Human Rights Impact Assessments of Trade Agreements*, School of Human Rights Research Series, v. 35 (Antwerp ; Portland: Intersentia, 2009), 39–42.
[110] Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', 7.

We posit that in the context of the GDPR's algorithmic governance regime, the DPIA should be understood as a nexus between the GPDR's two approaches to governing algorithmic decision-making. The DPIA links the GDPR's system of individual rights to its systemic governance of algorithms.

Understanding the DPIA in this way both clarifies its potential content, and leads us to observations about how the DPIA as Algorithmic Impact Assessment might be implemented and even improved. The DPIA is not a perfect Algorithmic Impact Assessment. As a tool in the GDPR's overall algorithmic governance regime, however, it has more potential than might initially meet the eye.

6.3. How Understanding the DPIA's Dual Role Helps Clarify Its Content

The DPIA has two roles: as a tool in the GDPR's systemic (and collaborative) governance regime, and as an element of the GDPR's protection of individual rights. Understanding the DPIA in this way—as a connection between the two regulatory subsystems—lets us better understand how it is meant to function as an Algorithmic Impact Assessment, even to the extent of further clarifying its content. It also leads us to some insights in the next Section (7) about the layers of algorithmic explanations produced by, and to be released according to, the GDPR.

When understood as part of the GDPR's collaborative governance of algorithms,[111] the DPIA is a form of monitored self-regulation. Binns has similarly identified the DPIA more generally as "meta-regulation."[112] Collaborative governance is centrally concerned with affecting management culture and creating meaningful changes within a company.[113] Monitored self-regulation attempts to change both company decision-making processes and decision-making heuristics.[114] The DPIA, in the context of algorithmic decision-making, tasks companies with considering risks of unfairness, error, bias, and discrimination, and with coming up with concrete ways of mitigating those risks. This affects firms' decisional heuristics by dictating, through the text, Recitals, and the Guidelines, what values a company must consider in building and overseeing algorithmic decision-making.

The process of conducting the DPIA—taking input from impacted individuals, consulting with an independent Data Protection Officer, consulting with a regulator where required, and involving both internal and external experts—is meant to change internal company processes.[115] As others have noted, baking in a compliance culture can be valuable, even where public oversight and input is not sought.[116] And the DPIA can also be understood in this context as a documentation requirement, or even the precursor to a reporting requirement, creating records that can later be sought and inspected by regulators under the GDPR's extensive information-forcing capabilities.[117]

The DPIA also, however, has an unexplored role in the GDPR's system of individual rights.

---

[111] Kaminski, 'Binary Governance', 57.
[112] Binns, 'Data Protection Impact Assessments', 23, 29 has similarly described DPIAs as 'meta-regulation,' which he characterizes as a narrower subset of co-regulation, "a means for the state to make corporations responsible for their own efforts to self-regulate".
[113] Alexander A. Boni-Saenz, 'Public-Private Partnerships and Insurance Regulation', *Harvard Law Review* 121 (2008): 1375; Freeman, 'The Private Role in the Public Governance'; Bamberger, 'Regulation as Delegation'.
[114] Bamberger, 'Regulation as Delegation', 435.
[115] See also Binns, 'Data Protection Impact Assessments', at 23.
[116] Bamberger, 'Regulation as Delegation', 467. See also Katyal, 'Private Accountability in the Age of Artificial Intelligence', 140.
[117] Article 29 Data Protection Working Party, Guidelines on Data Protection Impact Assessment (DPIA), 20. See also the investigatory powers of Data Protection Authority at Article 58(1) GDPR. See also Selbst & Barocas's call for documentation requirements in 'The Intuitive Appeal of Explainable Machines', 87 *Fordham Law Review* 1085 (2018) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3126971

First, the DPIA can serve as a source of material for the much-discussed disclosures to individuals about algorithmic decision-making: the individual notification and access rights. Remember, data subjects have a right to receive "meaningful information" about the "logic involved, as well as the significance and the envisaged consequences" of automated decision-making (Arts.13, 14, 15). A DPIA must contain, as mentioned above, "a systematic description of the envisaged processing operations and the purposes of the processing…"(Art.35(7)). If companies are already internally describing automated decision-making at a systematic level as part of the DPIA process, those internal descriptions could be disclosed to individuals, or at least serve as the basis for these disclosures, in addition to being released to the public in the form of summaries.

Similarly, a DPIA must include an assessment of "the risks to the rights and freedoms" of individuals, and individuals have a right in the context of automated decision-making to be informed of the "significance and envisaged consequences" of such decision-making (Arts.35, 13, 14, 15). Again, as a company conducting automated decision-making must conduct a DPIA, it should consider how the information it produces in that process might also feed into or even satisfy the individual rights requirements under the GDPR.

Second, despite other commentators' dismissal of DPIAs as failing to put in place individual due process,[118] the Guidelines on ADM explicitly envision the DPIA as an essential part of establishing suitable measures to safeguard individual rights, including a version of individual due process. The GDPR requires companies using solely automated decision-making, under the exceptions to its ban on such practices, to implement suitable measures to protect individual rights (Art.22). The Guidelines counsel that companies should use DPIAs to "identify what measures they will introduce to address the… risks involved".[119] Suggested measures include not just the use of audits or other forms of systemic accountability, but also a number of recognizable individual rights: informing individuals about the logic involved, explaining the significance and envisaged consequences of algorithmic decision-making, providing a way to contest a decision, and providing a way to express one's point of view.[120] The Guidelines on ADM counsel companies to import the various individual rights laid out in the GDPR that are restricted to *solely* automated decision-making in the text, as a form of risk management through the DPIA process even for algorithms that more significantly involve a human decision-maker.

In other words, the GDPR (or really, the interpreting Guidelines on ADM) envisions DPIAs, in the context of algorithmic decision-making, as serving as a form of commitment-making to protecting, or even enabling, individual algorithmic due process rights. By characterizing these individual rights as *risk mitigation measures*, it both provides a substantive backstop as to what must be included in a DPIA, and tasks companies with *constituting*—through the process of performing a DPIA—what these individual rights will look like in practice. Thus the DPIA as envisioned serves as a means of expanding company commitments, changing company decision-making heuristics to include an assessment of individual due process rights. It simultaneously serves as a collaborative governance mechanism used to involve companies in constituting the substance, in practice, of individual due process rights.[121]

Finally, the DPIA has a role in linking the GPDR's system of collaborative governance to its individual rights regime through the imposition of systemic accountability measures such as audits or external review. Remember, the general DPIA Guidelines only suggest, and do not mandate, consultation with external experts. In the context

[118] Reisman et al, *Algorithmic Impact Assessment*, 10.
[119] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 30.
[120] Id., 30.
[121] Kaminski, 'Binary Governance', 18.

of algorithmic decision-making, however, external expert involvement and oversight is more like a requirement. The use of external experts is framed by the Guidelines on ADM as a necessary risk-mitigation measure for algorithmic decision-making.

The reasoning goes as follows. Recital 71 requires, in the context of algorithmic decision-making, the use of "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…and that prevents, inter alia, discriminatory effects" (Rec. 71). Malgieri and Comandé observe that this requirement effectively expands the GDPR's "suitable safeguards" requirement from the series of individual due-process-like protections enumerated in the GDPR's text, to a far broader set of systemic accountability measures, including third-party auditing (Art.22).[122]

The Guidelines on ADM's list of best practices for suitable safeguards over algorithmic decision-making supports this interpretation, including recommendations that companies use both internal and external audits and external review boards.[123] When a company that deploys algorithmic decision-making conducts its DPIA, it will refer to the Guidelines on ADM's list of best practices in establishing risk-mitigation measures that are already regulator-approved. This means that in practice, in the context of algorithmic decision-making a company running through the cyclical DPIA process discussed above will likely incorporate external oversight and input at the risk mitigation stage, bringing external input into the cycle despite the fact that it is not a formal procedural requirement for DPIAs in general.

Conceptually, the implications of this are even broader. By characterizing third-party and expert oversight as a form of "suitable safeguard" or "suitable measure" to protect *individual rights* in the face of automated decision-making, the Guidelines on ADM link individual rights protection with collaborative governance techniques. Companies are tasked with coming up with ways to prevent error, bias, discrimination, and other harms to individual rights, and external oversight is imposed over how they choose to address these problems. That external oversight itself is simultaneously conceptualized as a crucial aspect of individual rights in the GDPR, standing in for individuals to ensure that they are not subjected to an unfair, arbitrary, discriminatory, or erroneous system.

A simpler way to say this is that expert oversight in the DPIA process serves two, or even three, roles: it watches the companies as they come up with ways of addressing problems with algorithmic decision-making, and it reassures individuals that their dignity and other rights are being respected by a fair system.[124] It also provides legitimation, or justification. As the mechanism through which this external oversight is implemented, the DPIA thus connects the two approaches to algorithmic governance in the GDPR.

6.4. Shortcomings of the DPIA

[122] Malgieri and Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation', 248.
[123] Article 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, 28, 32. See Casey et al., 'Rethinking Explainable Machine', 170-180, emphasizing the centrality of algorithmic audits.
[124] One of us has referred to these as first-order and second-order transparency. See Kaminski, 'Binary Governance', 28. Binns, 'Data Protection Impact Assessments', 32 discusses a similar notion in the regulatory theory literature, Gilad's concept of regulatory tiers. Gilad, 'It runs in the family: Meta-regulation and its siblings', *Regulation & Governance*, 4(4)485, 497 (2010).

The biggest shortcoming of the DPIA is that it does not include a mechanism for mandatory disclosure to the public.[125] Public disclosure, as discussed above, is understood by many to be an essential element of Impact Assessments as a policy tool.[126] Public-facing disclosure enables public feedback, both in the form of market feedback (enabling individuals to avoid companies with bad policies) and in the form of regulatory feedback over the longer term (enabling individuals to elect representatives who will put in place laws that will prevent bad company behaviour). By failing to mandate public disclosure, the GPDR's DPIA fails to trigger both of these mechanisms, which are essential components of a functioning collaborative governance regime.

This failure could be drastic. The GDPR puts a lot of faith in the behaviour of companies, and in the capacity of regulators. As discussed, the GDPR often tasks companies with coming up with the substance of (a) how individual rights will be implemented and (b) how to address unfairness, biases, and discrimination concerns about algorithms. In the absence of public oversight, how can we be sure that this hybrid system of individual rights and collaborative governance is working towards the public good?

One possible answer is to use regulatory oversight instead. But the GDPR's enforcers have not, historically, been well-resourced relative to the companies they regulate. Tasking regulators with extensive monitoring also forgoes some of the touted benefits of governing through public-private partnerships, including lowered costs and incorporating external third-party expertise. By failing to require public disclosure of Impact Assessments, the GDPR fails to activate necessary third parties in its governance regime, such as civil society actors or civic-minded experts who might not be recruited for auditing purposes. Similarly, the DPIA fails to involve serious stakeholder input, unless companies understand the Guidelines on ADM's emphasis on expert boards and third-party audits to be mandatory.[127]

Individual notification and access rights could instead do some of the work. This is somewhat more convincing. If companies indeed link their DPIA content to what they disclose to individuals (for example, disclosing the "logic involved" in a decision-making system), then it is likely that these disclosures will make their way to other third parties, who may be able to provide the expertise and oversight over company self-governance. For example, an individual who feels she has been discriminated against might disclose the information she has received about a system's decisional to a civil society group, which could in turn help publicize the story and the information, triggering market mechanisms or regulatory feedback from the public or oversight by external experts. This is a more attenuated way of getting at the same outcome as public disclosure, however, and risks failing entirely if companies significantly disaggregate the DPIA process from individual disclosure rights.

6.5. Lessons for Calls for Algorithmic Impact Assessments Generally

Our GDPR-specific analysis has implications for proposals for algorithmic impact assessments generally. Our research into the GDPR's version of AIAs suggests that the proposals discussed above have largely missed several important observations.

---

[125] Kloza et al, 'Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework towards a More Robust Protection of Individuals', 3; Michael Veale, Reuben Binns, and Jef Ausloos, 'When Data Protection by Design and Data Subject Rights Clash', *International Data Privacy Law* 8, no. 2 (1 May 2018): 118, https://doi.org/10.1093/idpl/ipy002.

[126] Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', 13. See also Selbst, 'Disparate Impact in Big Data Policing', 119. See also Froomkin, 'Regulating Mass Surveillance', 1790.

[127] See Binns, 'Data Protection Impact Assessments', at 33; Casey et al., 'Rethinking Explainable Machines', 180.

First, AIAs are not best understood as a stand-alone mechanism. In the context of the GDPR, they are one part of a much larger system of governance.[128] Only one author among the above—Katyal—considers how impact assessments interact with other tools in the regulatory toolkit (discussing the concurrent need for whistleblower protection and exemptions from trade secrecy law).[129] In the context of the GDPR, both Edwards and Veale and Casey et al. point to the DPIA's role in algorithmic accountability, but do not discuss at length its relationship to other accountability tools in the GDPR.[130] Our analysis suggests that impact assessments are just one tool in a larger regulatory ecosystem, and may work best when they are not deployed alone, and are instead understood as entwined with other regulatory tools.

Second, impact assessments can serve as a connection between collaborative governance and individual rights.[131] The information a company creates during the Impact Assessment process can feed into what it provides to individuals and to the public at large. The procedures an Impact Assessment puts in place can serve not just to prevent error, bias, and discrimination, but also to legitimize a system or even respect an individual's dignity within it. This dual role is exemplified by the GDPR's DPIA. In the GDPR context we found only one author, Binns, who identified that the GDPR's version of impact assessments is a kind of collaborative governance with the private sector (or what he identifies as "meta-regulation").[132] But this led Binns to critique the GDPR's version of the DPIA for inadequate public disclosure and stakeholder involvement, not to look to how the DPIA connects to the broader system of both collaborative governance tools and individual rights in the GDPR.

Third, as part of a larger system of governance, there are unexplored connections between the GDPR's DPIA and its underlying substantive individual rights and substantive principles. It is true that many of the GDPR's individual rights and principles about algorithmic decision-making are articulated in broad, sometimes aspirational, terms.[133] Unlike an EIS, the GDPR's version of the AIA has a substantive backstop, in for example Recital 71's admonishment that a data controller should minimize the risk of error and prevent discriminatory effects. The oddity is the GDPR's circularity: the AIA helps not just to *implement* but to *constitute* both these substantive backstops and the GDPR's individual rights. Thus there is a substantive backstop to company self-regulation through impact assessments—but it is a moving target, in part given meaning by affected companies themselves.

Finally, because the AIA links individual and systemic governance, we understand the GDPR's version of the AIA to be both the potential source of, and the mediator between, what we refer to below as "multi-layered explanations" contemplated in the GDPR. Several of the above scholars, including both Mantelero and Wright & Raab, emphasize the often collective dimensions of surveillance and data processing.[134] The GDPR's system of individual rights threatens by itself to miss the impact of surveillance, or in this case, automated decision-making, on groups,

[128] Edwards & Veale, 'Slave to the Algorithm', 77-80 understand this, as they discuss the DPIA in the context of many other rights in the GDPR. See also Kaminski, 'Binary Governance', 69.
[129] Katyal, 'Private Accountability in the Age of Artificial Intelligence', 117.
[130] Edwards & Veale, 'Slave to the Algorithm', 77-80; Casey et al., 'Rethinking Explainable Machines', 170-184.
[131] Only one proposal, to our knowledge, suggests using Impact Assessments to establish something resembling individual rights—a system of "enhanced due process mechanisms for affected individuals". Dillon Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', AI Now Institute, n.d., https://ainowinstitute.org/aiareport2018.pdf
[132] Binns, 'Data Protection Impact Assessments', 29 describing DPIAs as "enforced risk-assessment, and compliance with self-imposed, stakeholder-influenced policies… as an instance of 'meta-regulation".
[133] Mantelero, 'AI and Big Data', 765 (discussing how "Data protection laws adopt general principles… and general clauses… which are used to introduce non-legal social values into the legal framework").
[134] Mantelero, 'AI and Big Data', at 762-763; Wright & Raab, 'Constructing a Surveillance Impact Assessment', 615.

locations, and society at large.[135] A recent AI Now report provides an illustrative example of the problem: providing an individualized explanation for a single "stop and frisk" incident in New York City would have failed to reveal that over 80% of those subjected to stop and frisk by the NYPD were Black or Latino men.[136] But the Impact Assessment with its systemic approach to risk assessment and risk mitigation requires data controllers to analyze how the system impacts not just individuals but groups. We believe that systemic and group-based explanations uncovered during an AIA can and should be communicated to outside stakeholders, and that a case can be made that such release is required under the GDPR.

7. A Model Algorithmic Impact Assessment: Towards Multi-layered Explanations

We close this paper with a call for more work on establishing a model Algorithmic Impact Assessment specific to the GDPR that could serve also as a basis for what we call *multi-layered explanations* of algorithmic decision-making. This will involve interdisciplinary efforts: technologists to assess what risk-mitigation and accountability measures could be implemented, and lawyers and ethicists to think through how to better involve constituents and define problems. It will also involve a deeper exploration of how to link the material created during the DPIA process to the individual disclosures required under the GDPR.

To begin this conversation, we suggest that a Model Algorithmic Impact Assessment process should do at least the following. It should contemplate the involvement of civil society as a form of underused oversight. It should better involve and engage impacted individuals, not just through surveys but through representative boards, before an algorithm is deployed. It should contemplate requiring companies, or regulators, to help fund the involvement of both of the above, and provide technical expertise or the resources for obtaining technical expertise. It should involve not just external technical experts, but external experts in law and ethics to help define, or at least frame discussions of, what we mean by terms like "discrimination" or "bias."[137]

A Model Algorithmic Impact Assessment process should also deliberately widen the lens from algorithms as a technology in isolation, to algorithms as systems embedded in human systems—both those that design the technology, and those that use it.[138] There is a growing awareness that addressing problems of unfairness or bias in the technology in the abstract will be inadequate for mitigating these problems when an algorithm is implemented in practice. The risks come not just from the technology by itself, and not just from the humans who embed their values into the technology during its construction and training, but from how the humans using the algorithm are trained and constrained, or not constrained, in their use of it.[139] This connects to our suggestion that a Model Algorithmic Impact Assessment be truly continuous: a process that produces outputs, but also includes ongoing assessment and performance evaluation, especially for those algorithms that change quickly over time.

---

[135] There is a growing field of scholarship devoted to "collective data protection." See Mantelero, 'AI and Big Data', at 757, note 21; Linnet Taylor, Luciano Floridi, Bart van der Sloot (eds), Group Privacy: New Challenges of Data Technologies (Springer International Publishing, 2017).

[136] Reisman et al. at 19.

[137] For example, the COMPAS recidvisim risk assessment algorithm led to a significant public discussion over different ways of defining discrimination. Julia Angwin et al., 'Machine Bias', *Propublica* (May 23, 2016), https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (describing leading risk assessment tools forsentencing and corrections developed by Northpointe); Sam Corbett-Davies et al., A Computer Program Used for Bail and Sentencing Decisions Was Labeled Biased Against Blacks. It's Actually Not that Clear., WASH. POST, (Oct. 17, 2016) https://www.washingtonpost.com/news/monkey-cage/wp/2016/10/17/can-an-algorithm-be-racist-our-analysis-is-more-cautious-thanpropublicas.

[138] See Andrew D. Selbst et al., 'Fairness and Abstraction in Sociotechnical Systems', in *Proceedings of the Conference on Fairness, Accountability, and Transparency*, FAT* '19 (New York, NY, USA: ACM, 2019), 59–68, https://doi.org/10.1145/3287560.3287598.

[139] Id., 61 on COMPAS Case.

Substantively, a Model Algorithmic Impact Assessment could take advantage of the fact that it is conducted on a system-wide level to search for, and mitigate, social harms that go beyond impacted individuals.[140] For example, a Model AIA could be used to root out discrimination not just against particular individuals but against marginalized communities, identifying discrimination patterns that would be impossible to find through individual disclosures alone.[141] A Model Algorithmic Impact Assessment could explicitly require assessment of performance metrics, on a system-wide and ongoing basis, and require disclosure of these metrics to external experts.[142] This would not stretch the purpose of the DPIA, and would fill a current gap in the GDPR's current algorithmic accountability and disclosure regime.[143]

To make the Impact Assessment process meaningful, Data Protection Authorities must be willing to spot check and enforce against captured versions of it. While the GDPR does not require regulatory involvement in all DPIAs, DPAs could use the GDPR's broad information-forcing powers to inspect particular companies and check for compliance. This spot-checking might go not just to efficacy of process, but to substantive problems with algorithmic decision-making. DPAs might, too, over time establish more concrete best practices or support the establishment of sector-specific codes of conducts around algorithmic fairness, as suggested in the Guidelines on ADM. Several implementing Member States have already put in place substantive backstops around algorithmic decision-making, prohibiting decision-making based on particular factors, or that is discriminatory or biased. Slovenia, as mentioned, couples this substantive prohibition against discrimination with a required impact assessment process. This dual approach of linking impact assessments to substantive prohibitions may help to tether internal company risk mitigation measures to the public good.

7.1. Comparing DPIA content with Algorithmic Accountability requirements under the GDPR

Thus far few commentators have linked the Guidelines on Automated Decision-Making to the DPIA process.[144] Here, we connect the GDPR's text on DPIAs to these Guidelines on ADM, to show how the required content of DPIAs might serve as the basis for disclosures controllers are required to make to individuals.

Article 35(7) GDPR requires that a DPIA should contain:

1. "a systematic description of the envisaged processing operations and the purposes of the processing, (…)
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

[140] Mantelero, 'AI and Big Data', passim; L. Edwards, D. McAuley, and L. Diver, 'From Privacy Impact Assessment to Social Impact Assessment', in *2016 IEEE Security and Privacy Workshops (SPW)*, 2016, 53–57, https://doi.org/10.1109/SPW.2016.19; David Wright and Charles D. Raab, 'Constructing a Surveillance Impact Assessment', *Computer Law & Security Review* 28, no. 6 (December 2012): 613–26, https://doi.org/10.1016/j.clsr.2012.09.003; Charles Raab and David Wright, 'Surveillance: Extending the Limits of Privacy Impact Assessment. In: Wright David, De Hert Paul, Editors. Privacy Impact Assessment. Dordrecht'.

[141] Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', 18; See also Pauline Kim, 'Auditing Algorithms for Discrimination', *University of Pennsylvania Law Review Online* 166, no. 1 (1 January 2017): 196, https://scholarship.law.upenn.edu/penn_law_review_online/vol166/iss1/10.

[142] Edwards & Veale, 'Slave to the Algorithm?', 80.

[143] Reisman et al., 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability', 8. Edwards & Veale, 'Slave to the Algorithm?', 80.

[144] But see Casey et al, 'Rethinking Explainable Machines', 170-184, largely focused on practical compliance, discussing the Guidelines on Automated Decision-Making and DPIA.

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".

The Guidelines on DPIAs clarify that a *systematic description* should include: the nature, scope, context and purposes of the processing, (categories of) personal data, recipients and storage; a functional description of the processing operation and the assets on which personal data rely.[145]

If we compare the GDPR's algorithm accountability requirements with these requirements for DPIAs, the similarities are striking (see table 1):

Table 1. Comparison between DPIA duties and GDPR Algorithmic accountability duties under the GDPR

Content of DPIA (Art. 35(7)) GDPR	**GDPR algorithmic accountability disclosure duties Articles 13-15, 22 (WP Guidelines, p. 30)**
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.	Describing: a. *the categories of data* used c. *how any profile* used in the automated decision-making process is *built*, including any statistics used in the analysis;
2. an *assessment of the necessity and proportionality* of the processing operations in relation to the purposes	b. *why* the categories of data are *pertinent* d. *why* this profile is *relevant* to the automated decision-making process;
3. an *assessment of the risks to the rights and freedoms* of data subjects referred to in paragraph 1; and	e. *how it is used for a decision* concerning the data subject: ○ i.e. which kinds of legal or similarly significant effects under Art.22(1)
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.	*Which safeguards are adopted in compliance with article 22(3) and (4)*: ○ e.g. Contestation, human involvement, making representation, explanation, algorithm audit, etc.

In particular, the data controller's duty to systematically describe the processing operations in a DPIA (Article 35, GDPR and point "1" in Table 1) is similar to the algorithmic transparency duty to clarify the categories of personal data used in automated decision-making and how the algorithmic profiling is built (Articles 13-15 and 22 GDPR and points "a)" and "c)" of Table 1).

Analogously, the controller's duty to assess the necessity and proportionality of the processing operations in the DPIA (Article 35, GDPR and point "2" in Table 1) is similar to the algorithmic transparency duty to explain the pertinency of personal data used and the relevance of the profiling (Articles 13-15 and 22 GDPR, points b) and d) in Table 1). The controller's duty to assess the data processing risks and impacts on individuals (Article 35, GDPR and point "3" in Table 1) is similar to the transparency duty to explain the impact of the profiling use in automated decision-making (Articles 13-15 and 22 GDPR, points e) in Table 1). Lastly, the controller's duty to establish

[145] Article 29 Working Party, Guidelines on DPIA, 22.

safeguards of individual rights in case of automated decision-making (under Article 22(3) and (4) GDPR) is similar to the duty to find and describe measures envisaged to address the risks in DPIA (4).

In other words, in case of automated decision-making, the outcome of the DPIA steps in the GDPR (points 1-4 in Table 1) might correspond to profiling transparency duties in the GDPR (as interpreted by Article 29 Working Party, in Annex I).

7.2. Towards Multi-layered Explanations from Algorithmic DPIA

Our perspective on the DPIA as linking systemic governance to individual rights has implications, too, for the GDPR's overall approach to algorithmic explanations. The DPIA process seems in our view to suggest what we call "*multi-layered explanations*" for automated decision-making. These explanations may be crafted during the DPIA process, and should be released either directly to the public or to affected individuals.

We are not the first to observe that there are multiple layers of explanations of algorithmic decision-making required in the GDPR.[146] These stem from the GDPR's two types of individual transparency requirements, articulated in Articles 13, 14, and 15 on individual notice and access, and its algorithm-specific provisions in Article 22. Edwards & Veale in particular have suggested that individuals subject to algorithmic decision-making should be provided both of what they call "model-centric" and "subject-centric" explanations.[147] Model-centric explanations, they suggest, should include: the family of model, input data, performance metrics, and how a model was tested. Subject-centric explanations should include counterfactuals (that is, what changes would change the outcome of an individual decision), the characteristics of similarly classified individuals, and the confidence a system has in an outcome.[148]

With our perspective on the GDPR's two-pronged approach to algorithmic accountability, and our emphasis on the role of the DPIA, we understand there to be more layers: individual explanations, group explanations, and systemic explanations, both internal and external. And unlike Edwards & Veale, we have more optimism that these multi-layer explanations can be grounded either in the text or subtext of the GDPR.

Looking at the GDPR through the lens of individual rights reveals the by now familiar two layers of explanations: a right to an explanation of the model, and a right to an individual explanation of an individual decision. The GDPR requires disclosure to individuals of "meaningful information about the logic involved" in automated decision-

[146] Wachter et al, 'Why a Right to Explanation', 78; Selbst & Powles, 'Meaningful Information', cit., 241; Edwards & Veale, 'Slave to the Algorithm?', 52 ff. See also European Commission's High Level Expert Group on Artificial Intelligence, *Ethics Guidelines for Thrustworthy AI*, https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines, 15 ("The degree to which explicability is needed is highly dependent on the context and the severity of the consequences if that output is erroneous or otherwise inaccurate").

[147] Edwards & Veale, 'Slave to the Algorithms', 22. They helpfully add to the conversation about the kinds of explanations that could be provided: (A) **model-centric explanations** that disclose, for example, the family of model, input data, performance metrics, and how the model was tested; and (B) **subject-centric explanations** that disclose, for example, not just counterfactuals (what would I have to do differently to change the decision?) but the characteristics of others similarly classified, and the confidence the system has in a particular individual outcome.

[148] See also Sandra Wachter, Brent Mittelstadt, and Chris Russell, 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR', *Harvard Journal of Law & Technology*, 2018, *ArXiv:1711.00399 [Cs]*, 1 November 2017, http://arxiv.org/abs/1711.00399.

making on a systemic level.[149] It also establishes, we believe, the right to individual explanation of an individual decision.[150]

Looking through the lens of the DPIA as a nexus between systemic governance and individual rights, however, reveals something more. The DPIA process entails a whole web of explanations: to internal oversight bodies ranging from the DPO to internal auditors, to external third-parties such as auditors and expert boards, and as part of the overall assessment process. These explanations are of differing degrees of breadth, depth, and technological complexity. But they establish a complex system of information flows, beyond the individual transparency requirements of the GDPR. These information flows will often require intermediation—that is, explanation—not just disclosure of existing information.[151] As discussed above, these various disclosures and explanations likely will include not just systemic and individual analysis, but group-level analysis of how an algorithm might impact particular classes of individuals, or particular locations. Thus the DPIA process may address some of the concerns some scholars have about the DPIA's focus on individual rights, to the exclusion of groups.

Whether these explanations will go beyond the doors of companies is a different question. As discussed, a DPIA is not required to be made public, but its public disclosure is highly recommended, at least in the form of meaningful summaries.[152] We believe that analysis of how algorithms impact particular groups or places should be included in these public disclosures. This will help drive policy conversations in the way anticipated by most discussions of public disclosure of impact assessments. It will also go some of the way to addressing concerns about lack of stakeholder involvement and regulatory oversight over the impact assessment process, though we also counsel that companies aware of impacts on particular places or groups should seek out impacted individuals at an earlier stage of the process.

Moreover, there may be an argument for disclosure of group- or location-based explanations to individuals as part of the GDPR's individual transparency regime. That is, even if DPIAs are not required to be made public, and even if companies decide not to disclose to the public what they discover about the impact of algorithmic decision-making on particular groups, they may nonetheless have to do so to impacted individuals under Article 22. We understand the GDPR to suggest a connection between required DPIA analysis of systemic risks of unfairness and discrimination, and the individual rights to contestation, to express one's view, and to human intervention.[153] That is, for a person to be able to effectively invoke her right to contest an algorithmic decision, she may need to know whether she is being treated the same or differently as other similarly situated individuals. For the GDPR's series of individual rights to be meaningful, individuals need to know not just information about a particular stand-alone decision, but information about the algorithm's treatment of groups, and tendency towards bias and discrimination.

This group-based explanation, which we argue can be at least implied from - if not required by - the rights to contestation or to challenge the decision, could be created based on information on affinity or group profiling uncovered during a DPIA.

[149] Arts. 13 and 14 GDPR. See also Selbst & Powles, 'Meaningful Information', 241-242, discussing how this blends individualized and systemic explanations.
[150] Kaminski, 'The Right to Explanation', 199; Malgieri and Comandé , 'Why a Right to Legibility', 246.
[151] Frank Pasquale, 'Beyond Innovation and Competition: The Need for Qualified Transparency in Internet Intermediaries', *Northwestern University Law Review*, Vol. 104, n. 1, 2010, 105-174.
[152] Article 29 Working Party, Guidelines on DPIA, 17.
[153] Art. 22, GDPR.

Finally, some scholars have remarked that what is needed is not merely an explanation, but a legal *justification* of automated decisions taken.[154] The full concept of justification is not easy to address in the data protection framework and is beyond the scope of this paper. For the limited scope of this paper, however, justifying a decision means not merely explaining the logic and reasoning behind it, but also explaining why it is a correct, lawful and fair decision, i.e. that the decision is based on proportional and necessary data processing, using pertinent categories of data and relevant profiling mechanisms.

Again, connecting the DPIA to transparency requirements may clarify what this could mean. Language about the DPIA process suggests that in addition to technical explanations of a model, data controllers should produce *justificatory* explanations of a system during a DPIA. Under the DPIA process, data controllers must prove the legal proportionality and necessity of the data processing, and so also the legal necessity and proportionality of eventual automated decisions taken (Art. 35(7)(d)). This may constitute a form of justification of data use and profiling mechanisms. Similarly, the Guidelines on ADM recommend that data controllers (in order to comply with Articles 13-15) explain the pertinence of categories of data used and the relevance of the profiling mechanism.[155] Assessing whether the data used are pertinent and the profile is relevant for a decision, as well as assessing the necessity and proportionality of the data processing in an automated decision-making system, seem to constitute justification of automated decision systems. The purpose of such assessment is not transparency about the technology itself, but an explanation about the lawfulness, fairness, and legitimacy of certain decisions.

Combining the algorithmic DPIA process and the duty to disclose information about algorithmic decisions in coordinated actions would be beneficial not just for individuals but for data controllers[156]. Combining these tasks could benefit data controllers because:

1. they could optimize efforts that would be spent for the two different tasks of DPIA and disclosure requirements by taking compliance with DPIA duties (Article 35) and feeding them into transparency duties as imposed by Article 13-15 (and 22) of the GDPR;
2. publicly disclosing (at least some parts) of the DPIA as a basis for explaining automated decisions is considered a best practice recommended in the DPIA framework,[157] in line with the data protection by design principle (Article 25 GDPR);[158]
3. disclosing information about algorithmic data processing to data subjects and collecting their reactions (through, e.g., the right to contest, to have a new decision, to have human involvement, etc.)[159] could be considered compliant with the duty to seek the view of impacted data subjects (Article 35(9) GDPR), in the continuous cycle of DPIA framework;[160]

---

[154] Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI', *Columbia Business Law Review, 2019(2)*, https://papers.ssrn.com/abstract=3248829; Kaminski, 'Binary Governance', 12–17.
[155] See Article 29 Working Party, Opinion on Automated Individual Decision-Making, Annex, p. 30.
[156] On the list of positive externalities for data controller if they disclose a 'legibility' test on algorithms see Malgieri and Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation', 259–60.
[157] See Article 29 Working Party, Guidelines on DPIA, 17. See also Kloza et al., 'Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework towards a More Robust Protection of Individuals', 2.
[158] Veale, Binns, and Ausloos, 'When Data Protection by Design and Data Subject Rights Clash', 117–18.
[159] See Article 22(3) and recital 71. See also Antoni Roig, 'Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)', *European Journal of Law and Technology* 8, no. 3 (21 January 2018), http://ejlt.org/article/view/570.
[160] See on the importance of continuous engagement of involved subject in PIA Roger Clarke, 'An Evaluation of Privacy Impact Assessment Guidance Documents', *International Data Privacy Law* 1, no. 2 (1 May 2011): 112, https://doi.org/10.1093/idpl/ipr002.

4. in general terms, the dynamic merging of an algorithmic DPIA and of multi-layered explanation might be a "suitable safeguard" to protect fundamental rights and freedoms of individuals both under Article 22(3) and under Article 35(7)(d) of the GDPR;[161]
5. developing an algorithmic DPIA and explanation safeguards in parallel (intrinsically related to the right to contest a decision, right to human-in-the-loop, etc.) might be the best way to enrich transparency with accountability safeguards[162] and overcome the "transparency fallacy" through a virtuous cycle of algorithmic auditing and continuous detection/mitigation of unfair effects. [163]

The idea of at least partially merging algorithmic accountability duties with the DPIA process also seems useful considering the most advanced literature on explanations. As discussed above, a multi-layered and multi-step explanation would be a continuous *process*, not merely a product.[164]

8. Conclusion

There is a growing literature suggesting that Algorithmic Impact Assessments are a crucial tool in establishing algorithmic accountability. This paper addresses that tool as it is implemented in the GDPR. We find that the GDPR's version of an Algorithmic Impact Assessment serves as a central connection between its two approaches to regulating algorithms: individual rights and systemic governance. That framing allowed us to identify both value in, and shortcomings of, the GDPR's Impact Assessment regime as applied to algorithmic governance.

This analysis, we hope, will have value for other discussions of Algorithmic Impact Assessments beyond the GDPR. In particular, moving from individual transparency rights and governance accountability duties in the field of automated decision-making, we suggest a model of multi-layered explanations drawn from Algorithmic Impact Assessments. Since there are several layers of algorithmic explanation required by the GDPR, we recommend that data controllers disclose a relevant summary of a system, produced in the DPIA process, as a first layer of algorithmic explanation, to be followed by group explanations and more granular, individualized explanations. More research is needed, in particular about how different layers of explanations—systemic explanations, group explanations, and individual explanations—can interact each other and how technical tools can help in developing an Algorithmic Impact Assessment that might be re-used towards GDPR-complying explanations and disclosures.

REFERENCES

Ananny, Mike, and Kate Crawford. 'Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability'. *New Media & Society* 20, no. 3 (1 March 2018): 973–89. https://doi.org/10.1177/1461444816676645.

Bamberger, K. 'Regulation as Delegation: Private Firms, Decisionmaking, and Accountability in the Administrative State'. *Duke Law Journal*, 1 January 2006, 377.

Binns, Reuben. 'Data Protection Impact Assessments: A Meta-Regulatory Approach'. *International Data Privacy Law* 7, no. 1 (1 February 2017): 22–35. https://doi.org/10.1093/idpl/ipw027.

[161] About the link between 'risks to rights and freedoms' and impacts on individuals see Niels van Dijk, Raphaël Gellert, and Kjetil Rommetveit, 'A Risk to a Right? Beyond Data Protection Risk Assessments', *Computer Law & Security Review* 32, no. 2 (1 April 2016): 304, https://doi.org/10.1016/j.clsr.2015.12.017.

[162] Hildebrandt, 'The Dawn of a Critical Transparency Right for the Profiling Era', passim.

[163] Edwards and Veale, 'Slave to the Algorithm?', 65.

[164] Tania Lombrozo, 'The structure and function of explanations', *Trends in Cognitive Sciences* 10 (10) , (2006) 464–470. See also Tim Miller, 'Explanation in Artificial Intelligence: Insights from the Social Sciences', *Artificial Intelligence* 267 (1 February 2019): 6, https://doi.org/10.1016/j.artint.2018.07.007, 273 who explains that explanation has two processes: cognitive process and social process.

Boni-Saenz, Alexander A. ‘Public-Private Partnerships and Insurance Regulation’. *Harvard Law Review* 121 (2008): 1367.
Brkan, Maja. ‘Do Algorithms Rule the World? Algorithmic Decision-Making and Data Protection in the Framework of the GDPR and Beyond’. *International Journal of Law and Information Technology*. Accessed 24 April 2019. https://doi.org/10.1093/ijlit/eay017.
Bygrave, Lee A. ‘Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling’. *Computer Law & Security Review* 17, no. 1 (1 January 2001): 17–24. https://doi.org/10.1016/S0267-3649(01)00104-2.
Casey, Bryan, Ashkon Farhangi, and Roland Vogl. ‘Rethinking Explainable Machines: The GDPR’s “Right to Explanation” Debate and the Rise of Algorithmic Audits in Enterprise’. *Berkeley Technology Law Journal* 34, no. 2019 (19 February 2018). https://papers.ssrn.com/abstract=3143325.
Charles Raab, and David Wright. ‘Surveillance: Extending the Limits of Privacy Impact Assessment. In: Wright David, De Hert Paul, Editors. Privacy Impact Assessment. Dordrecht’. In *Privacy Impact Assessment*, edited by David Wright and Paul De Hert, 363–83. Dordrecht: Springer, 2012.
Citron, Danielle. ‘Technological Due Process’. *Faculty Scholarship*, 30 April 2009. https://digitalcommons.law.umaryland.edu/fac_pubs/1012.
Citron, Danielle, and Frank Pasquale. ‘The Scored Society: Due Process for Automated Predictions’. *Faculty Scholarship*, 1 January 2014. https://digitalcommons.law.umaryland.edu/fac_pubs/1431.
Clarke, Roger. ‘An Evaluation of Privacy Impact Assessment Guidance Documents’. *International Data Privacy Law* 1, no. 2 (1 May 2011): 111–20. https://doi.org/10.1093/idpl/ipr002.
Council of Europe. ‘Guidelines on the Protection of Individuals with Regard to the Processing of Personal Data in a World of Big Data’. Strasbourg, 23 January 2017.
Crawford, Kate, and Jason Schultz. ‘Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms’. *Boston College Law Review* 55, no. 1 (29 January 2014): 93.
Desai, Deven R., and Joshua A. Kroll. ‘Trust But Verify: A Guide to Algorithms and the Law’. *Harvard Journal of Law & Technology*, 27 April 2017. https://papers.ssrn.com/abstract=2959472.
Dijk, Niels van, Raphaël Gellert, and Kjetil Rommetveit. ‘A Risk to a Right? Beyond Data Protection Risk Assessments’. *Computer Law & Security Review* 32, no. 2 (1 April 2016): 286–306. https://doi.org/10.1016/j.clsr.2015.12.017.
Edwards, L., D. McAuley, and L. Diver. ‘From Privacy Impact Assessment to Social Impact Assessment’. In *2016 IEEE Security and Privacy Workshops (SPW)*, 53–57, 2016. https://doi.org/10.1109/SPW.2016.19.
Edwards, Lilian, and Michael Veale. ‘Enslaving the Algorithm: From a “Right to an Explanation” to a “Right to Better Decisions”?’ *IEEE Security & Privacy* 16, no. 3 (2018): 46–54.
———. ‘Slave to the Algorithm? Why a “Right to an Explanation” Is Probably Not the Remedy You Are Looking For’. *Duke Law & Technology Review* 16, no. 1 (4 December 2017): 18–84.
Freeman, Jody. ‘The Private Role in the Public Governance’. *New York University Law Review* 75 (2000): 543.
Gellert, Raphael. ‘We Have Always Managed Risks in Data Protection Law: Understanding the Similarities and Differences between the Rights-Based and the Risk-Based Approaches to Data Protection’. *European Data Protection Law Review (EDPL)* 2 (2016): 481.
Goodman, Bryce. ‘A Step Towards Accountable Algorithms？: Algorithmic Discrimination and the European Union General Data Protection’, 2016.
Hildebrandt, M. ‘The Dawn of a Critical Transparency Right for the Profiling Era’. In *Digital Enlightenment Yearbook*, edited by J. Bus, 2012th ed., 41–56. Amsterdam : IOS Press, 2012. https://repository.ubn.ru.nl/handle/2066/94126.
Kaminski, Margot E. ‘Binary Governance: Lessons from the GDPR’s Approach to Algorithmic Accountability’. *Southern California Law Review* 92, no. 6 (2019). https://papers.ssrn.com/abstract=3351404.
———. ‘The Right to Explanation, Explained’. *Berkeley Technology Law Journal*, 34, no. 1 (2019). https://papers.ssrn.com/abstract=3196985.
Kim, Pauline. ‘Auditing Algorithms for Discrimination’. *University of Pennsylvania Law Review Online* 166, no. 1 (1 January 2017). https://scholarship.law.upenn.edu/penn_law_review_online/vol166/iss1/10.
Kloza, Dariusz, Niels VAN Dijk, Raphaël Gellert, István Böröcz, Alessia Tanas, Eugenio Mantovani, and Paul Quinn. ‘Data Protection Impact Assessments in the European Union: Complementing the New Legal Framework towards a More Robust Protection of Individuals’. D.Pia.Lab Policy Brief No. 1/2017, n.d. https://cris.vub.be/files/32009890/dpialab_pb2017_1_final.pdf.
Kroll, Joshua, Joanna Huey, Solon Barocas, Edward Felten, Joel Reidenberg, David Robinson, and Harlan Yu. ‘Accountable Algorithms’. *University of Pennsylvania Law Review* 165, no. 3 (1 January 2017): 633.
Lum, Kristian, and William Isaac. ‘To Predict and Serve?’ *Significance* 13, no. 5 (2016): 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x.
Malgieri, Gianclaudio. ‘Automated Decision-Making in the EU Member States: The Right to Explanation and Other “Suitable Safeguards” in the National Legislations’. *Computer Law & Security Review*, 9 July 2019, 105327. https://doi.org/10.1016/j.clsr.2019.05.002.
Malgieri, Gianclaudio, and Giovanni Comandé. ‘Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation’. *International Data Privacy Law* 7, no. 4 (1 November 2017): 243–65. https://doi.org/10.1093/idpl/ipx019.

Mantelero, Alessandro. 'AI and Big Data: A Blueprint for a Human Rights, Social and Ethical Impact Assessment'. *Computer Law & Security Review* 34, no. 4 (1 August 2018): 754–72. https://doi.org/10.1016/j.clsr.2018.05.017.
Reisman, Dillon, Jason Schultz, Kate Crawford, and Meredith Whittaker. 'Algorithm Impact Assessment: A Practical Frameworks for Public Agency Accountability'. AI Now Institute, n.d. https://ainowinstitute.org/aiareport2018.pdf.
Roig, Antoni. 'Safeguards for the Right Not to Be Subject to a Decision Based Solely on Automated Processing (Article 22 GDPR)'. *European Journal of Law and Technology* 8, no. 3 (21 January 2018). http://ejlt.org/article/view/570.
Sandvig, Christian, Kevin Hamilton, Karrie Karahalios, and Cédric Langbort. 'Auditing Algorithms : Research Methods for Detecting Discrimination on Internet Platforms', 2014.
Selbst, Andrew D. 'Disparate Impact in Big Data Policing'. *Georgia Law Review* 52, no. 1 (19 February 2018): 3373.
Selbst, Andrew D., Danah Boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 'Fairness and Abstraction in Sociotechnical Systems'. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 59–68. FAT* '19. New York, NY, USA: ACM, 2019. https://doi.org/10.1145/3287560.3287598.
Selbst, Andrew D., and Julia Powles. 'Meaningful Information and the Right to Explanation'. *International Data Privacy Law* 7, no. 4 (1 November 2017): 233–42. https://doi.org/10.1093/idpl/ipx022.
Vanclay, Frank, Ana Maria Esteves, Ilse Aucamp, Equispectives Research, and Daniel M Franks. 'Social Impact Assessment: Guidance for Assessing and Managing the Social Impacts of Projects'. International Association for Impact Assessment, April 2015. https://www.iaia.org/uploads/pdf/SIA_Guidance_Document_IAIA.pdf.
Veale, Michael, Reuben Binns, and Jef Ausloos. 'When Data Protection by Design and Data Subject Rights Clash'. *International Data Privacy Law* 8, no. 2 (1 May 2018): 105–23. https://doi.org/10.1093/idpl/ipy002.
Veale, Michael, and Lilian Edwards. 'Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling'. *Computer Law & Security Review* 34, no. 2 (1 April 2018): 398–404. https://doi.org/10.1016/j.clsr.2017.12.002.
Wachter, Sandra, and Brent Mittelstadt. 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI'. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, 13 September 2018. https://papers.ssrn.com/abstract=3248829.
Wachter, Sandra, Brent Mittelstadt, and Luciano Floridi. 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation'. *International Data Privacy Law* 7, no. 2 (1 May 2017): 76–99. https://doi.org/10.1093/idpl/ipx005.
Wachter, Sandra, Brent Mittelstadt, and Chris Russell. 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR'. *ArXiv:1711.00399 [Cs]*, 1 November 2017. http://arxiv.org/abs/1711.00399.
Walker, Simon. *The Future of Human Rights Impact Assessments of Trade Agreements*. School of Human Rights Research Series, v. 35. Antwerp ; Portland: Intersentia, 2009.
Wright, David, and Michael Friedewald. 'Integrating Privacy and Ethical Impact Assessments'. *Science and Public Policy* 40, no. 6 (1 December 2013): 755–66. https://doi.org/10.1093/scipol/sct083.
Wright, David, and Emilio Mordini. 'Privacy and Ethical Impact Assessment'. In *Privacy Impact Assessment*, edited by David Wright and Paul De Hert, 397–418. Law, Governance and Technology Series. Dordrecht: Springer Netherlands, 2012. https://doi.org/10.1007/978-94-007-2543-0_19.
Wright, David, and Charles D. Raab. 'Constructing a Surveillance Impact Assessment'. *Computer Law & Security Review* 28, no. 6 (December 2012): 613–26. https://doi.org/10.1016/j.clsr.2012.09.003.