Special Issue Contribution

ELLJ

European Labour Law Journal
2023, Vol. 14(2) 124–151

DOI: 10.1177/20319525231167299
journals.sagepub.com/home/ell

Regulating algorithmic management: A blueprint

Jeremias Adams-Prassl
Professor of Law, Magdalen College, University of Oxford, Oxford, UK

Halefom Abraha
Postdoctoral Researcher, Bonavero Institute of Human Rights, University of Oxford, Oxford, UK

Aislinn Kelly-Lyth
Researcher, Bonavero Institute of Human Rights, University of Oxford, Oxford, UK

Michael 'Six' Silberman
Postdoctoral Researcher, Bonavero Institute of Human Rights, University of Oxford, Oxford, UK

Sangh Rakshita
Researcher, Bonavero Institute of Human Rights, University of Oxford, Oxford, UK

Abstract

The promise—and perils—of algorithmic management are increasingly recognised in the literature. How should regulators respond to the automation of the full range of traditional employer functions, from hiring workers through to firing them? This article identifies two key regulatory gaps—an exacerbation of privacy harms and information asymmetries, and a loss of human agency—and sets out a series of policy options designed to address these novel harms. Redlines (prohibitions), purpose limitations, and individual as well as collective information rights are designed to protect against harmfully invasive data practices; provisions for human involvement 'in the loop' (banning fully automated terminations), 'after the loop' (a right to meaningful review), 'before the loop' (information and consultation rights) and 'above the loop' (impact assessments) aim to restore human agency in the deployment and governance of algorithmic management systems.

Keywords

Algorithmic management, privacy harm, information asymmetry, human agency, data protection, information access, human in the loop, impact assessments

Corresponding author:
Jeremias Adams-Prassl, Professor of Law, Magdalen College, University of Oxford, Oxford, UK.
E-mail: jeremias.adams-prassl@law.ox.ac.uk

I. Introduction

Digitalisation is revolutionising the world of work. Fears of widespread technological unemployment driven by the rise of artificial intelligence continue to prove unfounded—but this does not detract from the fundamental impact the deployment of emerging technologies has on the organisation of work. The rapid pace of technological innovation has set the stage for the rise of algorithmic management (ARM): the potential automation of the full range of traditional employer functions, from hiring workers and managing the day-to-day operation of the enterprise through to the termination of the employment relationship. Understood in this way, the 'rewiring of the firm' poses just as much a threat to the Coasian 'entrepreneur-coordinator' as to her workforce.[1]

The origins of ARM are closely linked to the advent of the gig economy:[2] rather than acting as mere matchmakers between consumers and workers, platforms rely on increasingly sophisticated automated systems to condition all aspects of service provision, from explicit control over routing and price-setting to more subtle 'nudges' directing worker behaviour.[3] Boosted by the Covid-19 pandemic, ARM systems are quickly becoming omnipresent, not least through integration into existing systems, from word processing to enterprise management at large.[4] Vendors promise solutions covering the full range of employer functions in workplaces across the socio-economic spectrum.

The promise—and perils—of ARM have been extensively documented in the literature.[5] Policymakers and regulators are starting to take note of the need to tackle the problems associated with the deployment of ARM systems, from algorithmic opacity to mental and physical harm.[6] They face a range of complex questions: To what extent can existing norms evolve to address these problems? Which harms are genuinely novel? And more fundamentally, how can we foster genuine innovation whilst also protecting fundamental rights at work?[7]

In this article, we set out a comprehensive blueprint of regulatory options in response to the rise of ARM. Discussion is structured as follows. The remainder of the introduction defines ARM, and briefly sketches the case for regulating ARM by identifying two regulatory gaps: the exacerbation of privacy harms and information asymmetries, and the loss of human

1. Jeremias Adams-Prassl, 'What If Your Boss Was an Algorithm? Economic Incentives, Legal Challenges, and the Rise of Artificial Intelligence at Work' (2019) 41 Comparative Labor Law and Policy Journal 123.
2. Jeremias Prassl, *Humans as a Service: The Promise and Perils of Work in the Gig Economy* (Oxford University Press 2018); Antonio Aloisi & Valerio De Stefano, *Your Boss Is an Algorithm: Artificial Intelligence, Platform Work and Labour* (Hart 2022).
3. Alex Rosenblat and Luke Stark, 'Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers' (2016) 10 International Journal of Communication 3758.
4. Wolfie Christl, *Digitale Überwachung und Kontrolle am Arbeitsplatz* (Cracked Labs 2021).
5. For early comprehensive accounts, see the special issue of the *Comparative Labor Law & Policy Journal* (2019) 41.
6. For recent initiatives in the EU and North America, see, for example, the provisions on algorithmic management (Chapter III) of the proposed 'Platform Work Directive' (Proposal for a Directive of the European Parliament and of the Council on improving working conditions in platform work, COM (2021) 762 final (9 December 2021)); Government of Canada Directive on Automated Decision-Making 2019; California Bill AB-1651 Workplace Technology Accountability Act 2022 (California Bill AB-1651); Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People October 2022.
7. European Commission, 'Communication from the Commission to the European Parliament, The Council, The European Economic and Social Committee and the Committee of the Regions: Fostering a European Approach to Artificial Intelligence' COM (2021) 205 final, 8.

(especially, but not only, managerial) agency. We then set out eight concrete policy measures designed to address these gaps and provide a rationale explaining each of the regulatory choices involved. Section 2 is focused on protection against privacy harms and overcoming information asymmetries, with options ranging from explicit prohibitions ('redlines') and purpose limitations to information and data access rights.

Section 3 explores a range of options to re-establish agency for management as well as workers and their representatives. Instead of focusing exclusively on bans on fully automated decision-making (the oft-discussed requirement for a 'human in the loop'), we propose a series of interventions across the life cycle of ARM, from design and deployment to operations and review. This includes a role for 'humans before the loop', establishing requirements for ARM design and deployment; 'humans after the loop', viz, rights for workers to challenge, request explanations for, and request human review of, decisions that affect them; and 'humans above the loop', monitoring the broader implications of ARM through dedicated impact assessments. A brief conclusion reflects on the individual and institutional capacity required to realise the blueprint's proposals.

DEFINING ALGORITHMIC MANAGEMENT ('ARM') SYSTEMS

For purposes of subsequent discussion, we define algorithmic management systems as encompassing:

- the collection or creation of any information (whether identifiable or not) with a view to organising, monitoring, supervising, or evaluating work performance or behaviour; and/or
- the use of that information to support, augment, or fully automate decisions that affect working conditions, including access to work, earnings, occupational safety and health, working time, promotion and contractual status, and disciplinary as well as termination procedures.

A plethora of policy documents, enforcement actions, and legislative initiatives aimed at regulating ARM have emerged across the globe.[8] Some proposals call for targeted new legislation; others envisage a careful recalibration of existing norms. This divergence raises a crucial point: what is the case for characterising ARM as a distinct phenomenon in labour market regulation?

Sophisticated automation has fundamentally revolutionised traditional ways of monitoring and controlling workers. First, in terms of the ubiquity, constancy, and comprehensiveness of surveillance, and the granularity and intimacy of the information acquired.[9] Second, through the processing of the data thus collected, including a range of statistical and machine learning techniques which promise to reveal information that would not be accessible when examining individual datapoints in

8. See, for example, White House Office of Science and Technology Policy, 'Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People' (October 2022); California Bill AB-1651; Council of the District of Columbia, Stop Discrimination by Algorithms Act of 2021; Algorithmic Accountability Act of 2022, HR 6580, 117th Cong (2022); New York City Local Law 144 (2021), Automated Employment Decision Tools Law 2021; Illinois Artificial Intelligence Video Interview Act, 820 Ill Comp Stat 42 (2020); Chinese Internet Information Service Algorithm Recommendations 2021; Government of Canada Directive on Automated Decision-Making 2019; Platform Work Directive; Spanish Organic Law 3/2018 on Protection of Personal Data and Guarantee of Digital Rights.

9. Aiha Nguyen, 'The Constant Boss: Work Under Digital Surveillance' (Data & Society, 20 May 2021) 3–15; Matthew T Bodie and others, 'The Law and Policy of People Analytics' (2017) 88 University of Colorado Law Review 961, 987–98; Sam Adler-Bell and Michelle Miller, 'The Datafication of Employment: How Surveillance and Capitalism Are Shaping Workers' Futures without Their Knowledge' (*The Century Foundation*, 19 December 2018).

isolation.[10] This creates an ability to draw inferences and create action-guiding predictions not just in relation to individuals surveilled, but also future employees.[11]

A recognition of the technical complexities underlying ARM systems is important in informing policy proposals: there is little point in calling for standards or approaches that are inherently infeasible. Novelty in and of itself, however, does not necessarily make out the case for targeted regulatory intervention: existing regimes will frequently be able to adapt to novel challenges. Take algorithmic bias as an example:[12] upon careful inspection, the hitherto relatively rarely invoked prohibition on direct discrimination in European and UK law may well be able to address a range of technical causes of algorithmic bias.[13]

At the same time, however, there are clear limits to the protective scope of existing norms: subsequent sections explore two central challenges that require new forms of regulatory intervention. For each category, we first set out evidence for the specific problem, before identifying a number of corresponding policy responses and evaluating their respective advantages and drawbacks. Implementation details will vary depending on specific contexts, but the overall case is clear: in order to ensure the socially responsible deployment of ARM systems, regulatory measures need to protect workers against exacerbated privacy harms and rebalance excessive information asymmetries, as well as restoring agency for management, workers, and their representatives.

II. Protection against privacy harms and information asymmetries

The quantity and intimacy of data collected about workers creates significant privacy harms, far beyond violations of existing legal rights to privacy and data protection. These can include physical harms, economic harms, reputational and other social-psychological harms, direct psychological harms, individual 'autonomy harms' such as coercion and manipulation, collective harms such as 'chilling effects,' and the erosion of work/non-work boundaries and work-life balance.[14]

Take automated scheduling and task assignment as an example: ARM practices often rely on the aggregation of data collected from a wide range of sources to create metrics and qualitative insights to develop 'optimised' or 'just-in-time' schedules—which will frequently involve work intensification[15] and unsustainable and unpredictable working patterns.[16] Furthermore, if workers are aware of the metrics used to evaluate them and to organise their work, they may be incentivised to behave in a way that optimises their performance on these metrics, rather than using their professional judgment to work appropriately. Unless the metrics are very carefully designed, this may lead to

10. Bodie and others (n 9) 998–999.
11. Salomé Viljoen, 'A Relational Theory of Data Governance' (2021) 131 Yale Law Journal 370; Daniel Solove and Danielle Keats Citron, 'Privacy Harms' (2021) GW Law Faculty Publications & Other Works 1534, 21–22 <https://scholarship.law.gwu.edu/faculty_publications/1534> accessed 22 February 2023.
12. Aislinn Kelly-Lyth, 'Challenging Biased Hiring Algorithms' (2021) 41 Oxford Journal of Legal Studies 899.
13. Jeremias Adams-Prassl, Reuben Binns and Aislinn Kelly-Lyth, 'Directly Discriminatory Algorithms' (2023) 86 Modern Law Review 144.
14. See generally Danielle Keats Citron and Daniel J Solove, 'Privacy Harms' (2022) 102 Boston University Law Review 793; Future of Privacy Forum, 'Identifying algorithmic harms when creating DPIAs: a quick guide' (2018).
15. Annette Bernhardt, Lisa Kresge and Reem Suleiman, 'Data and Algorithms at Work: The Case for Worker Technology Rights' (UC Berkeley Labor Center 2021) 15–16 <https://laborcenter.berkeley.edu/data-algorithms-at-work/> accessed 25 May 2022.
16. Daniel Schneider and Kristen Harknett, 'Hard Times: Routine Schedule Unpredictability and Material Hardship among Service Sector Workers' (2021) 99 Social Forces 1682.

'disregarding safety rules and procedures and undermining professional work standards and ethics.'[17] In cases where workers may be assigned more or fewer hours, more or less lucrative work, or even terminated based on their performance on automatically computed metrics, these systems may significantly reduce the scope for workers' professional judgment, potentially creating situations where individuals are pressured to work in unsafe, illegal, or unethical ways.[18]

ARM systems furthermore exacerbate existing information asymmetries.[19] Firm decision-making is informed by increasingly granular data about the preferences and behaviour of each individual, while workers know little about what information is being collected or how it is being used, making it difficult to bargain effectively over data use.[20] Extensive and granular monitoring also facilitates tacit knowledge extraction, potentially reducing job mobility through routinisation.[21] Employers may also use monitoring to identify the 'least attached' (i.e., most economically mobile) workers and pay them more—leading to less economically mobile workers receiving lower remuneration.[22]

* * *

There are two main ways in which the law seeks to protect individuals against privacy harms and 'information asymmetry [that] is insidious': rules can mandate greater 'transparency, so that everyone has access to the relevant information, or [restrict] the power of organizations to create a one-sided game.'[23] Illustrations of the latter approach include outright bans, or functional limitations, on data collection, processing, and the exercise of algorithmic control, whereas rebalancing strategies are based on sharing information with workers and their representatives.

In this section, we explore four policy options to implement these responses: we first set out a short list of clear redlines (prohibited practices) in ARM and argue for a restriction of the legal grounds on—and purposes for—which data can be processed. We then turn to transparency requirements for employers and rights of information access for individuals, as well as collective information and data access rights.

2.1 Policy option 1: Redlines/prohibitions

Information asymmetries exacerbate existing asymmetries of bargaining power. This is particularly problematic where information is collected beyond the contractually established scope of the managerial prerogative, and/or deployed for particularly harmful (or downright illegal) purposes. In order to stop the unnecessary or particularly harmful deployment of ARM systems, regulators

17. Justin Nogarede, 'No Digitalisation without Representation: An Analysis of Policies to Empower Labour in the Digital Workplace' (FEPS Policy Study 2021) 11.
18. Nogarede (n 17); Ulrich Leicht-Deobald and others, 'The Challenges of Algorithm-Based HR Decision-Making for Personal Integrity' (2019) 160 Journal of Business Ethics 377.
19. Abi Adams, 'Technology and the Labour Market: The Assessment' (2018) 34 Oxford Review of Economic Policy 349, 355.
20. For details, see Zoe Adams and Johanna Wenckebach (elsewhere in this issue) and Dan Calacci and Jake Stein (elsewhere in this issue).
21. Bernhardt, Kresge and Suleiman (n 15) 15–16.
22. Nathan Newman, 'Reeingeering Workplace Bargaining: How Big Data Drives Lower Wages and How Reframing Labor Law Can Restore Information Equality in the Workplace' 85 University of Cincinnati Law Review 693. Newman (702) cites Richard B Freeman and James L Medoff, *What Do Unions Do?* (Basic Books 1984).
23. Geoffrey Lightfoot and Tomasz Piotr Wisniewski, 'Information Asymmetry and Power in a Surveillance Society' (2014) 24 Information and Organization 214.

should prohibit data collection and processing practices that pose particularly severe risks to human dignity and fundamental rights.

Redlines/Prohibitions

1. Prohibit employer monitoring of workers in particular contexts. This could include:
   (a) Outside of work (temporally or geographically).
   (b) Physical and relational contexts 'in work' where data collection or monitoring poses risks to human dignity or the exercise of fundamental rights, including:
      (i) in private spaces such as bathrooms; and
      (ii) in private conversations or communications, especially including conversations and communications with worker representatives.
2. Prohibit monitoring and collection and processing of any data (personal or non-personal) for purposes that pose risks to human dignity and fundamental rights, including:
   (a) Emotional or psychological manipulation;
   (b) Prediction of, or persuasion against, the exercise of legal rights, especially including the right to organise; and/or
   (c) Other context-specific harmful purposes.

2.1.1 Rationale. This policy option identifies two categories of redlines: a prohibition of worker monitoring in certain places and contexts; and a prohibition of ARM practices for certain purposes.

2.1.1.1 Prohibition of monitoring in specific contexts. Whilst monitoring employees is an inherent part of the employer's prerogative, its permissible extent has long been amongst the most common and controversial regulatory problems at work.[24] ARM exacerbates these controversies as new technologies enable employers to monitor employees at all times, across workplaces, and through many different systems.[25] This constant monitoring poses significant risks to the privacy and autonomy of workers: algorithmic planning, task allocation, and scheduling tools such as Preactor, for example, can severely limit workers' autonomy in the selection and ordering of their day-to-day tasks.[26]

This policy option places clear limits on the monitoring of workers outside working hours, including time on breaks or off-duty, to address the risks that arise due to the increasingly blurred boundary between the workplace and private life.[27]

24. See, for instance, *Bărbulescu v Romania* [2017] ECtHR 61496/08.
25. Article 29 Working Party, Opinion 2/2017 on data processing at work (adopted 8 June 2017) WP 249, 9.
26. Patrick Briône, 'Mind over Machines: New Technology and Employment Relations' (*Acas*, 14 June 2017) <https://www.acas.org.uk/mind-over-machines-new-technology-and-employment-relations/html> accessed 28 June 2022; Alex J Wood, 'Algorithmic Management: Consequences for Work Organisation and Working Conditions' (Joint Research Centre Working Papers on Labour, Education and Technology 2021-07, 2021). 2021–07 <https://ideas.repec.org/p/ipt/laedte/202107.html> accessed 4 March 2022.
27. For details on the harms of workplace monitoring, see Nogarede (n 17); Eurofound, 'Employee Monitoring and Surveillance: The Challenges of Digitalisation' (Publications Office of the European Union, Luxembourg, 2020); Bernhardt, Kresge, and Suleiman (n 15). The significant privacy and autonomy risks that arise from monitoring workers outside working hours have prompted some Member States to introduce new digital rights for employees

Closely related is a prohibition of the monitoring of private (physical and virtual) spaces in the workplace under all circumstances, including during working hours. In its 2017 Opinion, the Article 29 Working Party stated that employees must have 'certain private spaces to which the employer may not gain access under any circumstances.'[28] Several Member States follow the same approach and make it illegal for employers to monitor private zones such as bathrooms, changing rooms, rest areas, and bedrooms under all circumstances.[29] The protection of private space in the workplace includes private communications that are not intrinsically connected to work. Additionally, the monitoring of conversations with workers' representatives should be prohibited under all circumstances. Such a prohibitive approach is consistent with the proposed EU Platform Work Directive and other emerging regulatory initiatives.[30]

2.1.1.2 Prohibition of algorithmic management for high-risk purposes. The second category of redlines concerns ARM practices which pose particular risks to the human dignity, legitimate interests, and fundamental rights of workers.[31] This is particularly important where ARM systems are extremely intrusive, high-risk, or susceptible to high error rates.[32]

Emerging regulatory frameworks such as the Platform Work Directive and California's Workplace Technology Accountability Bill prohibit the processing of certain personal data whose processing poses high risks to the privacy of workers. The Platform Work Directive, in particular, addresses these risks by prohibiting the processing of certain types of personal data of platform workers, including any personal data on the emotional or psychological state of the platform worker.[33] California's Workplace Technology Accountability Bill also prohibits the use of ARM to make predictions about a worker's behaviour that are unrelated to the worker's essential job functions; to make predictions about a worker's emotions, personality, or other types of sentiments; and to identify, profile, or predict the likelihood of workers exercising their legal rights.[34]

---

such as the right to disconnect: see, e.g., Spanish Organic Law 3/2018 on Protection of Personal Data and Guarantee of Digital Rights, art 88.

28. Article 29 Working Party, Opinion 2/2017 on data processing at work (n 25).
29. The Member States that prohibit monitoring private spaces in the workplace include Germany, Croatia, Finland, Bulgaria, France, and Cyprus. For details on this, see Eurofound, 'Employee Monitoring and Surveillance' (n 27).
30. Article 6(5)(c) of the Platform Work Directive prohibits the processing of any personal data in relation to private conversations, including exchanges with platform workers' representatives. Similarly, section 1543(b) of California Bill AB-1651 prohibits, among other things, the monitoring of workers in order to identify, profile, or predict the likelihood of workers exercising their legal rights.
31. The proposed legal framework itself could rely on the wording of 'purposes that pose particular risks to human dignity and fundamental rights', with specific banned practices set out in an amendable Annex to ensure coverage of emerging harmful practices.
32. Bernhardt, Kresge and Suleiman (n 15); Lisa Feldman Barrett and others, 'Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements' (2019) 20 Psychological Science in the Public Interest 1. See also Valerio De Stefano and Mathias Wouters, 'AI and Digital Tools in Workplace Management and Evaluation: An Assessment of the EU's Legal Framework' (European Parliamentary Research Service, PE 729516, May 2022).
33. Platform Work Directive, art 6(5)(a).
34. California Bill AB-1651, s 1553.

Instead of focusing on particular categories of data, however, the present policy option establishes protections by prohibiting the collection of worker data for purposes that pose serious risks to human dignity and fundamental rights, especially the prediction of, or persuasion against, the exercise of their legal rights.[35] This crucially includes the right to organise. In the United States, for example, Whole Foods has relied on ARM systems to predict union activity in hundreds of stores and assign a unionisation 'risk score'.[36]

Such 'relational harms', in which data collected about one individual enables a potentially inaccurate or harmful inference to be drawn about another individual, are not necessarily covered by existing data protection rules.[37] While the full protections provided for 'special categories of personal data' under Article 9 of the General Data Protection Regulation (GDPR) would apply in the case of an employer attempting to infer current trade union membership, they might, for example, not apply to an employer trying to predict the likelihood of future organising activities. More challenging still are decisions taken at the plant or establishment level, such as deciding to close an entire store based on its 'unionisation risk.'[38] Such a decision might be based on aggregated and anonymised data, in which case the GDPR would not apply at all. Such difficulties could be overcome by prohibiting particularly egregious relational data harms because of their purpose, regardless of whether the data relied upon are personal or not.

2.2 Policy option 2: Specific and limited legal bases

Existing legal bases for the processing of workers' data in EU law, especially 'legitimate interest' and 'consent,' may legitimise the indiscriminate use of ARM systems, even when their use poses serious risks to human dignity or fundamental rights, or when their use is not proportionate or relevant to legitimate managerial aims. In combination with the absence of an explicit proportionality requirement, existing norms furthermore do not impose a clear obligation on employers to ensure that ARM systems are capable of serving their intended purpose (i.e., are technically 'valid').[39] This opens the door to the deployment of technically or organisationally ineffective or inappropriate systems, which can subject workers to serious harms.

There are several ways in which regulators could restrict the inappropriate and disproportionate uses of ARM and/or prohibit the use of systems that do not achieve their stated purpose. This includes a restriction of the available legal bases for data processing in the context of ARM, and a requirement that the use of ARM systems be proportionate, i.e., that ARM systems must be suitable to achieve the purpose for which they are being deployed, in the least intrusive manner possible.

35. Zoe Adams, Abi Adams-Prassl and Jeremias Adams-Prassl, 'Online Tribunal Judgments and the Limits of Open Justice' (2021) 41 Legal Studies 42.
36. Harmon Leon, 'Whole Foods Secretly Upgrades Tech to Target and Squash Unionizing Efforts' (*Observer*, 24 April 2020) <https://observer.com/2020/04/amazon-whole-foods-anti-union-technology-heat-map/> accessed 24 June 2022.
37. Salomé Viljoen, 'A Relational Theory of Data Governance' (2021) 131 Yale Law Journal 573.
38. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (GDPR), art 4(1).
39. For examples of 'invalid' but widely used ARM systems, see Alene Rhea and others, 'Resume Format, LinkedIn URLs and Other Unexpected Influences on AI Personality Prediction in Hiring: Results of an Audit', Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society < https://dl.acm.org/doi/10.1145/3514094.3534189>; more generally, see Frederike Kaltheuner (ed), *Fake AI* (Meatspace Press 2021).

Specific And Limited Legal Bases

1. Establish that the deployment and operation of an algorithmic management system shall be lawful only if:
   (a) The deployment and operation of the system meets at least one of the following requirements:
      (i) It is intrinsically connected to and strictly necessary for the performance of the contract of employment or in order to take steps which are strictly necessary for entering into a contract of employment;
      (ii) It is necessary for compliance with a legal obligation to which the employer is subject; or
      (iii) It is necessary in order to protect the vital interests of the worker or of another natural person.
   (b) The algorithmic management system is capable of achieving these goals in a proportionate manner.

2.2.1 Rationale. The proposed Platform Work Directive suggests that bar 'contractual necessity', the wide range of legal bases provided under Article 6 GDPR are not appropriate to justify ARM.[40] While the Platform Work Directive is a step in the right direction, it is also important to note that there are other grounds, such as 'legal obligation', that may be appropriate bases for deploying ARM in the employment context.

The purpose of this policy option is, therefore, to narrow the legal basis for ARM. Specifically, 'consent', 'legitimate interest', or 'public interest'[41] should not constitute valid bases for deploying ARM tools. Other processing of personal data would not be affected.

The effectiveness of consent, first, hinges on two fundamental assumptions: that the employee knows or is properly informed about what she is consenting to, and that the employee is in a position to negotiate and freely give, refuse, or revoke consent. The nature of employment relations and the increasing digitalisation of work challenges both assumptions at their core. The notion of 'freely given, specific, informed and unambiguous' consent within the meaning of Article 4(11) GDPR is largely theoretical in the employment context, given the inherent imbalance of power between worker and employer: 'employees are almost never in a position to freely give, refuse or revoke consent.'[42] Employees furthermore often 'lack a clear understanding of the extent of the data collected, of the technical functioning of the processing and therefore of what they are consenting to.'[43] The deployment of opaque and sophisticated ARM tools further obscures potential harms and limits employees' ability to make informed decisions.[44] Despite the widespread agreement on the invalidity of consent as a lawful basis in the employment context, repeated efforts to

40. See Platform Work Directive, art 6(5).
41. See Art. 6(1)(a), Art. 6(1)(f), and Art. 6(1)(e), respectively.
42. For instance, Article 29 Working Party, Opinion 2/2017 on data processing at work (n 25); European Data Protection Board, 'Guidelines 05/2020 on Consent under Regulation 2016/679 Version 1.1' (4 May 2020); Ifeoma Ajunwa, Kate Crawford and Jason Schultz, 'Limitless Worker Surveillance' (2017) 105 California Law Review 735.
43. Eurofound, 'Employee Monitoring and Surveillance' (n 27) 7.
44. On how algorithmic management in the workplace further shifts the existing power dynamics of employment relationships, see European Commission, 'Study to Support the Impact Assessment of an EU Initiative to Improve the Working Conditions in Platform Work' (22 October 2021) 68 <https://ec.europa.eu/social/main.jsp?catId=738&langId=en&pubId=8428&furtherPubs=yes> accessed 4 May 2022.

exclude consent as a ground for lawful processing of workers' personal data are not reflected in the GDPR.[45] Policy option 2 addresses this legal ambivalence.

Second, the proposed policy option excludes 'legitimate interest'. Including this broad basis could render the list nugatory, because almost all systems will be deployed for some legitimate interest. As the Article 29 Working Party has noted, the 'legitimate interest' ground is flexible and open-ended, which 'leaves much room for interpretation and has sometimes . . . led to lack of predictability and lack of legal certainty.'[46] In principle, employers will usually have an effectively unlimited legitimate interest in the 'smooth operation of the business', which could be used to justify data processing occasioning significant privacy or autonomy harms, or even violations of fundamental rights such as the right to organise. Permitting reliance on this range of ostensibly legitimate interests would make the line between acceptable and unacceptable deployment of ARM systems much harder to draw in practice, as well as denying workers the possibility to exercise their right to data portability.[47]

Third, the proposed policy option also excludes public interest as a means to legitimise the deployment and use of ARM in the employment context. This is because there is no relevant public interest consideration to deploy ARM systems. Public interest, as provided under Article 6(1)(e) GDPR is the general basis for the lawful processing of personal data for public sector purposes, and employers cannot use this legal basis to use ARM systems in the employment context.

In response to concerns about low-accuracy, invalid, or inappropriate systems, [48] the proposed policy option provides two further safeguards: first, a requirement that the system must be capable of achieving the goal related to the legal basis of the data processing. That is, the system must be valid for the purpose it is being introduced for. This excludes the deployment of systems for purposes they are not in fact capable of achieving (such as evaluating job candidates based on 30 seconds of video).[49] Second, a requirement that the system must be capable of achieving the goal in a proportionate manner; that is, the potential risks or threats to fundamental rights (e.g., privacy or autonomy harms) posed by a system must not be disproportionately large given the employer's purpose in its deployment.

45. See, for instance, European Commission, 'Second Stage Consultation of Social Partners on the Protection of Workers' Personal Data' (2002); European Commission, 'Staff Working Document: Impact Assessment Accompanying the Document Regulation of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data (General Data Protection Regulation) and Directive of the European Parliament and of the Council on the Protection of Individuals with Regard to the Processing of Personal Data by Competent Authorities for the Purposes of Prevention, Investigation, Detection or Prosecution of Criminal Offences or the Execution of Criminal Penalties, and the Free Movement of Such Data SEC (2012) 72 Final' 50, though cf Art. 88 read together with rec. 155, recital 55, and GDPR, Art. 9(2)a.
46. Article 29 Working Party Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC (9 April 2014) WP217.
47. GDPR, Art. 20. According to this provision, workers do not have the right to data portability if the processing is based on legitimate interest pursuant to point (f) of Article 6(1).
48. Arvind Narayanan, 'How to Recognize AI Snake Oil' (Center for Information Technology Policy, Princeton University) <https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf> accessed 22 February 2023; Bernhardt, Kresge and Suleiman (n 15) (recommending that employers should be prohibited from using algorithms that incorporate unproven, questionable, or particularly invasive technologies); Feldman Barrett and others (n 32); De Stefano and Wouters (n 32).
49. For details on this, see Narayanan (n 48); Manish Raghavan and others, 'Mitigating Bias in Algorithmic Hiring: Evaluating Claims and Practices' (*Arxiv.com*, 2020) <https://arxiv.org/abs/1906.09208> accessed 22 February 2023; Rhea and others (n 39).

2.3 Policy option 3: Individual notice obligations

Workers are frequently unaware of the existence of monitoring, data collection, and decision-making systems, or not fully informed of the operation of ARM. This problem can be addressed with an obligation on employers to make their workforce aware of the existence, activities, and purposes of automated monitoring, data collection, and decision-making systems, without, however, overloading individuals with information. The key goal is to ensure that employers provide workers with meaningful and relevant information.

Individual Notice Obligations

1. Establish obligations for employers to notify affected individuals of algorithmic management systems. Specifically:
   - (a) Information regarding the use of algorithmic management systems shall be provided by the employer to all affected individuals at three points in time in relation to the employment relationship:
     - (i) at the earliest technically feasible time during the employment application process;
     - (ii) at the point at which a contract of employment is offered to a prospective employee; and
     - (iii) at regular intervals throughout the duration of the employment relationship (at least once a year), in the event of a change in the risk posed by the system, and at any time upon request.
   - (b) The information to be provided could include:
     - (i) the existence of any algorithmic systems used in the process of monitoring, evaluating, or managing individuals or work, including, but not limited to, fully automated decision-making systems and scoring or evaluation systems whose outputs are used by human decision-makers;
     - (ii) the nature, purpose, and scope of the systems used, including the specific decisions and categories of decisions they take or support (such as selection, recruitment, assignment of tasks, productivity control, promotions);
     - (iii) all inputs, criteria, variables, correlations, and parameters used by the systems in producing those outputs;
     - (iv) the logic used by the systems to produce their outputs, including, but not limited to, weightings of different inputs and parameters;
     - (v) the outputs produced by the systems (e.g., decisions, recommendations, scores);
     - (vi) the consequences that the decisions taken or assisted by the algorithmic management systems may have on the individual;
     - (vii) the existence and extent of human involvement in decision-making processes involving the systems, and the competence, authority, and accountability of the human persons involved;
     - (viii) if the systems are provided by or sourced from a third party (e.g., a software vendor or an open-source software package), or operated by a third party, the name of the third party and the name or common description of the software;
     - (ix) information about individuals' rights to receive information about the systems and decisions (or other outputs produced by those systems) affecting them, to request human review of the decisions or other outputs, and to contest the decisions or other outputs; and information about how to exercise those rights;

(x) any other available avenues for recourse, such as rights to engage with relevant competent authorities (such as the data protection officer, worker representatives, data protection authority, labour body, or equality body) or to judicial remedy;
(xi) contact information for the relevant competent authorities.

2. The notice shall be concise, transparent, and intelligible, using clear and plain language, and made available in an easily and continuously accessible electronic format.

2.3.1 Rationale. Transparency has become an organising principle of AI regulation. The EU's Ethics Guidelines for Trustworthy Artificial Intelligence, for instance, identify the principle of transparency as 'a crucial component of achieving Trustworthy AI'.[50] The scale and granularity of data collection coupled with the opacity of ARM tools lead to the risk of workers' personal data being processed, and consequential decisions being made, unknown to the workers themselves.[51] A lack of transparency in ARM systems furthermore obscures responsibility and complicates attribution, making it difficult for workers to exercise their rights.[52]

As Mateescu and Nguyen have pointed out, 'algorithmic management can create power imbalances that may be difficult to challenge without access to how these systems work as well as the resources and expertise to adequately assess them.'[53] Countering this risk requires both a strengthening, updating, and particularising of existing tools, as well as new worker rights and protections. Algorithmic transparency requirements would disincentivise employers from developing and deploying opaque systems and ensure accurate responsibility attribution in ARM.

Policy option 3 establishes transparency obligations at three stages of the employment relationship: at the job application stage, once the employment contract is offered, and during the employment relationship. During the employment relationship, the policy would require employers to provide workers with detailed and meaningful information about the data collected, and the existence of the ARM system and its functionalities at least once a year, at any time where there is a change in the risk posed by the system, and at any time upon request by the individual worker.

The GDPR already provides a list of transparency requirements in the form of information and access rights at the individual level. Workers can leverage these rights to counterbalance information asymmetry, exercise their rights, and voice their concerns. Furthermore, adequate information and access rights can serve as 'organizing and power-building tools' for workers.[54] However, the transparency requirements specifically applicable to ARM are limited in scope and detail.[55] For instance, the two significant safeguards provided under Article 15(1)(h) GDPR do not apply to semi-automated decision making. Unless a decision is fully automated, workers do not have (i) the right to know the existence of the decision-making system, or (ii) the right to obtain

50. High-Level Expert Group on Artificial Intelligence, 'Ethics Guidelines for Trustworthy AI' (2019) <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai> accessed 22 February 2023.
51. Robin Allen and Dee Masters, 'Technology Managing People—The Legal Implications' (Trades Union Congress 2021).
52. Maranke Wieringa, 'What to Account for When Accounting for Algorithms: A Systematic Literature Review on Algorithmic Accountability', Proceedings of the 2020 ACM Conference on Fairness, Accountability, and Transparency, 2 <https://doi.org/10.1145/3351095.3372833> accessed 22 February 2023.
53. Alexandra Mateescu and Aiha Nguyen, 'Algorithmic Management in the Workplace' (Data & Society, February 2019).
54. Calacci and Stein (elsewhere in this issue).
55. GDPR, Arts. 13(2)(f), 14(2)(g), and 15(1)(h), in conjunction with Art. 22 and Art. 5.

meaningful information about the logic, significance, and consequences of the decision-making system.[56]

The proposed information and transparency obligation would expand the scope and detail of the information and access rights provided under existing law. While Article 15(4) GDPR provides that data subjects have the right to receive a copy of their personal data as long as it does 'not adversely affect the rights and freedoms of others', this provision has been used to refuse to provide information to subjects of ARM.[57] Furthermore, the transparency requirements provided under the GDPR are limited to information about the existence of decision-making systems, and information about the logic, significance and consequences of the systems.[58] The GDPR does not, on the other hand, require controllers to provide other information that may be decisive for workers' ability to understand how decisions were arrived at, how to ensure that the decisions were appropriate, and what access to recourse they have if they were not.[59] Policy option 3 addresses this gap by clarifying and extending the information and transparency obligations of the GDPR.

The policy option also extends the scope of the algorithmic transparency requirements of the Platform Work Directive. In addition to the algorithmic transparency requirements under Article 6 of the Platform Work Directive, it also requires employers to provide affected workers with information about all inputs, criteria, variables, correlations, and parameters used by the systems in producing those outputs; the existence and extent of human involvement; and any third parties involved in developing or operating the system. Policy option 3 also clarifies the manner and means by which the information listed above should be provided (i.e., the notice should be concise, transparent, and intelligible, using clear and plain language, and made available in an easily and continuously accessible electronic format). Beyond the substantive content of the information to be provided to workers, it also requires employers to inform workers of their rights and available avenues for recourse, in line with emerging initiatives such as California's Workplace Technology Accountability Bill (AB-1651), which requires similar transparency and notice obligations.

Given the additional scope and detail of transparency obligations being proposed here, 'information overload' for workers is a concern.[60] Crucially, workers may be overloaded with information that is not meaningful. Meaningful information in the context of data protection law is information that (i) makes data subjects aware of the existence of the data processing; (ii) enables them to verify the lawfulness of that processing; and (iii) enables them to exercise their rights in relation to the processing.[61] To ensure

56. Some data protection authorities, such as the Austrian DPA, are of the opinion that the specific transparency obligations under Article 15(1)(h) are not limited to solely automated decisions but encompass other automated decisions, even if they do not meet the high threshold established by Article 22. For details on this, see Halefom Abraha (elsewhere in this issue). See also Sebastião Barros Vale and Gabriela Zanfir-Fortuna, 'Automated Decision-Making Under the GDPR: Practical Cases from Courts and Data Protection Authorities' (Future of Privacy Forum 2022).
57. Michael 'Six' Silberman and Hannah Johnston, 'Using GDPR to Improve Legal Clarity and Working Conditions on Digital Labour Platforms' (2020) European Trade Union Institute Working Paper 2020.05, 12, 24.
58. See GDPR, Arts. 13(2)(f), 14(2)(g), and 15(1)(h).
59. For more detail, see Calacci and Stein (elsewhere in this issue), arguing that a central purpose of data access rights in the workplace context is to allow workers to understand what decisions are being made, how they are made, and how they affect workers.
60. See, e.g., Yannic Meier, Johanna Schäwel, and Nicole C Krämer, 'The Shorter the Better? Effects of Privacy Policy Length on Online Privacy Decision-making' (2020) 8 Media and Communication 291.
61. Abraha (elsewhere in this issue).

that notices do not become too long or too overburdened with 'meaningless' information, policymakers could impose requirements on the type, clarity, and amount of information to be included.[62]

2.4 Policy option 4: Collective notice obligations and data access

One of the core tasks of worker representative bodies is to counterbalance employers' prerogatives and address collective risks and harms through social dialogue. Given the complexity and opacity of ARM systems, however, worker representatives will only be able to fulfil these roles with clear, concise, relevant, and timely information concerning the planned deployment, intended uses, anticipated effects, and operation of ARM systems.

There are multiple ways in which this goal could be achieved, including an obligation on employers using ARM systems to notify worker representatives of the existence and functioning of those systems, or a right of access for worker representatives to all individual-level data collected, used, processed, or created by ARM systems (provided that the individuals involved consent), including in machine-readable format via automatic data transfer. Finally, a right could be created for worker representatives to bring collective litigation and file collective complaints on behalf of groups of workers. Providing data access rights to worker representatives will also help mitigate information overload at the individual level.

Collective Notice and Data Access

1. Establish obligations for employers to notify worker representatives regarding algorithmic management systems. Specifically, establish that:
   (a) The employer shall, on an ongoing basis, provide to the worker representatives all system level information to be provided to individual workers and applicants, including all information regarding algorithmic management systems detailed in Policy option 3.
   (b) Should the worker representatives request additional information within the scope of content set out in policy option 3, but beyond that provided in the notice provided to individuals, the employer shall provide the requested information within a period of time not to exceed one month.
2. Establish a right of access for worker representatives to individual-level data collected, used, processed, or created by algorithmic management systems, provided that the individuals to whom the data relate consent. Additionally, establish that:
   (a) Worker representatives shall further have the right to receive such data on an ongoing basis in a machine-readable format, including, on the worker representatives' request, through a secure electronic data transfer procedure.
3. Establish a right for worker representatives to bring collective litigation against the employer and submit collective complaints to relevant data protection authorities on behalf of groups of workers.

To take account of diverse industrial relations traditions, the definition of 'worker representatives' should be provided for by national laws and/or practices.

62. See, e.g., the discussion of 'main parameters' in European Commission, 'Commission Notice: Guidelines on ranking transparency pursuant to Regulation (EU) 2019/1150 of the European Parliament and of the Council' (2020/C 424/01), 8–18.

2.4.1 Rationale. Individual data access rights are not enough to mitigate the exacerbation of existing information asymmetries and power imbalances occasioned by the introduction of ARM, for at least three reasons. First, ARM may produce inaccurate or unlawful decisions or other outputs. A systematic pattern of such errors, however, will not be immediately cognisable by individuals as erroneous; it will become visible only in the aggregate.[63] This makes a collective right of access to aggregate data essential.

Second, in some cases, the harms suffered by individuals may be relatively small compared to the time and effort cost of exercising data access rights to investigate and challenge possibly erroneous decisions,[64] or the harm may be large, but individuals may not realise that they have suffered a harm (e.g., when a job applicant is not offered an interview). In such cases, only a collective or representative body with an overarching mandate of ensuring the lawfulness and accuracy of ARM systems and broad rights of access to information about, and data used by, those systems, is positioned and incentivised to intervene to correct or prevent the harms.

Third, direct access to information and data about what is happening in the workplace is crucial to the effective performance of worker representatives' 'protective' function.[65] Consider, for example, individual pay negotiations: over time, this may lead to a gender pay gap. An employer may not notice this, or may accept it as regrettable but struggle to take corrective action.[66] Worker representatives with a mandate to be alert for discrimination risks, however, are likely to notice it if they have access to data on all worker pay, and are likely to be well positioned to negotiate

63. See generally Susan Sturm, 'Second Generation Employment Discrimination: A Structural Approach' (2001) 101 Columbia Law Review 458, 460, 470–71.
64. On one view, considering algorithmic management as a new phenomenon, this scenario offers a new case of the classic collective action problem (Mancur Olson, *The Logic of Collective Action: Public Goods and the Theory of Groups* (Harvard University Press 1971); for later, more sophisticated elaborations of the problem, see, e.g., Dan M Kahan, 'The Logic of Reciprocity: Trust, Collective Action, and Law' (2003) 102 Michigan Law Review 71 and Elinor Ostrom and James Walker (eds), *Trust and Reciprocity: Interdisciplinary Lessons for Experimental Research* (Russell Sage Foundation 2005)). Considering algorithmic management as merely new means in achieving the same old ends of greater managerial control over a workforce, however, this scenario can be understood as a continuation by new means of an old strategy, namely, divide and conquer. The regulatory solution is therefore also the same: to facilitate the aggregation of information across potentially affected individuals and empower a representative to act in their collective interests. See further, Calacci and Stein (elsewhere in this issue) and Adams and Wenckebach (elsewhere in this issue).
65. See, e.g., discussing in detail the data needs of employee representatives in the German industrial relations context, Michael Grauvogel, Detlef Hase, and Dietmar Röhricht, *Wirtschaftsausschuss und Betriebsrat: Praxiswissen für Betriebsräte* (Forba, Luchterhand Verlag 1996). A summary document may be found at <https://tinyurl.com/48z8xz5b> accessed 22 February 2023. 'It often happens [when employee representatives request these data] that the employer implies that it has no specific personnel plans, and no explicit personnel budget. How can this be verified? The employer refuses to provide information about planned investments; documents regarding finances are not provided [to the employee representatives]. What can be done?' (Summary document, 2; authors' translation). For considerations of worker uses of data in the data-driven workplace in the US context, see, e.g., Vera Khovanskaya and Phoebe Sengers, 'Data Rhetoric and Uneasy Alliances: Data Advocacy in US Labor History,' Proceedings of the 2019 ACM Conference on Designing Interactive Systems, 1391–1403; and Vera Khovanskaya, Phoebe Sengers, and Lynn S Dombrowski, 'Bottom-up Organizing with Tools from on High: Understanding the Data Practices of Labor Organizers,' Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.
66. See, e.g., the case based on empirical research and described in Sturm (n 63) 469–72, especially 470: 'In part because the firm aggressively recruits women at the entry level and fails to track patterns in work assignment and promotion, the firm's management has been largely unaware of any problem [relating to biased promotion patterns] until . . . complaints arose.'

corrective action with the employer. No individual worker is likely to have access to enough data to identify this pattern. Additionally, if the representatives do not have an explicit right of access to this data, assembling this data will require collection from individual workers—an unnecessary burden that impairs the ability of the representatives to fulfil their mandate.[67]

Worker representatives should therefore have the right to initiate collective litigation or file collective complaints to a data protection authority so that system level problems can be resolved at the system level, rather than leaving it to individual workers to perceive and litigate systematically biased or erroneous patterns of decision-making that harm entire groups of workers. Requiring individuals to file litigation or complaints individually (a) imposes unnecessary administrative burdens on all parties, including both the competent authority (i.e., the courts or data protection authority) and the employer, and (b) creates a risk that the root cause of the problem is not surfaced and addressed. For example, an employer may choose to settle litigation with individual employees rather than address the problem that occasioned the litigation in the first place.[68]

III. Restoring human agency

Most firms deploying ARM systems rely on third-party software. This can pose significant challenges for management's ability fully to control—and in some cases even to understand—the systems which take or guide key managerial decisions.[69] This loss of managerial agency has significant implications for work quality, as the 'rhetoric of algorithmic management can distance companies from the effects of their business decisions.'[70] Human agency is central to flexibility and empathy; taking managerial decisions without human agency risks alienation and violations

67. Litigation in the platform economy, in particular against ride-hail platforms Uber and Ola, demonstrates the challenges that workers' representatives face when collecting and aggregating data from individual workers. Uber refused data to workers who filed Article 15 GDPR data subject access requests, on the grounds that the workers were making the requests for the purpose of aggregating the data to gain insights for negotiating over working conditions, and that this was not a valid use of workers' Article 15 GDPR rights. The court disagreed, indicating that the intended use of the data had no bearing on the platforms' obligation to provide it. See *Rechtbank Amsterdam*, Case C/13/687315 HARK 20-207, ECLI:NL:RBAMS:2021:1020 (11 March 2021) <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1020> accessed 13 December 2022 and *Rechtbank Amsterdam*, Case C/13/689705 HARK 20-258, ECLI:NL:RBAMS:2021:1019 (11 March 2020) <https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2021:1019> accessed 13 December 2022. For unofficial English translations, see <https://ekker.-legal/en/2021/03/13/dutch-court-rules-on-data-transparency-for-uber-and-ola-drivers/> accessed 15 December 2022.
68. See, e.g., Steven Shavell, 'The Level of Litigation: Private versus Social Optimality of Suit and of Settlement' (1999) 19 International Review of Law and Economics 99. The author argues that while generally settlement prior to trial is 'socially [more] desirable' than going to trial due to the costs of litigation, 'a trial may be socially beneficial' if 'the cause of harm will be unclear unless it is resolved at trial' (see p 112, n 41). Relatedly, the fundamental question at issue may not be legally resolved; in *Otey v CrowdFlower*, for example, an early case in California revolving around the alleged misclassification of ostensibly self-employed platform workers, many stakeholders and policymakers eagerly awaited the judgment—but the case was settled in 2015 and the legal question was not resolved. Had it been even partially resolved, some of the significant legal and political energies expended later in California around the topic of false self-employment of platform workers could perhaps have been put to some other purpose.
69. Geoffrey Lightfoot and Tomasz Piotr Wisniewski, 'Information Asymmetry and Power in a Surveillance Society' (2014) 24 Information and Organization 214; Corentin Curchod and others, 'Working for an Algorithm: Power Asymmetries and Agency in Online Work Settings' (2020) 65 Administrative Science Quarterly 644; Calacci and Stein (elsewhere in this issue).
70. Alexandra Mateescu and Aiha Nguyen, 'Explainer: Algorithmic Management in the Workplace' (Data & Society 2019) 14 <https://datasociety.net/library/explainer-algorithmic-management-in-the-workplace/> accessed 25 May 2022.

of workers' individual dignity. The 'erosion of the personal relationship changes the role and dynamics between the worker and the corporation', signalling a 'deeper shift away from a sense of care and responsibility'.[71]

The newly intermediated control over workers which results from ARM is, furthermore, fundamentally distinct from prevailing bureaucratic structures which rely on individuals to take management decisions in the workplace.[72] This poses a challenge to the operation of legal accountability mechanisms, which are structured around the human exercise of managerial prerogatives, by obfuscating the location of responsibility for decisions made or supported by algorithms. This diffusion of responsibility for managerial decisions into the supply chain—specifically, into the vendor firms operating ARM software—threatens the operation of regulatory systems such as employment law, designed to condition managerial agency through a combination of *ex ante* incentives and *ex post* liability attached to the exercise of employer functions.[73]

In addition to the loss of managerial agency, ARM also narrows the space for worker participation in individual and strategic decisions, including decisions regarding the implementation of new systems that affect workers' rights and working conditions. As a result, it erodes the agency enjoyed by workers and their representatives—individually and collectively.[74]

In addition to overcoming information asymmetries and tackling privacy harms, the second main goal of regulating ARM should therefore be a restoration of human agency across all relevant stages of a firm's decision-making processes. To this end, the policy options set out in this section include a limited ban on fully automated decision-making, supplemented by human involvement both before and after the loop of automated decision-making, as well as regular impact assessments above the loop.

3.1 Policy option 5: Humans in the loop (prohibition of automated termination)

The full automation of significant decisions can lead to a complete and immediate loss of human agency; for example, when management have no control over termination decisions in their particular office or plant. [75] Existing restrictions on fully automated decision-making, such as Article 22 GDPR, are limited in scope[77] and unclear in practice. At the same time, excessive involvement of

71. Katherine C Kellogg, Melissa A Valentine, and Angèle Christin, 'Algorithms at Work: The New Contested Terrain of Control' (2020) 14 The Academy of Management Annals 366.
72. Abigail Gilbert and Anna Thomas, 'The Amazonian Era—The Gigification of Work' (Institute for the Future of Work 2021) 39.
73. Jeremias Prassl, *The Concept of the Employer* (Oxford University Press 2015), ch 1.
74. Adams and Wenckebach (elsewhere in this issue).
75. Colin Lecher, 'How Amazon Automatically Tracks and Fires Warehouse Workers for "Productivity"' (*The Verge*, 25 April 2019) <https://www.theverge.com/2019/4/25/18516004/amazon-warehouse-fulfillment-centers-productivity-firing-terminations> accessed 23 June 2022; Cansu Safak and James Farrar, 'Managed by Bots: Data-Driven Exploitation in the Gig Economy' (Worker Info Exchange Report 2021).
76. Microsoft Viva Insights <https://www.microsoft.com/en-us/microsoft-viva/insights> accessed 16 December 2022; see further Wolfie Christl, 'Digitale Überwachung Und Kontrolle Am Arbeitsplatz: Von Der Ausweitung Betrieblicher Datenerfassung Zum Algorithmischen Management?' (Cracked Labs, September 2021) <https://crackedlabs.org/daten-arbeitsplatz> accessed 17 January 2022.
77. Although Article 22(1) prohibits automated decision-making that significantly affects workers, employers can still circumvent this prohibition by relying on the series of exceptions provided by Article 22. For details, see Abraha (elsewhere in this issue).

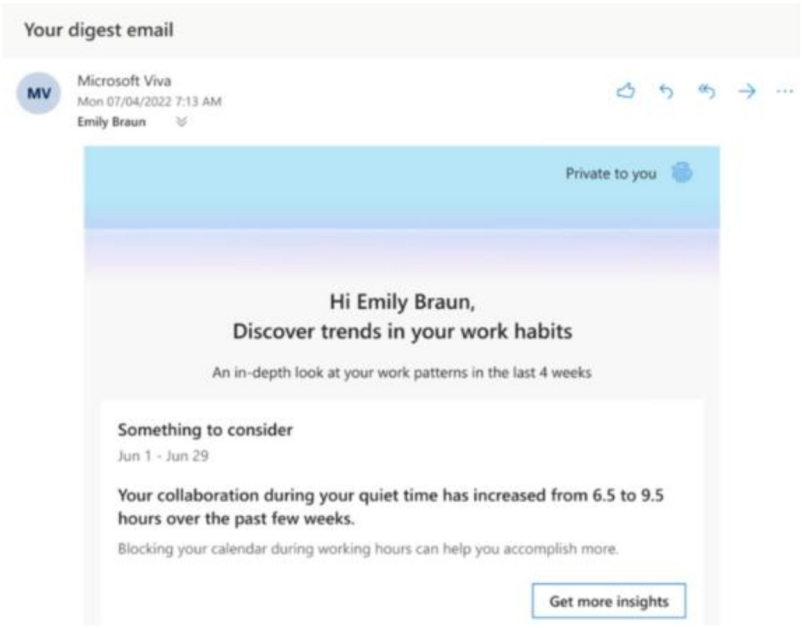

Figure 1. A screenshot from a video explaining the features of 'Microsoft Viva,' a tool that integrates into Microsoft Teams and Microsoft Outlook.[76] Although the tool is currently described as 'privacy-protected' and the email in the screenshot indicates that the email is 'private to' the recipient, the information in the email could easily be sent to managers on their request. (Or, assuming the email is received on a workplace account, managers may simply have direct access to it.) In jurisdictions where it is not legally required that workers be given notice about the generation of such insights and their purpose, management could use them for performance evaluation without workers' knowledge—and without any opportunity for workers to contest inaccurate or misleading insights.

human decision-makers can lead to fatigue and the rubber-stamping of ARM systems' recommendations. Legal regulation should therefore ensure meaningful human agency at the most consequential, and potentially irrecoverable, moment in the employment relationship: termination.

Prohibition of Automated Termination

Prohibit automated termination of the employment relationship. Specifically:

1. Establish a requirement for *meaningful* human involvement in termination decisions.

3.1.1 Rationale. Perhaps the most intuitive response to concerns about an absence of human agency is to reinstate a human decision-maker: high-stakes decisions should not be made on a solely automated basis.[78] Any such bright line quickly breaks down, however, as has become clear in discussions

78. For a concerning example, see Lecher (n 75); documents obtained by *The Verge*, available at <https://cdn.vox-cdn.com/uploads/chorus_asset/file/16190209/amazon_terminations_documents.pdf> accessed 22 February 2023; Safak and Farrar (n 75); 'Dutch & UK Courts Order Uber to Reinstate "Robo-Fired" Drivers' (*Worker Info Exchange*, 14 April

surrounding Article 22 GDPR, which provides a data subject with 'the right not to be subject to a decision based solely on automated processing . . . which produces legal effects concerning him or her or similarly significantly affects him or her'.[79] A series of questions immediately arise: how is 'significance' to be understood, for example,[80] and what does it mean for a decision to be 'solely' automated?[81] Moreover, how do these considerations change when decision-making processes are multi-stage, as where job applicants are triaged before human review and the 'bottom' 10% receive less attention, for example?[82]

Such ambiguities may be unavoidable when designing an omnibus human in the loop provision. Employment-specific legislation, on the other hand, facilitates greater precision: concrete decisions which require human involvement can be identified. Limiting the prohibition to a small set of clearly defined decisions ensures a proportionate approach and negates the need for a vaguer reference to 'significance'. The narrow scope of application also makes meaningful human oversight a more realistic prospect: where a large proportion of decisions require human involvement, it is likely that rubber stamping will be the *de facto* norm.[83]

We propose termination of the employment relationship as the sole decision requiring meaningful human involvement because: (i) it is a uniquely harmful decision; (ii) existing law already makes it necessary to identify a moment of decision-making;[84] and (iii) despite the existing ban on significant automated decision-making, various carve-outs in Article 22 GDPR mean that the lawfulness of automated dismissal is still ambiguous.[85]

2021) <https://www.workerinfoexchange.org/post/dutch-uk-courts-order-uber-to-reinstate-robo-fired-drivers> accessed 20 November 2022.

79. The scope of application of this right, and its interplay with the proposals made here, is discussed briefly at section 4.
80. For an example of exploration in the context of personalised pricing, see Benjamin Wong, 'Online Personalised Pricing as Prohibited Automated Decision-Making under Article 22 GDPR: A Sceptical View' (2021) 30 Information & Communications Technology Law 193, 200–201. For an attempt at guidance, see Article 29 Working Party, 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' (22 August 2018) WP 251 rev.01, 21–22. These Guidelines have been adopted by the European Data Protection Board (EDPB).
81. For discussion, see Article 29 Working Party, 'Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679' 20–21; Diana Sancho, 'Automated Decision-Making under Article 22 GDPR: Towards a More Substantial Regime for Solely Automated Decision-Making' in Martin Ebers and Susana Navas (eds), *Algorithms and Law* (Cambridge University Press 2020) 142–144. Both of these ambiguities are long-standing; for early discussion, see Lee Bygrave, 'Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling' (2001) 17 Computer Law & Security Review 17, 19–20. They persist nonetheless; both issues were raised by stakeholders in the UK government's consultation on reforming data protection law, but there was 'no clear consensus' on how to resolve them. UK Department for Digital, Culture, Media & Sport, 'Data: a new direction—government response to consultation' (updated 23 June 2022) <https://www.gov.uk/government/consultations/data-a-new-direction/outcome/data-a-new-direction-government-response-to-consultation> accessed 22 February 2023. For further critique on Article 22, see Tal Z Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 47 Seton Hall Law Review 995, 1015–1018. There is also ongoing debate about whether Article 22 bestows a right which must be exercised or implies an automatic prohibition; see Luca Tosoni, 'The Right to Object to Automated Individual Decisions: Resolving the Ambiguity of Article 22(1) of the General Data Protection Regulation' (2021) 11 International Data Privacy Law 145.
82. Reuben Binns and Michael Veale, 'Is That Your Final Decision? Multi-Stage Profiling, Selective Effects, and Article 22 of the GDPR' (2021) 11 International Data Privacy Law 319.
83. For the concept of 'meaningful' human involvement, see Article 29 Working Party (n 79) 21.
84. Clarity about the moment of a dismissal decision may be necessary for purposes such as pay and notice period. For an example of a definition, see UK Employment Rights Act 1996 s 97. Although dismissal decisions may still be multi-stage (in that dismissal may be preceded by disciplinary processes), a final decision on dismissal can thus ordinarily be identified, and human involvement in that decision can be mandated.
85. For a description of the relevant carve-outs, see Henni Parviainen, 'Can Algorithmic Recruitment Systems Lawfully Utilise Automated Decision-Making in the EU?' (2022) 13 European Labour Law Journal 225. Reliance on these carve-outs will be even more difficult in the context of termination, but the same ambiguities apply in principle.

The Article 22 right is not universal: it only prohibits significant decisions which are not necessary for entering into, or the performance of, a contract between the data subject (worker) and data controller (employer).[86] Recall policy option 2's proposal to permit the deployment of ARM only where necessary for performance of the employment contract, compliance with another legal obligation, or protection of the vital interests of the worker or another natural person. On that basis, an employer would have to show that automated termination is necessary for the performance of the employment contract, leaving Article 22 GDPR unable to provide affected workers with a right not to be subject to such decision-making.[87] As a result, it is necessary to explicitly specify the decisions which it is not permissible to fully automate. Termination is the most significant decision in the employment relationship, the 'tail [which] wags the whole dog of the employment relation'.[88] Fully automated termination should therefore be explicitly prohibited under all circumstances, regardless of the legal basis for the processing involved.

3.2 Policy option 6: Humans after the loop (right to human review)

The loss of human agency and organisational accountability resulting from the automation of significant decisions furthermore threatens to leave workers without a meaningful avenue of contestation. Existing protections are limited and legally uncertain, thus requiring the creation of clear avenues to contest and correct inappropriate automated decisions. To this end, regulators should establish a series of rights with respect to decisions taken or supported by ARM systems, including a right to request and receive an explanation of the decisions, a right to contest the decisions, and a right to request human (managerial) review—and, if appropriate, correction—of the decisions.

A Right to Human Review

A right to decisions based on accurate facts and valid decision-making processes should be realised in the context of algorithmic management. To this end:

1. Establish individual rights regarding decisions taken or supported by algorithmic management systems; *viz*, the rights for an individual affected by such a decision to:
   (a) request and receive a written explanation of the facts, circumstances, and reasons leading to the decision;
   (b) contest the decision;
   (c) discuss, supplement, and clarify these facts, circumstances, and reasons with a competent authorised human;
   (d) request and receive a human review of the decision in light of the above; and
   (e) have the decision rectified if the facts, circumstances, and/or reasons leading to it are found to be erroneous or unlawful.

3.2.1 Rationale. Individuals should have a right not to be subjected to decisions based on decision-making processes that are flawed (e.g., as a result of a technical or design fault, or biased or

86. GDPR, Art. 22(2)(a).
87. GDPR, Art. 22(3) would require the employer to 'implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision', but such protections would be afforded to a broader range of decisions under this proposal in any case.
88. Hugh Collins, *Justice in Dismissal* (Oxford University Press 1992) 270.

incomplete training or input data), fundamentally unfit for purpose (e.g., a system whose purveyor claims that it can hire the best candidates for a job, but in fact chooses randomly, or based on criteria that cannot credibly be claimed to predict job performance[89]), or based on incorrect or incomplete facts.[90]

Existing laws aim to offer such protection in a variety of ways. Data protection law provides a right to rectify inaccurate or incomplete personal data in Article 16 GDPR. Domestic labour laws restrict the lawful grounds for employee sanction and dismissal, as well as providing a variety of procedural protections.[91] ARM systems, however, expose a number of gaps and weaknesses in existing protections. Existing regimes, for example, do not appear to provide a right to valid or fit for purpose decision-making processes. This is a particularly salient problem because empirical research reveals that flawed ARM systems are common, especially in automated hiring.[92]

Policy option 3, above, proposed a requirement that the deployment of ARM must be linked to specific goals and must be capable of achieving those goals. The present policy option creates individual rights that tie together existing protections and requirements in a way that should protect not only against decisions made on the basis of inaccurate or incomplete data (as covered by the GDPR) but also against flawed, invalid, or otherwise unfit decision-making processes. Such individualised protections are a necessary check against the ubiquitous deployment of different systems which are flawed in similar ways (e.g., in discriminating against particular demographic groups as a result of having been trained on data that embodies existing biases), which can result in the same individuals repeatedly suffering the same detriment.[93]

If a worker is given a formal warning as a result of a performance score produced by an algorithmic system, she could exercise the right established in this policy option to receive clarification of the algorithmic process that produced the score. If the system is found not to be capable of serving the purpose for which it was deployed—for example, because it uses inappropriate criteria—then this policy option establishes a further right to have the decision rectified, surfacing the potential unlawfulness of the system.[94]

Alternatively, if the system is found to be capable (i.e., valid), but the specific decision is found to have been made on the basis of inaccurate or incomplete data, policy option 6 establishes a right

89. See, e.g., Rhea and others (n 39); Narayanan (n 48); Kaltheuner (n 39). See also policy option 2, above.
90. On the need for policymakers to consider fitness for purpose, see Inioluwa Deborah Raji and others, 'The Fallacy of AI Functionality', Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency <http://arxiv.org/abs/2206.09511> accessed 7 July 2022.
91. See generally Bernd Waas and Guus Heerma van Voss (eds), *Restatement of Labour Law in Europe, vol III: Dismissal Protection* (Hart Publishing 2022).
92. See, e.g., Raghavan and others (n 49) 469–481; Rhea and others (n 39).
93. Kathleen Creel and Deborah Hellman, 'The Algorithmic Leviathan: Arbitrariness, Fairness, and Opportunity in Algorithmic Decision-Making Systems' (2022) 52 Canadian Journal of Philosophy 26.
94. Although the GDPR recognises the right to express one's point of view and the right to contest the decision made by algorithms, there is a strong disagreement over whether and to what extent the GDPR also creates the right to explanation; cf Bryce Goodman and Seth Flaxman, 'European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation"' <https://arxiv.org/abs/1606.08813> accessed 14 December 2022; 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' (2017) 7 International Data Privacy Law 76; Gianclaudio Malgieri and Giovanni Comandé, 'Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation' (2017) 7 International Data Privacy Law 243; Lilian Edwards and Michael Veale, 'Slave to the Algorithm? Why a "Right to an Explanation" Is Probably Not the Remedy You Are Looking For' (2017) 16 Duke Law & Technology Review 18.

for the worker to have the decision re-computed once the data have been corrected. Notably, while data protection law (specifically Article 16 GDPR) establishes a right to rectify or complete inaccurate or incomplete personal data, it is not clear whether this implies a right to have decisions made on the basis of inaccurate data re-made; indeed, some have argued that the courts have interpreted data protection law in such a way to exclude the existence of such a right.[95]

Finally, this policy option clarifies that, at least in the context of ARM, a right to an explanation does (and should) exist, and applies to any decision taken or supported by ARM systems.

3.3 Policy option 7: Humans before the loop (information and consultation rights)

In addition to reviewing specific decisions *ex post facto*, the restoration of human agency also requires a space for meaningful contestation concerning the deployment, configuration, and changes to ARM systems. In order to ensure meaningful human involvement in such decisions, ARM systems should be included explicitly in the scope of existing information and consultation rights and obligations.

Information and Consultation Rights

Establish a formal right to information and consultation (for worker representatives) regarding the design, configuration, and deployment of algorithmic management systems, as well as regarding any changes to configuration that trigger individual notifications, as set out in policy option 3. In the EU context, this could be achieved by adding a new point (d) to Article 4(2) of directive 2002/14, such as:

> (d) information and consultation on decisions regarding the development, procurement, configuration, and deployment of algorithmic management systems, as well as any changes to the system or its configuration that affect, or can be expected to affect, working conditions.

Information and consultation rights are a minimum requirement. In Member States where existing worker governance rights, such as codetermination, go beyond information and consultation, algorithmic management should be explicitly added to the obligatory scope of those rights.

3.3.1 Rationale. The introduction of new technology in the workplace is a fundamental moment of change, including, but not limited to, changes in skill requirements for different roles and workers' levels of autonomy and discretion.[96] It frequently presents an opportunity for the reorganisation of existing practices and processes, which in turn often occasions the renegotiation of previously agreed upon rights and obligations—and, occasionally, conflict between employers and employees and their representatives.[97] In a particularly stark example, an industry study in 2021 found that

95. See, e.g., Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-thinking Data Protection Law in the Age of Big Data and AI' (2019) 2019(2) Columbia Business Law Review 494, 529–30.
96. See, e.g., Peter H Cappelli, 'Skill Gaps, Skill Shortages, and Skill Mismatches: Evidence for the US,' (2014) (US NBER Working Paper 20382, 29–30, and citations therein. Now published at (2015) 68 Industrial & Labor Relations Review 251 <https://journals.sagepub.com/doi/full/10.1177/0019793914564961> accessed 22 February 2023.
97. See, e.g,. Valeria Cirillo and others, 'Trade Unions' Responses to Industry 4.0 amid Corporatism and Resistance' (2020) Laboratory of Economics and Management Working Paper Series No 2020/21, esp 13–15 <https://www.econstor.eu/bitstream/10419/228160/1/1727909011.pdf> accessed 22 February 2023.

88% of 1,250 US employers surveyed 'terminated workers after implementing monitoring software', presumably because the monitoring software revealed employees spending work time on 'non-work activities'.[98]

Given these stark findings, all workers whose working conditions are likely to be impacted by an ARM system should in principle be involved in its design.[99] In practice, this may not always be possible, especially given that the majority of ARM systems are not designed in house, but rather are purchased from external vendors. It is important not to overstate this point, however, as most such systems are not deployed out of the box, but rather are customised prior to or during deployment. The moments of customisation and deployment are therefore a key point in the life cycle of an ARM tool or system in a particular workplace where strong collective information and consultation rights can and should attach.

For example, it would likely be impracticable to stipulate that providers of standard office and collaboration software such as Microsoft Teams consult every employee in every workplace the software will be used, or even every representative body in such workplaces. However, a workplace deploying such software should be obligated to provide information and to consult with relevant employee representatives when deciding which of the productivity monitoring (i.e., ARM) features of such software (see, e.g., Figure 1) should be activated and how they will be used.

An explicit right to information and consultation regarding ARM creates the possibility to proactively anticipate and mitigate possible harms or risks that may arise from the deployment of these systems. The policy options in this blueprint are designed to create safeguards that make it highly likely for such harms and risks to be surfaced and corrected after the systems have been deployed. However, investigating and correcting faulty or unexpectedly unlawful systems may be time-consuming and costly.[100] Establishing explicit rights to information and consultation makes it more likely that potential harms and risks can be anticipated and mitigated before deployment.[101] The proposal aims thereby to mitigate long-term harms and reduce the total lifecycle costs of operating legally compliant and socially responsible ARM systems.[102]

98. '6 in 10 Employers Require Monitoring Software for Remote Workers' (*Digital.com*, 2022) <https://digital.com/6-in-10-employers-require-monitoring-software-for-remote-workers> accessed 22 February 2023.

99. Best practices on the design of technology in the workplace generally have been developed under the heading 'participatory design'; in the context of participatory design of algorithmic management specifically, see, e.g., Min Kyung Lee and others, 'Participatory Algorithmic Management: Elicitation Methods for Worker Well-being Models,' Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society, 715–726.

100. See generally Donald Langevoort, 'Monitoring: The Behavioral Economics of Corporate Compliance with Law' (2002) 2002(1) Columbia Business Law Review 71.

101. The desirability of surfacing and attempting to mitigate risks prior to the deployment of technology is well-established in the literature on occupational safety and health; see, for example, Bruce K Lyon and Georgi Popov, 'Communicating and Managing Risk: The Key Result of Risk Assessment' (2017) 62 Professional Safety 35. In the discipline of risk management generally, stakeholder consultation is well-established as a key practice. ISO 31000 ('Risk management—principles and guidelines'), for example, emphasises that 'communication and consultation with external and internal stakeholders should take place during all stages of the risk management process.'

102. In software engineering, there is an extensive literature on lifecycle costs and the ways in which surfacing requirements early in the software life cycle can significantly reduce costs; see, for example BW Boehm, 'Verifying and Validating Software Requirements and Design Specifications' (1984) 1 IEEE Software 75; and, more recently, Klaus Ehrlenspiel and others, *Cost-Efficient Design* (Springer 2007), especially ch 4 ('Influencing the lifecycle cost'). More broadly, the financial returns to stakeholder engagement have been studied extensively in the management literature; see, e.g.,

3.4 Policy option 8: Humans above the loop (impact assessments)

The organisational impact of technically and logically complex ARM systems is often difficult to predict and sometimes, due to a lack of data on impacts and structured high-level oversight, difficult to manage *ex post facto*. Limited or unclear employer obligations regarding prospective impact assessments, such as Article 35 GDPR, do little to address this problem. Regulators should therefore seek to increase the scope and quality of *ex ante* and *ex post* human deliberation regarding potential risks of ARM systems, and consideration and implementation of appropriate risk mitigation strategies by ensuring internal collection of data on the impacts of ARM systems. This could include substantial annual impact assessments for all ARM systems, and the involvement of worker representatives in the production and publication of such assessments.

Impact Assessments

Employers should carry out annual algorithmic management impact assessments (ARMIA) to evaluate the impacts of algorithmic management systems on working conditions.

1. The ARMIA could include:
   (a) all system level information which is to be provided to individual employees and applicants;
   (b) a description and evaluation of the relevant impacts and risks, by reference to quantitative information about the operation of the systems where relevant;
   (c) a description and assessment of any retained or new safeguards adopted to mitigate those impacts and risks;
   (d) an evaluation of the effectiveness of new and existing safeguards, including an assessment of whether they are appropriate for the impacts and risks identified;
   (e) a description of the consultation(s) carried out with workers and their representatives, and of the changes made in response to views expressed.
2. Working conditions should be defined to include at least:
   (a) workers' access to work assignments, their earnings, their occupational safety and health, their working time, their promotion and their contractual status;
   (b) evaluation of risks to the safety and health of workers, in particular regarding possible risks of work-related accidents, psychosocial, and ergonomic risks;
   (c) other working conditions regulated in domestic law.
3. Employers should consult worker representatives when identifying the risks and possible safeguards, and consider and include the views of worker representatives as part of the ARMIA.
4. There should be clear publication requirements for the ARMIA:
   (a) The ARMIA is to be made publicly available, subject to redaction of confidential technical and commercial detail.
   (b) The full (unredacted) ARMIA is to be available to worker representatives and regulatory bodies, with suitable measures to protect confidentiality.

Witold J Henisz, Sinziana Dorobantu, and Lite J Nartey, 'Spinning Gold: The Financial Returns to Stakeholder Engagement' (2014) 35 Strategic Management Journal 1727.

3.4.1 Rationale. Limiting and regulating the use of ARM systems in general means that the most widespread harms can be identified and prevented at the regulatory level. This approach can only go so far, however: algorithmic management is used in a wide variety of contexts, and a legislative instrument which seeks to capture all harms will ultimately be overdeterminative. Indeed, the limits of generalised regulation are evident within the current blueprint: we propose to create redlines, but these redlines are necessarily limited to contexts in which it would never be appropriate to use ARM; we propose to restrict the lawful grounds for use of ARM, but we note that while the contractual necessity ground must be retained, it carries obvious risks; and we propose to prohibit automated termination, but we recognise that a wider ban on solely automated decision-making would create too much uncertainty. Since harms could still arise within uncaptured areas, consideration turns to the potential for context-specific risk mitigation.[103]

One way to achieve such mitigation is by requiring employers to self-regulate, thus explicitly building harm prevention into the process of deploying and operating ARM systems.[104] This approach has long been adopted in the environmental context, via the mechanism of environmental impact assessments.[105] Data governance has similarly mandated impact assessments at an early and ongoing stage,[106] and impact assessments are now a key feature of many proposals on the regulation of AI and ARM.[107]

103. Alessandro Mantelero and Maria Samantha Esposito, 'An Evidence-Based Methodology for Human Rights Impact Assessment (HRIA) in the Development of AI Data-Intensive Systems' (2021) 41 Computer Law & Security Review 105561, 7 noting in relation to the EU's proposed AI Act that case-specific assessment is 'more effective in terms of risk prevention and mitigation than using risk presumptions based on an abstract classification of "high-risk sectors and high-risk uses of purposes"'; see also Margot E Kaminski and Gianclaudio Malgieri, 'Algorithmic Impact Assessments under the GDPR: Producing Multi-Layered Explanations' (2021) 11 International Data Privacy Law 125, 139. For additional detail on a proposed ARM-specific impact assessment, see Aislinn Kelly-Lyth and Anna Thomas (elsewhere in this issue).
104. See Gunther Teubner, 'Substantive and Reflexive Elements in Modern Law' (1983) 17 Law & Society Review 239, coining the context of 'reflexive law'.
105. Council Directive 85/337/EEC of 27 June 1985 on the assessment of the effects of certain public and private projects on the environment (1985) OJ L175/40; see now Directive 2011/92/EU of the European Parliament and of the Council of 13 December 2011 on the assessment of the effects of certain public and private projects on the environment (2011) OJ l26/1, and Directive 2014/52/EU of the European Parliament and of the Council of 16 April 2014 amending Directive 2011/92/EU on the assessment of the effects of certain public and private projects on the environment (2014) OJ L124/1.
106. GDPR, Art. 35; for discussion of the lessons from environmental law, see A Michael Froomkin, 'Regulating Mass Surveillance as Privacy Pollution: Learning from Environmental Impact Statements' (2015) 2015 University of Illinois Law Review 1713. See also Reuben Binns, 'Data Protection Impact Assessments: A Meta-Regulatory Approach' (2017) 7 International Data Privacy Law 22.
107. For examples of enacted regulation, see Chinese Internet Information Service Algorithm Recommendation Management Regulations (2021) Art. 8, the scope of which includes algorithmic recommendation technology for scheduling and decision-making; Canadian Directive on Automated Decision-Making (2019) and accompanying Algorithmic Impact Assessment tool; New York City Local Law 144 (2021), Automated Employment Decision Tools Law 2021 s 20-871 (1) and (2) (entering into force 1 January 2023). For examples of proposals, see Sophie Farthing and others, *Human Rights and Technology* (Australian Human Rights Commission, 2021); Platform Work Directive, art 7(1); UK Public Authority Algorithm HL Bill (2021–2022) 73, cl 4; UK All-Party Parliamentary Group on the Future of Work, 'The New Frontier: Artificial Intelligence at Work' (2021) 12 <www.futureworkappg.org.uk> accessed 22 February 2023; Algorithmic Accountability Act (n 8) s 4(5); California Bill AB-1651, s 1560; Trades Union Congress, 'Dignity at Work and the AI Revolution' (2021) 8; Stephanie Sheir and others, 'Policy Briefing: Algorithmic Impact Assessments—Building a Systematic Framework of Accountability for Algorithmic Decision Making' (Institute for the Future of Work 2021) <https://perma.cc/3H9W-ZPHU>; Bernhardt, Kresge and Suleiman (n 15) 25.

Although the GDPR's data protection impact assessment (DPIA) provision is a good starting point for mandatory algorithmic impact assessments,[108] it suffers from several shortcomings: the DPIA obligation may not apply to all ARM systems,[109] and there are no transparency obligations attached to the DPIA. Transparency is necessary for assessments' full benefits to be realised, both within the firm and more broadly. Appropriate transparency supports worker participation by ensuring that worker representatives can verify the results of the consultation process and enabling oversight by regulators and civil society.[110] Publication of the impact assessment also facilitates genuine transparency, as it should include 'the algorithmic system's reason for existence, the context of the development, [and] the effects of the system'.[111] Identification of such effects will include identification of systemic arbitrary decision-making—one of the key concerns about the demise of human agency.[112]

This policy option builds on the DPIA obligation by more concretely specifying procedural requirements. In contrast with the GDPR, this instrument would better clarify the contents of the impact assessment,[113] and would specify a stronger consultation requirement.[114] The potential for participatory co-construction of impacts and mitigations, explored in depth above, is one of the benefits of impact assessments.[115] Effective consultation would require in-depth information to be shared with worker representatives, including the full version of the final ARMIA. While this may raise concerns about confidentiality or trade secrets, there are already established approaches within EU law for protecting confidentiality in such circumstances.[116]

IV. Conclusion

The policy options set out in the preceding sections have presented different ways of tackling specific instances of privacy harms, information asymmetries, and the loss of human (especially, but not only, managerial) agency. In order to realise the proposal's ambitions, one further, overarching, set of policies is required: capacity building for management, workers, and worker representatives in order to facilitate meaningful engagement with ARM systems.

108. Kaminski and Malgieri (n 103); for a good example, see Ministerie van Justitie en Veiligheid, 'Public DPIA Teams OneDrive SharePoint and Azure AD' (16 February 2022) <https://www.rijksoverheid.nl/documenten/publicaties/2022/02/21/public-dpia-teams-onedrive-sharepoint-and-azure-ad> accessed 24 June 2022.
109. For example, if completely anonymised data are processed to make macro-level decisions about the organisation, that processing may have impacts for workers but could nonetheless fall outside the scope of the GDPR's DPIA obligation if the data are not considered 'personal data': GDPR, Art. 4(1) and recital 26.
110. Directive 2011/92/EU, recital 16, notes that public participation in environmental impact assessments increases accountability and transparency, thus 'contributing to public awareness of environmental issues and support for the decisions taken'. The absence of transparency for data protection impact assessments (DPIA) under Article 35 GDPR has been identified as the biggest shortcoming of the DPIA and its point of greatest divergence from the model algorithmic impact assessments proposed in the literature: Kaminski and Malgieri (n 103) 130–133.
111. Wieringa (n 52) 7.
112. Creel and Hellman (n 92); Sandra Wachter, 'The Theory of Artificial Immutability: Protecting Algorithmic Groups Under Anti-Discrimination Law' (Tulane Law Review, forthcoming 2022).
113. The requirements of Article 35(7) GDPR are less specific.
114. Article 35(9) GDPR merely requires data controllers to 'seek the views of data subjects or their representatives on the intended processing' 'where appropriate', an obligation which has proven very weak in practice.
115. See, for example, Directive 2011/92/EU, recitals 16 and 17.
116. Directive 2002/14/EC of the European Parliament and of the Council of 11 March 2002 establishing a general framework for informing and consulting employees in the European Community (2002) OJ L80, art 6.

This point has been well-recognised in the context of discussion on humans in the loop for data protection law, with official guidance on Article 22 GDPR suggesting that human intervention should be 'carried out by someone who has the authority and competence to change the decision'.[117] Proposed criteria for determining whether oversight is meaningful include the degree of liability for failure (with higher liability being proposed as a mechanism for greater engagement), the level of support available for the task, and the level of information access provided to the decision-maker.[118] Legislation could specify criteria for ensuring that humans in and after the loop are able to reject algorithmic outputs or overturn algorithmic decisions without fear of disproportionate consequences in the event of human error—or of arbitrary retaliation.

Legislative proposals have also sought to create autonomy and capacity for impact assessors: the proposed Platform Work Directive, for example, would require persons responsible for monitoring impacts to have 'the necessary competence, training and authority to exercise that function', and to 'enjoy protection from dismissal, disciplinary measures or other adverse treatment for overriding automated decisions or suggestions for decisions',[119] while the proposed US Algorithmic Accountability Act would require 'ongoing training and education' for relevant individuals on documented impacts in analogous cases and any improved methods on conducting impact assessments.[120] The enacted Canadian Directive on Automated Decision-Making, which requires impact assessments for algorithmically informed public sector decisions, provides for 'employee training in the design, function, and implementation of the Automated Decision System' to ensure individuals' capacity to carry out their tasks.[121] California's Workplace Technology Accountability Bill similarly requires that human reviewers should be (i) granted sufficient authority, discretion, resources, and time to corroborate the ARM outputs, (ii) have sufficient expertise and understanding of the ARM in question to interpret its outputs as well as the results of relevant algorithmic impact assessments, and (iii) have the education, training, or experience sufficient to allow them to make a well-informed decision.[122]

Autonomy for employee-side human decision-makers should be similarly ensured, whilst protecting worker autonomy and capacity. The requirement to respond to, and engage with, workers' representatives' views, as included in consultation requirements imposed before and above the loop, is designed to provide a forum for substantive worker participation in the impact assessment process.[123] Worker representatives should also have access to expert assistance and training as necessary to engage with consultation processes.[124] Where human decision-makers

117. Article 29 Working Party (n 79) 21.
118. Ben Wagner, 'Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems' (2019) 11 Policy & Internet 104.
119. Platform Work Directive, art 7(3).
120. US HR 6580 s 4(5)).
121. Canadian Directive on Automated Decision-Making (2019), section 6.3.5; particularised at Appendix C.
122. California Bill AB-1651, s 1555.
123. These include collective bargaining and co-determination rights, for example; cf California AB 1651 (n 6), amending Division 2 of the Labor Code s 1563 and providing workers with a right to anonymously dispute impact assessments and 'request that the labor agency conduct an investigation of the employer' if certain criteria are satisfied.
124. For an example, see Platform Worker Directive, Art. 9(3) or Directive 2009/38/EC of the European Parliament and of the Council of 6 May 2009 on the establishment of a European Works Council or a procedure in Community-scale undertakings and Community-scale groups of undertakings for the purposes of informing and consulting employees (Recast) (2009) OJ L122/28, Art. 7 and Annex I.

are tasked with changing, overturning, or limiting automated decisions, the law must ensure that they have the capacity and autonomy to do so.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

We acknowledge funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement No 947806), and welcome feedback and discussion: ai.work@law.ox.ac.uk. The usual disclaimers apply.