Special Issue Contribution

ELLJ

European Labour Law Journal
1–30

DOI: 10.1177/20319525231167981
journals.sagepub.com/home/ell
SAGE

From access to understanding: Collective data governance for workers

Dan Calacci
MIT Media Lab, Massachusetts, MA 02139, USA

Jake Stein
Department of Computer Science, University of Oxford, Oxford, UK

Abstract
Regulating data collection and use in the workplace is now more a matter of regulating working conditions than data protection. This article argues that existing data protection law fails workers for precisely this reason. We examine how workers currently use data protection law, labour law, and technology to access and leverage the data they produce at work and identify key ways in which current regulation falls short. Existing regulations primarily aim to protect worker privacy, an approach that ignores the fact that data use now often defines the fundamental conditions of work, particularly in the gig economy. This is because a key limitation of modern data protection law for workers is its myopic focus on the individual 'data subject', whose rights to data stem from a right to privacy or data protection. Instead, data regulation in the workplace requires a framework that acknowledges the core interest workers have in accessing their data: to collectively exert greater agency and control at work. We argue that workplace data regulation should largely be a matter of workplace governance and worker co-determination, an approach rooted in workers' rights, to negotiate the terms of their employment agreements and specific working environments.

Keywords
data protection, data access, worker privacy, data subjects, collective agency, workplace governance, co-determination, collective rights

Corresponding author:
Dan Calacci, MIT Media Lab, Massachusetts, MA 02139, USA
Email: dcalacci@media.mit.edu

1. Introduction: Current approaches and limits

By appropriating information about 'data subjects' through user interactions, surveillance, and the general instrumentation of daily life, many modern firms transform data into new kinds of value, ranging from more accurate advertising pipelines to highly engineered warehousing processes.[1] Defined variously as informational, surveillance, or data capitalism, current relations between firms, workers, and consumers are heavily characterised by the use of data.[2] While the extraction of data from consumers might indicate a new mode of production, the flow of data in the workplace is thicker, more opaque, and more personally consequential than that of the consumer.[3] Its lack of attention from critics can be attributed not to its non-existence, but instead to the fact that rather than transforming workplace relations, it simply deepens them.

This article takes the perspective that the use of workplace data by firms does not fundamentally transform the capitalist political economy. Instead, we view the modern use of worker data as an extension of already-existing employment relations. Firstly, to control and discipline workers (workplace surveillance).[4] Secondly, to optimise production processes—for example, to hone the algorithms that control ridesharing market supply or to set production quotas for warehouse workers.[5] Finally, workplace data is used to produce value that is secondary to the primary goal of the firm, such as companies 'dogfooding' their own products or ride-hailing firms selling mobility datasets.[6] It is this final characteristic of data use in the consumer context that makes so many critics eager to describe our political economy as capitalism in new clothes.[7] However, when viewed from the lens of the worker, it does not necessarily produce a foundational shift in relations, but an escalation of asymmetric control.[8]

This perspective suggests that workplace data relations, and technology use generally, should be treated as an aspect of modern workplace conditions. If this is the case, then the question of how best to regulate data and algorithms at work is directly related to the rights of workers within the

1. Amazon, for example, optimises warehouse worker behaviour based on thick workplace surveillance: 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 9 November 2021.
2. For more on data relations, see Julie E Cohen, *Between Truth and Power* (Oxford University Press 2019). Some have even suggested that modern data relations have returned the west to a political economy more akin to feudalism than capitalism; see Evgeny Morozov, 'Critique of Techno-Feudal Reason' (2022) 133 New Left Review 89; surveillance capitalism, platform capitalism, informational capitalism; Phil Jones, *Work without the Worker: Labour in the Age of Platform Capitalism* (Verso Books 2021).
3. Historically, employers have had almost a 'limitless' prerogative to surveil workers; see Ifeoma Ajunwa, Kate Crawford, and Jason Schultz, 'Limitless Worker Surveillance' (2017) 105 California Law Review 735.
4. ibid; see also Ifeoma Ajunwa and Daniel Greene, 'Platforms at Work: Automated Hiring Platforms and Other New Intermediaries in the Organization of Work' in Steven P Vallas and Anne Kovalainen (eds), *Work and Labor in the Digital Age* (Emerald Publishing Limited 2019).
5. Julia Laing Gordon, 'Under Pressure: Addressing Warehouse Productivity Quotas and the Rise in Workplace Injuries' (2021) 49 Fordham Urban Law Journal 149.
6. 'Dogfooding' refers to firms testing their own products internally before releasing them on the market; for more on value creation through data assets, see Niels van Doorn and Adam Badger, 'Platform Capitalism's Hidden Abode: Producing Data Assets in the Gig Economy' (2020) 52 Antipode 1475.
7. Lehdonvirta reaches back even further, examining platform-consumer relations using a lens from a variety of power relations: Vili Lehdonvirta, *Cloud Empires: How Digital Platforms Are Overtaking the State and How We Can Regain Control* (The MIT Press 2022).
8. Kathleen Griesbach and others, 'Algorithmic Control in Platform Food Delivery Work' (2019) 5 Socius.

employment relationship. This article builds on recent legal scholarship that re-imagines the role of workplace governance in creating and enforcing employer regulations and limits.[9] Enabling and enhancing worker governance rights within firms and across industries can shift burdens away from employers, grant additional agency to workers, help solve the problem of defining data use related harms *ex ante*, and expand data protection to include collective data rights.

The rest of this introduction examines how workers are currently collectively accessing and using data, with a particular focus on platform work. We argue that data subject rights and data protection should not be the basis for data rights in the workplace, instead supporting nascent approaches grounded in labour law. Section 2 takes these approaches and limitations and concretises them in a speculative case study examining what affordances different approaches might offer to call centre workers. In section 3, we outline how a combination of access rights, liability mechanisms, and worker co-determination could help support workers in algorithmic and data-driven workplaces.

Below, we outline three broad strategies workers are using to protect, access, and benefit from their own data: making collective data subject access requests; enacting new, sector-specific labour law; and constructing 'data intermediaries'.

1.1 Data protection and the GDPR

To begin, it is worth asking why workers might seek access to their data in the first place. Examining existing approaches that workers are taking reveals that working groups are generally seeking access to data in order to aggregate it, aiming to further the goals of worker organisations and interest groups. This approach is particularly salient in platform work, where aggregate data about a collective of workers can create value, help balance marketplaces, and provide helpful regulatory functions.[10] Rather than focusing on privacy, effective regulation for workers must grant meaningful data *understanding*. This means that even if the range of data made accessible to workers is expanded, focusing on only privacy or data subject rights may not be sufficient.[11]

Worker data rights are constrained for two main reasons. First, data protection is primarily understood as a way to preserve the agency of individual data subjects.[12] As a result, data protection law over-relies on identifiability as a basis from which to define which data should be accessible to data subjects. This basis significantly limits the potential of data protection rights to facilitate

9. Matthew T Bodie, 'The Law of Employee Data: Privacy, Property, Governance' (2022) 97 Indiana Law Journal 707.
10. For example, data collected from platforms could be used by drivers to optimise earnings; see Harshal A Chaudhari, John W Byers, and Evimaria Terzi, 'Putting Data in the Driver's Seat: Optimizing Earnings for On-Demand Ride-Hailing' in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (Los Angeles 2018) (Association for Computing Machinery 2018). Data access can also facilitate crowdsourced algorithm audits, such as in Dan Calacci and Alex Pentland, 'Bargaining With The Black-Box: Designing and Deploying Worker-Centric Tools to Audit Algorithmic Management' (2022) 6 Proceedings of the 21st ACM International Conference on Human-Computer Interaction 428.
11. Recent case law in the EU is expanding the reach of data subject access; see Case C-434/16 *Nowak v Data Protection Commissioner* [2017] ECLI:EU:C:2017:994.
12. Rather than address the collective and relational nature of data collection and use, the GDPR and other data protection laws generally focus myopically on the individual: Salomé Viljoen, 'A Relational Theory of Data Governance' (2021) 131 Yale LJ 573 ('The focus on individual selfhood is expressed in the canonical purpose of data governance: informational self-determination').

collective insight.[13] Second, data that should potentially be covered under a comprehensive data access right is generally subject to employer intellectual property rights and protections.[14]

Some of these limits have become clear in case law. In 2021, an international coalition of workers' associations brought a series of lawsuits against Uber and Ola to the Amsterdam Civil Court. These lawsuits, filed by the App Drivers and Couriers Union (ADCU), the International Alliance of App Transport Workers (IAATW), and Worker Info Exchange (WIE), requested insight into the automated decision systems and data processing used by Uber and Ola to allegedly automatically deactivate drivers' accounts, as well as other algorithms used on the platforms.[15] These cases demonstrate some of the specific affordances and limits of the GDPR's ability to grant collective rights to data and information about working conditions. There are three important takeaways from this case law relevant to workers seeking to leverage collective subject access requests.

First, a clear lesson is that the GDPR can offer some useful protections for workers. Although the basis for data subject access rights under Article 15 of the GDPR is motivated by individual rights, the district court of Amsterdam found that exercising rights to gain a collective good—in this case, insight into algorithmic functioning for trade union activities—is lawful.[16] The lack of specificity in Article 15 facilitates using data aggregated from subject access requests for organising and union activity, despite the protests of ride-hailing companies like Ola.[17]

Second, it is currently unclear to what extent various rights extended through the GDPR offer workers access to data about or used by automated systems that increasingly define conditions of work. The various systems targeted by workers' requests, including ride-matching, ride pricing (and, accordingly, driver pay), work scheduling, suspension, deactivation, driver ratings, and penalties, collectively create the algorithmic management systems that define their working conditions, yet the Court rejected most of their requests for information about how these systems

13. ibid 638–639 (on how data as a collective resource facilitates new, more socially functional lines of inquiry and social choice).
14. Daniel Gill and Jakob Metzger, 'Data Access Through Data Portability' (2022) 8 European Data Protection Law Review 221, 11 (noting that trade secret or IP law may protect 'personal data in relation to other data subjects'); also see Ruth Janal, 'Data Portability under the GDPR: A Blueprint for Access Rights?' (Nomos Verlagsgesellschaft mbH & Co KG 2021) 334–335 <https://www.nomos-elibrary.de/10.5771/9783748924999-319/data-portability-under-the-gdpr-a-blueprint-for-access-rights?page=1> accessed 17 October 2022 (although the author is wary of the risk trade secret law presents for the exercise of portability rights, they note that 'the interest of the controller to protect an existing trade secret may be considered under Article 20(4) GDPR').
15. *Applicants v Uber BV*, Rb Amsterdam, 11 March 2021 RvdW 2021, C/13/687315 m.nt (on alleged fraudulent activity leading to automatic driver deactivation, brought by ten drivers from the UK); *Applicants v Ola Netherlands BV*, Rb Amsterdam, 11 March 2021 RvdW 2021, C/13/689705 m.nt (on Ola's various 'scores' assigned to drivers, including fraud probability, matching, and other systems). For a summary of the data, systems, and outcomes of the cases, see Christina Hießl, 'Case Law on Algorithmic Management at the Workplace: Cross-European Comparative Analysis and Tentative Conclusions' (2021) SSRN Electronic Journal <https://doi.org/10.2139/ssrn.3982735>.
16. *Uber BV* (n 15).
17. *Ola Netherlands BV* (n 15) (Ola cited GDPR's focus on the individual to prevent collective data aggregation by workers, noting 'The request to transfer personal data in a certain format stems from the wish of [applicants] to have this data entered directly in a WIE database for analysis with the aim of improving the negotiating position of platform workers. Recital 68 of the GDPR states that the right to data portability serves to strengthen the data subject's control over their own data.')

operate.[18] In these cases, limited access to details about the 'logic involved' in such systems equates to denying workers access to details of their working conditions.

Third, the GDPR's reliance on the concept of the data subject can risk excluding crucial contextual information about data-driven systems that can significantly impact workers. In all cases, Article 20 GDPR, which provides data portability, was interpreted as being limited to data created by the data subject or observed data about the data subject. This interpretation is notably different to recent EU decisions that include inferences and computations made from a subject's data as falling within the category of 'personal data' and, therefore, subject to the right of access; the scope of data to which Article 20 applies is narrower than the scope of the other access rights.[19] This data could include ratings, reliability scores, and weighted customer ratings that are used for meaningful decisions by the systems catalogued above.[20] While this information might still be covered under Article 15, the value of this data for workers submitting requests in aggregate is only realised if it can also be analysed and processed in aggregate. If the data are not made available in machine readable format (i.e., consistent with the provisions of Article 20), aggregate analysis may be prohibitively complex or costly. The limitation of the right to data portability to data provided by data subjects therefore is a significant impediment to workers' ability to use the GDPR to understand and improve their working conditions.

Access is further complicated by Article 15(4), which stipulates that access must not interfere with other rights, 'including trade secrets or intellectual property and in particular the copyright protecting the software'.[21] This is interpreted to mean both the intellectual property and trade secret rights of a platform as well as the privacy rights of customers.[22]

The imprecise nature of trade secrets' application to algorithmic systems means that a wide variety of data, especially when aggregated, may be legally protected as employer property. For example, while truck drivers may be able to individually collect data on the routes they take in order to, e.g., track mileage, the aggregation of a fleet of truck drivers' routes may encompass a trade secret insofar as it might reveal the locations an employer services, a well-established area

18. The Court's reasoning in rejecting one request for data used to profile drivers was not that the request did not fall under Article 22; it was because the request was not 'sufficiently specified'; as others have noted, this suggests that future requests that *are* sufficiently specified may be considered. See 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) 23 <https://fpf.org/wp-content/uploads/2022/05/FPF-ADM-Report-R2-singles.pdf> accessed 21 November 2022. While data *about* a decision system is limited to those that fulfil the requirements in Articles 22(1) and (4), data about the *existence* of a system is required in all cases; see Hießl, 'Case Law on Algorithmic Management' (n 15) 19–21.
19. Case C-184/20 *OT v Vyriausioji tarnybinės etikos komisija* [2022] ECLI:EU:C:2022:601.
20. On the meaning and impact of such scores, ratings, and profiling, see Sylvie Delacroix and Michael Veale, 'Smart Technologies and Our Sense of Self: Going beyond Epistemic Counter-Profiling' in Mireille Hildebrandt and Kieron O'Hara (eds), *Life and the Law in the Era of Data-driven Agency* (Edward Elgar Publishing 2020).
21. 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).
22. While the limits that IP and trade secret law place on access requests were not explored in these cases, driver access to full star ratings of rides was limited by reason of customer privacy. This is ostensibly because the decision systems where it might have been applicable were not considered covered under Article 22. It could be the case that obtaining a copy of the 'logic' of Uber's matching system would violate trade secret law, triggering Article 15(4). See Hießl, 'Case Law on Algorithmic Management' (n 15) 21 ('the protection of interests under Article 15 (4) GDPR could also include the protection of intellectual property and copyright').

of trade secret law. Aggregate collection of wage or salary data has also been the aim of trade secret litigation, although the strength of the trade secret argument to salary transparency has been questioned by legal scholars.[23]

There are other provisions within the GDPR that could be used by workers seeking data understanding. Article 88 empowers Member States to 'provide for more specific rules to ensure the protection of the rights and freedoms' of workers while processing their personal data.[24] The language of paragraph two even suggests that collective data might also be made available.[25] Despite the promise of these rights granting workers a claim to further contextualised personal data, implementations of Member State specific laws under Article 88 have yet to live up to their potential.[26]

Drivers seeking access to workplace data are typically concerned primarily with their working conditions, rather than their personal privacy. Yet these systems are more than just the terms and conditions of work. They have hidden impacts on worker autonomy and the power relationship between platforms and workers that can drastically impact platform working conditions. Driver ratings and matching algorithms can result in a gamified work environment that places significant stress on workers.[27] Matching and pricing algorithms form an environment of information asymmetry, where workers are manipulated into making decisions primarily based on an information environment controlled by the platform.[28] The resulting atomisation of work that these systems create can also impose a chilling effect on potential worker solidarity and organising.[29] How an employer uses worker data has significant downstream impacts on a worker's basic conditions of employment. This raises the question of whether workplace data collection and use should be fundamentally a labour issue, rather than one of data protection.

1.2 Approaches grounded in labour law

Workers and advocates in jurisdictions with weak data protection laws are testing how labour law can regulate data and technology use in the workplace. One approach taken in California is to advance protections for workers through amending jurisdictional labour codes, rather than

23. In both the US and EU, trade secret law is intended to protect firms from private misappropriation of information, not to hinder worker power; see Cynthia Estlund, 'Extending the Case for Workplace Transparency to Information about Pay' (2014) 4 UC Irvine Law Review 781. The US National Labor Relations Act (NLRA) is also clear about workers' ability to discuss wages and work conditions, including aggregation, with each other and allies; see NLRA, 29 USC secs 151–159.
24. GDPR art 88.
25. ibid. Article 88 explicitly targets 'the transparency of processing, the transfer of personal data within a group of undertakings, or a group of enterprises engaged in a joint economic activity and monitoring systems at the workplace'.
26. Halefom H Abraha, 'A Pragmatic Compromise? The Role of Article 88 GDPR in Upholding Privacy in the Workplace' (2022) 12 International Data Privacy Law 276 (noting that Article 88 'invites further fragmentation, legal uncertainty, and inconsistent enforcement').
27. On characterising a driver's perception of his work on a delivery platform, see Griesbach and others, 'Algorithmic Control in Platform Food Delivery Work' (n 8) ('His work on Postmates is a game, a competition between him and others on the platform.').
28. See, e.g., Alex Rosenblat and Luke Stark, 'Algorithmic Labor and Information Asymmetries: A Case Study of Uber's Drivers' (2016) 10 International Journal of Communication 3758; for an investigation into similar work done remotely, see Alex J Wood and others, 'Good Gig, Bad Gig: Autonomy and Algorithmic Control in the Global Gig Economy' (2019) 33 Work, Employment and Society 56.
29. Alex J Wood and Vili Lehdonvirta, 'Antagonism beyond Employment: How the "Subordinated Agency" of Labour Platforms Generates Conflict in the Remote Gig Economy' (2021) 19 Socio-Economic Review 1369.

consumer data privacy bills. The California Consumer Privacy Act (CCPA) and its extension, the California Privacy Rights Act (CPRA), grant consumers broad data rights such as extensive notice requirements of personal data processing and expanded private rights of action for privacy harms.[30] However, both the CCPA and CPRA limit rights under employment through a specific exemption for employee data set to expire in 2023.[31]

Not content to wait until this exemption is removed, workers and advocates have passed and proposed amendments to the state Labor and Government Codes that provide protections for workers in specific sectors and workers more broadly. Assembly Bill 701 (AB-701), passed in 2021, provides warehouse workers and advocates with greater transparency over the use of productivity quotas and automated management systems in warehouses and regulates their use by employers.[32] While this Bill can be read primarily as a workplace safety measure, it is generally understood as a response to the widely criticised data-driven and automated decision systems that Amazon, one of California's largest employers, uses to manage its warehouse workers.[33]

The Bill is designed to help workers who, often due to financial constraints and power asymmetries, 'prioritize quota compliance over their own safety' while working under surveillance, quota, and automated dismissal systems similar to Amazon's.[34] Importantly, it offers workers the ability to request data about their work speed and quotas using a third party, significantly reducing overhead for worker advocates and acknowledging the collective nature of workplace data. AB-701 offers a model of sectoral technology regulation that works to counter specific algorithmic harms, and can be thought of as a kind of pragmatic *ex post* approach to regulating algorithmic harm.[35] However, the Bill does more than just address labour violations: it provides workers with the tools—data

30. For notice requirements, see generally Cal Civ Code §1798.100; for private rights of action, see Cal Civ Code §1798.150(a)(1) ('Any consumer whose . . . personal information . . . is subject to an unauthorized access').
31. Cal Civ Code §1798.145(n) (specifying an exclusion for some obligations in cases where a consumer is acting as 'an employee, owner, director, officer, or independent contractor of' a business). While this exclusion was widely expected to be extended past 2023, the California legislature has, as of the time of writing, adjourned for the 2022 season without extending this exemption, seemingly leaving workers covered under the CCPA and CPRA starting January 2023, see Gary Drenik, 'Privacy Regulations Expand Beyond Ad-Tech Industry' (*Forbes*, 11 November 2022) <https://perma.cc/3WVQ-3HMR>.
32. AB-701 provides workers with the ability to request data about any productivity quotas from their employer if they suspect their quotas have led to injury or workplace violations, places restrictions on what behaviours can be tracked, and requires employers to inform workers about any disciplinary actions taken due to automated quota systems; see AB-701, 2020–2021 Reg Sess (Cal 2021).
33. California houses a higher number of Amazon employees than any other state, and its scale rivals other significant employers in California, such as the University of California. Amazon employs over 170,000 Californians as of 2022; see Amazon, 'Investing in the U.S.' (*About Amazon*, 2021) <https://www.aboutamazon.com/investing-in-the-u-s> accessed 14 October 2022; Amazon's automated management has been blamed for their warehouses' disproportionately high injury rate; see Noam Scheiber, 'California Bill Could Alter Amazon Labor Practices' (*The New York Times*, 22 September 2021) <https://www.nytimes.com/2021/09/06/business/economy/amazon-california-warehouse-labor.html> accessed 14 October 2022. Other investigations have documented various negative impacts of the system on worker well-being; see Spencer Soper, 'Fired by Bot at Amazon: "It's You Against the Machine"' (*Bloomberg.com*, 28 June 2021) <https://www.bloomberg.com/news/features/2021-06-28/fired-by-bot-amazon-turns-to-machine-managers-and-workers-are-losing-out> accessed 9 November 2021.
34. 'Assemblywoman Gonzalez Introduces Bill to Protect Warehouse Workers from Hazardous Working Conditions' (press release, 16 February 2021) <https://perma.cc/D5MM-VH52>; for details about Amazon's automatic dismissal system, see Lecher (n 1).
35. AB-701 is an *ex post* approach, in that it both provides liability mechanisms and utilises *informational advantage* to define harms that should be targeted by regulation; see Brian Galle, 'In Praise of *Ex Ante* Regulation' (2015) 68 Vanderbilt Law Review 1715.

access and explanations of the quota system—needed to gain a more comprehensive picture of their working conditions. With the associated work speed data used to produce the quotas and similar metrics, worker groups could interrogate the systems used by Amazon and other employers, an option unavailable to workers in other US states that may have to rely on limited (or non-existent) data protection law alone.

California Assembly Bill 1651 (AB-1651), introduced in 2022, modifies the state Labour Code to provide workers with rights similar to those granted to consumers by the CPRA, and would be the first comprehensive Bill in the US that would establish worker data rights.[36] Designed specifically for the workplace and based in labour law, it provides more expansive rights and protections in some areas than the CPRA or the GDPR. It adopts an expansive definition of worker data that explicitly includes inferences relating to a worker, offers specific provisions for third-party vendors, registers all workplace productivity algorithms with the state labour agency, and establishes an Occupational Health and Safety review process for systems using algorithms.[37]

Related Bills in other states, such as the Illinois Employee Security Act or Massachusetts' Information Privacy and Security Act, are omnibus Bills that attempt to regulate electronic surveillance in the workplace (Illinois) or general consumer privacy (Massachusetts).[38] These Bills offer comparatively limited provisions for workers, with the recently-introduced Massachusetts Bill actually striking all employee related language that was present in the original Bill.[39] Instead of granting access rights, these Bills limit data collection and regulate how employee data can later be used.[40] Interestingly, the Illinois Bill prohibits any dismissal determined using any data from electronic monitoring systems.[41]

AB-1651 also explicitly acknowledges third-party vendor agreements, a productive first step towards recognising the technical challenges inherent in using data protection rights to address the harms of workplace information asymmetries. The workplace data ecosystem is becoming increasingly complex: data is held not only by employers themselves, but is accessed via federated systems or protocols—leaving such data outside the immediate access or control of employers.

36. See AB-1651, 2021–2022 Reg Sess (Cal 2022), and, for an overview of the bill, see Lil Kalish, 'A Lawmaker Wants to Protect Workers from Surveillance' (*Calmatters.org*, 19 April 2022) <https://calmatters.org/california-divide/2022/04/california-workers-surveillance/> accessed 28 November 2022.
37. For defining worker data, see AB-1651, sec 1521(d) (defining worker data as data that 'relates to, describes, . . . a particular worker, regardless of how the information is collected, inferred, or obtained'); for third-party vendor provisions, see sec 1531(a) (detailing rights including 'a vendor acting on behalf of an employer'); for algorithmic provisions, see sec 1553(b) ('Before an employer or a vendor . . . uses a productivity system that uses algorithms, the employer shall submit a summary of the system to the labor agency').
38. For Massachusetts, see Bill H 4514, 192nd General Ct, 2022 Sess (Mass 2022); for Illinois, see SB 2332, 102d General Assemb, Reg Sess (Ill 2021). At the time of writing, both bills are being considered by their respective state legislatures and have not yet passed.
39. See Bill H 142, 192nd General Ct, 2021 Sess (Mass 2021).
40. The fact that they limit later *use* of employee data actually may place their purview beyond that of the EU's competencies, which limit it to regulating what *kinds* of data are collected. For example, the EU's proposed Platform Work Directive limits collection of 'emotional' data, but not, e.g., *using* data to infer emotions. 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) (Platform Work Directive).
41. This shows an alternate approach to regulating automated decision systems—such systems cannot be put in place if data collected from workers cannot be used in decision-making at all; see SB 2332 (n 38) s 10(i) (holding that an employer may not rely on electronic monitoring data in 'discharging, disciplining, or promoting an employee', and that such decisions must be made using 'human-based information sources' such as co-workers or supervisor assessments).

Federated identity systems and other data products routinely provide access to or redirect subject data in ways that blur the lines between one workplace and the next.[42] In these cases, data may not be immediately identifiable to an individual subject, and thus determining data's provenance and tracking down where it has been shared may be disproportionately difficult—if not impossible.

1.3 Collective bargaining agreements

Professional athletes are some of the most intensely surveilled and quantified workers in the world.[43] The sheer amount of capital involved in professional sports globally has incentivised and enabled sports leagues to measure athletes essentially constantly, both on and off the field.[44] This comprehensive data includes most biometric data that is technically possible to measure, and is routinely used in management decisions and processed by third-party software systems.[45]

Importantly, many highly paid and high-profile professional athletes, especially in the US, are represented by unions with creative bargaining agreements that address data collection and use.[46] In this regard, some athlete unions have been on the forefront of 'negotiating the algorithm' through democratic worker representation for years.[47] Including terms of data use, access, and licensing in collective bargaining agreements (CBAs) is one clear way athlete workers leverage existing labour law to gain a level of control over how their data is used. Both the National Football League and the National Basketball Association's agreements establish committees dedicated solely to the supervision and monitoring of how athlete data is used, with clear grievance procedures and consequences outlined if a team or the league violates the included provisions.[48]

These agreements grant individual athletes the right to define the 'representatives who will have access' to their biometric and other data.[49] While this has been vaguely interpreted as a presumption that athletes have a kind of 'ownership' right over their data, this extended right to control access instead presumes more inalienable, fundamental rights, similar to the normative concept that the

42. Federated identity systems operated by Google and Facebook are facilitated by web protocols like OAuth to allow consumers to 'log-in with Google,' granting authorisation for data transfers from a data subject's Google account to a third party, see: Nate Barbettini, 'OAuth 2.0 and OpenID Connect in Plain English!' (dir OktaDev, 2018) <https://www.youtube.com/watch?v=0VWkQMr7r_c> accessed 9 May 2022; and San-Tsai Sun and Konstantin Beznosov, 'The Devil Is in the (Implementation) Details: An Empirical Analysis of OAuth SSO Systems' (2012) Proceedings of the 2012 ACM Conference on Computer and Communications Security <http://dl.acm.org/citation.cfm?doid=2382196.2382238> accessed 19 April 2022.
43. See Bodie (n 9) 738–743; Skyler R Berman, 'Bargaining Over Biometrics: How Player Unions Should Protect Athletes in the Age of Wearable Technology' (2019) 85 Brooklyn Law Review 543.
44. 'The NBA's Adam Silver: How Analytics Is Transforming Basketball' (*Knowledge at Wharton*, 1 June 2017) <https://perma.cc/J6CP-Q8HW> accessed 15 October 2022.
45. Nick Busca, 'As Biometrics Boom, Who Owns Athletes' Data? It Depends on the Sport.' *Washington Post* (2 February 2021) <https://www.washingtonpost.com/sports/2021/02/02/athletes-biometrics-data-privacy/> accessed 15 October 2022 (on the detailed extent of what data is collected from athletes, including glucose levels, heart rate, skin conductance, etc); Berman (n 43) (on how data is shared and processed by third parties).
46. Berman (n 43) 545, 557.
47. The NBA's CBA included provisions limiting the league's use of wearable sensor data as early as 2017; see Rian Watt, 'The New NBA CBA Addresses Wearable Technology, But What Does That Mean?' (*Vice*, 1 February 2017) <https://perma.cc/ZY6Q-EMJS> accessed 15 October 2022.
48. Berman (n 43) 558–560.
49. ibid 560: '[T]here seems to be presumption of player ownership in data, given the players can request that their team "restrict or expand the list of representatives who will have access to such [biometric] information and data."'

GDPR adopts.[50] This notion is not concretely specified in the CBAs discussed here, but advocates have argued for a core 'bill of rights' to be included in future default CBAs that address biometric data ownership and use.[51] This approach hybridises labour and data protection regulation and could be expanded to other industries, creating a more normative approach to data rights for workers that have union representation.

In the EU, enabling data and AI regulation through workplace collective bargaining agreements is quickly becoming a go-to answer of advocates and legal scholars.[52] The general consensus is that Article 88 GDPR cements jurisdictional collective bargaining agreements as a fundamental protector of data rights in the workplace, and scholars have called for 'default' rights that should be present in agreements.[53] While there are a number of EU Member States with laws that allow collective bargaining over data protection within the workplace, they are primarily concerned with regulating processing, rather than providing access.[54]

There are limits to what bargaining can achieve. First, labour laws in both the US and EU only legally require employers to bargain over core mandatory aspects of the employment relationship, which may not include data protection or technology use.[55] Second, the abysmal union density in the US restricts the potential reach of union approaches to worker data protections.[56] Third, the rapid advancement of fissured work, such as platform work, is largely excluded from company-specific bargaining agreements.[57]

1.4 Data institutions and workers' inquiry

The changes in data protection and labour law outlined above are improving workers' ability to access their personal data, and increased oversight from legal mechanisms like California's AB-1651 or the EU's proposed Platform Work Directive promise to create a 'floor' for conditions of algorithmically mediated work. These are positive steps towards protecting workers, but what workers want in these contexts is greater agency at work; access to data and algorithmic restrictions

50. ibid.
51. ibid 567–569.
52. See Emanuele Dagnino and Ilaria Armaroli, 'A Seat at the Table: Negotiating Data Processing in the Workplace. A National Case Study and Comparative Insights' [2019] SSRN Electronic Journal <https://www.ssrn.com/abstract=3403729> accessed 15 October 2022; Valerio De Stefano and Simon Taes, 'Algorithmic Management and Collective Bargaining' (2022) Transfer: European Review of Labour and Research 8 <https://doi.org/10.1177/10242589221141055> ('Collective bargaining and trade union action are arguably the most effective tools for tackling rapid technological developmens in algorithmic management').
53. Valerio De Stefano, '"Masters and Servers": Collective Labour Rights and Private Government in the Contemporary World of Work' (2020) 36 International Journal of Comparative Labour Law and Industrial Relations 425; for a discussion of 'default' rights, see: Dagnino and Armaroli (n 52).
54. See Abraha, 'A Pragmatic Compromise?' (n 26) 284 (17 EU Member States have some combination of regulatory modes outside of data protection law that carry employee data processing provisions).
55. For a US-centric analysis, see Steven E Abraham and Bart D Finzel, 'New Technology in Unionized Firms: Advantages of Mandatory Bargaining' (1997) 10 Employee Responsibilities and Rights Journal 37.
56. At most, these approaches would reach a little over 6% of the private US workforce, and using the rate at which EU CBAs include data protection provisions as a guideline, they would likely reach a small fraction of those workers in reality. For US union density, see Bureau of Labor Statistics, 'Union Members: 2020' (2021) <https://www.bls.gov/news.release/pdf/union2.pdf>; for the percentage of CBAs with data provisions, see Dagnino and Armaroli (n 52) 186 (noting that 'the proportion of company-level collective agreements . . . dealing with employee data management and protection . . . now stands at 4.2%').
57. Stefano and Taes (n 52) 16.

are one step towards that goal. Nowhere is this clearer than in the cohorts of workers organising together in data institutions and worker coalitions, often outside of labour union structures, that are actively aggregating and leveraging data in a kind of digital 'worker inquiry' in order to counter workplace narratives or audit algorithm behaviour.[58]

Information gathering, sensemaking, and communication are cornerstones of worker organising and coalition-building.[59] In the union context, data about worker conditions and other information allows workers to more effectively bargain with employers and can dramatically impact resulting employment agreements.[60] Predestining responses to modern algorithmic management, independent data collection and auditing were crucial to workers responding to scientific management in the twentieth century.[61] Historic practices such as auditing employer productivity measurements bear a striking resemblance to the 'worker data science' some modern workers use to audit gig economy platforms and produce collective knowledge.[62] 'Worker data science' refers to the practice of workers using data to develop a better understanding of their working conditions, audit technological systems to which they are subject at work, and build organising power through producing knowledge. Worker data science, digital workerism, and digital worker inquiry are all terms that represent variations on this central concept.[63] These projects, like data cooperatives, generally depend on workers collecting their own data about their work, rather than leveraging data already collected by an employer.

Participants in these projects collect their own data through a variety of means, usually through either low- or little-tech systems or by custom technology developed by advocates or

58. For an introduction to worker inquiry, see Jamie Woodcock, 'The Workers' Inquiry from Trotskyism to Operaismo: A Political Methodology for Investigating the Workplace' (2014) 14 Ephemera 489, 493.
59. Richard B Freeman and James L Medoff, 'What Do Unions Do' (1984) 38 ILR Review 244. Karl Marx's famous *Enquête Ouvrière* published in the *Revue Socialiste* in 1880 can be seen as a form of information gathering and worker's inquiry; see Hilde Weiss, 'Karl Marx's "*Enquête Ouvrière*"' in Kevin B Anderson and Bertell Ollman (eds), *Karl Marx* (Routledge 2012).
60. Nathan Newman, 'Reengineering Workplace Bargaining: How Big Data Drives Lower Wages and How Reframing Labor Law Can Restore Information Equality in the Workplace' (2017) 85 University of Cincinnati Law Review 693, 703 (unions leveraging information on worker 'voice' to understand both worker and employer interests); see also Freeman and Medoff (n 59) 65 (on fighting for the 'median' worker rather than accepting conditions amenable to the 'least attached' worker in a group). For more on worker voice and its historical role in labour unions, see Albert O Hirschman, *Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States*, vol 25 (Harvard University Press 1970).
61. Strategies in response to scientific management practices such as quantified 'time studies' included fighting for data transparency in order to audit employer wage calculations that used employer-collected work speed data; see Vera Khovanskaya and others, 'The Tools of Management: Adapting Historical Union Tactics to Platform-Mediated Labor' (2019) 3 Proceedings of the ACM on Human-Computer Interaction 1. For scientific management, see Frank B Gilbreth, *Primer of Scientific Management* (Van Nostrand 1914).
62. For a summary of 'worker data science' and related projects, see Karen Gregory, '"Worker Data Science" Can Teach Us How to Fix the Gig Economy' (*Wired*, 7 December 2021) <https://www.wired.com/story/labor-organizing-unions-worker-algorithms/> accessed 8 December 2021.
63. Digital worker inquiry and digital workerism both refer to a broader approach to political education through self-research, while 'worker data science' is more focused on the practice of using data analysis to further worker goals. For 'digital workerism' and 'digital worker's inquiry', see Jamie Woodcock, 'Towards a Digital Workerism: Workers' Inquiry, Methods, and Technologies' (2021) 15 NanoEthics 87; for 'digital workerism' related to computing specifically: Dan Calacci, 'Organizing in the End of Employment: Information Sharing, Data Stewardship, and Digital Workerism' (2022) *Symposium on Human-Computer Interaction for Work*.

researchers.[64] In the *Shipt Calculator* project, MIT researchers and worker-organisers co-designed a system that aggregated data from delivery workers to audit a platform's shift from a transparent pay scheme to a black-box algorithm, producing a report that revealed that the change cut the pay of over 40% of workers.[65] The tool was part of a coordinated campaign that culminated in a driver protest.[66] Similar projects that track working time and wage theft exist for other delivery platforms such as Deliveroo, as well as in more traditional employment relationships.[67]

However, it is not sustainable for worker groups to either build their own technological tools or depend on outside involvement when new topics of inquiry arise. Without shared resources, coalitions are left to rebuild social, organisational, and technical infrastructure with each new initiative. Small-scale projects like these are unburdened by a larger ambition to create governance or institutions and so reveal what workers want data for most: to document working conditions, ask questions of their workplaces, and build organisational power to access labour rights that may otherwise be difficult to exercise.[68]

One other way workers band together to share data is in the form of data cooperatives.[69] Data cooperatives, distinct from platform cooperatives, are legal constructs designed to facilitate the

64. For examples of low-tech tools being used to generate data on working conditions, see Lynn Dombrowski, Adriana Alvarado Garcia, and Jessica Despard, 'Low-Wage Precarious Workers' Sociotechnical Practices Working towards Addressing Wage Theft' (2017) Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems; tools like Turkopticon are built by researchers and created for worker use; see Lilly C Irani and M Six Silberman, 'Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk' (2013) Proceedings of the SIGCHI Conference on Human Factors in Computing Systems <https://dl.acm.org/doi/10.1145/2470654.2470742> accessed 13 July 2022.

65. For a description of the system, see Calacci and Pentland, 'Bargaining with the Black-Box' (n 10); for details on the report and advocacy, see 'Shopper Transparency Calculator 2.0' (*Coworker.org*) <https://home.coworker.org/shiptcalc/> accessed 10 March 2022.

66. Chris Mills Rodrigo, 'Shipt Workers Organize Most Targeted Protest Yet over New Pay Model' (*The Hill*, 15 October 2020) <https://thehill.com/policy/technology/521290-shipt-workers-organize-most-targeted-protest-yet-over-new-pay-model> accessed 18 August 2021.

67. For the Deliveroo system, see 'Deliveroo Unwrapped' (*Ridersroovolt.com*) <https://data.ridersroovolt.com/> accessed 10 January 2022. The Time Project is a work tracker for TV workers in the UK; see 'Time Project' <https://thetimeproject.co.uk/> accessed 10 March 2022; see also the WeClock project, an app that is like a 'strava for work': 'WeClock' <https://weclock.it/> accessed 10 January 2022.

68. For example, while wage theft is an enormous problem in low-wage work, it can be difficult to document. Projects that aim to make wage theft reporting easier are effectively working to facilitate better access to already-existing rights, rather than create new ones.

69. Data cooperatives are both technical and legal constructs, and so the obligations and structure of cooperatives vary widely by jurisdiction. Importantly, data cooperatives do not have to be entirely novel, separate entities from existing worker and citizen organisations; an existing credit union, labour union, or platform cooperative could each be a data cooperative if it implements the right technical or legal structures; see Alex Pentland and Thomas Hardjono, 'Data Cooperatives', in Alex Pentland, Alexander Lipton, and Thomas Hardjono (eds), *Building the New Economy* (MIT Press 2020) <https://wip.mitpress.mit.edu/pub/pnxgvubq/release/1> accessed 28 September 2020. For more on different forms of data governance, see Ana Brandusescu and Jonathan van Geuns, 'Shifting Power through Data Governance' (Mozilla Insights 2020) <https://foundation.mozilla.org/en/data-futures-lab/data-for-empowerment/shifting-power-through-data-governance/>.

pooling of data to collectively benefit their members.[70] The general operating principle of cooperatives (co-ops), rooted in agricultural collectives, is to return benefits to members on the basis of their membership.[71] A major obstacle for historic and modern co-ops is acquiring early investment or equity to fund operations.[72] Even if one does consider personal data a commodity, it is not nearly as liquid as the agricultural products that co-op structures were originally designed to support.[73] Marketing the personal data of co-op members is a difficult task, particularly with limited membership (which limits the scale of a co-op's data) and limited capital for engineering talent (which limits the development of data products that can be sold).

The goal of groups like Driver's Seat, a co-op of on-demand ride-hail drivers, is not to directly profit from driver data, but instead is to help drivers optimise earnings by gaining 'insights that are usually kept secret by employers like Uber.'[74] This is a significant challenge when one considers the capital-intensive operations intrinsic to data co-ops: data analysis, app development, data warehousing, etc. Innovations in cooperative law have enabled groups like Driver's Seat to maintain membership and capital by inviting outside investor-members, providing them the runway needed to transform their members' personal data into data products or other commodities.[75]

However, this fundamental dynamic leaves worker data co-ops in a precarious position. Activities such as performing third-party algorithmic audits of platforms or creating products that provide members with insights about their work conditions may benefit members, but they do not create revenue needed to sustain an association. Worker data co-ops are also in a disadvantageous position to sell products that are derivative of their members' data. The data that they can generally collect, such as a drivers' location history, is *also* already collected and organised by the platforms their members work for. Requiring data co-ops or similar data collectives to sustain themselves in an open marketplace without additional regulatory or institutional support limits their reach and ability to empower members.

70. Data cooperatives are distinct from 'platform cooperatives' in that they shift power to data subjects and generally offer transparency into data collection, use, and value, rather than just offering governance mechanisms; for example, FairBnB <https://fairbnb.coop>, a cooperative alternative to AirBnB, is a platform cooperative, but it does not strictly empower its users or the hosts of its platform to share in the value that their data provides. For a primer on platform cooperatives, see Trebor Scholz, 'Platform Cooperativism: Challenging the Corporate Sharing Economy' (Rosa Luxemburg Stiftung 2016).
71. The history of cooperatives as a legal structure have roots in agricultural collectives that sought to market the commodities of their patrons at the highest prices while keeping collective operating costs low; see James B Dean and Thomas Earl Geu, 'The Uniform Limited Cooperative Association Act: An Introduction' (2008) 13 Drake Journal of Agricultural Law 63, 71.
72. ibid 73.
73. Whether personal data is a commodity is a topic of thorough debate. Regarding its liquidity, some scholars have argued that personal data *only* becomes a commodity when transformed into 'big data' through significant processing. See Jim Thatcher, David O'Sullivan, and Dillon Mahmoudi, 'Data Colonialism through Accumulation by Dispossession: New Metaphors for Daily Data' (2016) 34 Environment and Planning D: Society and Space 990.
74. Brandusescu and van Geuns (n 69).
75. Driver's Seat is established as a Limited Cooperative Association (LCA) under Colorado law; LCAs allow for 'investor members' that are members by virtue of contributing capital as well as the usual 'patron members', significantly lowering the barrier to receiving investment to cover start-up costs, see Colo Rev Stat 7-58-405 (US).

2. A case study in the technical complexity of worker data autonomy

We have surveyed the existing legislation and strategies workers adopt to gain greater autonomy over technology use in the workplace through data access. While recent interpretations of data protection law expand the extent of workers' data access rights, data protection alone may not be sufficient to counteract the harms of algorithmic technologies used in the workplace. This section explores a case study that illustrates how the increasingly fragmented deployment of data-intensive systems in the workplace can render data protection-oriented approaches intractable or ineffective.

Certain forms of work that are more 'digitally legible' are subject to particularly high levels of workplace surveillance and algorithmic management.[76] These jobs are not entirely automatable, yet are repeatable enough to be continuously optimised through modelling and analytics. They extend the Taylorist concept into more complex modes of work, including warehouse jobs, platform work, and settings like call centres. In these contexts, third-party software (beyond core workplace systems) often provides business-to-business systems tailored to highly specified elements of management.[77] Call centres that focus on customer support or sales operations are subject to seamless loops of data capture, analysis, and optimisation.[78] Such algorithmic management processes require intermediation by third-party services, blurring the boundaries of storage, ownership, and processing. This fragmented technological landscape makes regulating workplace technology use from the perspective of data access particularly fraught.

2.1 Impacts on working conditions

The working conditions of call centre workers are influenced by algorithmic management on several levels. Workers might be held to monthly or quarterly quotas, weekly sales qualification or successful call quotas, and daily or even hourly call quantity figures. An additional level might monitor and respond to the outcomes of individual calls to match which targets might be assigned to which call centre employee based on past success of pairings between similar customers or similar workers.[79] Finally, a spate of machine learning powered tools may be used to analyse the speech and content of calls, coaching workers to change tone, the length of questions or responses and word order, or the content of what workers are meant to say.[80]

2.2 Contending with a shifting technical landscape

Each of these levels of algorithmic optimisation may occur through the combination of workplace systems, which in turn may store and process data both inside the boundaries of a workplace via

76. Legibility here refers to the ease with which a job or action can be quantified and predicted; most famously, the on-demand driver and delivery economy. See Jamie Woodcock and Mark Graham, *The Gig Economy: A Critical Introduction* (Polity 2019); also see Sarah Kessler, *Gigged: The End of the Job and the Future of Work* (St Martin's Press 2018).
77. Wolfie Christl, 'Digitale Überwachung Und Kontrolle Am Arbeitsplatz. Von Der Ausweitung Betrieblicher Datenerfassung Zum Algorithmischen Management? Eine Studie von Cracked Labs' (Cracked Labs 2021).
78. Jamie Woodcock, *Working the Phones: Control and Resistance in Call Centres* (Pluto Press 2016).
79. Min Kyung Lee and others, 'Working with Machines: The Impact of Algorithmic and Data-Driven Management on Human Workers' (2015) Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems <https://dl.acm.org/doi/10.1145/2702123.2702548> accessed 5 September 2022.
80. 'Gong for Sales Teams' (*Gong.io*) <https://www.gong.io/sales/> accessed 4 September 2022.

cloud providers, or operate entirely via licensed third-party tools, which in turn may store data with their own cloud services. Certainly, systems like these will most often rely on customer relationship management (CRM) tools, but will invariably also make use of enhancing applications from the independent application 'ecosystems' that orbit CRM software.[81]

Systems that were previously integrated within workplaces or organisations are now often fragmented and outsourced.[82] This service-based business model allows third-party software providers to capitalise on pooling data across their customer base to offer more powerful out-of-the-box ML and AI enabled software. For workers, however, it means the systems that shape working conditions are actually constituted by a constellation of technologies and services which use data sourced from workers or businesses that share nothing in common besides that their managers have elected to use similar software. While such a federated ML production process allows for the creation of more powerful models, the resulting technological landscape seems likely to deeply confound the exercise of agency for workers via data access rights by introducing conflicting rights to privacy and obscuring where and by whom data is held.

2.3 Data subject access requests and workplace quota transparency

Data subject access requests (DSARs) may seem a promising route to gain access to an individuals' records from their employer, particularly in the case of work done over the phone, as past case law preceding the GDPR has set precedent for such data to be considered the personal data of the worker.[83] However, even a successful DSAR may not provide clarity into how algorithmic decisions were made by an employer. If the models used by an employer belong to a third-party vendor, its details are outside the intellectual property claims of even the employer. Gong.io is one popular solution for call optimisation, allowing sales managers to review calls with automated content labelling for coaching on speech patterns. Though analysis of workers' calls may take place within the systems owned by that workplace, for example via a plug-in connected to the firm's CRM instance, Gong's audio analysis tools would only be licensed for use. Gong.io would be a 'processor' of employee personal data, and therefore required by Article 28 GDPR to assist the controller—the employer—to fulfil its obligation to data subjects. In practice, however, the technical and organisational complexity of the processing arrangement, combined with employers' tendencies to resist disclosing this kind of information, might lead to workers' data access rights not being fulfilled.[84]

2.4 Seeking access to models directly

Workers might also attempt to gain access to models themselves. Model parameters, statistical artifacts, and profiles constructed from personal data by ML currently become non-personal when

81. Amrit Tiwana, *Platform Ecosystems: Aligning Architecture, Governance, and Strategy* (Morgan Kaufmann 2014).
82. Tobias Fiebig and others, 'Heads in the Clouds: Measuring the Implications of Universities Migrating to Public Clouds' (*arXiv*, 27 July 2021) <http://arxiv.org/abs/2104.09462> accessed 20 June 2022.
83. Further, GDPR recital 63 extends coverage to the logics through which data were processed, perhaps even entitling some share of transparency to the models used on or trained through use of subjects' data; for case law on similar data, see *Copland vs United Kingdom* App no 62617/00 (ECtHR 2 April 2007).
84. For examples of how anonymisation can work against the interest of data subjects, see Omer Tene and Jules Polonetsky, 'Big Data for All: Privacy and User Control in the Age of Analytics' (2012) 11 Northwestern Journal of Technology & Intellectual Property 239, 263–264.

divorced from obvious identifiers such as a person's name.[85] These kinds of data, which have been creatively referred to as 'bastard data', are often the data that the most impactful algorithmic systems leverage, yet they are often excluded from current data protection regimes.[86] Some approaches propose to apply principles from differential privacy to further strip details from data in the name of principles of data minimisation.[87]

Despite attempts to sanitise data, recent research in ML security and vulnerability illustrates that models produced via deep learning predictably leak personal data when prompted strategically in adversarial attacks.[88] This revelation, even in a privacy-derived regime of data protection, provides strong arguments for model access by data subjects as it demonstrates models may retain and leak personal data used for training. However, model access or transparency is not necessarily in line with workers' goals when seeking autonomy over data they generate at work. Much of what might be meaningful to workers in contextualised data might not be clear until workers see the data itself.

Otherwise personal data can also be transformed into data which may be subject to labour-specific regulation. Consider, for example, transforming images or video into heart rate information, or using data about one's environment to infer emotional state.[89] On one hand, as long as these data are not anonymised or aggregated in such a way that they no longer 'relate to' any identifiable individual, they remain 'personal data' (and potentially among the 'special categories' of personal data, subject to the more stringent requirements set out by Article 9 GDPR)—and therefore, in theory at least, normal data access rights apply. Here again, however, the technical and organisational complexity typical of contemporary processing arrangements, combined with the by now relatively well-documented compliance and enforcement deficits with respect to workplace data protection rights,[90] pose significant practical challenges for workers aiming to gain access to these inferred or transformed data. And, in the EU context at least, labour and Member State legislation typically does not aim to create or clarify rights of access for them.[91]

85. Although pseudonymous data (data without explicit identifiers) is considered personal data under Article 2(5) GDPR, once transformed using common statistical techniques it often no longer becomes pseudonymous and no longer falls under the usual categories defined under Article 4(1) as identifiable, either directly or indirectly.
86. 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.
87. Abigail Goldsteen and others, 'Data Minimization for GDPR Compliance in Machine Learning Models' (2022) 2 AI and Ethics 477.
88. Nicholas Carlini et al, 'The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks', *(arXiv*, 16 July 2019) <http://arxiv.org/abs/1802.08232>.
89. See, for example, Daniel McDuff, Sarah Gontarek, and Rosalind Picard, 'Remote Measurement of Cognitive Stress via Heart Rate Variability' (2014) *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*; for prediction of emotional state using environmental data, see Cristina Bustos and others, 'Predicting Driver Self-Reported Stress by Analyzing the Road Scene' (2021) *9th International Conference on Affective Computing and Intelligent Interaction*.
90. Justin Nogarede, 'No Digitalisation without Representation: An Analysis of Policies to Empower Labour in the Digital Workplace' (2021) Foundation for European Progressive Studies Policy Study November 2021, 15–18 <https://feps-europe.eu/wp-content/uploads/2021/11/No-digitalisation-without-representation.pdf> accessed 13 December 2022.
91. An exception to this may be the proposed EU Platform Work Directive—emotional data inferred from other data is still emotional data, and so may be subject to the directive's restrictions. For legislation outside the EU, see Antonio Aloisi and Valerio De Stefano, 'Between Risk Mitigation and Labour Rights Enforcement: Assessing the Transatlantic Race to Govern AI-driven Decision-making through a Comparative Lens' (elsewhere in this issue).

The increasingly complex and modular multi-agent workflow for training and deploying ML models complicates the process of determining where subjects' data ends and models begin. Notably, Wachter and Mittlestadt called for a right to reasonable inferences, but the distributed footprint of ML systems among several companies may make explanations of inferences difficult to achieve.[92] Martens, for instance, draws attention to the 'wide legal no-man's land' that surrounds model access, which results from the effects of the 1996 Database Directive's intersection with GDPR.[93] In the present technical context ruled by an ecosystem of microservices and third party SaaS providers, the circulation of data in raw, 'bastard', or model form can be impossible to trace or disproportionately costly to track down. Together, the complications added by ML to the equation of worker data access make strong arguments for regulating data as neither assets nor private information, but rather as a condition of work which shapes employer-worker relations. Though some researchers have put forward approaches which delegate the right to lodge DSARs in order to aggregate data and access models, the technical complications elaborated here reflect that DSARs still suffer significant technical and institutional obstacles in empowering workers.[94]

Finally, call centre workers may have strong claims to use the labour-specific laws discussed in this article's introduction, as they do not premise access on identifiability or other similar requirements.[95]

2.5 Worker self-inquiry

If workers are unable to access data through the rights defined in data protection regulation, perhaps self-inquiry might yield greater success. Self-tracking apps have been used in platform work to help workers track their mileage, expenditure, pay rates, and to optimise their income across platforms. Further, tools like WeClock can be used in tracking movements to reveal harmful working conditions on warehouse floors.[96] However, the technical landscape of call centres poses a number of difficulties for self-tracking activities.

Though call centre work presents a case in which goal attainment over time and general performance figures are often made available to workers as a tactic to gamify work, provoke competition, or motivate quota attainment, those data (even in aggregate) are unlikely to be sufficient as the basis for organised worker inquiry.[97]

Both the technical nature of the algorithms used and the institutional dynamics of third-party ML tools restrict workers' ability to gain sufficient data access to address algorithmic management

92. Sandra Wachter and Brent Mittelstadt, 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI' (2019) 2 Columbia Business Law Review 494.
93. Bertin Martens, 'The Importance of Data Access Regimes for Artificial Intelligence and Machine Learning' [2018] SSRN Electronic Journal <https://www.ssrn.com/abstract=3357652> accessed 15 October 2022.
94. Hadi Asghari, Thomas van Biemen and Martijn Warnier, 'Amplifying Privacy: Scaling up Transparency Research through Delegated Access Requests' (*arXiv*, 12 June 2021) <https://arxiv.org/abs/2106.06844>.
95. Recent case law, specifically *Nowak* (n 11), points out that data protection is meant to be understood under a 'wide scope' of data: 'EU legislature to assign a wide scope to that concept, which is not restricted to information that is sensitive or private, but potentially encompasses all kinds of information, not only objective but also subjective, in the form of opinions and assessments, provided that it 'relates' to the data subject.'
96. 'WeClock' <https://weclock.it/> accessed 23 November 2022.
97. Janaki Kumar, 'Gamification at Work: Designing Engaging Business Software' in Aaron Marcus (ed), *Design, User Experience, and Usability. Health, Learning, Playing, Cultural, and Cross-Cultural User Experience* (Springer 2013) <http://link.springer.com/10.1007/978-3-642-39241-2_58> accessed 15 October 2022.

through self-inquiry. Technically speaking, models and algorithms—for instance, to recognise speech patterns amongst call centre employees—will make it unachievable for individual workplaces to gather a sufficiently representative data set themselves to invert algorithms used for management. Instead, attempts to invert or clarify models via data aggregation would likely require data sourced amongst all the workplaces from which the third-party providers' model was derived. Looking forward, the contracts necessary for workplaces to license third-party tools, and the interoperability such tools require, provide handy opportunities for workers to advocate for greater access in instances where they do have bargaining or co-determination available.

3. Collectively regulating the collection and use of worker data

Workers are leveraging data protection law and other instruments to gain meaningful access to information known about them by employers to exercise greater autonomy at work. This section argues that the rights and instruments available to workers should reflect this goal, rather than a general right to privacy or data protection. To do this, this section outlines technical and legal steps that can help facilitate meaningful worker data access and agency. Section 3.1 lays out our broad argument, informed by our analysis in sections 1 and 2. Section 3.2 identifies specific ways the GDPR can be expanded or altered to increase its usefulness in the employment context. Section 3.3 discusses how labour law and data protection law could be hybridised through instruments like the Platform Work Directive to benefit workers. Finally, Section 3.4 explores how certain legal tools, including data trusts and works councils, could be adapted to provide increased worker protections and strengthen the workers' voice regarding data and technology use at work.

3.1 Principles for worker data regulation

Workers' legal rights should reflect their interest in using their data to exert agency over their working conditions. Platform workers collect data independently from their employer to organise more effectively.[98] On-demand drivers leverage collective subject access requests to hold employers' use of technology to account. Meanwhile, traditional workers use data about their workday to document working conditions more generally.[99] These actions demonstrate that the primary reason workers aim to use the data they produce at work is to exert increased agency, rather than react to perceived privacy harms or concerns. This pattern is consistent with modern critiques of data collection and uses that frame data as part of a collective political economy instead of as personal information about individuals.[100]

This is not to suggest that regulation should ignore the dignitarian or material harms that arise from invasive data collection at work.[101] We address this below by discussing the potential for

98. See ns 62, 65, and 67 (worker data science projects that independently collect and analyse platform worker data).
99. See ns 67 and 96 (tools that allow more traditional workers to collect and aggregate data).
100. Cohen, *Between Truth and Power* (n 2) 49 ('The raw materials consist of data identifying or relating to people, and the public domain made up of those materials is biopolitical—rather than, say, personal or informational'). See also Zoe Adams and Johanna Wenckebach, 'Collective Regulation of Algorithmic Management' (elsewhere in this issue).
101. Ajunwa, Crawford, and Schultz, 'Limitless Worker Surveillance' (n 3) (on the dignitarian and material harms of privacy breaches under employment).

private rights of action based on 'privacy harm' (rather than regulatory breach) in the employment context.[102] We argue for a hybrid approach, based on existing behaviour of workers and advocates.

First, tools and instruments aimed at regulating data and technology in the workplace should primarily focus on creating circumstances that increase worker power, rather than defining *ex ante* harms or assigning liability *ex post*. This perspective values approaches that increase worker participation in technology use decisions in the workplace.

Second, ideas of data access at work should be replaced with the goal of data understanding. This can be accomplished by expanding GDPR's access rights and explicitly extending the range of rights mandatable to worker representatives in the EU and US.

Third, worker co-determination and worker representation should be a primary way that technology is regulated in workplaces. This general principle can help address the collective nature of data governance, create flexible data protection regulation in EU Member States, and solve the *ex ante* problem of defining appropriate data use for all labour contexts.[103]

3.2 Expanding the GDPR for workers

Purtova has outlined an impending state of affairs in which personal data is collected everywhere, and as a result GDPR and other data protection legislation become 'laws of everything' too.[104] In the workplace, datafication has happened at an even faster rate—with workplace systems surrounding everyone from salaried knowledge workers to the most casualised platform workers with loops of surveillance, optimisation, and automation. In this regard, the constellation of technologies that now populates the workplace have brought about Purtova's future much earlier than in the consumer context, while the consumer-oriented, privacy-focused data protection regulations she mentioned have yet to evolve with them. Accordingly, workers' needs for meaningful data access are urgently underserved by GDPR as it stands, and there is a pragmatic reason for its extension.

3.2.1 Embrace some extensions to the GDPR and the legal basis of data protection. Many scholars warn against expanding GDPR for fear of diluting its protections, or that GDPR's focus on personal data would preclude the ability of its potential extensions' potency from the get-go.[105] On the other hand, instruments such as those promulgated under Article 88 GDPR[106] and the proposed Platform Work Directive do extend worker data rights in a manner that fits into the existing regime established by the GDPR.

The Platform Work Directive's approach is particularly promising, given its side-stepping of the primary arguments presented against the expansion of data protection rights. Purtova notes that

102. For more on harm-based liability rather than regulatory liability in the context of information and privacy, see Ignacio Cofone, 'Beyond Data Ownership' (2021) 43 Cardozo Law Review 501, 556–557 ('For liability to be most effective, private rights of action must be based on harm, not based on regulatory breach').
103. Phoebe Moore for the European Parliament Panel for the Future of Science and Technology, 'Data Subjects, Digital Surveillance, AI and the Future of Work' (European Parliamentary Research Service Scientific Foresight Unit 2020) <https://www.europarl.europa.eu/thinktank/en/document/EPRS_STU(2020)656305> accessed 17 October 2022.
104. Nadezhda Purtova, 'The Law of Everything: Broad Concept of Personal Data and Future of EU Data Protection Law' (2018) 10 Law, Innovation and Technology 40.
105. Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701.
106. See Abraha (n 26).

'scholars have been critical of the concept of personal data as growing too broad,' primarily because of the 'concept of personal data, ie identifiability of a person, and the corresponding strand of technological development, ie re-identification and de-anonymisation algorithms.'[107] To the contrary, the Platform Work Directive lists personal data as only one category of data among others that should be specifically regulated. While today, stripping data of its identifiable features may be a commonplace tactic, under the Platform Work Directive, whether or not it is personal would not determine whether the law applies.

3.2.3 Decouple identifiability from access. The decoupling of identifiability from access rights is a step forward for which the Platform Work Directive provides precedent, yet it leaves essential elements of data access in models unaddressed. Most important among them are workers' access to models and AI systems. We mentioned earlier that ML model access under data protection regimes tends to focus on inferences and their explainability. When regarded through the lens of collective working conditions, however, such individual explanations of isolated decisions are similarly inadequate as those of using only individual personal data to clarify data which is only meaningful in aggregate form. Indeed, such explanations might be self-fulfilling or protected by intellectual property laws. In workplaces making extensive use of external software services, it is unlikely to be clear to workers (or even managers) *when* ML-driven decisions are being made. Indeed, many third-party SaaS systems, including Gong.io, discussed in the case study in section 2, illustrate that ML training may be happening outside the scope (or the knowledge) of workplace managers themselves. Together, these technical and legal ambiguities demand clear terms for groups of workers and their representatives to access models trained on data collected from their actions or which impact their working conditions.

Machine learning models frequently cross institutional bounds of individual workplaces when purchased as software or services. When machine learning services are purchased from external providers by management, model, parameter, and training data access rights cannot be meaningfully fulfilled by management alone. As models are increasingly used across workplaces, sectoral, platform-specific, and regional worker cooperation becomes more necessary. Existing scholarship emphasises the need for data subject controls over which data can be used for model training.[108] However, stripping models of identifying data does little to ameliorate alterations they may make to workplace conditions. Approaches using collective bargaining, as we will explore later in this section, may provide a meaningful alternative.

3.2.3 Extend third-party and authorised agent capabilities to enable trusts. One under-explored approach to collective data protection and governance is the concept of data trusts, especially for extensively surveilled workers.[109] Data trusts are mechanisms and legal instruments that ensure that personal information collected for one purpose can be shared and used for a different

107. Purtova, 'The Law of Everything (n 104) 41.
108. Goldsteen and others (n 87).
109. See Part 1D of this article, and also Chris Reed and Irene Ng, 'Data Trusts as an AI Governance Mechanism' (2019) SSRN Electronic Journal <https://papers.ssrn.com/abstract=3334527> accessed 15 October 2022 and Jack Hardinges, 'What Is a Data Trust?' (*The ODI*, 10 July 2018) <https://theodi.org/article/what-is-a-data-trust/> accessed 15 October 2022.

purpose, while protecting an individual's interest in privacy and other fundamental rights.[110] This is primarily accomplished through trust law, mainly present in the UK, US, and Canada.[111] Trusts have historically been used to appoint a steward (trustee) to 'manage an asset for a purpose… on behalf of a beneficiary or beneficiaries who own the asset.'[112] In the case of a data or digital trust, the assets held by the trust are a subset of beneficiaries' digital rights—such as data rights—or actual digital assets, such as code or data itself.[113]

Trusts are attractive for collectively managing data because they allow beneficiaries to pool the rights they have over their personal data while maintaining the legal option to hold trustees partially liable if the data is misused.[114] However, data trusts are still a nascent tool for consumers as well as workers interested in exerting control over their data. Although the concept of using trust law for data has been discussed at least since 2004, data trusts actually in operation are difficult to find.[115] In a March 2022 report, the Open Data Institute (ODI) noted that they were unable to find a candidate data trust mature enough to have had meaningful impact.[116]

So, we are left to speculate: how might workers use data trusts in practice to collectively regulate data and technology use in the workplace? Trusts could be established for workforces that are highly surveilled. Such a trust would follow the ODI's definition, where the data itself is held by a data controller (in this case, likely the employer) from which the trust is independent.[117] Trustees appointed by workers would take on a legally binding responsibility to ensure that the data is used and shared (with employers) for the benefit of the workers. The actions of the trust could be limited to labour-related actions, to avoid commercial conflicts of interest such as the trust sharing or selling data to a third-party buyer (while sharing profits with beneficiaries) that, for example, reveals an employer trade secret.[118] Ideally, this would be a civic trust that includes governance mechanisms controlling the trustee's decisions.[119] This arrangement could be set up within a specific company or more widely across workers within a specific sector, such as ride-hailing.

This arrangement would provide obvious benefits. First, workers could gain significant bargaining power: workers with democratic control over how their data is used through a civic data trust

110. This definition is paraphrased from Reed and Ng (n 109) and Hardinges (n 109).
111. Brandusescu and van Geus (n 69).
112. Sean Martin McDonald, 'Reclaiming Data Trusts' (*Centre for International Governance Innovation*, 5 March 2019) <https://www.cigionline.org/articles/reclaiming-data-trusts> accessed 25 November 2019.
113. ibid.
114. Jack Hardinges and others, 'ODI Report: Data Trusts: Lessons from Three Pilots' (Open Data Institute 2019) <https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit?usp=sharing&usp=embed_facebook> accessed 17 August 2021.
115. Lilian Edwards, 'The Problem with Privacy' (2004) 18 International 263.
116. Aditya Singh and Jack Hardinges, 'Measuring the Impact of Data Institutions' (Open Data Institute 2022) <https://theodi2022.wpengine.com/wp-content/uploads/2022/03/2022_ODI_Measuring-the-impact-of-data-institutions.pdf>.
117. For ODI's definition, see Hardinges and others (n 114) 6; as Delacroix and Lawrence note, separating the data trust from the controller is also crucial to maintaining a trust's fiduciary obligation to its beneficiaries in practice: Sylvie Delacroix and Neil D Lawrence, 'Bottom-up Data Trusts: Disturbing the "One Size Fits All" Approach to Data Governance' (2019) 9 International Data Privacy Law 236, 241.
118. In an even more convoluted arrangement, the employer (also the data controller) could be named as a beneficiary as well, although this could create obvious problems: Kieron O'Hara, 'Data Trusts' (2020) 6 European Data Protection Law Review 484, 487.
119. McDonald (n 112).

could gain a kind of data leverage that would provide meaningful power in a relationship with a data-dependent employer.[120] Second, workers could use the data collected by their employer to run their own analyses and reports to understand their collective conditions without collecting their own data, making the labour negotiating process much more efficient. Third, although a trust would not by default give workers the right to co-determine what kinds of data collection or algorithms they are subject to, it could end up having that effect. Data collected through worker surveillance is not useful to an employer if workers do not agree to the employer using that data, a precondition to employer access under a data trust regime.

However, for data trusts to succeed, data protection laws must include certain affordances. The most important of these is allowing data subjects to grant third-party entities the ability to exercise a limited set of their data protection rights in particular circumstances.[121] Such circumstances could be limited to entities that have a fiduciary relationship to the data subject, as in the case of trusts, or those that have an existing 'trusted' relationship to the data subject, such as a union.[122]

The opening clause in Article 88(1) GDPR permits collective agreements to address data protection within EU Member States and encourages regulatory 'experimentation'. Member State regulations based on Article 88 GDPR should test provisions that allow data subjects a limited ability to transfer some rights under GDPR, making some rights 'mandatable'. Scholars exploring this idea have noted that it is unlikely that national legislation making certain rights mandatable would be in breach of EU law.[123] This would allow workers to focus on the specific values and goals of negotiations in the workplace, while leaving decisions about how those values translate into specific workplace data decisions to a trustee.

The intermediation of data processing and storage in the technical example of section 2 makes access to data that would be placed in trusts difficult in the present regulatory paradigm. However, the high degree of interoperability in CRM ecosystems could simplify future access for worker collective action. Data trust infrastructure could piggyback off of standardised data formats and integration standards already in place for CRM 'ecosystems', removing some technical boundaries for data trusts.

A worker data trust would also require significant reporting and transparency requirements about how data is stored, used, and shared to carry out its fiduciary duties.[124] A related problem in the employment context is that of meaningful consent. It is unclear how civic worker trusts could

120. Data leverage, specifically in the consumer context, refers to the ability to withhold or change data sharing or access to achieve a desired goal; here, it could involve workers collectively withholding collected data if an employer refuses to negotiate over, e.g., a new productivity scoring algorithm; see: Nicholas Vincent and others, 'Data Leverage: A Framework for Empowering the Public in Its Relationship with Technology Companies' (2021) Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency.

121. Article 80 GDPR affords data subjects the right to mandate a third party the ability to exercise only a small subset of available rights, mainly relating to judicial remedy and complaints; for data trusts to be practical, other rights must be easily mandatable; see Article 80(1) GDPR. On the transference of the right to portability, see Gill and Metzger (n 14) 16 ('. . . the text does not mention any possibility to transfer the right').

122. Circumstances such as having an existing, trusted relationship with a trustee such as a union or works council in a clear context circumvents the paradox of trusts outlined in O'Hara (n 118) (pointing out that while trust frameworks assume subjects are unable to provide informed consent to a data controller due to insufficient knowledge, they assume that subjects are able to understand their relationship with the trustee).

123. Delacroix and Lawrence (n 117) 236 ('such intervention would not necessarily extend (nor diminish) the scope of controllers' obligations').

124. However, proposed regulations such as AB-1651 already impose reporting requirements on employers that would undoubtedly be similar to those required in a trust arrangement; see Gill and Metzger (n 14) 15–17.

implement democratic governance systems that can work flexibly within the employment context.[125] Governance systems would need to balance providing worker control over trustee decisions while meeting the needs of the modern employment relationship.[126] Solving this tension might require new technologies for data governance—a major barrier, particularly when such trusts might span multiple workplaces or operate on the trade level.[127]

3.3 Hybridise labour law and data protection

How expanded data protection provisions should be made available to workers is still being decided. California's AB-701 and AB-1651 offer specific data access rights and protections based on the legal basis of workers' labour rights, although they are inspired by consumer data protection law. This approach provides certain affordances and protections that are crucially needed within the workplace. Like the proposed EU Platform Work Directive, AB-701 provides additional protections and procedures for even more specific types of data and areas of work.[128]

The proposed Platform Work Directive has a 'dual' legal basis. It is based broadly on Article 153 of the Treaty on the Functioning of the European Union (TFEU), which empowers the EU to take measures to improve working conditions. However, the provisions relating to algorithmic management have an additional legal basis in Article 16 TFEU, which empowers the EU to take measures to protect individuals' personal data. As a result, the Platform Work Directive's provisions regarding workplace technology use are framed fundamentally as data protection rights.[129] For example, one of the most stringent restrictions placed on employers by the Platform Work Directive stipulates that digital labour platforms may not 'process any personal data on the emotional or psychological state of the platform worker.'[130] The intent of this section is likely to limit the emotional manipulation of workers, but the current draft text, animated by its 'data protection' orientation, leaves open questions about common management practices that are arguably emotional or psychological manipulation, but take place without processing data explicitly 'about' the emotional or psychological state of the worker.[131] For example, a worker's location data, required by delivery platforms,

125. Valerio De Stefano, '"Negotiating the Algorithm": Automation, Artificial Intelligence, and Labor Protection' (2019) 41 Comparative Labor Law and Policy Journal 15, 32 ('Managerial control that 'skips' employee consent to make decision-making under uncertainty more efficient is 'one of the key legal bricks of the modern firm').
126. For additional thoughts on this tension more generally, see the discussion of consent in Christine Rinik, 'Data Trusts: More Data than Trust? The Perspective of the Data Subject in the Face of a Growing Problem' (2020) 34 International Review of Law, Computers & Technology 342, 347–348.
127. Dan Wu and others, 'How Data Governance Technologies Can Democratize Data Sharing for Community Well-Being' (2021) 3 Data & Policy e14, 8.
128. See n 35 (on California AB-701).
129. For the basis of the Platform Work Directive, see specifically section 2 of the Platform Work Directive explanatory memorandum, specifying that the directive is based in TFEU Art. 153(1)(b) and Art. 153(2)(b)); see also Consolidated Version of the Treaty on European Union [2008] OJ C326/47 Art. 153(1)(b) (granting the EU power to 'support and complement' activities of Member States in the improvement of working environments and working conditions).
130. Platform Work Directive, art 6(5)(a).
131. Emotional manipulation is a hallmark of digital platform work; see Valerio De Stefano and Wouters M, 'Embedding Platforms in Contemporary Labour Law' in Jan Drahokoupil and Kurt Vandaele (eds), *A Modern Guide To Labour and the Platform Economy* (Edward Elgar Publishing 2021) 129 and Ifeoma Ajunwa, 'Algorithms at Work: Productivity Monitoring Applications and Wearable Technology as the New Data-Centric Research Agenda for Employment and Labor Law' (2018) 63 St Louis University Law Journal 21.

could be used to estimate if a worker is in a surge area or near their home. Should offering the worker an incentive to continue working be prohibited, even though the decision involved data necessary to the contract that would not be categorised as an emotional inference? On the other hand, what about workers for whom emotional data could be argued to be 'strictly necessary', such as teletherapists or healthcare workers?[132]

In the American context, there is further need for an approach that seriously considers both labour law and data protection law. Several recent examples in US case law concretely illustrate how state consumer data protection laws clash with federal labour law when workers try to exercise their rights. Illinois' Biometric Information Privacy Act (BIPA) regulates the use and handling of individuals' biometric data and allows for civil suits in cases where a collector mishandles that data.[133] The state has since seen over a dozen class action cases filed against employers on behalf of employees claiming a mishandling of biometric data.[134] In 2022, an employee of Roosevelt University in Illinois filed a civil action suit seeking damages, claiming that the university had not obtained informed consent when collecting his handprint as evidence of his clocking in to work.[135] In this case and others, the state court has held that for unionised workers, the legal question becomes whether the union's CBA includes biometric information and whether the employee exhausted available grievance procedures, not whether the employer violated BIPA.[136]

This raises two main problems for unionised employees working under state privacy laws in the US. First, it effectively exempts the employment relationship from privacy claims brought under state law. Most CBAs include broad management rights clauses that, like in *Walton*, cover most terms and conditions of employment, whether mandatory or permissive.[137] Current labour law interpretations in the US do not consider privacy separately from other terms and conditions of employment.[138] This means that American regulations relying on *ex post* claims of harm are basically toothless in the employment context: these claims must be settled through grievance procedures governed under a CBA, not by rights of action. Unionised workers without data protection or surveillance provisions in their CBAs are then left unprotected.

Second, legal provisions allowing authorised agents to provide consent on behalf of workers may actually limit worker agency if implemented poorly. In a 2020 case, also in

132. For more on this and other open questions in the proposed Platform Work Directive text, see Michael Veale, M Six Silberman, and Reuben Binns, 'Fortifying the Algorithmic Management Provisions in the Proposed Platform Work Directive' (elsewhere in this issue).

133. See BIPA (740 ILCS 14/1) generally, and BIPA sec 14/20 for rights of action specifically.

134. There are several relevant opinions in Illinois state case law that deal with an employer's obligations regarding employee biometric data under BIPA. While we will not list all of them here, we direct the reader to, for example, *Bryant v Compass Group USA,* 958 F3d 617 (7th Cir 2020); *Fernandez v Kerry*, 14 F4th 644 (7th Cir 2021); *Miller v Southwest Airlines*, 926 F3d 898 (7th Cir 2019). For an overview of state litigation related to BIPA, see Michael McMahon M, 'Illinois Biometric Information Privacy Act Litigation in Federal Courts: Evaluating the Standing Doctrine in Privacy Contexts' (2021) SSRN Electronic Journal <https://papers.ssrn.com/abstract=3929645> accessed 15 October 2022.

135. *Walton v Roosevelt University*, 2022, IL App (1st) 210011.

136. ibid 21. In general, this is due to the fact that US federal labour legislation pre-empts any claim under state law brought by employees who are represented by a union. The two relevant pieces of federal legislation are the Railway Labor Act (RLA), 45 USC sec 152 and the Labor Management Relations Act, 29 USC sec 185(301) (LMRA).

137. ibid 9.

138. *Miller* (n 134) para 904 ('That biometric information concerns workers' privacy does not distinguish it from many other subjects, such as drug testing, that are routinely covered by collective bargaining and on which unions give consent on behalf of the whole bargaining unit.').

Illinois, an employee argued that only she could provide meaningful consent under BIPA, not her union.[139] This claim was rejected. BIPA allows for an 'authorized agent' to consent on behalf of an individual, and unions act as workers' 'sole and exclusive' agents under federal labour law.[140] This meant that the union's apparent consent through their agreement's management rights clause overrode any individual members' lack of consent.[141]

Although unionised workers' rights to action under BIPA are limited, non-union workers seem to have standing. According to another recent case, *Figueroa v Kronos,* non-union employees' claims regarding consentless biometric data collection are not pre-empted by any federal labour law, and so will stay.[142] Interestingly, *Figueroa v Kronos* also demonstrates that BIPA allows employees a right of action against third-party collectors that an employer contracts with, such as AI and surveillance vendors.[143] The Court held that to avoid liability, vendors must require an employer to obtain written consent before using their system to collect employee data, demonstrating at least one coherent approach to liability in employment contexts.[144]

When federal data protection legislation is enacted in the US, it must explicitly pre-empt parts of federal labour law.[145] Without doing so, it risks severely hampering workers' right of action, a core instrument of proposed US data protection policy.

3.4 Incentivise worker co-determination through altering rules

How can the EU overcome its limited competencies regarding labour law to ensure data collection and use protections for workers? Worker co-determination should be considered a serious solution.[146] Article 88(1) GDPR and the proposed EU Platform Work Directive each enable collective agreements to regulate employee data processing.[147] However, the involvement of worker representatives in negotiating technology use in workplaces varies significantly between EU Member States.[148]

Co-determination should be valued over union representation, as unions often have limited bargaining rights over decisions core to business operations, such as technology introduction and use. Co-determination also provides worker representatives with important information that is usually readily available only to employers, which is crucial for governing worker data use. German and Austrian works councils are of special note, as they already have the general responsibility and right not only to be informed of employer operations, but also to veto and participate in meaningful decision-making regarding technology at work—not just data protection.[149] Co-determination

139. *Peatry v Bimbo Bakeries USA,* No 19 C 2942, 2020 WL 919202, at *3–4 (ND Ill February 26 2020).
140. See BIPA, sec 14/15 (for the authorised agent provision) and NLRA, secs 158(1)(5), 158(2) and 159(a).
141. *Peatry* (n 139) *3–4.
142. *Figueroa v Kronos*, 454 F Supp 3d 772 (ND Ill 2020).
143. ibid *783.
144. ibid.
145. Specifically, see n 136.
146. Moore (n 103) (several sections argue explicitly for expanded or even mandated worker co-determination across Member States as a way to regulate workplace data use and collection).
147. Platform Work Directive.
148. Abraha (n 26).
149. ibid 12. German works councils hold dual co-determination at both the shop floor level and the company level, where worker representatives hold positions on company boards; see Moore (n 146).

should be encouraged or incentivised by EU Member States as a mechanism to enforce flexible and effective workplace data protection regulation.[150]

One way this could happen is by using a default approach.[151] Employer data processing could be subject to a set of default requirements and restrictions, such as those laid out in the Platform Work Directive. Article 88 GDPR provides for such a default, granting a floor to employee data protection rights through GDPR that can then further be defined by Member States. The trick is to design 'altering rules' that grant employers and sectors the ability to opt out of certain additional provisions that may be defined at Member State level when they participate in some desired behaviour, such as collective bargaining.[152] In this case, the desired behaviour would be implementing worker representation and co-determination regarding (at minimum) workplace data collection and use. This approach has been argued as a mechanism to incentivise workplace governance more generally in US labour relations, and there is no reason it would not also work at Member State level in the EU.[153]

This avoids the EU directly regulating labour relations in Member States and so could avoid overstepping the EU's supporting competencies.[154] Such altering rules could be made available by Member States to firms or sectors where it is determined that meaningful worker representation is present. The provisions employers should be able to opt out of through such altering rules should be those that pose a burden on employers but do not impact the floor of fundamental rights granted to workers under the GDPR or a directive like the Platform Work Directive. For example, additional reporting and impact assessment requirements established by Member State legislation enacted under Article 88 GDPR could be made optional in the case that worker governance structures implement suitable and adequate measures for monitoring and mitigating the impact of automated decisions.[155] Similar solutions have been shown to work in areas of workplace safety, where employers in the US with 'adequate internal compliance programs' receive less stringent enforcement under federal and state regulation.[156]

While this could help provide new forms of worker co-determination across EU Member States, how might it empower workers with greater agency over workplace technology use? Existing reports provide some clues, and we summarise some of the requirements that the EU could mandate of such worker governance groups below.[157] First, employers should, as is outlined in Platform Work Directive Article 9, have information and consultation obligations to worker representatives. As we argue above, these obligations should include more than just information about use of data, but also include access to workers' personal and contextual information. This obligation

150. Moore (n 146) p 89; Adams and Wenckebach (n 100).
151. Default approaches are probably most commonly understood in the context of US contract law: Steven J Burton, 'Default Principles, Legitimacy, and the Authority of a Contract Symposium on Default Rules and Contractual Consent' (1993) 3 Southern California Interdisciplinary Law Journal 115.
152. Ian Ayres, 'Regulating Opt-Out: An Economic Theory of Altering Rules' (2011) 121 Yale Law Journal 2032.
153. Brett H McDonnell and Matthew T Bodie, 'From Mandates to Governance: Restructuring the Employment Relationship' (2021) 81 Maryland Law Review 887.
154. A closer analysis of this question is needed, but such an instrument would not violate TFEU article s153(4) or 153(6), and would appear to be within the competencies granted through Articles 153(1) and (2). See Consolidated Version of the Treaty on European Union [2008] OJ C326/47.
155. See Platform Work Directive arts 7(1)–(3).
156. McDonnell and Bodie (n 153) 943.
157. See Moore (n 146) 88–94 for clear ideas on how collective governance in the workplace should apply to data governance.

is one important step to empowering worker groups to detect harm or data misuse and to continually develop a clear understanding of working conditions.[158]

Second, worker representatives should have an active role in defining and negotiating what data is 'necessary' for carrying out work as referenced in Article 6(5) of the Platform Work Directive or Article 6(1)(b) GDPR. This definition is crucial for establishing common understanding regarding data processing and protection between workers and employers and establishing context-specific data rules.

Third, collective consent and rights approaches should be considered, whereby, e.g., worker representatives are able to provide 'collective' consent on behalf of the workers they represent.

Fourth, worker representatives could be considered trustees of worker data by default. These provisions, along with meaningful co-determination with respect to 'platforms' like Salesforce or other central systems for data aggregation in the workplace, would provide strong protections regarding use of worker data while reducing undue burdens regarding data access and use that employers might face under individual consent or constant data governance regimes.[159]

3.5 Ex post *approaches*

Ex ante worker governance may not prevent employers from using employee data in ways that disempower workers. Can *ex post* liability approaches offer additional protections? We discuss some ways to handle the main harms that can occur due to data use and collection in the workplace below: harms due to privacy violations, including invasive inferences, and harms due to population-level disparate impact.

Harms due to data collection and use are notoriously difficult to both characterise and detect.[160] Tort claims in the US brought due to privacy violations must generally demonstrate significant injury; violations of expectations to privacy do not meet muster.[161] As an alternative, the recent concept of 'privacy harms' is an attractive way to model how both data collection and its use can cause harm.[162] Privacy harms effectively operationalise the idea of privacy as an issue of context and social norms.[163] The general idea is that party A is subject to a privacy harm if information about them that was disclosed or inferred by party B increases their, or another party's, understanding of A more than A expected.[164]

158 See also Adams-Prassl and others, 'Regulating Algorithmic Management: A Blueprint' (elsewhere in this issue), especially 'Policy Option 7.'

159. For example, a trust model would allow for quick, workplace-level decision-making regarding worker data use while not requiring employers or worker representatives to collect the consent of every individual worker.

160. For a more general approach, see Julie E Cohen, 'Examined Lives: Informational Privacy and the Subject as Object' in Paul Schiff Berman (ed) *Law and Society Approaches to Cyberspace* (Routledge 2007); for specific arguments as to why tort law and strict liability do not sufficiently address privacy harms, see Omri Ben-Shahar, 'Data Pollution' (2019) 11 Journal of Legal Analysis 104, 129–131. This is also the case under the GDPR: German courts have rejected GDPR violation claims under articles 79 and 82 that do not present material damages; see: Amtsgericht [local court] 7 November 2018 8 C 130/18 AG Diez <https://openjur.de/u/2116788.html>.

161. The exception to this is in recent cases in Illinois under BIPA, where claims only need to demonstrate that a violation of privacy rights under the act occurred, not that any material or 'concrete' harm ensued: *Rosenbach v Six Flags Entertainment*, 129 NE 3d 1197, para 22 (2019).

162. See Ignacio N Cofone and Adriana Z Robertson, 'Privacy Harms' (2017) 69 Hastings Law Journal 1039.

163. Although not an exact mapping, the idea of privacy as contextual integrity is closely related: Helen Nissenbaum, 'Privacy as Contextual Integrity' (2004) 79 Washington Law Review 119.

164. For a full description, see Cofone and Robertson (n 162) p 1385–1387

This concept is helpful both in the EU and US, as it concretises the conditions under which violations have occurred. Employee privacy protections in the US generally operate by distinguishing between personal and business information.[165] The concept of privacy harm ignores this distinction, as it is based not on an expectation of 'privacy', but rather an expectation of what the employer knows about the employee.[166] In the EU, other legal regimes such as tort and labor law could provide private rights of action beyond the protections offered by the GDPR through this concept. While the employer prerogative generally presumes surveillance in the US, the spread of that surveillance into the home with the advent of remote work as well as the depth of information that AI can now infer challenges what is normative for an employer to 'know' about a worker.[167]

Bodie offers the example of a 'smart' office chair cushion that 'records bad posture, heart rates, and time away from the chair.'[168] While a worker might understand that their boss can now see their heart rate, the worker might not expect that affective computing technologies could process the cushion's data to infer their emotional state. This example highlights both the limits and power of the concept of privacy harm in the workplace. Without control over whether the smart cushion is installed (due to limited worker power and the employer prerogative), the expectation now becomes that their boss will know their heart rate. However, the employer would still be liable for the harm created when they infer the worker's emotional state without notice.[169] This would be true whether the affective technology in question was provided by a third party or if the employer performed the inference in-house.

In this way, the concept of privacy harm would allow for additional protection not only from data collection, but also from data use with respect to inferences. This approach would also allow for the kinds of class action suits brought by groups of employees under BIPA, an important allowance for platform workers and organised workers alike.[170] In workplaces with a representative worker entity, privacy harms could also be considered at the level of the workforce, rather than just the individual. The principle still holds: inferences about a group, such as predictions of a group's overall personality traits or computed behavioural metrics, could be argued to violate that group's expectations of what is known about them in aggregate.

With respect to disparate impact, a major roadblock to regulatory enforcement and private right of action is, like privacy harms, detection. The case of disparate impact under automated decision-making is even more difficult to detect, as it requires an analysis of past decisions rather than just

165. Bodie (n 9) 36 ('U.S. law protects privacy within employment by drawing a line between personal information and business-related information').
166. Cofone and Robertson (n 162) 1050.
167. See Bodie (n 9) 36.
168. ibid.
169. Ignacio Cofone, 'Privacy Standing' (19 January 2021) 1398 SSRN Electronic Journal <https://papers.ssrn.com/abstract=3782887> accessed 16 October 2022.
170. Except that claims under privacy harm would be based on harm, rather than regulatory breach; for the BIPA cases, see (n 139) (class action suit under BIPA by employees of a bakery that engaged in biometric data collection without consent).

information about inferences, and because individual automated decisions are not evidence of systemic disparate impact.[171] Worker representation can help solve this *ex post* by serving as a detection mechanism, but only with the right information and resources. Representatives should not only be granted information rights like we argue above, but also have the right to use expert knowledge to test automated systems for disparate and systemic impacts.[172]

4. Conclusion

A groundswell of recent scholarship criticises the pervasive role of surveillance and data collection within dominant modes of economic production. However, when considering the effects of and potential countermeasures to increasingly datafied modes of production, there is tremendous danger in overemphasising its surveillant nature at the expense of attention to the relations of production.[173] In this article, we argue that approaches to workers' data rights have made precisely that mistake and, as a result, risk leaving workers with fewer options to assert autonomy over their work. Current approaches paint with the individualistic brushstrokes of data protection, fundamentally mischaracterising the role that data collection and analytics technologies play in the contemporary workplace, as well as the goals of workers who strive for greater data autonomy. Instead, we interpret asymmetrical access to big-data-derived information and models as a material condition of labour.

Following our reframing of information access as a labour issue, we illustrate how a privacy-focused regime of consumer-oriented data access is often mismatched to the categorically different needs of workers: collective and contextual information access. Our analysis also recognises that the differences between goals of consumer and worker data access underscores the urgency and complexity of implementing appropriate regulatory instruments. The fact that flows of data in the workplace are thicker, more opaque, and more personally consequential also presents greater challenges in creating generalisable, yet still contextually applicable regulation. Above and beyond the fragmentation inherent in applying labour law across diverse jurisdictions, forms of work, and technical configurations, our case study demonstrates the further complexity posed by the distributed technological nature of contemporary workplace data analytics systems.

Given the unforeseeable nature of the harms that can arise from data collection and use, we suggest that regulations focus on increasing worker agency and participation, rather than pre-empting specific harms in *ex ante* constraints or (only) providing remedies via *ex post* liability. We reject regulation that conditions data access on qualities like identifiability or content, advocating that access is useless to workers without also providing the sufficient context, expertise and tooling to understand it. Finally, we advocate for prioritising worker participation, bargaining,

171. For example, detecting disparate impact in ride-hailing platforms requires statistical and mathematical competence that many worker representatives likely lack; see Akshat Pandey and Aylin Caliskan, 'Disparate Impact of Artificial Intelligence Bias in Ridehailing Economy's Price Discrimination Algorithms' (2021) Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society <https://dl.acm.org/doi/10.1145/3461702.3462561> accessed 4 October 2022.

172. This is similar to the provision under the Platform Work Directive that provides for the hiring of an expert who worker representatives can consult, with the cost 'borne by the digital labor platform' (Platform Work Directive, art 9(3)). See also Aislinn Kelly-Lyth, 'Algorithmic Discrimination at Work' (elsewhere in this issue).

173. Evgeny Morozov, 'Capitalism's New Clothes' (*The Baffler*, 4 February 2019) <https://thebaffler.com/latest/capitalisms-new-clothes-morozov> accessed 6 May 2022; Adams and Wenckebach (n 100).

and co-determination in regulation, given their flexibility to capture diverse workplace systems, local regulatory environments, and labour contexts.

We suggest immediate routes to include principles in regulation. We argue for expanding data protection and access rights to common-sense categories of worker data following the example of the Platform Work Directive, while abandoning ill-fitting definitions of personal and identifiable data. These measures may help get data into the hands of workers and their representatives, but further interventions are still needed, like those explicitly empowering workers to authorise representatives. While these measures might open new opportunities for worker self-inquiry, the fastest and most effective route is still via worker co-determination over data collection, use, and in license agreements with third-party software or ML providers. Together, these recommendations validate information asymmetries as material conditions of labour rather than privacy or property concerns, and regulate them as such. At the same time, we acknowledge the complexity of implementing a labour law basis for worker data autonomy, opting to build on existing regulation in a hybrid approach to more quickly supply much needed tools for worker empowerment.

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

This work was funded in part by the Oxford Martin school under the Ethical Web and Data Architectures Project.