International Journal of Advanced Artificial Intelligence Research (IJAAIR)
eISSN: 3087-4270
Volume03 Issue11 | Page No: - 24-48

Received: 36 Oct 2026, Revised: 13 Oct 2026, Accepted: 22 Oct 2026, Published: 03 Nov 2026

DOI: https://doi.org/10.55640/ijaair-v03i11-03

Provenance For Autonomous Agents: What Regulated Enterprises Must Reconstruct After An AI Agent Acts

Sashank Siwakoti
Independent Researchers USA

ABSTRACT: Organizational decisions are linked to actions through autonomous AI agents which influence employees' work and operational delivery. The aim of this study is to propose an integrated framework for provenance in regulated enterprises across human resource management (HRM), labour management (LM), operations management (OM) and industrial relations (IR). The research is an integrative synthesis of 40 different research publications between 2001 and 2026 and descriptive secondary analysis of 20 publicly reported survey estimates. The literature is structured around four themes: Provenance and accountability, HR and organizational decisions, labor and industrial relations and human-AI operational performance. The synthesis sets out four interconnected requirements: the same context for both sources and execution, explicit delegation and human contribution, linkages between employment arrangements and operational delivery, and traceable participation and evaluation. These requirements are translated into eight proposed evidence dimensions that include mandate, rules, inputs, configuration, execution, participation, outcomes and stewardship. The differences in the shares reporting positive workforce and operational outcomes, based on public employer consultation data, range from 9.30 to 15.00 percentage points. Implementation context is provided by training, managerial decision and governance indicators. The main analytical result is the link between technical execution and institutional arrangements for people and work that is established through meaningful provenance. The study provides a 40-study evidence matrix, four research figures, a public data comparison that can be replicated, and seven proposed evaluation measures to help guide future workforce decisions and operational coordination and evaluate reconstruction quality in enterprises.

KEYWORDS: autonomous agents, provenance, human resource management, labor management, operations management, industrial relations, worker participation, AI governance.

1. Introduction

1.1 Research context

Autonomous agents can be used to fetch information, choose actions, call tools and modify business records towards a business goal. In this study, an autonomous agent is defined as a software system that makes and takes decisions on the steps to be taken within the delegated authority, with human intervention being arranged in a way depending on the decision. The definition allows for both single agents and integrated workflows and situates responsibility in the organization. Once such a system acts, a regulated enterprise must have an account that relates the purpose, authority, information used, sequence in which it is executed and the resulting organizational state.

In recent years there has been a growing body of practitioner research that reports on the use of agentic systems in organizations and the need to reconstruct consequential action pathways (OECD, 2026). This puts the relationship between technical implementation and managerial accountability on the agenda of substantive research questions.

This requirement has an employment dimension. Staffing decisions may affect service continuity, skill

utilization, overtime opportunities and working hours. A work-order assignment relates an operational demand to the skills and availability of an employee. A recruitment workflow can link the candidate evidence, selection criteria, managerial assessment and the final decision. Human resource management, labor management, operations management and industrial relations are the key elements that play a major role in the interpretation of autonomous action in these applications.

In this study, the term provenance is used to refer to the linked evidence of the processes involved in authorizing, informing, carrying out and reviewing an organizational action. A reconstructed account requires evidence sufficient to explain the complete organizational sequence. The exact duties to be performed during a deployment are set out in applicable legislation, sectoral requirements and employment agreements.

The central research question is: What evidence can a regulated enterprise use to investigate an autonomous-agent action and analyses its consequences for workforce decisions, operational performance, and industrial relations?

1.2 Research objectives

The overall aim is to develop an integrated provenance model that connects autonomous-agent execution with management and control, working arrangements, and proven results for the organization.

The objectives are:

To compile 40 relevant research studies and determine what evidence is needed to build an agent's mandate, decision context, execution and verified outcome.
To set up conditions for the recruitment, development of employees, evaluation of their performance, assignment of tasks, working time and remuneration.
To relate workforce choices to operational relationships and resource allocation and process completion.
To define the connection of worker information, consultation, collective agreements and human review to autonomous action.
To examine the publicly available indicators of the decisions made by managers, training of the workforce, consultation and governance, and how these relate to the results of the four management domains and the requirements for proof of provenance.
To create an analysis and measurement protocol for quality of reconstruction and workforce and operational outcomes.

1.3 Research contribution

The study provides a proposed eight-dimensional evidence framework based on a 40-study integrative synthesis, a replicable analysis of public workplace data and a comparison across four management domains and an enterprise evaluation protocol. Its main argument is that the significance of an independent action is contingent on the use and operating procedures of the action. Implementing those arrangements in a traceable way sets the groundwork for an informed review, organizational learning, and the subsequent empirical evaluation.

The framework and seven evaluation measures are proposed outputs of literature synthesis and secondary descriptive analysis; the study does not report an enterprise implementation or a tested autonomous-agent reconstruction.

2 Literature Review

2.1 Technical provenance and agent execution

Provenance research has been shown to have established that a reconstruction question dictates the type of evidence needed to answer the question. Buneman et al. (2001) separate inputs that lead to a result from the input sources that provide specific values. Cheney et al. (2009) tie these questions to the derivation structures that link inputs to outputs. This distinction is relevant for making organizational decisions, because it determines what evidence is taken into account, what records are used and how they are processed to become an action. Simmhan et al. (2005) categorize provenance systems based on their purpose, subject, representation, storage and dissemination. Davidson and Freire (2008) also differentiate between workflow specifications and execution histories, and describe the role of human annotations.

Herschel et al. (2017) link the information forms and system requirements necessary to support provenance applications. All of these studies provide a guideline: the granularity, relationships and accessibility of the record to be retained should be determined by the intended review and reconstruction task. Pasquier et al.

(2017) show an implementation path with CamFlow that can capture information flows in the system and allows for capture policies for specific applications. They are involved in the correlation of operational events and the broader context in which these events are processed. This is applicable if an enterprise action spans applications and needs to be supported by the activity of the agent as well as the systems to which it is sending requests.

AI documentation research is extended to artefacts that control the computational behavior. Model cards link model versions and uses to evaluation context (Mitchell et al., 2019). The composition of the datasets, their collection, preparation and maintenance are described in datasheets (Gebru et al., 2021). Raji et al. (2020) structure the internal algorithmic auditing process into scoping, mapping, documentation, testing and reflection. These contributions represent a set of records that can be used to provide an organizational meaning to an execution trace. Souza et al. (2025) present a preliminary distributed implementation of PROV-AGENT, which connects model invocations, prompts, responses, tools and workflow tasks directly to the agents. Its contribution is the linkage with interactions and the broader workflow. This study builds on that relationship by clarifying the employment criteria and delegated authority and institutional arrangements that enable a reconstructed enterprise action to be interpreted.

2.2 Human resource management and organisational decisions

HR research puts the focus on the purpose of an analysis and the employment context in the center of AI-supported decision making. Tambe et al. (2019) note the challenges of personnel decisions and suggest that data science should be applied by using causal reasoning, experimentation and employee contribution. Marler and Boudreau (2017) classify HR analytics according to the processes, theories, and outcomes of HR. They are relevant together because there is a clear connection between the data from the workforce, the interpretation of the data by management, the action chosen, and the outcome measured. Leicht-Deobald et al. (2019) also develop a conceptual framework for personal integrity in making algorithmic HR decisions and identify data literacy, ethical consciousness, and involvement in system design as relevant organizational practices.

Bankins and Formosa (2023) describe meaningful work as task integrity, skill development, task significance, autonomy, and belonging. Deployment decisions and work distribution are central to interpretation, and the analysis of these factors drives those decisions. These views extend beyond what staff consider criteria and considerations that would detract from or impact HR provenance. Langer et al. (2019) investigate the acceptance of automated job interviews in a preregistered video-vignette experiment. The study analyzes interview format, applicant experience, and the stakes of recruitment decisions. It contributes to the current system is the process context that needs to be included with candidate assessments: why an interview was held, how automation was applied, and what information was provided and how the human interaction was orchestrated.

A substantive research theme is also the organization of human and AI contributions. In Jarrahi (2018) complementary analytical and human judgement capacities are conceptualized. Shrestha et al. (2019) differentiate delegation, sequential decision and aggregation structures. Raisch and Krakowski (2021) describe the ongoing interaction of automation and augmentation, and Raisch and Fomina (2025) propose three types of hybrid problem-solving, autonomous, sequential and interactive. These studies confirm that it is important to document who did what, who contributed and the authority of the final decision. Kellogg et al. (2020) link algorithmic direction, evaluation and consequent workforce decisions with worker responses. Their synthesis demonstrates that there is a need for different organizational interpretations of recommendations, assessments and implemented employment actions. The present study is based on the idea that this distinction can be represented in a common action record that maintains the computational contribution, managerial decision and employment arrangement in place at the time.

2.3 Labor management and industrial relations

The literature on platform work offers comprehensive evidence of the link between algorithmic allocation and workers' experiences. This paper by Lee et al. (2015) looks at the drivers' understanding and reactions to algorithmic management. Rosenblat and Stark (2016) link differences in access to information with the organization of work. In this article, Wood et al. (2019) explore concepts of autonomy and algorithmic control

in the context of remote platform work. In these studies, the allocation criteria, performance information, and workers' interpretation of the work that digital systems coordinate are central to understanding the work.

Their implications need to be taken into consideration in the context of employment. Wood and Lehdonvirta (2021) discuss worker responses and collective action outside of the traditional work context. Vallas and Schor (2020) describe how platforms structure economic activity and wield power in different organizational structures. Duggan et al. (2020) link app-based work with algorithmic management and HR responsibilities and the employment relationship. The literature provides the support to record the process of establishing, communicating and experiencing work rules and to interpret the findings on the platform in their original institutional context.

Organizational application is complemented by work design research. Parent-Rocheleau and Parker (2022) link algorithmic management mechanisms to work features and worker outcomes. Parker and Grote (2022) view autonomy, skill use, feedback, social relationships and job demands as design options for digital work. Meijerink and Bondarouk (2023) discuss the enabling and constraining aspects of algorithmic management. All of these studies contribute to the consideration of an assignment in terms of the resources, discretion and working arrangements it entails.

Rani et al. (2024) broaden the evidence to regular workplaces in the case studies of logistics and healthcare. They relate algorithmic systems to work organization, tasks, skills and industrial relations. It offers a link between the platform research and the organizational processes of regulated companies where staffing and operations are also based on professional responsibility and employment conditions. The provenance contribution was built on this literature and it records the allocation rules, communicated expectations, actual work and participation of workers. In labor management, the record is used to record how an assignment has become working time and a payable entitlement. It provides a description of how information, consultation and agreements influenced the rules that were enforced through technology in the field of industrial relations. The following interpretations are complementary to the same organizational episode.

2.4 Human and AI contributions to operational performance

Empirical studies show how crucial it is to define the task and comparison when assessing the work that is supported by AI. Brynjolfsson et al. (2025) study a gradual rollout of a conversational assistant in the workplace and find that productivity improvements differ by experience and capability of the workers. In a randomized professional-writing experiment, Noy and Zhang (2023) discover improvements in task time and assessed quality. The alignment of the task and AI skills is recognized as a crucial factor by Dell'Acqua et al. (2026) in understanding productivity and quality. These studies are supportive of defining tasks, user contributions, and assessed outputs and then linking them in an operational evaluation.

The evidence also helps to make sense of the concept of human-AI complementarity. Dellermann et al. (2019) conceptualize the deliberate combination of human and artificial capabilities and their learning relationships. In this work, Fügener et al. (2022) explore productive delegation by means of controlled classification tasks, such as simulated AI-to-human allocation. Bansal et al. (2021) differentiate between explanation provision and acceptance of recommendations and the accuracy of the decisions. Vaccaro et al. (2024) combine human-AI performance and differentiate between human-AI performance that is augmented compared to human-only performance and human-AI synergy that is greater than the sum of its constituent parts. This distinction needs to be clearly established with baselines and cannot be used to replace each organizational outcome with an improvement in one measure.

There is a need for management constructs to complement these task studies and workplace survey evidence provides these constructs. Lane et al. (2023) review training, consultation and reported work place results. Milanez et al. (2025) categorize managerial information, decision speed, decision quality and autonomy, as well as governance practices. The measures are used to analyze public data as described in Section 4.2. The OECD (2026) study on the deployment of agentic AI provides direct evidence of organizational use. Interviews with 25 organizations share practices with scoped permissions, human decision checkpoints, employee upskilling, agent

registries and observing actions and data changes. These qualitative results link agent execution and organizational structures required for agent interpretation. In this study, they are used as direct contextual supports for constructing consequential action sequences, while the scope of the research in this study is preserved as illustrations.

2.5 Institutional foundations for reconstruction

Four additional documents link the research synthesis to existing concepts of provenance and governance. PROV-DM is the conceptual data model underlying the W3C PROV family of specifications. It defines entities, activities and responsible agents as well as relationships of use, derivation, and delegation (Moreau and Missier, 2013). Its vocabulary provides a common language to link organizational documents to the activities and responsibilities associated with creating them.

The NIST AI Risk Management Framework relates responsibilities, personnel competence, human oversight and feedback to organizational governance (Tabassi, 2023). The Generative AI Profile includes advice on provenance documentation and privacy and stakeholder expectations (Autio et al., 2024). The following voluntary frameworks are used to assist in the allocation of evidence responsibilities to the process owners, technical teams and workforce functions.

Adopted EU AI Act 2024 provides a jurisdiction-specific example of a text document. Articles 12, 14 and 26 and Annex III are dedicated to specific employment uses and logging, human oversight and employer information obligations (European Parliament and Council, 2024). The employment applications and reviewable records relationship is analyzed using that identified edition in this study. The duties to be performed on deployment are determined by the legal version, jurisdiction and employment conditions of the deployed action in question.

2.6 Research contribution and analytical position

The 40 studies provide complementary perspectives on the same organizational issue. Technical research is an explanation of how records and activities can be linked. HR and decision research describe the authority, the criteria and the human contributions which give meaning to a decision. Labour and industrial relations research provides an understanding of the allocation and control of labour. The task context, comparison baselines and verified outcomes are important when evaluating performance because they are explained by operational studies.

Their research contribution is the integration, at the level of the action episode. This study considers the evidence linking an autonomous action to its employment purpose, applicable criteria, authorized roles, operational context, affected workforce and review outcome as "workforce decision provenance". The proposed framework thus provides the possibility to reconstruct the computational sequence and the organizational interpretation of the computation.

There are different roles of historical reconstruction, computational replay and business explanation in this design. Reconstruction assembles recorded authority, context, sequence and outcome. When the execution conditions are present, Replay will run the process again. Business explanation is the criteria, evidence and documented justification of a decision. The study emphasizes reconstruction and provides explanations to make the record useful to managers, workers and representatives. Its main analytical thesis is that the continuity of these relationships over the entire episode is what defines organizational provenance.

3 Methodology

3.1 Research design

This study is an integrative review of 40 different research publications and descriptive secondary analysis of the publicly reported workplace survey data. The review proposes an organizational provenance framework, which brings together technical research with evidence in the areas of human resource management, labour management, operations management and industrial relations. The publications selected cover the time period 2001 - 2026. The evidence role of empirical studies and research reviews, conceptual articles and technical investigations remain intact.

The corpus was purposively selected in order to get complementary coverage in four analytical groups, each consisting of 10 studies. These assignments are not estimates of the distribution of the broader literature, but organize the synthesis. Once a publication is added to the corpus, it is added to all the management

domains to which it informs. The following four normative/technical documents offer institutional context: W3C PROV-DM, the NIST AI Risk Management Framework, the NIST Generative AI Profile and the adopted 2024 EU AI Act text. They are distinguished from the 40 research studies. The OECD Employment Outlook workbook provides another way of summarizing the survey evidence presented by Lane et al. (2023) for 2022.

3.2 Study selection and evidence collection

The targeted primary-source searches and reference tracing were completed on 30 September 2026. Retrieval was provided by publisher websites, author repositories, university repositories, scholarly repositories and the original OECD and ILO-JRC websites. The four topic groups from Table 1 were used in the searches, which also included the following key terms: algorithmic management, author name and provenance, article titles, work design, and HR decision-making and human-AI collaboration terminology. The substantive link with the research aims was given high priority, as was the ability to identify the author(s) and the original source (which was inspectable) and the clearly identifiable publication version.

Eligible publications made an empirical finding, technical method or developed theoretical argument on one or more of the following: Reconstruction of Data and Actions; Organizational Delegation and Human Judgement; Management of People and Work; Evaluation of Human-AI Performance. When the same study was reported in a working paper and a final article, the publisher final versions were preferred. Promotional material and duplicate versions were excluded from the research corpus as were editorial introductions. The outcome of the review is a bounded integrative synthesis for the purpose of developing a proposed framework.

Table 1 Composition of the 40-study research corpus

Analytical group	Included studies and evidence	Contribution to the synthesis
Provenance and accountability	T01-T10; 10 studies of database and workflow provenance, system capture, model and data documentation, audit and agent workflows	Establishes how sources, configurations, execution and responsibility become traceable
HR and organisational decisions	H01-H10; 10 studies of HR analytics, algorithmic personnel decisions, human judgement, meaningful work and applicant perceptions	Connects technical evidence with job criteria, managerial authority, competence and employment decisions
Labour and industrial relations	L01-L10; 10 studies of algorithmic management, platform and regular employment, work design and worker responses	Connects allocation and operational rules with autonomy, working arrangements and worker participation
Human-AI operational performance	O01-O10; 10 studies comprising workplace surveys, experiments, conceptual synthesis, meta-analysis and agentic deployment interviews	Establishes the importance of task context, delegation, workforce capability and verified outcomes

The 40-row evidence matrix is included in Appendix A. The study is identified in each row, and the research design, substantive finding and the implication for the provenance are developed in this study. Full text of all 40 studies was inspected to determine relevance, and published articles and full author manuscripts or proofs were inspected, the identity of which was verified by matching with the final bibliographic record. Empirical details of samples are only included when directly verified. Conceptual propositions do not become observations and findings from platform work do not become their employment context.

Appraisal of evidence included source identity and clarity of the reported design, traceability of the extracted finding and relevance to the proposed organizational application. Reconciliation of publications reporting the same underlying research was done by the title, author team, DOI and version. Conceptual integration was carried out by reviewing findings and the studies that make up the findings were not included in the corpus count. This leaves the difference between 40 research publications and the number of independent datasets they contain intact.

3.3 Public data and analytical populations

Twenty published estimates (18 percentages, 2 percentage-point contrasts) are contained in the quantitative extract. The 12 estimates are based on the original OECD survey reports and the eight group percentages on the official Statlink workbook. The value, unit, analytical population, fieldwork period and source location are retained for each record. There are 10 descriptive comparisons: four between managerial indicators, two between governance indicators and four between the categories of employer consultation.

Table 2 Public survey data and analytical populations

Source and fieldwork	Full survey sample	Indicators used	Population for the selected estimates
Milanez et al. 2025; June-August 2024	6,047 establishment responses in six countries; businesses with at least 20 employees	Four managerial outcomes and four governance indicators	Managers in establishments using algorithmic management; published adopter aggregates
Lane et al. 2023; January-February 2022	5,334 workers in finance and manufacturing across seven countries	Training provision and consultation-associated performance differences	AI users; training bases 1,231 finance and 912 manufacturing; consultation comparison bases 1,131 and 842
OECD 2023 workbook; the 2022 employer survey also reported by Lane et al.	2,053 employers in finance and manufacturing across the same seven countries	Four positive outcome assessments by consultation status	AI-adopting employers in each consultation category; original combined aggregates retained

The 2024 survey covers France, Germany, Italy, Japan, Spain and the United States; its aggregation uses sector and size weights and equates country weights. The 2022 surveys cover Austria, Canada, France, Germany, Ireland, the United Kingdom and the United States. Worker weights include age, gender and education, and employer weights, firm size. The estimates are made using the full samples, the description of which

represents the survey coverage; the population of each estimate is determined by the eligibility of the items (Lane et al., 2023, pp. 21-23; Milanez et al., 2025, Annex A).

3.4 Synthesis and analytical procedure

The study-level extraction record defines a literature unit of analysis. The authorized task/trigger, execution, confirmed result, and review are the organizational unit that builds the proposed framework. At the episode level, an episode can include multiple model invocations, tool calls, and transactions/human interventions. Different identifiers keep the episode in touch with the events that are contained.

The study developed eight provisional codes for synthesis: mandate and authority, rules for governance, lineage of inputs, system configuration, effects on process and change of state, human engagement, workforce and operational outcomes, and stewardship of evidence. Four management domains were identified with additional tags. Several codes might also be applied to a study, and its primary group identified in Table 1 did not change.

The provisional codes were not eight independent induction discoveries, but originated deductively in relationship to the six research objectives and the four analytical groups. The HR, labour, industrial-relations, and organizational decision literature (H01-H10, L01-L10, O01-O10) sensitized mandate and authority, governing rules, human participation, and workforce and operational outcomes. Database, workflow documentation, and agent-trace research entities (T01-T06, T08-T10) informed lineage, system configuration, execution, and state change. Evidence stewardship encompassed efforts for capture (T02), auditing (T05), access, and documentation (T07-T09). The technical and governance vocabularies provided by the W3C PROV-DM and the NIST documents were not included in the 40 studies.

The analysis process had four stages. The first step was to interpret each finding within the study design and its original context. Then results were compared across and within groups to identify complementary evidence needs. Third, these requirements were stated as records and relationships to reconstruct an action episode. Finally, the proposed framework was applied to the four management domains and operationalized in seven proposed evaluation measures. This evidence matrix differentiates the source finding from its use in the didactic application developed in this study.

During the comparison stage, each preliminary category was evaluated for whether it matched the finding and derivation application from the source; in each case, the study was recorded in Appendix A, with multiple categories acceptable for one study. The eight categories were retained as proposed final dimensions as each catapulted a different reconstruction requirement, and was supported in the selected corpus: authority differs from an applicable rule; a rule differs from an input; configuration differs from an executed event; a logged attempt differs from verified state change; participation differs from the measured outcome; and stewardship governs the integrity and accessibility of all of these records.

To represent the four domain requirements from Section 4.1, the study did not need to delete or combine any categories. This evidence mapping is an analytical confirmation, not proof of inductive saturation, independent coder agreement, or enterprise validity, and appears in Table 4 and in Supplement S1, the dimension codebook. A consistency review was carried out; sources were confirmed as solid published versions, findings were verified against the cited claim(s), and no duplicate records were found. The synthesis evaluates evidence for relevance or complementarity. The quality of the chosen corpus is therefore not to be viewed as descriptive of effectiveness in the number of studies or concepts presented, but rather as descriptive.

3.5 Quantitative analysis and scope of inference

The precision of estimates published remains unchanged, as does the meaning of the population or measure. Percentage points are calculated by subtraction. For each employer response, difference scores are computed as the percentage difference between the consultation group reporting a positive impact and the no-consultation group reporting a positive impact. Marginal response shares under the same adopter eligibility condition are used to compare managerial or governance indicators. Each result is independent of others.

The analysis used is descriptive. Training rates represent observations of reported provision;

governance represents observations of reported practices; consultation represents document-observed associations. The results of the source that the sector difference in provision of training is not statistically significant are maintained. Comparisons of aggregates are evaluated through their design that has been published, and no further significance tests and causal estimates are supplied.

The empirical evidence has two functions. Both study specifications and reported results in workplace AI are grounded in relevant management constructs. It provides evidence for the multi-step execution from direct agent research that informs organizational governance. These contributions are woven into the proposed framework, and the evaluation protocol in Section 4.9 addresses questions about the framework's impact on reconstruction quality and enterprise performance. This maintains alignment with observed evidence from the analysis and its implications for organizational applications.

3.6 Data availability and reproducibility

Appendix A includes all 40 included studies, along with how they were identified, and references with the DOI or original publication link. The public employer consultation data are available from OECD StatLink 34bepy, which accompanies OECD Employment Outlook 2023, Figure 7.3 Panel B. This study uses version 1 of the workbook, from 11 July 2023; worksheet g7-3, cells B40:E41, retrieved on 30 September 2026. Values are accurate to 2 decimal places, with floating-point serialization noise removed. The supplemental files include the reported source cell reference and a checksum of the supplied source cell extract. Figures 1–4 and Table 3 summarize the twenty survey estimates and the computed comparisons, and the accompanying text discusses them.

The 40-study extraction sheet, 20-estimate data extract, and proposed-dimension evidence map are contained in Supplement S1 (Supplement_S1_Study_and_Survey_Extraction.xlsx). The eight two-decimal employer consultation values and the StatLink worksheet/cell coordinates are recorded in Supplement S2 (Supplement_S2_Source_Cell_Extract.csv). The four employer contrasts and six managerial/governance comparisons are included in Supplement S3 (Supplement_S3_Reproducible_Calculations.xlsx), which contains the source data inputs and explicit formulas performed in the spreadsheet. In Supplement S4 (Supplement_S4_SHA256_Manifest.txt), exact supplementary material files are provided, and their corresponding checksums for hash function sha256 are reported. These materials were provided as a supplement to this SJR submission and are therefore not necessarily the materials that accompany the published version of the article; the original will be maintained at the site described in the submission's supplemental materials listing and is freely available to read at https://stat.link/34bepy.

The eight employer-reported values are in cells B40:E41 on the "g7-3" worksheet in the OECD Excel workbook. However, source-cell locations have not been independently verified, and the reported numerical data have not been separately verified against OECD data from the original OECD workbook. As a result, the provided SHA-256 checksum was used only to validate the source-cell transcription in Supplement S2, and the extraction records and calculations are not designed to replicate the outcome of the original OECD workbook or to validate any corresponding source file.

4 Results and Analysis

4.1 Findings from the integrative synthesis

The 40-study synthesis determines 4 complementary needs for the autonomization of work. Each requirement is linked to a technical record and a management relationship, and each emerges from bodies of research of various natures, though they are identifiable by their evidentiary contexts. The outcome is an all-inclusive narrative about the enterprise information required for an action and the records in which that information is preserved.

Sources and Execution Need a Shared Organizational Context

Database and workflow research encourages outputs to be clearly distinguishable from their transformation and execution (Buneman et al., 2001; Cheney et al., 2009; Davidson & Freire, 2008). Audit research and documentation provide context for the model, data set, and review information (Mitchell et al., 2019; Gebru et al., 2021; Raji et al., 2020). This relationship is shaped by the purpose and identity of the action itself. Source values are meaningful only when they are linked to a

specific job requirement and an employee record in the decision period to which they apply. The derived requirement then forms a technical chain from the source to the action, and its meaning is the effect for the organization.

Delegation and Human Contribution Are Part of the Evidence

Organizational decision research shows that modes of cooperation between humans and AIs can involve delegation, sequential contribution, and aggregation (Shrestha et al., 2019; Raisch & Fomina, 2025). Both the allocation rule and the human response shape performance in collaboration studies (Fügener et al., 2022; Bansal et al., 2021). The synthesis is thereby recorded and is granted authority and participation. The record shall include who authorized the task, what action the agent may have taken, what evidence a reviewer may have considered, and how the action was authorized. This requirement is associated with managerial accountability and HR competence and is connected to the overall coordination of the operations.

Work allocation combines employment and service-delivery aspects. Work design research is about the relationship between algorithmic management, task characteristics, and workers' outcomes (Parent-Rocheleau & Parker, 2022; Parker & Grote, 2022). Research on meaningful work and regular workplace factors includes skill utilization, job design, and labor organization (Bankins & Formosa, 2023; Rani et al., 2024). The synthesis creates an action-level connection between demand, qualification, assignment, work, and outcomes. A maintenance allocation, for example, needs to show the task and asset the person is working on, then the technician's skills, availability, and applicable working arrangements. This same episode may thus be used to support an employment review as well as an operational assessment.

Participation and Evaluation Complete the Reconstruction

Workplace studies suggest that employee participation and feedback are linked to the deployment of algorithmic systems (Leicht-Deobald et al., 2019; Lane et al., 2023). Task fit and comparison baselines are fundamental to performance assessment in performance research (Dell'Acqua et al., 2026; Vaccaro et al., 2024). Direct agents research provides an action-reconstruction relation that connects traces and the governance of action sequences (Souza et al., 2025; OECD, 2026). The derivation requirement makes the implemented action dependent on the legislation that underlies it and on the results that are later evaluated. This provides a recorded path for the role of industrial relations, with information, consultation, agreed rules, and recorded responses. These four requirements lead to the eight proposed evidence dimensions presented in Section 4.3. The public-data analysis that follows provides quantitative context for managerial decisions, capability, governance and participation. These combine to establish the records and relationships to be evaluated when the proposed framework is applied within an enterprise.

4.2 Secondary analysis of public workplace data

The quantitative analysis establishes four complementary empirical themes: managerial decision resources, workforce capability, organizational governance and consultation-related outcomes. Figures 1-4 retain the original populations and units and link the findings to the study's management aims.

Managerial Decisions and Operational Coordination

In establishments with algorithmic management, published improvements are 70% for information availability, 62% for decision speed, 60% for decision quality, and 50% for managerial autonomy (Milanez et al., 2025, pp. 30-31). These individual assessments are displayed in Figure 1.

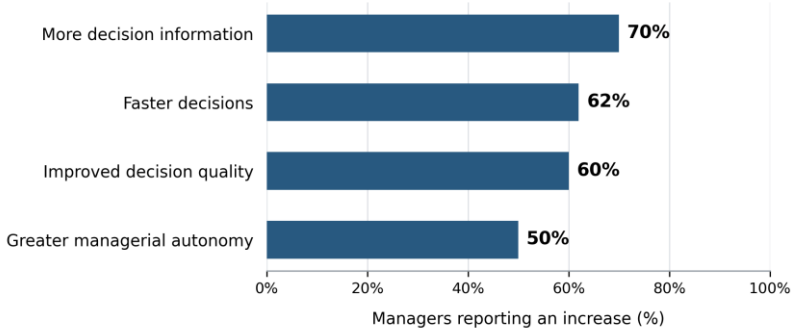


Figure 1 Reported increases in managerial decision indicators

Information is 10 points higher than the rounded value for decision quality, 8 points higher for speed, and 20 points higher for autonomy; speed is 2 points higher than the rounded value for quality. Each contrast helps evaluate information lineage, decision timing, managerial judgment, and delegated authority. The associated evidence for operations management links demand and capacity inputs to the final transactions. In HR/labor management, it links decision criteria and decision-maker authority to allocation or employment decisions.

The relationship between HR capability and the industrial-relations dimension is shown in Figure 2. Of those using AI in finance and insurance, 58% have reported some training provision, and 53% in manufacturing have reported the same. The interpretation emphasizes training as a relevant feature in both industries. Differences across industrial sectors are statistically insignificant (Lane et al., 2023, pp. 69-70).

Workforce Training and Employee Participation

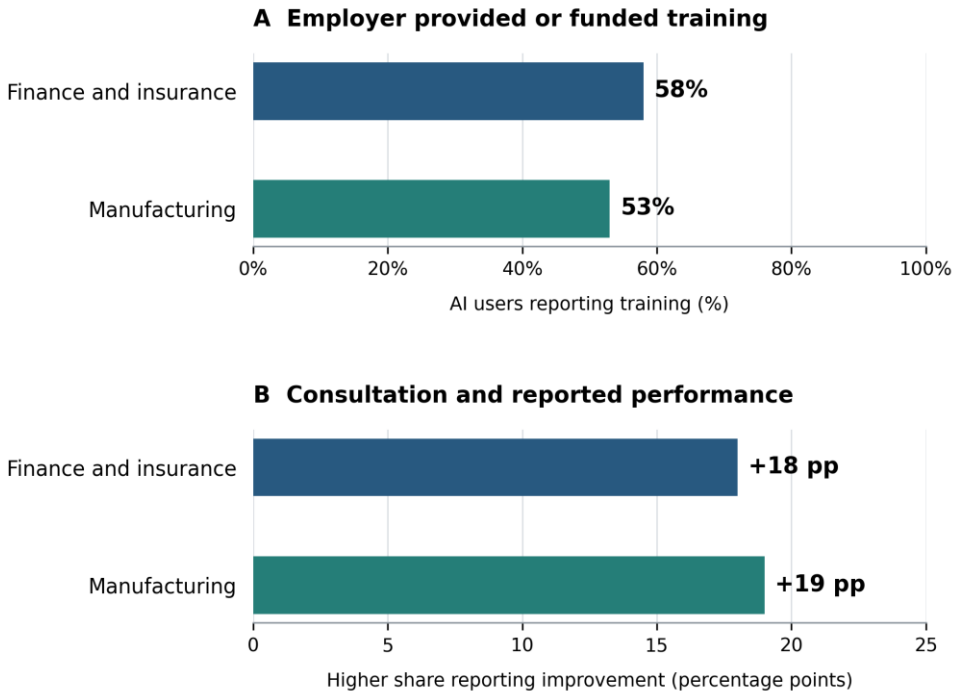


Figure 2 Workforce training and consultation

Differences associated with the consultations are 18 points in finance and 19 points in manufacturing. Each indicates a higher percentage of AI users who reported positive overall performance when they used AI than when they did not. This study uses the result to establish 'partnership participation records' as an analytical input: the record would link the issue discussed, the information supplied, the contribution of the workforce, the management response, and the outcome rule. In the context of human resource management, the

provenance can be used to establish a traceable relationship between training of employees, competencies acquired, and their application in the work assignments or review responsibilities. Training records may also show what capability was built and, for subsequent learners, who completed the training activity and which assignments or review activities used that capability.

Governance Practices as Foundations for Provenance

Recent employer performance data show that 64% of adopter establishments have implementation guidelines and use guidelines, 63% have worker and representative consultation, and 56% have regular audits (Milanez et al., 2025, p. 10). The three specific practices are illustrated in Figure 3. Several practices may be present, and each percentage pertains to its own indicator.

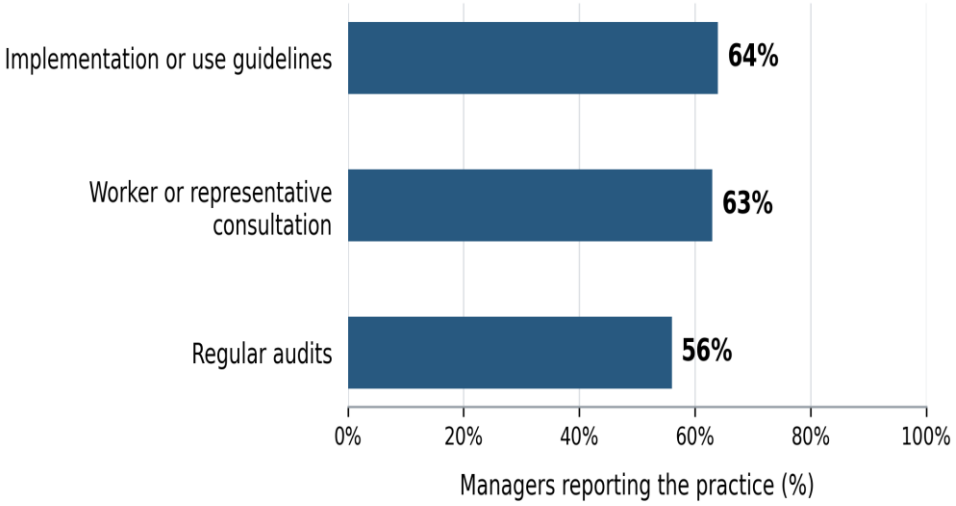


Figure 3 Reported governance practices

The descriptive differences are 8 percentage points between guidelines and audits and 7 points between consultation and audits. These practices matter in analysis because they generate evidence: guidelines capture evidence of applicable rules, consultation records capture evidence of worker participation, and audit-related records capture evidence of organizational oversight and review. This study serially couples these practices, making it possible to evaluate the relationship between an organizational arrangement and the action implementing that arrangement in the same event. The governance rates position these practices relative to each other in terms of relevance; the proposed provenance design determines how evidence for them can be linked.

Consultation-Group Comparison Using the Public Workbook

The official StatLink workbook gives participants a reproducible comparison of four employer-reported outcomes. The total group of employers for all the countries and sectors surveyed and published in 2022 is shown in Figure 4. The specific values and this study's calculations are shown in Table 3 (OECD, 2023, Figure 7.3, Panel B).

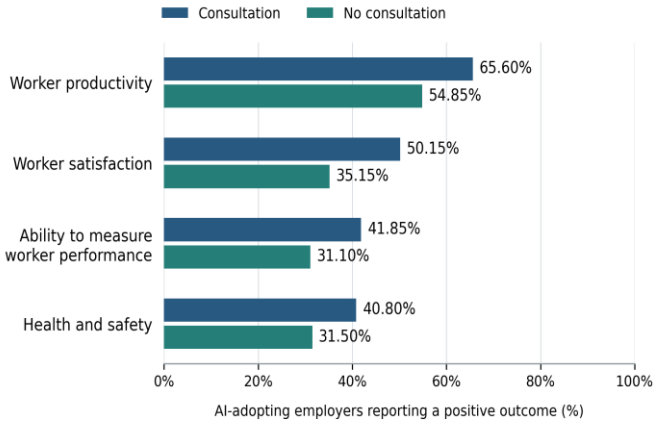


Figure 4 Employer assessments by consultation status

The four published outcome shares remain higher in the consultation than in other categories. The largest contrast is in worker satisfaction, at 15.00 percentage points. Productivity differs from the capacity to measure performance by 10.75 points, and health and safety by 9.30 points. The ordering specifies the estimates to be displayed and maintains each outcome as a separate measure.

Table 3 Employer consultation outcomes and calculated differences

Employer-reported positive AI outcome	Consultation (%)	No consultation (%)	Difference (percentage points)
Worker productivity	65.60	54.85	10.75
Worker satisfaction	50.15	35.15	15.00
Managers' ability to measure worker performance	41.85	31.10	10.75
Health and safety	40.80	31.50	9.30

Source: OECD (2023), StatLink 34bepy. In this study, percentage-point differences are the difference between the percentage of employers reporting positive outcomes without consultation and the percentage of employers reporting positive outcomes with consultation. The percentages shown are the shares of employers who positively evaluated the impact of AI on each outcome. The sector-country aggregation remains unchanged, and the comparisons are descriptive associations.

The results link industrial relations to the organization of work in four ways. Worker productivity applies to service provision. Satisfaction gives HR a measure of workforce experience. The ability to measure performance relates to evidence of performance management. Health and safety links labor organization to task-related working conditions. A typical implication record of provenance that accompanies participation, applicable rules, allocation decisions, and subsequent outcomes. This study analytically interprets the public's findings. Methodological and descriptive quantitative contribution. It clarifies relevant workplace relationships and shows a clear path from the source

cells in the public to the comparison and management interpretation cells. The eight-dimensional framework proposed below translates these relationships into the evidence requirements for autonomous agents' actions.

4.3 Proposed integrated provenance framework

The main analysis product is an eight-dimensional provenance framework proposed. When technical events are connected to the authority, the rules governing staff work, and the actual achievements these rules uphold, they become organizationally interpretable. This translates to a practical evidence structure as shown in **Table 4.** Table 4 shows eight dimensions that are the deductive categories that were retained in the study based on the study-to-requirement mapping shown in Appendix A and in Supplement S1; these dimensions have not been independently validated through enterprise deployment or through reconstruction of the execution logs, or via a formal completeness test. The table lists proposed record classes and relationships that are proposed for future evaluation and are not verifiable controls or results of implementation.

Table 4 Proposed integrated provenance framework

Dimension	Evidence to reconstruct	Management connection
Mandate and authority	Purpose, initiating request, accountable role, delegation scope and approval conditions	Establishes responsibility for employment and operational decisions
Governing rules	Applicable policy, procedure and agreement versions, effective dates and decision criteria	Connects action to job requirements, working arrangements and negotiated provisions
Input lineage	Source records, versions, retrieval times and transformations	Identifies the demand, availability, qualification and performance information used
System configuration	Agent and model identifiers, relevant instructions, tools and configuration versions	Identifies the authorised technical configuration under which the action occurred
Execution and state change	Ordered actions, transaction identifiers, responses and confirmed resulting records	Reconciles instructions with assignments, work orders and administrative changes
Human participation	Review authority, evidence considered, interventions, notifications and consultation references	Connects managerial judgement and worker participation with the episode
Workforce and operational outcomes	Final decisions, work allocation, recorded hours, remuneration effects and process completion	Supports joint assessment of employment decisions and operational delivery

Dimension	Evidence to reconstruct	Management connection
Evidence stewardship	Record origin, integrity checks, access permissions, retention and disclosure rules	Makes the account retrievable and appropriate for each authorised reviewer

The evidence matrix supports the derivation. Mandate and human participation are supported by organizational delegation and internal auditing (H05-H08, T07, and O08-O10). Input lineage, configuration, and execution link the database and workflow provenance to the model and dataset documentation (T01-T06 and T08-T10). The measurement of governing rules and workforce outcomes is based on HR (H01-H04) and work design and employment research (H09-H10 and L01-L10) and supplemented by the outcome context of the survey and task studies (O01-O06). This also includes capture, documentation, and authorized review (T02, T05-T09), along with additional and supplementary documents from the W3C and NIST. These are the mappings that were developed by this study in light of the findings in Appendix A.

Linkage is the theme upon which the proposed framework is built. A policy's unique identifier is linked to the version in use when the action occurred. A permission is linked to a specific determination and the authority of the person making the determination. A transaction relates to its resultant confirmed impact on the target system. These items should then be easily retrievable and pertinent to the proper episode. The account should include four distinct events representing what was captured: intended action, attempted action, tool-acknowledged action, and verified business change action. A shift change should be verified by referring to the change in the official roster record. To claim a change in the payroll amount, refer to the relevant approved state payroll. This separation lays out the evidence required to show readers what was actually implemented.

4.4 Human resource management analysis

The HR analysis shows that employment criteria are necessary to give meaning to workforce provenance, as supported by literature in personnel analytics, integrity, and meaningful work (Tambe et al., 2019; Leicht-Deobald et al., 2019; Bankins & Formosa, 2023). Technical records indicate which data and rules were applied; HR defines whether these are technical records relating to the job, employee, and decision period. The proposed framework therefore gives HR the authority to set authoritative employment parameters and the organization's interpretation of personnel evidence. For Recruitment, the episode should link to the approved role profile, published selection criteria, information to be provided to candidates at assessment, agent results, reviewer assessment and decision.

Effective dates are provided for criteria changes, and each decision is linked to a specific version of the criterion. This establishes an auditable record of how the organization has handled job-related requirements during the selection process. For performance management, the record should take a line-of-sight approach that links the assessment period to the work, resources, complexity, and the employee's contribution. A supervisor's assessment is documented and grounded in clear, identifiable, evidence-based information. Any comments made by an employee or an employer-authorized change are linked to the same review. This design allows analysis of performance decisions in light of comparable work and the circumstances involved in the operation.

Human resource development is an additional application. A skills recommendation must be linked to the observed task requirement, validated competency assessment, development discussion, and learning opportunity provided. The remaining records can serve as proof of completion and proficiency and then show eligibility for later assignments. Provenance therefore adds a temporal dimension to assessing how autonomous support processes contribute to skills development and internal mobility.

HR also plays a direct role in managerial capability. Preparation for reviewers should include understanding of the source records and

understanding of agent recommendations and documenting those decisions that have been justified. The research contribution links the records representing workforce decisions to the competencies needed to review them. Its effectiveness can be determined by the quality of the review and how evidence is used across the review.

4.5 Labor management analysis

The labor-management analysis aligns with the evidence chain used for operational allocation decision-making. This linkage emerges from the work design and employment lines of work in Parent-Rocheleau and Parker (2022), Meijerink and Bondarouk (2023), and Rani et al. (2024). When the rules of employment are still connected with the transaction, scheduling, task assignment, and remuneration can be rebuilt. Appropriate scheduling factors include agreed availability, qualifications, current assignments, and applicable working-time arrangements. A record of the decision should be made on the scheduling record, and the approved change communicated to the worker, as well as the final allocation.

The applicable provisions become identifiable inputs to the episode where there are rest periods and overtime allocation/shift preferences in the arrangement. This structure will enable analysis of meaningful allocation patterns. The overtime distribution can be compared to eligible workers and available jobs. Training may be tied to the requirements of the job and the individual growth plan. Comparisons may be drawn by taking into account availability, job classification, location, and any relevant qualifications. So, the comparative unit determines the range for the employment choice under review. The remuneration procedure must be a series of steps and must link the agreed assignment, the recorded work, and the calculation of the approved remuneration.

Authorized adjustment and payroll posting are intended to be separate, related documents that include authorized hours and recorded hours. The subsequent narrative summarizes how the project was turned into a claimable benefit in line with the agreement. Labor and management have unambiguous work allocation, and there is a clear link between employment entitlements and work allocation, as envisioned in the proposed framework. It also outlines the information that will be analyzed regarding scheduling predictability, allocation consistency, and administrative correctness. Theodore Hertz students provide workforce-focused enterprise with a direct impact.

4.6 Operations management analysis

The operations analysis framework documents evidence of continuity across business systems, based on workflow provenance and first-hand research with agents (Davidson & Freire, 2008; Souza et al., 2025; OECD, 2026), demonstrating the need for reconstruction. The demand planning process can link to staffing, maintenance, inventory, and service delivery in an action episode. The relevant corresponding individual handovers should therefore link the handover action with the handover reception and the organizational state to be reached.

The proposed record for a maintenance assignment integrates all evidence of asset condition, work-order priority, technician qualifications, assignment shift availability, authorization, and completion evidence. In the case of service allocation, it is connected to the extent of need and commitment, competencies, and capacity. These relationships allow evaluation of a decision within the decision-making process. Records can be correlated by a common episode identifier; each system has its own identifier for episode transactions, and each system is an authority.

In these timing data sequences, events are visible, and dependency links show where one step's output is used in the next. When an agent delegates work to other agents/many human agents, the delegated log shows the delegated agents/workers' responsibilities and acceptance conditions. Workforce indicators can then be linked to operational evaluation to connect service completion, first-pass acceptance, schedule changes, and resource use. The analytical contribution is a joint performance perspective: the same evidence supports both delivery and the working arrangements through which delivery was achieved. It also provides the opportunity to keep the influence of human inspection in an automated process.

4.7 Industrial relations analysis

The issue of worker participation is substantively addressed in the context of the irradiated food supply chain in the industrial-relations analysis, which is based on literature regarding involvement of employees and worker responses and consultation (Leicht-Deobald et

al., 2019; Wood and Lehdonvirta, 2021 and Lane et al., 2023). The rules governing autonomous work have been established and applied, as evidenced by worker information, representative consultation, and collective agreements. This study proposes a consultation chain linking planned organizational change, the consulting trigger, information provided, representative involvement, and subjects for discussion, management's response, and the final decision. It is structured to illustrate how participation has been used by agents to develop their rules.

The consultation outcome and whether an agreement was reached are recorded separately. Collective agreements must meet the same accuracy level required for the equipment. They should include implementation dates, whom the clause applies to (which level), and how it will be interpreted based on the organizational process and the operational rule implementing it. There may be an arrangement that corresponds to a scheduling or allocation episode for that time period. This makes negotiated provision a factor in the system's behavior. There are review and grievance steps that can be added to the sequence. An employee enquiry can tie into the questioned decision, including supporting evidence, the employee's narrative, the reviewer's assessment, and the reviewer's response.

An authorized adjustment generates a new event that is linked to its predecessor. The result provides a consistent record of the organization's response and the implementation of any accepted change. A collective level analysis of suitably consolidated evidence can be included as a basis for discussion of workload distribution, skill development, scheduling arrangements and the use of operational gains. Access to be controlled by discussion purpose and representation applicable. This generates an evidence framework for informed dialogue between participation and implementation in the workplace.

4.8 Comparative analysis across management domains

The domain analysis seeks to identify multiple management meanings in an action episode. A scheduled task is an operational assignment, a utilization of employees' capabilities, an application of working-time arrangements, and an implementation of rules established by the organizational participation activity. These relationships are explained in Table 5.

Table 5 Comparative management analysis

Management domain	Core relationship reconstructed	Outcome for enterprise evaluation
Human resource management	Job criteria and competency evidence linked to authorised employment decisions	Consistency of assessment, development access and review quality
Labour management	Allocation rules linked to assignments, actual hours and remuneration	Scheduling predictability, entitlement accuracy and allocation patterns
Operations management	Demand and resource requirements linked to executed and confirmed transactions	Process completion, service quality and continuity across handovers
Industrial relations	Information, consultation and agreements linked to implemented rules and review outcomes	Explanation usefulness, response timeliness and completion of agreed actions

A common analytical approach is to follow the applicable rule from the supporting evidence to the action taken. This relationship provides a concrete basis for comparisons across similar cases. Comparability in HR is determined by the job and the criteria. In labor administration, comparability depends on eligibility and the allocation of opportunities. Once the operation is underway, it considers operational conditions and task requirements. Comparability in industrial relations depends on the nature of the organizational change, the consultation needs, and the agreed consultation arrangements needed to make the change effective.

These relationships are established through evidence stewardship. The NIST Generative AI Profile links provenance documentation to privacy and security, as well as stakeholders' expectations (Autio et al., 2024). This study supports this principle through purpose-specific views that make employee evidence available for individual decisions, relevant evidence available for collective decisions, and process evidence available for operational reviews. The quality of each reconstruction is enhanced by integrity checks and corroboration with authoritative business systems, as well as with records of origin. This organizational-level comparison would resolve the research question. The complete report should bring together technical performance and the practice and institutional forms in which it resides. The desired outcome is a quantifiable data set that can be compared with work experiences to observe operational performance.

4.9 Proposed measurement framework and data analysis protocol

The analysis suggests a measurement scale for further assessment of the enterprise. Data collection begins with a given decision class, such as recruitment assessment, shift allocation, assigning workers to training, adjusting compensation, or scheduling maintenance. Before sampling, the system provides episode boundaries and evidence requirements/review conditions for each. Executing a transaction provides a listing of the executed transaction and its end state in transaction records. Human resource and labor reports include information on role criteria, qualifications, availability and employment outcomes. The rules that are active during a given period are identified in policy and agreement registers.

The system verifies and consults human and institutional contributions based on review and consultation records. Structured interviews or questionnaires capture employee, manager, and representative perspectives on explanation clarity, review usefulness, and the practical value of participation. The denominator for each percentage is the number of eligible persons in the respective population. Applicability is determined in advance of scoring. Evidence coverage reports are provided for the percentage of episodes that meet all critically required evidence items, such as authority and required review. Reconstruction time and completion status are reported for each selected episode. These reporting decisions make the measures comparable and preserve their essence even when a reconstruction is successful. The workforce outcome should be assessed at an appropriate level for each process. Recruitment analysis can be used to compare the application of documented practices. Labor analysis can be made on notice periods, labor allocation, recorded hours, and payroll reconciliation. HR career development analysis can examine educational access and access to qualified roles.

Industrial-relations analysis may include response timeliness, explanation utility, and performance of agreed actions. A prospective design can provide baseline and implementation periods for comparable decision classes and, when possible, comparable operating units, for subsequent field evaluation. Analysis should consider demand, staffing, job mix, seasonality, and other organizational changes. Descriptive distributions develop a performance pattern; a suitable comparison design is required to estimate the contribution of the provenance intervention. Interview analysis can account for how participants use these records and which components help them understand or review them.

Table 6 lists seven proposed evaluation measures that emerged from the framework development in this study. Using them would produce enterprise data that could measure improvements in decision quality, operational coordination, and employee experience.

Table 6 Proposed evaluation measures for enterprise evaluation

Measure	Operational definition	Research objective
Evidence coverage	Percentage of applicable required items that are retrievable, verifiable and linked to the correct episode	Completeness of the reconstruction record
Reconstruction success	Percentage of all selected episodes meeting a predefined reconstruction checklist	Usability of evidence for organisational review
Reconstruction time	Median and 90th-percentile time from authorised case access to checklist completion among completed cases, accompanied by the success rate and elapsed time and status of unfinished cases	Practical efficiency of review
Workforce rule traceability	Percentage of decisions requiring rule and authority links that retain both the applicable rule version and required authority	Reviewable HR and labour decisions
Verified transaction coverage	Percentage of executed actions requiring confirmation that are linked to authoritative resulting records	Operational continuity and confirmed outcomes
Review quality	Percentage of sampled completed reviews recording evidence considered, disposition and reviewer justification	Substance of human oversight
Consultation traceability	Percentage of changes meeting predefined consultation triggers with a complete participation and response chain	Informed industrial relations

Each of these seven proposed evaluation measures is yet to be calibrated and tested through enterprise reconstruction cases. The proposed field protocol must empirically evaluate each measure's denominator and threshold, interrater agreement, and sensitivity to missing or conflicting records. The main evaluation connection, truth be told, is between reconstruction quality and management results. Evidence coverage and verified transaction coverage look at the record, not the transaction. Review quality assesses how the record is applied; workforce and operational indicators assess the decision-making and working arrangements the record supports. This delineating and layered interpretation explains a specific research purpose for each measure.

4.10 Interpretation through an organizational application

The application logic is shown in a shift and maintenance assignment. When it receives the authorized scheduling request, an agent recalls the asset requirement, maintenance procedure, technician qualifications, and technician availability records. The allocation criteria relate to the operational priority and to the relevant working arrangement. The recommendation record lists the proposed assignment and business rationale. The evidence is reviewed by a supervisor, and the decision or appropriate adjustment is noted. The roster and work order are confirmed by the scheduling and maintenance systems. Following data links work done to verifying the work completed and possible entitlement to remuneration. Allocation rules are linked to the institutional process by which they were agreed to using relevant consultation and agreement references.

Four analyses can be undertaken for the same episode: HR examines qualification aspects; labor management examines working time and entitlements; operations reviews assignment and completion factors; and industrial relations analyses agreement aspects and workforce input treatment. This illustrates the proposed analytical framework. Its organizational effects would be measured via the suggested area protocol. Coherent recording of the organizational record needs to be shared among the relevant management functions. The episode is coordinated by a process owner; authority for employability criteria, employment history, operational confirmations, and participation references is held by relevant functions. Technical teams provide the means and tools to collect and obtain organizational evidence. The results of the reconstruction activities can then be used to train, design, and clarify work rules.

5. Key Findings

They result from integrating 40 research studies and descriptively analyzing public workplace-survey aggregates, producing seven interconnected findings. These findings are grouped into four thematic areas: some reflect established research evidence, while others present the proposed framework and applications for empirical validation.

5.1 Action-Level Reconstruction and Delegated Authority

The basic unit for reconstructing actions, organizational decisions, participants' responsibilities, and verified outcomes is an action episode. A reviewable episode is defined as a set of information linking an authorized purpose with rules, information sources, system configuration, execution sequence, human interventions, and confirmed results. Technical accountability records also do not confirm whether an action was approved, what institutional conditions applied, or whether it effectively produced the hoped-for change. The study offers eight dimensions for integrating mandate and authority, governing rules, lineage of inputs, system configuration, execution and state change, human activity, workforce output, operational output, and evidence stewardship, and for integrating them into a coherent reconstruction.

It is also important to clearly record delegated authority and the human factor. Existing research has identified three types of decision-making systems: autonomous, sequential, and collaborative, and suggests the importance of defining and assigning decision rights and responsibilities. Managerial accountability, oversight mechanisms, and organizational interpretation of the autonomous agent's actions are considered through task allocation, reviewer authority, evidence considered, human interventions, and final dispositions as a suggested pathway.

5.2 Workforce Decisions and Employee Participation

Workforce evidence is a vital component of operational provenance. Employment criteria and employee qualifications are used to inform decisions made about the organization and its implications for employees and operational delivery. Assignment rules, working-time arrangements, and remuneration align with employment criteria and employee qualifications. Linking these elements by a common episode identifier allows for an analytical foundation to be conceived regarding recruitment, scheduling, performance assessment, employee development, and payroll decision-making. These relationships are based on suggested evidence rather than actual improvements in workforce outcomes.

To understand the scope of autonomous decisions, one must also consider roles and responsibilities of employees and the institutional set-up. Favorable employee-reported performance outcomes have been published with the consultation of workplace research associates. Similarly, where consultation took place, public employer-survey data reveal more positive assessments in organizations reporting this for worker productivity, satisfaction, performance measurement, and health and safety. The differences calculated in this study range from 9.30 to 15.00 percentage points. These observational correlations cannot be used to determine causal relationships. However, they agree that consultation documentation, management response, and collective agreement and workplace rules implemented can have analytic value and should be maintained in the proposed provenance format.

5.3 Operational Performance and Governance Requirements

Different measures are needed to evaluate operational performance and have well-understood conditions for

comparison. Reported improvements in information availability are separated from reported improvements in decision-making speed, decision-making quality, and managerial autonomy using published managerial indicators. Studies on human-AI collaboration also highlight task appropriateness, worker competency, delegation plans, and baseline worker performance as key factors in successful collaboration. Therefore, the proposed design focuses on task conditions, computer contributions, human contributions, execution records, and human-confirmed task outcomes. This approach would facilitate measuring the impact of performance once it is complete, but it does not assume that measuring it will lead to organizational improvement.

Organizations that have been established may have governance practices that can be used to reconstruct the organization. Workplace indicators have been published that identify implementation guidelines, worker consultation, and regular auditing as governance structures. Connecting policy versions, consultations, audit outcomes and authorizations, and specific agent actions would enable reflection on how organizational requirements are applied. However, governance, but not necessarily full reconstruction, accountability, or operational effectiveness, can be deduced only from the implementation of governance.

5.4 Proposed Evaluation Measures and Empirical Validation

The proposed framework represents an enterprise evaluation protocol that can be empirically studied in the future. The seven proposed evaluation measures are: evidence coverage, success of reconstruction, reconstruction time, traceability of workforce rules, verified transaction coverage, review quality, and review consultation traceability. Clearly defined eligibility criteria and appropriate denominators are required for each measure, and evidence requirements and assessment procedures must be consistent between organizational episodes.

Future empirical studies might focus on assessing whether actions of which the agents themselves are truly autonomous can be reconstructed by an authorized viewer based on the dimensions of evidence mentioned. In enterprise implementation studies, records or their absence can be evaluated; the accuracy of their reconstruction, and the efficiency and helpfulness of the review can be studied, and organizational explanations can be assessed. Future studies might further explore factors such as relationships between recursive quality, workforce decision-making, employee experience, and operational performance, as well as differences between organizational contexts. This framework offers a structured, analytic approach to autonomous-agent provenance issues in a regulated enterprise. Future implementation studies and systematic empirical validation are still to be conducted to assess its practical effectiveness, measurement reliability, and generalizability.

6. Limitations

The literature review is bounded, integratively collected, purposefully assembled, and synthetic, not a systematic census of studies on provenance, HR, labour, operations, and industrial relations. The 10 publications were chosen for each analytical group, favoring complementary views, while potentially excluding relevant evidence or contrary evidence. These eight dimensions are retained in this corpus; whether they are exhaustive and stable across the fuller body of research for the supported constructions cannot be inferred. Independent dual-coder agreement testing and formal saturation analysis are not documented; evidence mapping should be interpreted as coders' interpretation of and rationale for re-creating the instrument, rather than as a validated coding instrument.

The quantitative analysis follows the published overall percentages and comparisons, rather than individual participants. Therefore, this study cannot reweight estimates, make adjustments for group differences due to differences in workforce composition and organizational characteristics, estimate uncertainty separately, or measure intrastructure variation. Productivity and satisfaction measures are self-perceptions by employers and workers and may include memory and experience biases, interpretation or response biases, and may not reflect independently verified performance. In particular, the percentage-point employer consultation contrasts from 9.30 to 15.00 are associations, not causal links, meaning they do not show that better consultation necessarily resulted in better outcomes, or that the proposed provenance framework caused better outcomes.

Much of the workforce work focused on support for AI or algorithmic management, but not necessarily on delegated autonomous agents that manipulate equipment or make persistent changes to business records. The directly agent-related research in the corpus consists of early technical demonstrations and practitioner talks, which provide context but do not test this framework, as they occur in an unregulated enterprise (Souza et al., 2025; OECD, 2026). Its transferability will vary depending on the industry, jurisdiction, collective arrangements, and the size of the organization, system integration, record retention, and privacy requirements. Therefore, the proposed dimensions might need to be further refined for specific decision classes and legal contexts.

Enterprise implementation, actual reconstruction trial, independent reviewer scoring exercise, and the validation of seven proposed evaluation measures were not conducted. Further research is needed to prespecify episode and sample requirements, calibrate retrieval and reconstruction with authentic multi-system agent records, quantify projected reviewer agreement and failure rates, and demonstrate workforce-operational outcomes by decision class and site. Until these steps are taken, the framework and measures presented are suggested ways to test empirically, not proven conformity monitoring measures or performance interventions.

7. Conclusion

This research synthesizes 40 research papers to provide an integrated provenance of the actions autonomous agents perform in regulated enterprises. It also conducts a descriptive secondary analysis of publicly available workplace information. It proposes a scaffold of 8 potential dimensions for reconstruction: mandate, authority, governing rules, input lineage, system configuration, execution and state change, human participation, workforce and operational outcomes, and evidence stewardship. This main element centers on the duality of technical execution and the structures of people and work. Public data analysis provides solid context when making management decisions. The employer consultation comparison shows that employers are 10.75 percentage points more likely to agree that the productivity assessment is good, that workers are happy, and that they can measure worker performance, health, and safety.

Additional capabilities and processes are identified by training, managerial decision, and governance indicators for relevance to implementation. These descriptive findings inform the incorporation of workforce capability, participation, and/or review in the proposed provenance framework, and the broader research indicates the significance of task context and the partition of combined human and AI contributions. The implications are domain-specific for the four domains of management. In HRM, provenance links job requirements, assessment evidence, and development opportunities to authorized decisions. In labor management, allocation, actual work, and remuneration relate to applicable employment arrangements. Agent execution and handovers in operations management relate to confirmed outcomes.

Information and consultation, agreements, and organizational responses in industrial relations are linked to the rules put in place in daily work. A methodological contribution that can be reproduced is also offered by the study. The 40-study evidence matrix distinguishes published findings from the organizational applications that have been drawn from them. The public-data calculations are traceable using source-cell references, and seven proposed evaluation measures convert the proposed framework into an enterprise research protocol. Implementation records and participants' perspectives added to this protocol can be used in subsequent studies to measure reconstruction quality, workforce decision process, and operational performance. The proposed framework could be evaluated by regulated enterprises to ascertain whether it facilitates the ability to reconstruct autonomous actions: what the agent did, who authorized the action, what the evidence and arrangements were that informed the decision, how parties contributed, and what organizational outcomes resulted from such actions

References

1. Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile. NIST AI 600-1. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.600-1

2. Bankins, S., & Formosa, P. (2023). The ethical implications of artificial intelligence (AI) for meaningful work. Journal of Business Ethics, 185, 725–740. https://doi.org/10.1007/s10551-023-05339-7

3. Bansal, G., Wu, T., Zhou, J., Fok, R., Nushi, B., Kamar, E., Ribeiro, M. T., & Weld, D. S. (2021). Does the whole exceed its parts? The effect of AI explanations on complementary team performance. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, 1–16. Association for Computing Machinery. https://doi.org/10.1145/3411764.3445717

4. Brynjolfsson, E., Li, D., & Raymond, L. (2025). Generative AI at work. The Quarterly Journal of Economics, 140(2), 889–942. https://doi.org/10.1093/qje/qjae044

5. Buneman, P., Khanna, S., & Tan, W.-C. (2001). Why and where: A characterization of data provenance. In J. Van den Bussche & V. Vianu (Eds.), Database Theory—ICDT 2001 (Lecture Notes in Computer Science, Vol. 1973, pp. 316–330). Springer. https://doi.org/10.1007/3-540-44503-X_20

6. Cheney, J., Chiticariu, L., & Tan, W.-C. (2009). Provenance in databases: Why, how, and where. Foundations and Trends in Databases, 1(4), 379–474. https://doi.org/10.1561/1900000006

7. Davidson, S. B., & Freire, J. (2008). Provenance and scientific workflows: Challenges and opportunities. In Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data (pp. 1345–1350). Association for Computing Machinery. https://doi.org/10.1145/1376616.1376772

8. Dell'Acqua, F., McFowland, E., III, Mollick, E., Lifshitz, H., Kellogg, K. C., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2026). Navigating the jagged technological frontier: Field experimental evidence of the effects of artificial intelligence on knowledge worker productivity and quality. Organization Science, 37(2), 403–423. https://doi.org/10.1287/orsc.2025.21838

9. Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. Business & Information Systems Engineering, 61(5), 637–643. https://doi.org/10.1007/s12599-019-00595-2

10. Duggan, J., Sherman, U., Carbery, R., & McDonnell, A. (2020). Algorithmic management and app-work in the gig economy: A research agenda for employment relations and HRM. Human Resource Management Journal, 30(1), 114–132. https://doi.org/10.1111/1748-8583.12258

11. European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Adopted text of 13 June 2024, Articles 12, 14 and 26 and Annex III. Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

12. Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive challenges in human–artificial intelligence collaboration: Investigating the path toward productive delegation. Information Systems Research, 33(2), 678–696. https://doi.org/10.1287/isre.2021.1079

13. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2021). Datasheets for datasets. Communications of the ACM, 64(12), 86–92. https://doi.org/10.1145/3458723

14. Herschel, M., Diestelkämper, R., & Ben Lahmar, H. (2017). A survey on provenance: What for? What form? What from? The VLDB Journal, 26(6), 881–906. https://doi.org/10.1007/s00778-017-0486-1

15. Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organizational decision making. Business Horizons, 61(4), 577–586. https://doi.org/10.1016/j.bushor.2018.03.007

16. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. Academy of Management Annals, 14(1), 366–410. https://doi.org/10.5465/annals.2018.0174

17. Lane, M., Williams, M., & Broecke, S. (2023). The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers. OECD Social, Employment and Migration Working Papers, No. 288. OECD Publishing. https://doi.org/10.1787/ea0a0fe1-en

18. Langer, M., König, C. J., & Papathanasiou, M. (2019). Highly automated job interviews: Acceptance under the influence of stakes. International Journal of Selection and Assessment, 27(3), 217–234.

https://doi.org/10.1111/ijsa.12246

19. Lee, M. K., Kusbit, D., Metsky, E., & Dabbish, L. (2015). Working with machines: The impact of algorithmic and data-driven management on human workers. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (pp. 1603–1612). ACM. https://doi.org/10.1145/2702123.2702548

20. Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I., & Kasper, G. (2019). The challenges of algorithm-based HR decision-making for personal integrity. Journal of Business Ethics, 160(2), 377–392. https://doi.org/10.1007/s10551-019-04204-w

21. Marler, J. H., & Boudreau, J. W. (2017). An evidence-based review of HR Analytics. The International Journal of Human Resource Management, 28(1), 3–26. https://doi.org/10.1080/09585192.2016.1244699

22. Meijerink, J., & Bondarouk, T. (2023). The duality of algorithmic management: Toward a research agenda on HRM algorithms, autonomy and value creation. Human Resource Management Review, 33(1), Article 100876. https://doi.org/10.1016/j.hrmr.2021.100876

23. Milanez, A., Lemmens, A., & Ruggiu, C. (2025). Algorithmic management in the workplace: New evidence from an OECD employer survey. OECD Artificial Intelligence Papers, No. 31. OECD Publishing. https://doi.org/10.1787/287c13c4-en

24. Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model cards for model reporting. In Proceedings of the Conference on Fairness, Accountability, and Transparency (pp. 220–229). Association for Computing Machinery. https://doi.org/10.1145/3287560.3287596

25. Moreau, L., & Missier, P. (Eds.). (2013). PROV-DM: The PROV Data Model. W3C Recommendation, 30 April 2013. https://www.w3.org/TR/prov-dm/

26. Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. Science, 381(6654), 187–192. https://doi.org/10.1126/science.adh2586

27. OECD. (2023). OECD Employment Outlook 2023: Artificial Intelligence and the Labour Market. OECD Publishing. Figure 7.3, Panel B. https://doi.org/10.1787/08785bba-en. Associated data: https://stat.link/34bepy.

28. OECD. (2026). Agentic AI in organisations: Early insights from practitioner interviews. OECD Artificial Intelligence Papers, No. 65. OECD Publishing. https://doi.org/10.1787/1257a26f-en

29. Parent-Rocheleau, X., & Parker, S. K. (2022). Algorithms as work designers: How algorithmic management influences the design of jobs. Human Resource Management Review, 32(3), Article 100838. https://doi.org/10.1016/j.hrmr.2021.100838

30. Parker, S. K., & Grote, G. (2022). Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. Applied Psychology, 71(4), 1171–1204. https://doi.org/10.1111/apps.12241

31. Pasquier, T., Han, X., Goldstein, M., Moyer, T., Eyers, D., Seltzer, M., & Bacon, J. (2017). Practical whole-system provenance capture. In Proceedings of the 2017 Symposium on Cloud Computing (pp. 405–418). Association for Computing Machinery. https://doi.org/10.1145/3127479.3129249

32. Raisch, S., & Fomina, K. (2025). Combining human and artificial intelligence: Hybrid problem-solving in organizations. Academy of Management Review, 50(2), 441–464. https://doi.org/10.5465/amr.2021.0421

33. Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. Academy of Management Review, 46(1), 192–210. https://doi.org/10.5465/amr.2018.0072

34. Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. In Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (pp. 33–44). Association for Computing Machinery. https://doi.org/10.1145/3351095.3372873

35. Rani, U., Pesole, A., & González Vázquez, I. (2024). Algorithmic management practices in regular

workplaces: Case studies in logistics and healthcare (EUR 31846 EN; JRC136063). Publications Office of the European Union. https://doi.org/10.2760/712475

36. Rosenblat, A., & Stark, L. (2016). Algorithmic labor and information asymmetries: A case study of Uber's drivers. International Journal of Communication, 10, 3758–3784. https://ijoc.org/index.php/ijoc/article/view/4892

37. Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. California Management Review, 61(4), 66–83. https://doi.org/10.1177/0008125619862257

38. Simmhan, Y. L., Plale, B., & Gannon, D. (2005). A survey of data provenance in e-science. ACM SIGMOD Record, 34(3), 31–36. https://doi.org/10.1145/1084805.1084812

39. Souza, R., Gueroudji, A., DeWitt, S., Rosendo, D., Ghosal, T., Ross, R., Balaprakash, P., & Ferreira da Silva, R. (2025). PROV-AGENT: Unified provenance for tracking AI agent interactions in agentic workflows. In Proceedings of the 2025 IEEE International Conference on e-Science (pp. 467–473). IEEE. https://doi.org/10.1109/eScience65000.2025.00093

40. Tabassi, E. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.AI.100-1

41. Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. California Management Review, 61(4), 15–42. https://doi.org/10.1177/0008125619867910

42. Vaccaro, M., Almaatouq, A., & Malone, T. (2024). When combinations of humans and AI are useful: A systematic review and meta-analysis. Nature Human Behaviour, 8(12), 2293–2303. https://doi.org/10.1038/s41562-024-02024-1

43. Vallas, S., & Schor, J. B. (2020). What do platforms do? Understanding the gig economy. Annual Review of Sociology, 46, 273–294. https://doi.org/10.1146/annurev-soc-121919-054857

44. Wood, A. J., & Lehdonvirta, V. (2021). Antagonism beyond employment: How the 'subordinated agency' of labour platforms generates conflict in the remote gig economy. Socio-Economic Review, 19(4), 1369–1396. https://doi.org/10.1093/ser/mwab016

45. Wood, A. J., Graham, M., Lehdonvirta, V., & Hjorth, I. (2019). Good gig, bad gig: Autonomy and algorithmic control in the global gig economy. Work, Employment and Society, 33(1), 56–75. https://doi.org/10.1177/0950017018785616