

OPERATIONS · HAULAGE, WAREHOUSING AND DISTRIBUTION

The Protected Constraint

Artificial intelligence across freight and warehouse operations, and the limits it must never be allowed to negotiate

Paul Forrest

Fractional Head of AI, Ecaveo
September 2026

About this paper

This paper opens the operations strand of the Ecaveo Responsible AI Series, alongside a trilogy on professional firms covering private equity, accounting and law. Each paper stands alone. All of them argue one thing, which is that the durable advantage comes from decisions that can be reconstructed afterwards, and that systems producing answers without producing a record destroy more value than they create.

Logistics differs from the three professional papers in a way that governs everything here. When a professional firm gets an output wrong, somebody is misadvised. When a haulier or a warehouse operator gets an output wrong, a vehicle moves, a load is secured, a person walks into an aisle. The constraints are physical, they are enforced by regulators with prosecuting powers, and no optimiser is entitled to trade against them.

Editorial balance. Roughly seven parts in ten of this document is evidence and sector analysis, two parts in ten is original framework, and one part in ten concerns how I would run the work. Where the evidence is thin, the paper says so, and in this sector it is thin in places the market does not admit.

HOW TO READ THE EVIDENCE TAGS

Every substantive claim carries a tag stating what kind of evidence supports it.

REGULATOR A regulator, government department or official statistics publisher speaking in its own voice.

STANDARD A published standard, statutory instrument or formal guidance document.

CASE LAW A judgment, tribunal decision or enforcement action.

PEER-REVIEWED A primary study, systematic review, legislative investigation or independent verification body working from primary documents.

SURVEY A survey or dataset with disclosed method and sample.

VENDOR RESEARCH Research or a survey from a body with a commercial or membership interest in the answer. Directional only.

PRACTICE Practitioner opinion, including my own.

EVIDENCE GAP A question the available evidence does not answer.

ECAVEO My own framework, proposed rather than proven.

Legal, regulatory and safety notice. This paper describes legislation, regulatory guidance, safety duties and enforcement action throughout, and Chapters 2 and 5 in particular. Nothing in this document constitutes legal advice, and nothing in it is a substitute for a competent person's risk assessment. The provision of legal advice sits outside the terms of any engagement with the author or with Ecaveo, and the material is presented to support discussion and further review by qualified advisers.

Statements about what a source says are attributed to that source. Where I draw a practical implication the source does not itself state, the text says so. Guidance and timetables in this area have moved repeatedly during 2026.

Contents

A note from the author 4

Executive summary 5

ONE Forecast, fact and instruction 8

TWO The constraints that do not move 11

THREE Where the value sits across the operation 16

FOUR Six ways this fails in an operation 19

FIVE The constraint library, and the control model 22

SIX Proving it, and the evidence gate 26

SEVEN The economics of verification 28

EIGHT Ninety days, then nine months 30

NINE The session to run with the board 32

TEN Where the evidence is thin 33

References 36

THE ARGUMENT IN ONE PARAGRAPH

Artificial intelligence is good at predicting what a logistics operation will face and poor at knowing what it is forbidden to do about it. The strongest deployments therefore keep three things apart, which are the forecast, the operational fact and the instruction, and they hold the hard limits on driving hours, load security, site safety and vehicle condition outside the model where a person owns them and a test proves they hold. The recommendation in this paper is to build the constraint library before buying the optimiser, because a recommendation that breaches a limit is worse than no recommendation, and because the limits are the part a regulator will ask about.

ILLUSTRATIVE MATERIAL

Every operator, site, customer and person in the worked examples is invented. Each is marked as an illustrative scenario where it appears. Published research, official statistics and enforcement decisions are cited and numbered. Invented material is never presented as evidence.

CITATION

Forrest, P. (2026). *The Protected Constraint: Artificial Intelligence in Haulage, Warehousing and Distribution*. Ecaveo Responsible AI Series, Operations. Version 1.0, September 2026.

ALSO IN THIS SERIES

The Traceable Thesis. Private equity.

Sufficient and Appropriate. Accounting and audit.

Verified Authority. Law firms.

A NOTE FROM THE AUTHOR

The part that cannot be negotiated

Most sectors can absorb a confident wrong answer. An operation that moves 44 tonnes at 56 miles an hour cannot, and that single fact should shape the whole of the deployment.

I came to this through business process work in the 1990s and moved into machine learning and language models from 2014. I build small, specific models in preference to large general ones, because a model trained on everything is calibrated for nothing, and because an operation rewards narrow reliability far more than it rewards breadth.

What struck me when I started looking properly at this sector was the gap between two bodies of material. On one side sits a large and confident technology market selling optimisation and autonomy, with case studies and adoption figures attached. On the other sits the official record, which is unusually good in this sector because freight is measured. The Department for Transport counts the tonne kilometres. The Traffic Commissioners count the licences. The Health and Safety Executive counts the people who were killed.^{16,19,12} **REGULATOR** Set the two side by side and the second is a great deal less flattering than the first.

That is the honest starting point. Empty running across UK road freight stood at 30 per cent of vehicle distance in the year to March 2026, and both goods moved and vehicle distance fell by 1 per cent over a period of heavy promotional claims about optimisation.¹⁶ **REGULATOR** Whatever the technology is doing, it has not yet shown up in the national numbers.

This paper argues that it can, and that the route to it runs through a discipline the sector already understands better than most. Operations people know what a hard limit is. They know the difference between a plan and what actually happened. Applied to the technology, that instinct gets a firm most of the way. The rest is the tedious part, which is writing the limits down.

TWO PAGES THAT CIRCULATE ON THEIR OWN

Executive summary

The adoption claims and the measured outcomes are irreconcilable. The regulatory ground is firmer than most operators realise, and it is about to get firmer still.

Three things are true at once in this sector, and holding all three is the beginning of a sensible plan. The technology genuinely predicts better than what came before. The operational benefit of that prediction is undemonstrated outside vendor case studies. The constraints that govern what an operator may do with a prediction are legally enforceable, physically real and, in one important respect, weaker in Great Britain than in the European Union.

1

The adoption numbers in circulation cannot both be right. The Office for National Statistics, surveying 38,637 businesses in June 2026, put artificial intelligence use among UK businesses with 10 or more employees at about 35 per cent.¹³ **REGULATOR** A vendor survey of 300 European decision-makers, 75 in each of four markets, reported that only 3 per cent of shippers and logistics providers did not use it.¹⁴ **VENDOR RESEARCH** A gap of that size is a measurement problem. Geography does not explain it.

2

Buyers themselves cannot see the return. A survey of 394 supply chain professionals published in August 2026 found that 55 per cent of chief supply chain officers were unclear on the return from their investments.¹⁵ **VENDOR RESEARCH** A separate global study of 1,719 respondents reported that while roughly nine in ten organisations used artificial intelligence somewhere, 37 per cent could attribute any effect on earnings to it, and 6 per cent met the study's definition of a high performer, meaning they attributed at least 5 per cent of earnings before interest and tax to it and described the effect as significant.¹⁷ **VENDOR RESEARCH**

3

The prediction itself is real, and it was proved on a forecasting problem. In a blind forecasting competition over 42,840 daily sales series, entered by 5,507 teams, the winning machine learning method beat the best statistical benchmark by 22.4 per cent, and all of the top 50 entries beat it by more than 14 per cent.¹⁸ **PEER-REVIEWED** That is a genuine result about accuracy. It says nothing about on-time delivery.

4

The safety regulator has stated its position, and it is technology-neutral by design. In a policy statement of 12 June 2026 the Health and Safety Executive set out how it regulates artificial intelligence, noting that the goal-setting nature of the 1974 Act means it applies regardless of the technology in use, and that duty holders must assess the risk and control it so far as is reasonably practicable.⁸ **REGULATOR** No new duty has appeared. The existing one already reaches this.

5

Worker monitoring is the live legal question, and the rules are moving in two directions at once. The Data (Use and Access) Act 2025 widened the lawful bases available for significant automated decisions, while the safeguards on transparency, human intervention and contestability remain.¹⁰ **REGULATOR** Meanwhile a government consultation on workplace monitoring technologies, open until 30 September 2026, is testing a statutory duty to consult workers about them.¹¹ **REGULATOR**

6

Driverless freight is not a near-term operating decision in this country. The most advanced operator had accumulated about 440,000 driverless miles by June 2026 on its own account,²⁸ **VENDOR RESEARCH** against 19.0 billion kilometres run by UK heavy goods vehicles in a single year.¹⁶ **REGULATOR** Britain's self-driving pilot scheme, opened in March 2026, is framed around passenger services and does not provide a route for goods vehicles.²⁷ **STANDARD**

What follows for an operator

Build the constraint library first. Driving hours, load security, site rules, vehicle condition, equipment interlocks and site-specific prohibitions belong in a versioned, tested, owned artefact that sits outside any model and that no optimiser can overrule. That artefact is worth more than the optimiser; it outlives every vendor, and it is the thing a Traffic Commissioner or an inspector will ask to see.

Then separate the three kinds of output. A forecast is a statement about the future with a confidence attached. An operational fact is something that happened, with a source. An instruction is a thing a person is being asked to do. Systems that blur these, and most do, are how a prediction becomes an order without anybody deciding that it should.

Start where a wrong answer is recoverable and a check is cheap. Demand and capacity planning qualifies, and so does exception explanation. Anything that paces a person, controls equipment or touches a roadworthiness decision does not belong in a first wave.

Chapter 8 sets out a 90-day proof on three bounded workflows, run in shadow mode before any live use. Chapter 10 states what would change the position taken here.

CHAPTER ONE

Forecast, fact and instruction

Almost every failure in this field begins when one of these three is mistaken for another. Keeping them apart is a design decision, and it is cheap to make at the start.

A logistics system deals in three kinds of statement and treats them very differently. A forecast says that a vehicle will probably arrive late. An operational fact says that the vehicle left the depot at 06.42, which is either true or wrong and has a source behind it somewhere. An instruction says that the planner should reallocate the drop. The three carry different burdens of proof and different consequences, and a screen that renders them in the same typeface has already lost the distinction.

The point holds across every application in the sector. A late-arrival prediction, a maintenance alert or a recommended pick path can each improve a decision, and each can equally propagate bad master data, miss a safety constraint or optimise one node at the expense of the network.

What makes this harder in operations than in an office is the pace. A planner handling 300 loads has seconds per decision and will take the recommendation because the recommendation is there. The question is not whether people will over-rely on these systems. They will, under time pressure, as they always have with every previous generation of planning tool. So the question is what the system does to make the basis of a recommendation visible in the seconds actually available. Few of them do much.

THE CORE CLAIM

A forecast becomes an instruction the moment somebody acts on it without deciding to. That transition is where the accountability goes missing.

1.1 The cycle this series uses

Every paper in this series describes the same five-stage cycle, because it is the sequence a decision follows whether or not anybody writes it down. In an operation the third stage carries an obligation the professional papers do not have, which is that the judgement is bounded by limits somebody else has set and a regulator can enforce.

THE CYCLE EVERY PAPER IN THIS SERIES EXAMINES

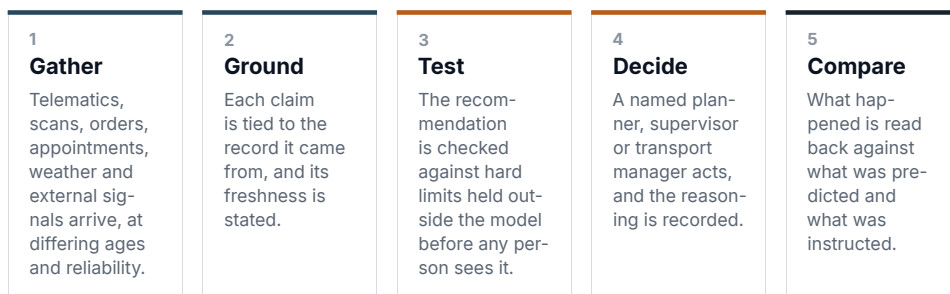


FIGURE 1 The third stage is the one this paper is named after. A system that reaches stage four without passing stage three is not assisting a decision, it is making one.

Ecaveo framework, used across the series. **ECAVEO**

1.2 What the technology is genuinely good at

There is one rigorous, independent result worth knowing, and it deserves to be stated accurately because it is routinely overstated. In the M5 forecasting competition, 5,507 teams forecast 42,840 daily unit sales series drawn from 3,049 products across 10 stores, evaluated blind against a held-out period. The winning submission improved on the best exponential smoothing benchmark by 22.4 per cent, and every one of the top 50 submissions beat that benchmark by more than 14 per cent.¹⁸ **PEER-REVIEWED**

Four qualifications travel with that result. It measured forecast accuracy and no operational outcome at all. The data was United States grocery retail, a long way from UK freight or warehousing. Competition conditions gave entrants clean data, unlimited effort and no operating constraints to work around. The gains came overwhelmingly from gradient-boosted tree methods, which bear little resemblance to the generative systems currently being sold.

The best independent evidence in this field proves that machines forecast well. It does not prove that anything arrives sooner.

1.3 What the national numbers show

Freight is measured, which makes this sector unusually testable. In the year to March 2026, goods vehicles registered in Great Britain moved 164 billion tonne kilometres, lifted 1.56 billion tonnes and covered 19.0 billion vehicle kilometres. Goods moved and vehicle distance each fell by 1 per cent. Empty running stood at 30 per cent of total vehicle distance, amounting to 5,714 million kilometres run empty.¹⁶ **REGULATOR**

That last figure is the most useful number in this paper. It is officially measured, it represents the clearest available headroom in UK road freight, and it has proved stubborn through a decade of optimisation software. Empty running sits at 30 per cent, or 5,714 million kilometres, and it has been flat for years.

Government has identified the ground. The freight plan published in June 2022 set out five priority areas, one of which is data and technology, together with a £7m innovation fund to accelerate adoption.⁴⁷ **REGULATOR** A successor multimodal plan was announced in April 2025 and had not been published as at the date of this paper. **EVIDENCE GAP**

The sector itself is large enough for the question to matter. Workforce jobs in transport and storage stood at about 1.92 million in the first quarter of 2026.²⁰ **REGULATOR** There were 66,222 goods vehicle operator licences in force in Great Britain in the year to 31 March 2025, authorising 376,044 vehicles.¹⁹ **REGULATOR** The trade association Logistics UK puts the sector's contribution at £175bn and 8 per cent of the workforce, although it does not publish the derivation and the figure should be attributed to the association and not treated as an official statistic.²¹ **VENDOR RESEARCH**

1.4 A statistic to stop repeating

One number appears in almost every discussion of this sector and should not. The claim that the United Kingdom is short of 100,000 heavy goods vehicle drivers originated in a trade association letter of June 2021, built from a fall in test passes, a fall in European Union national drivers and an asserted pre-existing shortage of 60,000. The first two components trace to official data. The third does not, and when Full Fact asked the association to explain it, it could not.²² **PEER-REVIEWED**

The replacement is better and more useful. The Department for Transport asks operators directly. In the first quarter of 2026, 616 of 2,424 businesses surveyed, or 25 per cent, reported driver vacancies, and 266 had raised wages in response.¹⁶ **REGULATOR** That is a measured figure with a stated sample, it points to a workforce renewal problem, and it is the number a board should plan against.



Demand and capacity planning. The forecasting gains are real and they were demonstrated on accuracy, not on service.

CHAPTER TWO

The constraints that do not move

The obligations here are older than the technology, enforced by regulators with prosecuting powers, and in one respect looser in Great Britain than across the Channel.

A reminder on register. This chapter summarises what published sources say and grades how firmly they say it. I am not a lawyer and this is not a statement of the law. Where I draw an implication a source does not itself state, the sentence says so, and any point bearing on an operator's own position should be confirmed with qualified advisers.

2.1 Safety, and the duty that already reaches this

The Health and Safety Executive published a policy statement on its regulatory approach to artificial intelligence on 12 June 2026. It describes three routes by which the technology falls within its remit, which are workplace use affecting health and safety, product safety for artificial intelligence in workplace machinery where the Executive acts as market surveillance authority, and effects on its wider remit. On the legal position it is direct, noting that the goal-setting nature of the Health and Safety at Work etc. Act 1974 means it applies regardless of the technology being used.⁸ **REGULATOR**

Two days earlier the Executive announced that it would produce Britain's first joint industry guidance on collaborative robotics, working with the Regulatory Innovation Office, Automate UK and the Manufacturing Technology Centre, with the first stage due to launch in the summer of 2026. I could find no published guidance as at the date of this paper.⁹ **REGULATOR** The stated problem is instructive for anyone planning a warehouse automation programme. There is no legal barrier to adopting these systems, and the Executive's assessment is that fear of non-compliance is limiting adoption.

The underlying workplace transport framework is unchanged and remains the reference point for any site where vehicles and people share space. The Executive groups the duties into three main areas, which are the site in its design and its activity, the vehicle, and the driver; covering risk assessment, site layout and traffic routes, separation of pedestrians from vehicles, driver competence, vehicle selection and maintenance, and the management of reversing.⁷ **REGULATOR**

The scale of what that framework is holding back is worth stating. In 2025 and 2026, on provisional figures published in June 2026, 126 workers were killed in work-related accidents in Great Britain, of whom 24 were struck by a moving vehicle.¹² **REGULATOR** In the preceding year 59,219 employee injuries were reported under the reporting regulations, against a Labour Force Survey estimate of 680,000 workers sustaining an injury, a divergence the Executive attributes largely to under-reporting.¹² **REGULATOR**

REGULATOR

Workplace transport guidance groups the duties into three areas, which are the site, the vehicle and the driver. My own view is that an automation programme should be assessed against all three, and not just the one it touches directly.

2.2 Drivers' hours, and the vocabulary that has changed

The limits are unchanged in substance and the language has moved. Government guidance no longer refers to European Union rules, and now describes the assimilated rules, previously called the EU rules.²³ **STANDARD** A 2026 operating procedure that still says otherwise is showing its age.

THE LIMITS AN OPTIMISER MAY NOT NEGOTIATE

LIMIT	FIGURE
Daily driving	Nine hours, extendable to 10 hours twice a week
Weekly driving	56 hours
Fortnightly driving	90 hours in any two consecutive weeks
Break	A break or breaks totalling at least 45 minutes after no more than four and a half hours of driving
Daily rest	11 hours, reducible to nine hours three times between any two weekly rest periods
Weekly rest	45 hours unbroken, reducible to 24 hours every other week

TABLE 1 These are the assimilated rules for vehicles over 3.5 tonnes. Separate domestic rules apply outside that scope. Any routing or scheduling recommendation that breaches one of these is a defect.

Recording has moved on as well. Since 1 July 2026 a full smart tachograph 2 must be fitted to goods vehicles over 2.5 tonnes used for international journeys for hire and reward, with transitional units permitted in vehicles registered before 24 December 2025.²⁴ **STANDARD** Vehicles in purely domestic operation are outside that retrofit requirement.

Better recording is not better governance, and a case decided in July 2026 makes the point more sharply than any argument could. A Traffic Commissioner revoked an operator's licence and disqualified its sole director and transport manager indefinitely after he repeatedly falsified drivers' hours records by using another driver's tachograph card, conduct the Commissioner described as serious and deliberate falsification striking at the heart of the drivers' hours enforcement regime.²⁵ **CASE LAW** The recording technology worked exactly as designed throughout.

2.3 The transport manager, and why the duty is personal

Statutory guidance from the Senior Traffic Commissioner, at version 17 since March 2025, sets out what continuous and effective management means. The need for a proper, active transport manager is described as not a mere formality but a serious requirement. Continuous and effective responsibility is said to mean just that. A transport manager must be more than a transport manager in name.²⁶ **STANDARD**

The listed duties include managing, auditing and reviewing compliance systems to ensure they are effective, implementing effective fleet management functions to ensure safe operation, and downloading data from vehicle units at least every 90 days and from driver cards at least every 28 days.²⁶ **STANDARD** An assessment annex asks what technology and management systems will be used.

What the guidance does not do is prescribe telematics, automated analysis or any particular system, and I found no statutory guidance and no tribunal decision addressing automated compliance systems directly. **EVIDENCE GAP** My own view, which the guidance does not state, is that this silence should be read as neutral. The duty is to exercise continuous and effective

management. Technology is a means of discharging it and has never been recognised as a substitute for it.

2.4 Roadworthiness, and the limit of predictive maintenance

This is where the sector's enthusiasm most often outruns the rulebook. The guide to maintaining roadworthiness requires a meaningful walkaround check by a driver or responsible person, preferably immediately before a vehicle is used and at minimum once in every 24 hours of operation, prompt written reporting of defects, and retention of records for at least 15 months. Safety inspections should, where practical, follow a time-based pattern, with indicative intervals running from four weeks for arduous or off-road work to 13 weeks for lightly loaded operation.⁵ **STANDARD**

The one continuous monitoring technology the guidance recognises is electronic braking performance monitoring, for trailers, and it is worth being precise about what that recognition extends to. A trailer fitted with such a system requires a laden brake test only where there is insufficient data to produce a report, or where work has been done on the braking system, whereas a trailer without one is expected to undergo a minimum of four laden brake tests spread evenly across the year.⁵ **STANDARD** The system therefore stands in place of the laden roller brake test. It does not stand in place of the safety inspection, and the guidance requires an evaluation of the performance report by a competent person before every one of them.^{5,6} **STANDARD**

The practical consequence is that predictive maintenance in road transport is currently an efficiency and an early warning, not a replacement for a regime. An operator planning to extend inspection intervals on the strength of condition monitoring is planning something the published guidance does not support, and would be well advised to test that with the Traffic Commissioner's office before doing it.

WORTH HIGHLIGHTING

No published guidance permits a predictive maintenance system to displace a scheduled safety inspection. An operator told otherwise by a supplier should ask to see the source.

2.5 Worker monitoring, where the ground is moving

Two things changed in opposite directions and both matter. The Data (Use and Access) Act 2025 replaced the previous restriction on solely automated decision-making, and the Information Commissioner's own explanation is that the Act opens up the full range of lawful bases for significant automated decisions, potentially including legitimate interests, so long as appropriate safeguards continue to apply and excepting special category data.¹⁰ **REGULATOR** The safeguards themselves did not go.

The Commissioner's guidance on monitoring workers through solely automated processes, still marked as under review because of that Act, requires that workers are told they are subject to such decision-making and given meaningful information about the logic involved and the significance and envisaged consequences. It requires alternatives for workers who ask for human intervention, and that those who ask are not disadvantaged compared with those who do not. Where a human is involved, the involvement must be meaningful, and reviewers should not just routinely apply the automated recommendation but should have meaningful influence, including the authority and competence to go against it. It adds that those providing oversight should remain engaged, critical and able to challenge the system's outputs.⁴ **REGULATOR**

The Commissioner has enforced this reasoning, although not yet in this sector. In February 2024 it issued enforcement notices requiring a leisure operator and associated trusts to stop using facial recognition and fingerprint scanning to monitor the attendance of more than 2,000 employees across 38 sites, finding that workers had not been proactively offered an

alternative, that the system was presented as a requirement in order to get paid, and that less intrusive means were available.²⁹ **CASE LAW**

Read across to a distribution centre and the reasoning transfers without much strain. I could find no enforcement action by the Commissioner concerning logistics, haulage or delivery workers, in a sector where handheld scanners, in-cab cameras and rate monitoring are close to universal. **EVIDENCE GAP**

2.6 What is coming, and the asymmetry with Europe

Three developments should be on a board agenda now.

A government consultation on workplace monitoring technologies opened on 8 July 2026 and closes on 30 September 2026. It covers technologies used to monitor, manage or make decisions relating to workers, expressly including algorithmic management and automated decision-making, and it is testing three options, one of which is a legislative duty to consult or negotiate with workers about them.¹¹ **REGULATOR** An operator with in-cab cameras and rate monitoring has a direct interest in the outcome and a narrow window in which to say so.

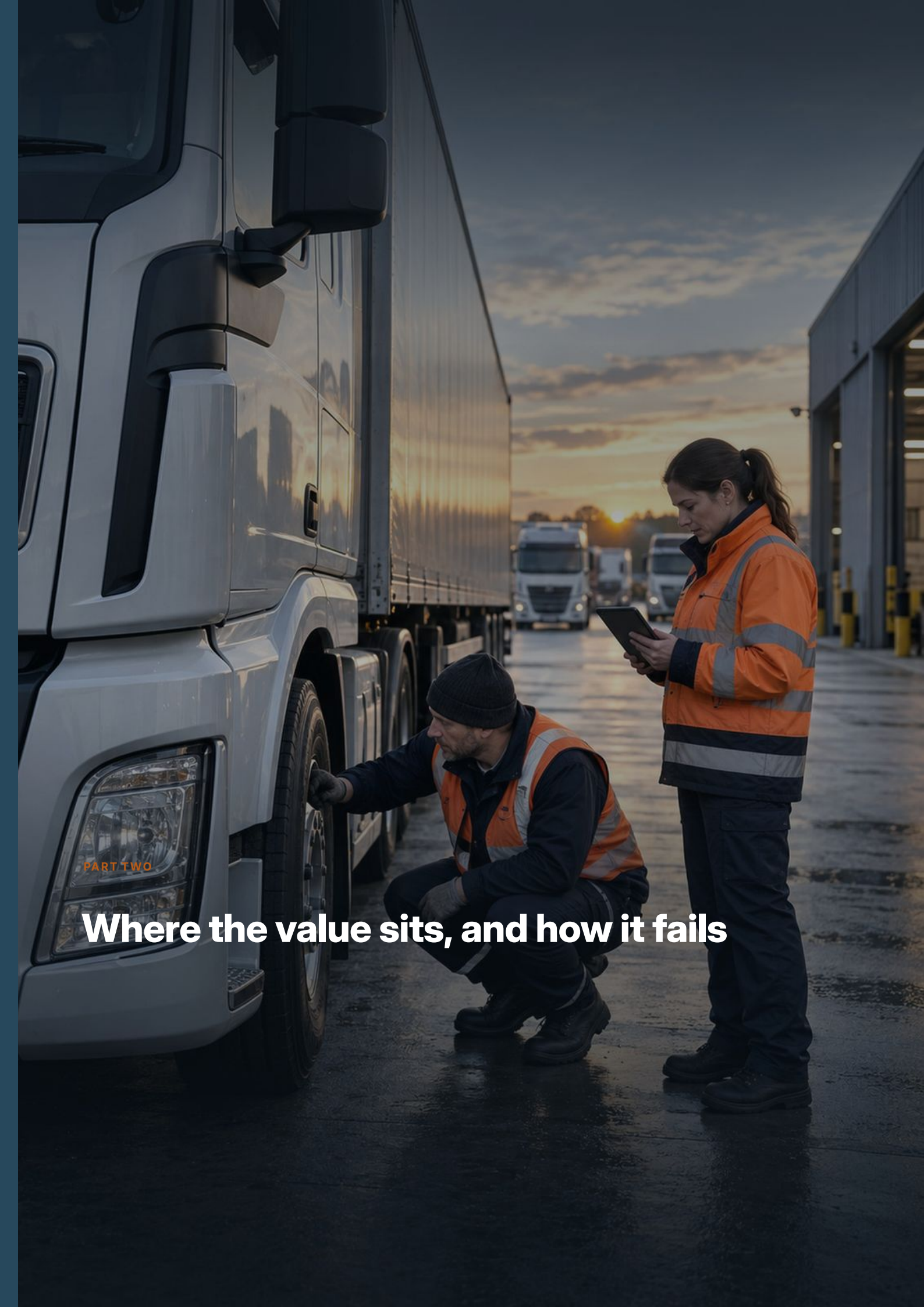
The European rules on high-risk systems are the ones most likely to reach a logistics operator, and the relevant category is employment. Annex III point 4(b) covers systems intended to make decisions affecting terms of work-related relationships, to allocate tasks based on individual behaviour or personal traits or characteristics, or to monitor and evaluate the performance and behaviour of persons in such relationships.² **STANDARD** That language describes warehouse task allocation, pick-rate monitoring and driver scoring rather closely. There is no standalone road transport category, and the critical infrastructure category is narrower than it looks, reaching safety components in the management and operation of road traffic, which does not describe an operator's own routing.

The timetable moved in July 2026. Regulation (EU) 2026/1744, in force from 27 July 2026, deferred the requirements for Annex III high-risk systems to 2 December 2027, with 2 August 2028 for systems embedded in regulated products.² **STANDARD** The prohibitions, the literacy duty, the general-purpose model rules and the transparency obligations are in force now.

THE ASYMMETRY A BRITISH OPERATOR SHOULD UNDERSTAND

From 20 January 2027 the European machinery regulation places safety components with fully or partially self-evolving behaviour using machine learning into a category requiring third-party conformity assessment, so self-certification is not available.³ **STANDARD** That regulation does not apply in Great Britain, where the Supply of Machinery (Safety) Regulations 2008 continue to apply. An amending instrument made on 22 July 2026 will permit machinery conforming to the European regulation onto the British market, and it comes into force on 20 January 2027 and not before.³ **STANDARD** The European regulation applies in Northern Ireland under the Windsor Framework.

My own reading, which no source states in these terms, is that a British operator buying automation faces a lighter formal assurance regime than a competitor in Dublin or Belfast, and that the sensible response is to ask suppliers for the European assessment anyway.



PART TWO

Where the value sits, and how it fails

CHAPTER THREE

Where the value sits across the operation

Rank the portfolio by two questions. Can a wrong answer be undone, and can a person check it in the time the job allows?

Those two tests sort this sector better than any taxonomy by function. A demand forecast that is wrong costs a planning cycle and can be revised. A task instruction that is wrong moves a person. The distance between those two is the whole of the risk gradient, and a portfolio built along it will produce value before it produces an incident.

THE OPPORTUNITY PORTFOLIO

WORKFLOW	WHAT THE TECHNOLOGY DOES WELL	THE VALUE TEST
Demand and capacity	Forecasts volume by lane, customer, product and window, and tests disruption scenarios	Lower avoidable overtime, idle capacity and service failure against a transparent baseline
Exception control	Joins orders, inventory, transport and external signals to explain a risk and propose recovery	Time to a validated explanation, and fewer manual status chases
Document and data work	Extracts from consignment paperwork, proof of delivery, claims and customs documents	Recall of known fields, error rate and the time a clerk spends agreeing back to source
Transport planning	Recommends routes, loads and sequences within legal, vehicle and customer limits	Fewer empty miles and late deliveries, with no recommendation ever breaching a limit
Fleet reliability	Combines inspections, defects, history and telematics to prioritise investigation	Earlier intervention and higher availability, with engineers deciding roadworthiness
Warehouse flow	Improves slotting, replenishment, waves, dock appointments and congestion forecasts	Throughput and travel distance, with safe separation of people, vehicles and machinery preserved

TABLE 2 Ordered by how recoverable a wrong answer is. The value test is stated in place of the benefit. A benefit measured against nothing will not survive a peak.

3.1 Demand and capacity is the right place to start

Forecasting is where the independent evidence is strongest, where a wrong answer is recoverable, and where the check is cheap because last year's actuals exist. An operator can take 24 months of history, hold out the final three, and measure the new method against whatever it uses now. That is a fortnight of work and it produces a number.

It is also where the benefit shows up as an absence, which makes it awkward to claim. Better volume forecasting reduces the agency shifts that were booked and not needed, the trunk that ran half full, the overtime authorised on Friday afternoon. None of those appear in a benefits case unless somebody baselines them first, and that is the single most common reason a good forecasting pilot fails to survive its first review. **PRACTICE**

3.2 Exception control, where the technology fits the shape of the job

Exception management is retrieval and explanation, which makes it unusually well matched to what these systems actually do. A customer asks where an order is. The answer requires joining an order record, a stock position, a vehicle location, a site appointment and possibly a weather or traffic signal, each held in a different system at a different age.

ILLUSTRATIVE SCENARIO

The pallet that was in three places

An invented regional distributor deploys an exception assistant across its order, warehouse and transport systems. Asked about a late consignment, it reports confidently that the pallet is on a vehicle two hours from delivery.

The pallet had been short-picked the previous evening and a replacement raised. The warehouse system knew this. The transport system was still carrying the original allocation, because the integration ran every four hours and the last run had been at 05.00. The assistant read both systems correctly and reconciled them by preferring the more recent record, which was the wrong one.

No component failed. The assistant was never told how old each source was, and so it could not say what it did not know. A freshness stamp on every retrieved field would have turned a confident answer into a useful one. The integration ran every four hours because that was the default when it was configured in 2021, and nobody had revisited it.

The design lesson generalises. Any system reading across operational sources should state, for every field it uses, where the value came from and how old it is. Data freshness is not a technical detail in a logistics deployment. It is the difference between an answer and a guess.

Underneath freshness sits identity, and the sector solved that problem long ago without always using the solution. The logistic label standard makes the serial shipping container code the single mandatory element on a label, giving every logistic unit a unique identifier that any party in the chain can read.⁴⁸ **STANDARD** An operation that identifies units consistently gives a retrieval system something to reason about. One that does not will find that the cleverest model available cannot tell it where a pallet is.

3.3 Transport planning, where the constraints do the work

Routing and load optimisation is the oldest application in this sector and it long predates the current wave. What has changed is the ability to incorporate messier inputs, such as site access notes written in free text, historical dwell times by customer and driver-reported access problems that previously lived in nobody's system.

The requirement is that every recommendation is checked against the hard limits before a planner sees it. Driving hours, vehicle dimensions and weights, licence categories, site curfews, hazardous goods rules, customer booking windows. Chapter 5 describes how I would hold those. The relevant point here is that a planning system which can produce a non-compliant plan will produce one, on a Friday in December, and the planner who accepts it will be the one answering for it.

3.4 What to leave alone in a first wave

Three things. Anything that sets or enforces the pace of a person's work, for reasons Chapter 4 sets out and because the legal ground is actively moving. Anything that controls equipment directly, which is a functional safety problem with its own standards and its own specialists. Anything that contributes to a roadworthiness decision, because the published guidance does not support substituting a prediction for an inspection.

None of those is a permanent prohibition. Each describes work where a wrong answer is hard to undo, hard to detect, or unsupported by the regime the operator is licensed under.



Loading dock flow. Congestion forecasting is valuable and advisory. The moment it controls a barrier or an appointment slot, the tier changes.

CHAPTER FOUR

Six ways this fails in an operation

Two of these six put people in front of vehicles. The other four cost money quietly, which is why they survive longer.

Risk registers in this field are usually organised by cause. The taxonomy below is organised by what the operation actually experiences, because that is what a depot manager can be asked to watch for. I am proposing the grouping rather than reporting it. **ECAVEO**

THE FAILURE TAXONOMY

FAILURE	WHAT ACTUALLY HAPPENS	ANNOUNCES ITSELF
Constraint breach	A route, schedule, load or task recommendation violates a legal, vehicle or site limit	Yes, if a person catches it
Cyber physical disruption	A compromised connector, vendor or agent reaches a warehouse or transport system and operations stop	Yes, immediately
Stale data confidence	The system reconciles sources of different ages and reports the result without stating what it did not know	No
Local optimisation	One site, lane or shift improves while the network absorbs the cost somewhere it is not measured	No
Pacing drift	Monitoring introduced for one purpose becomes the basis for evaluating, ranking or disciplining people	No, until a claim arrives
Manual mode decay	Planners and supervisors lose the ability to run the operation when the system is unavailable	No, until the day it matters

TABLE 3 The right-hand column is the point of the table. Only the first two announce themselves, which is why the other four need detection designed in. Vigilance will not find them.

4.1 Constraint breach, and why instructions are not advice

The first failure is the one the title of this paper addresses. A system that can recommend a non-compliant plan has a defect, whatever its accuracy on everything else. The remedy is architectural, and Chapter 5 sets it out.

What is worth saying here is that the usual mitigation fails. Training planners to spot non-compliant recommendations puts the check in exactly the place the system was bought to relieve. Under pressure, the recommendation wins.

4.2 Local optimisation, the failure nobody logs

A depot improves its dock turnaround by 14 minutes by holding inbound vehicles in the yard. The vehicles are now late for their next collection. The measure improves, the network worsens, and because the two are owned by different managers with different reports, nobody reconciles them.

This failure predates the technology by decades and the technology accelerates it, because an optimiser will exploit a local objective far more thoroughly than a person would. The control is to state the counter-measure when the objective is set, and to report the two together. It costs nothing at design time and is close to impossible to retrofit.

DESIGN RESPONSE

Every optimisation objective should name the measure it might damage, and that measure should be reported next to it.

4.3 Pacing drift, and what the strongest evidence shows

The most substantial evidence on productivity pacing and injury comes from a United States Senate committee investigation published in December 2024, which had access to internal company documents. It reported required rates including stowing 350 units an hour, picking 450 and packing 600 in some roles, enforced through time-on-task monitoring, and it found injury rates substantially above the industry average.³⁰ **PEER-REVIEWED**

The material that matters most is the company's own research. A 2020 internal project found that pausing disciplinary action for rate and giving workers more time reduced injury risk, and the committee reported that senior leadership declined to adopt it over concern about negatively impacting rate and productivity. A 2021 project identified a maximum number of repetitions per shift for a specified warehouse role, aimed at reducing back injury, which the committee reported workers routinely exceeded and which the company declined to implement.³⁰ **PEER-REVIEWED**

Two cautions belong with that. The committee was chaired by a legislator with a declared adversarial position towards the company, so the documentary findings are strong evidence and the characterisations are advocacy, and the company disputes the findings. Separately, a safety regulator citation of the same company in January 2023 concerned ergonomic and struck-by hazards under the general duty clause, with the regulator stating that the inspections found work processes designed for speed but not safety, and that action did not concern robotics.³¹ **REGULATOR**

Worth highlighting is what is missing. I could find no peer-reviewed epidemiological work on pacing and injury in warehousing, and no UK enforcement action involving robotics, automated guided vehicles or an algorithmically driven work process. **EVIDENCE GAP** The contrast between substantial American regulatory and legislative attention and no visible British equivalent is itself a finding, at a time when automated warehousing here is expanding.

4.4 Manual mode decay

The sixth failure is the one that does its damage on a single day. An operation that has planned by system for three years contains planners who have never planned without one, and a vendor outage, a ransomware event or a failed integration then meets a team without the muscle memory to run manually.

The evidence that this matters is not hypothetical. The independent body that categorises cyber events assessed the April 2025 attacks on two British retailers as a category two systemic event costing between £270m and £440m across affected parties, with one retailer's average daily consumer spend falling 22 per cent and online sales dropping to near zero. It specifically identified fragility in just-in-time inventory systems, difficulty rerouting goods under exclusive supply contracts, and concentrated risk in rural areas served by a single provider.³² **SURVEY**

WHERE EACH FAILURE IS CAUGHT

A	B	C	D
Before the output	In the output	At the boundary	In the operation
The constraint library rejects any recommendation that breaches a hard limit, before a person ever sees it.	Freshness stamps, stated confidence and named counter-measures expose stale data and local optimisation.	Separation of operational from corporate systems, least privilege and brokered access contain cyber physical disruption.	Consultation, purpose limits, rotation and manual-mode drills address pacing drift and skill decay.

FIGURE 2 Each failure has one natural interception point. Operators concentrate on the third column and under-invest in the first, which is the one that reaches the road.

Ecaveo framework. **ECAVEO**

4.5 A distinction worth preserving about automation incidents

Two incidents at the same automated grocery operator are routinely cited together and they are different in kind. A fire at one site in February 2019, which destroyed a facility handling roughly a tenth of the company's UK capacity at a reported cost of about £110m, was attributed jointly by the company, its insurer and the fire service to an electrical fault in a first-generation battery charging unit.³³ **CASE LAW** A fire at a second site in July 2021 was attributed by the company to the collision of three robots on the grid.³⁴ **CASE LAW**

Only the second is evidence about automated decision-making. The first is an electrical failure in a charging system, which could have happened to any fleet of battery-powered equipment. Conflating them overstates the case against automation, and a paper arguing for careful evidence should not do it.

CHAPTER FIVE

The constraint library, and the control model

One artefact matters more than any model in this sector. Build it first, own it internally, and let no supplier hold it.

A constraint library is a versioned, tested, human-owned statement of everything an operation is forbidden to do, held outside every model and applied to every recommendation before a person sees it. It is the most valuable thing a logistics operator can build in this field, and almost nobody has one. **ECAVEO**

5.1 What goes in it

Six categories cover most operations. Statutory limits, meaning driving hours, weights and dimensions, licence categories and dangerous goods rules. Vehicle and equipment limits, meaning what each asset can physically and lawfully do. Site rules, meaning curfews, access restrictions, banned manoeuvres, pedestrian separation and one-way systems. Customer and contractual commitments, meaning booking windows and service definitions. Competence rules, meaning what a given person is qualified and authorised to do. Absolute prohibitions, meaning the small set of things the operation will never do whatever the optimiser calculates.

Three properties make it work. It is versioned, so that a recommendation made in March can be tested against the rules that applied in March. It is testable, so that a regression set of known-bad plans is rerun after every change to a model, a supplier or the library itself. It is owned by a named person inside the operation, because a rule held only in a vendor's configuration is a rule the operator cannot evidence.

ILLUSTRATIVE SCENARIO

The rule that lived in a configuration file

An invented contract logistics operator changes transport management supplier after four years. During migration the team discovers that a rule preventing a particular trailer type from being planned into two customer sites, added after a near miss in 2022, exists only as a constraint in the old supplier's configuration.

Nobody in the operation can say who added it or why. The near-miss report is in a different system and does not mention the trailer. The migration team has three days and a choice between carrying a rule it cannot justify and dropping a rule that may exist for a good reason.

The rule was correct. It cost two days of investigation to establish that, and it would have cost nothing at all had it been written down outside the tool with a reason attached.

5.2 Risk tiers

Controls should scale with what the output touches and how reversible it is. Sophistication is beside the point.

RISK TIERS AND MINIMUM CONTROLS

TIER	TYPICAL SCOPE	MINIMUM CONTROLS
One	Internal analysis and drafting, with no operational effect	Approved tool, training, source checking and human review before anything is issued
Two	Commercially confidential information, or advisory output within a planning workflow	Registered use case, data-flow review, evaluation, logging and a named reviewer
Three	Recommendations acted on by planners, supervisors or drivers, and anything monitoring a person	Constraint checking before display, freshness statements, a data protection impact assessment, worker consultation and a named operational owner
Four	Anything controlling equipment, setting a person's work rate, or contributing to a roadworthiness or safety decision	A senior accountable owner, independent verification, functional safety assessment, monitoring, tested fallback to manual and a rehearsed rollback

TABLE 4 Tier is set by consequence and reversibility. A system used at more than one tier is controlled at the highest.

5.3 Operational technology, and the boundary that matters most

A warehouse contains operational technology, meaning technology that interfaces with the physical world. Conveyors, sortation, automated storage and retrieval, and the control systems behind them belong to a different discipline from corporate information technology, and the national cyber authority is explicit that its priorities are different, because safety, reliability and availability carry physical consequences when they fail.³⁵ **STANDARD**

Its secure connectivity principles for operational technology, published in January 2026, set out eight principles covering the balance of risk and opportunity, limiting exposure, centralising and standardising connections, secure protocols, hardening the boundary, limiting the impact of compromise, logging and monitoring all connectivity, and establishing an isolation plan.³⁵ **STANDARD** In August 2026 the authority warned of increased targeting of operational technology across multiple sectors globally including the United Kingdom, and advised that organisations should not assume their operational technology is inaccessible from the internet without verifying it.³⁶ **REGULATOR**

That warning names neither transport nor logistics, so it should be read as a general one. A sector-specific warning does exist. In May 2025 a joint advisory co-sealed by agencies from 11 countries including the national cyber authority reported a Russian state campaign targeting logistics and technology organisations, including air traffic control, airports, maritime and port systems, assessed as beginning in February 2022.³⁷ **REGULATOR**

The design consequence for any operator adding intelligence to a warehouse is a hard separation between the layer that advises and the layer that actuates, with least privilege, brokered access and a defined fail-safe state. A recommendation engine that can reach a programmable logic controller has collapsed a boundary that took the industry 20 years to establish.

PRACTICE

Ask a supplier what happens to the conveyor when their cloud service is unreachable. The answer tells you whether you bought a tool or a dependency.

5.4 The six technical controls

Underneath the tiers sit six controls, consistent with the security guidelines co-sealed by agencies from 18 countries, the generative risk profile published by the American standards body, and data protection accountability expectations.^{1,38,39} **STANDARD**

1. Enterprise identity and site, customer or contract entitlements are enforced before retrieval and before any tool call.
2. Contracts define training use, retention, hosting region, subprocessors, deletion, incident notification and audit rights, and the absence of any one of them is treated as a finding.
3. Agents operate with least privilege and short-lived credentials, and any consequential action requires explicit approval and can be reversed.
4. Every production system records the model, prompt, retrieval sources, tools, permissions and evaluation version behind each output.
5. Prompt injection, hostile documents, sensitive-data disclosure and unsafe output handling are tested before scale, using material of the kind the system will meet.
6. Generated work passes the same review and records controls as work produced without it.

An operator wanting a certifiable management system has ISO/IEC 42001, published in December 2023.⁴⁹ **STANDARD** My own view is that certification is worth having only after the constraint library exists, because certifying a management system that governs nothing in particular is an expensive way to produce a certificate.



Resilient freight. The constraint library is the artefact that survives a change of supplier, a change of system and a change of management.



PART THREE

Proof, economics and delivery

CHAPTER SIX

Proving it, and the evidence gate

Run every operational pilot in shadow mode first. The recommendation is produced, recorded and compared, and nobody acts on it until the record says they should.

Shadow mode is the single most useful practice in this sector and it is rarely used. The system produces its recommendation. The operation carries on as it would have done, and both are recorded. After four weeks there is a dataset showing what the system would have done, what the operation actually did, and what happened. No customer was affected and the comparison is real. **ECAVEO**

THE SIX DIMENSIONS OF AN EVIDENCE GATE

DIMENSION	WHAT IS MEASURED	THE GATE
Outcome	Service level, cycle time, throughput, empty running or utilisation	Improves the intended measure against a real baseline, with the counter-measure unharmed
Constraints	Recommendations tested against the constraint library across normal and edge conditions	Zero breaches. A single breach fails the pilot. It does not trigger a tuning exercise
Judgement	Override rates, challenge quality, escalation and whether users can explain the recommendation	Users retain independent judgement and manual competence is maintained
Safety and people	Near misses, injuries, monitoring purpose limits, consultation and worker feedback	No adverse safety signal, and monitoring used only for the purpose consulted on
Security and resilience	Operational technology separation, permission leakage, red-team findings and fallback testing	No unresolved critical finding, and fallback to manual tested under load
Economics	Licence, integration, assurance, review, support and change cost	Total cost justified by verified benefit, including the cost of checking

TABLE 5 A pilot passes when every dimension clears its gate. The constraint dimension is specific to this sector and its gate is zero.

6.1 Three pilots worth running

An arrival and exception predictor, scoped to selected lanes, showing the operational factors behind each alert and making no change to any plan. Measured on precision, recall, warning time, the rate at which planners act, the resulting service change and the burden of false alarms.

A warehouse congestion forecaster, scoped to dock or aisle pressure from permissioned systems and appointment data, controlling no equipment. Measured on congestion minutes and queue time, on throughput, and on whether the supervisor trusts it by week four.

A maintenance evidence assistant, scoped to producing source-linked defect and service summaries for engineering review, deciding nothing. Measured on recall of known defects, the rate of unsupported claims, review time and audit traceability, with roadworthiness decisions remaining with the engineer throughout.

6.2 The minimum assurance record

Six things, for every production system. Purpose, intended users, benefit, owner and review date. Allowed and prohibited uses, data classes, sites and actions. Model, supplier, prompt, retrieval sources, connectors, permissions and dependencies. Risk assessments, the data protection impact assessment where people are monitored, contractual controls and any safety decisions taken. Evaluation sample, method, results, limitations and approval. Monitoring, feedback, incidents, changes, rollback and retirement.

6.3 Scale, redesign or stop

Scale where the benefit is achieved, the constraint gate is clean, residual risk has been accepted by a named owner and ownership is funded. Redesign where a value signal exists and quality, review burden, workflow fit or controls miss the gate. Stop where no verified benefit appears, the risk is unacceptable, the fit is weak, or a simpler method performs better.

In an operation the simpler method is a serious competitor and deserves to be tested honestly. A published appointment discipline, a changed cut-off time or a renegotiated delivery window will often beat an optimiser, cost nothing to run and require no assurance at all.

CHAPTER SEVEN

The economics of verification

This technology moves cost from producing a plan to checking one. In an operation the checking is done under time pressure by people who are already fully occupied.

A business case that counts planner hours saved and stops there is counting the wrong side of the ledger. Generating a plan becomes nearly free, and the binding constraint moves to verification, which in a transport office happens in the minutes before a vehicle leaves. **ECAVEO**

The test applies to any proposal in about a minute. Take the cost of producing the work by hand, the cost of generating it, and the cost of checking the generated version to the standard the work requires. Where generation plus verification comes in below production, the case holds.

WHY SOME USE CASES PAY AND OTHERS DO NOT

1 Machine checkable A plan tested automatically against the constraint library. The check costs milliseconds and a person reviews only what fails.	2 Person checkable An exception explanation with sources and freshness stamps. A supervisor opens two records and knows.	3 Unverifiable in time A recommendation with no stated basis, arriving at 16.50. There is nothing to check it against before the vehicle goes.
---	---	---

FIGURE 3 The constraint library is what moves a use case from the third column into the first. That is its economic function, quite apart from its safety function.

Ecaveo framework. **ECAVEO**

Read that way, the constraint library stops being a compliance artefact. It is the mechanism that makes optimisation affordable, because it converts human verification into machine verification for everything except the exceptions.

7.1 The cost lines that are missing

Three costs are routinely absent from a pilot case. Supervisory review time, described above. Assurance work, meaning evaluation, red-teaming and supplier diligence, the data protection impact assessment where people are monitored, and the records that support all of it. The cost of change, meaning the regression testing against the constraint library that follows every material model or supplier update, which recurs indefinitely.

My own rule of thumb, offered as practitioner judgement rather than measured fact, is that a tier three deployment carries ongoing assurance and review costs of a similar order to its licence cost. **PRACTICE**

7.2 The automation question, answered with the numbers that exist

Boards ask how far automation should go, and the honest answer is that the published evidence is thinner than the marketing implies. The best structured survey I could find, of more than 120 decision-makers drawn from a trade readership, reported the proportion of operations described as fully automated at 24 per cent for labelling, 13 per cent for packaging, 12 per cent for picking, 11 per cent for storage, 10 per cent for conveyance and 3 per cent for retrieval.⁴⁰ **VENDOR RESEARCH**

That last figure deserves a moment. Among a self-selected, automation-interested readership, the physically hardest part of the job is almost entirely manual. It is a weak survey. It also cuts against the commercial interest of the publication carrying it, which is worth something.

On cost and payback there is nothing to report. I could find no independent published data on what warehouse automation costs or when it pays back, and no independent United Kingdom data on adoption rates for automated storage and retrieval or goods-to-person systems. **EVIDENCE GAP** The figures in circulation are market research commissioned by suppliers or content recycled without attribution. A sector is being urged to invest at scale on an evidence base that does not exist publicly.

7.3 Driverless freight, stated at its actual size

The most advanced operator of driverless trucks reported approximately 440,000 cumulative driverless miles from launch to June 2026, with no collisions attributed to its system, running between Laredo and Dallas and between Phoenix and Fort Worth. It reported quarterly revenue of \$2m against a net loss of \$270m.²⁸ **VENDOR RESEARCH**

Set beside the 19.0 billion vehicle kilometres covered by UK heavy goods vehicles in a single year, that cumulative total is roughly four thousandths of one per cent of one year of British operation.¹⁶ **REGULATOR** That is 0.004 per cent of one year's British operation.

In Great Britain the position is further back still. The self-driving vehicle pilot scheme opened on 31 March 2026 and provides the first lawful route to operation without a safety driver, and it is framed around passenger services.²⁷ **STANDARD** Regulations permitting automated passenger service permits came into force in May 2026.²⁷ **STANDARD** The core authorisation framework of the Automated Vehicles Act 2024 is not in force and is expected from late 2027.²⁷ **STANDARD** There is no commercial driverless goods vehicle operation in this country and no regulatory route specifically opened for one.

CHAPTER EIGHT

Ninety days, then nine months

The first month is spent in depots and control rooms. A programme that starts with a vendor demonstration ends with a tool nobody is allowed to act on.

THE DELIVERY SEQUENCE

PHASE	TIMING	OUTPUTS
Discover	Weeks 1 to 4	Control rooms, depots and warehouses observed; constraints, data latency, systems, safety cases and baselines mapped; current use inventoried
Prove	Weeks 5 to 10	Three advisory pilots run in shadow mode before any controlled live use, with written stop conditions
Standardise	Weeks 11 to 18	Constraint library, model and use-case register, data contracts, operational technology access rules and role-based training
Scale	Months 5 to 9	Proven workflows extended by site and lane, with local owners, network measures and tested manual fallback

TABLE 6 The Standardise phase is where the constraint library, the data contracts and the operational technology access rules become the default.

8.1 The first 30 days

Observe real work on nights and at peak, because the system will be relied on hardest when the operation is under pressure and that is the condition worth understanding. Interview transport managers, shift managers, planners and drivers separately. Inventory what is already in use, including the capability embedded in the transport and warehouse management systems the operator already licences, which is invariably a longer list than the technology function expects.

Map the constraints before anything else, because that work is useful whatever happens next. Map data latency alongside it, recording for each source how old the data is when a decision is taken. Most operations have never written this down and are surprised by the answer.

Publish interim guidance on safe use within the first fortnight, along with a route for proposing an experiment. Select three pilots with baselines, shadow-mode periods and written stop conditions. Demonstrate one bounded workflow end to end so that the board sees capability and constraint checking in the same room.

PRACTICE

Where people are monitored, start the consultation in month one. The government consultation on workplace monitoring closes on 30 September 2026 and a statutory duty to consult is one of the options being tested.

8.2 What success looks like at nine months

Scaled workflows with measured improvement against the baselines set in month one, and with the named counter-measure unharmed. A constraint library owned internally, versioned and regression tested. A single intake route, approved systems, documented use cases and monitored supplier changes. Planners and supervisors who can explain a recommendation, challenge it and run the operation without it. A tested fallback to manual working, rehearsed and not merely written down. A backlog that separates proven opportunities from research and from uses the operator has decided against.

CHAPTER NINE

The session to run with the board

Eight questions, applied to any proposed use. A use case that survives all eight is worth funding.

1. What observable outcome should improve, for whom, and by when?
2. What is the current baseline, and which measure might this damage while improving that one?
3. What data does it use, how old is that data when the decision is taken, and what obligations attach to it?
4. What error would matter, who could be harmed, and how would anyone detect it?
5. Which hard limits apply, where are they held, and who owns them?
6. Who reviews the output, who decides, who can override, and who answers for the result?
7. What simpler method should this be compared against, and has anybody tried it?
8. What evidence would justify scaling this, and what evidence would stop it?

The fifth question is the one this paper adds to the series, and in most operations it has no answer at the start. The hard limits exist in a transport manager's head, in a supplier's configuration file and in a laminated notice by a loading door. Until they exist in one place that a person owns, there is nothing for a recommendation to be tested against. Expect the first attempt at writing them down to take longer than anyone budgets, and to surface at least one rule nobody can now justify.

TWO THINGS TO DECIDE IN THE ROOM

Name the person who owns the constraint library. Not the technology, not the supplier relationship, the statement of what this operation will not do. Then decide whether anything in the programme paces, ranks or evaluates a person, and if it does, start the consultation before the pilot.

CHAPTER TEN

Where the evidence is thin

Five gaps, and the first of them undermines most of what this sector is currently being sold.

10.1 Nobody has shown that this improves operations

As at September 2026 I could find no peer-reviewed field evaluation, government evaluation or regulator analysis establishing a causal operational benefit from artificial intelligence in United Kingdom road freight or warehousing. **EVIDENCE GAP** **PRACTICE** What exists is benchmarking on historical datasets, reviews of that benchmarking, and a large volume of vendor case studies without counterfactuals. The Federal Reserve's own monitoring of adoption across the American economy omits transportation and warehousing entirely.⁴¹ **REGULATOR**

That absence sits awkwardly beside the adoption claims, and it is the reason this paper argues for shadow mode and baselines. A single operator running a properly baselined comparison across 20 depots would produce better evidence than anything currently published. If such a study appeared and showed a clear operational effect, the risk appetite in Chapter 5 would be too conservative and I would say so.

10.2 Nobody has measured whether monitoring drives drivers away

Driver retention is the sector's central workforce problem and in-cab camera monitoring is among its most contested interventions. A systematic review of in-vehicle telematics for monitoring driving behaviour was published in a peer-reviewed journal in 2024 and is the right calibre of source.⁴² **PEER-REVIEWED** Beyond it, essentially all available material on effectiveness is published by the companies selling the systems, and I found no independent United Kingdom evidence on collision rates and none at all on retention. **EVIDENCE GAP** **PRACTICE**

An operator could answer the retention question for itself within a year, by comparing leaver rates and exit interview themes across depots with and without cameras. I am not aware of anyone publishing such a comparison.

10.3 The regulator has not yet arrived

I could find no enforcement action by the Information Commissioner concerning algorithmic management or monitoring of logistics, haulage or delivery workers, and no United Kingdom tribunal decision on the subject. **EVIDENCE GAP** European regulators have been considerably more active. An Italian authority fined a delivery platform in November 2024 over algorithmic management, including gamified shift allocation and dismissals without appeal, unlawful use of biometric data, and monitoring of workers' movements while the application ran in the background outside working hours.⁴³ **CASE LAW** A Dutch appeal court held in April 2023 that platform operators had to disclose the criteria behind automated decisions affecting drivers.⁴⁴ **CASE LAW**

A British operator reading those decisions should expect the reasoning to arrive here eventually, and should also note that one widely repeated claim about a large fine for algorithmic dismissal is wrong. The substantial Dutch penalty imposed in 2024 concerned unlawful transfers of driver data to the United States, which is a different matter entirely.⁴⁵ **REGULATOR**

10.4 The exposure figures are nine years old

The best official United Kingdom estimate of automation exposure was published in 2019 using 2017 data, put 7.4 per cent of jobs in England at high risk, and did not break out drivers, warehouse operatives or elementary storage occupations.⁴⁶ **REGULATOR** It predates the current generation of technology entirely and uses a methodology criticised for assuming whole occupations are automated when only tasks are.

For a sector employing 1.92 million people, the absence of a current official estimate is a gap in the evidence available to policymakers as much as to operators. **EVIDENCE GAP**

10.5 The one number I would most like

Whether better forecasting reduces empty running is testable and untested. Empty running has sat near 30 per cent of vehicle distance for years and is measured nationally every quarter.¹⁶ **REGULATOR** If collaborative load matching and forecasting work at the scale claimed for them, that figure should move, and it has not.

I hold open the possibility that the gains are real and are being absorbed elsewhere, in service levels, in fuel, in reduced overtime, none of which appear in that series. That is a defensible position and it is also exactly the kind of claim this paper asks operators to stop accepting without a baseline.



Service and stewardship. Everything in this paper exists to protect the moment at which a person decides that a vehicle may go.

References

Guidance and timetables in this area moved repeatedly during 2026 and remain subject to change. Where a source is a survey from a body with a commercial or membership interest in the result, the reference says so.

1. Information Commissioner's Office. *Guidance on AI and data protection*. Page last updated 15 March 2023 and flagged by the Commissioner as under review. <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/>
2. European Union. *Regulation (EU) 2024/1689 (Artificial Intelligence Act)*, Annex III points 2 and 4, as amended by *Regulation (EU) 2026/1744 (Digital Omnibus on AI)* of 8 July 2026, in force 27 July 2026, deferring Annex III high-risk obligations to 2 December 2027. <https://eur-lex.europa.eu/eli/reg/2026/1744/oj/eng>
3. European Union. *Regulation (EU) 2023/1230 on machinery*, applying from 20 January 2027, Annex I points 5 and 6. Great Britain continues to apply the *Supply of Machinery (Safety) Regulations 2008*, amended by SI 2026/867, made 22 July 2026 and coming into force 20 January 2027. <https://www.legislation.gov.uk/uksi/2026/867/made>
4. Information Commissioner's Office. *What do we need to do if we use monitoring tools that use solely automated processes?* Page dated 16 June 2026 and marked as under review following the Data (Use and Access) Act. <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/employment/monitoring-workers/what-do-we-need-to-do-if-we-use-monitoring-tools-that-use-solely-automated-processes/>
5. Driver and Vehicle Standards Agency. *Guide to maintaining roadworthiness: commercial goods and passenger carrying vehicles*. Published 22 April 2014, last updated 28 April 2025. <https://www.gov.uk/government/publications/guide-to-maintaining-roadworthiness>
6. Driver and Vehicle Standards Agency. *Electronic braking performance monitoring system (EBPMS): industry standard specification*. Published 12 May 2023, last updated 10 September 2024.
7. Health and Safety Executive. *Workplace transport, and HSG136: A guide to workplace transport safety*, which groups the duties into site, vehicle and driver. <https://www.hse.gov.uk/workplacetransport/about.htm>
8. Health and Safety Executive. *HSE's regulatory approach to artificial intelligence*. Policy statement, 12 June 2026. <https://press.hse.gov.uk/2026/06/12/hses-regulatory-approach-to-artificial-intelligence-ai/>
9. Health and Safety Executive. *Britain's health and safety regulator to create its first ever joint industry guidance for collaborative robotics in the workplace*. 10 June 2026. <https://press.hse.gov.uk/2026/06/10/britains-health-and-safety-regulator-to-create-its-first-ever-joint-industry-guidance-for-collaborative-robotics-in-the-workplace/>
10. Information Commissioner's Office. *The Data (Use and Access) Act 2025: what does it mean for organisations?* Page dated 19 June 2026. Automated decision-making provisions at section 80 commenced 5 February 2026. <https://www.legislation.gov.uk/ukpga/2025/18/section/80>
11. Department for Business and Trade. *Make Work Pay: workplace monitoring technologies*. Consultation opened 8 July 2026, closing 30 September 2026. <https://www.gov.uk/government/consultations/make-work-pay-workplace-monitoring-technologies>
12. Health and Safety Executive. *Work-related fatal injuries in Great Britain, 2025/26*, provisional, published 30 June 2026, and *Non-fatal injuries at work in Great Britain, 2024/25*. <https://www.hse.gov.uk/statistics/>
13. Office for National Statistics. *Artificial intelligence in UK businesses: 2023 to 2026*. Published 20 July 2026. Business Insights and Conditions Survey, 38,637 responding businesses, fieldwork 15 to 28 June 2026. Self-reported use, with artificial intelligence undefined for respondents. <https://www.ons.gov.uk/businessindustryandtrade/business/businessservices/articles/artificialintelligenceinukbusinesses/2023to2026>
14. Descartes Systems Group. Survey of 300 senior decision-makers across Germany, the United Kingdom, France and the Benelux countries, 75 in each, with fieldwork described as conducted during 2025 and reported in 2026. Vendor-commissioned survey; the headline is reached by taking the union of loosely defined use categories.
15. Gartner. *Survey finds majority of chief supply chain officers unclear on AI investment returns*. 5 August 2026. 394 supply chain professionals, fieldwork November 2025 to February 2026, non-probability sample. <https://www.gartner.com/en/newsroom/press-releases/2026-08-05-gartner-survey-finds-majority-of-chief-supply-chain-officers-unclear-on-ai-investment-returns>
16. Department for Transport. *Domestic road freight statistics, United Kingdom, April 2025 to March 2026*. Published 5 August 2026. Tables RFS0101, RFS0117 and RFS0301. <https://www.gov.uk/government/statistics/domestic-road-freight-statistics-april-2025-to-march-2026>
17. McKinsey & Company. *The State of AI*. Published 25 August 2026. 1,719 respondents across 97 nations, fieldwork 4 May to 8 June 2026, non-probability panel, not logistics-specific. <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>
18. Makridakis, S., Spiliotis, E., and Assimakopoulos, V. *The M5 accuracy competition: results, findings and conclusions*. International Journal of Forecasting, 2022. 42,840 series, 5,507 teams, blind out-of-sample evaluation.

19. Traffic Commissioners for Great Britain. *Annual Report 2024-25*. <https://www.gov.uk/government/publications/traffic-commissioners-annual-report-2024-to-2025>
20. Office for National Statistics. *UK workforce jobs*, series JWS4, transport and storage, first quarter 2026. <https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/timeseries/jws4/lms>
21. Logistics UK. *The Logistics Report 2026*. 3 June 2026. Trade association publication; the derivation of the gross value added figure is not disclosed.
22. Full Fact. *Is there a shortage of 100,000 HGV drivers?* 28 September 2021. <https://fullfact.org/economy/fuel-shortage-hgv-drivers/>
23. UK Government. *Drivers' hours: assimilated rules*. Last updated 9 September 2026. <https://www.gov.uk/drivers-hours/eu-rules>
24. Driver and Vehicle Standards Agency. *Approved Tachograph Centre Special Notice 02-26*, 26 May 2026, and *Drivers' hours: goods vehicles, tachograph rules*, last updated 1 July 2026.
25. Driver and Vehicle Standards Agency. *Haulier faces licence revocation after director admitted tachograph fraud*. 23 July 2026. <https://www.gov.uk/government/news/haulier-faces-licence-revocation-after-director-admitted-tachograph-fraud>
26. Senior Traffic Commissioner. *Statutory Document No. 3: Transport Managers*, version 17.0, commencing March 2025. <https://www.gov.uk/government/publications/traffic-commissioners-transport-managers-november-2018>
27. Centre for Connected and Autonomous Vehicles and Department for Transport. *Self-driving vehicle pilot scheme: information for applicants*, published 31 March 2026; *The Automated Vehicles (Permits for Automated Passenger Services) Regulations 2026*, SI 2026/439, in force 15 May 2026; *Automated Vehicles Act 2024*, core provisions not yet commenced.
28. Aurora Innovation. *Second quarter 2026 results*, 29 July 2026. Company reporting on its own operations, with the mileage and safety record unverified independently. <https://ir.aurora.tech/news-events/press-releases>
29. Information Commissioner's Office. *Enforcement notices issued to Serco Leisure Operating Limited and associated trusts*, 23 February 2024. <https://ico.org.uk/action-weve-taken/enforcement/2024/02/serco-leisure-operating-limited-and-relevant-associated-trusts/>
30. United States Senate Committee on Health, Education, Labor and Pensions. *Majority staff report on warehouse injuries and productivity quotas, December 2024*. A legislative investigation with document access, conducted under a chair with a declared adversarial position; the company disputes the findings. https://www.help.senate.gov/imo/media/doc/amazon_investigation.pdf
31. United States Department of Labor, Occupational Safety and Health Administration. *Citations issued at three warehouse facilities*, 18 January 2023. *Ergonomic and struck-by hazards under the general duty clause; the action did not concern robotics*. <https://www.dol.gov/newsroom/releases/oshsha20230118>
32. Cyber Monitoring Centre. *Statement on ransomware incidents in the retail sector*. 20 June 2025. Category two systemic event, £270m to £440m, modelled from transaction-level spending data and public disclosures. <https://cybermonitoringcentre.com/2025/06/20/cyber-monitoring-centre-statement-on-ransomware-incidents-in-the-retail-sector-june-2025/>
33. Reported cause of the February 2019 fire at an automated grocery fulfilment centre, attributed jointly by the operator, its insurer and the fire service to an electrical fault in a first-generation battery charging unit. Reported damage of approximately £110m.
34. Reported cause of the July 2021 fire at a second automated grocery fulfilment centre, attributed by the operator to the collision of three robots on the grid.
35. National Cyber Security Centre. *Operational Technology* collection, published 18 March 2024, and *Secure connectivity principles for operational technology*, 14 January 2026. <https://www.ncsc.gov.uk/collection/operational-technology>
36. National Cyber Security Centre. *Disruptive cyber activity highlights risk from internet-exposed systems and edge devices*. 27 August 2026. The advisory does not name transport or logistics specifically. <https://www.ncsc.gov.uk/news/disruptive-cyber-activity-highlights-risk-from-internet-exposed-systems-and-edge-devices>
37. Joint cyber security advisory AA25-141A on Russian state targeting of logistics and technology organisations, 21 May 2025, co-sealed by agencies from 11 countries including the National Cyber Security Centre. <https://www.cisa.gov/news-events/cybersecurity-advisories/aa25-141a>
38. National Cyber Security Centre. *Guidelines for secure AI system development*. 27 November 2023, co-sealed by agencies from 18 countries. <https://www.ncsc.gov.uk/collection/guidelines-secure-ai-system-development>
39. National Institute of Standards and Technology. *Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile*, NIST AI 600-1.26 July 2024. <https://doi.org/10.6028/NIST.AI.600-1>
40. Peerless Research Group for *Modern Materials Handling*. *2026 Automation Study*. More than 120 respondents drawn from a trade magazine readership; self-selecting, small, and predominantly United States. Field dates not stated.
41. Allen, J. S. *Monitoring AI adoption in the U.S. economy*. Board of Governors of the Federal Reserve System, FEDS Notes, 3 April 2026. Transportation and warehousing is absent from the industries covered. <https://www.federalreserve.gov/econres/notes/feds-notes/monitoring-ai-adoption-in-the-u-s-economy-20260403.html>
42. Boylan, J., Meyer, D., and Chen, W. S. *A systematic review of the use of in-vehicle telematics in monitoring driving behaviours*. *Accident Analysis and Prevention*, volume 199, 2024, article 107519. <https://doi.org/10.1016/j.aap.2024.107519>
43. Garante per la protezione dei dati personali. *Decision against a food delivery platform*, 13 November 2024. Breaches found across multiple provisions, including algorithmic ranking and account

blocking without adequate appeal, unlawful use of biometric data and background-mode location monitoring.

44. Amsterdam Court of Appeal. Judgments concerning platform drivers and automated decision-making, 4 April 2023.
45. Autoriteit Persoonsgegevens. Fine of €290m imposed August 2024 for unlawful transfers of driver data to the United States. A data transfer case rather than an algorithmic management case. <https://www.autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-imposes-a-fine-of-290-milli-on-euro-on-uber-because-of-transfers-of-drivers-data-to-the-us>
46. Office for National Statistics. *Which occupations are at highest risk of being automated?* 25 March 2019, analysing 2017 data. Uses a task-based methodology criticised for overstating displacement, and does not break out logistics occupations. <https://www.ons.gov.uk/employmentandlabour>
- market/peopleinwork/employmentandemployeetypes/articles/whichoccupationsareathighestriskofbeingautomated/2019-03-25
47. Department for Transport. *Future of freight plan*. 15 June 2022. Five priority areas including data and technology, and a £7m Freight Innovation Fund. A successor multimodal freight plan was announced in April 2025 and had not been published as at September 2026. <https://www.gov.uk/government/publications/future-of-freight-plan>
48. GS1. *GS1 Logistic Label Guideline*, release 1.3, ratified July 2019. The Serial Shipping Container Code is the single mandatory element of a logistic label. https://www.gs1.org/docs/tl/GS1_Logistic_Label_Guideline.pdf
49. International Organization for Standardization and International Electrotechnical Commission. *ISO/IEC 42001:2023, Artificial intelligence management system*. December 2023.

ECAVEO RESPONSIBLE AI SERIES

The Protected Constraint

Operations strand. Version 1.0, September 2026.

ALSO IN THIS SERIES

The professional firms trilogy runs alongside this paper. The Traceable Thesis covers private equity and Sufficient and Appropriate covers accounting and audit.

Verified Authority covers law firms. Each paper stands alone and all four share one framework.

THE AUTHOR

Paul Forrest is a fractional Head of AI working with private equity and venture capital portfolio businesses on the deployment of artificial intelligence.