

PAPER ONE OF THREE · PRIVATE EQUITY

The Traceable Thesis

Artificial intelligence across the investment lifecycle, and the controls that keep judgment, confidence and capability claims where they belong

Paul Forrest

Fractional Head of AI, Ecaveo
July 2026

About this paper

This is the first of three papers on deploying artificial intelligence inside professional firms that carry personal accountability for their work. The second concerns accounting and audit firms, the third law firms. Each paper stands alone. Together they argue one thing, which is that the durable advantage in professional work comes from decisions that can be reconstructed afterwards, and that systems which produce answers without producing a record destroy more value than they create.

Private equity sits first in the series because its exposure is the widest. A general partner owns an investment decision, a set of portfolio interventions, a valuation and a series of statements to investors about how all of that was arrived at. Each of those is a place where an unverified machine output can do real damage.

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. Where it points against the industry's preferred story, the paper says that too.

HOW TO READ THE EVIDENCE TAGS

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

REGULATOR A regulator, central bank, supervisory body or international financial institution speaking in its own voice.

STANDARD A published standard, framework or professional requirement.

PEER-REVIEWED A primary study, working paper or systematic review.

SURVEY A survey or published dataset with disclosed method and sample.

VENDOR RESEARCH Research or a survey from a firm with a commercial 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 and regulatory notice. Parts of this paper describe legislation, regulatory guidance and supervisory expectations, in particular in Chapter 2. Nothing in this document constitutes legal advice. 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. Regulatory timetables in this area have moved more than once.

Contents

	A note from the author	4
	Executive summary	5
ONE	What has actually changed	7
TWO	What the supervisors say, and what they leave open	9
THREE	Where the value sits across the lifecycle	14
FOUR	Six ways this fails in a fund	17
FIVE	The control model	20
SIX	Proving it, and the evidence gate	24
SEVEN	The economics of verification	26
EIGHT	Ninety days, then nine months	27
NINE	The session to run with the partnership	28
TEN	Where the evidence is thin	29
	References	31

THE ARGUMENT IN ONE PARAGRAPH

Artificial intelligence compresses the information work that surrounds an investment decision, and it does so most reliably in origination, diligence and portfolio monitoring. The compression is worth having. It becomes dangerous at the point where a general partner allows a synthesis to stand in place of the evidence behind it, because the fund then owns a conclusion nobody can reconstruct. The recommendation in this paper is to buy compression and pay for traceability, and to treat any workflow that cannot produce a source-linked record as unfinished, whatever its speed.

ILLUSTRATIVE MATERIAL

Every firm, fund, portfolio company and person in the worked examples is invented. Each is marked as an illustrative scenario where it appears, and none describes a real manager or holding. Published research is cited and numbered. Invented material is never presented as evidence.

CITATION

Forrest, P. (2026). *The Traceable Thesis: Artificial Intelligence Across the Investment Lifecycle*. Ecaveo Responsible AI Series, Paper One. Version 1.0, July 2026.

ALSO IN THIS SERIES

Paper two. *Sufficient and Appropriate*. Artificial intelligence, audit evidence and professional scepticism in accounting firms.

Paper three. *Verified Authority*. Artificial intelligence, privilege and the duty to the court in law firms.

A NOTE FROM THE AUTHOR

The memory of a fund

Private equity is unusually good at recording what it decided and unusually poor at recording why. That asymmetry is the subject of this paper.

I have spent my working life on how organisations turn what they know into what they do, first through business process work in the 1990s and since 2014 through machine learning and language models. I build small, specific models in preference to large general ones, because the problems worth solving in a fund are narrow and the cost of a confident wrong answer is high.

What prompted this paper was a pattern I kept meeting in investment committees. The papers were excellent. The debate was genuine. Six months later, when a portfolio company missed its plan, nobody could say with confidence which of the diligence findings had actually carried the decision, which had been noted and discounted, or who had held the contrary view. The investment memo survived. The reasoning did not.

Artificial intelligence makes that problem worse before it makes it better. A well-built retrieval system will produce a fluent, plausible synthesis of a data room in minutes, and a fluent synthesis is exactly the kind of artefact that displaces the evidence it was drawn from. The firms that will do well out of this are the ones that treat the synthesis as a draft and the source trail as the asset.

This paper argues for a specific and slightly unfashionable position, which is that the return on artificial intelligence in a fund comes from traceability. Speed without it is a liability the partnership has not yet priced. I have graded the evidence throughout. Where the research contradicts what the industry likes to say, I have quoted the research.

TWO PAGES THAT CIRCULATE ON THEIR OWN

Executive summary

The near-term opportunity is real, narrow and unevenly evidenced. The near-term risk is concentrated in confidentiality, in claims made to investors, and in the quiet substitution of synthesis for evidence.

A general partner deploying artificial intelligence is making three separate bets, and they fail in different ways. The first is that the technology will compress information work, which the available evidence supports. The second is that compressed information work will improve investment outcomes, which the available evidence does not yet show. The third is that the firm can say what it is doing without misdescribing it, and that is where enforcement has already landed.

1

Adoption in UK financial services is already broad, and the supervisors know it. The third Bank of England and Financial Conduct Authority survey of artificial intelligence in UK financial services, published in November 2024, reported that 75 per cent of the 118 responding firms were already using it, with a further 10 per cent planning to within three years.¹² **SURVEY** The question facing a partnership is no longer whether to start.

2

The regulatory perimeter is deliberately technology-neutral, which transfers the work to the firm. The FCA states that it does not plan to introduce extra regulations for artificial intelligence and will rely on existing frameworks.⁵ **REGULATOR** That is a lighter rulebook and a heavier evidential burden, because the firm must show that an existing outcome was met. There is no box to tick.

3

Misdescribing capability has produced enforcement; misusing capability has not yet. The SEC has charged advisers and issuers over false artificial intelligence claims in a run of matters from March 2024 onward, with penalties of \$225,000 and \$175,000 in the first pair alone.⁸ **REGULATOR** I could find no enforcement action anywhere against a private fund adviser for the underlying use of the technology. **EVIDENCE GAP**

4

The performance evidence is associational and comes mostly from firms that sell advice. The strongest dataset I located covers 471 private-equity-backed companies and reports that those at the highest artificial intelligence maturity carried a median revenue multiple of roughly 31 times against 13 times at the lowest.¹⁵ **VENDOR RESEARCH** The study does not isolate cause from company quality or sector, and a reader should treat it as an association. Calling it a return goes further than the authors do.

5

General partners themselves are more sceptical than the marketing suggests.

In a survey of 103 investment and investor-relations professionals conducted over the turn of 2025 and 2026, 39 per cent expected no material financial impact on portfolio companies during 2026, and respondents saw the clearest return in diligence and sourcing. Value creation came some way down the list.¹⁴ **SURVEY**

6

The industry has an exposure to this technology that has nothing to do with using it. The IMF reported in April 2026 that roughly one-fifth of all direct-lending loans sit in the software sector, and that leveraged private credit vehicles may carry software-borrower exposure above half of net asset value.¹⁸ **REGULATOR**

A fund can decline every tool on the market and still be long the same theme.

What follows for a general partner

The recommendation is less exciting than the sales cycle and more useful. A firm should join up its own records before buying any intelligence, because a system that reasons across four incompatible data rooms will reason confidently about the wrong company. It should insist that every generated claim carries a link to the material it came from, states what it could not find, and offers the contrary reading where one exists. It should place a named individual between any output and any decision, and record who that was.

Above all it should record decisions. A decision record costs almost nothing at the moment of the decision and cannot be reconstructed afterwards at any price. It is the only asset in this field that appreciates, and it is the one that a change of deal team currently destroys.

Chapter 8 sets out a 90-day proof on three bounded workflows, with baselines and stop conditions. Chapter 10 describes what would have to be true for me to change the position taken here.

CHAPTER ONE

What has actually changed

Funds have had analytical technology for 20 years. What arrived recently is the ability to read unstructured material at scale, and that is a narrower and more useful change than the market describes.

For most of the past two decades the analytical advantage in private markets came from structured data, and it was competed away in the ordinary manner. Better company databases, and better market data behind them. Every serious manager now buys comparable structured data from a small number of comparable suppliers, which means whatever advantage it once conferred has largely gone.

The change that matters is the arrival of systems that read unstructured material with tolerable reliability. A data room is unstructured. So is a management presentation, and so are four years of board packs, a customer contract set and an investment committee's own history of theses and post-mortems. Until recently none of it could be queried. The work of reading it was done by junior professionals under time pressure, and the coverage achieved depended on how many of them were available that fortnight.

Compression is the honest word for what is on offer. A retrieval system does not know which customer concentration matters. It makes it possible to ask, across 3,000 documents in an afternoon, a question that nobody would previously have bothered to ask, because asking it would have cost a week. Broader coverage changes what a diligence team can consider, though whether considering more produces better decisions is a separate question and Chapter 10 admits I cannot answer it.

THE CORE CLAIM

The advantage has moved from having the data to being able to reconstruct why a decision was taken. That property cannot be bought from a vendor.

1.1 The cycle this series uses

Every paper in this series describes the same five-stage cycle, because it is the sequence a professional decision actually follows whether or not anybody writes it down. Information is gathered, converted into evidence, weighed as a judgement, delivered as a decision, and compared afterwards against what happened. Private equity performs the first and fourth stages well. The second, third and fifth are where institutional knowledge either accumulates or evaporates.

THE CYCLE EVERY PAPER IN THIS SERIES EXAMINES

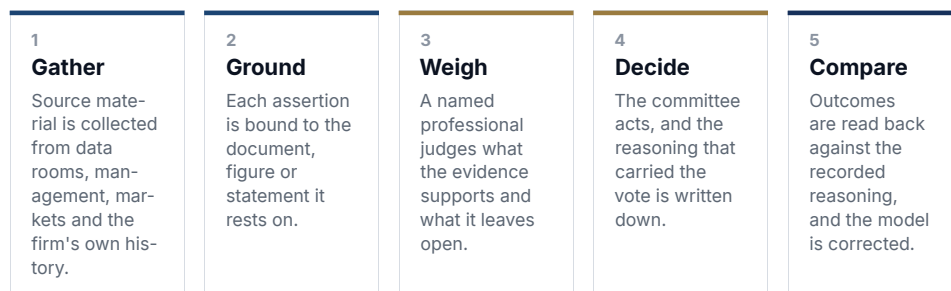


FIGURE 1 Technology has improved stage one for 20 years. Stages two, three and five are where a fund's institutional memory is made or lost.

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

1.2 Adoption is broad, and shallower than it looks

Two figures should be read together. The Bank of England and the FCA reported in November 2024 that 75 per cent of 118 responding financial services firms were already using artificial intelligence, with a further 10 per cent intending to start within three years.¹² **SURVEY** Against that, a survey of 200 fund and operating leaders published in May 2026 found that only 36 per cent of portfolio companies used it across more than one application, and that 7 per cent operated at what the authors described as enterprise scale.¹⁶ **VENDOR RESEARCH**

The second survey also reported that 95 per cent of funds considered their initiatives to be meeting or exceeding the original business case. That figure and the 7 per cent figure appear in the same document, and the tension between them is instructive. Success measured against a firm's own initial expectation is a statement about the expectation. My reading is that the 7 per cent is the more informative number, and that a reader should treat self-assessed success rates in this market as sentiment.

There is a related discipline worth applying to every statistic in this field. Most of the adoption figures in private markets are produced by advisory firms and vendors with a commercial interest in the answer, and very few publish a sampling frame or a response rate. That does not make them worthless. It makes them directional, and a partnership setting a budget against them should know which kind of number it is holding.

WORTH HIGHLIGHTING

A widely circulated claim that 95 per cent of enterprise artificial intelligence pilots fail has thin methodological disclosure and contested authorship. I have left it out.

1.3 The uncomfortable finding

The evidence that artificial intelligence improves investment outcomes in private equity is weak, and saying so plainly is more useful than the alternative. The most substantial quantitative work I located segments 471 private-equity-backed companies across 31 industries by artificial intelligence maturity and reports large differences in revenue multiples, with the highest maturity group at roughly 31 times against roughly 13 times for the lowest. The same study reports median revenue per employee of about \$180,000 at the highest maturity level against about \$118,000 at the level below it, and it publishes no dollar figure for the lowest.¹⁵ **VENDOR RESEARCH**

Read carefully, that is an association between two things that both correlate with company quality. Better-run companies adopt new operating technology earlier, and better-run companies command higher multiples, so the finding is consistent with artificial intelligence causing value, with value causing adoption, and with a third factor causing both. The authors do not claim to have separated them, and the research is published by a firm that sells transformation advice.

No independent study has yet shown that artificial intelligence improves private equity returns. The absence is the most defensible thing a general partner can say about it.

I have looked for a counterweight from a source without a commercial position, meaning an academic study, a central bank analysis or a regulator-commissioned review that isolates a causal effect in private markets. I did not find one. **EVIDENCE GAP** That absence should shape how a partnership talks to its investors, because a claim of improved returns is currently a claim without a published basis, and Chapter 2 describes what happens to firms that make claims of that kind.

CHAPTER TWO

What the supervisors say, and what they leave open

The United Kingdom regulates outcomes and leaves the technology itself alone. That is a genuine easing of the rulebook and a genuine tightening of what a firm must be able to show.

A reminder on register. This chapter summarises what published sources say and grades how firmly they say it. It does not state the legal position of any firm. Where I draw an implication a source does not itself state, the sentence says so.

2.1 The FCA position

The FCA states that it does not plan to introduce extra regulations for artificial intelligence, and that it will instead rely on existing frameworks which it considers already mitigate many of the associated risks.⁵ **REGULATOR** The page first appeared in September 2025 and carries an editorial amendment dated February 2026, so it is a standing position. Nothing changed direction in February. The frameworks it points to include the Consumer Duty and the Senior Managers and Certification Regime, which between them locate responsibility in named individuals.

For a private fund manager the practical consequence sits in the second of those. Where a senior manager holds a prescribed responsibility, the use of a system inside that area does not move the responsibility. My own view, which goes beyond what the FCA page states, is that a partnership should expect any future supervisory interest to begin with the question of which senior manager understood the tool well enough to own its output, and that a firm which cannot answer that question quickly will find the conversation unpleasant.

The FCA has also built a testing route. Its AI Live Testing service, described in a feedback statement following 67 responses to the original engagement paper, ran a first cohort with applications closing in September 2025 and a second between January and March 2026, with findings expected in the first quarter of 2027.⁶ **REGULATOR** A manager with a genuinely novel deployment has a supervised route available, which is a different proposition from waiting for guidance.

2.2 The AIFM regime is being rewritten around the firm, and it is silent on this

In July 2026 the FCA consulted on a new UK regime for alternative investment fund managers, proposing a three-tier structure of small, medium and large firms in place of the current thresholds, and a consolidated sourcebook to replace the rules assimilated from the European directive.⁷ **REGULATOR** A second consultation carrying draft rules is planned, with a policy statement to follow, and the reported target for implementation is 2028. As at the date of this paper no policy statement has been published, so the regime should be described as proposed.

What the consultation does not do is address artificial intelligence, and that silence is the most important feature of the regulatory position for this audience. The European guidance that does exist, published by ESMA in May 2024, is framed around MiFID II retail investment services and its organisational and conduct requirements.⁹ **REGULATOR** It does not address alternative investment fund managers. A firm looking for a rule written for its own activity will not find one, and will have to reason from general systems and controls obligations instead.

EVIDENCE GAP

Neither the FCA nor ESMA has published guidance addressing artificial intelligence in the hands of an alternative investment fund manager, as distinct from a retail-facing investment firm.

2.3 Where enforcement has actually landed

The pattern in enforcement is consistent and worth stating precisely, because it is frequently misdescribed. Regulators have punished false statements about artificial intelligence. They have not yet punished the underlying use of it.

In March 2024 the SEC charged two investment advisers with making false and misleading statements about their use of the technology, one having claimed it could predict which companies were about to succeed and the other having described itself as the first regulated artificial intelligence financial adviser, with civil penalties of \$225,000 and \$175,000.⁸ **REGULATOR** Further matters followed. An adviser charged in October 2024 over claims of automated artificial intelligence trading had raised around \$4m from 45 investors, and the principal was ordered to pay approximately \$463,611 alongside an industry bar.¹⁰ **REGULATOR** In April 2025 the founder of a shopping application that claimed artificial intelligence checkout faced parallel civil and criminal charges after raising over \$42m, on the allegation that the work was substantially done by people.¹¹ **REGULATOR**

Two features of that run deserve attention. Every matter concerns a description of capability. Not one turns on a failure of it, and none involves a private fund adviser. **EVIDENCE GAP** **PRACTICE** I would not read the second point as comfort. The examination priorities published for 2026 state that the Division will review the accuracy of representations about artificial intelligence capabilities, assess whether firms have adequate policies to supervise their use, examine the training and security controls applied against artificial intelligence and polymorphic malware threats, and test whether operations and controls are consistent with the disclosures made to investors.¹³ **REGULATOR**

THE SUPERVISORY POSITION AT A GLANCE

SOURCE	WHAT IT SAYS	FOR A UK MANAGER
FCA approach to AI	No separate AI rules are planned, and existing outcome-focused frameworks apply ⁵	Binding indirectly, through Consumer Duty, SM&CR and systems and controls
FCA CP26/28	Proposes a three-tier AIFM regime and a consolidated sourcebook ⁷	Consultation stage only, and silent on artificial intelligence
ESMA guidance, May 2024	Organisational and conduct expectations for AI in investment services ⁹	Scoped to MiFID II retail services, so it does not reach an AIFM
EU AI Act, Annex III	High-risk categories in finance cover creditworthiness of natural persons and life and health insurance pricing ¹⁹	Investment decisioning is outside the high-risk list. A portfolio company may be inside it on its own facts
PRA SS1/23	Principles-based model risk management, applied to AI and machine learning ²⁰	Reaches PRA-regulated firms. An FCA-only manager has no equivalent model risk standard
SEC enforcement, 2024 to 2025	Penalties for false or exaggerated AI claims ^{8,10,11}	Directly relevant to any manager marketing to US investors

TABLE 1 Read down the right-hand column. The obligations that bind a UK fund manager today are general ones, and the specific instruments either do not apply or do not yet exist.

2.4 Three gaps a partnership should know about

The first gap concerns inside information. The FCA has worked on the use of these tools to detect market abuse, through a surveillance technology sprint among other things. What neither it nor ESMA has published is guidance on how such systems interact with the market abuse regime as a source of risk, or with information barriers, which leaves firms applying general principles to a class of system that can surface material across boundaries in ways a person would not. **EVIDENCE GAP** My own view is that this is the single most likely area for a first supervisory intervention in private markets, because the failure mode is easy to describe and hard to defend.

The second concerns model risk. The Prudential Regulation Authority's supervisory statement on model risk management applies principles-based expectations to models including machine learning, and the Authority convened chief risk officers from 21 regulated firms in October 2025 to discuss applying it to this technology.²⁰ **REGULATOR** A fund manager regulated only by the FCA sits outside that statement. Nothing equivalent has been published for it, so the standard a manager applies to a valuation model touched by a generated assumption is currently a standard it sets itself.

The third concerns investors. The leading limited partner association has published a responsible artificial intelligence guide for asset owners, which sits inside its environmental, social and governance series and carries nine questions an asset owner might put to a manager.²⁵ **STANDARD** Those questions are framed around artificial intelligence as a portfolio governance matter. They do not reach a manager's own use of it in diligence, monitoring or valuation, and I could find no standard diligence content that does. **EVIDENCE GAP** UK Private Capital, which adopted that name in January 2026 having previously been the British Private Equity and Venture Capital Association, offers training on the subject and has published no position paper.

Limited partners are therefore asking about a manager's own use of these tools in ad hoc terms. A manager that publishes a clear, evidenced position ahead of the question will be answering on its own terms.

2.5 The obligations that do bind, today

Underneath the absence of sector-specific rules sit obligations that apply now. Data protection law reaches any processing of personal information inside a diligence exercise, and the Information Commissioner's guidance on artificial intelligence and data protection remains the principal reference, although the page carries a 2023 update and the Commissioner has flagged it as under review.¹ **REGULATOR** The Data (Use and Access) Act 2025 brought revised provisions on automated decision-making into force in February 2026.²¹ **STANDARD**

On security, the National Cyber Security Centre's guidelines for secure artificial intelligence system development, co-sealed by agencies from 18 countries, set expectations across design, development, deployment and operation.² **STANDARD** The American standards body's generative artificial intelligence profile sets out 12 risk categories and suggested actions, and remains the most usable checklist I know of for a technical review.³ **STANDARD** Firms wanting a certifiable management system have ISO/IEC 42001, published in December 2023.²² **STANDARD** For the systemic picture, the Financial Stability Board's assessment of the financial stability implications of this technology, summarised for supervisors by the Bank for International Settlements, is the most useful single document.²³ **REGULATOR** **PRACTICE** The UK government's own playbook sets out ten principles which, although written for departments, translate cleanly to a partnership.²⁴ **STANDARD**

None of that is specific to private equity, and all of it is enforceable against a private equity firm. The Commissioner's own guidance carries a notice saying it is under review, which is an awkward position for everyone and not one a firm can wait out.



Origination and market sensing. Broader coverage of unstructured material changes what a team can consider, and leaves the consideration where it was.

A photograph of three business professionals (two men and one woman) sitting around a dark, polished table in a modern office setting. They are all looking at laptops, suggesting a collaborative meeting. The man on the left is resting his chin on his hand. The woman in the center is looking down at her laptop. The man on the right is looking at his laptop. The table is cluttered with papers, pens, and a cup of coffee. A large window in the background shows a cityscape with a prominent dome. A brass lamp is visible behind the woman. The overall mood is professional and focused.

PART TWO

Where the value sits, and how it fails

CHAPTER THREE

Where the value sits across the lifecycle

Six workflows carry most of the available return. They are ranked here by how confidently the value can be demonstrated.

A portfolio of use cases is a set of hypotheses and should be treated as one. Priority should follow workflow observation, data classification, investor expectations and the cost of an error. The ranking below reflects the second of those tests most heavily, because a workflow whose output can be checked against a known answer can be proved, and one whose output cannot be checked can only be believed.

THE OPPORTUNITY PORTFOLIO

WORKFLOW	WHAT THE TECHNOLOGY DOES WELL	THE VALUE TEST
Diligence workbench	Extracts obligations, customer themes, pricing evidence and open questions from a data room, with links back to source	Wider review and faster issue identification, with every finding traceable to a document
Portfolio monitoring	Explains variance against plan, surfaces operational anomalies and drafts the questions a chair should ask	Time to a validated explanation, with finance owners retaining the numbers and the narrative
Firm knowledge	Retrieves prior theses, sector work, playbooks and post-mortems subject to permissions	Search time and answer support, measured against permission leakage
Origination intelligence	Maps themes, companies, executives and prior interactions across approved sources	Coverage and continuity, with no automated recommendation attached
Committee support	Builds source-linked chronologies, scenario summaries and challenge questions	More time for debate, and no automated recommendation or vote
Value creation	Identifies repeatable commercial, procurement and service patterns across holdings	Portfolio-wide pattern recognition, implemented company by company

TABLE 2 Each row states the value test. A benefit that nobody measured against a baseline will not survive its first budget review.

3.1 Diligence is where the case is strongest

Diligence has the property that makes evaluation possible, which is a known answer. A completed transaction has a set of issues that were found, a set that were missed and later surfaced, and a documentary record of both. A firm can therefore take three closed deals, rebuild the data rooms, and measure what a system finds against what the team found, how long each took, and what each missed.

That is an unusually clean test and very few firms run it. The survey evidence supports the intuition. In the outlook study covering 103 investment and investor-relations professionals, respondents identified diligence and deal sourcing as the areas where they saw the highest return.¹⁴ **SURVEY** My own experience is consistent with that, with the qualification that the measured gain concentrates in coverage. Teams do not finish diligence earlier. They ask more of it in the same time, which is a good outcome if the extra questions were worth asking and an expensive one if they were not, and I have seen both.

ILLUSTRATIVE SCENARIO

The data room that answered the wrong question

An invented mid-market manager runs a retrieval system across a 4,100-document data room for a business services acquisition. The system reports no material change-of-control provisions in the top 20 customer contracts.

The finding is correct and the conclusion drawn from it is wrong. Three of the top 20 relationships operate under framework agreements held at parent level, which were not in the data room, and a fourth had been novated in a manner the uploaded document set did not reflect. The system reported accurately on what it was given, and nobody established what it had not been given.

The control that would have caught this costs nothing. A system that states its own coverage, saying which contracts it read and which named counterparties it found no document for, converts a confident answer into a useful one.

3.2 Monitoring is where the value is most often left on the table

A portfolio company's monthly pack arrives late, in a format that changed last quarter, with a commentary written by the person whose performance it describes. An operating partner covering seven holdings reads them in sequence and forms a view. The failure mode is not inattention. It is that variance explanations are compared against nothing, because the previous explanation is in a different document that nobody opens.

Retrieval across a portfolio's own reporting history changes that. A system can compare this month's explanation of a gross margin movement against the six previous explanations, and note that input costs have now been blamed four times while the volume mix has moved the other way. The output is a question. Its value lies in arriving before the board meeting.

Finance owners should keep the numbers and the narrative. My own position, which goes further than most vendors would like, is that a monitoring system should be prohibited from restating a financial figure and permitted only to point at one. The moment a generated number can enter a board pack, the fund has acquired a reconciliation problem it did not have before.

PRACTICE

The best monitoring implementations I have seen produce a list of questions for the chair, and refuse to produce a judgement about the company.

3.3 Value creation across a portfolio, honestly described

The pitch for portfolio-wide value creation is that a fund can find a commercial pattern in one holding and apply it across 15. The pitch is sound and the execution rarely is, because the pattern usually depends on the specifics of a business the fund does not control day to day. The survey evidence points the same way. Nearly 40 per cent of general partners in the outlook study expected no material financial impact on portfolio companies during 2026, and those expecting impact leaned towards cost savings.¹⁴ **SURVEY**

There is a version of this that works, and it is narrower than the pitch. Procurement, service-desk deflection and contract administration repeat across holdings in recognisable form, and a fund that builds one good implementation and redeploys it is doing something real. Identifying the pattern is cheap. Implementing it 15 times is the whole of the work, and it is done by people inside each company who have a day job and a plan of their own, which is the part the portfolio-wide slide never shows.



Traceable diligence. A finding that cannot be opened back to its source is an opinion with a citation style.

3.4 What to leave alone for now

Two categories should stay out of the first wave. Automated investment recommendation belongs there because the output cannot be validated against anything until the position is realised, which places the feedback loop years beyond the evaluation window. Anything touching valuation belongs there because a generated assumption inside a model becomes indistinguishable from a derived one within a single version, and the firm has no model risk standard of its own to catch it.

Neither is a permanent prohibition. In both, the cost of an error is high and the ability to detect it is low, and no supervisory framework currently tells a manager what good looks like. Work of that description can wait a year and lose nothing.

CHAPTER FOUR

Six ways this fails in a fund

Failures in this field are not random. They fall into six recognisable patterns, and five of the six are invisible at the moment they occur.

Risk registers in this area tend to list technologies. A more useful taxonomy lists the ways in which a fund ends up holding a position it cannot defend, because those are the events that cost money. The six below are drawn from the published risks in the standards literature and from what I have seen go wrong in practice. I am proposing the grouping rather than reporting it. **ECAVEO**

THE FAILURE TAXONOMY

FAILURE	WHAT ACTUALLY HAPPENS	VISIBLE AT THE TIME
Confidential leakage	Deal, portfolio, lender or investor material enters a system outside the firm's contractual perimeter	No
Coverage illusion	The system answers accurately on the material it was given, and nobody establishes what it was not given	No
Synthesis displacement	A fluent summary is circulated, and the evidence behind it stops being read	No
Barrier crossing	Retrieval surfaces material across fund, deal or portfolio boundaries that a person would have been walled from	No
Model drift in a live process	A supplier update changes behaviour during a transaction, and the later output is compared against the earlier one as though nothing moved	No
Capability overstatement	Marketing, investor reporting or a due diligence questionnaire describes more than the firm does	Yes, eventually

TABLE 3 The right-hand column is the point of the table. Five of the six failures produce no signal at the time, which is why detection has to be designed in.

4.1 Coverage illusion is the one to worry about first

Of the six, coverage illusion is the most likely to affect an investment decision and the least likely to be caught. A system asked whether a data room contains a particular class of risk will answer about the data room. The question a partner believes they asked is whether the risk exists.

The remedy is a design requirement. Telling people to be careful will not do it. Any system used in diligence should report its own denominator, meaning the documents it read, the documents it could not parse, the entities it expected to find material for and did not, and the date range it covered. Ask a supplier whether their product does this. Most do not, and the answer usually arrives as an explanation of why it is difficult.

4.2 Synthesis displacement, and why good writing makes it worse

The better the summary, the fewer people open the source. That is a statement about human attention, and it was true of well-written diligence reports long before any of this. What changed is the volume. A team can now generate a polished synthesis of every workstream, and the marginal cost of producing one is low enough that they will.

A fluent summary is the most efficient device ever invented for stopping people reading the evidence.

The design response is to make the evidence trail cheaper to open than the summary is to trust. Where every assertion carries an inline link to the paragraph it came from, spot-checking becomes a 10-second act and reviewers do it. Where the evidence sits in a separate appendix, they do not. That is a user-interface decision with a governance consequence, and it is usually made by somebody who was not told it was a governance decision.

THE DESIGN RESPONSE

Make the source one click from the claim, and the claim will be checked. Put it in an appendix and it will not.

WHERE EACH FAILURE IS CAUGHT

A At the boundary Contracts, tenancy and entitlements stop confidential material and barrier crossings before they happen.	B In the output Coverage statements, source links and stated uncertainty expose illusion and displacement.	C In the pipeline Version records and regression tests catch supplier drift during a live transaction.	D Before publication A claims register and compliance approval catch overstatement.
--	--	--	---

FIGURE 2 Each failure has one natural interception point. A control placed anywhere else costs more and catches less.

Ecaveo framework. **ECAVEO**

4.3 Barrier crossing

Conflicts and information barriers are deterministic controls, and retrieval systems are probabilistic ones. Placing the second in charge of the first is the error. A model asked to find relevant precedents will find relevant precedents, including in the deal the firm is walled from, unless entitlements are enforced before retrieval.

The control is architectural. Permissions are resolved at the point the index is queried, the identity of the person asking is carried through every hop, and no component in the chain holds a broader view than the person it is answering. Systems that retrieve broadly and then suppress results are common, cheaper to build, and unsuitable for a firm with information barriers.

4.4 Capability overstatement, which is the one that has produced fines

The fifth failure is the only one with a published enforcement record, and it is also the easiest to prevent. A firm that maintains a register of what it actually does, who owns the evidence for each claim, and which claims have been approved for external use, will not accidentally tell an investor something untrue.

The discipline extends to due diligence questionnaires. My own view is that answers about artificial intelligence in a questionnaire should be drafted by the person accountable for the system and approved by compliance, in the same manner as a performance claim, because that is what a regulator will treat it as.

CHAPTER FIVE

The control model

Controls should scale with the consequence of an error. How clever the system is does not come into it. Four tiers cover everything a fund does.

The organising question is who bears the cost when the output is wrong. A memo summarised badly costs an hour. A diligence finding missed badly costs a fund. Tiering by consequence produces a model that people can apply without convening a committee every time, which is worth more in practice than a more complete scheme nobody uses.

RISK TIERS AND MINIMUM CONTROLS

TIER	TYPICAL SCOPE	MINIMUM CONTROLS
One	Public or non-confidential internal material, with reversible assistance	Approved tool, training, source checking and human review before anything is published
Two	Confidential information, or decisions internal to a workflow	Registered use case, data-flow review, evaluation, logging and a named reviewer
Three	Investor, deal or portfolio work product	Formal privacy, security and vendor assurance, with testing on representative material
Four	Investment decisions, valuations, investor communications and autonomous material actions	A senior accountable owner, independent verification, strict entitlements, monitoring and a tested rollback

TABLE 4 Tier is set by the consequence of an error and by who bears it. The sophistication of the underlying system does not enter the assessment.

5.1 The six technical controls that carry the weight

Most of the assurance burden is met by six controls, and they are consistent with the security guidance from the National Cyber Security Centre, the American standards body's generative profile, and the Information Commissioner's accountability expectations.^{1,2,3} **STANDARD**

1. Enterprise identity and deal, fund or portfolio 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 corpus, 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.

5.2 The control that matters most is a record, not a restriction

Five of those six are restrictions. The fourth is a record, and my own view is that it does more work than the other five together. A fund that can reconstruct which model, which documents and which permissions produced a given finding can investigate an error. A fund that cannot is left arguing about recollection.

The decision record extends the same logic to the committee. What carried the vote, what was noted and discounted, who dissented and on what grounds. None of that is expensive to capture at the time and none of it can be recovered later. Whether a fund that starts now will actually read those records back in three years is a different matter, and depends on habits no framework can install.

ILLUSTRATIVE SCENARIO

The model that changed during the exclusivity period

An invented buyout team runs a contract analysis across a target's customer base in March and repeats it in May after a further tranche of documents arrives. The May output flags 14 problematic termination provisions against nine in March, and the team reads the difference as a deterioration in the new tranche.

The supplier had updated the underlying model in April. The change in behaviour accounted for most of the difference, and the new documents accounted for little of it. Nobody recorded which version produced the March result, so the comparison could not be rerun.

A version record and a small regression set would have cost a day to build and would have answered the question in an hour.

5.3 Vendor assurance, in the order that matters

Vendor questions in this market tend to run long and land softly. Four questions do most of the work. What happens to our material, and where in the contract does it say so. How is the system's behaviour versioned and how are we told when it changes. What does the system do when it does not know, and can we see that state. Who at your firm is accountable when it is wrong.

A supplier that answers those four crisply is usually competent across the rest. My own practice is to ask the third question first, because a system that cannot express uncertainty will not be safe in a tier three workflow whatever else it does well.



Portfolio monitoring. The output that earns its place is a question for the board, arriving before the meeting.



PART THREE

Proof, economics and delivery

CHAPTER SIX

Proving it, and the evidence gate

An evaluation protocol written after people have used the tool will confirm what they already believe. Write it first.

Adoption and enthusiasm are useful signals and neither proves value or safety. A pilot should begin with a protocol that states the baseline, the test material, the measures, the thresholds and the conditions under which the work stops. Three of those five are routinely omitted, and the omitted ones are always the last two.

THE SIX DIMENSIONS OF AN EVIDENCE GATE

DIMENSION	WHAT IS MEASURED	THE GATE
Outcome	Cycle time, coverage, search time or throughput	Improves the intended workflow against a real baseline
Quality	Accuracy, completeness, support for each claim, corrections and rework	No material regression, with critical errors below a stated threshold
Judgement	Challenge quality, override behaviour, escalation and reviewer understanding	Users retain independent judgement and can explain the result
Security and privacy	Permission leakage, sensitive-data exposure, unsafe actions and red-team findings	No unresolved critical finding, with residual risk accepted by a named owner
Economics	Licence, compute, integration, assurance, review and support cost	Total cost is justified by verified benefit, including the cost of checking
Operations	Reliability, supplier changes, monitoring, support and rollback	A named owner and tested operational controls are in place

TABLE 5 A pilot passes when every dimension clears its gate. Partial passes are how organisations end up scaling something that works for the people who chose it.

6.1 Three pilots worth running

A diligence evidence assistant, scoped to source-linked extraction and question answering over a permissioned or synthetic data room, tested on recall of known issues, the rate of unsupported claims, reviewer time and security behaviour.

A portfolio variance investigator, scoped to drafting explanations and investigation questions from approved monthly packs, tested on time to a validated explanation, correction rate and the confidence of the finance owner who has to stand behind the numbers.

A firm knowledge assistant, scoped to retrieval of prior theses, sector work, playbooks and policies under existing permissions, tested on search time, answer support, permission leakage and whether people come back to it unprompted in week six.

6.2 The minimum assurance record

Each production system should carry a record covering six things. The purpose, intended users, benefit, owner and review date. The allowed and prohibited uses, data classes, jurisdictions and actions. The model, supplier, prompt, retrieval configuration, connectors, permissions and dependencies. The risk assessments, contractual controls and any professional or regulatory decisions taken. The evaluation sample, method, results, limitations and approval. The monitoring, feedback, incidents, changes, rollback and retirement plan.

That list is not a compliance exercise. It is the set of facts a partner will want on the afternoon something has gone wrong, and assembling it then is impossible.

6.3 Scale, redesign or stop

Three outcomes, stated in advance. Scale where the benefit is achieved, quality is maintained, 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.

The third outcome needs a sponsor with enough standing to use it. A programme that has never stopped anything is not being evaluated.

CHAPTER SEVEN

The economics of verification

This technology moves cost from producing work to checking it. Value appears only where checking is cheaper than producing was.

Most business cases in this area count the hours saved in production and stop there. That is the wrong side of the ledger. Generation becomes nearly free, and the binding constraint moves to verification, which is performed by senior people whose time is the most expensive in the firm. **ECAVEO**

The consequence is a test that can be applied to any proposed use 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 use case pays. Where it does not, the firm has bought a faster way to create review burden.

WHY SOME USE CASES PAY AND OTHERS DO NOT

1 Cheap to check Retrieval with source links. A reviewer opens three citations and knows. Verification cost stays low, so the saving survives.	2 Costly to check An unsourced synthesis of a data room. Checking it means reading the data room, which is the work it replaced.	3 Impossible to check A recommendation with no stated basis. There is nothing to verify against until the position is realised.
---	---	--

FIGURE 3 The verification cost of an output is a design property. Source links move a use case from the second column to the first.

Ecaveo framework. **ECAVEO**

Read that way, source links stop being a nicety. They are the mechanism that makes a use case economic, because they convert verification from re-performance into spot-checking. Buy a tool without them and you have bought the second column, at the price of the first.

7.1 The cost line nobody budgets

Three costs are routinely missing from the case. The senior review time described above. The assurance work, meaning evaluation, red-teaming and vendor diligence, with the records that support them. The cost of supplier change, meaning the regression testing and re-approval that follows every material model update, which recurs indefinitely and which no pilot business case I have seen includes.

My own rule of thumb, offered as practitioner judgement rather than as measured fact, is that a serious tier three deployment carries ongoing assurance and review costs of a similar order to its licence cost. A partnership budgeting only for licences will be surprised twice, once by the bill and once by the quality of what it got for free. **PRACTICE**

CHAPTER EIGHT

Ninety days, then nine months

The first month is observation and inventory. Nothing durable follows from a programme that starts by choosing a vendor.

THE DELIVERY SEQUENCE

PHASE	TIMING	OUTPUTS
Discover	Weeks 1 to 4	Workflow observation, data map, inventory of approved and unapproved use, baselines and pilot selection
Prove	Weeks 5 to 10	Three controlled pilots with source-linked evaluation, red-team tests and a written stop condition for each
Standardise	Weeks 11 to 18	Approved patterns, model register, single intake route, vendor terms and role-based training
Scale	Months 5 to 9	Proven workflows extended across deals and holdings, with monitored benefits and a named local owner for each

TABLE 6 The Prove phase runs on three narrow workflows with baselines and stop conditions, and the Standardise phase exists so that what worked becomes the default.

8.1 The first 30 days

Interview the partners and the operating team, and observe real work across origination, diligence, portfolio and investor relations. Inventory what is already in use, including the systems embedded in software the firm already owns, which is invariably a longer list than anyone expects. Map data boundaries, contractual duties, entitlements and the people who actually decide.

Publish interim guidance on safe use within the first fortnight, along with a simple route for proposing an experiment, because the alternative is a shadow estate that grows while the programme is still designing itself. Select three pilots with baselines, representative test material and written stop conditions. Demonstrate one bounded workflow end to end, so that partners see capability and controls in the same room.

8.2 What success looks like at nine months

Scaled workflows with measured quality, time or service improvement against the baselines set in month one. A single intake route, approved systems, documented use cases and monitored supplier changes. Users who can explain what the system cannot do and escalate a failure when they meet one. Business owners who hold the controls and the benefits after the programme team has gone. A backlog that separates proven opportunities from research and from uses the firm has decided against.

The last of those is the one that signals maturity. A roadmap with nothing on the prohibited list has not been thought about.

PRACTICE

The inventory is the step most often skipped and the one that most often changes the plan. Assume the firm is already using more than it knows.

CHAPTER NINE

The session to run with the partnership

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 where does the work wait, repeat or fail today?
3. What data is involved, and what contractual or investor obligations apply to it?
4. What error would matter, who would be harmed, and how would anyone detect it?
5. Who reviews the output, who decides, who can override, and who owns the result?
6. What simpler method should this be compared against?
7. How will a change in the model or the supplier be detected during use?
8. What evidence would justify scaling this, and what evidence would stop it?

The fourth and the eighth do most of the filtering. A proposal whose sponsor cannot describe the error that would matter has not been thought through, and a proposal with no stop condition has already committed the firm.

ONE THING TO DECIDE IN THE ROOM

Name the person who will own the decision record. The tooling and the vendor relationship can sit elsewhere. What matters is that somebody owns the record of what the firm decided and why, and in most partnerships nobody currently does.

CHAPTER TEN

Where the evidence is thin

Four things in this paper would change if the evidence changed, and one of them would change the whole argument.

A paper that grades its evidence should say what would overturn it. The following are the points where I am least confident, stated in descending order of how much the argument depends on them.

10.1 The performance question

No independent study establishes that artificial intelligence improves private equity returns. **EVIDENCE GAP** The strongest evidence is associational and published by advisory firms. If a credible causal study appeared, showing for instance that funds using source-linked diligence systems produced measurably better loss ratios on a matched sample, the case for investment would move from process efficiency to alpha, and the risk appetite in Chapter 5 would be too conservative.

I would want to see a matched sample, disclosed selection, a pre-registered outcome measure and an author without a transformation practice. Until then, a manager telling investors that this improves returns is making a claim it cannot support.

10.2 The supervisory question

The current UK position relies on general frameworks, and the specific instruments either do not apply to alternative managers or do not yet exist. That could change quickly. The proposed AIFM regime is at consultation stage with a second consultation and a policy statement to follow, and the results of the FCA's live testing cohorts are expected in the first quarter of 2027.^{6,7} **REGULATOR** A firm that has documented its use cases will find a new rulebook cheap to comply with, and one that has not will be doing the documentation under deadline, which is the expensive way to do it.

10.3 The systemic question

The exposure described in the executive summary is real and I have deliberately kept it separate from the operational argument. The Financial Policy Committee noted in July 2026 that equity market concentration is driven in part by a narrow set of artificial-intelligence-related firms, that related credit has grown rapidly including in private credit and leveraged finance, and that correlated, leveraged positioning could amplify a fall.¹⁷ **REGULATOR** The Bank for International Settlements reported in January 2026 that private credit exposure to this theme had grown from near zero to over \$200bn.⁴ **REGULATOR**

My own view is that a fund's operating decisions about this technology and its balance-sheet exposure to the theme should be governed by different people and reported separately, because conflating them produces a portfolio view that flatters both.

10.4 The question I cannot answer

Whether broader coverage in diligence produces better decisions, or simply more findings to discount, is not something the available evidence resolves. **EVIDENCE GAP** It is testable. A firm with a decision record covering three years, and a discipline of reading outcomes back against recorded reasoning, could answer it for itself within one fund cycle. That is the strongest argument in this paper for starting the record now, and it is the reason Chapter 9 ends where it does.

References

Guidance in this area remains subject to jurisdiction, to the facts of an engagement and to later revision. Where a source is a survey run by a firm with a commercial 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. 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>
3. 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.A.1.600-1>
4. Aldasoro, I., Doerr, S., and Rees, D. *Financing the AI boom: from cash flows to debt*. BIS Bulletin No. 120, Bank for International Settlements, 7 January 2026. <https://www.bis.org/publ/bisbull120.pdf>
5. Financial Conduct Authority. *AI and the FCA: our approach*. First published 8 September 2025, editorially amended 13 February 2026. <https://www.fca.org.uk/firms/innovation/ai-approach>
6. Financial Conduct Authority. *FS25/5: AI Live Testing*. Feedback statement following 67 responses to the April 2025 engagement paper. <https://www.fca.org.uk/publications/feedback-statements/fs25-5-ai-live-testing>
7. Financial Conduct Authority. *CP26/28: The UK AIFM Regime*. 14 July 2026. Consultation stage, with a second consultation and a policy statement to follow. <https://www.fca.org.uk/publications/consultation-papers/cp26-28-uk-aifm-regime>
8. US Securities and Exchange Commission. *SEC Charges Two Investment Advisers with Making False and Misleading Statements About Their Use of Artificial Intelligence*. Press release 2024-36, 18 March 2024. <https://www.sec.gov/newsroom/press-releases/2024-36>
9. European Securities and Markets Authority. *ESMA provides guidance to firms using artificial intelligence in investment services*. 30 May 2024. Scoped to MiFID II investment services. <https://www.esma.europa.eu/press-news/esma-news/esma-provides-guidance-firms-using-artificial-intelligence-investment-services>
10. US Securities and Exchange Commission. *SEC Charges Rimar Capital Entities and Principals with Fraud and Other Violations Involving Purported AI Trading*. Press release 2024-167, 10 October 2024. <https://www.sec.gov/newsroom/press-releases/2024-167>
11. US Securities and Exchange Commission. *SEC Charges Founder of AI Shopping App Nate with Defrauding Investors*. 9 April 2025, with parallel criminal charges brought by the Department of Justice. <https://www.sec.gov/enforcement-litigation/litigation-releases/lr-26282>
12. Bank of England and Financial Conduct Authority. *Artificial intelligence in UK financial services 2024*. 21 November 2024. Third survey in the series, 118 responding firms. <https://www.bankofengland.co.uk/report/2024/artificial-intelligence-in-uk-financial-services-2024>
13. US Securities and Exchange Commission, Division of Examinations. *2026 Examination Priorities*. Artificial intelligence named across claims, governance, resilience and disclosure consistency.
14. Bain & Company with StepStone. *Private Equity's Reality Check: GP Outlook 2026*. 2 March 2026. Survey of 103 investment and investor-relations professionals, fielded December 2025 to January 2026. Advisory-firm survey, self-selecting sample.
15. McKinsey & Company. *Beyond productivity: How AI creates value in private equity*. 23 June 2026. Dataset of 471 private-equity-backed companies across 31 industries. Advisory-firm research, associational rather than causal.
16. FTI Consulting. *2026 Private Equity AI Radar*. 19 May 2026. Survey of 200 fund and operating leaders. Advisory-firm survey with self-assessed success measures and undisclosed field dates.
17. Bank of England. *Financial Policy Committee Record, July 2026*. <https://www.bankofengland.co.uk/financial-policy-committee-record/2026/july-2026>
18. International Monetary Fund. *Global Financial Stability Report*, April 2026, Chapter 1. <https://www.imf.org/-/media/files/publications/gfsr/2026/april/english/ch1.pdf>
19. European Union. *Regulation (EU) 2024/1689 (Artificial Intelligence Act)*, Annex III. High-risk financial services categories cover creditworthiness of natural persons and life and health insurance pricing. Amendments deferring the high-risk obligations were agreed during 2026 and, as at July 2026, awaited formal adoption and publication in the Official Journal.
20. Prudential Regulation Authority. *SSI/23: Model risk management principles for banks*. May 2023, with chief risk officer roundtables on applying it to artificial intelligence and machine learning held in October 2025 and published 24 November 2025. <https://www.bankofengland.co.uk/prudential-regulation/publication/2025/november/praholds-model-risk-management-roundtable-on-ai>
21. United Kingdom. *Data (Use and Access) Act 2025*. Provisions on automated decision-making commenced 5 February 2026.
22. International Organization for Standardization and International Electrotechnical Commission. *ISO/IEC 42001:2023, Artificial intelligence management system*. December 2023.

-
23. Financial Stability Board. *The Financial Stability Implications of Artificial Intelligence*. 14 November 2024, summarised for supervisors by the Bank for International Settlements Financial Stability Institute on 26 June 2025. <https://www.fsb.org/uploads/P14112024.pdf>
24. UK Government. *Artificial Intelligence Playbook for the UK Government*. 10 February 2025. <https://www.gov.uk/government/publications/ai-playbook-for-the-uk-government/artificial-intelligence-playbook-for-the-uk-government-html>
25. Institutional Limited Partners Association. *Responsible AI Quick Guide for Asset Owners*, produced with Malk Partners as part of the ESG Quick Guide series. <https://ilpa.org/wp-content/uploads/2023/10/Responsible-AI-Quick-Guide-for-Asset-Owners.pdf>

ECAVEO RESPONSIBLE AI SERIES

The Traceable Thesis

Paper one of three. Version 1.0, July 2026.

ALSO IN THIS SERIES

Paper two. Sufficient and Appropriate. Artificial intelligence, audit evidence and professional scepticism in accounting firms.

Paper three. Verified Authority. Artificial intelligence, privilege and the duty to the court in law firms.

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.