STUDY NO. 7 · AI AND OPERATIONAL SOVEREIGNTY

ISO 55001 and the AI agent: who signs before the auditor?

From Yokkaichi to the audit room: what the standard already allows you to demand of an AI agent on a critical asset

Nizar Younes Mqam · ORCID 0009-0008-5549-7620

Study · 6 October 2026 · Cahiers de technopolitique, no. 7

Since 2022, artificial intelligence agents have been operating valves, writing work orders and preparing maintenance decisions on industrial assets, sometimes without prior human validation. This study rereads the 2024 edition of ISO 55001 with a single question in mind: when an agent prepared or made the decision, what do you show the auditor? It examines nine real cases documented by open sources, from Japan to Saudi Arabia and from California to Morocco, and finds that none of them says who answers for a decision made by the machine, that only one describes the trace of its decisions, and that seven rely on the technology of an identified supplier, outside the operator in at least six cases. It shows that the protocol, the identities and the rules governing these agents are being written beyond the reach of the asset owner, and then that the standard already contains, clause by clause, what is needed to demand five items of any agent touching an asset: a named person responsible, a declared scope of autonomy, the trace of every decision, the data and its source, and the contract with the supplier. An agent that cannot be audited is not progress in asset management: it is one requirement of the standard less.

The telephone at Yokkaichi · A single question · The other shore · Nine machine rooms · What I found nowhere · Who owns the hand? · What the standard already knows how to demand · Back to the telephone · References

The telephone at Yokkaichi

In the winter of 2022, in a chemical plant on Ise Bay, Japan, a distillation column separates butadiene. Two valves govern everything there: the liquid level, the quality of the product, the heat that is recovered or lost. Conventional automatic control had failed on those two valves, and the operators moved them by hand, frequently intervening whenever a disturbance occurred. On 17 January 2022, the hand is taken from them. For thirty-five days, eight hundred and forty hours, an artificial intelligence of a particular kind, a reinforcement learning agent, operates the valves itself [1].

This is not a recommendation on a screen, nor an alert to be validated. The machine acts. And this is where the scene becomes interesting: during the trial, operators pick up the telephone to ask a question that says everything about what we are living through. Can we stop it?

The answer exists, and it is a good one: "we can simply stop the AI and go back to normal DCS operation". The safety systems stayed in place. The trial lasts nearly a year; a new version of the model

is commissioned in April 2022, after a scheduled shutdown; in March 2023, the system is officially adopted, presented as a world first, and it is now part of the site's routine. The operators, for their part, describe manoeuvres that no human would have imagined [1].

That scene stayed with me all autumn. Because it contains, in a few lines, everything an asset manager should demand of a machine that decides: a scope, two valves; reversibility, the return to conventional control; and people who know they can stop it. And because it also contains what nobody, in that file or in any of the eight others I read, says: who answers for what the machine did during those eight hundred and forty hours.

A single question

I reread ISO 55001 in its 2024 edition, from end to end, with a single question in mind: **if an artificial intelligence agent prepared or made the decision, what do I show the auditor?**

For this standard never speaks of AI. It speaks of decisions, evidence and responsible persons. It even requires, and this is new in 2024, that the organization equip itself with a decision-making framework (clause 4.5); it has added knowledge management (7.7) and renamed the former preventive action as predictive action (10.3), which is precisely the ground where agents are settling. Everything in this text rests on a principle that asset managers know by heart and that the world of AI is discovering: a requirement met without evidence is worth nothing before an auditor. What is not demonstrated is not justifiable.

The question is therefore neither "is AI effective?", the press releases answer that at length, nor "is AI dangerous?", others take care of that. It is a question of evidence: who decided, on what data, who answers for it, and where is the trace.

The other shore

Justice must first be done to those already working on these machines, because they are numerous and serious. The European regulation on artificial intelligence, Regulation (EU) 2024/1689, known as the AI Act, requires of high-risk systems logs and traceability, human oversight entrusted to persons "who have the necessary competence, training and authority", and it classes as high-risk the safety components of critical infrastructure: digital infrastructure, road traffic, water, gas, heating, electricity [2]. The legislator itself asked for time: these obligations, planned for August 2026, have just been postponed to December 2027 by the simplification regulation of July 2026, for want of harmonised standards ready in time [2]. The countdown is known, and it is short.

My working library overflows from that shore. OWASP, the Open Worldwide Application Security Project, whose vulnerability lists serve as a reference for security practitioners, published in June 2026 a hundred-and-thirty-nine-page state of agent security and governance. Three figures there give the scale of the problem. In most companies, non-human identities, that is, the service accounts and API keys through which machines authenticate, already outnumber humans one hundred to one. Among them, 97% hold broader rights than necessary. And "ghost agents", launched for one task and never decommissioned, continue to carry valid identities. The conclusion of its section on identity fits in one sentence, "identity is the new perimeter", and it changes the question one asks of a machine. It

is no longer enough to check that it holds a valid key; one must be able to trace back, action by action, to the person who delegated to it the authority to act [3].

In Singapore, the Infocomm Media Development Authority, the public authority for digital matters, gathered more than twenty lawyers to examine who answers when an agent causes harm. Their first finding is that autonomy "diffuses" responsibility: the more the machine decides on its own, the more the actors multiply along the chain, the harder it becomes to say where the fault lay, and the question of who bears the harm of an unforeseeable action remains, in their own text, open [4]. In Europe, a team of nine researchers and practitioners gathered around Luca Nannini, between Piccadilly Labs and the University of Santiago de Compostela, with members of SaferAI and ForHumanity Europe, mapped what European law requires of agent providers, and concludes that a high-risk system whose behavioural drift is not traceable cannot, as things stand, satisfy the essential requirements of the AI Act; it notes in passing that the regulation's official helpdesk describes its own considerations on agents as "preliminary" [5]. In France, the Council for AI and Digital Affairs devotes an entire note to agentic AI, in which it sees a rupture in the circulation of data and a dilution of the user's control [6].

That world, then, looks at the AI system, and it looks at it better and better. But nobody, to my knowledge, looks from the other shore: that of the asset manager, who must demonstrate, clause by clause of ISO 55001, that their decisions remain auditable when an agent prepares them. I ran the test on the most complete mapping I know, a reference work that crosses the GDPR, the European regulation, the American NIST framework and the ISO/IEC 42001 standard over hundreds of requirements applied to agents: ISO 55001 does not appear in it once, nor does asset management [7]. The academic works that display the word governance govern maintenance policy by means of agents, not the agents themselves [8]; the only text that frankly crosses management systems and AI defends the opposite thesis to mine, that of a mature system serving as the foundation for agents [9]; and seventeen searches of the scientific literature returned nothing else [10]. That crossing was the one left open. So I went to see what is actually happening, public files in hand.

Nine machine rooms

Poughkeepsie, New York State. In a critical data centre, 164 air handling units, 124 computer room air conditioners, 11 chillers, 5 boilers, 28 heat exchangers. An agent built on a large language model works there in daily operations: it reads alerts and open work orders, correlates them, forecasts the next ones, groups tasks over two weeks, diagnoses causes, and pushes its conclusions into the maintenance software. With one reservation, written in black and white: the work orders it recommends are sent back "for review and approval" [11]. The site's engineers credit it with diagnoses 25 to 40% faster. The formal evaluation, for its part, rests on thirty-six representative queries; the failure tables cite hallucinations and code errors that blocked the retrieval of work orders. Remember this case: it is the only one of the nine where the agent's outputs and logs are archived "for auditability". It is also a case where the supplier of the tool is the author of the study that evaluates it.

Fadhili, Saudi Arabia. In the acid gas removal unit of a processing plant, several agents of the same technical family as at Yokkaichi, reinforcement learning, control the process "directly and autonomously" [12]. The introduction took place in three phases, with a simulator of the plant built first to train the agents and assess their reliability, then integration into the existing control system to rely on its safety functions. The announced gains are real on paper: 10 to 15% less amine and steam,

about 5% less electricity. But everything comes from a supplier's press release, the implementation is "under detailed evaluation", and on human oversight there is a single indication: manual interventions have decreased.

California. The grid operator processes every day hundreds of requests to take equipment out of service, lines, substations, power plants, for works, maintenance or faults. Before connecting the slightest agent to it, it did the opposite of the great leap: months of listening to more than two hundred operators, fabricated test cases, then each agent tried separately, one by one, before integration [13]. The agents extract keywords from requests, flag discrepancies, retrieve similar outages from the history, write overnight summaries. The operator, for their part, still does the assessing. The national laboratory that documented the pilot speaks of a "co-pilot" approach; the supplier, for its part, announces processing time cut from six hours to thirty minutes, a figure that no independent source confirms [14].

Denver, Colorado. A water utility is testing a conversational assistant that gives the operators of a new plant access to their operating instructions, scattered through long documents. A purely informational role, and a line drawn by the director of business technology and project management: "a machine shouldn't be making those decisions for us" [15]. The detail that matters lies elsewhere: the same utility has used for some twenty years a machine learning model to target the renewal of its pipes, age, location, soil, corrosion, repair history, frost, and renews about 1% of its network each year. Twenty years of prediction without an agent, then an agent that decides nothing: that too is a policy.

Osaka. In an equipment plant, when a technician discovers a fault on a round, an agent gives them the probable cause and the corrective actions in under ten seconds. The equipment drawings were converted into knowledge graphs, crossed with maintenance histories and a root cause analysis method [16]. Announced accuracy: over 90%. Measured on five fault situations, rated by two technicians. North Sea: specialised agents monitor the production of an offshore field, check data quality, spot wells that are candidates for intervention, inspect pipelines by vision, and 30% less manual effort on inspection is announced, 80% on well selection; only the abstract could be read, the full text is paywalled [17]. Norway: an oil operator that collects about one million time series but monitors only a small share of them has launched an anomaly detection programme in which agents are to link the anomaly to the work order, drawings and history gathered, "before a human even picks up the phone" [18].

And Morocco. The maintenance subsidiary of the OCP group has practised since 2020 an approach it calls "bionic", historical data and expert knowledge combined, monitors its oils online, and insists, through its chief executive, on "Humanware", human expertise; it recommends starting with small pilots before scaling up, and warns against using safety sensors for predictive maintenance [19]. Sound caution, thin documentation: neither deployment scope nor figures. It is not an agent, it is a predictive model, and I keep it precisely for that: it shows where one starts from.

To measure what separates all this from assumed autonomy, one last stop, outside artificial intelligence. In the Australian Pilbara, iron ore trains have been running without drivers since 2018: about two hundred locomotives, more than one thousand seven hundred kilometres of track, sixteen mines, four port terminals, under camera surveillance from an operations centre [20]. Total autonomy, on a heavy asset, that worries nobody. Because the perimeter is closed, the rules

deterministic, the surveillance centralised, and one knows who operates. Everything the agents do not yet have.

There remains the plumbing, and it is going in fast. IBM Maximo, IBM's maintenance management suite present in major industries around the world, announces in its version 9.2 "agentic workflows": a supervisor agent that plans and executes multi-step tasks, creation and modification actions in natural language, "Take AI-assisted actions (create, update, modify)", and a server using the MCP protocol that allows outside agents to be plugged in [21]. It is the tool in which the Poughkeepsie agent was industrialised. And on that protocol, it is the American National Security Agency that published, in May 2026, its design considerations: "MCP's rapid proliferation has outpaced the development of its security model", and it inverts the usual pattern, since it is the servers that query and sometimes execute on behalf of clients [22]. In other words: the door through which a third-party agent will tomorrow write a work order in your CMMS is already fitted, and those who know it best are worried about it.

What I found nowhere

Over the course of these readings, the same hollow keeps coming back, and it ends up making more noise than all the rest. **None of the nine files says who answers for a decision made or prepared by the agent.** Who signs the work order at Poughkeepsie? Who answers for a valve movement at Yokkaichi? Not a name, not a role, anywhere. A single file describes the trace of its decisions, IBM's; a second, California, describes it halfway. Five out of nine describe human oversight, in very different forms: an approval in a piece of software, a stop button, an operator who assesses, a principle asserted by a director, a word, "Humanware". And seven out of nine rely on the technology of an identified supplier, outside the operator in at least six cases, who is sometimes the evaluator of its own agent, and with whom, in California, the agents and the models are hosted.

This hollow is not peculiar to my nine files. The AI Agent Index, the first public index of deployed agents, published in February 2025 by an international team gathered around Stephen Casper of MIT and Noam Kolt of the Hebrew University, quantifies the same imbalance at scale: seven systems out of ten publish documentation, but fewer than one in five displays a formal safety policy, and fewer than one in ten reports safety testing [23]. One documents what the agent can do; one does not document what has been verified. Meanwhile, in simulation, prompt injections already propagate from agent to agent in robotic systems, through the prompt structures they share, to the point of inducing adverse actions [24]. Between a valve at Yokkaichi and a robotic arm, the difference is not one of kind.

In other words: autonomy is already here, two real installations run without prior human validation, and the plumbing that will generalise it is in place. Evidence, for its part, has not followed. This is not a finding about technology. It is a finding about what these organisations could demonstrate, tomorrow morning, to an auditor.

Who owns the hand?

Let us return one last time to Yokkaichi, but to look at something other than the valves. Let us look at the hands. Until then, those of the operators held the column, frequently intervening whenever a

disturbance occurred, with that knowledge which is not written down and which the standard calls, since 2024, knowledge. Then the machine took over, and the operators saw manoeuvres pass that no human would have imagined [1]. Read that sentence twice. It is admiring and, to my mind, it is terrible. The operators add that they learn by watching the machine at work. This is how I hear it: there now exist in that plant gestures that no human devised, and those who held the column have become their pupils. The plant is Japanese, the supplier of the model is Japanese, the control system is Japanese, and the guidelines on AI reliability to which the operator's general manager of production engineering had himself contributed are Japanese [1]. That is a rare alignment. Almost everywhere else, the hand that decides has come from elsewhere.

Take the inventory with me. Seven of my nine files rely on the technology of an identified supplier, outside the operator in at least six cases. In California, the agents that prepare the work of the grid operators run in the supplier's data centres, where the grid manager's data already resides [13]. At Poughkeepsie, the one who designed the agent is also the one who evaluated it [11]. At Fadhili, process safety rests on the supplier's control system, into which its own agents have been integrated [12]. A single utility, in Denver, takes the trouble to say that its data stays with it, under the terms of the licence agreement [15]: that is proof that the clause exists, and the silence of the others is all the more audible for it.

Then there is the plumbing, and this is where the matter becomes properly technopolitical. The protocol through which agents talk to each other and to your software, the one IBM has just opened in Maximo [21], was popularised by a San Francisco company in November 2024 [22]. Eighteen months later, it is neither a standards body nor an industrial authority that publishes its security considerations: it is the American National Security Agency, which observes that "MCP's rapid proliferation has outpaced the development of its security model", and which describes a protocol in which it is the servers that query the clients and sometimes execute on their behalf [22]. The door through which an agent will tomorrow write into the CMMS of a Moroccan pumping station or a French district heating network was drawn in California, and its security considerations were published at Fort Meade. The owner of the station, for their part, was consulted by nobody.

Behind the door, the keys. The three OWASP figures cited above, one hundred machine identities for every human, 97% excessive rights, ghost agents never decommissioned, take on their industrial meaning here, and the same report tells what that gives when it goes wrong: a hijacked coding agent harvesting credentials "at machine speed", with humans intervening at only four to six decision points per campaign; and a production database wiped by an agent without any attacker being involved, because the extent of its rights governed the extent of the damage, whatever its cause [3]. Remember the authors' formula: the scope of the identity is the blast radius. In a plant, the blast radius is not a metaphor. And that identity, who issues it, who revokes it, who holds its register? Almost never the operator.

Finally, the rules. They are being written right now, and they are being written without us. The vulnerability list carries authority from a community of practitioners, that of OWASP [3]. The most advanced thinking on the civil liability of agents comes from a regulator in Singapore [4]. Brussels legislated, then asked for time: its obligations for high-risk systems, including critical infrastructure, are postponed to December 2027, for want of harmonised standards ready in time [2], and a report by The Future Society, relayed by the European team cited above, estimates that those standards will

likely fail to fully address the risks specific to agents [5]. The Singapore lawyers put their finger on the mechanics of this value chain: those who develop the models are best placed to shape their capabilities, but have only a limited view of the final use; downstream, the use is precise, but the possibility of acting on the agent's basic behaviour is reduced [4]. Knowledge is upstream, responsibility is downstream. And all the way downstream, at the end of the chain, there is the asset owner: the one who holds the structure, the operating permit, the neighbours, and the accident. They have only one table where they sit in their own name, that of the ISO committee preparing the specification on decision making in asset management. They would do well to sit there knowing what is being decided at the other tables.

Seen from here, the word "Humanware", which the chief executive of OCP's maintenance subsidiary places at the heart of his company's AI, alongside software intelligence [19], changes meaning. It is not the caution of a latecomer. It is the lucid posture of someone who knows that the brain will come from elsewhere, and that the only sovereignty accessible to an adopter is the one they keep in their own people: start small, understand what you are plugging in, keep the expertise that allows the machine to be contradicted. South as north, when one writes neither the protocol, nor the model, nor the rule, two things remain that one can still hold: the competence of one's operators, and the evidence.

This Cahier thus extends what I wrote about maintenance AI, which leaves you ownership and takes control from you [25]: the agent is the next step, the one where you keep the asset and give up the decision. It also extends what I wrote about delegated operation, where the works come back and the capability does not [26]: at Yokkaichi, "simply stop the AI" is true today; in ten years, when the last operators who drove the column by hand have gone, who will know how to take the valves back? And it extends what I wrote about the cloud that catches fire [27]: the Poughkeepsie agent lives in a data centre, and we saw in March 2026 what becomes of one hundred and nine digital services when drones reach the buildings that host them. The hand is no longer a hand. It is a key, hosted with a third party, obeying a foreign protocol, under rules written by others.

That is why the five items I am now going to ask of the standard are not auditor's paperwork. A name, a scope, a trace, sources, a contract: these are the only things left to the asset owner when everything else is rented. Audit is not the opposite of innovation. In a world where the brain of operations belongs to someone else, it is the last act of operational sovereignty that remains within reach of the one who bears the risk. The hand belongs to whoever can prove they have it.

What the standard already knows how to demand

For this is where this reading of the standard takes on its meaning: there is no need for a new standard, nor for a moratorium. The 2024 edition already contains everything, and it has contained it since before the agents. Clause 5.3 requires top management to assign responsibilities and authorities by name: it made no exception for decisions that came from a machine. Clause 7.2 knows only "persons" whose competence must be proven: an agent's competence has no box, and that is precisely the problem of whoever will have to answer for it. Clause 7.5.2 lists the author among what identifies a document: who is the author of a work order written by an agent? Clause 7.6 requires that every item of data have a specification and a source: that covers everything that feeds the machine. Clause 8.1 asks for a level of evidence sufficient to have confidence that the processes have been carried out

as planned: without a trace of what the agent did, that confidence is an act of faith. Clause 8.3 expressly targets technologies provided by external providers: the agent is almost always one of them. And clause 10.2 asks that the causes of an incident be traced back, which presupposes reconstructing the decision that preceded it: without traceability, it becomes inapplicable. As to whether a retrained model, or an agent whose behaviour has drifted, is a change to be managed within the meaning of clause 8.2, I leave the question open; it deserves its own debate.

I have already described elsewhere what this standard requires clause by clause [28], and how maintenance AI shifts control without shifting ownership [25]. This text is about something else: evidence. And laid end to end, these clauses outline a requirement that any client can formulate today, in a specification, in a contract, in a management review. Of any agent touching an asset, I would ask five items: **who answers for its decisions**, by name; **its scope of autonomy**, declared, from mere advice to action without validation; **the trace of every decision**, retained and usable; **the data that feeds it and its source**; **the contract with its supplier**, interfaces and responsibilities included.

What strikes me is that the vocabulary already exists on the other shore, and that it says the same thing. What security calls a delegation chain and an audit trail, the standard calls authority and documented information. What the European team cited above calls untraceable behavioural drift, clause 8.2 calls an unmanaged change. What the Singapore lawyers call diffuse responsibility, clause 5.3 calls an unassigned authority. The two shores describe the same bridge, each from its own side, and neither has yet crossed. European operators also have a second lever: what the standard allows one to demand today, the law will in turn impose on high-risk systems, on the horizon of December 2027, logs and human oversight entrusted by name included [2]. And auditors are tooling up: the German Federal Office for Information Security is submitting an AI security audit methodology based on ISAE 3000, the very standard of assurance audits [29].

None of this is unprecedented. Traceability exists at Poughkeepsie, reversibility at Yokkaichi, the introduction method in California, framed autonomy in the Pilbara. What is missing is never the technique. It is the requirement. Nobody yet asks for the five items together, and the ISO committee currently preparing a specification on decision making in asset management [30] would do useful work by remembering it.

One honesty, to finish. I measured nothing myself, and these files are, for the most part, written by those who sell the agents. Failures, withdrawals, incidents are not published; the corpus shows what organisations want to show. And when a source says nothing on a point, that does not prove the thing does not exist. But that is precisely my conclusion, and it is the standard's: in asset management, what is not demonstrated is not justifiable, and what is not justifiable does not pass the audit.

Back to the telephone

The best news from this inquiry lies in the scene that opens it. At Yokkaichi, the operators' question had an answer: yes, the AI can be stopped, and one can return to the previous way of running the plant. Someone had thought of it, had built it, and knew how to say so. That is exactly what

managing an asset is: not preventing the machine from deciding, but being able to show, at any moment, who decides, within what limits, and how one takes the hand back.

One line is still missing at the bottom of that file, as at the bottom of the eight others: the name of the person who answers. As long as it stays empty, "yes, we can stop it" will be only half an answer to the operators' question. The other half, the one an auditor will ask, comes down to three questions: stopped by whom, in the name of what authority, and provable through what chain, from the order given to the machine to the person who assumes responsibility for it? A plant that can answer those three questions does not merely have a stop button. It has governance.

An agent that cannot be audited is not progress in asset management. It is one requirement of the standard less.

References

All sources were consulted and verified between 2 and 6 October 2026. The [n] cross-references in the text point to this list.

[1] Yokogawa, "In a World First, Yokogawa's Autonomous Control AI Is Officially Adopted for Use at an ENEOS Materials Chemical Plant", press release of 30 March 2023; "In a World First, Yokogawa and JSR Use AI to Autonomously Control a Chemical Plant for 35 Consecutive Days", press release of 22 March 2022 (integration with Yokogawa's CENTUM VP control system); and customer reference page "'A world's first' is now routine – ENEOS Materials Successfully Achieves Autonomous Control at a Plant Using AI". 35-day trial (840 hours) from 17 January to 21 February 2022, new version of the model in April 2022 after a scheduled shutdown, adoption in March 2023; project subsidised by METI. Read on 2 October 2026, except the 2022 press release, read on 7 October 2026. https://www.yokogawa.com/us/news/press-releases/2023/2023-03-30/ https://www.yokogawa.com/us/news/press-releases/2022/2022-03-22/ https://www.yokogawa.com/us/library/resources/references/successstory-eneos-materials/

[2] Regulation (EU) 2024/1689 of 13 June 2024 (OJEU of 12 July 2024, EUR-Lex text 02024R1689-20240712, initial version, reread on 6 October 2026): record-keeping and traceability (Article 12), human oversight (Article 14), obligations of deployers including human oversight entrusted to persons with the necessary competence, training and authority (Article 26); high-risk systems of Annex III, including safety components of critical infrastructure. Dates reset by Regulation (EU) 2026/1744 of 8 July 2026 (OJEU of 24 July 2026, amended Article 113, reread on 6 October 2026): obligations of Sections 1 to 3 of Chapter III applicable from 2 December 2027 for Annex III and from 2 August 2028 for Annex I. See also ISO/IEC 42001:2023, artificial intelligence management systems. https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02024R1689-20240712 https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32026R1744

[3] OWASP GenAI Security Project, "State of Agentic AI Security and Governance", version 2.01, June 2026, 139 pp., CC BY-SA 4.0 licence. Section "Agent Identity vs Non-Human Identity": ratios of 100 to 1 and up to 500 to 1 between non-human identities and humans, 97% of non-human identities with excessive privileges (Entro Security 2025 report cited), ghost agents, delegation chains preserving the principal's context. Read on 6 October 2026. https://genai.owasp.org/resource/state-of-agentic-ai-security-and-governance/

[4] Infocomm Media Development Authority (Singapore), "Legal Responsibility for AI Agents", discussion paper, May 2026, 36 pp.: working group of more than twenty lawyers; the autonomy and decision-making of agents "diffuse" responsibility for outcomes; responsibility for unforeseeable actions along the value chain is left as an open question. Read on 6 October 2026. https://www.imda.gov.sg/-/media/imda/files/about/emerging-tech-and-research/artificial-intelligence/agents-legal-responsibility.pdf

[5] Nannini L., Smith A. L., Maggini M. J., Panai E., Feliciano S., Tiulkanov A., Maran E., Gealy J., Bisconti P., "AI Agents under EU Law: A Compliance Architecture for AI Providers", working paper of 7 April 2026, arXiv:2604.04604: conclusion on untraceable behavioural drift; the regulation's official helpdesk considerations

described as preliminary; harmonised standards M/613 (CEN/CENELEC JTC 21) in progress. Read on 6 October 2026. https://arxiv.org/abs/2604.04604

[6] Conseil de l'IA et du numérique (CIANum) and CNIL, note "IA agentique et protection des données personnelles : équation à inconnues multiples pour les utilisateurs" (Agentic AI and personal data protection: an equation with multiple unknowns for users), July 2026, 17 pp. Read on 6 October 2026. https://www.conseil-ia-numerique.fr/nos-travaux/ia-agentique-et-protection-des-donnees-personnelles-equation-inconnues-multiples-pour

[7] Kenney N. M. (2026), "Governing Agents: A Practitioner's Cross-Framework Reference. Mapping GDPR, the EU AI Act, NIST AI RMF, and ISO/IEC 42001 to AI Agents", Digital 520, 1st edition. No occurrence of ISO 55001 or of asset management in the whole work (verified on the full text on 6 October 2026).

[8] Crespo Márquez A., Gómez Fernández J. F. (2026), "Agentic AI for autonomous preventive maintenance policy governance: a multi-agent framework for dynamic industrial environments", Expert Systems with Applications. https://doi.org/10.1016/j.eswa.2026.131767

[9] McMurray B. G. (June 2026), "AI: A Case for Safety & Operational Excellence", white paper. https://bryangmcmurray.com/

[10] Searches carried out on 2 October 2026: 17 OpenAlex queries (keywords and full text) crossing ISO 55001/55000 with AI and agents, 14 abstracts read, plus the professional literature via metasearch engine. Corpus of nine cases assembled the same day, all sources open and read on that date. https://openalex.org/

[11] Zhou N., Patel D., Bhattacharyya A. (2026), "Deployed AI Agents for Industrial Asset Management: CodeReAct Framework for Event Analysis and Work Order Automation", Proceedings of the AAAI Conference on Artificial Intelligence, 40(47). POKMAIN site, Poughkeepsie; 36 evaluation queries; model comparison. https://doi.org/10.1609/aaai.v40i47.41453

[12] Yokogawa, "Aramco and Yokogawa Achieve a Major Milestone with Commissioning of Multiple Autonomous Control AI Agents at Major Gas Facility", press release of 29 October 2025. Single supplier source. https://www.yokogawa.com/us/news/press-releases/2025/2025-10-29/

[13] Allsup M., "Under the hood of CAISO's AI outage management pilot", Latitude Media, 28 July 2025; Stewart E. et al., "Adoption of AI in the Utility T&D Sector: Use Cases, Consequence, Assessment and Benefits", Idaho National Laboratory, INL/RPT-25-87526, September 2025. https://www.latitudemedia.com/news/under-the-hood-of-caisos-ai-outage-management-pilot/ https://inl.gov/content/uploads/2026/02/Adoption-of-AI-in-the-Utility-TD-Sector.pdf

[14] OATI, programme of the OATI Energy Conference 2026, CAISO session, and "AI Genie" product page. Supplier's figure, not confirmed by an independent source. https://www.oati.com/events/oati-energy-conference-2026/ https://www.oati.com/ai-genie/

[15] Miller E., "Rethinking the AI-Water Nexus", Headwaters Magazine (Water Education Colorado), Spring 2026 issue, 26 June 2026. https://watereducationcolorado.org/publications-and-radio/headwaters-magazine/spring-2026-the-artificial-intelligence-issue/rethinking-the-ai-water-nexus/

[16] Hitachi and Daikin, "Through Collaborative Creation, Daikin and Hitachi Begin Trial Operation of AI Agent Supporting Equipment Failure Diagnostics in Factories", press release of 22 April 2025. No source read says whether the trial concluded. https://www.hitachi.com/en/press/articles/2025/04/0422/

[17] Shekhawat D., Barua J., Bhatia K., Saumya S. (2025), "Offshore Production Surveillance and Intervention Using Multi Agent AI", SPE Offshore Europe, SPE-226728-MS. Abstract only. https://doi.org/10.2118/226728-MS

[18] Cognite, "Cognite and NVIDIA Operationalize NV-Tesseract with Aker BP to Transform Industrial Anomaly Detection and Safeguard Critical Assets", press release of 23 March 2026. https://www.cognite.com/en/company/newsroom/cognite-and-nvidia-operationalize-nv-tesseract-with-aker-bp-to-transform-industrial-anomaly

[19] Hespress English, "OCP maintenance solutions harnesses AI for predictive factory maintenance", 17 September 2025 (remarks by the chief executive at the 37th technical conference of the Arab Fertilizer Association). https://en.hespress.com/121030-ocp-maintenance-solutions-harnesses-ai-for-predictive-factory-maintenance.html

[20] Rio Tinto, "World-first autonomous trains deployed at Rio Tinto's iron ore operations", press release of 28 December 2018; Hitachi Rail, case study "Australia: World First - Fully Autonomous Heavy-Haul Rail Network", 30 July 2024. Railway automation, not an AI agent. Read on 2 October 2026. https://www.riotinto.com/en/news/

releases/2018/world-first-autonomous-trains-deployed https://www.hitachirail.com/case-studies/australia-world-first-fully-autonomous-heavy-haul-rail-network/

[21] IBM, "Working smarter with agentic workflows in Maximo Assistant" (Maximo Application Suite 9.2), IBM Think, 12 August 2026. Read on 2 October 2026. https://www.ibm.com/think/perspectives/working-smarter-agentic-workflows-maximo-assistant

[22] National Security Agency, "Model Context Protocol (MCP): Security Design Considerations for AI-Driven Automation", Cybersecurity Information Sheet U/OO/6030316-26, May 2026, version 1.0, 17 pp. Read on 6 October 2026. https://www.nsa.gov/Portals/75/documents/Cybersecurity/CSI_MCP_SECURITY.pdf

[23] Casper S., Bailey L., Hunter R. et al. (2025), "The AI Agent Index", arXiv:2502.01635: 70.1% of the systems surveyed publish documentation, 49.3% their code, 19.4% a formal safety policy, 7.5% external safety testing and 9% developer testing. Read on 6 October 2026. https://arxiv.org/abs/2502.01635

[24] Nagaraja N., Bagari A., Bahsi H. (2026), "When Prompts Control Robots: Prompt Injection Attacks in Multi-Agent Robotic Systems", arXiv:2608.00747. Read on 6 October 2026. https://arxiv.org/abs/2608.00747

[25] Mqam N. Y. (2026), "L'IA de maintenance : vous gardez la propriété, vous perdez la maîtrise" (Maintenance AI: you keep ownership, you lose control), Cahiers de technopolitique, no. 2. https://doi.org/10.5281/zenodo.23207203

[26] Mqam N. Y. (2026), "Exploitation déléguée : les ouvrages reviennent, la capacité ne revient pas" (Delegated operation: the works come back, the capability does not), Cahiers de technopolitique, no. 4. https://doi.org/10.5281/zenodo.23207209

[27] Mqam N. Y. (2026), "Quand le cloud prend feu" (When the cloud catches fire), Cahiers de technopolitique, no. 1: drone strikes of 1 and 2 March 2026 on AWS data centres in the United Arab Emirates and Bahrain, more than 109 AWS services disrupted in the Emirates. https://doi.org/10.5281/zenodo.23206998

[28] Mqam N. Y. (2026), "ISO 55001 : ce que la norme exige vraiment, clause par clause" (ISO 55001: what the standard really requires, clause by clause), Études MARAMM, no. 4. https://doi.org/10.5281/zenodo.22977896

[29] Bundesamt für Sicherheit in der Informationstechnik (BSI), "A5 Audit Methodology based on ISAE 3000", community draft v0.9 of 30 June 2026, 27 pp. Read on 6 October 2026. https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/A5/A5_node.html

[30] ISO/TC 251, draft technical specification ISO/TS 55014, "Guidance for asset management decision making", stage 30.60, record consulted on 2 October 2026. https://www.iso.org/standard/89182.html

To cite this study

Mqam, N. Y. (2026). *ISO 55001 and the AI agent: who signs before the auditor? From Yokkaichi to the audit room: what the standard already allows you to demand of an AI agent on a critical asset.* Cahiers de technopolitique, no. 7. https://doi.org/10.5281/zenodo.23207212

Cahiers de technopolitique, no. 7 · Study · 6 October 2026
Publisher and publication director: Nizar Younes Mqam · Place of publication: Tourcoing, France
DOI: 10.5281/zenodo.23207212
Contact: y.nizar@proton.me · https://technopolitique.eu/cahiers/n7/en/
Author's ORCID: https://orcid.org/0009-0008-5549-7620
© 2026 Nizar Younes Mqam. Published in open access under the Creative Commons Attribution – NonCommercial – NoDerivatives 4.0 International licence (CC BY-NC-ND 4.0). Sharing and citation permitted with attribution; modification and commercial use prohibited. creativecommons.org/licenses/by-nc-nd/4.0/