ÉTUDE N° 7 · VOLET IA ET SOUVERAINETÉ OPÉRATIONNELLE

ISO 55001 face à l'agent IA : qui signe devant l'auditeur ?

De Yokkaichi à la salle d'audit : ce que la norme permet déjà d'exiger d'un agent IA sur un actif critique

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

Étude · 6 octobre 2026 · Cahiers de technopolitique, n° 7

Depuis 2022, des agents d'intelligence artificielle manœuvrent des vannes, rédigent des ordres de travail et préparent des décisions de maintenance sur des actifs industriels, parfois sans validation humaine préalable. Cette étude relit la norme ISO 55001 dans son édition de 2024 avec une seule question : quand c'est un agent qui a préparé ou pris la décision, que montre-t-on à l'auditeur ? Elle examine neuf cas réels documentés par des sources ouvertes, du Japon à l'Arabie saoudite, de la Californie au Maroc, et constate qu'aucun ne dit qui répond d'une décision prise par la machine, qu'un seul décrit la trace de ses décisions, et que sept reposent sur la technologie d'un fournisseur identifié, extérieur à l'exploitant dans six cas au moins. Elle montre que le protocole, les identités et les règles qui gouvernent ces agents s'écrivent hors de portée du propriétaire de l'actif, puis que la norme contient déjà, clause par clause, de quoi exiger cinq pièces de tout agent touchant un actif : un responsable nommé, un périmètre d'autonomie déclaré, la trace de chaque décision, les données et leur source, le contrat avec le fournisseur. Un agent qu'on ne peut pas auditer n'est pas un progrès de la gestion d'actifs : c'est une exigence de la norme en moins.

Le téléphone de Yokkaichi · Une seule question · L'autre rive · Neuf salles de machines · Ce que je n'ai trouvé nulle part · À qui appartient la main ? · Ce que la norme sait déjà demander · Retour au téléphone · Références

Le téléphone de Yokkaichi

L'hiver 2022, dans une usine chimique de la baie d'Ise, au Japon, une colonne de distillation sépare du butadiène. Deux vannes y règlent tout : le niveau de liquide, la qualité du produit, la chaleur qu'on récupère ou qu'on perd. Le contrôle automatique classique avait échoué sur ces deux vannes, et les opérateurs les manœuvraient à la main, en intervenant souvent dès qu'une perturbation survenait. Le 17 janvier 2022, on leur retire la main. Pendant trente-cinq jours, huit cent quarante heures, une intelligence artificielle d'un genre particulier, un agent d'apprentissage par renforcement, actionne les vannes elle-même[1].

Ce n'est pas une recommandation sur un écran, ni une alerte à valider. La machine agit. Et c'est là que la scène devient intéressante : pendant l'essai, des opérateurs décrochent leur téléphone pour poser une question qui dit tout de ce que nous sommes en train de vivre. Est-ce qu'on peut l'arrêter ?

La réponse existe, et elle est bonne : « we can simply stop the AI and go back to normal DCS operation », on peut simplement arrêter l'IA et revenir à la conduite habituelle. Les sécurités sont

restées en place. L'essai dure près d'un an ; une nouvelle version du modèle est mise en service en avril 2022, après un arrêt programmé ; en mars 2023, le système est adopté officiellement, présenté comme une première mondiale, et il fait aujourd'hui partie de la routine du site. Les opérateurs, eux, racontent des manœuvres qu'aucun humain n'aurait imaginées[1].

Cette scène m'a accompagné tout l'automne. Parce qu'elle contient, en quelques lignes, tout ce qu'un gestionnaire d'actifs devrait exiger d'une machine qui décide : un périmètre, deux vannes ; une réversibilité, le retour au contrôle conventionnel ; et des gens qui savent qu'ils peuvent l'arrêter. Et parce qu'elle contient aussi ce que personne, dans ce dossier ni dans aucun des huit autres que j'ai lus, ne dit : qui répond de ce que la machine a fait pendant ces huit cent quarante heures.

Une seule question

J'ai relu la norme ISO 55001 dans son édition de 2024, de bout en bout, avec une seule question en tête : **si c'est un agent d'intelligence artificielle qui a préparé ou pris la décision, qu'est-ce que je montre à l'auditeur ?**

Car cette norme ne parle jamais d'IA. Elle parle de décisions, de preuves et de responsables. Elle exige même, et c'est nouveau en 2024, que l'organisme se dote d'un cadre de prise de décision (clause 4.5) ; elle a ajouté la gestion des connaissances (7.7) et rebaptisé l'ancienne action préventive en action prédictive (10.3), qui est précisément le terrain où les agents s'installent. Tout, dans ce texte, repose sur un principe que les gestionnaires d'actifs connaissent par cœur et que le monde de l'IA découvre : une exigence respectée sans preuve ne vaut rien devant un auditeur. Ce qui n'est pas démontré n'est pas justifiable.

La question n'est donc ni « l'IA est-elle efficace ? », les communiqués y répondent à longueur de colonnes, ni « l'IA est-elle dangereuse ? », d'autres s'en chargent. C'est une question de preuve : qui a décidé, sur quelles données, qui en répond, et où est la trace.

L'autre rive

Il faut d'abord rendre justice à ceux qui travaillent déjà sur ces machines, parce qu'ils sont nombreux et sérieux. Le règlement européen sur l'intelligence artificielle, le règlement (UE) 2024/1689 que l'on appelle l'AI Act, impose aux systèmes à haut risque des journaux et une traçabilité, un contrôle humain confié à des personnes « qui disposent des compétences, de la formation et de l'autorité nécessaires », et il range dans le haut risque les composants de sécurité des infrastructures critiques : numérique, trafic routier, eau, gaz, chauffage, électricité[2]. Le législateur lui-même a demandé du temps : ces obligations, prévues pour août 2026, viennent d'être reportées à décembre 2027 par le règlement de simplification de juillet 2026, faute de normes harmonisées prêtes[2]. Le compte à rebours est connu, et il est court.

Ma bibliothèque de travail déborde de cette rive-là. L'OWASP, l'Open Worldwide Application Security Project, dont les listes de vulnérabilités servent de référence aux praticiens de la sécurité, a publié en juin 2026 un état de la sécurité et de la gouvernance des agents de cent trente-neuf pages. Trois chiffres y donnent l'échelle du problème. Dans la plupart des entreprises, les identités non humaines, c'est-à-dire les comptes de service et les clés d'API par lesquels les machines s'authentifient, sont déjà

cent fois plus nombreuses que les humains. Parmi elles, 97 % détiennent des droits plus étendus que nécessaire. Et des « agents fantômes », lancés pour une tâche puis jamais décommissionnés, continuent de porter des identités valides. La conclusion de sa section sur l'identité tient en une phrase, « identity is the new perimeter », l'identité est le nouveau périmètre, et elle change la question que l'on pose à une machine. Il ne suffit plus de vérifier qu'elle détient une clé valide ; il faut pouvoir remonter, action par action, jusqu'à la personne qui lui a délégué l'autorité d'agir[3].

À Singapour, l'Infocomm Media Development Authority, l'autorité publique du numérique, a réuni plus de vingt juristes pour examiner qui répond quand un agent cause un dommage. Leur premier constat est que l'autonomie « diffuse » la responsabilité : plus la machine décide seule, plus les acteurs se multiplient le long de la chaîne, plus il devient difficile de dire où était la faute, et la question de savoir qui porte le dommage d'une action imprévisible reste, dans leur propre texte, ouverte[4]. En Europe, une équipe de neuf chercheurs et praticiens réunie autour de Luca Nannini, entre Piccadilly Labs et l'université de Saint-Jacques-de-Compostelle, avec des membres de SaferAI et de ForHumanity Europe, a cartographié ce que le droit européen exige des fournisseurs d'agents, et conclut qu'un système à haut risque dont la dérive comportementale n'est pas traçable ne peut pas, en l'état, satisfaire aux exigences essentielles de l'AI Act ; elle relève au passage que le guichet officiel du règlement qualifie ses propres considérations sur les agents de « préliminaires »[5]. En France, le Conseil de l'IA et du numérique consacre à l'IA agentique une note entière, qui y voit une rupture dans la circulation des données et une dilution de la maîtrise de l'utilisateur[6].

Ce monde-là regarde donc le système d'IA, et il le regarde de mieux en mieux. Mais personne, à ma connaissance, ne regarde depuis l'autre rive : celle du gestionnaire d'actifs, qui doit démontrer, clause par clause d'ISO 55001, que ses décisions restent auditables quand un agent les prépare. J'ai fait le test sur la cartographie la plus complète que je connaisse, un ouvrage de référence qui croise le RGPD, le règlement européen, le cadre américain du NIST et la norme ISO/IEC 42001 sur des centaines d'exigences appliquées aux agents : ISO 55001 n'y apparaît pas une fois, la gestion d'actifs non plus[7]. Les travaux universitaires qui affichent le mot gouvernance gouvernent la politique de maintenance par des agents, pas les agents eux-mêmes[8] ; le seul texte qui croise franchement systèmes de management et IA défend la thèse inverse de la mienne, celle d'un système mature servant de socle aux agents[9] ; et dix-sept recherches dans la littérature scientifique n'ont rien rendu d'autre[10]. C'est ce croisement qui était libre. Je suis donc allé voir ce qui se passe réellement, dossiers publics à l'appui.

Neuf salles de machines

Poughkeepsie, État de New York. Dans un centre de données critique, 164 centrales de traitement d'air, 124 armoires de climatisation, 11 groupes froids, 5 chaudières, 28 échangeurs. Un agent fondé sur un grand modèle de langage y travaille au quotidien : il lit les alertes et les ordres de travail non soldés, les corrèle, prévoit les prochains, regroupe les tâches sur deux semaines, diagnostique des causes, et pousse ses conclusions dans le logiciel de maintenance. Avec une réserve, écrite noir sur blanc : les ordres de travail qu'il recommande sont renvoyés « for review and approval », pour revue et approbation humaine[11]. Les ingénieurs du site le créditent de diagnostics 25 à 40 % plus rapides. L'évaluation formelle, elle, repose sur trente-six requêtes représentatives ; les tableaux d'échec citent des hallucinations et des erreurs de code qui ont bloqué la récupération d'ordres de travail. Retenez ce

cas : c'est le seul des neuf où l'on archive les sorties et les journaux de l'agent « for auditability », pour l'auditabilité. C'est aussi un cas où le fournisseur de l'outil est l'auteur de l'étude qui l'évalue.

Fadhili, Arabie saoudite. Dans l'unité d'élimination des gaz acides d'une usine de traitement, plusieurs agents de la même famille technique qu'à Yokkaichi, l'apprentissage par renforcement, contrôlent « directly and autonomously » le procédé[12]. L'introduction s'est faite en trois phases, avec un simulateur de l'usine construit d'abord pour entraîner les agents et évaluer leur fiabilité, puis une intégration au système de contrôle existant pour s'appuyer sur ses sécurités. Les gains annoncés sont réels sur le papier : 10 à 15 % d'amine et de vapeur en moins, environ 5 % d'électricité. Mais tout vient d'un communiqué du fournisseur, la mise en œuvre est « en cours d'évaluation détaillée », et sur la supervision humaine, une seule indication : les interventions manuelles ont baissé.

Californie. Le gestionnaire du réseau électrique traite chaque jour des centaines de demandes de retrait d'ouvrages, lignes, postes, centrales, pour travaux, maintenance ou défaut. Avant d'y brancher le moindre agent, il a fait l'inverse du grand saut : des mois d'écoute de plus de deux cents opérateurs, des cas de test fabriqués, puis chaque agent essayé séparément, un à un, avant intégration[13]. Les agents extraient les mots-clés des demandes, signalent les écarts, retrouvent les retraits semblables dans l'historique, rédigent des synthèses nocturnes. L'opérateur, lui, instruit toujours. Le laboratoire national qui a documenté le pilote parle d'une approche « co-pilote » ; le fournisseur, lui, annonce un traitement passé de six heures à trente minutes, chiffre qu'aucune source indépendante ne confirme[14].

Denver, Colorado. Une compagnie d'eau teste un assistant conversationnel qui donne aux opérateurs d'une usine neuve l'accès à leurs consignes d'exploitation, dispersées dans de longs documents. Rôle purement informatif, et une ligne fixée par le directeur des technologies métier et de la gestion de projets : « a machine shouldn't be making those decisions for us », une machine ne devrait pas prendre ces décisions à notre place[15]. Le détail qui compte est ailleurs : la même compagnie utilise depuis une vingtaine d'années un modèle d'apprentissage automatique pour cibler le renouvellement de ses conduites, âge, emplacement, sol, corrosion, historique des réparations, gel, et renouvelle environ 1 % de son réseau par an. Vingt ans de prédiction sans agent, puis un agent qui ne décide de rien : c'est aussi une politique.

Osaka. Dans une usine d'équipements, quand un technicien découvre une panne en ronde, un agent lui indique la cause probable et les actions correctives en moins de dix secondes. Les plans des équipements ont été convertis en graphes de connaissances, croisés avec les historiques de maintenance et une méthode d'analyse de cause[16]. Précision annoncée : plus de 90 %. Mesurée sur cinq situations de panne, notées par deux techniciens. Mer du Nord : des agents spécialisés surveillent la production d'un champ offshore, contrôlent la qualité des données, repèrent les puits candidats à intervention, inspectent les pipelines par vision, et l'on annonce 30 % d'effort manuel en moins sur l'inspection, 80 % sur la sélection des puits ; seul le résumé a pu être lu, le texte complet est payant[17]. Norvège : un exploitant pétrolier, qui collecte environ un million de séries temporelles mais n'en surveille qu'une petite part, a lancé un programme de détection d'anomalies où des agents doivent relier l'anomalie à l'ordre de travail, plans et historique rassemblés, « before a human even picks up the phone », avant même qu'un humain ne décroche[18].

Et le Maroc. La filiale maintenance du groupe OCP pratique depuis 2020 une approche qu'elle appelle « bionique », données historiques et savoir des experts mêlés, surveille ses huiles en ligne, et insiste,

par la voix de son directeur général, sur le « Humanware », l'expertise humaine ; elle recommande de commencer par de petits pilotes avant de généraliser, et met en garde contre l'usage des capteurs de sécurité pour la maintenance prédictive[19]. Prudence de bon aloi, documentation mince : ni périmètre de déploiement, ni chiffres. Ce n'est pas un agent, c'est un modèle prédictif, et je le garde justement pour cela : il montre d'où l'on part.

Pour mesurer ce qui sépare tout cela de l'autonomie assumée, un dernier arrêt, hors intelligence artificielle. Dans le Pilbara australien, des trains de minerai de fer roulent sans conducteur depuis 2018 : environ deux cents locomotives, plus de mille sept cents kilomètres de voie, seize mines, quatre terminaux portuaires, sous surveillance par caméras depuis un centre d'opérations[20]. Une autonomie totale, sur un actif lourd, qui n'inquiète personne. Parce que le périmètre est fermé, les règles déterministes, la surveillance centralisée, et que l'on sait qui exploite. Tout ce que les agents n'ont pas encore.

Reste la plomberie, et elle s'installe vite. IBM Maximo, la suite de gestion de maintenance d'IBM présente dans les grandes industries du monde entier, annonce dans sa version 9.2 des « agentic workflows » : un agent superviseur qui planifie et exécute des tâches en plusieurs étapes, des actions de création et de modification en langage naturel, « Take AI-assisted actions (create, update, modify) », et un serveur au protocole MCP qui permet de brancher des agents extérieurs[21]. C'est l'outil dans lequel l'agent de Poughkeepsie a été industrialisé. Et sur ce protocole, c'est l'Agence nationale de sécurité américaine qui a publié, en mai 2026, ses considérations de conception : « MCP's rapid proliferation has outpaced the development of its security model », sa prolifération a devancé son modèle de sécurité, et il inverse le schéma habituel, puisque ce sont les serveurs qui interrogent et parfois exécutent pour les clients[22]. Autrement dit : la porte par laquelle un agent tiers écrira demain un ordre de travail dans votre GMAO est déjà posée, et ceux qui la connaissent le mieux s'en inquiètent.

Ce que je n'ai trouvé nulle part

Au fil de ces lectures, un même creux revient, et il finit par faire plus de bruit que tout le reste. **Aucun des neuf dossiers ne dit qui répond d'une décision prise ou préparée par l'agent.** Qui signe l'ordre de travail de Poughkeepsie ? Qui répond d'une manœuvre de vanne à Yokkaichi ? Pas un nom, pas un rôle, nulle part. Un seul dossier décrit la trace de ses décisions, celui d'IBM ; un second, la Californie, la décrit à moitié. Cinq sur neuf décrivent une supervision humaine, sous des formes très différentes : une approbation dans un logiciel, un bouton d'arrêt, un opérateur qui instruit, un principe affirmé par un directeur, un mot, « Humanware ». Et sept sur neuf reposent sur la technologie d'un fournisseur identifié, extérieur à l'exploitant dans six cas au moins, qui est parfois l'évaluateur de son propre agent, et chez qui, en Californie, les agents et les modèles sont hébergés.

Ce creux n'est pas propre à mes neuf dossiers. L'AI Agent Index, premier index public des agents déployés, publié en février 2025 par une équipe internationale réunie autour de Stephen Casper, du MIT, et de Noam Kolt, de l'Université hébraïque, chiffre le même déséquilibre à grande échelle : sept systèmes sur dix publient une documentation, mais moins d'un sur cinq affiche une politique de sécurité formelle, et moins d'un sur dix fait état de tests de sécurité[23]. On documente ce que l'agent sait faire ; on ne documente pas ce qu'on a vérifié. Pendant ce temps, en simulation, des injections de prompt se propagent déjà d'agent en agent dans des systèmes robotiques, par les structures de

prompt qu'ils partagent, jusqu'à induire des actions adverses[24]. Entre une vanne de Yokkaichi et un bras robotique, la différence n'est pas de nature.

Autrement dit : l'autonomie est déjà là, deux installations réelles tournent sans validation humaine préalable, et la plomberie qui la généralisera est posée. La preuve, elle, n'a pas suivi. Ce n'est pas un constat sur la technique. C'est un constat sur ce que ces organisations pourraient démontrer, demain matin, à un auditeur.

À qui appartient la main ?

Revenons une dernière fois à Yokkaichi, mais pour regarder autre chose que les vannes. Regardons les mains. Jusque-là, celles des opérateurs ont tenu la colonne, en intervenant souvent quand survenait une perturbation, avec ce savoir qui ne s'écrit pas et que la norme appelle, depuis 2024, des connaissances. Puis la machine a pris le relais, et les opérateurs ont vu passer des manœuvres qu'aucun humain n'aurait imaginées[1]. Lisez cette phrase deux fois. Elle est admirative, et, à mes yeux, elle est terrible. Les opérateurs ajoutent qu'ils apprennent en regardant la machine faire. Je l'entends ainsi : il existe désormais dans cette usine des gestes qu'aucun humain n'a conçus, et ceux qui tenaient la colonne en sont devenus les élèves. L'usine est japonaise, le fournisseur du modèle est japonais, le système de contrôle est japonais, et les lignes directrices sur la fiabilité de l'IA auxquelles le directeur de l'ingénierie de production de l'exploitant avait lui-même contribué sont japonaises[1]. C'est un alignement rare. Presque partout ailleurs, la main qui décide est venue d'ailleurs.

Faites l'inventaire avec moi. Sept de mes neuf dossiers reposent sur la technologie d'un fournisseur identifié, extérieur à l'exploitant dans six cas au moins. En Californie, les agents qui préparent le travail des opérateurs du réseau tournent dans les centres de données du fournisseur, là où résident déjà les données du gestionnaire[13]. À Poughkeepsie, celui qui a conçu l'agent est aussi celui qui l'a évalué[11]. À Fadhili, la sécurité du procédé repose sur le système de contrôle du fournisseur, dans lequel ses propres agents ont été intégrés[12]. Une seule compagnie, à Denver, prend la peine de dire que ses données restent chez elle, aux termes de l'accord de licence[15] : c'est la preuve que la clause existe, et le silence des autres n'en est que plus audible.

Puis il y a la plomberie, et c'est là que l'affaire devient proprement technopolitique. Le protocole par lequel les agents se parlent et parlent à vos logiciels, celui qu'IBM vient d'ouvrir dans Maximo[21], a été popularisé par une entreprise de San Francisco en novembre 2024[22]. Dix-huit mois plus tard, ce n'est ni un organisme de normalisation ni une autorité industrielle qui en publie les considérations de sécurité : c'est l'Agence nationale de sécurité américaine, qui constate que « MCP's rapid proliferation has outpaced the development of its security model », et qui décrit un protocole où ce sont les serveurs qui interrogent les clients et parfois exécutent pour eux[22]. La porte par laquelle un agent écrira demain dans la GMAO d'une station de pompage marocaine ou d'un réseau de chaleur français a été dessinée en Californie, et ses considérations de sécurité ont été publiées à Fort Meade. Le propriétaire de la station, lui, n'a été consulté par personne.

Derrière la porte, les clés. Les trois chiffres de l'OWASP cités plus haut, cent identités de machine pour un humain, 97 % de droits excessifs, des agents fantômes jamais décommissionnés, prennent ici leur sens industriel, et le même rapport raconte ce que cela donne quand cela tourne mal : un agent de programmation détourné qui moissonne des identifiants « at machine speed », des humains

n'intervenant qu'à quatre à six points de décision par campagne ; et une base de données de production effacée par un agent sans qu'aucun attaquant soit impliqué, parce que l'étendue de ses droits gouvernait l'ampleur du dégât, quelle qu'en soit la cause[3]. Retenez la formule des auteurs : la portée de l'identité est le rayon de l'explosion. Dans une usine, le rayon de l'explosion n'est pas une métaphore. Et cette identité, qui la délivre, qui la révoque, qui en détient le registre ? Presque jamais l'exploitant.

Enfin, les règles. Elles s'écrivent en ce moment, et elles s'écrivent sans nous. La liste des vulnérabilités fait autorité depuis une communauté de praticiens, celle de l'OWASP[3]. La réflexion la plus avancée sur la responsabilité civile des agents vient d'un régulateur de Singapour[4]. Bruxelles a légiféré, puis a demandé du temps : ses obligations pour les systèmes à haut risque, dont les infrastructures critiques, sont reportées à décembre 2027, faute de normes harmonisées prêtes[2], et un rapport de The Future Society, relayé par l'équipe européenne citée plus haut, estime que ces normes risquent de ne pas couvrir pleinement les risques propres aux agents[5]. Les juristes de Singapour mettent le doigt sur la mécanique de cette chaîne de valeur : ceux qui développent les modèles sont les mieux placés pour en façonner les capacités, mais n'ont qu'une vue limitée de l'usage final ; en aval, l'usage est précis, mais la possibilité d'agir sur le comportement de base de l'agent est réduite[4]. Le savoir est en amont, la responsabilité est en aval. Et tout en aval, au bout de la chaîne, il y a le propriétaire de l'actif : celui qui détient l'ouvrage, l'autorisation d'exploiter, les riverains, et l'accident. Il n'a qu'une seule table où il siège en son nom propre, celle du comité ISO qui prépare la spécification sur la prise de décision en gestion d'actifs. Il ferait bien de s'y asseoir en sachant ce qui se décide aux autres tables.

Vu d'ici, le mot « Humanware », que le directeur de la filiale maintenance de l'OCP place au cœur de l'IA de son entreprise, à côté de l'intelligence logicielle[19], change de sens. Ce n'est pas la prudence d'un retardataire. C'est la posture, lucide, de quelqu'un qui sait que le cerveau viendra d'ailleurs, et que la seule souveraineté accessible à un adoptant est celle qu'il garde dans ses propres gens : commencer petit, comprendre ce qu'on branche, conserver l'expertise qui permet de contredire la machine. Au sud comme au nord, quand on n'écrit ni le protocole, ni le modèle, ni la règle, il reste deux choses qu'on peut encore tenir : la compétence de ses opérateurs, et la preuve.

Ce Cahier prolonge ainsi ce que j'écrivais de l'IA de maintenance, qui vous laisse la propriété et vous prend la maîtrise[25] : l'agent est l'étape suivante, celle où l'on garde l'actif et où l'on cède la décision. Il prolonge aussi ce que j'écrivais de l'exploitation déléguée, où les ouvrages reviennent et où la capacité ne revient pas[26] : à Yokkaichi, « simply stop the AI » est vrai aujourd'hui ; dans dix ans, quand les derniers opérateurs qui ont conduit la colonne à la main seront partis, qui saura reprendre les vannes ? Et il prolonge ce que j'écrivais du cloud qui prend feu[27] : l'agent de Poughkeepsie vit dans un centre de données, et nous avons vu en mars 2026 ce que deviennent cent neuf services numériques quand des drones atteignent les bâtiments qui les hébergent. La main n'est plus une main. C'est une clé, hébergée chez un tiers, obéissant à un protocole étranger, sous des règles écrites par d'autres.

C'est pourquoi les cinq pièces que je vais maintenant demander à la norme ne sont pas de la paperasse d'auditeur. Un nom, un périmètre, une trace, des sources, un contrat : ce sont les seules choses qui restent au propriétaire de l'actif quand tout le reste est loué. L'audit n'est pas le contraire de l'innovation. Dans un monde où le cerveau de l'exploitation appartient à quelqu'un d'autre, c'est le

dernier acte de souveraineté opérationnelle qui reste à la portée de celui qui porte le risque. La main appartient à celui qui peut prouver qu'il l'a.

Ce que la norme sait déjà demander

Car c'est ici que cette lecture de la norme prend son sens : il n'y a besoin ni d'une norme nouvelle, ni d'un moratoire. Celle de 2024 contient déjà tout, et elle le contient depuis avant les agents. La clause 5.3 exige que la direction attribue nommément responsabilités et autorités : elle n'a pas prévu d'exception pour les décisions venues d'une machine. La clause 7.2 ne connaît que des « personnes » dont il faut prouver la compétence : la compétence d'un agent n'a pas de case, et c'est bien le problème de celui qui devra en répondre. La clause 7.5.2 cite l'auteur parmi ce qui identifie un document : qui est l'auteur d'un ordre de travail rédigé par un agent ? La clause 7.6 exige que chaque donnée ait une spécification et une source : cela couvre tout ce qui alimente la machine. La clause 8.1 demande un niveau de preuve suffisant pour avoir confiance dans le fait que les processus ont été suivis comme prévu : sans trace de ce que l'agent a fait, cette confiance est un acte de foi. La clause 8.3 vise expressément les technologies fournies par des prestataires : l'agent est presque toujours l'une d'elles. Et la clause 10.2 demande de remonter aux causes d'un incident, ce qui suppose de reconstituer la décision qui l'a précédé : sans traçabilité, elle devient inapplicable. Quant à savoir si un modèle réentraîné, ou un agent dont le comportement a dérivé, est un changement à maîtriser au sens de la clause 8.2, je laisse la question ouverte ; elle mérite son propre débat.

J'ai déjà décrit ailleurs ce que cette norme exige clause par clause[28], et comment l'IA de maintenance déplace la maîtrise sans déplacer la propriété[25]. Ce texte-ci porte sur autre chose : la preuve. Et mises bout à bout, ces clauses dessinent une exigence que tout donneur d'ordre peut formuler dès aujourd'hui, dans un cahier des charges, dans un contrat, dans une revue de direction. À tout agent touchant un actif, je demanderais cinq pièces : **qui répond de ses décisions**, nommément ; **son périmètre d'autonomie**, déclaré, du simple conseil à l'action sans validation ; **la trace de chaque décision**, conservée et exploitable ; **les données qui l'alimentent et leur source** ; **le contrat avec son fournisseur**, interfaces et responsabilités comprises.

Ce qui me frappe, c'est que le vocabulaire existe déjà sur l'autre rive, et qu'il dit la même chose. Ce que la sécurité appelle chaîne de délégation et piste d'audit, la norme l'appelle autorité et informations documentées. Ce que l'équipe européenne citée plus haut appelle dérive comportementale non traçable, la clause 8.2 l'appelle un changement non maîtrisé. Ce que les juristes de Singapour appellent responsabilité diffuse, la clause 5.3 l'appelle une autorité non attribuée. Les deux rives décrivent le même pont, chacune de son côté, et aucune n'a encore traversé. Les exploitants européens disposent d'ailleurs d'un second levier : ce que la norme permet d'exiger dès aujourd'hui, la loi l'imposera à son tour aux systèmes à haut risque, à l'horizon de décembre 2027, journaux et contrôle humain nommément confié compris[2]. Et les auditeurs s'outillent : l'office fédéral allemand de la sécurité de l'information soumet une méthodologie d'audit de la sécurité de l'IA fondée sur ISAE 3000, la norme même des audits d'assurance[29].

Rien de tout cela n'est inédit. La traçabilité existe à Poughkeepsie, la réversibilité à Yokkaichi, la méthode d'introduction en Californie, l'autonomie encadrée dans le Pilbara. Ce qui manque n'est jamais la technique. C'est l'exigence. Personne ne demande encore les cinq pièces ensemble, et le

comité ISO qui prépare en ce moment une spécification sur la prise de décision en gestion d'actifs[30] ferait œuvre utile en s'en souvenant.

Une honnêteté pour finir. Je n'ai rien mesuré moi-même, et ces dossiers sont, pour la plupart, écrits par ceux qui vendent les agents. Les échecs, les retraits, les incidents ne se publient pas ; le corpus montre ce que les organisations veulent montrer. Et quand une source ne dit rien d'un point, cela ne prouve pas que la chose n'existe pas. Mais c'est précisément ma conclusion, et c'est celle de la norme : en gestion d'actifs, ce qui n'est pas démontré n'est pas justifiable, et ce qui n'est pas justifiable ne passe pas l'audit.

Retour au téléphone

La meilleure nouvelle de cette enquête tient dans la scène qui l'ouvre. À Yokkaichi, la question des opérateurs avait une réponse : oui, on peut arrêter l'IA, et revenir à la conduite d'avant. Quelqu'un y avait pensé, l'avait construit, et savait le dire. C'est exactement cela, gérer un actif : non pas empêcher la machine de décider, mais pouvoir montrer, à tout moment, qui décide, dans quelles limites, et comment on reprend la main.

Il manque encore une ligne au bas de ce dossier, comme au bas des huit autres : le nom de celui qui répond. Tant qu'elle restera vide, « oui, on peut l'arrêter » ne sera qu'une moitié de réponse à la question des opérateurs. L'autre moitié, celle qu'un auditeur posera, tient en trois questions : arrêtée par qui, au nom de quelle autorité, et prouvable à travers quelle chaîne, de l'ordre donné à la machine jusqu'à la personne qui en assume la responsabilité ? Une usine qui sait répondre à ces trois questions n'a pas seulement un bouton d'arrêt. Elle a une gouvernance.

Un agent qu'on ne peut pas auditer n'est pas un progrès de la gestion d'actifs. C'est une exigence de la norme en moins.

Références

Toutes les sources ont été consultées et vérifiées entre le 2 et le 6 octobre 2026. Les renvois [n] du texte pointent vers cette liste.

[1] Yokogawa, « In a World First, Yokogawa's Autonomous Control AI Is Officially Adopted for Use at an ENEOS Materials Chemical Plant », communiqué du 30 mars 2023 ; « In a World First, Yokogawa and JSR Use AI to Autonomously Control a Chemical Plant for 35 Consecutive Days », communiqué du 22 mars 2022 (intégration au système de contrôle CENTUM VP de Yokogawa) ; et page de référence client « “A world's first” is now routine – ENEOS Materials Successfully Achieves Autonomous Control at a Plant Using AI ». Essai de 35 jours (840 heures) du 17 janvier au 21 février 2022, nouvelle version du modèle en avril 2022 après arrêt programmé, adoption en mars 2023 ; projet subventionné par le METI. Lues le 2 octobre 2026, sauf le communiqué de 2022, lu le 7 octobre 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] Règlement (UE) 2024/1689 du 13 juin 2024 (JOUE du 12 juillet 2024, texte EUR-Lex 02024R1689-20240712, version initiale, relu le 6 octobre 2026) : journaux et traçabilité (article 12), contrôle humain (article 14), obligations des déployeurs dont le contrôle humain confié à des personnes disposant des compétences, de la formation et de l'autorité nécessaires (article 26) ; systèmes à haut risque de l'annexe III, dont les composants de sécurité des infrastructures critiques. Dates recalées par le règlement (UE) 2026/1744 du 8 juillet 2026 (JOUE du 24 juillet 2026, article 113 modifié, relu le 6 octobre 2026) : obligations des sections 1 à 3 du chapitre III applicables au 2 décembre 2027 pour l'annexe III et au 2 août 2028 pour l'annexe I. Voir aussi ISO/IEC 42001:2023, systèmes

de management de l'intelligence artificielle. https://eur-lex.europa.eu/legal-content/FR/TXT/HTML/?uri=CELEX:02024R1689-20240712 https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=CELEX:32026R1744

[3] OWASP GenAI Security Project, « State of Agentic AI Security and Governance », version 2.01, juin 2026, 139 p., licence CC BY-SA 4.0. Section « Agent Identity vs Non-Human Identity » : ratios de 100 pour 1 et jusqu'à 500 pour 1 entre identités non humaines et humains, 97 % d'identités non humaines à privilèges excessifs (rapport Entro Security 2025 cité), agents fantômes, chaînes de délégation préservant le contexte du mandant. Lu le 6 octobre 2026. https://genai.owasp.org/resource/state-of-agentic-ai-security-and-governance/

[4] Infocomm Media Development Authority (Singapour), « Legal Responsibility for AI Agents », document de discussion, mai 2026, 36 p. : groupe de travail de plus de vingt juristes ; l'autonomie et la prise de décision des agents « diffusent » la responsabilité des résultats ; la responsabilité des actions imprévisibles le long de la chaîne de valeur est laissée en question ouverte. Lu le 6 octobre 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 », document de travail du 7 avril 2026, arXiv:2604.04604 : conclusion sur la dérive comportementale non traçable ; considérations du guichet officiel du règlement qualifiées de préliminaires ; normes harmonisées M/613 (CEN/CENELEC JTC 21) en cours. Lu le 6 octobre 2026. https://arxiv.org/abs/2604.04604

[6] Conseil de l'IA et du numérique (CIANum) et CNIL, note « IA agentique et protection des données personnelles : équation à inconnues multiples pour les utilisateurs », juillet 2026, 17 p. Lu le 6 octobre 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, 1re édition. Aucune occurrence d'ISO 55001 ni de la gestion d'actifs dans tout l'ouvrage (vérifié sur le texte intégral le 6 octobre 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. (juin 2026), « AI: A Case for Safety & Operational Excellence », livre blanc. https://bryangmcmurray.com/

[10] Recherches menées le 2 octobre 2026 : 17 requêtes OpenAlex (mots-clés et texte intégral) croisant ISO 55001/55000 avec l'IA et les agents, 14 résumés lus, plus la littérature professionnelle via métamoteur. Corpus de neuf cas constitué le même jour, toutes sources ouvertes et lues à cette 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). Site POKMAIN, Poughkeepsie ; 36 requêtes d'évaluation ; comparaison de modèles. 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 », communiqué du 29 octobre 2025. Source unique du fournisseur. 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 juillet 2025 ; Stewart E. et coll., « Adoption of AI in the Utility T&D Sector: Use Cases, Consequence, Assessment and Benefits », Idaho National Laboratory, INL/RPT-25-87526, septembre 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 de l'OATI Energy Conference 2026, session CAISO, et page produit « AI Genie ». Chiffre du fournisseur, non confirmé par une source indépendante. 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), numéro Spring 2026, 26 juin 2026. https://watereducationcolorado.org/publications-and-radio/headwaters-magazine/spring-2026-the-artificial-intelligence-issue/rethinking-the-ai-water-nexus/

[16] Hitachi et Daikin, « Through Collaborative Creation, Daikin and Hitachi Begin Trial Operation of AI Agent Supporting Equipment Failure Diagnostics in Factories », communiqué du 22 avril 2025. Aucune source lue ne dit si l'essai s'est conclu. 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. Résumé seul. 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 », communiqué du 23 mars 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 septembre 2025 (propos du directeur général à la 37e conférence technique de l'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 », communiqué du 28 décembre 2018 ; Hitachi Rail, étude de cas « Australia: World First - Fully Autonomous Heavy-Haul Rail Network », 30 juillet 2024. Automatisme ferroviaire, pas un agent d'IA. Lues le 2 octobre 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 août 2026. Lu le 2 octobre 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, mai 2026, version 1.0, 17 p. Lu le 6 octobre 2026. https://www.nsa.gov/Portals/75/documents/Cybersecurity/CSI_MCP_SECURITY.pdf

[23] Casper S., Bailey L., Hunter R. et coll. (2025), « The AI Agent Index », arXiv:2502.01635 : 70,1 % des systèmes recensés publient une documentation, 49,3 % leur code, 19,4 % une politique de sécurité formelle, 7,5 % des tests de sécurité externes et 9 % des tests par le développeur. Lu le 6 octobre 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. Lu le 6 octobre 2026. https://arxiv.org/abs/2608.00747

[25] Mqam N. Y. (2026), « L'IA de maintenance : vous gardez la propriété, vous perdez la maîtrise », Cahiers de technopolitique, n° 2. https://doi.org/10.5281/zenodo.23207203

[26] Mqam N. Y. (2026), « Exploitation déléguée : les ouvrages reviennent, la capacité ne revient pas », Cahiers de technopolitique, n° 4. https://doi.org/10.5281/zenodo.23207209

[27] Mqam N. Y. (2026), « Quand le cloud prend feu », Cahiers de technopolitique, n° 1 : frappes de drones des 1er et 2 mars 2026 sur des centres de données d'AWS aux Émirats arabes unis et au Bahreïn, plus de 109 services AWS perturbés aux Émirats. https://doi.org/10.5281/zenodo.23206998

[28] Mqam N. Y. (2026), « ISO 55001 : ce que la norme exige vraiment, clause par clause », Études MARAMM, n° 4. https://doi.org/10.5281/zenodo.22977896

[29] Bundesamt für Sicherheit in der Informationstechnik (BSI), « A5 Audit Methodology based on ISAE 3000 », projet communautaire v0.9 du 30 juin 2026, 27 p. Lu le 6 octobre 2026. https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Kuenstliche-Intelligenz/A5/A5_node.html

[30] ISO/TC 251, projet de spécification technique ISO/TS 55014, « Guidance for asset management decision making », stade 30.60, fiche consultée le 2 octobre 2026. https://www.iso.org/standard/89182.html

Pour citer cette étude

Mqam, N. Y. (2026). *ISO 55001 face à l'agent IA : qui signe devant l'auditeur ? De Yokkaichi à la salle d'audit : ce que la norme permet déjà d'exiger d'un agent IA sur un actif critique.* Cahiers de technopolitique, n° 7. https://doi.org/10.5281/zenodo.23207211

Cahiers de technopolitique, n° 7 · Étude · 6 octobre 2026
Éditeur et directeur de la publication : Nizar Younes Mqam · Lieu d'édition : Tourcoing, France
DOI : 10.5281/zenodo.23207211
Contact : y.nizar@proton.me · https://technopolitique.eu/cahiers/n7/
ORCID de l'auteur : https://orcid.org/0009-0008-5549-7620
© 2026 Nizar Younes Mqam. Publié en accès ouvert sous licence Creative Commons Attribution – Pas d'Utilisation Commerciale – Pas de Modification 4.0 International (CC BY-NC-ND 4.0). Partage et citation autorisés avec attribution ; modification et usage commercial interdits. creativecommons.org/licenses/by-nc-nd/4.0/deed.fr