# IA : la supervision comme impératif de gouvernance

***GBR Patrick PERROT***
*Inspecteur Général*
*Inspection Générale de la Gendarmerie Nationale*

Émancipée de son statut de curiosité expérimentale, l'intelligence artificielle (IA) essaime aujourd'hui ses applications dans tous les secteurs d'activité, marquant ainsi sa transition définitive du laboratoire vers le théâtre opérationnel. Les utilisations se multiplient que ce soit dans le domaine industriel ou celui du particulier. Dans le même temps, les nations comme les instances internationales peinent à mettre en œuvre une régulation qui fait consensus. Les exploitations de l'IA se développent trop souvent de manière anarchique sans réelle gouvernance alors que celle-ci demeure le socle d'une utilisation responsable. En sécurité intérieure, l'enjeu de la responsabilité est majeur tant le risque de mésusage peut être à l'origine d'impacts considérables sur les droits fondamentaux : biais non contrôlés, opacité des modèles et des données, détournements d'usage, manque de robustesse cyber. Face à ces enjeux, la nécessité de superviser les systèmes d'intelligence artificielle par l'évaluation et l'audit est devenue à la fois nécessaire et indispensable. L'IA ne pourra s'imposer en sécurité intérieure qu'à la condition de la confiance du citoyen. Il s'agit alors d'assurer une intégration de l'IA en veillant à une gouvernance placée au bon niveau, à un cadre d'usage qui correspond au juste besoin, à une convergence de compétences, garante de la singularité de l'IA. La facilité croissante de développement des applications d'IA, y compris par des acteurs non spécialisés, renforce encore la nécessité d'une supervision capable d'auditer les systèmes. C'est tout l'objet de mettre en place au sein des organisations sans attendre que cela ne vienne de l'extérieur une instance de supervision, de contrôle et d'audit des systèmes d'IA. L'objectif est d'assurer à la fois une exploitation légitime et de confiance, et une protection de l'institution face aux dérives et mésusages.

## Confiance et légitimité : un double impératif

L'IA suscite de nombreuses inquiétudes quant à une exploitation débridée, notamment au sein des entreprises et des institutions. Alors que le citoyen l'utilise sans réelle exigence dans sa vie quotidienne, il demeure beaucoup plus défiant lorsque cette IA est exploitée par les institutions étatiques ou au sein de l'entreprise. Ce paradoxe oblige les organisations privées comme publiques à mettre en place des mesures assurant une légitimité aux usages. Et ces derniers, pour être durables, doivent répondre à l'exigence de la confiance comme préalable à la légitimité. En effet, confiance et légitimité sont les deux faces d'une même pièce. Alors que la première permet l'adhésion du citoyen, la seconde assure le cadre d'usage dans le respect des valeurs sociétales choisies. Ne pas envisager ces deux volets, c'est prendre le risque d'une désapprobation de l'usage, d'une crise de confiance du citoyen et in fine d'une dépendance aux sirènes venues de l'extérieur.

La célérité de développement de l'IA et de son adoption dans la vie quotidienne a fait oublier que le socle nécessaire à l'adoption est la confiance. Celle-ci ne se décrète pas, elle se gagne et s'entretient. Elle repose sur la connaissance qui permet la compréhension des méthodes trop souvent considérées comme opaques. La connaissance, c'est la capacité à développer et utiliser des systèmes plus

Document publié le 26/09/2025 – L'utilisation de tout ou partie de ce texte doit être soumise à autorisation ©CyberCercle

fiables, plus robustes et donc plus responsables. Les systèmes sont conçus pour des tâches spécialisées et ne peuvent être généralisés sans impact sur la performance et la sécurité. Il s'agit de s'assurer qu'ils fonctionnent dans des environnements maîtrisés, notamment aux niveaux des biais et des risques cyber. En d'autres termes, connaître les méthodes mathématiques sous-jacentes aux systèmes constitue le meilleur moyen pour lutter contre les discriminations et prévenir le détournement des systèmes, tant au niveau des phases d'apprentissage que de décision.
L'impératif de confiance et de légitimité prévient les usages anarchiques et ne doit pas être considéré comme un frein au développement de l'IA. Mettre en place les garanties de supervision est une condition stratégique à une exploitation opérationnelle, efficace, responsable et durable.

## Vers une approche intégrée de la supervision de l'IA en sécurité intérieure

La supervision des systèmes d'IA répond à une exigence légale, même si la mise en oeuvre pratique est encore assez peu précise ou tout au moins laisse une liberté de manœuvre dont il faut savoir se saisir. En sécurité intérieure, l'enjeu est double : il s'agit de permettre d'exploiter des systèmes d'IA, protecteur de notre modèle de société, tout en demeurant très vigilant sur l'impact en termes de droits fondamentaux et de libertés individuelles. A cette fin, une approche intégrée qui rassemble les aspects organisationnels, scientifiques et techniques, opérationnels, éthiques et réglementaires apparaît la plus responsable. Chaque aspect couvre un champ d'audit qui apporte confiance et légitimité.

### *L'audit organisationnel : s'engager en responsabilité*

L'adoption de l'IA au sein d'une organisation répond à une exigence de responsabilité en prévenant le développement d'applications anarchiques dénuées de gouvernance comme de réels cadres d'usage. L'objectif de l'audit organisationnel est de s'assurer de la présence d'une gouvernance engagée et de contrôler les garanties de responsabilité dans la chaîne de subsidiarité. Il s'agit d'examiner le niveau de la structure de gouvernance, la répartition des rôles, les mécanismes de priorisation comme de réponse aux incidents mais aussi l'ensemble de la cohérence structurelle. Alors qu'il est souvent clamé que la technologie est neutre, il ne peut en être dit autant des instances de gouvernance mettant en œuvre une technologie. L'audit organisationnel prévient le risque d'une adoption techno-centrée et replace l'humain au sein de l'enjeu de responsabilité.

### *L'audit scientifique et technique : comprendre pour une meilleure maîtrise*

L'objectif de l'audit scientifique et technique est de garantir le niveau de performance des systèmes de la phase de conception à celle de production. Mais l'examen du seul niveau de performance ne peut suffire : il s'agit aussi d'examiner la robustesse des systèmes à des environnements diversifiés en s'assurant de leur résilience en cas de généralisation comme d'attaques cyber. Les systèmes d'IA peuvent très vite devenir des vecteurs de vulnérabilité. Il est essentiel de contrôler que les systèmes fonctionnent dans les limites définies et conformément aux exigences de performance comme de résilience. A titre d'illustration, dans le cadre d'un outil de classification d'images, il s'agira de vérifier la qualité des données d'apprentissage, de validation et de test, le niveau de faux positifs et de faux négatifs en phase de généralisation, la performance requise au point de fonctionnement choisi et de s'assurer de la robustesse face aux manipulations potentielles des données (empoisonnement, deepfakes, attaques adversariales...).

### *L'audit opérationnel : garantir une action efficace*

L'usage opérationnel pour être efficace doit respecter un besoin réel et non s'inscrire dans un usage artificiel. Il est essentiel en IA de s'assurer que le système respecte les limites du cadre d'usage envisagé, est utilisé conformément aux prescriptions

Document publié le 26/09/2025 – L'utilisation de tout ou partie de ce texte doit être soumise à autorisation ©CyberCercle

requises et prévient toute baisse de compétence professionnelle ou toute perte de responsabilité humaine. En sécurité intérieure, la garantie d'un usage en connaissance est essentielle, c'est une exigence de redevabilité vis à vis du citoyen. Tout utile que soit une application d'IA, celle-ci ne doit pas entraîner une perte de compétence à court, moyen ou long terme. L'IA en milieu professionnel se conçoit comme un outil d'assistance et non d'assistanat. Il s'agit par exemple en ordre public de conserver la main sur la conception de manœuvre, comme en police judiciaire sur la capacité à qualifier une infraction. Même si elle est efficace, ce n'est pas à la machine de le faire. La performance d'un système d'IA ne signifie pas nécessairement l'efficacité de son usage. Il s'agit aussi de vérifier le niveau de compétence à l'exploitation des utilisateurs et de contrôler l'évolution du système au fur et à mesure de son utilisation.

### *L'audit éthique et réglementaire : transformer le risque de dérive en protection stratégique*

Les dérives éthiques comme réglementaires sont inévitables si elles ne sont pas appréhendées au plus tôt et en connaissance. Trop d'organisations enferment l'IA au sein des DSI (Direction des Systèmes d'Information) en négligeant la pluralité des exigences de l'IA, notamment éthique et réglementaire. Ce type d'audit doit garantir au citoyen un usage de l'IA qui contrôle les biais, qui prévient la discrimination, qui respecte les libertés individuelles tout en garantissant une protection collective de l'individu. L'efficacité opérationnelle ne doit pas se faire sans contrôle au détriment des droits fondamentaux. Un système d'IA mal configuré, mal utilisé, voire mal conçu, peut influencer directement des décisions impactant illégitimement le citoyen. L'automatisation peut en outre généraliser le mésusage et accroître considérablement l'étendue de l'impact. Cet audit a donc pour objectif de garantir le respect des principes éthiques fondamentaux et de s'assurer de l'application du règlement européen sur l'IA.

Ainsi, en sécurité intérieure, l'IA représente une rupture profonde qui ouvre des perspectives inédites d'efficacité et d'anticipation. Mais elle exige dans le même temps une responsabilité accrue. Dans un contexte où la défiance peut fragiliser l'action publique, seule une gouvernance éclairée, reposant sur une supervision fondée sur des audits rigoureux, permettra de profiter de la puissance de l'IA comme atout stratégique plutôt que la voir constituer une menace.

Loin d'être un frein, l'audit constitue le levier d'une innovation durable offrant la transparence nécessaire à la confiance, la robustesse indispensable à la résilience, et la légitimité sans laquelle aucune technologie ne peut s'ancrer dans le temps. L'avenir de l'IA en sécurité intérieure ne réside donc pas uniquement dans la performance technique, mais bien dans la capacité des institutions à l'intégrer en responsabilité, en prenant en compte sa singularité.

Superviser l'IA, c'est donc faire le choix stratégique de ne pas subir, de protéger sans aliéner, de renforcer sans affaiblir, et de placer la confiance citoyenne au cœur de la puissance publique. C'est à cette condition que l'IA deviendra non seulement un outil opérationnel, mais aussi et avant tout un véritable levier stratégique de sécurité et de souveraineté.

Document publié le 26/09/2025 – L'utilisation de tout ou partie de ce texte doit être soumise à autorisation ©CyberCercle