

Focus normalisation

DES PRODUITS SÛRS AVEC UNE INTELLIGENCE ARTIFICIELLE COMPLEXE ?

Le législateur est confronté au défi de définir des exigences pour des systèmes dont le comportement est imprévisible.

**IS PRODUCT SAFETY COMPATIBLE WITH COMPLEX ARTIFICIAL INTELLIGENCE ? –
For systems with unpredictable behaviour, defining requirements presents legislators
with a challenge.**

CORRADO
MATTIUZZO
KAN

Bien qu'il n'existe pas de définition généralement acceptée de l'intelligence artificielle (IA), il est bien clair que les différentes méthodes qu'elle met en œuvre ont pour but d'aider l'individu à prendre des décisions – voire de se charger de les prendre à sa place. Il reste à préciser, en revanche, dans quels cas ou dans quelles conditions on peut permettre que les décisions d'un système qui ont une incidence sur la sécurité soient influencées ou automatisées par les méthodes de l'intelligence artificielle.

Les risques inhérents à un produit doivent être évalués et réduits à un niveau acceptable avant la mise à disposition sur le marché. Les spécifications concernant le niveau élevé de protection à respecter se trouvent dans les directives et règlements du Marché intérieur européen. Les produits et équipements de travail qui ne relèvent pas du domaine harmonisé sont soumis à des réglementations nationales.

Selon la hiérarchie des mesures de protection, un produit doit être conçu de manière à exclure tout danger dès le départ. Lorsque cela n'est pas réalisable, des dispositifs de protection doivent réduire les risques à un point tel qu'il ne subsiste que des risques résiduels acceptables. Et enfin, l'utilisateur doit être informé de ces risques. Les dispositifs de commande, s'ils sont utilisés pour exécuter les fonctions de sécurité d'un produit, constituent l'un des éléments déterminants de ce concept.

Il est essentiel que les fabricants soient en mesure d'évaluer les risques susceptibles d'émaner de leurs produits. Or, cette condition ne serait pas remplie en voulant par exemple se fier à un dispositif de commande assisté par apprentissage automatique¹

pour empêcher que des personnes soient mises en danger par des pièces mobiles d'une machine : les concepteurs de systèmes qui reposent sur les méthodes complexes de l'intelligence artificielle (comme par exemple, l'apprentissage automatique avec réseaux neuronaux) ne sont pas, à ce jour, en mesure d'expliquer de manière satisfaisante, même rétrospectivement, pourquoi leur système s'est comporté de telle ou telle manière.

La technique de sécurité en terrain inconnu

Les bases et hypothèses techniques sur lesquelles repose la technique de sécurité traditionnelle n'ont pas été élaborées en vue d'être appliquées à des systèmes qui prennent des décisions ayant une incidence sur la sécurité, de manière automatisée et par le biais de méthodes complexes de l'intelligence artificielle. C'est pourquoi des études sont menées actuellement sur des méthodes d'évaluation. Il est prévu que les résultats en soient exploités immédiatement pour le travail de normalisation². L'objectif est de déterminer comment il est possible d'utiliser *a priori* l'IA, en relation avec des systèmes qui ont une incidence sur la sécurité.

Une approche permettant de démontrer de manière fiable la sécurité de systèmes très complexes consiste à définir des « arguments », destinés à fournir des indices « forts » déduits de manière inductive (mais non pas la preuve absolue). C'est l'approche adoptée depuis longtemps pour des technologies très complexes, notamment dans la technique nucléaire ou dans l'aérospatiale, mais aussi pour vérifier si un logiciel convient à une application ayant une incidence sur la sécurité.

KAN Brief
KOMMISSION ARBEITSSCHUTZ UND NORMUNG

Cet article est issu du Bulletin d'information KANBrief n° 1/21 (consultable sur : www.kan.de/fr) de la Kommission für Arbeitsschutz und Normung (KAN).
The English version of this article is accessible at www.kan.de/en.

On s'efforce actuellement d'utiliser ces approches – qui relèvent plutôt du « management du risque » – afin d'élaborer des catalogues de critères pour un niveau acceptable de risques, visant également des méthodes de l'intelligence artificielle. Ces catalogues de critères peuvent contenir des indications sur la spécification et la modélisation, l'« explicabilité » et la « compréhensibilité » des décisions, la « transposabilité » à différentes situations, la vérification et la validation du système, la surveillance pendant l'exécution, l'interaction homme – machine, la sécurisation et la certification des processus, ainsi que l'éthique et la protection des données liées à la sécurité. C'est également dans ce sens que va la demande du Parlement européen en faveur d'un règlement (UE) relatif aux principes éthiques pour le développement, le déploiement et l'utilisation de l'IA, de la robotique et des technologies connexes. Le Parlement y propose de tels critères pour l'évaluation de la conformité.

Une telle approche signifie que la sécurité n'est pas définie en premier lieu d'après des caractéristiques produits vérifiables, mais par des critères de processus vérifiables. Pour se rapprocher d'un niveau élevé de sécurité allant dans le sens des réglementations européennes sur la sécurité des produits et de l'idée fondamentale de la prévention au travail, il faudrait d'abord prouver que les critères des « arguments » évoqués plus haut sont complets et fiables. C'est pourquoi, à proprement parler, les prescriptions qui définissent pour cela le cadre et les exigences essentielles ne peuvent, elles aussi, être établies qu'une fois que les hypothèses

sur lesquelles elles se basent ont fait leurs preuves de façon fiable.

Les premières approches réglementaires

La norme ISO/TR 22100-5³, qui vient d'être publiée, tente de déterminer les limites à l'intérieur desquelles, en vertu de l'état actuel de la législation et de la normalisation, l'apprentissage automatique pourrait être intégré dans un système de commande de machine. La Commission européenne présente actuellement une proposition concernant non seulement la révision de la directive « Machines » n° 2006/42/CE, mais aussi un règlement sur l'IA qui contiendraient des cadres juridiquement contraignants pour l'utilisation de l'intelligence artificielle. Ces cadres doivent contenir des exigences complètes, claires et vérifiables sur les cas et les conditions dans lesquelles il est permis que des décisions d'un système ayant une incidence sur la sécurité soient influencées ou automatisées par des méthodes de l'intelligence artificielle. C'est maintenant aux experts qu'il revient d'examiner si ces critères sont respectés. ●

1. L'apprentissage automatique consiste à donner aux ordinateurs la capacité d'apprendre à effectuer une tâche à partir de données, et non pas en étant explicitement programmés ou entraînés à partir de règles compréhensibles par l'être humain.
2. Par exemple, le projet ISO/TR 5469 « Artificial intelligence – Functional safety and AI systems » (dans l'ISO/IEC JTC 1/SC 42/WG 3).
3. ISO/TR 22100-5:2021 – Sécurité des machines – En relation avec l'ISO 12100 – Partie 5 : Implications de l'intelligence artificielle pour l'apprentissage automatique.



Application permettant à des chauffeurs-livriers de disposer d'informations à l'approche d'une entreprise cliente.

© Grégoire Maisonneuve pour l'INRS/2021