

# L'intégration de l'IA dans la protection de la population

Qu'il s'agisse de prévoir la propagation mondiale du nouveau coronavirus originaire de Chine ou de détecter les incendies de forêt en Californie, l'utilisation de l'intelligence artificielle dans la protection de la population doit permettre d'améliorer la prévention des catastrophes, la réponse qui leur est apportée ainsi que le rétablissement après une catastrophe.

Par Kevin Kohler et Benjamin Scharte

L'intelligence artificielle (IA) est un terme générique qui renvoie à un domaine de recherche et d'applications. Aujourd'hui, le terme est essentiellement utilisé pour désigner le sous-domaine de l'apprentissage automatique (aussi appelé «apprentissage machine»), qui décrit un ensemble d'algorithmes reposant sur l'apprentissage statistique. Ces dernières années, les progrès en matière de puissance de calcul, de disponibilité des données et des algorithmes évolutifs ont conduit à une série d'avancées. L'IA est ainsi devenue une technologie habilitante ayant des applications dans tous les secteurs. La présente analyse se concentre sur l'utilisation de l'IA dans la protection de la population. La protection de la population recouvre la gestion de toute la gamme des risques collectifs découlant des dangers environnementaux, sociétaux et techniques, tels que les phénomènes météorologiques extrêmes, les pandémies, les attentats terroristes ou les accidents industriels graves, afin de limiter les dommages, de protéger la population et de préserver ses moyens de subsistance.

Bien appliquée, l'IA peut aider à réduire les pertes humaines et économiques causées par les catastrophes et s'intègre ainsi dans le paradigme «AI for Good» (l'IA au service d'un monde meilleur), qui vise à tirer parti de l'intelligence artificielle pour atteindre les objectifs de développement du-



Dans certains pays, les forces de police utilisent des véhicules aériens sans pilote afin d'assurer le respect des mesures de sécurité prises afin de prévenir la propagation du coronavirus. *Irakli Gedenidze / Reuters*

nable des Nations Unies. En septembre 2020, l'Union internationale des télécommunications devrait ainsi accueillir à Genève le quatrième sommet mondial «AI for Good», un sommet soutenu par le gouvernement suisse. Dans le même temps, le fait que de mauvaises décisions en matière de protection de la population puissent coûter des vies accentuent le besoin de recourir à des outils algorithmiques assurant une aide à la décisions et une prise de décision robustes, sûres, équitables et interprétables.

## Un large spectre d'applications

Compte tenu de l'étendue des missions de protection de la population et du caractère habilitant de l'IA, il existe des utilisations potentielles de l'IA pour de nombreux risques et ce à toutes les étapes du cycle de gestion des catastrophes. Par exemple, des modèles d'IA sont utilisés pour l'analyse dynamique des risques, ce qui permet d'aider à identifier les vulnérabilités, à détecter les dangers à un stade précoce et à en prévoir l'évolution. Ainsi, des prévisions locales

des inondations et des glissements de terrain sont régulièrement mises à jour permettant de diffuser à temps des alertes de sécurité pour les zones concernées. L'IA peut également être appliquée à la gestion des infrastructures critiques. Les réseaux électriques intelligents («smart grids») intègrent notamment des applications permettant d'assurer un contrôle et une optimisation distribués dans les microréseaux, de catégoriser les pannes de réseau selon leur type et leur gravité et de prévoir la demande, le prix de l'électricité ainsi que la production d'électricité des installations photovoltaïques et éoliennes.

En outre, l'analyse vidéo basée sur l'IA, associée à l'omniprésence des caméras, accroît rapidement la lisibilité des foules et des espaces publics. Si ces aspects soulèvent de graves préoccupations en matière de pro-

## Dans les sociétés démocratiques, tout programme de surveillance devra s'accompagner d'un débat politique sur la protection des données et les droits civils.

tection de la vie privée, une meilleure connaissance de la situation peut aider à gérer les menaces terroristes, les pandémies et les migrations de masse.

L'IA peut aussi soutenir les efforts de recherche et de sauvetage après une catastrophe. Par exemple, des systèmes sans pilote utilisant l'IA peuvent fournir en toute sécurité des informations très précises sur leurs environnements opérationnels. Dans le même temps, les progrès réalisés dans le traitement automatique du langage naturel offrent des capacités croissantes de traduction en temps réel susceptible de s'avérer utiles dans les situations impliquant des minorités linguistiques et des touristes ou dans les missions d'aide internationale.

La présente analyse n'a pas vocation à rendre compte des utilisations de l'IA dans toute la gamme de risques. Elle s'appuiera essentiellement sur deux exemples actuels: la pandémie mondiale de coronavirus et les incendies dévastateurs qui ont récemment sévi en Californie et en Australie.

### La surveillance des maladies

Les systèmes hybrides qui s'appuient à la fois sur des données traditionnelles et sur le traitement automatique du langage naturel pour analyser les signaux médiatiques sont les plus prometteurs. Des entreprises

### L'IA à l'Office fédéral de la protection de la population (OFPP)

Les laboratoires de microbiologie ont été parmi les premiers à intégrer l'IA dans leurs flux de travail en appliquant des algorithmes à des tâches telles que la détection chromogénique ou le comptage de colonies. Aujourd'hui, les laboratoires doivent fournir des résultats rapides avec des ressources limitées et des défis croissants pour la société dans son ensemble, tels que la mobilité accrue des nouveaux agents pathogènes et les bactéries résistantes aux antibiotiques. Le laboratoire Spiez, l'institut fédéral pour la protection ABC au sein de l'OFPP, utilise actuellement l'IA dans certains domaines d'analyse, de détection et de diagnostic tels que la sélection de biomarqueurs dans les spectres de masse, la différenciation et l'identification d'agents pathogènes ou l'identification d'espèces vectrices dans les projets de surveillance. À moyen terme, l'OFPP pourrait recourir à de nouvelles méthodes de traitement de données pour renforcer l'efficacité des services de conseil, optimiser la prise des appels d'urgence pendant les crises ou améliorer la fiabilité et les délais de réponse des systèmes d'alerte.

comme BlueDot et Metabiota, ou encore HealthMap au Boston Children's Hospital, proposent une surveillance mondiale des maladies infectieuses associant un filtrage des nouveaux articles par l'IA, l'analyse des schémas de déplacement et des évaluations d'experts. Toutes ont pu avertir leurs clients de l'apparition d'un nouveau virus à Wuhan avant le Centre des États-Unis pour le contrôle et la prévention des maladies (CDC) et l'Organisation mondiale de la Santé (OMS). Cependant, les rumeurs qui circulaient dans le milieu médical chinois ont aussi gagné l'Occident par le biais des communications personnelles et il a encore fallu recourir au jugement humain pour évaluer la gravité des alertes automatisées.

La crise actuelle du coronavirus a également donné lieu à un grand nombre de nouvelles applications de l'IA dont l'efficacité reste difficile à déterminer à ce stade. Des systèmes d'IA sont par exemple utilisés pour le triage et le diagnostic. Des entreprises comme Infervision, YITU et AliBaba ont mis au point des systèmes d'IA pour diagnostiquer la maladie à coronavirus (COVID-19) par tomographie et assister ainsi le personnel médical surchargé. Babylon Health, qui travaille avec le National Health Service britannique, et beaucoup d'autres start-up ont intégré les critères de diagnostic du COVID-19 dans leurs chatbots médicaux. Toutefois, du fait du manque de données d'entraînement précises sur le COVID-19, ces systèmes d'analyse de symptômes s'appuient actuellement sur des règles établies à la main, et non sur l'apprentissage automatique. L'IA est également déployée dans la recherche de traitements. Deepmind a publié des prédictions informatiques des structures protéiques liées au nouveau coronavirus pour

aider les chercheurs à mieux le comprendre. SRI International et Itkos ont annoncé l'établissement d'un partenariat visant à utiliser une technologie de modélisation générative basée sur l'IA pour la découverte automatique de médicaments antiviraux.

Ces derniers temps, des pays comme la Chine, la Russie et Israël se sont tournés vers leurs systèmes nationaux de surveillance pour faire appliquer les consignes de distanciation physique et pour suivre les cas de COVID-19 suspects ou confirmés. Les sociétés chinoises de reconnaissance faciale, telles que SenseTime et Megvii, ont adapté leurs systèmes d'IA pour identifier les personnes malgré leur masque et pour détecter celles qui n'en portent pas ou qui ont de la température. Dans certaines villes chinoises, les forces de police portent même des «casques intelligents» qui repèrent les personnes ayant de la fièvre et l'application Health Code utilise l'historique de géolocalisation des téléphones mobiles pour attribuer aux individus un code de risque vert, jaune ou rouge qui détermine s'ils sont autorisés à entrer dans les bâtiments et les espaces publics ou obligés de rester en quarantaine. Compte tenu des coûts économiques faramineux et de la restriction drastique des libertés individuelles associés à un confinement sans discernement, un nombre croissant de sociétés envisageront probablement de recourir à des dispositifs automatisés pour surveiller de manière fine de larges pans de la population. De telles mesures peuvent toutefois s'avérer sujettes à controverse car elles constituent aussi des moyens de contrôle social que les gouvernements pourraient être tentés de prolonger au-delà de l'état d'exception initial. Dans les sociétés démocratiques, tout programme de ce type doit donc s'accompagner d'un débat politique sur la vie privée, la protection des données et les droits civils.

## Maîtriser les incendies

Dans le contexte des feux de forêt, l'analyse d'images satellites, des véhicules aériens sans pilote ou des webcams peuvent aider à détecter automatiquement les incendies et réduire ainsi les délais de réaction, ce qui augmente les chances de pouvoir éteindre le foyer. Les pompiers s'appuient également sur des outils de modélisation tels que FARSITE pour prévoir l'évolution des feux de forêt en fonction de la topographie, de la météo et du combustible. Ces dispositifs peuvent les aider à planifier des incendies maîtrisés et à créer des coupe-feux sans combustible pour limiter l'ampleur des incendies non planifiés. En Californie, le projet WIFIRE mené par la Fondation nationale pour la science (NSF) et le San Diego Supercomputer Center a mis au point Firemap, une plateforme web qui s'appuie sur les données pour simuler, prévoir et visualiser en temps réel le comportement des feux de forêt. Les pompiers de Los Angeles et plus d'une centaine de services de secours locaux ont recours à ce système. Il est important de noter que l'apprentissage automatique n'est pas la seule méthode utilisée par cet ensemble d'outils numériques. Ainsi, la NASA et le Système européen d'information sur les feux de forêts (EFFIS) s'appuient tous deux sur des algorithmes faits à la main pour la détection active des incendies à partir de données satellitaires. L'IA est cependant utilisée pour catégoriser la couverture terrestre, pour détecter la présence de fumée et le périmètre du feu sur les webcams, pour prévoir l'accumulation de combustible et les conditions météorologiques dangereuses et, de façon plus expérimentale, pour analyser les communications radio et les publications sur les réseaux sociaux.

## Des risques accentués

Si l'IA offre des possibilités, elle crée ou accentue également certains risques pour la société. Les risques généraux liés aux applications de l'IA incluent les questions relatives à la confidentialité de données, la perpétuation des biais raciaux et de genre contenus dans les jeux de données d'entraînement, la robustesse et la sécurité des classificateurs, le caractère interprétable et vérifiable de la logique décisionnelle de l'IA, ainsi que la responsabilité et la redevabilité du processus de prise de décision fondé sur des algorithmes. L'utilisation de l'IA dans la protection de la population recèle les mêmes risques, souvent intensifiés par l'environnement décisionnel. Les équipes de recherche et de sauvetage, par exemple, opèrent dans un environnement très stressant et à forts enjeux. Elles ont donc besoin

de technologies fiables et éprouvées. De nombreux systèmes sans pilote utilisant l'IA n'ont pas encore atteint le niveau technologique requis pour fonctionner correctement dans des situations de communication dégradées et des conditions météorologiques extrêmes. Par conséquent, il convient de soumettre ces associations homme-machine à des tests de résistance simulant des catastrophes.

D'autre part, l'utilisation de l'IA dans le secteur public appelle à des exigences accrues en matière de transparence, de prévention des discriminations et d'intégrité, la légitimité des gouvernements reposant sur le fait qu'ils servent le bien public. La plupart des applications de l'IA dans le domaine de la protection de la population sont plutôt au service des professionnels que du grand public. Néanmoins, lorsque les processus publics d'aide à la décision et de prise de décision fondés sur des algorithmes ont des conséquences majeures pour les personnes, il est important de fournir des informations compréhensibles sur leur logique décisionnelle.

En outre, les infrastructures critiques doivent être résilientes face aux situations exceptionnelles. Comme le souligne le rapport du groupe de travail interdépartemental sur l'IA (GTI IA) publié en décembre 2019, la complexité et l'opacité croissantes des systèmes d'IA menacent également de réduire les compétences et les connaissances du personnel travaillant dans les infrastructures critiques. Le recours à l'IA pour surveiller et réguler des systèmes tels que le réseau électrique, notamment, place les humains au rang d'acteurs plus passifs. Le personnel humain devra toutefois conserver une compréhension globale du système pour pouvoir le manipuler activement et réagir de façon compétente aux dysfonctionnements, ce qui peut s'avérer particulièrement utile dans le scénario peu probable, mais non négligeable, d'une cyberattaque à grande échelle contre des infrastructures critiques.

## Les défis de la mise en œuvre

L'IA n'est pas une panacée. La vague actuelle d'IA repose sur l'apprentissage statistique à partir de grands ensembles de données. Malgré ses nombreuses applications passionnantes, l'IA possède des capacités extrêmement limitées à apprendre des concepts abstraits à partir d'un petit nombre d'exemples, à comprendre les causalités, à transposer les éléments appris d'un domaine à un autre, à gérer les structures hiérarchiques et à faire preuve d'un

## Pour aller plus loin

Bullock, J., Luccioni, A., Pham, K., Lam, C., & Luengo-Oroz, M. (2020). *Mapping the Landscape of Artificial Intelligence Applications against COVID-19*. <https://arxiv.org/pdf/2003.11336.pdf>

Commission européenne (2020), *Livre blanc: Intelligence artificielle – Une approche européenne axée sur l'excellence et la confiance*, COM(2020) 65 final.

GFDRL (2018). *Machine Learning for Disaster Risk Management*. Washington, DC: GFDRL.

GTI IA (2019), *Défis de l'intelligence artificielle*, Rapport du groupe de travail interdépartemental «Intelligence artificielle» au Conseil fédéral.

Simonsen, L., Gog, J., Olson, D. et Viboud, C. (2016), Infectious Disease Surveillance in the Big Data Era: Towards Faster and Locally Relevant Systems, *The Journal of Infectious Diseases*, 214(4), p. 380–385.

Stanley, J. (2019), *The Dawn of Robot Surveillance AI, Video Analytics, and Privacy*, Union américaine pour les libertés civiles.

raisonnement de bon sens. En l'absence de données historiques ou d'environnement d'entraînement virtuel approprié, on ne peut pas s'attendre à ce que les systèmes d'IA soient performants. Par exemple, la précision et la résolution des prévisions de la demande vis-à-vis du réseau électrique dépendent de la disponibilité de données détaillées issues des compteurs intelligents. Ce manque de données d'entraînement est particulièrement prononcé pour les événements à faible probabilité et à fort impact, ainsi que pour les risques liés aux technologies émergentes.

L'agrégation et la standardisation des données constituent un autre grand défi. Les données concernant les dangers ou les services de protection de la population sont souvent conservées en silos. Pour créer des jeux de données d'entraînement d'une qualité et d'un volume suffisants, les données doivent être partagées et étiquetées de manière standardisée. Afin de surmonter le problème des structures de données compartimentées, la Western Fire Chiefs Association aux États-Unis, par exemple, fournit à 29 corps de sapeurs-pompiers desservant plus de 20 millions de personnes une plateforme logicielle permettant de réaliser des analyses de performance au sein d'un cadre commun avec un ensemble de données centralisé et normalisé, associé à des flux de travail clairs. L'ensemble de

données intégré est anonymisé et partagé sur le cloud via le Fire Data Lab.

Enfin, les compétences dans les domaines de la science des données et de l'apprentissage automatique sont insuffisantes. Les talents en matière d'IA sont concentrés dans les grandes entreprises technologiques et les universités, tandis que les gouvernements n'ont pas assez de capacité en interne pour mettre au point leurs propres modèles d'IA. Le secteur public doit donc collaborer avec les universités et le secteur privé, ce qui peut impliquer de mettre les ensembles de données concernés à la disposition de

## Les défis liés à la disponibilité des données et à l'expertise du secteur public sont particulièrement forts dans les structures politiques décentralisées.

ces acteurs. Compte tenu des exigences accrues en matière de robustesse, de transparence et de résilience, les solutions génériques ne constituent pas toujours une option réaliste. Et même si c'est le cas, la procédure de passation de marché peut nécessiter d'être en mesure de tester et de comprendre les solutions du commerce. Aux États-Unis, par exemple, le National Institute of Standards and Technology (NIST) a réalisé des tests pour évaluer différents fournisseurs de systèmes de reconnaissance faciale.

### La Suisse doit unir ses forces

Les défis liés à la disponibilité des données et à l'expertise du secteur public en matière d'IA sont particulièrement forts dans les structures politiques décentralisées. En Suisse, la protection de la population est répartie entre les organisations partenaires de la police, les sapeurs-pompiers, les services de santé, les opérateurs d'infrastructures critiques et la protection civile. Comme en Allemagne et en Autriche, la gestion des catastrophes en temps de paix incombe en grande partie aux niveaux infranationaux. Cependant, le gouvernement fédéral est responsable de la recherche et du développement. Il peut coordonner, en accord avec les cantons, voire diriger la gestion des catastrophes, si nécessaire.

Le rapport du GTI IA souligne la nécessité d'identifier les processus dans l'ensemble de l'administration fédérale et de garantir un accès transversal aux données. Dans le contexte de la protection de la population, il est sans doute encore plus important de constituer des ensembles de données partagés et normalisés entre les organisations partenaires au niveau local, les cantons, le gouvernement fédéral et les partenaires internationaux.

Le manque de compétences techniques dans l'administration publique peut être comblé par la mutualisation des expertises, une collaboration étroite avec les universités et l'instauration de partenariats public-privé. Le rapport du GTI IA préconise notamment d'étudier la possibilité de créer un réseau de compétences pour l'application de l'IA au sein de l'administration fédérale. Dans la mesure où l'essentiel de la recherche et du développement associée à la protection de la population se déroule dans les universités, cette coopération est également cruciale pour tester et élargir les applications dans des domaines prometteurs. Lorsqu'une application arrive à maturité, les partenariats public-privé représentent une bonne alternative aux produits du commerce.

Enfin, la sûreté et la sécurité sont plus importantes que la vitesse de déploiement dans le contexte des infrastructures critiques. Selon le livre blanc de la Commission européenne sur l'IA, un cadre réglementaire doit inclure des normes obligatoires pour les applications à haut risque en ce qui concerne les données d'entraînement, la conservation des données et des dossiers, les informations à fournir, la robustesse et la précision, ainsi que le contrôle humain. Il reste à voir si la Suisse envisagera une réglementation similaire pour les applications sensibles de l'IA dans le domaine de la protection de la population.

### La mise en commun des données

À l'échelle internationale, l'accent a été mis, ces dernières décennies, sur la mise en place de structures comparables et la création d'un ensemble mondial de données sur les impacts des catastrophes. Le Cadre de

Sendai pour la réduction des risques de catastrophe 2015–2030 prévoit que les pays rendent compte des pertes économiques et humaines, des stratégies et des aides au développement liées aux catastrophes. La Suisse soutient activement ces efforts en tant que donateur volontaire du Bureau des Nations Unies pour la réduction des risques de catastrophe, et la surveillance définie dans le Cadre de Sendai est importante pour suivre les progrès et inciter les États à agir. Les gouvernements ont pourtant du mal à exploiter avantageusement ces données pour prévoir l'évolution des risques ou déterminer les modalités de préparation et d'intervention les plus efficaces.

La pandémie mondiale de coronavirus met en évidence, une fois de plus, l'intérêt du partage de données et de la coopération scientifique au niveau international pour protéger les populations. Les initiatives de coopération ponctuelles, telles que l'ensemble de données de recherche ouvert sur le COVID-19 (CORD-19), peuvent faire la différence. Mais il est tout aussi important de constituer systématiquement des ensembles de données sur les risques à l'échelle mondiale avant la prochaine catastrophe. Le programme post-Sendai de réduction des risques de catastrophe doit donc donner la priorité au partage de données sur les risques en eux-mêmes, ainsi qu'aux politiques publiques destinées à répondre à ces risques. L'initiative de la communauté «AI for Good» visant à conceptualiser, constituer et mettre en commun des données au niveau mondial pourrait être le vecteur de ces efforts.

Voir le [site thématique du CSS](#) pour en savoir plus sur la résilience sociotechnique et la préparation aux catastrophes.

**Kevin Kohler** est chercheur au sein de l'équipe «Risk and Resilience» du Center for Security Studies (CSS) à l'ETH de Zurich.

**Benjamin Scharte** est chef de l'équipe «Risk and Resilience» au Center for Security Studies (CSS).

Les analyses de politique de sécurité du CSS sont publiées par le Center for Security Studies (CSS) de l'ETH Zurich. Deux analyses paraissent chaque mois en allemand, français et anglais. Le CSS est un centre de compétence en matière de politique de sécurité suisse et internationale.

Editeurs: Julian Kamasa, Fabien Merz, Lisa Watanabe, Benno Zogg  
Traduction: Interserv; Relecture: Fabien Merz  
Layout et graphiques: Miriam Dahinden-Ganzoni  
ISSN: 2296-0228; DOI: 10.3929/ethz-b-000408270

Feedback et commentaires: [analysen@sipo.gess.ethz.ch](mailto:analysen@sipo.gess.ethz.ch)  
Téléchargement et abonnement: [www.css.ethz.ch/cssanalysen](http://www.css.ethz.ch/cssanalysen)

Parus précédemment:

**Ukraine: la dimension religieuse du conflit** No 259  
**Le piège colombien: une autre paix partielle** No 258  
**Les opérations militaires en milieu urbain** No 257  
**Sécurité spatiale: la prochaine décennie** No 256  
**Les nouvelles technologies et les frontières en Europe** No 255  
**Le Moyen-Orient et la nouvelle route de la soie** No 254