# L'IA POUR LE RENSEIGNEMENT : SON APPORT ET LA STRICTE RÉGULATION NÉCESSAIRE


**En septembre dernier, le *Special Competitive Studies Project* (SCSP) a publié un rapport très intéressant sur l'impact de l'IA dans les techniques de renseignement : *« The Digital case officer »*.**

**Il s'agit certes d'une organisation américaine mais ses appels à vigilance restent néanmoins toujours d'actualité car ils nous renseignent sur les possibles utilisations de cette technologie et les inquiétudes (juridiques, éthiques et même politiques) nourries auprès de la population américaine.**

Pour évoluer dans ce nouveau paysage, la communauté du renseignement américaine (IC) doit maîtriser les technologies mêmes qui la menacent. L'IA avancée, en particulier les systèmes agents capables d'agir de manière autonome, peut désormais renforcer chaque phase du cycle de recrutement des agents de renseignement humains avec une ampleur et une précision inaccessibles aux seuls humains.

Ces « agents de renseignement numériques » peuvent :
- Cibler : synthétiser de vastes ensembles de données pour identifier et hiérarchiser les agents potentiels en fonction de leur accès, de leur motivation et de leur vulnérabilité.
- Évaluer et développer : établir des profils psychologiques détaillés à partir d'empreintes numériques et engager des conversations personnalisées et à long terme avec les cibles afin d'établir une relation et une confiance, en utilisant des personnages hyperréalistes. L'IA peut gérer simultanément des centaines de conversations de ce type, une tâche impossible pour un agent humain.
- Recruter et gérer : proposer des arguments de recrutement personnalisés en se référant aux griefs ou motivations spécifiques d'une cible et fournir des conseils de sécurité opérationnelle en temps réel aux agents une fois qu'ils ont été recrutés.

L'objectif n'est pas de remplacer les agents humains, mais de leur donner plus de moyens. L'avenir du renseignement humain réside dans l'équipe homme-machine, où l'IA gère l'énorme volume de traitement des données et la prise de contact initiale, libérant ainsi les agents chargés des dossiers pour qu'ils puissent se concentrer sur des tâches à forte valeur ajoutée telles que la prise de décisions nuancées, la gestion de la relation psychologique entre l'agent et son informateur, et la supervision d'opérations à haut risque. L'intelligence artificielle peut également aider les collecteurs humains à atteindre une échelle sans précédent en repérant un plus grand nombre de cibles, en développant simultanément leurs profils et en menant des approches virtuelles.

La puissance de l'IA dans le domaine du renseignement humain soulève également d'importantes questions juridiques et éthiques qui exigent un cadre de gouvernance solide. L'utilisation de systèmes autonomes présente un paysage complexe de questions juridiques et éthiques que la communauté du renseignement devra aborder, notamment la manière dont elle protégera la vie privée des citoyens américains dont les informations pourraient se trouver dans l'ensemble de données utilisées pour entraîner les modèles d'IA, la manière de garantir la transparence et la responsabilité des actions et des recommandations de l'IA, et les garde-fous nécessaires pour s'assurer que les IA du renseignement ne poursuivent pas des stratégies de manipulation amorales qu'un agent humain rejetterait. Ces préoccupations peuvent être quelque peu apaisées par les progrès continus dans la sophistication des modèles d'IA. Néanmoins, la pierre angulaire d'une solution viable est le principe du contrôle humain significatif (MHC). À chaque moment critique, en particulier la décision finale de recruter, la mission d'un agent ou les actions qui présentent un risque important pour l'agent ou les intérêts de la sécurité nationale américaine, un humain responsable doit être en mesure d'exercer un jugement final.

**Synthèse des recommandations**

Pour relever les défis d'une nouvelle ère technologique et conserver un avantage décisif en matière de renseignement, les États-Unis doivent agir avec détermination et urgence. Pour y parvenir, il ne suffit pas de disposer de nouveaux outils, il faut également repenser en profondeur le rôle des agents de terrain, en amplifiant le jugement humain grâce à l'intelligence artificielle. Ce rapport fournit un cadre stratégique aux décideurs politiques de haut niveau et aux dirigeants de la communauté du renseignement afin de guider la transformation de l'entreprise HUMINT nationale.

Voici quelques-unes des mesures clefs proposées :

- Établir des lignes directrices claires sur l'utilisation de l'IA pour le renseignement humain : le directeur de la CIA, en coordination avec le directeur du renseignement national, le ministère de la Justice et les autres responsables des agences de la communauté du renseignement, devrait publier des directives officielles sur l'utilisation de l'IA pour le renseignement humain. Ces directives devraient définir les cas d'utilisation approuvés, les vérifications et autorisations requises, ainsi que les protocoles de coordination interagences. Elle pourrait par exemple exiger que tout déploiement opérationnel d'une IA, tel que le *Digital Case Officer*, soit approuvé par les hauts responsables et documenté de la même manière que les autres activités de collecte sensibles. Elle devrait également établir des limites claires, interdisant explicitement l'utilisation de l'IA pour le recrutement de catégories spécifiques d'individus (par exemple, des ressortissants américains ou des fonctionnaires alliés) sans autorisation explicite d'un niveau supérieur.

- Mettre en œuvre des exigences en matière de contrôle humain : les agences de la communauté du renseignement devraient codifier les points de contrôle *« human-in-the-loop »* qui se sont imposés comme les meilleures pratiques en matière de développement. Par exemple, elles pourraient exiger qu'aucune IA ne puisse recruter entièrement une source humaine sans l'accord d'un responsable humain de haut niveau et la signature d'un responsable du contre-espionnage numérique (CI). Elle pourrait également exiger que toute tâche confiée à une source par l'IA au-delà d'un certain seuil de risque (par exemple, demander à la source d'entreprendre une activité physique ou d'accéder à des informations classifiées) soit approuvée par un contrôleur humain.

- Développer un cadre interagences *« HUMINT-As-A-Service »* : la communauté du renseignement, sous l'impulsion de la CIA, devrait envisager d'offrir le service d'agent de liaison numérique à d'autres agences et éventuellement à des partenaires étrangers proches. Dans ce modèle, la CIA conserverait la propriété du système tout en le déployant pour le compte d'autres entités selon des conditions convenues. Un cadre politique complet devrait délimiter le processus de demande de soutien opérationnel de l'IA, définir des protocoles pour le partage des données et des résultats, et établir des responsabilités de surveillance claires lorsque, par exemple, le FBI utilise un outil développé par la CIA. Ce cadre garantirait également la mise en place des autorités compétentes, en fonction de la nature de l'opération. Essentiellement, la capacité d'IA devrait être traitée comme un atout national, pouvant être prêté selon des règles définies, plutôt que chaque agence ne développe des systèmes d'IA redondants avec des normes disparates. La CIA, en tant qu'agent exécutif de la communauté du renseignement pour les relations de renseignement étranger,

devrait prendre l'initiative de déterminer quelles technologies et plateformes la communauté du renseignement américaine devrait partager de manière proactive avec les partenaires des *Five Eyes* et d'autres services amis.

- Red Teaming continu : avant le lancement de toute nouvelle version ou fonctionnalité majeure, soumettre l'IA à une nouvelle attaque de *red team* (lire à ce sujet l'article de la CNIL : https://www.cnil.fr/fr/definition/red-teaming)
- En outre, prévoir des audits périodiques de *red team* (peut-être tous les six mois), même en l'absence de changements majeurs, afin de détecter les problèmes émergents ou les dérives de performance. La communauté du renseignement devrait faire appel à des experts externes (universitaires, hackers éthiques) sous contrat afin de garantir des perspectives nouvelles. Les conclusions devraient être communiquées à la fois aux développeurs et aux autorités de contrôle. Cela permettra d'institutionnaliser les tests adversaires, garantissant ainsi que l'IA ne reste jamais statique face à l'évolution des menaces.

Intégrer l'IA dans la formation aux techniques d'espionnage : la CIA, le FBI et la DIA devraient modifier les programmes de formation existants aux techniques d'espionnage HUMINT et intégrer des cursus axés sur la création, l'intégration et la gestion des agents de terrain numériques dans les opérations.

- Responsabilité, supervision et rapports : désigner officiellement un « responsable » pour chaque opération d'IA. La communauté du renseignement devrait désigner un responsable de la surveillance de l'IA ou un point de contact dans chaque agence concernée, qui aurait accès aux dossiers de l'IA et pourrait effectuer des inspections ou des examens inopinés. S'engager à fournir des mises à jour périodiques aux organismes de surveillance, en particulier sur l'utilisation de l'IA dans les opérations. Il pourrait s'agir d'une annexe annuelle aux rapports existants. Elle devrait inclure des mesures anonymisées telles que le nombre d'opérations, les types généraux de cibles, les problèmes de conformité éventuels et les résultats.

Nicolas CORSI
Membre de l'AASSDN