

Tout le monde ajoute de l'IA. Peu de gens se demandent ce qu'elle ajoute à leur risque.

Maintenance prédictive. Inspection automatisée. Aide à la décision pour la logistique et le renseignement. L'IA s'implante rapidement dans la défense et l'aérospatiale, souvent portée par la feuille de route d'un fournisseur plutôt que par une décision délibérée de l'organisation qui l'utilise. Rares sont ces systèmes livrés avec une réponse claire à une question simple : que se passe-t-il si quelqu'un lui fournit délibérément des données erronées.

Tournez la page pour découvrir le panneau stop qui a trompé un ordinateur



La véritable exposition

Là où l'IA t'ouche à l'aide à la décision ou aux systèmes autonomes, ce n'est pas un détail secondaire. C'est un risque qu'il faut évaluer avant qu'une capacité ne soit validée, pas après.

01

Les attaques adverses contre l'apprentissage automatique ne

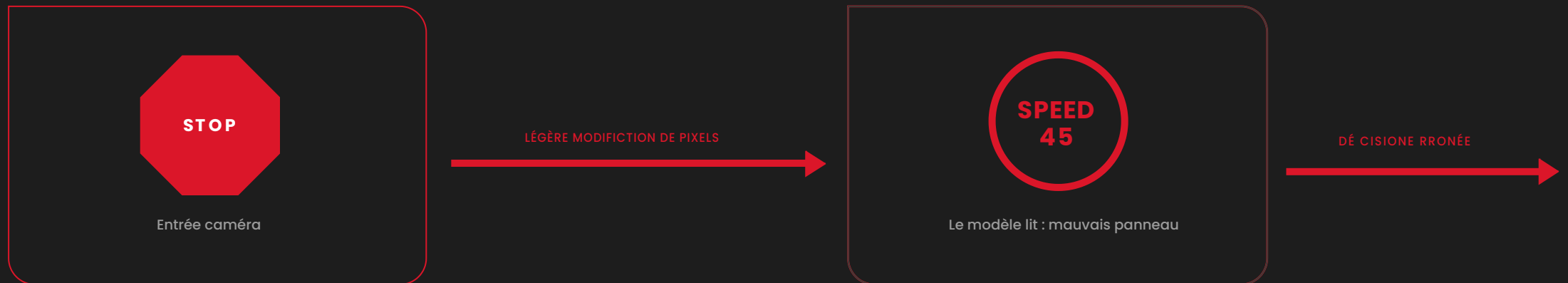
Il a été démontré qu'une modification minimale et délibérée d'une image ou d'un flux de capteur peut tromper des systèmes de classification, d'une manière qu'un observateur humain ne manquerait jamais.

2018

En 2018, des chercheurs de l'université du Michigan et de l'université de Washington ont montré que quelques petits autocollants placés sur un panneau stop pouvaient amener un système de vision par ordinateur à le lire comme un panneau de limitation de vitesse, de façon fiable, sous plusieurs angles et à plusieurs distances. Rien dans l'apparence du panneau n'aurait alerté un passant.

02

Le type de défaillance qu'une revue de sécurité classique ne détectera pas.



Pourquoi c'est important au salon AAD

Les organisations de la défense et de l'aérospatiale sont confrontées à une version plus exigeante d'un problème que tous les secteurs rencontrent désormais : les capacités d'IA arrivent plus vite que la gouvernance nécessaire pour les encadrer, et dans ce secteur, les conséquences d'une erreur sont plus lourdes qu'ailleurs. Nous aidons les organisations à construire la couche d'évaluation et de gouvernance avant la mise en service, et non après coup, une fois qu'un problème s'est déjà produit.

03

Ce que nous faisons

Notre pratique en IA s'inscrit aux côtés de nos activités de cybersécurité et de gouvernance plutôt qu'à part, car dans la pratique, les deux ne peuvent pas vraiment être dissociées.



Évaluations des risques et de la gouvernance liés à l'IA, couvrant la provenance des données, l'intégrité des modèles et la responsabilité des décisions automatisées.



Tests adverses des systèmes d'apprentissage automatique, à la recherche de modes de défaillance qu'une revue de sécurité classique ne détecterait pas.



Feuilles de route pour une adoption sécurisée de l'IA, destinées aux organisations qui évaluent des outils d'IA fournisseurs avant leur mise en production.



Cadres de politique et de gouvernance pour un usage responsable de l'IA, alignés sur les normes internationales émergentes.

**Les organisations qui se font
surprendre par l'IA sont
rarement celles qui posent trop
de questions avant de l'adopter.
Ce sont généralement celles qui
ont supposé que le fournisseur
s'était déjà posé ces questions à
leur place.**

DISystems – Conseil en cybersécurité et GRC en
Afrique du Sud, dans la région de la SADC et à
l'international, en partenariat avec **Cyberantix™**

CONTACTEZ-NOUS À L'ADRESSE :
tsepang@disystems.co.za
pour convenir d'un rendez-vous et en savoir plus.

