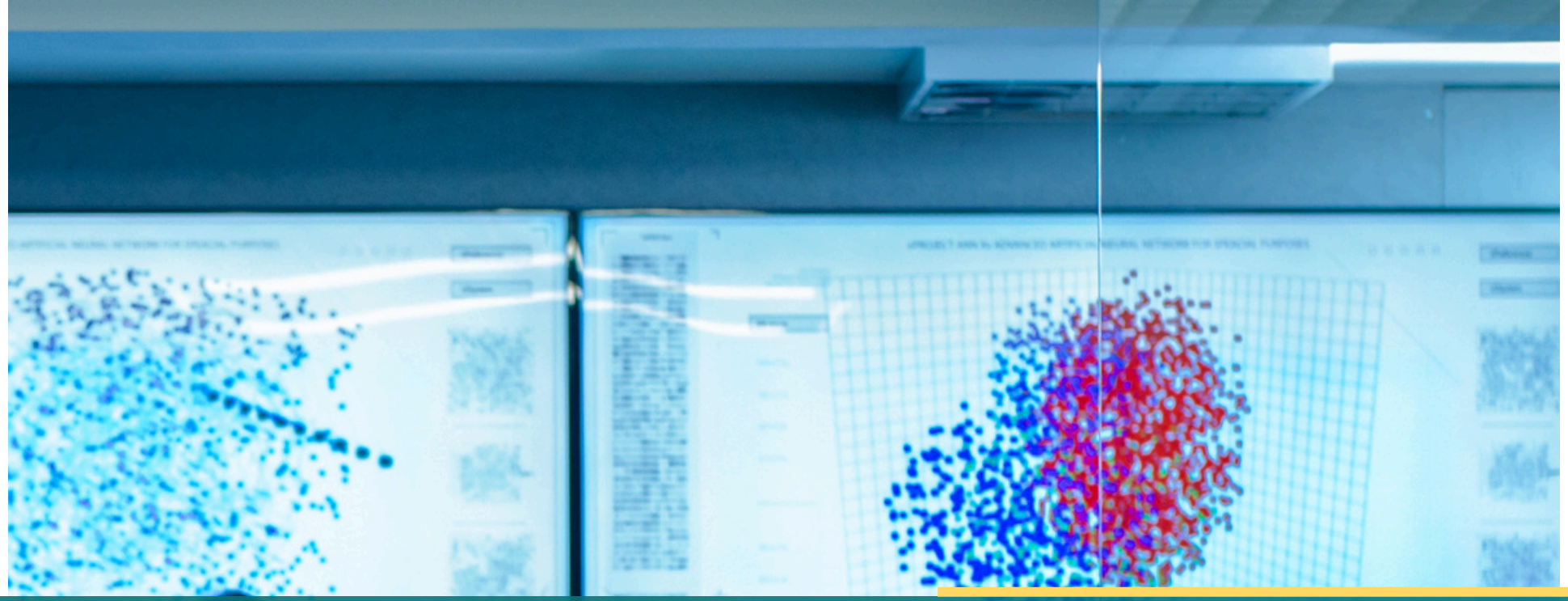
LUMIÈRE SUR … LES RISQUES LIÉS À LA CONFIDENTIALITÉ DES DONNÉES UTILISÉES POUR L'ENTRAÎNEMENT DES MODÈLES D'INTELLIGENCE ARTIFICIELLE

CADA, Avis n°20230314 du 30 mars 2023 sur le modèle d'Intelligence Artificielle de la Cour de Cassation pour pseudonymiser les décisions de justice avant leur publication



1 POURQUOI VEILLER À LA CONFIDENTIALITÉ DES DONNÉES D'ENTRAÎNEMENT ?

Impact organisationnel : Des tiers, dont vos concurrents, **pourraient récupérer vos données confidentielles**

Impact réputationnel : Des tiers pourraient récupérer des données personnelles et les utiliser afin de **nuire aux personnes**

Impact financier : Une sanction de la CNIL pourrait être imposée si les mesures adéquates n'ont pas été mises en œuvre afin de garantir la confidentialité des données personnelles d'entraînement du modèle

2 UN EXEMPLE

Le modèle d'intelligence artificielle utilisé par la Cour de Cassation lui permet de pseudonymiser automatiquement ses décisions de justice avant publication

..... Mme EXEMPLE mariée à M. BON .....
..... née le 4 mars 1970 ..... à Paris .....

Décision de justice non-pseudonymisée

Modèle d'intelligence artificielle (apprentissage automatique)

Pseudonymisation des décisions de justice

..... Mme Y. mariée à M. X .....
née en 1970 ..... en France .....

Décision de justice pseudonymisée

Ce modèle d'intelligence artificielle fait l'objet **d'un apprentissage en continu sur les corrections et annotations apportées par les vérificateurs**, afin d'améliorer ses performances au cours du temps

Modèle d'intelligence artificielle (apprentissage automatique)

Pseudonymisation des décisions de justice

..... Mme Y. mariée à M. X.....
née en 1970 ..... en France .....

Décision de justice pseudonymisée

Vérification de la pseudonymisation pour publication en ligne

Entraînement du modèle IA sur la base des corrections des erreurs et de l'annotation des décisions de justice

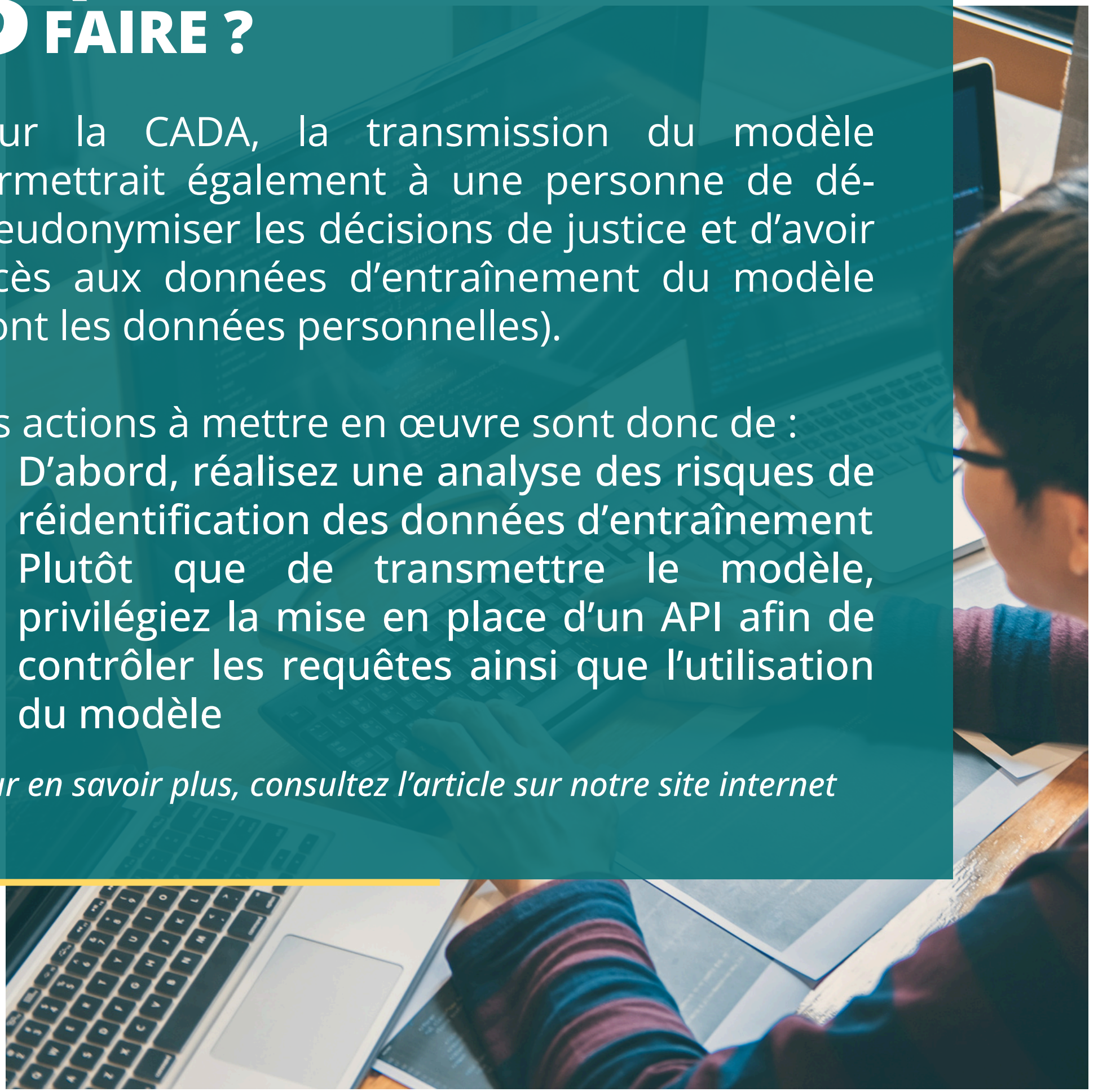
3 QUE FAIRE ?

Pour la CADA, la transmission du modèle permettrait également à une personne de dé-pseudonymiser les décisions de justice et d'avoir accès aux données d'entraînement du modèle (dont les données personnelles).

Les actions à mettre en œuvre sont donc de :

- D'abord, réalisez une analyse des risques de réidentification des données d'entraînement
- Plutôt que de transmettre le modèle, privilégiez la mise en place d'un API afin de contrôler les requêtes ainsi que l'utilisation du modèle

Pour en savoir plus, consultez l'article sur notre site internet



inside