

IA au travail : Ce que nos données peuvent révéler sans que nous ne l'ayons jamais dit



Septembre 2026 – IL – Fonctions Corporate

Edito de Thierry TABOY - CFE CGC AI federal referent

Nous laissons chaque jour dans l'entreprise des dizaines de traces numériques : connexions aux outils, agendas, messageries, compétences, déplacements professionnels, accès à des services ou à des bâtiments.

Prises séparément, ces informations répondent souvent à une fonction précise et peuvent paraître anodines.

L'intelligence artificielle change progressivement la manière dont il faut les regarder. Elle permet de rapprocher beaucoup plus facilement des informations dispersées et d'en déduire des éléments qui ne figurent explicitement dans aucune base de données.

Et cela déplace une partie de la question de la vie privée.

Des dossiers conçus pour ne plus être identifiables

Des chercheurs de Cornell Tech et de Weill Cornell Medicine ont récemment travaillé sur de véritables notes cliniques après leur passage par des systèmes de désidentification. Nom, adresse, coordonnées, identifiants : les informations permettant de rattacher directement le dossier à une personne avaient été retirées.

Autrement dit, les chercheurs sont partis de documents dans lesquels, en principe, **on ne devait plus pouvoir retrouver le patient**.

Ils ont ensuite construit un dispositif chargé de faire exactement l'inverse.

Avec ClinicalBERT, le plus performant des trois systèmes de désidentification testés, leur dispositif est encore parvenu à réidentifier correctement le patient dans 9 % des notes étudiées.

Près d'une note clinique sur dix a donc pu être rattachée à son patient alors même qu'elle avait préalablement été traitée pour empêcher cette identification.

Ce résultat ne signifie évidemment pas que 9 % de tous les dossiers médicaux anonymisés seraient aujourd'hui identifiables. Il montre autre chose : retirer les éléments directement identifiants ne fait pas disparaître tous les indices.

Un âge, une pathologie rare, des traitements, une chronologie, quelques événements personnels : leur combinaison peut finir par raconter une histoire suffisamment singulière.

Quand les petits indices deviennent un portrait

Une autre étude, présentée cette année à USENIX Security par des chercheurs notamment issus de l'ETH Zurich et d'Anthropic, s'est intéressée aux personnes s'exprimant sous pseudonyme sur Internet.

Au fil des années, quelqu'un mentionne son métier, une ville, une école, un changement professionnel, quelques centres d'intérêt. Aucun de ces éléments ne donne nécessairement son identité.

Ensemble, ils commencent à dessiner un portrait.

Les chercheurs ont utilisé de grands modèles de langage pour extraire ces indices, rechercher des profils compatibles et raisonner sur les candidats. Dans certaines configurations expérimentales, leur méthode atteint 68 % de rappel pour 90 % de précision. Sur environ 89 000 candidats, elle atteint encore 55 % de rappel à ce niveau de précision.

Cela ne signifie pas que 68 % des internautes utilisant un pseudonyme pourraient être identifiés. L'enseignement est plus intéressant : **une partie du travail de rapprochement qui exigeait beaucoup de temps humain peut désormais être automatisée.**

La donnée n'a pas changé. La puissance de celui qui la regarde, si.

Ce qui nous protégeait aussi : le temps

Notre vie privée bénéficie depuis longtemps d'une protection assez banale : personne n'a nécessairement le temps de réunir toutes les pièces du puzzle.

Les chercheurs parlent de *practical obscurity*. Une information peut être accessible en théorie tout en restant difficile à exploiter parce que l'effort nécessaire est trop important.

Les grands modèles réduisent cet effort. Ils peuvent parcourir de grandes quantités de textes, repérer des indices faibles, les rapprocher et répéter l'opération à grande échelle.

La singularité de nos traces était d'ailleurs connue bien avant l'IA générative. Dès 2013, une étude portant sur quinze mois de données de mobilité de 1,5 million de personnes montrait que quatre points spatio-temporels — quatre lieux associés à quatre moments — suffisaient à distinguer 95 % des individus dans le jeu de données étudié.

Nos traces sont souvent beaucoup plus singulières que nous ne l'imaginons.

Dans l'entreprise, le salarié est déjà identifié

C'est ici que le sujet devient particulièrement important dans le monde du travail.

Car le problème n'est généralement même plus de retrouver l'identité d'une personne derrière des données anonymisées.

Le salarié est déjà identifié.

L'entreprise connaît son identité, son poste et son organisation de rattachement. Les outils numériques peuvent par ailleurs produire d'autres traces : connexions, agendas, usages d'outils collaboratifs, données professionnelles, informations de mobilité ou de localisation lorsqu'un service le nécessite.

Chacune peut avoir une finalité parfaitement légitime.

Ce qui mérite notre attention est ce qu'il devient possible de **déduire de leur rapprochement**.

Imaginons un système disposant de données techniques permettant de connaître certains lieux ou moments de connexion de salariés.

Prises séparément, ces informations peuvent répondre à un besoin précis. Répétées dans le temps, elles peuvent faire apparaître des habitudes. Rapprochées entre plusieurs personnes, elles peuvent révéler des proximités ou des rencontres régulières.

Ajoutons un agenda, un organigramme ou d'autres informations professionnelles : il devient potentiellement possible de reconstruire une partie de relations ou de comportements qui n'ont jamais été directement renseignés.

Aucune base ne contenait nécessairement cette information.

Le système l'a déduite.

Et certaines déductions peuvent toucher directement à la vie privée et aux libertés individuelles et collectives.

L'agent n'a rien piraté. Il a simplement déduit

Cette question va prendre davantage d'importance avec l'arrivée des agents d'intelligence artificielle dans les organisations.

Nous avons appris à gérer les droits d'accès : qui peut consulter quelle application, quelles données sont accessibles, à quelles bases un système peut se connecter.

Mais imaginons un agent parfaitement autorisé à consulter plusieurs systèmes.

Il peut respecter tous ses droits d'accès et ne contourner aucune sécurité. Pourtant, le rapprochement de ce qu'il consulte peut lui permettre de produire une connaissance nouvelle.

Il n'a franchi aucune barrière.

Il a inféré.

Les spécialistes des données savent depuis longtemps effectuer ce type de rapprochement. L'IA n'invente pas le phénomène. Elle en modifie l'échelle, la vitesse et l'accessibilité.

Une question supplémentaire à poser

Lorsqu'un nouvel outil numérique arrive dans l'entreprise, nous savons demander : quelles données collecte-t-il ? Pour quelle finalité ? Qui peut y accéder ? Combien de temps sont-elles conservées ? Quels sont les droits des salariés ?

Ces questions restent fondamentales.

Une autre doit désormais entrer dans nos réflexes :

Que peut-on apprendre sur un salarié à partir de ces données, seules ou combinées avec celles que l'entreprise possède déjà ?

Cela ne revient pas à considérer chaque nouvel outil comme un dispositif de surveillance. L'intelligence artificielle peut améliorer des services, simplifier des tâches et ouvrir de nombreuses possibilités aux salariés.

Mais ces usages supposent de la confiance.

Un salarié peut comprendre qu'une donnée soit utilisée pour une finalité professionnelle clairement définie. Il doit aussi pouvoir avoir confiance dans le fait qu'elle ne servira pas, par rapprochement avec d'autres, à produire sur lui une connaissance sans rapport avec cette finalité.

Le droit européen dispose déjà de principes solides : finalité, proportionnalité, minimisation des données, protection de la vie privée, information des personnes.

Ce qui change rapidement, ce sont les capacités techniques.

Face à un système d'intelligence artificielle, demander « **quelles données utilisez-vous ?** » restera indispensable.

Il faudra de plus en plus souvent ajouter :

« Et avec ces données, qu'êtes-vous capable de savoir sur moi que je ne vous ai jamais dit ? »



**CHOISISSEZ
CEUX
QUI
AGISSENT !**



Thierry Taboy – 06 85 12 60 55
Awataf Temagoulte – 06 32 58 82 79

cfecgc-orange.org

Pour recevoir nos publications : <https://cfeora.org/S0iLels>

