

Les signaux faibles dans le travail social

L'IA prédictive, un outil de maltraitance institutionnelle ?

Adrien Guionie

Mars 2026

1. Ce que j'entends sur le terrain, et ce qui m'a fait basculer

À chaque intervention que je mène auprès des structures sociales et médico-sociales, la même demande revient avec une régularité troublante. L'intelligence artificielle pour détecter les signaux faibles. Les directeurs en parlent avec espoir, les cadres avec pragmatisme, les élus avec l'évidence de ceux qui ont trouvé la solution miracle.

Cette demande n'a rien de surprenant. Elle ne vient pas d'une volonté de nuire, mais d'une intention profondément sincère face à des difficultés que je connais bien et que je prends au sérieux. L'augmentation des situations complexes, la diminution des effectifs, la pression sur les délais, la responsabilité juridique qui pèse de plus en plus lourd sur les professionnels.

Les arguments avancés sont souvent les mêmes, et ils sont loin d'être dénués de sens pour qui les entend la première fois : augmenter l'accès au droit, faire davantage de prévention, repérer les situations à risque avant qu'il ne soit trop tard. Repérer plus vite pour optimiser les moyens rares et faire des économies.

L'intelligence artificielle prédictive apparaît alors comme une bouée de sauvetage, une réponse technique à des tensions structurelles, un outil providentiel qui viendrait compenser nos manques chroniques de moyens humains et de temps.

Pendant longtemps, cette demande me paraissait logique. Bienveillante même.

Jusqu'à ce qu'une phrase, totalement banale, prononcée lors d'un échange avec un établissement, ne vienne tout éclairer d'une lumière différente.

"On aimerait utiliser les signaux faibles pour classifier les urgences de nos signalements."

Une phrase anodine, presque technique, mais qui porte en elle toute l'ambiguïté de ce que ces outils promettent et ce qu'ils font concrètement.

Classifier les urgences automatiquement, sur la base d'un score, décider qui mérite notre attention immédiate et qui peut attendre, déléguer à un algorithme le discernement professionnel qui fait le cœur de notre métier.

C'est à ce moment précis que j'ai compris que cette bienveillance reposait sur une méconnaissance totale de ce que ces outils font réellement.

Parce que pendant que nous fantasmons sur les promesses des signaux faibles, d'autres secteurs les ont déjà déployés, testés, mesurés, documentés. Et dans certains cas, les ont abandonnés après des scandales retentissants dont les conséquences humaines sont aujourd'hui établies.

Alors avant de déployer ces dispositifs dans le travail social, regardons ce qui s'est passé ailleurs, regardons les faits, regardons les chiffres, et surtout, décryptons les mécanismes techniques qui font que ces algorithmes ciblent toujours les mêmes personnes.

Les plus fragiles.

2. Quatre cas concrets : ce que les signaux faibles font vraiment

2.1. La CAF : trente-deux millions de personnes notées chaque mois

PREMIER CAS : L'ALGORITHME DE LA CAF

Depuis 2010, chaque allocataire de la CAF se voit attribuer un score. Un chiffre entre 0 et 1. Plus vous vous rapprochez de 1, plus l'algorithme vous considère comme suspect. Dépassez un certain seuil, et un contrôle se déclenche automatiquement.

Vous ne le savez pas. Personne ne vous prévient. Mais tous les premiers du mois, l'algorithme recalcule votre score. À partir de plus de mille variables analysées lors de sa conception, il utilise maintenant plusieurs dizaines de critères pour vous noter.

Quels critères ? Être bénéficiaire du RSA. Toucher l'AAH. Être parent isolé. Habiter dans un quartier classé défavorisé. Percevoir l'AAH tout en travaillant. Avoir des revenus qui varient d'un mois sur l'autre.

Vous avez compris le problème ? L'algorithme ne cible pas des comportements frauduleux. Il cible des situations de précarité.

Les chiffres sont sans appel :

- 32 millions de personnes concernées, dont 13 millions d'enfants
- Les familles monoparentales, 82% de mères seules, sont massivement surreprésentées dans les contrôles
- Résultat financier ? 0,2% du montant total des prestations versées
- 15 organisations, dont Amnesty International et l'ANAS, ont saisi le Conseil d'État

L'algorithme ne coche pas la case "femme". Il n'en a pas besoin. Il agrège des variables qui, mises bout à bout, créent un proxy statistique parfait pour cibler les mères seules : parent isolé plus RSA plus quartier défavorisé égale, dans 82% des cas, une mère seule.

Ce n'est pas un bug. C'est le fonctionnement normal de l'algorithme.

2.2. Les Pays-Bas : quand le gouvernement doit démissionner

DEUXIÈME CAS : LE SCANDALE NÉERLANDAIS

Entre 2013 et 2019, l'administration fiscale néerlandaise a utilisé un système de détection automatique des fraudes aux allocations familiales. L'objectif : identifier les bénéficiaires payés à tort.

Vingt-six mille familles ont été accusées à tort, intégralement, toutes.

Les montants réclamés atteignaient parfois plusieurs dizaines de milliers d'euros : saisies bancaires, expulsions, séparations familiales, et dans plusieurs cas documentés, tentatives de suicide.

Qu'est-ce qui a mal tourné ? L'algorithme ciblait de manière disproportionnée les familles d'origine immigrée, comme l'a confirmé l'enquête parlementaire. La variable "double nationalité" servait de signal de risque.

Un profilage racial, systémique, automatisé.

Les conséquences :

- Démission du gouvernement Rutte III en janvier 2021
- Reconnaissance officielle de discrimination institutionnelle
- Indemnisation de 30 000 euros par famille (sur fonds publics)
- Abandon définitif du système

Mais le plus terrible, c'est ce que l'enquête a révélé sur le comportement des agents. Ils validaient les décisions de l'algorithme, systématiquement, sans examen critique. Le score disait "fraudeur", donc la personne était traitée comme un fraudeur.

La machine avait remplacé le jugement.

2.3. La DGFIP : cinquante pour cent de contrôles, quatorze pour cent de résultats

TROISIÈME CAS : LE DATAMINING FISCAL

La Direction générale des finances publiques a massivement investi dans le croisement automatique de données pour compenser la baisse des effectifs. L'objectif : que le datamining représente la moitié des contrôles fiscaux programmés. Cet objectif a été atteint dès 2022 pour les professionnels.

Et alors ? Quel est le résultat ?

La Cour des comptes vient de publier son rapport. Décembre 2025, et le constat est sans appel.

Les chiffres de la Cour des comptes :

- En 2023, les contrôles issus du datamining représentaient 44% de l'ensemble des contrôles
- Ils n'ont généré que 13,8% des droits et pénalités mis en recouvrement
- Le rendement est nettement inférieur aux contrôles traditionnels ciblés par expertise humaine
- Résultats globaux du contrôle fiscal : stagnants depuis 2015, alors que les recettes fiscales totales ont progressé de 44%

La juridiction financière le dit explicitement : "*L'efficacité de la stratégie reste toutefois à démontrer.*" Elle parle de "résultats décevants".

Pourquoi ce paradoxe ? Parce que l'algorithme génère un volume massif de contrôles, mais il cible préférentiellement des dossiers à faible enjeu, des incohérences déclaratives mineures. Pendant ce temps, les fraudes sophistiquées passent sous le radar : montages offshore, prix de transfert manipulés, optimisations agressives.

L'algorithme contrôle beaucoup. Mais il contrôle mal.

2.4. Le Japon : quand l'algorithme dit "non" et qu'un enfant meurt

QUATRIÈME CAS : HONOKA NAKABAYASHI

En 2023, au Japon, une fillette de 4 ans, Honoka Nakabayashi, décède. Malgré plusieurs informations préoccupantes transmises par la crèche qui ne la voyait plus, malgré des demandes d'aide répétées de la mère elle-même, aucun dossier n'avait été constitué par les services de protection de l'enfance.

Personne n'avait vu cet enfant depuis un an.

Le département utilisait depuis 2020 un système d'intelligence artificielle pour évaluer les informations préoccupantes. L'algorithme analysait 130 000 données historiques selon 21 critères. Pour des situations similaires à celle d'Honoka, l'algorithme recommandait le placement dans 39% des cas.

Autrement dit : dans 61% des cas, l'algorithme disait "non".

Le jour de sa mort, Honoka pesait 4 kg de moins que la moyenne pour son âge. Elle vivait seule dans une pièce à part, séparée de sa mère et de ses sœurs. Elle n'était pas lavée. Ses ongles n'avaient jamais été coupés.

Ce que révèle l'affaire :

- L'algorithme reposait sur un questionnaire à choix multiples ("y a-t-il une blessure au visage ?")
- Il ne prenait pas en compte l'abandon à la naissance ni le placement en pouponnière
- Le média ITmedia parle de "sacralisation de l'IA" : acceptation sans critique des résultats fournis
- Mars 2025 : le Japon abandonne l'IA en protection de l'enfance (7 millions d'euros investis)

Le rapport d'évaluation post-mortem du département souligne que l'IA n'était qu'un outil censé soutenir et non remplacer le jugement humain. Mais dans les faits, le système guidait les décisions. Et personne n'est allé voir l'enfant.

L'algorithme avait dit "non". Honoka est morte.

3. Décryptage technique : pourquoi ça ne peut pas marcher autrement

Alors maintenant, analysons les mécanismes internes de ces systèmes. Comprenons pourquoi ces quatre cas ne sont pas des accidents, ni des problèmes de paramétrage ou de mauvaises données, mais des conséquences inévitables du fonctionnement même de ces algorithmes.

3.1. Les biais algorithmiques et la confiance aveugle dans les chiffres

Analysons maintenant les mécanismes internes de ces systèmes. Comprenons pourquoi ces quatre cas ne sont pas des accidents, mais des conséquences inévitables du fonctionnement même de ces algorithmes.

Premier mécanisme : le biais de représentation

Le biais de représentation

L algorithmes apprennent sur des données qui ne reflètent pas équitablement la réalité, mais surreprésentent certaines catégories au détriment d'autres. Ils reproduisent ensuite cette surreprésentation dans leurs prédictions.

Reprenons le cas de la CAF. En 2010, l'algorithme a été entraîné sur 7 000 contrôles aléatoires. Le problème ? Les allocataires précaires sont mécaniquement surreprésentés dans ces données historiques : ils cumulent plus de prestations, effectuent plus de démarches, leurs situations changent plus souvent.

Résultat : l'algorithme "apprend" que les profils à risque ressemblent aux profils des personnes précaires, non pas parce qu'elles fraudent davantage, mais parce que leurs dossiers sont plus visibles dans le système.

On ne peut pas corriger un biais algorithmique quand le biais est dans la réalité sociale elle-même.

Deuxième mécanisme : le cercle vicieux et la confiance aveugle

Le cas néerlandais illustre un phénomène encore plus pervers. L'algorithme cible les familles immigrées, qui font l'objet de contrôles approfondis détectant des irrégularités mineures (retards, erreurs de formulaire). Ces détections alimentent l'algorithme qui "apprend" que sa prédiction était correcte et renforce son ciblage. Le système se valide lui-même.

Pendant ce temps, les familles non ciblées ne font l'objet d'aucun contrôle. Leurs éventuelles irrégularités ne sont jamais détectées.

Mais le plus grave ? Les agents validaient quasi systématiquement les décisions de l'algorithme. Ce phénomène, documenté en psychologie cognitive sous le nom d'**automation bias**, désigne la tendance à accorder une confiance excessive aux systèmes automatisés au point de négliger son propre jugement.

On l'observe chez les pilotes qui suivent aveuglément le pilote automatique, chez les médecins qui suivent les recommandations d'un système d'aide à la décision, chez nous-mêmes quand nous suivons Waze même si le trajet semble absurde.

Aux Pays-Bas, le score de risque élevé créait une présomption de culpabilité. À la CAF, les agents justifient les contrôles par "le système a détecté une anomalie", sans pouvoir expliquer laquelle. Au Japon, l'algorithme disait "non" et personne n'est allé voir l'enfant.

Pourquoi cette confiance aveugle ? Parce que l'algorithme produit un score qui possède une aura d'objectivité. Il est quantifié, calculé, reproductible. Il semble neutre, rationnel, scientifique. L'humain ne remet pas en doute une donnée statistique, mathématique. C'est ce qu'on nous a appris à l'école : les chiffres disent le vrai.

Sauf que ce score ne mesure pas la réalité d'une fraude ou d'un danger. Il mesure la ressemblance statistique d'un profil avec des profils passés ayant fait l'objet de contrôles. C'est une probabilité conditionnelle calculée sur des données historiques biaisées.

Le score a remplacé l'analyse.

Troisième mécanisme : l'invisibilité structurelle des situations complexes

Les quatre cas ont un point commun : l'algorithme ne détecte pas les situations les plus graves, mais les plus fragiles.

À la CAF, les allocataires ciblés présentent des "anomalies" facilement détectables mais parfaitement légitimes (reprise d'emploi en temps partiel, activité saisonnière). Les fraudes organisées, les montages complexes ne génèrent aucun signal et nécessitent une expertise humaine.

À la DGFIP, même problème : le datamining détecte des incohérences mineures mais passe à côté des optimisations fiscales agressives, d'où le faible rendement (13,8% des recettes pour 44% des contrôles).

Au Japon, l'algorithme disait "non" pour Honoka parce que sa situation ne correspondait pas aux patterns statistiques majoritaires. Sa singularité la rendait invisible.

Question pour les décideurs

Si l'algorithme ne détecte que ce qui est facilement détectable, et que ce qui est facilement détectable correspond massivement à des situations de précarité légitimes, quel est l'intérêt réel de l'algorithme ? À part renforcer la surveillance des plus fragiles.

La question de la fiabilité

Valeur Prédictive Positive (VPP)

Probabilité qu'une personne identifiée comme "à risque" par l'algorithme le soit réellement.
 $VPP = \text{Vrais Positifs} / (\text{Vrais Positifs} + \text{Faux Positifs})$.

Le scandale néerlandais nous donne une illustration dramatique. Vingt-six mille familles accusées. Toutes à tort. Si l'on suppose que l'algorithme a aussi détecté des cas réels (donnée non publique), le taux de faux positifs était manifestement très élevé, probablement supérieur à 50%.

À la DGFIP, le faible rendement (13,8% des recettes pour 44% des contrôles) suggère la même chose. La majorité des dossiers ciblés ne présentaient pas d'enjeu significatif.

Alors maintenant, transposons au travail social.

Imaginez un algorithme de signaux faibles en protection de l'enfance avec une VPP de 50%. Cela signifie qu'une famille sur deux ciblée comme "à risque" ne l'est pas réellement.

Les conséquences humaines ? Stigmatisation, intrusion, rupture de confiance, Déstabilisation d'une famille déjà fragile.

Pour rien.

L'impossibilité du "risque zéro"

Les promoteurs des signaux faibles nous font une promesse. "Ne rien laisser passer." "Détecter avant qu'il ne soit trop tard." "Anticiper les situations de danger."

Cette promesse est techniquement impossible.

Tout algorithme prédictif doit arbitrer entre deux types d'erreurs :

- **Les faux négatifs** : ne pas détecter une situation réellement à risque
- **Les faux positifs** : détecter à tort une situation comme à risque

Diminuer les faux négatifs implique mécaniquement d'augmenter les faux positifs. Et inversement. C'est un équilibre technique incontournable.

Dans le contexte du travail social, les deux types d'erreurs ont des conséquences graves :

- Un faux négatif = une situation de danger non détectée
- Un faux positif = une famille stigmatisée, contrôlée, déstabilisée à tort

L'algorithme ne résout pas ce dilemme. Il le déplace. Il le technicise. Et surtout, il le rend invisible.

Le professionnel qui suit les recommandations algorithmiques ne voit pas les faux négatifs (par définition, ils ne sont pas ciblés). Et il ne questionne pas les faux positifs (l'algorithme a dit "à risque", donc on contrôle).

Question éthique fondamentale

Qui assume la responsabilité de l'erreur ? Le professionnel qui a suivi l'algorithme ?
L'institution qui l'a déployé ? L'éditeur du logiciel ? Le data scientist qui l'a paramétré ?
Ou personne ?

3.4. De l'accompagnement à la surveillance : le glissement sémantique

Le changement de paradigme

Regardez ce qui s'est passé à la CAF. Un glissement sémantique majeur. De la "maîtrise des risques" (vocabulaire de gestion) à la "lutte contre la fraude" (vocabulaire répressif).

Ce glissement n'est pas un détail. Il transforme la nature même de la relation entre l'institution et l'utilisateur.

Historiquement, la CAF était un organisme de versement de droits. Les contrôles existaient, mais dans une logique de régularisation. Vérifier que les prestations correspondent aux situations. Corriger les erreurs. Récupérer les trop-perçus.

De l'accompagnement à la surveillance

À la CAF, un glissement sémantique majeur s'est opéré : de la "maîtrise des risques" à la "lutte contre la fraude". Chaque allocataire devient un suspect potentiel. Le score transforme la présomption d'innocence en présomption de fraude. Trente-deux millions de personnes notées chaque mois, sans le savoir, sans pouvoir se défendre, sans recours.

Un outil qui surveille au lieu de protéger, ce n'est plus du travail social. C'est du contrôle social.

4. Transposition au travail social : les risques que nous prenons

4.1. Ce qui fait la spécificité de notre secteur

Le travail social repose sur un paradoxe délicat : intervenir auprès de personnes en situation de vulnérabilité, parfois de manière contrainte, tout en maintenant une relation de confiance suffisante pour que l'accompagnement produise du changement. Sans cette confiance, rien n'est possible.

Comment établir une alliance de travail si la personne sait qu'elle a été ciblée par un algorithme comme "à risque" ? Le cas néerlandais nous donne la réponse : les familles accusées à tort ont rompu tout lien avec les institutions sociales. La confiance, une fois brisée, ne se reconstruit pas.

Les données du travail social ne se quantifient pas sans perte massive d'information. Un travailleur social note : "La mère semble dépassée par la situation." Cette observation est contextualisée, nuancée, fait partie d'une évaluation globale. Extraite de son contexte, transformée en variable booléenne, elle devient un signal qui alimentera un score. Le sens est perdu, le contexte est perdu, la nuance est perdue, mais le chiffre demeure.

4.2. Les risques spécifiques que nous prenons

Si nous reproduisons le schéma observé à la CAF, un algorithme de signaux faibles ciblera mécaniquement les familles qui laissent le plus de traces : familles suivies par l'ASE, bénéficiaires du RSA, familles monoparentales, habitants de quartiers prioritaires.

Ce ciblage reproduira une réalité que nous connaissons : les mesures de protection concernent majoritairement des familles pauvres, non parce que la maltraitance est absente des milieux aisés, mais parce que ces familles sont plus visibles. L'algorithme transformera cette surreprésentation statistique en surreprésentation algorithmique.

Pendant ce temps, les situations de maltraitance dans les milieux aisés resteront invisibles. Ces familles consultent en libéral, scolarisent en privé, ne génèrent aucune donnée exploitable.

Le score algorithmique fige une situation dans le temps. Une famille notée "à risque élevé" à un instant T le restera tant que les variables n'évoluent pas. L'algorithme crée une forme de déterminisme en contradiction frontale avec les principes du travail social.

4.3. L'incompatibilité structurelle

L'analyse de ces quatre cas concrets et de leurs mécanismes techniques me conduit à une conclusion : les algorithmes de signaux faibles sont structurellement incompatibles avec les principes du travail social.

Cette incompatibilité n'est pas conjoncturelle. Elle est intrinsèque : les algorithmes reproduisent les biais des données d'entraînement qui surreprésentent les populations précaires. Le ciblage algorithmique renforce cette surreprésentation. La confiance dans les scores produit une paresse cognitive. Les faux positifs et négatifs ont des conséquences graves. La logique de surveillance entre en contradiction avec la logique d'accompagnement.

Aucun de ces problèmes ne peut être résolu par un meilleur algorithme.

5.1. L'intelligence artificielle dans le travail social : délimiter les usages

Je ne suis pas contre l'intelligence artificielle dans le travail social. Je travaille chaque jour à développer des outils pour notre secteur.

Mais il faut délimiter précisément ce qui est légitime et ce qui ne l'est pas.

Les usages légitimes de l'IA dans le travail social

L'IA peut être un outil d'aide à la rédaction, un assistant augmenté qui aide le professionnel à structurer sa pensée, à reformuler, à synthétiser des écrits complexes. C'est ce que fait DictIA : il accompagne l'écriture professionnelle sans jamais prendre la décision à la place du travailleur social.

L'IA peut être un support d'accompagnement pour les personnes elles-mêmes, comme Mon Éval' et Mon Histoire : des outils qui aident les jeunes à structurer leur réflexion, à préparer leurs rencontres avec les professionnels, à garder une trace de leur parcours. L'IA est au service de la personne, pas de son contrôle.

L'IA peut être un vecteur d'innovation sociale, permettant aux professionnels de terrain de créer des outils directement liés aux besoins qu'ils identifient, sans passer par des circuits institutionnels lourds. C'est ce que permet l'IA générative : démocratiser la création d'outils adaptés.

Caractéristique commune à ces usages : **l'IA reste un outil au service du professionnel ou de la personne accompagnée. Elle ne prend aucune décision. Ne produit aucun score. N'oriente aucune priorisation.**

Le jugement professionnel demeure central. La relation humaine reste première.

Les usages que je refuse catégoriquement

- **Scoring des personnes** : attribution d'un score de risque, de priorité, de dangerosité
- **Priorisation algorithmique** : orientation des interventions sur la base de calculs prédictifs
- **Détection automatique de signaux faibles** : croisement de données pour identifier des situations "à risque"
- **Aide à la décision contraignante** : systèmes dont les recommandations orientent la décision

Caractéristique commune : **l'IA oriente, filtre, ou prend en charge une partie de la décision professionnelle.**

Elle transforme l'accompagnement en surveillance.

5.2. Les lignes rouges

Au regard de tout ce que nous avons vu, je pose plusieurs lignes rouges :

1. **Interdiction du scoring des personnes** : Aucun système ne doit attribuer un score numérique à une personne accompagnée
2. **Interdiction des systèmes prédictifs** : Aucun algorithme ne doit prédire un comportement, un risque, une évolution
3. **Supervision humaine effective obligatoire** : Tout usage d'IA doit faire l'objet d'un contrôle humain réel, informé, critique
4. **Transparence totale** : Les personnes doivent être informées de tout usage d'IA les concernant
5. **Droit d'opposition** : Possibilité de refuser un traitement algorithmique et d'obtenir un examen humain

5.3. Conclusion

Les signaux faibles nous promettent une révolution. Cette promesse repose sur une triple illusion.

L'illusion technique : l'algorithme serait plus objectif que l'humain.

L'illusion prédictive : on pourrait anticiper les situations à risque.

L'illusion gestionnaire : on optimiserait l'allocation des moyens.

L'analyse des quatre cas concrets démontre le contraire.

Les algorithmes reproduisent et amplifient les biais humains. Ils ciblent les plus fragiles, pas les plus à risque. Ils génèrent un volume considérable de faux positifs. Leur rendement est décevant. Ils transforment l'accompagnement en surveillance. Ils détruisent la relation de confiance.

Un outil qui cible les plus fragiles au lieu de les protéger, ce n'est pas de l'aide à la décision.

C'est de la maltraitance institutionnelle assistée par ordinateur.

Le jour où un algorithme décidera qui mérite notre attention et qui peut attendre, nous aurons perdu l'essence même de notre métier.

le regard direct, le lien humain, le discernement professionnel, la capacité à douter, la liberté de nous tromper

Certaines technologies ne sont pas faites pour être adoptées.

Elles sont faites pour être refusées.

Les signaux faibles en font partie.

Références

Défenseur des droits - *Algorithmes, systèmes d'IA et services publics : pour que le droit n'oublie personne*, novembre 2024.
https://www.defenseurdesdroits.fr/sites/default/files/2024-11/DDD_rapport_algorithmes-systemes-d-IA-et-services-publics_2024_20241025.pdf

Cour des comptes - *La lutte contre la fraude fiscale*, décembre 2025.
<https://www.ccomptes.fr/sites/default/files/2025-12/20251216-Lutte-contre-la-fraude-fiscale.pdf>

La Quadrature du Net - *Scoring of welfare beneficiaries: the indecency of CAF's algorithm now undeniable*, novembre 2023.
<https://www.laquadrature.net/en/2023/11/27/scoring-of-welfare-beneficiaries-the-indecency-of-cafs-algorithm-now-undeniable/>

La Quadrature du Net - *Notation des allocataires : la CAF étend sa surveillance à l'analyse des revenus en temps réel*, mars 2024.
<https://www.laquadrature.net/2024/03/13/notation-des-allocataires-la-caf-etend-sa-surveillance-a-lanalyse-des-revenus-en-temps-reel/>

CNAPE - *Le recours à l'intelligence artificielle en protection de l'enfance. Note de positionnement*, décembre 2025.
<https://www.cnape.fr/documents/lintelligence-artificielle-en-protection-de-lenfance/>

Le Média Social - *Protection de l'enfance : pourquoi le Japon abandonne l'IA, et ce que la France devrait en retenir*, juillet 2025.
https://www.lemediasocial.fr/protection-de-lenfance-pourquoi-le-japon-abandonne-lia-et-ce-que-la-france-devrait-en-retenir_YaOgKp

Collectif Changer de Cap - *Témoignages d'allocataires contrôlés par l'algorithme de la CAF*.
<https://www.changerdecap.net/>

