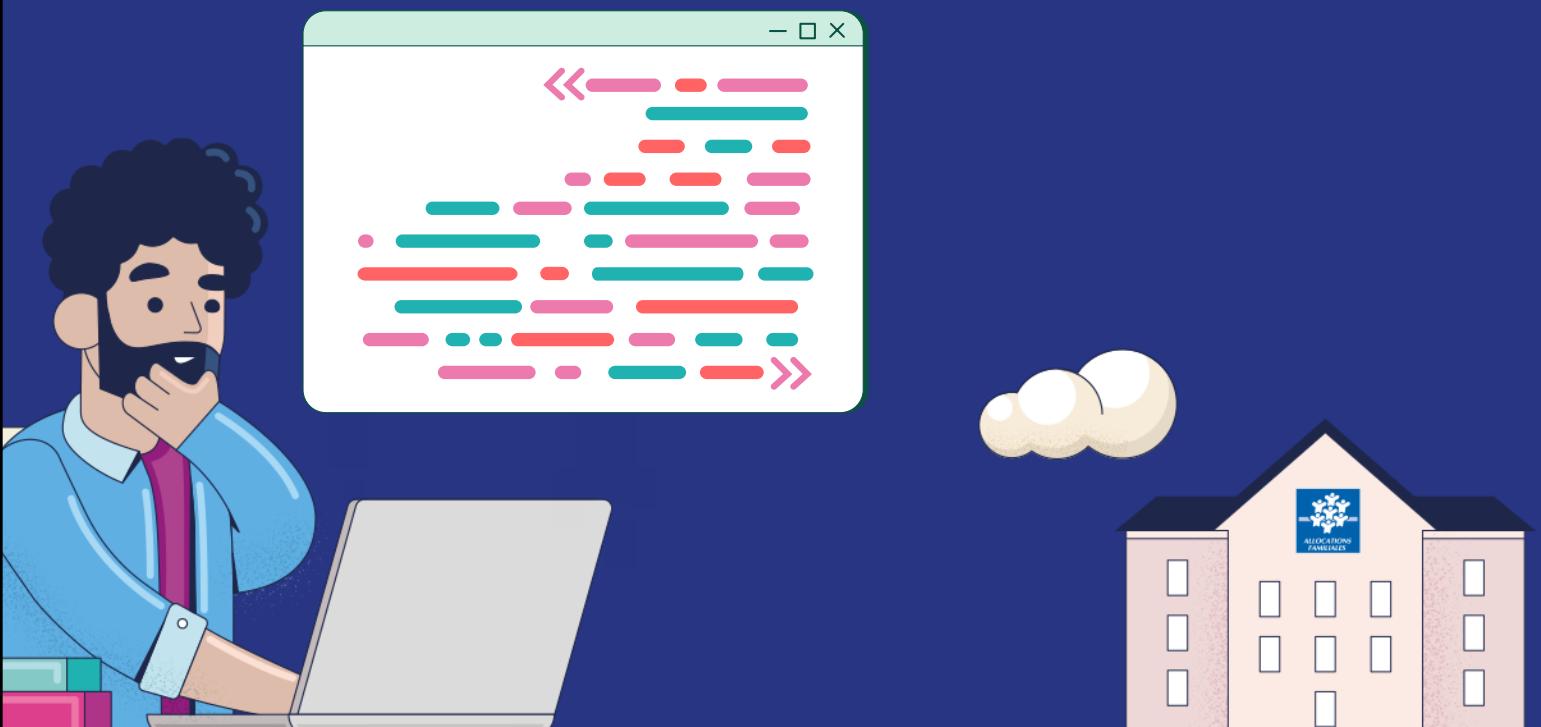


# Transparence et garanties éthiques autour de l'algorithme d'appui aux contrôles des CAF

Mise en œuvre du modèle  
datamining données entrantes 2026 (DMDE)



**Dossier de presse**  
Janvier 2026

## Édito : Nicolas Grivel, directeur général de la Cnaf

« *Conjuguer efficacité des contrôles avec la transparence et l'éthique dans l'usage des données* ».

### 1 | Une démarche éthique structurée sur l'utilisation des données et des algorithmes

- Une démarche rendue nécessaire par l'évolution des outils numériques
- Une nouvelle gouvernance mise en place en 2025
- Le comité et la Charte
- Une volonté de transparence et de pédagogie

### 2 | Contrôles Caf : un nouvel algorithme plus transparent

- Pourquoi les Caf utilisent-elles le datamining ?
- Comment les Caf identifient-elles les dossiers à contrôler ?
- Des choix concrets pour réduire les risques de biais
- La solidarité à la source : un modèle adapté à un contexte déclaratif qui change
- Les critères pris en compte par le DMDE 2026

## Contacts presse

**Virginie RAULT**  
07 78 95 49 90

**Julien PRADINES**  
06 15 95 78 79

[presse@cnafr.fr](mailto:presse@cnafr.fr)



“

## **Conjuguer efficacité des contrôles avec la transparence et l'éthique dans l'usage des données ”**



Dans un contexte où l'ouverture des données et des algorithmes publics suscite débats et attentes, l'utilisation des outils numériques doit être envisagée comme une opportunité, et non comme une menace.

À la Cnaf, nous faisons le choix d'une démarche fondée sur la transparence, l'éthique et la responsabilité. Les outils numériques que nous construisons ne sont pas une fin en soi : ils sont un levier pour garantir le paiement au juste droit, améliorer la qualité de service et anticiper les besoins des allocataires.

Cette ambition s'inscrit dans une logique maîtrisée, avec des garanties fortes : respect du Règlement général de protection des données (RGPD), protection des données personnelles, démarche de limitation des biais potentiels et publication d'une documentation.

Le portail Cafdata, dont la nouvelle ergonomie était présentée aux médias l'an dernier, illustre depuis plusieurs années cette volonté : rendre accessibles des données fiables à tous – chercheurs, journalistes, citoyens – tout en préservant la confidentialité pour les allocataires.

L'ouverture des données et algorithmes de la Cnaf suppose une gouvernance exigeante, incarnée par un comité et une charte d'éthique, pour que chaque innovation – qu'il s'agisse d'algorithmes de ciblage ou de systèmes décisionnels modernisés – soit pensée dans le respect des droits.

Notre responsabilité est double : garantir la performance des politiques publiques et renforcer la confiance des citoyens. Dans un monde où la donnée est partout, nous affirmons qu'elle doit être au service de la solidarité et de l'équité.

**Ouvrir, oui. Mais ouvrir avec sens, méthode et transparence.**

**Nicolas Grivel, directeur général de la Cnaf**

# 1

Une démarche éthique structurée sur l'utilisation des données et des algorithmes

## Une démarche éthique structurée sur l'utilisation des données et des algorithmes

### Une démarche rendue nécessaire par l'évolution des outils numériques

L'utilisation d'algorithmes, et la mobilisation à terme des technologies d'intelligence artificielle, est riche en opportunités pour les organismes de protection sociale. Afin de mobiliser ces leviers de manière éclairée et responsable, la Cnaf souhaite développer une stratégie basée sur la confiance, qui place l'éthique au centre de ses préoccupations. Cette stratégie doit permettre d'améliorer la qualité du service rendu aux allocataires et aux partenaires, mais aussi les conditions de travail des agents des Caf tout en respectant les valeurs de service public.

Cette démarche a abouti concrètement à la mise en place d'une gouvernance : un comité éthique, constitué de représentants de la Cnaf et des Caf, mais aussi d'associations d'usagers, de membres du Conseil d'administration et d'experts en droit, nouvelles technologies... Une charte a également été formalisée, pour encadrer le développement et l'usage des algorithmes et de l'IA.

Enfin, depuis le début de l'année 2026, un nouveau modèle de datamining est mis en œuvre, dont la conception était rendue nécessaire par les évolutions réglementaires et la généralisation de la solidarité à la source depuis mars 2025. Sa démarche d'élaboration a été particulièrement guidée par la prise en compte des risques de discrimination liés aux variables retenues, en excluant les variables les plus sensibles ou contestables, selon une méthode explicable. Enfin, dans un souci de transparence sur son fonctionnement, le code source est pour la première fois accessible en ligne sur le [site caf.fr](http://site.caf.fr).

### Une nouvelle gouvernance mise en place en 2025

Le [comité d'éthique sur les usages des données, des algorithmes et de l'intelligence artificielle de la Cnaf](#), réuni pour la première fois en mars 2025, a pour mission d'appuyer la Cnaf afin d'identifier les risques et les moyens de les atténuer, notamment sur les sujets sensibles liés aux libertés publiques.

Il est composé\* :

- d'experts des nouvelles technologies, du droit, des sciences de la donnée et de la réflexion éthique et philosophique,
- de représentants des usagers des Caf et du Conseil d'administration de la Cnaf,
- de professionnels de la Cnaf et des Caf.

\* La liste de ses membres et le compte-rendu des réunions sont en ligne sur [caf.fr](http://caf.fr) :  
[Le comité d'éthique sur les usages des données, des algorithmes et de l'intelligence artificielle de la Cnaf](#)

Une démarche éthique structurée sur l'utilisation des données et des algorithmes

Le comité s'appuie sur un texte fondateur : [La Charte éthique sur l'usage des algorithmes et de l'intelligence artificielle de la Cnaf](#)

Elle rappelle dans son préambule la nécessité d'un respect absolu de la vie privée des allocataires et de leurs données personnelles. Elle vise ainsi à couvrir tant les algorithmes que les technologies d'intelligence artificielle et de big data dont la création, l'usage, et les règles sont définis par la Cnaf et relèvent de sa responsabilité.

La Charte formalise ainsi les engagements de la Cnaf et des Caf pour :

- des finalités de l'utilisation des algorithmes et de l'IA en adéquation avec leur mission de service public et la responsabilité sociétale ;
- des innovations au service de l'humain ;
- un principe de transparence ;
- la sécurité au cœur des données et des solutions ;
- un cadre responsable et durable limitant l'impact environnemental.

### Une volonté de transparence et de pédagogie

En 2025, la Cnaf s'est attachée à communiquer vers les allocataires et le grand public sur l'utilisation des données, via l'article sur caf.fr [La Caf et vos données personnelles](#).

L'objectif est d'expliquer comment les Caf traitent leurs données personnelles, nécessaires pour verser des prestations et garantir la qualité du service, et de préciser les mesures de sécurité mises en œuvre pour les protéger, dans un contexte où les menaces de cyberattaques se multiplient.

La conception du nouveau modèle de datamining, utilisé dans le cadre de la politique de contrôle, est une première réalisation concrète de cette démarche d'éthique et de transparence.



# 2

## Contrôles Caf : un nouvel algorithme plus transparent

Depuis janvier 2026, dans le cadre de sa politique de réduction des trop-perçus et des rappels de prestations, la Cnaf utilise une nouvelle version de son modèle de datamining : le DMDE 2026 (DataMining Données Entrantes 2026). Cet algorithme est un outil statistique d'aide au ciblage de contrôles réalisés par les Caf, afin de contribuer au paiement au juste droit : corriger des erreurs, limiter les indus (trop-perçus) et identifier aussi des rappels de droits en faveur des allocataires. Concrètement, le modèle datamining DMDE 2026 va permettre d'identifier les situations d'allocataires les plus exposées à la présence d'un indu de plus de 600 € sur une période de 6 mois.

Pour mener ce projet, la Cnaf a fait le choix d'une démarche éthique dès la conception : elle encadre l'usage des données, renforce la transparence, prévoit un suivi des effets du modèle et vise à limiter les risques de discrimination en s'appuyant sur le comité d'éthique.

### Pourquoi les Caf utilisent-elles le datamining ?

Chaque mois, les 101 Caisses d'allocations familiales versent des prestations à 13,8 millions de foyers, couvrant 32 millions de personnes, soit presqu'un habitant sur deux. En 2024, les montants versés ont représenté plus de 108,6 milliards d'euros.

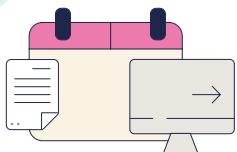
La mission des Caf est également de veiller au juste droit, c'est-à-dire que le montant adapté à la situation de l'allocataire soit versé au bon moment, en respectant l'égalité de traitement entre les allocataires.

Les Caf réalisent donc des contrôles (31,5 millions en 2024) pour vérifier l'exactitude des dossiers, détecter les erreurs, informer et conseiller l'allocataire sur ses droits et repérer les fraudes. En effet, les prestations reposant en partie sur les déclarations de ressources, de situation familiale, de logement..., avec une réglementation parfois complexe, des erreurs peuvent se produire.

En fonction des résultats du contrôle, la Caf peut ajuster les prestations versées. Cela peut entraîner des rappels de droits non perçus ou des demandes de remboursement en cas de trop-perçus (indus).

Les contrôles peuvent être automatiques (via l'échange de données avec les autres services publics), sur pièces (en demandant des pièces justificatives auprès de l'allocataire), ou sur place, au domicile ou à l'accueil de la Caf.

→ Plus d'infos sur Caf.fr : [Tout savoir sur les contrôles à la Caf](#)



## Comment les Caf identifient-elles les dossiers à contrôler ?

La Caf peut s'appuyer sur des incohérences détectées lors du traitement du dossier ou à la suite du signalement d'un partenaire, d'un tiers ou encore d'un contrôle aléatoire..

Pour moins de 1 % de l'ensemble des contrôles réalisés par la Cnaf et les Caf ont recours au modèle datamining pour identifier les dossiers les plus exposés à un risque d'indu élevé.

Le contrôleur de la Caf s'appuie sur ces « scores de risque », parmi d'autres informations, pour se concentrer sur les dossiers les plus susceptibles de comporter des erreurs. **Cependant, la décision d'un contrôle reste humaine** : le contrôleur peut aussi tenir compte d'autres informations et de contraintes opérationnelles (l'organisation des tournées de contrôles sur place par exemple).

→ [Plus d'infos sur Caf.fr : Contrôles et datamining à la Caf - Des prestations au juste droit pour tous](#)



**Thomas Desmoulins, en charge du département prévention, contrôle et lutte contre la fraude de la Cnaf :**

*« Les dossiers ciblés via l'algorithme représentent moins d'1 % des 31,51 millions de contrôles réalisés en 2024, et 6,5 % des 2,5 millions de contrôles sur pièces ou sur place, réalisés par les contrôleurs, soit une part minoritaire. »*

## Des choix concrets pour réduire les risques de biais

Conformément à la réglementation lors de la conception de ce nouveau modèle de datamining, certaines variables ont été écartées pour limiter les risques de discrimination directe ou indirecte : nationalité, genre, adresse, lieu de résidence.

La situation économique de l'allocataire étant inévitable, elle est intégrée au modèle datamining.

La démarche exclut également des données de « comportement » (connexions à l'espace « Mon compte », contacts, échanges de pièces...) ainsi que les informations relatives à des contrôles ou contentieux antérieurs.



**Lucie Gonzalez, directrice des statistiques, des études et de la recherche de la Cnaf :**

*« Le modèle conserve une méthode statistique de type régression logistique, choisie pour sa lisibilité et son explicabilité (impact de chaque variable via des coefficients). Les données utilisées sont issues du système d'information de la CNAF, et du dispositif de ressources mensuelles. »*

## La solidarité à la source : un modèle adapté à un contexte déclaratif qui change

La solidarité à la source, généralisée depuis mars 2025, simplifie la déclaration trimestrielle des ressources pour le versement du Rsa et de la prime d'activité via des déclarations pré-remplies.

Le DMDE 2026 tient compte de ce changement et prévoit une articulation de deux modèles sur la période de transition : un modèle pour les périodes avant pré-remplissage et un modèle pour les périodes après pré-remplissage, avec une cohabitation temporaire jusqu'à la disparition progressive du modèle « avant » (horizon 2027).

→ [Plus d'infos sur Caf.fr : La Caf innove pour simplifier l'accès aux droits](#)

## Quels critères (variables) le DMDE 2026 prend-il en compte ?

Le DMDE 2026 comprend deux modèles pendant la période de transition (avant/après solidarité à la source). Les listes ci-dessous présentent les critères utilisés par chacun.

### Modèle 1 (périodes avant solidarité à la source) – 17 variables

- Situation d'activité du responsable de dossier et du conjoint (dont présence d'un conjoint sans activité)
- Présence d'au moins un enfant âgé de plus de 19 ans dans le foyer
- Activité(s) des enfants (dont enfant étudiant, enfant sans activité, etc.)
- Présence d'une tutelle morale (dossier géré par une tutelle)
- Montant moyen de prestations perçues sur les 12 derniers mois
- Montant moyen d'allocations familiales perçues sur les 12 derniers mois
- Montant moyen de complément familial perçu sur les 12 derniers mois
- Montant moyen de prime d'activité perçue sur les 12 derniers mois
- Montant moyen de RSA perçu sur les 12 derniers mois
- Niveau de revenu (par rapport au Smic)
- Présence d'abattement ou de neutralisation appliquée sur le dossier (24 derniers mois)
- Présence de pension alimentaire dans les déclarations trimestrielles (24 derniers mois)
- Changements dans la pratique d'une activité non salariée (24 derniers mois)
- Fait génératrice de confirmation de logement étudiant
- Projet d'accompagnement – RSA droits et devoirs
- Fait génératrice volet social lié au Complément mode de garde (CMG)
- Nombre de changements de droit au RSA (aller-retour dans la prestation – 24 derniers mois)

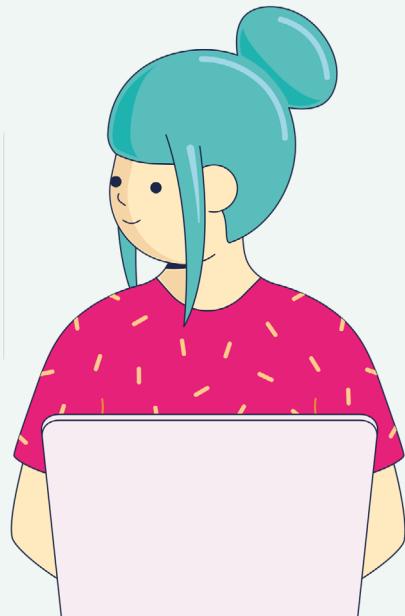
## Modèle 2 (périodes après solidarité à la source) – 11 variables

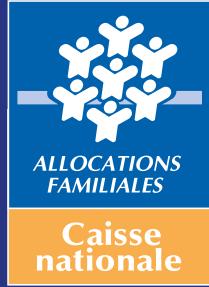
- Situation d'activité du responsable de dossier et du conjoint (dont présence d'un conjoint sans activité)
- Présence d'au moins un enfant âgé de plus de 19 ans dans le foyer
- Présence d'une tutelle morale (dossier géré par une tutelle)
- Montant moyen de prestations perçues sur les 12 derniers mois
- Montant moyen de prime d'activité perçue sur les 12 derniers mois
- Niveau de revenu (par rapport au Smic)
- Pratique d'une activité non salariée (24 derniers mois)
- Signalements : corrections à la baisse des montants pré-affichés dans la déclaration trimestrielle (selon fréquence et montant)
- Fait générateur de confirmation de logement étudiant
- Fait générateur volet social lié au Complément mode de garde (CMG)
- Nombre de changements de droit au RSA (aller-retour dans la prestation – 24 derniers mois)

## Transparence : mise à disposition en open source du code informatique du DMDE 2026

Dans la continuité du travail de transparence engagé avec le comité d'éthique, la Cnaf met à disposition le [code source du DMDE 2026](#). Cette publication permet de consulter :

- la liste des variables utilisées par l'algorithme ;
  - les coefficients/pondérations associés à chacune ;
  - la manière dont ces éléments sont combinés pour produire un score de risque.
- Plus d'infos sur Caf.fr : [Transparence et garanties éthiques autour de l'algorithme](#)





Caisse nationale des Allocations familiales  
32 avenue de la Sibelle - 75014 Paris

 @cnaf\_actus

 [facebook.com/cafallocationsfamiliales](https://facebook.com/cafallocationsfamiliales)

 [youtube.com/c/AllocationsfamilialesFrance](https://youtube.com/c/AllocationsfamilialesFrance)

 [linkedin.com/company/caf-allocations-familiales](https://linkedin.com/company/caf-allocations-familiales)

