

Commission nationale de l'informatique et des libertés

Délibération n° 2014-045 du 30 janvier 2014 portant avis sur un projet d'arrêté portant création par la direction générale des finances publiques d'un traitement automatisé de lutte contre la fraude dénommé « ciblage de la fraude et valorisation des requêtes » (demande d'avis n° 1726052)

NOR : CNIX1405066X

La Commission nationale de l'informatique et des libertés,

Saisie par le ministre délégué auprès du ministre de l'économie et des finances, chargé du budget, d'une demande d'avis concernant un projet d'arrêté portant création par la direction générale des finances publiques d'un traitement automatisé de lutte contre la fraude ;

Vu la convention n° 108 du Conseil de l'Europe pour la protection des personnes à l'égard du traitement automatisé des données à caractère personnel ;

Vu la directive 95/46/CE du Parlement européen et du Conseil du 24 octobre 1995 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données ;

Vu le livre des procédures fiscales, notamment ses articles L. 10, L. 81 et L. 229 à L. 231 ;

Vu le code pénal, notamment ses articles 313-1 et suivants ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés, notamment son article 26-1 (2°) ;

Vu le décret n° 2005-1309 du 20 octobre 2005 modifié pris pour l'application de la loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés ;

Vu l'arrêté du 4 novembre 2013 portant création par la direction générale des finances publiques d'un entrepôt de données dédié aux études fiscales nationales et internationales dénommé « EDEN » ;

Vu la délibération n° 2013-302 du 15 octobre 2013 portant avis sur un projet d'arrêté portant création d'un entrepôt de données dédié aux études fiscales nationales et internationales dénommé « EDEN » ;

Après avoir entendu M. Didier GASSE, commissaire, en son rapport et M. Jean-Alexandre SILVY, commissaire du Gouvernement, en ses observations,

Emet l'avis suivant :

La commission a été saisie par le ministre délégué auprès du ministre de l'économie et des finances, chargé du budget, d'une demande d'avis relative à un projet d'arrêté portant création par la direction générale des finances publiques (DGFIP) d'un traitement automatisé de lutte contre la fraude.

La création de ce traitement, dénommé « ciblage de la fraude et valorisation des requêtes » (CFVR) et basé sur du datamining, s'inscrit dans un contexte global de lutte contre la fraude fiscale. Il utilisera les données issues de onze traitements différents et innovera en ce qu'il utilisera une technique dite de « datamining » consistant en une fouille approfondie des données s'appuyant sur des méthodes exploratoires basées sur la statistique et des algorithmes et permettant de modéliser des comportements.

Dans la mesure où le traitement projeté a notamment pour objet la prévention, la recherche, la constatation et la poursuite d'infractions pénales, sa mise en œuvre doit être autorisée par arrêté ministériel, pris après avis motivé et publié de la commission, en application des dispositions de l'article 26-1 (2°) de la loi du 6 janvier 1978 modifiée.

Dans ce contexte de priorité donnée à la fraude fiscale, la présente demande fait suite à la création par l'arrêté du 4 novembre 2013, publié au *Journal officiel* du 12 décembre 2013, d'un entrepôt de données dédié aux études fiscales dénommé « EDEN » réunissant les données de 19 traitements différents. Il doit permettre, par l'amélioration des capacités d'analyse de l'administration fiscale, de mieux identifier les situations potentiellement frauduleuses en mettant en évidence des incohérences ou des défaillances déclaratives dans les dossiers des contribuables, améliorant ainsi le ciblage des contrôles à la suite de requêtes appropriées.

Il s'agit d'une expérimentation dont la durée devrait être de six mois.

Sur la dénomination du traitement

Le projet d'arrêté ne procède ni dans son titre ni dans ses différentes dispositions à la dénomination du traitement. Pour des raisons de commodité évidentes, la commission demande que le traitement soit effectivement dénommé conformément à l'article 29 de la loi de 1978.

Sur les finalités du traitement

Aux termes de l'article 2 du projet d'arrêté, le traitement projeté permet « *une modélisation des comportements frauduleux, qui se base notamment sur les caractéristiques des cas de fraudes avérées, afin de mener des actions de prévention, de recherche, de constatation ou de poursuite d'infractions pénales ainsi que des opérations de constatation ou de poursuite de manquements fiscaux* ».

La finalité du traitement projeté est légitime dans la mesure où les éléments qui en seront issus n'auront qu'une valeur de signalement parmi d'autres à la disposition des services fiscaux et ne conduiront en aucun cas à une programmation automatique des contrôles ni *a fortiori* à des décisions de redressement directement opposables aux contribuables.

Ce traitement sera mis en œuvre par la nouvelle mission dénommée « Requêtes et valorisation », rattachée au service du contrôle fiscal et au service des systèmes d'information de la DGFIP, créée dans cette perspective de rationalisation du ciblage et de valorisation des requêtes effectuées sur l'ensemble des fichiers fiscaux.

Plus précisément, le mode de fonctionnement de l'outil « CFVR » résulte d'un programme se basant sur des équations mathématiques de probabilité. En aucun cas, les calculs opérés ne constitueront le résultat final du traitement. Ceci constitue une garantie selon laquelle l'outil constitue bien une aide au ciblage et à l'optimisation de la détection de la fraude et non pas un outil de profiling destiné à identifier directement des fraudeurs potentiels.

Ainsi, les éléments qui en seront issus n'auront qu'une valeur de signalement parmi d'autres à la disposition des services fiscaux et ne conduiront en aucun cas à une programmation automatique des contrôles ni *a fortiori* à des décisions de redressement directement opposables aux contribuables.

A cet égard, la commission prend acte que le traitement n'a pas vocation à contenir de listes nominatives de personnes à contrôler, mais doit uniquement permettre la modélisation de comportement frauduleux. Ces listes devront alors être présentées aux services de la DGFIP qui valoriseront, confirmeront ou infirmeront le risque, avant de prendre toute décision de contrôle, contribuant ainsi à préciser les méthodologies existantes.

Dans cette mesure, le traitement est conforme aux dispositions de l'article 10 de la loi du 6 janvier 1978 modifiée.

La réalisation du prototype, pour permettre une analyse ultérieure objective des données, devra reprendre les principales étapes du projet global qui sont les suivantes :

- inventaire des données nécessaires ;
- l'examen de la qualité de ces données ;
- la recherche de critères de départ pour qualifier les fraudes ;
- le périmètre modélisation sur la base de ces données et de ces critères, l'examen du résultat sous un angle métier ;
- l'itération avec une possible modélisation enrichie soit des données, soit de critères afin de sortir un résultat lisible, explicable et utile.

Ce projet expérimental se déroulera sur une période de six mois.

Ces travaux expérimentaux devraient permettre ainsi de dégager des secteurs à risque et de comprendre l'évolution de certaines données fiscales. En effet, la détection de fraudes récurrentes a vocation, dans une démarche de datamining, à générer des indicateurs pour valoriser les applicatifs existants. Plusieurs essais seront donc nécessaires pour améliorer la modélisation de l'outil, qui a lui-même vocation à améliorer la performance des autres outils existants, tel qu'EDEN.

En raison des enjeux soulevés par le datamining, la combinaison et le croisement croissant de différents outils de détection et de lutte contre la fraude, ainsi que le caractère innovant de l'outil présenté, la commission estime que des garanties fortes, notamment au travers des mesures de sécurité, doivent être prises afin d'assurer la protection des données à caractère personnel, en particulier en ce qui concerne le nombre des agents habilités à émettre des requêtes et celui des destinataires.

Compte tenu de ces observations, la commission considère que la finalité du traitement est déterminée, explicite et légitime, conformément aux dispositions de l'article 6 (2°) de la loi du 6 janvier 1978 modifiée.

Sur le caractère expérimental du traitement

La commission demande que ce caractère soit précisé dans le texte de l'arrêté et que la DGFIP lui transmette le bilan de cette expérimentation avant toute généralisation du traitement, qui devrait elle-même faire l'objet d'un nouvel arrêté.

Dans la mesure où la commission ne dispose d'aucun élément sur le fonctionnement du traitement lui-même, le bilan transmis devra comporter :

- une description des conditions de mise en œuvre technique et opérationnelle des phases du prototype ;
- des éléments chiffrés sur le contenu de la base, sur le nombre de consultations, sur les conditions de mises à jour des données et sur les demandes de droits d'accès indirect ;
- une description précise des différentes interconnexions d'EDEN avec d'autres traitements et des modalités d'échanges avec ces autres traitements ;
- des éléments de conclusion généraux relatifs au fonctionnement du traitement, aux éventuelles difficultés rencontrées, aussi bien juridiques que techniques.

Sur la nature des données traitées

Aux termes de l'article 3-I du projet d'arrêté, les données à caractère personnel traitées relèvent de quatre catégories :

- identification des personnes physiques et éléments de situation professionnelle et économique ;
- identification des entreprises et éléments de situation professionnel et économique ;
- informations d'ordre économique et financier des entreprises ;
- résultats obtenus par le traitement.

Les trois premières catégories concernant l'identification se réfèrent aux données issues des onze traitements suivants : ADELIE, MEDOC, FNDP, Obligation déclarative des domiciliants, BODACC, REBECA, TSE, SIR, ALPAGE, COMPAS (présence d'un compte bancaire à l'étranger - O/N) et SIRIUS-PRO.

Comme pour EDEN, la combinaison de données personnelles et de données professionnelles est justifiée par le fait qu'une analyse conjointe facilite la détection de montages fiscaux frauduleux.

La quatrième catégorie de données concerne les résultats du traitement et en fait la spécificité par rapport à EDEN. C'est la catégorie la plus sensible avec le numéro SIREN, le numéro SIRET, la dénomination, l'adresse et les critères à risques.

Au regard de ces éléments, la commission considère que ces catégories de données sont adéquates, pertinentes et non excessives au regard de la finalité poursuivie, conformément aux dispositions de l'article 6 (3°) de la loi du 6 janvier 1978 modifiée.

Toutefois, face au volume important de données destinées à être traitées, la commission souhaite que des garanties fortes soient apportées quant à l'utilisation qui pourra être faite des données à caractère personnel par l'outil « CFVR ».

Sur la durée de conservation des données

Aux termes de l'article 4 du projet d'arrêté, les données seront conservées dans le traitement pendant une durée de six mois, c'est-à-dire le temps de l'expérimentation.

Cette durée de conservation est adaptée au regard de la finalité poursuivie par le traitement et de la nature expérimentale de ce dernier.

Sur les destinataires des données

L'article 5 du projet d'arrêté prévoit que « les informations traitées seront consultables sur un serveur de valorisation de la DGFIP par les personnels habilités de la mission "Requêtes et valorisation" ».

Il est également précisé que seuls les agents des services – au maximum sept personnes – de la DGFIP chargés de la recherche et du contrôle fiscal sont destinataires des résultats obtenus par la mission « Requêtes et valorisation ».

La commission estime que la limitation du nombre de destinataires du traitement constitue une garantie forte du respect de la protection des données des personnes concernées et qu'il convient de ne pas y déroger.

Sur les droits des personnes concernées

L'article 6 du projet d'arrêté est consacré aux droits des personnes concernées par le traitement. En application du VI de l'article 32 de la loi de 1978, le droit d'information ne s'applique pas à ce traitement, puisque ce dernier a notamment pour objet la prévention, la recherche, la constatation ou la poursuite d'infractions pénales.

En application de l'article 38 de la même loi, le droit d'opposition ne s'applique pas davantage.

Enfin, l'arrêté prévoit que les droits d'accès et de rectification s'exerceront de manière indirecte, dans les conditions prévues par l'article 42 de la loi de 1978.

En effet, permettre à toute personne d'obtenir l'ensemble des informations la concernant contenues dans ce traitement, et notamment les critères de fraude retenus, compromettrait l'objectif recherché, car les entités en ayant connaissance pourraient organiser leur fraude en fonction de ces informations.

Sur la sécurité des données et la traçabilité des actions

La commission estime que la dimension importante du traitement projeté, la nature des données traitées ainsi que la volonté d'automatiser la détection de la fraude sont de nature à accroître les risques en termes de protection des droits des personnes concernées. Ces risques appellent des précautions renforcées en matière de sécurité du traitement.

A cet égard, la commission note que les échanges de données sont réalisés au moyen de canaux sécurisés et, notamment, que les données transmises sont chiffrées.

De plus, la commission relève que des profils d'habilitation pour accéder au traitement ont été définis et que l'authentification des utilisateurs est basée sur des mots de passe dont la complexité est conforme aux préconisations de la CNIL.

La commission relève qu'une fonctionnalité de journalisation des opérations de consultation, création, mise à jour et suppression a été définie, étant précisé que les données journalisées sont supprimées au bout de quatre ans (conservation pendant un an en ligne puis pendant trois ans en archives).

Les autres mesures de sécurité décrites par le responsable de traitement sont conformes à l'exigence de sécurité prévue par l'article 34 de la loi du 6 janvier 1978 modifiée.

La commission rappelle toutefois que cette obligation nécessite la mise à jour des mesures de sécurité au regard de la réévaluation régulière des risques.

Pour la présidente :
Le vice-président délégué,
E. DE GIVRY