

Commission nationale de l'informatique et des libertés

Délibération n° 2020-124 du 10 décembre 2020 portant avis sur un projet de décret portant modalités de mise en œuvre par la direction générale des finances publiques et la direction générale des douanes et droits indirects de traitements informatisés et automatisés permettant la collecte et l'exploitation de données rendues publiques sur les sites internet des opérateurs de plateformes en ligne (demandes d'avis n° 2218895 et 2218896)

NOR : CNIX2102956V

La Commission nationale de l'informatique et des libertés,

Saisie par le ministère de l'économie, des finances et de la relance d'une demande d'avis concernant un projet de décret portant modalités de mise en œuvre par la direction générale des finances publiques et la direction générale des douanes et droits indirects de traitements informatisés et automatisés permettant la collecte et l'exploitation de données rendues publiques sur les sites internet des opérateurs de plateformes en ligne ;

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 le règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE ;

Vu la directive (UE) 2016/680 du Parlement européen et du Conseil du 27 avril 2016 relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel sur les autorités compétentes à des fins de prévention et de détection des infractions pénales, d'enquêtes et de poursuites en la matière ou d'exécutions de sanctions pénales, et à la libre circulation de ces données, et abrogeant la décision-cadre 2008/977/JAI du Conseil ;

Vu la loi n° 78-17 du 6 janvier 1978 modifiée relative à l'informatique, aux fichiers et aux libertés ;

Vu la loi n° 2019-1479 du 28 décembre 2019 de finances pour 2020 notamment son article 154 ;

Vu le décret n° 2019-536 du 29 mai 2019 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 la décision n° 2019-796 DC du 27 décembre 2019 du Conseil Constitutionnel ;

Vu la délibération n° 2019-114 du 12 septembre 2019 portant avis sur le projet d'article 9 du projet de loi de finances pour 2020 ;

Après avoir entendu M. Philippe-Pierre CABOURDIN, commissaire en son rapport, et M. Benjamin TOUZANNE, commissaire du Gouvernement, en ses observations,

Emet l'avis suivant :

L'article 154 de la loi n° 2019-1479 du 28 décembre 2019 de finances pour 2020 autorise, à titre expérimental pour une durée de trois ans, les administrations fiscales et douanières à collecter et exploiter les contenus librement accessibles et manifestement rendus publics par les utilisateurs sur les sites internet des opérateurs de plateforme en ligne mentionnés au 2° du I de l'article L. 111-7 du code de la consommation (ci-après les « plateformes et réseaux sociaux »). Cette collecte doit permettre de rechercher des indices relatifs à la commission de certaines infractions limitativement énumérées par la loi.

La Commission rappelle qu'elle s'est déjà prononcée sur le dispositif de collecte envisagé dans sa délibération n° 2019-114 du 12 septembre 2019 susvisée.

La Commission relève que le législateur a depuis, fixé certaines caractéristiques des traitements projetés, notamment les durées de conservation, les modalités d'exercice des droits d'accès et d'opposition ainsi que certaines des garanties à mettre en œuvre telles que la transmission d'une analyse d'impact relative à la protection des données à caractère personnel (AIPD) ou l'interdiction de recourir à un système de reconnaissance faciale.

Si, dans ce contexte, la Commission n'entend dès lors pas revenir dans le détail sur ces différents points, elle rappelle qu'il est expressément prévu que l'expérimentation devra faire l'objet d'une première évaluation dont les résultats lui seront transmis ainsi qu'au Parlement au plus tard dix-huit mois avant son terme d'une part, et qu'un bilan définitif devra également leur être transmis six mois avant son terme d'autre part. Elle rappelle qu'elle sera particulièrement attentive aux contenus des documents transmis ainsi qu'aux suites que le ministère entendra donner à l'expérimentation menée. La Commission renvoie à sa délibération n° 2019-114 susvisée s'agissant des informations qu'elle souhaite *a minima* lui être communiquées.

Sur la présentation d'ensemble du dispositif prévu à l'article 154 de la loi de finances pour 2020

De manière générale, le dispositif mis en œuvre sera composé de deux phases distinctes pour chacune des administrations concernées : une phase d'apprentissage et de conception suivie d'une phase d'exploitation des données.

Les articles 5 et 6 du projet de décret se rapportent à la phase d'apprentissage et de conception du dispositif. Les traitements de données à caractère personnel mis en œuvre doivent permettre de développer des outils de collecte et d'analyse des données (tels que des outils d'identification de personnes titulaires de comptes, d'identification des données de localisation ou des outils permettant la suppression des données), et d'élaborer des « critères de pertinence » ou « indicateurs » susceptibles de caractériser des manquements et infractions recherchés.

Les articles 7 à 9 du projet de décret se rapportent à la phase d'exploitation des données. Il s'agit pour les administrations concernées, à partir des indicateurs précédemment élaborés et des outils de collecte et d'analyse développés, de collecter des contenus susceptibles de révéler un manquement ou une infraction en matière fiscale ou douanière. La Commission rappelle que les données ainsi collectées par le dispositif de collecte ne pourront donner lieu à aucune décision automatisée des administrations fiscales et douanières à l'encontre des personnes concernées d'une part et que ces données, si elles révèlent par la suite un éventuel comportement frauduleux, ne pourront être utilisées que dans le respect des règles relatives à la procédure de contrôle mentionnée au titre II du code des douanes ou au chapitre premier du titre II de la première partie du livre des procédures fiscales, d'autre part.

La Commission relève que les données ainsi collectées seront transmises aux traitements « ciblage de la fraude et valorisation des requêtes » (CFVR) de la Direction générale des finances publiques (DGFIP) et « valorisation des données pour l'analyse de risque » de la Direction générale des douanes et droits indirects (DGDDI). Cette transmission doit permettre de déterminer si les données collectées constituent des indices caractérisant un manquement ou une infraction en matière fiscale ou douanière.

Dans la mesure où cette transmission de données conduit notamment à l'enregistrement de nouvelles données dans les traitements précités CFVR et « valorisation des données pour l'analyse de risque », la Commission a également été saisie pour avis, de manière concomitante au présent projet de décret, des projets d'actes réglementaires modifiés encadrant ces deux traitements. Les modifications de ces traitements font l'objet de deux délibérations distinctes du même jour.

Ces éléments généraux rappelés, la Commission entend formuler les observations suivantes.

Sur les finalités et le régime juridique applicable

En premier lieu, la Commission relève que le dispositif prévu à l'article 154 de la loi de finances repose, en pratique, sur la mise en œuvre de plusieurs traitements de données à caractère personnel correspondant de manière générale aux deux phases présentées précédemment.

L'article 4 du projet de décret précise à cet égard que « *les traitements* » autorisés sont mis en œuvre par les administrations concernées, en vue :

« a) *de la collecte des données et de la sélection des données pertinentes durant la phase prévue aux articles 5 et 6 et au I des articles 7, 8 et 9 ;*

b) *du transfert de données pour les traitements visés au II des articles 7, 8 et 9. »*

La Commission considère qu'une telle rédaction vise davantage à expliciter les grandes fonctionnalités des traitements mis en œuvre que leurs finalités. A ce titre, elle observe que d'autres dispositions du projet de décret qui lui est soumis apportent davantage de précisions sur les finalités effectivement poursuivies par les traitements mis en œuvre lors de chacune des phases du dispositif. Sur ce point, la Commission prend acte de l'engagement du ministère de modifier l'article 4 du projet de décret afin de faire apparaître de manière explicite les finalités des différents traitements mis en œuvre.

En second lieu, la Commission relève que le ministère entend faire application des seules dispositions de la directive (UE) 2016/680 susvisée (ci-après « la directive »), s'agissant de l'ensemble du dispositif expérimental. Elle observe toutefois que les deux phases du dispositif poursuivent en première intention des objectifs distincts dont il convient de tirer toutes les conséquences s'agissant du régime juridique applicable.

La phase d'apprentissage et de conception a en effet uniquement pour objectif primaire de développer des outils techniques et d'établir des indicateurs. Sur ce point, la Commission constate qu'à l'issue de la phase d'apprentissage et de conception, l'ensemble des données à caractère personnel collectées sont supprimées et qu'aucun ciblage de personnes physiques ou morales ou envoi d'informations à un service de contrôle ou de gestion n'est réalisé par les administrations concernées.

Au regard de ces éléments la Commission considère que la phase d'apprentissage et de conception du dispositif devrait relever du régime juridique du règlement européen sur la protection des données (ci-après le « RGPD »). La phase d'exploitation a directement pour finalité la recherche d'infractions fiscales et douanières, elle relève donc bien de la directive. La Commission prend acte que le ministère entend tirer toutes les conséquences de cette distinction et modifier le projet de décret en conséquence.

Sur le périmètre du dispositif et les données traitées

A titre liminaire, la Commission rappelle qu'il revient au ministère de s'assurer de la stricte proportionnalité des données collectées lors de chacune des deux phases du dispositif mis en œuvre.

– Sur les notions de contenus « librement accessibles » et « manifestement rendus publics »

L'article 3 du projet de décret rappelle que, conformément à l'article 154 de la loi de finances pour 2020, seuls les contenus librement accessibles et manifestement rendus publics sur les sites internet des opérateurs de plateforme en ligne pourront être collectés.

Dans sa décision n° 2019-796 DC du 27 décembre 2019, le Conseil constitutionnel a précisé que « [...], les données susceptibles d'être collectées et exploitées doivent répondre à deux conditions cumulatives. D'une part, il doit s'agir de contenus librement accessibles sur un service de communication au public en ligne d'une des plateformes précitées, à l'exclusion donc des contenus accessibles seulement après saisie d'un mot de passe ou après inscription sur le site en cause. D'autre part, ces contenus doivent être manifestement rendus publics par les utilisateurs de ces sites. Il en résulte que ne peuvent être collectés et exploités que les contenus se rapportant à la personne qui les a, délibérément, divulgués. [...] ».

La Commission considère tout d'abord qu'il résulte des termes mêmes de cette décision que les « contenus librement accessibles » doivent être entendus comme les contenus auxquels un utilisateur non inscrit ou sans enrôlement préalable (création de compte, fourniture de certaines informations pour créer un identifiant ou toute autre forme d'inscription) sur une plateforme ou un réseau social pourrait avoir accès, sans saisie préalable d'un mot de passe. A cet égard, elle relève d'emblée que le recours à des identités d'emprunts ou la création de comptes spécialement créés à cet effet par les administrations fiscale et douanière est bien exclue par le présent projet de décret.

La Commission relève qu'il ressort de la documentation transmise que le ministère entend utiliser des API (interfaces de mise à disposition des données des sites) proposées par les plateformes ou les réseaux sociaux, et/ou des techniques de « *webscraping* » (techniques d'extraction du contenu de sites, via des scripts ou des programmes automatisés) pour collecter les données des plateformes et des réseaux sociaux. Elle observe que ces deux modalités de collecte nécessitent toutefois de se créer un compte développeur dans le cas des API ou un compte utilisateur pour le « webscraping ».

Interrogé sur ce point, le ministère a indiqué que les « contenus librement accessibles » désignent, selon son interprétation, les données publiées sur les plateformes et les réseaux sociaux sans paramètre de confidentialité spécifique ou avec un paramétrage de confidentialité public, à savoir les données qui ne sont pas publiées en mode privé ou en accès restreint à un cercle de contacts, quel que soit les modalités techniques utilisées pour les collecter.

La Commission considère que si le ministère peut utiliser des API, en se créant un compte développeur, les contenus accessibles *via* ces API doivent être limités à ceux accessibles par un utilisateur qui ne disposerait pas de compte sur la plateforme et sans saisie préalable d'un mot de passe. Les sites ou plateformes nécessitant une inscription telle que définie précédemment ou un mot de passe ne peuvent donc être accessibles conformément à la loi telle qu'interprétée par le Conseil constitutionnel. Le décret devra être modifié en conséquence.

En ce qui concerne la notion de contenus manifestement rendus publics, la Commission relève que le projet de décret permet, dans certaines conditions, la collecte de commentaires rédigés par des tiers.

Interrogé sur le caractère manifestement rendu public des commentaires rédigés par des tiers sur la page personnelle d'un individu, le ministère a fait valoir le caractère public – par nature – des commentaires publiés sur les sites marchands au regard de leur modèle économique ainsi que la connaissance, par les utilisateurs de ces sites marchands, des paramètres de confidentialité retenus.

S'agissant des réseaux sociaux, le ministère estime que les utilisateurs ont également la possibilité de configurer les paramètres de confidentialité de leurs pages, en l'absence de paramétrage spécifique, de suppression des contenus concernées ou de signalement effectué auprès de la plateforme concernée, l'utilisateur les divulguent délibérément.

La Commission estime néanmoins, au regard de la décision du Conseil constitutionnel précitée, que pour être manifestement rendus publics les contenus doivent être délibérément divulgués par la personne titulaire du compte ou de la page ce qui implique incontestablement une action volontaire de sa part. *A contrario*, elle estime que la simple absence de mise en place d'un paramétrage de confidentialité spécifique par exemple ne suffit pas à caractériser qu'une personne a délibérément divulgué un contenu.

Compte tenu de ce qui précède, la Commission demande que les commentaires des tiers ne soient pas collectés dans le cadre du dispositif envisagé. Elle insiste également pour que le décret ne soit pris que lorsque le ministère aura identifié des moyens techniques permettant de limiter sa collecte aux seules données se rapportant à la personne concernée et à elle seule, délibérément divulguées pour les rendre manifestement publiques au sens des dispositions applicables.

En ce qui concerne les données collectées lors des phases d'apprentissage, de conception et d'exploitation, elle relève de manière générale que des garanties ont été mises en place. En particulier, il est prévu que les indicateurs devant permettre de caractériser une probabilité de fraude, qui seront d'abord élaborés automatiquement par des algorithmes, fassent ensuite l'objet d'une analyse humaine afin d'écarter ceux qui impliqueraient la collecte de données sensibles qui malgré les filtres prévus auraient été collectées ou le développement d'outils techniques permettant la suppression automatique de telles données. De la même manière, le recours à un système de reconnaissance faciale visant à identifier des personnes à partir des photographies mises en ligne est exclu par le décret. Si des précisions ont été apportées par le ministère sur ce que recoupe la catégorie des données relative aux « *contenus de toute nature, y compris diffusés en temps réel* », qui indique que cela recouvre principalement les « *hashtag* » et l'ensemble des publications quel que soit leur format informatique (par exemple des codes chiffrés, des algorithmes, ...) dès lors qu'ils sont librement accessibles et manifestement rendus publics par l'utilisateur de la plateforme, elle estime toutefois que cette formulation mériterait d'être précisée dans le projet de décret.

– Sur les données collectées durant la phase d'apprentissage et de conception

La Commission observe que durant la phase d'apprentissage et de conception, l'élaboration des différents outils permettant notamment de collecter, d'analyser ou encore de créer des indicateurs à partir des contenus librement accessibles et manifestement rendus publics seront développés à partir d'un échantillon de données.

Elle relève que pour l'administration fiscale, l'échantillon de données est composé par des données d'identification d'entreprises préalablement sélectionnées (pour la recherche d'activité occulte) et des données d'identification de personnes physiques également préalablement identifiées (pour la recherche de fausse domiciliation à l'étranger). Cet échantillon sera par ailleurs limité à environ une centaine d'entreprises pour la recherche d'activité occulte et à une liste réduite de personnes pour la fausse domiciliation à l'étranger (une dizaine de personnes). Des garanties ont été mises en place s'agissant de la collecte des contenus (écrits, images, photographies, sons, signaux, vidéos, ou contenus de toute nature, y compris diffusés en temps réel) au travers par exemple d'une durée de conservation limitée des données ou encore de l'absence de ciblage des individus à partir de ces dernières.

Par ailleurs la Commission prend acte que l'ampleur de l'échantillon constitué par l'administration des douanes et droits indirects sera limité à ce qui est strictement nécessaire pour développer des indicateurs et critères de pertinence et que les données seront collectées uniquement sur des pages internet « *en lien avec le tabac* ».

– Sur les données collectées durant la phase d'exploitation

La Commission relève qu'il est prévu, pendant la phase d'exploitation, la possibilité de collecter l'ensemble des contenus énumérés pendant la phase d'apprentissage (écrits, vidéo, photographies, etc.) correspondant aux indicateurs, pour la recherche de l'exercice d'une activité occulte ou les délits douaniers visés par la loi, ou correspondant aux outils d'identification de personnes et les outils d'identification géographique pour la recherche d'infraction aux règles de domiciliation fiscale.

Elle relève ainsi que ce n'est qu'à l'issue de la phase d'apprentissage et de conception, c'est-à-dire lors de l'identification des indicateurs et la configuration des outils de collecte et d'analyse, qu'il sera possible de déterminer précisément les données à caractère personnel traitées dans le cadre de la phase d'exploitation. La Commission considère que, si les indicateurs retenus à l'issue de la phase d'apprentissage et de conception sont raisonnables et efficaces, les données traitées dans le cadre de la phase d'exploitation pourront être considérées comme pertinentes et proportionnées au regard de l'objectif général du traitement, à savoir la lutte contre la fraude. Elle relève néanmoins que le choix du gouvernement de réglementer dans un même décret, dès le départ, les traitements utiles à la conception du dispositif expérimental et ceux correspondant au fonctionnement de l'expérimentation elle-même, conduit à ne pas pouvoir énumérer de façon précise dans l'acte réglementaire les données collectées durant la phase de fonctionnement, alors même que, à juste titre, pour minimiser les données traitées, ne seront collectées que les données correspondant aux indicateurs élaborés.

Indépendamment de ce qui précède, la Commission souligne la nécessité d'encadrer la phase d'exploitation par un acte réglementaire spécifique.

Sur la transmission des données aux traitements « CFVR » et « valorisation des données pour l'analyse de risque »

Les articles 7 à 9 du projet de décret prévoient que certaines données collectées feront l'objet d'un « *rapprochement* » avec les données enregistrées dans les traitements « CFVR » et « valorisation des données pour l'analyse de risque ».

La Commission relève que le « *rapprochement* » projeté constitue en réalité une transmission de ces données suivie d'une comparaison de celles-ci avec les données contenues dans « CFVR » et « valorisation des données pour l'analyse de risque », laquelle est opérée directement au sein de ces traitements-ci. Elle observe que les résultats de ce rapprochement ne sont pas intégrés au dispositif global de collecte de données sur les réseaux sociaux, l'article 4 b du projet de décret faisant uniquement mention du « *transfert de données pour les traitements visés au II des articles 7, 8 et 9* ». La Commission considère ainsi que la terminologie employée aux articles 7 à 9 du projet de décret mériterait d'être explicitée. Sur ce point, la Commission prend acte de l'engagement du ministère d'explicitier les opérations de transfert et de comparaison des données dans le projet de décret.

En ce qui concerne les modalités de ce « *rapprochement* », la Commission prend acte des éléments apportés par l'administration fiscale selon lesquels seules les tables d'indicateurs associant des références de comptes (un titulaire de compte) et la présence des indicateurs seront transmises à « CFVR » à l'exclusion des données « *brutes* » collectées sur les plateformes et réseaux sociaux.

En outre, sur le périmètre des données pouvant faire l'objet d'un rapprochement avec le traitement « valorisation des données et analyse des risques », le ministère a précisé que le transfert des données sera strictement limité aux données correspondant aux indicateurs, expurgées des données sensibles et des données qui ne sont pas de nature à concourir à la constatation des infractions et manquements visés dans le cadre de l'article 154 de la loi de finances précitée, et que celles-ci seront par la suite sélectionnées sur la base de scores. La Commission considère que le contrôle effectué sur les données avant leur transmission doit permettre de garantir la proportionnalité de celles-ci au regard de l'objectif poursuivi.

Sur les mesures de sécurité

La Commission relève que des solutions de chiffrement permettant d'assurer un niveau adéquat de protection des données et de respect de leur confidentialité sont mises en œuvre tant pour l'administration fiscale que pour l'administration des douanes et droits indirects.

La Commission observe par ailleurs que des mesures sont mises en place afin de garantir un strict accès aux données et que seules les personnes dûment habilitées et dans la limite du besoin d'en connaître pourront y accéder comme l'indique l'article 11 du projet de décret.

Concernant la traçabilité des actions, la Commission relève que l'article 10 du projet de décret prévoit que les opérations de collecte, de modification, de consultation, de communication, d'interconnexion et d'effacement des données font l'objet d'une journalisation et que ces données de journalisation ne seront consultables que par les personnes habilitées.

Les autres mesures de sécurité n'appellent pas de remarques de la Commission.

La présidente,
M.-L. DENIS