

3

Les données du ministère de l'agriculture et de l'alimentation : un potentiel à mieux exploiter

PRÉSENTATION

Les données sont au cœur de la société de l'information et de la transformation numérique des administrations. Elles sont une composante essentielle des systèmes d'information grâce auxquels, désormais, les politiques publiques sont mises en œuvre.

Depuis la loi pour une République numérique de 2016, les données sont considérées comme un élément central de la démarche d'État-plateforme : l'ouverture des données publiques et la facilitation de toutes leurs utilisations (appelée aussi « valorisation ») sont les deux objectifs principaux de la politique de l'État en matière de données, car la valeur de ces dernières réside exclusivement dans leur usage.

La Cour des comptes a déjà souligné la nécessité d'une gouvernance cohérente de l'ouverture des données publiques de trois opérateurs relevant du ministère de la transition écologique et solidaire, dans un référé adressé au Premier ministre le 11 décembre 2018 intitulé « La valorisation des données de Météo-France, de l'Institut géographique national (IGN) et du Centre d'études et d'expertise sur les risques, l'environnement, la mobilité et l'aménagement (Céréma) : l'enjeu de l'ouverture des données publiques ».

Elle a conduit en 2019 une enquête sur la transformation numérique du ministère de l'agriculture et de trois de ses opérateurs, l'Agence de services et de paiement (ASP), FranceAgriMer (FAM) et le Centre national de la propriété forestière (CNPFF).

Le ministère de l'agriculture et de l'alimentation ainsi que ses opérateurs sont à la fois producteurs et utilisateurs de données (sur les exploitations agricoles et les produits alimentaires, sur les aides agricoles européennes, sur les cultures et les cheptels...). Au cœur des systèmes d'information, ces données sont à la fois des outils au service des politiques publiques et la matière première de nouvelles utilisations (exploitations scientifiques ou commerciales, nouveaux services).

La diversité croissante des données, la forte augmentation de leur volume et la complexité de leurs régimes juridiques imposent au ministère de l'agriculture la mise en œuvre d'une politique cohérente de la donnée qui concilie circulation des données et protection de la vie privée, obligation renforcée par le règlement général sur la protection des données (RGPD) entré en vigueur en 2018.

À l'issue de son enquête, la Cour fait le constat que l'exploitation des données agricoles, produites en grande quantité, est très encadrée (I). Cette exploitation est indispensable pour améliorer l'action du ministère de l'agriculture et de l'alimentation (II). Des progrès restent cependant à accomplir pour pleinement valoriser ces données (III).

I - Des quantités massives de données dont l'exploitation est encadrée par la loi

A - Des données très diverses produites en masse

Le ministère de l'agriculture et de l'alimentation (MAA) est détenteur d'un patrimoine de données¹²⁷ important : le dernier plan de transformation numérique, à l'été 2018, identifiait plus de 300 jeux de données. Les données de la sphère agricole sont collectées et produites par de nombreux acteurs : outre les services du ministère, centraux ou déconcentrés, et les opérateurs du ministère (ASP, FAM, CNPF, Institut national de l'origine et de la qualité (INAO), Office national des forêts (ONF), Institut géographique national (IGN), des données sont également produites par des organismes professionnels (par exemple les filières) et des acteurs de terrain, vétérinaires, vignerons, éleveurs, exploitants agricoles, pêcheurs, etc.

¹²⁷ Les données sont des informations de toute nature : mesures, observations, résultats de calculs, renseignements déclaratifs, statistiques, fréquemment conservées électroniquement sous la forme de bases de données, destinées à être exploitées aux fins de prise de décision, notamment par des traitements automatisés informatiques. On rencontre fréquemment les expressions « *big data* » pour données en masse et « *open data* » pour données publiques.

Les données sont constituées en bases thématiques, riches et souvent complexes, qui relèvent de nombreuses catégories, parmi lesquelles :

- la politique agricole commune (PAC) : bénéficiaires d'aides, catalogue des mesures, montant des aides, indicateurs de résultat, parcelles agricoles et leur couvert, résultats de contrôles sur place ;
- la connaissance des animaux (cheptels, origine, race, âge) ;
- les statistiques relatives aux productions et filières (prix, coûts de production, résultats des entreprises, exportations) ;
- la pêche et l'aquaculture (stocks, captures, production, prix) ;
- la forêt (parcelles, nature et âge du peuplement, dénivelés, production) ;
- les résultats d'inspections ou contrôles sanitaires et vétérinaires.

Ces données ont des usages multiples : production de statistiques annuelles et connaissances scientifiques ; aide à la décision, suivi des marchés et des filières ; évaluation de politique publique ; contrôle et suivi des politiques mises en œuvre (sécurité sanitaire, encadrement des droits de plantation de vignes, contrôles vétérinaires) ; gestion et suivi des aides européennes¹²⁸ et nationales ; obligations liées à des politiques intégrées (pêche, eau, environnement, biodiversité), etc.

Les données : nombre et diversité

Certaines bases de données sont à la fois volumineuses par la quantité d'entrées (données en masse) et diversifiées par les attributs enregistrés pour chaque entrée.

Ainsi, la base de données nationale des identifiants (BDNI) des animaux professionnels compte-t-elle, pour les seuls bovins, plus de 20 millions de bêtes.

Le MAA et ses opérateurs répertorient dans leurs bases de données un grand nombre d'utilisateurs : 7 700 navires de pêche, 15 600 collectivités territoriales ou établissements publics, 83 000 entreprises agro-alimentaires, 560 000 exploitants, 3,5 millions de propriétaires forestiers.

¹²⁸ Les aides européennes relèvent de la politique agricole commune (PAC) et sont financées par le Fonds européen agricole de garantie (FEAGA) et le Fonds européen de développement rural (FEADER). Pour la politique commune de la pêche (PCP), elles relèvent du Fonds européen pour les affaires maritimes et la pêche (FEAMP).

Le registre parcellaire graphique (RPG), base de données consacrée à la gestion des aides de la PAC versées aux agriculteurs en fonction des surfaces cultivées, recense neuf millions de parcelles agricoles, pour chacune desquelles plus d'une dizaine de variables sont actualisées chaque année par l'Agence de services et de paiement. L'une de ces variables, la culture pratiquée, peut être déclinée en 300 valeurs différentes.

Couplées aux systèmes d'informations (SI)¹²⁹, les données constituent à présent le support de mise en œuvre de toutes les politiques publiques du MAA à mesure que la dématérialisation et l'automatisation des procédures progressent.

B - La recherche d'un équilibre entre principe d'ouverture des données et protection des secrets

Le cadre juridique relatif à l'utilisation des données doit concilier le principe d'ouverture¹³⁰ des données publiques, inscrit dans le code des relations entre le public et l'administration (CRPA) en son article L. 311-1 introduit par la loi pour une République numérique du 7 octobre 2016, et le respect de la vie privée et des multiples secrets protégés par la loi¹³¹. La protection des données personnelles a été récemment renforcée par l'entrée en vigueur des dispositions du règlement général sur la protection des données (RGPD). Le RGPD précise les obligations d'information des tiers quant aux données détenues les concernant, leur donne un droit d'accès et de rectification, et impose le principe d'un recueil systématique et renouvelé de leur consentement pour toute réutilisation de ces données.

¹²⁹ Un système d'information est un ensemble organisé de ressources (matériels, logiciels, personnels, données et procédures) qui permet de regrouper, classer, traiter et diffuser de l'information dans un environnement donné.

¹³⁰ Pour mémoire, l'ouverture des données publiques poursuit trois objectifs : renforcer la transparence de l'action administrative et de la vie publique, identifier des leviers d'amélioration de la gestion publique et susciter l'innovation économique.

¹³¹ Sont notamment protégés le secret statistique, le secret de la vie privée, les secrets des affaires, et ne sont communicables qu'à l'intéressé les informations portant une appréciation ou un jugement de valeur sur une personne physique, ou faisant apparaître un comportement dont la divulgation pourrait lui porter préjudice (article L. 311-6 du CRPA).

L'ouverture des données publiques du MAA s'effectue selon les dispositions législatives en vigueur : les administrations et les opérateurs publics doivent publier un registre des informations publiques (RIP) répertoriant toutes les bases de données détenues et ouvertes au public. Si le MAA respecte globalement cette obligation, il n'en va pas systématiquement de même pour les trois opérateurs contrôlés par la Cour.

Les règlements¹³² des aides européennes (FEAMP, FEADER et FEAGA) imposent la publication annuelle, en ligne, des informations suivantes : nom du bénéficiaire d'aides, commune d'enregistrement, montant des paiements des différentes mesures, type et description des mesures. C'est la contrepartie des aides allouées¹³³.

L'ouverture par principe des données publiques agricoles doit cependant s'opérer dans le respect de la protection des secrets, avec des distinctions sensibles selon la nature des données.

L'application des règles de protection du secret par la Commission d'accès aux documents administratifs (CADA)

La protection des secrets (par exemple celui de la vie privée ou celui des affaires) est plus ou moins forte selon la qualification des données concernées.

¹³² Article 111 du règlement (UE) n° 1306/2013 du Parlement et du Conseil du 17 décembre 2013 relatif au financement, à la gestion et au suivi de la politique agricole commune.

¹³³ Les systèmes d'information et de gestion des aides de la PAC (dénommés ISIS et OSIRIS) ont connu des dysfonctionnements majeurs, sources d'importants retards de paiement des aides aux agriculteurs, ces cinq dernières années. Ils ont fait l'objet de plusieurs travaux des juridictions financières : Cour des comptes, *La chaîne de paiement des aides agricoles (2014-2017) : une gestion défaillante, une réforme à mener*, communication à la commission des finances du Sénat, juin 2018 ; *Bilan du transfert aux régions de la gestion des fonds européens structurels et d'investissement (FESI)*, communication à la commission des finances, de l'économie générale et du contrôle budgétaire de l'Assemblée nationale, mai 2019, disponibles sur www.ccomptes.fr. La Cour y reviendra dans un rapport sur la conduite par l'État des grands projets numériques, à remettre à la commission des finances du Sénat en juin 2020.

Ainsi, la CADA a rendu des avis très protecteurs du respect de la vie privée : elle a recommandé que la diffusion du registre parcellaire graphique ne permette pas d'identifier les exploitants des îlots (ensemble de parcelles culturales contiguës) alors que les parcelles et leurs propriétaires sont des données accessibles au public *via* le cadastre¹³⁴. Cet avis rejoint le RGPD selon lequel les données de localisation sont des données personnelles.

Dans d'autres décisions, ne concernant pas le registre parcellaire graphique, elle s'est prononcée pour la publication de documents qui, en principe, ressortissent de la catégorie des informations qui ne sont communicables qu'à l'intéressé. Elle a ainsi estimé que la demande formulée par une association d'avoir communication des résultats d'inspection d'un élevage porcin était recevable, au motif que ces documents relevaient de la catégorie des informations environnementales, donc communicables sans restriction à toute personne en faisant la demande (article L. 124-2 du code de l'environnement) nonobstant les règles de protection de la vie privée¹³⁵.

Les données environnementales et celles relatives aux émissions de substances dans l'environnement sont donc publiables avec peu de restrictions. Dès lors, l'État pourrait améliorer l'information du public en rendant accessibles chaque année ces données de manière organisée et en veillant à les mettre en valeur, par exemple sous une forme cartographique.

La valorisation des données publiques agricoles par des acteurs publics ou privés doit cependant tenir compte des exigences du RGPD. Le respect de l'obligation de recueillir le consentement des tiers à la réutilisation de données les concernant, ainsi que de celle de conserver la trace de ce consentement et de l'authentifier, se heurte pour l'heure à des difficultés techniques qui ne sont que partiellement surmontées et freinent les initiatives en ce sens. C'est la Commission nationale de l'informatique et des libertés (CNIL) qui veille à la bonne application des dispositions issues du RGPD et à l'utilisation conforme des systèmes d'information nominatifs.

Ainsi, afin de pouvoir diffuser les résultats d'inspection de restaurants, en principe couverts par le secret (au titre de l'article L. 311-6 du CRPA relatif aux « informations de nature à porter préjudice aux intéressés »), c'est une disposition législative expresse du code rural qui

¹³⁴ Avis CADA n° 20163985 du 20 octobre 2016 ASP, délégation du Limousin.

¹³⁵ Conseil CADA n° 20172154 du 22 juin 2017, direction départementale de la protection des populations du Finistère.

autorise désormais leur publication aux fins de bonne information du public¹³⁶.

Le MAA et ses opérateurs, comme tous les acteurs publics, sont encore dans une phase de recherche d'un équilibre entre protection de la vie privée et ouverture des données¹³⁷.

II - Exploiter les données, une nécessité pour améliorer l'action du ministère

L'utilisation des données s'articule autour de trois axes : la gestion efficace des données en interne au sein de l'administration dans le cadre des procédures existantes, l'exploitation des données en masse pour innover et améliorer les politiques publiques et, enfin, la mise à disposition auprès de tiers extérieurs à l'administration pour favoriser les réutilisations.

A - Mieux employer les données pour moderniser et simplifier l'administration

L'action du MAA et de ses opérateurs se traduit, pour les usagers, par le recours à environ 400 procédures qui requièrent de nombreuses informations et documents à transmettre par formulaires. La simplification des procédures vise à limiter les informations et pièces demandées et, notamment, à réutiliser celles déjà fournies : c'est la démarche intitulée « Dites-le nous une fois ». L'amélioration de la relation avec les usagers passe aussi par des procédures automatisées en ligne, l'assistance au renseignement des formulaires et des réponses instantanées sur la recevabilité de certaines demandes.

¹³⁶ L'article 45-1 de la loi n° 2014-1170 du 13 octobre 2014 d'avenir pour l'agriculture, l'alimentation et la forêt a introduit une disposition à l'article L. 231-1 du code rural autorisant la publication des données des contrôles vétérinaires selon des modalités définies réglementairement (décret n° 2016-1750 du 15 décembre 2016 organisant la publication des résultats des contrôles officiels en matière de sécurité sanitaire des aliments).

¹³⁷ La CADA et la CNIL, en association avec les services d'Etalab, ont élaboré un guide portant sur l'ouverture des données qui n'est paru qu'en mai 2019 (CNIL, CADA, Etalab, *Guide pratique de la publication en ligne et de la réutilisation des données publiques* (« open data »), *présentation du cadre juridique de l'ouverture des données*).

Cette modernisation repose sur l'utilisation automatique des données au moyen de petits programmes informatiques appelés « interfaces de programmation applicatives » (API)¹³⁸ dans les systèmes d'information. La mise en commun des bases de données publiques, leur accessibilité et le bon fonctionnement des API permettent aux applications de rechercher, de manière autonome, les informations utiles *via* internet.

Identifier et authentifier des usagers : SIRIUS et BACUS

Les usagers du MAA et de ses opérateurs ont des identifiants propres à chaque organisme : un numéro « Numagrin » dans la base de données nationale des usagers (BDNU) du MAA, un numéro dit « PACAGE » dans le SI de la PAC « Telepac » et un numéro SIRET auprès de l'Insee pour les entreprises.

Jusqu'en 2018, ces administrations ne partageaient pas leurs données, ce qui aboutissait pour les usagers à des demandes redondantes. Pour y remédier, le MAA a développé le « système d'indexation et de référencement interne des usagers » (SIRIUS) déployé, en septembre 2018, dans les services d'économie agricole des directions départementales des territoires et de la mer. SIRIUS attribue un identifiant unique aux usagers du MAA et permet, *via* des API (notamment l'« API Entreprises » développée par la DINUM), la mise à disposition automatisée de données du répertoire SIRENE de l'Insee, du répertoire national des associations (RNA), de la Mutualité sociale agricole ainsi que de données fiscales.

En complément, un deuxième système assure l'authentification des tiers et leur inscription : BACUS (« base d'authentification des usagers ») sécurise les authentifications et permet aux usagers de consulter ou modifier leurs informations personnelles, ainsi que de suivre l'avancement de leurs démarches. BACUS n'est toutefois pas encore ouvert aux entreprises car la livraison du module d'identification (« Pro Connect »), développé par la direction interministérielle du numérique (DINUM), se fait attendre.

Les échanges entre les bases de données du ministère, celles de ses opérateurs et celles d'autres administrations (base SIRENE de l'Insee, données fiscales DGFIP par exemple) ont ainsi permis de faire évoluer certaines procédures par l'automatisation du renseignement des formulaires et par l'amélioration de leur ergonomie.

¹³⁸ Une interface de programmation applicative (ou API) est un ensemble de fonctions informatiques par lesquelles deux logiciels interagissent et échangent automatiquement des données par internet en flux continu.

Des applications nomades facilitent les démarches des usagers comme, par exemple, dans l'enseignement agricole où deux applications « ARPENT » permettent les inscriptions en ligne aux examens, le renseignement automatisé des formulaires et la consultation des résultats à distance.

VITIPANTATION : utilisation automatisée de données interministérielles pour simplifier les démarches des viticulteurs

Fruit de la collaboration des services du Premier ministre, de FranceAgriMer, de l'INAO, du MAA, de la direction générale des douanes et droits indirects (DGDDI) et de l'IGN, VITIPANTATION est un « e-service » qui permet aux viticulteurs de faire des demandes d'autorisation de plantation.

Cette téléprocédure fonctionne grâce à des API avec d'autres SI : celui de l'IGN (Géoportail) et le casier viticole informatisé (CVI) de la DGDDI. Elle repose aussi sur l'identification des exploitants, *via* leur numéro SIRET et leur inscription préalable auprès de FAM.

Les déclarations, dématérialisées, bénéficient d'une aide à la saisie en ligne avec un outil cartographique (superposition des parcelles cadastrales et d'un fond cartographique, ainsi que des aires des indications géographiques protégées). L'instruction des demandes est rapide, voire automatique pour la majorité des demandes avec, dans ce cas, une délivrance instantanée des autorisations.

Pour la campagne 2017, plus de 40 000 décisions ont été rendues. Le gain est notable pour les usagers : saisie ergonomique, formulaire pré-rempli, contrôle de cohérence automatique des champs et indication immédiate du résultat de l'instruction pour un certain nombre de dossiers. Le guichet unique permet aux exploitants de retrouver toutes les informations utiles sans démarche physique auprès de différents services. Pour les agents, la charge de travail liée à l'instruction et à la réconciliation des informations est allégée (l'application permet de délivrer de l'ordre de 40 000 autorisations par an depuis la mise en œuvre de l'obligation d'autorisation en 2015). Le développement de l'application a représenté en 2019 une dépense d'investissement de 1,3 M€.

Ces quelques réalisations illustrent bien les évolutions que l'on est en droit d'attendre de l'exploitation efficace des données. Le MAA doit encore progresser : ainsi le principe « Dites-le nous une fois » n'est pas encore appliqué pour toutes les démarches et la simplification de l'ensemble des procédures administratives du MAA reste à venir.

B - Exploiter les données, une source d'innovation pour les politiques publiques

Au-delà des procédures administratives, l'exploitation des données détenues par le MAA et ses opérateurs peut être source d'innovations et d'amélioration de l'action publique. Ainsi, le traitement de données de masse par des algorithmes élaborés, voire par de l'intelligence artificielle, permet désormais des réalisations jusqu'alors impossibles. Quelques réalisations, parfois encore expérimentales, offrent des exemples des potentialités d'exploitation des données.

Tel est le cas du projet « iAlim », lauréat en 2018 d'un appel à projets de la DINSIC, devenue depuis DINUM, et de la direction interministérielle de la transformation publique (DITP) : abrité par la direction générale de l'alimentation du MAA et destiné à améliorer le ciblage des contrôles des restaurants dans le domaine de la sécurité sanitaire des aliments, iAlim vise à utiliser l'intelligence artificielle pour traiter les données textuelles accessibles librement sur internet (les commentaires des utilisateurs des établissements de restauration) afin de repérer des établissements sensibles.

Le suivi automatisé des cultures dans la PAC post-2020 : le « *monitoring* » grâce à l'intelligence artificielle

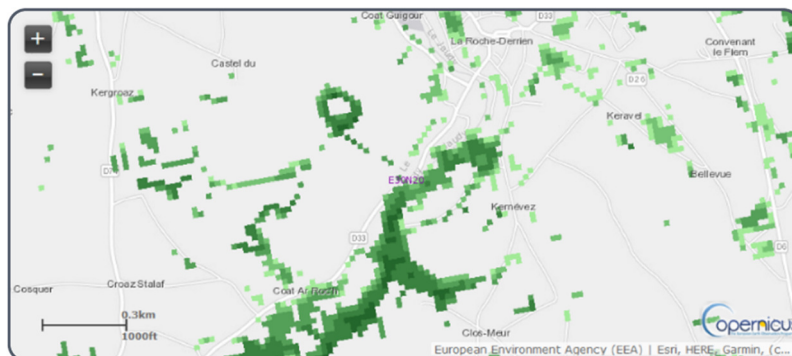
Dans le cadre de la programmation de la PAC post-2020, à la demande de la Commission européenne, l'Agence de services et de paiement devra mettre en œuvre un suivi automatisé du couvert des parcelles aidées au titre des mesures du FEAGA. Nommée « monitoring », cette opération sera réalisée par le MAA, l'IGN et l'ASP au moyen du traitement automatisé par l'intelligence artificielle d'images satellites issues du programme Copernicus. Le système d'information, encore au stade expérimental à l'heure actuelle, devra reconnaître automatiquement le type de couvert végétal des parcelles et déterminer si l'activité agricole ainsi menée est bien éligible aux dispositifs d'aides sollicités.

**Carte n° 1 : trois traitements de données satellites
(La Roche Derrien, Côtes d'Armor)**

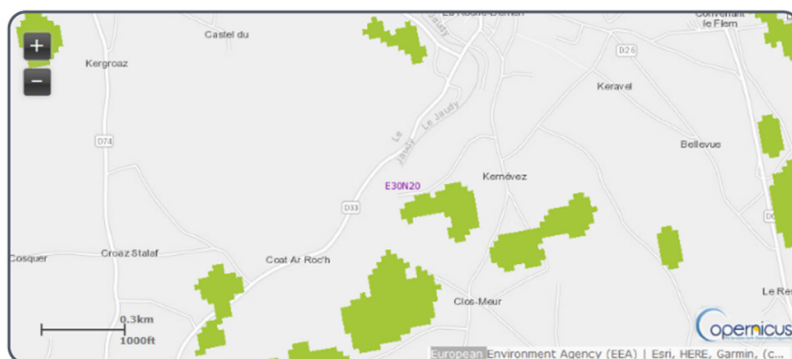
Mosaïque d'images haute résolution (couleurs naturelles) (2012)



Densité du couvert d'arbres (2015)



Prairies (2015)



Sources : Images Copernicus/ Service de surveillance terrestre

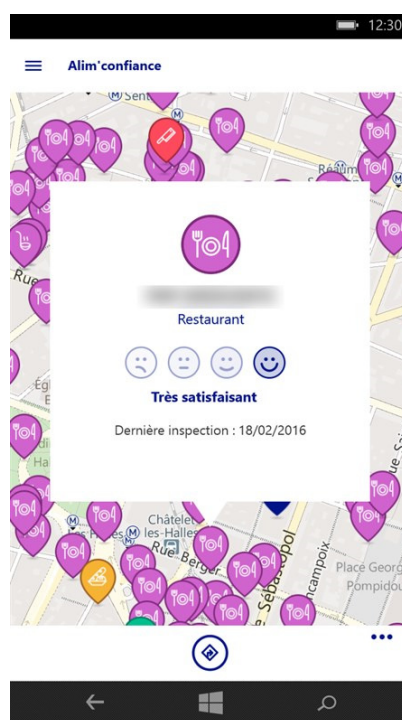
L'application Alim'confiance est un service à destination du public. Cette application lui permet de consulter le résultat des inspections sanitaires des établissements de restauration (y compris de restauration collective) et des commerces alimentaires à partir d'une simple géolocalisation¹³⁹. Les résultats d'inspection, disponibles pendant un an, sont présentés sous forme de quatre pictogrammes allant de « très satisfaisant » à « à améliorer d'urgence ». Le ministère a fait le choix, pour cette application, d'une solution achetée « sur étagère » plutôt que de développer un produit spécifique¹⁴⁰.

Le MAA, appliquant le principe d'ouverture des données publiques, a aussi publié ces données sur le portail de l'État data.gouv.fr. Des utilisateurs les ont alors reportées sur des cartes en format ouvert (« *open street map* ») et ont compilé les résultats d'inspection au-delà de la dernière année, ce qui permet de suivre l'évolution dans le temps des établissements. Il existe donc désormais, à disposition des consommateurs, deux offres de visualisation de ces données, l'une ministérielle, l'autre issue de la société civile.

¹³⁹ La Cour des comptes a récemment examiné, dans son rapport annuel, les conditions de contrôle de la sécurité sanitaire de l'alimentation (Cour des comptes, « Le contrôle de la sécurité sanitaire de l'alimentation : des progrès à consolider », in *Rapport public annuel 2019*, La Documentation française, février 2019, disponible sur www.ccomptes.fr).

¹⁴⁰ Le coût de ce service varie entre 30 000 et 50 000 € par an selon le volume des données.

Capture d'écran Alim'Confiance



C - Favoriser une réutilisation, en toute sûreté, des données privées

Dans le domaine agricole, « l'internet des objets »¹⁴¹ produit des données en masse pour les cultures et l'élevage grâce aux capteurs embarqués sur le matériel agricole, les drones, les équipements d'élevage et les bovins. Ces données privées (par exemple celles relatives au suivi des rendements, aux intrants, à l'état sanitaire du bétail, etc.) pourraient contribuer à améliorer la conduite des politiques publiques, notamment en les croisant avec des données publiques. Les grandes entreprises industrielles et de services du secteur agricole ont quant à elles déjà saisi l'enjeu de l'exploitation des données, parfois même au détriment des intérêts des producteurs auprès desquels elles les collectent.

¹⁴¹ « L'internet des objets » désigne le réseau de dispositifs physiques et d'éléments intégrés aux appareils, logiciels ou capteurs, permettant à ces objets d'être connectés et d'échanger des données *via* internet.

À l'étranger, de grandes entreprises commercialisent des engins dotés de capteurs : les données ainsi collectées appartiennent à l'entreprise, qui en autorise l'usage aux agriculteurs mais qui les valorise surtout pour son propre bénéfice, notamment auprès de compagnies d'assurance ou d'intervenants sur les marchés financiers. Cet usage des données peut ne pas être dans l'intérêt des agriculteurs, qui par exemple se retrouvent en position de faiblesse lors de la négociation de leurs prix par rapport aux acheteurs, du fait de l'asymétrie d'information entre eux. En France, de grandes compagnies internationales offrent déjà des outils intégrés collectant et analysant l'ensemble des données disponibles sur une exploitation, issues de matériels ou de bases de données extérieures.

Il y a donc un intérêt public à favoriser un cadre assurant la sûreté et la souveraineté¹⁴² de la réutilisation des données agricoles publiques et privées, garantissant la protection des intérêts de l'État et des agriculteurs.

C'est pourquoi le MAA a lancé une initiative destinée à faire connaître, dans le monde agricole, les enjeux de partage et réutilisation de leurs données.

**Une initiative de valorisation des données de la sphère agricole :
la société API-AGRO**

API-AGRO est lauréat de l'appel à projets de 2013 « Recherche finalisée et innovation », financé par le compte d'affectation spéciale « Développement agricole et rural » (CASDAR), réunissant la majorité des instituts techniques agricoles et des organismes de recherche-développement (Inra, Agro-Transfert Ressources et Territoires). Cette SAS, créée le 8 novembre 2016, regroupe aujourd'hui une trentaine d'acteurs : organismes de recherche appliquée (instituts techniques agricoles notamment), organismes de développement (avec le réseau des chambres d'agriculture) et une quinzaine d'acteurs privés.

À destination de toutes les filières agricoles (grandes cultures, élevage, viticulture, arboriculture, maraîchage, horticulture), API-AGRO agit comme courtier ou intermédiaire entre producteurs de données et utilisateurs. L'entreprise offre une plateforme permettant de diffuser et d'exploiter facilement des données et des services. Les fournisseurs de données restent maîtres de la visibilité et de l'accessibilité qu'ils souhaitent

¹⁴² La souveraineté numérique repose sur l'implantation sur le territoire national des serveurs hébergeant les bases de données et sur l'utilisation de solutions logicielles françaises (notamment les algorithmes).

y donner, répondant à des cas d'usage de diffusion large (telles que les données publiques ouvertes), à des ventes d'accès à de la donnée ou à des abonnements, mais également aux projets de recherche et/ou commerciaux nécessitant un espace sécurisé et privatif d'échanges de données.

API-AGRO offre également des services pour faciliter la réutilisation des données mises à disposition à travers la plateforme. API-AGRO a fait le choix de solutions informatiques françaises (logiciels et infrastructures d'hébergement) afin de garantir leur « souveraineté ».

Le modèle économique ne repose pas sur la vente des données, puisque certains jeux sont publics et donc gratuits, mais sur la rémunération des services *via* des abonnements souscrits par les entreprises.

III - Des données encore trop peu valorisées

A - Parvenir à une bonne gestion des données

La stratégie du MAA pour les données est déclinée dans son schéma directeur des systèmes d'information (SDSI) pour la période 2018-2022 et développée dans son plan de transformation numérique (2018-2022) qui vise à « développer une administration des données cohérente entre les membres de la sphère agricole et la mettre au service des politiques publiques ministérielles ». Le ministère entend faire de la gestion transversale des données une source de progrès : en améliorant les politiques publiques, en fiabilisant les procédures, en augmentant la performance et l'efficacité des SI et en simplifiant les politiques publiques grâce au principe « Dites-le nous une fois » déjà évoqué.

Le MAA avait identifié les points à améliorer pour atteindre ses objectifs : son manque de connaissance des données déjà exploitées par les opérateurs agricoles, qui pourraient faire l'objet d'échanges entre SI, et son absence de vision globale de celles déjà publiées et des modalités de leur publication. Aussi ne peut-il pas appuyer l'élaboration et l'évaluation des politiques publiques sur des analyses fondées sur le croisement de données, du fait de leur disponibilité insuffisante et d'un manque de compétences internes pour leur analyse.

La Cour a d'ailleurs constaté, lors d'un précédent contrôle portant sur le FEAGA, que les services chargés de la PAC au sein de la direction générale de la performance environnementale et économique des entreprises (DGPE) n'avaient pas un accès aisé et permanent aux données issues des SI de gestion des aides de l'Agence de services et de paiement,

mais à de simples extractions périodiques. La conception initiale de l'outil ne prévoyait pas un tel accès, et il n'a pas été jugé prioritaire de développer cette fonctionnalité, la priorité ayant été donnée, ces dernières années, aux adaptations nécessaires pour le paiement des aides.

L'objectif d'achèvement de ces chantiers est fixé à 2022. Toutefois, à mi-2019, ces divers projets n'en étaient qu'à leur commencement, les calendriers détaillés n'étaient pas arrêtés et les moyens spécifiques ne leur étaient pas affectés. Il est vrai que la délégation au numérique du ministère ne compte que quatre agents.

Les réutilisations de données sont encore rares. Sur les 290 jeux de données agricoles rendus publics sur le site Etalab, seuls 15 ont fait l'objet d'une réutilisation extérieure documentée¹⁴³. Il n'existe pas de mesure ou d'évaluation de la valeur créée par les réutilisations de données publiques, ni même d'inventaire exhaustif de ces réutilisations.

L'exploitation des potentialités des données au sein du ministère reste modeste, sinon inexistante, puisque ce dernier ne dispose pas de compétences expertes en traitement des données de masse ou en intelligence artificielle. De telles compétences seraient utiles à la DGPE, notamment pour évaluer les effets des aides publiques et le suivi de l'économie des filières et des revenus agricoles.

B - Réunir les conditions préalables à la valorisation des données

La gouvernance des données de la sphère agricole reste embryonnaire : le catalogue de l'ensemble des données disponibles n'est pas achevé, leur interopérabilité et l'harmonisation des formats des bases de données ne sont pas garanties et l'outil public d'agrégation des bases de données internes au MAA n'est pas encore livré.

De plus, à mesure que les applications sont interconnectées et échangent des informations, les interdépendances accroissent la fragilité du système. Il suffit qu'une API, parmi d'autres, soit indisponible pour que la téléprocédure qui l'utilise soit bloquée.

L'« État-plateforme », caractérisé par des flux de données échangées par les administrations, gouvernées par des API, nécessite donc un véritable encadrement interministériel : qualité des données, délais de réponse des API, garantie de disponibilité, procédure en cas de défaut de

¹⁴³ Selon les informations du site d'Etalab www.data.gouv.fr de septembre 2019.

service, etc. Des conventionnements et un suivi des engagements sont nécessaires pour réguler les échanges, *via* des API, entre opérateurs qui relèvent parfois de ministères différents. Or, la gouvernance interministérielle des API fait aujourd'hui défaut : la Cour estime qu'il faudrait remédier à cette lacune.

Enfin, les moyens et compétences consacrés au pilotage et à la gouvernance des données au sein du MAA semblent peu proportionnés aux défis à relever. Les crédits affectés à la délégation au numérique et à la donnée s'élevaient à 66 900 € en 2018 quand l'ensemble des dépenses informatiques était de 35 M€. Le budget informatique du ministère comme les ressources humaines des services informatiques sont mobilisés aux deux tiers par les dépenses contraintes : fonctionnement, dépenses incompressibles et maintien en condition opérationnelle des systèmes. Le MAA ne dispose d'expertise ni en intelligence artificielle ni en traitement des données en masse, expertise qui lui serait absolument nécessaire.

C - Faciliter la réutilisation des données publiques

La valorisation des données de la sphère agricole se traduit par un foisonnement d'initiatives d'ampleur inégale, souvent modestes, menées rapidement et parfois selon la méthode « agile »¹⁴⁴, qui contrastent avec les grands projets informatiques structurants du ministère, caractérisés par leur longue durée, leurs retards fréquents et leur coût élevé. En comparaison, les projets relatifs à la donnée permettent des progrès ponctuels dans une certaine dispersion.

Si les données constituent un bien public et un gisement de valeur, la réalisation de ces potentialités repose sur leur exploitation active, que ce soit par l'administration ou par des acteurs privés.

La seule disponibilité sur internet des données ne suffit pas à garantir leur réutilisation : une démarche active de facilitation des réemplois, voire de pédagogie, est nécessaire. Ainsi pour favoriser l'exploitation des données, il peut être nécessaire de les qualifier, de créer les conditions économiques et juridiques de leur réutilisation, ou encore d'illustrer leurs potentialités, par exemple grâce à des visualisations.

¹⁴⁴ La méthode agile est un mode de développement de projets (fréquemment informatiques) reposant sur des itérations fréquentes avec les commanditaires. En se dispensant de cahiers des charges détaillés, la méthode agile recherche la satisfaction des besoins du commanditaire avec une forte mobilisation de ce dernier dans les tests des versions successives du produit final.

CONCLUSION ET RECOMMANDATIONS

Si les données de la sphère agricole constituent un potentiel de valeur pour l'administration, les usagers et les acteurs privés, leur exploitation nécessite au préalable des investissements pour les qualifier et les mettre à disposition des futurs utilisateurs. Le MAA est engagé dans une stratégie destinée à lui permettre de tirer le plus grand parti de ses données de masse, particulièrement diversifiées, mais la démarche n'en est encore qu'à ses débuts et dotée de faibles moyens. Pour l'heure, les profits de cette ouverture ne sont pas encore pleinement mesurés.

Le MAA et ses opérateurs valorisent leurs données à travers diverses initiatives ponctuelles, une approche systémique et articulée restant à mettre en œuvre.

À défaut de pouvoir stimuler la valorisation de ces données par les acteurs du monde agricole, le MAA aurait tout intérêt à renforcer ses propres capacités d'exploitation de ses données, en recourant à des ressources expertes en intelligence artificielle et en traitement de données de masse.

La Cour formule donc les recommandations suivantes à l'État :

- 1. accroître le recours à des ressources expertes en matière d'intelligence artificielle et d'exploitation des données de masse (MAA) ;*
 - 2. réguler et sécuriser les échanges via des interfaces de programmation applicative (API) entre les opérateurs publics (DINUM).*
-

Réponses

Réponse du Premier ministre	166
Réponse du président de la commission d'accès aux documents administratifs (CADA)	169

Destinataires n'ayant pas d'observation

Directrice générale de l'établissement national des produits de l'agriculture et de la pêche (FranceAgrimer)

RÉPONSE DU PREMIER MINISTRE

Vous m'avez communiqué le chapitre sur les données du ministère de l'agriculture et de l'alimentation (MAA) qui illustre l'importance des enjeux liés aux données, ainsi que la complexité de la politique de leur valorisation.

Les recommandations sont pleinement cohérentes avec les orientations du Gouvernement en la matière, dont la mise en œuvre opérationnelle est en cours au travers notamment des plans de transformation numérique ministériels, notamment celui du MAA, et de la stratégie TECH.GOUV de la direction interministérielle du numérique (DINUM), qui a succédé à la DINSIC en vertu du décret n° 2019-1088 du 25 octobre 2019 relatif au système d'information et de communication de l'État et à la direction interministérielle du numérique.

1- La circulation des données à l'intérieur de l'administration permet à la fois de simplifier la vie des usagers en ne leur demandant pas des informations déjà connues et de diminuer la charge d'instruction par l'utilisation de données déjà validées par l'administration.

Des données présentes dans de très nombreuses procédures ont été identifiées par la DINUM qui a démontré la possibilité et l'intérêt de leur circulation à travers notamment la mise en place de la plateforme « API entreprises ».

Le MAA utilise l'API entreprises qui est l'une des sources alimentant le système SIRIUS cité par la Cour et met en œuvre d'ores et déjà le « Dites-le nous une fois » sur les données concernées. L'enjeu actuel pour le MAA est, en capitalisant sur ce succès, d'identifier les autres données, plus spécifiques à ses métiers, présentes dans plusieurs de ses procédures ou de celles de ses opérateurs et auxquelles il sera pertinent d'étendre ce mode de fonctionnement. Ce travail d'identification est en cours.

Comme le souligne la Cour, ce développement des interactions entre systèmes d'information, dont le fonctionnement est de plus en plus dépendant de la circulation effective des données, impose aussi une gouvernance interministérielle forte. La gouvernance interministérielle du numérique a été en conséquence renforcée dans le cadre de la stratégie TECH.GOUV. Elle se trouve désormais facilitée à deux niveaux : stratégique, au niveau des secrétaires généraux des ministères (Comité d'orientation stratégique interministériel du numérique) et opérationnelle, au niveau des directions du numérique (DNUM) et DSI (Comité interministériel du numérique).

De manière opérationnelle, dans le cadre de la stratégie TECH.GOUV, la DINUM a identifié le sujet de la circulation des données comme une priorité. C'est la raison pour laquelle une équipe dédiée à la mise en œuvre du principe du « Dites-le nous une fois » accompagnera techniquement et juridiquement l'ensemble des administrations dans la simplification des démarches en ligne et la suppression des pièces justificatives. L'offre de service portée par la DINUM sera publiée dans le courant du mois de décembre 2019.

Par ailleurs, la fluidification de la circulation des données n'est possible qu'à travers la définition d'un cadre général partagé par l'ensemble des fournisseurs de données. Ainsi dans le cadre de la stratégie TECH.GOUV, la DINUM travaillera sur la mise en place d'une plateforme d'échange de données inter-administrations.

Ce projet permettra :

- *de définir une architecture technique cible partagée par les ministères ;*
- *d'identifier des briques techniques qui devront être standardisées et d'autres qui devront être portées par les ministères afin que le modèle perdure ;*
- *de développer des briques existantes qui permettront de faciliter l'échange de données inter-administration (comme l'outil SIGNUP pour les habilitations) ;*
- *de créer un cadre commun autour du développement d'API, de leurs délais de réponse, des garanties de disponibilité afin de définir les contours d'une gouvernance interministérielle des API orchestrée par la DINUM ;*
- *de refondre la plateforme d'échange et de confiance (PEC) qui véhicule aujourd'hui près de 4 millions de flux issus des démarches usagers provenant de Service-public.fr vers près de 9 000 partenaires (dont 7 800 collectivités).*

La DINUM porte également une attention toute particulière à la diffusion des données de référence, que le MAA utilise quotidiennement (ex. registre parcellaire graphique ou encore la base SIRENE). Ces bases de données sont essentielles pour le bon fonctionnement des politiques publiques et notamment celles portées par le MAA. C'est la raison pour laquelle l'État s'engage sur la diffusion de ces données mises à jour et sur des niveaux de performance et de disponibilité exigeants.

De même, l'identité numérique est aussi un moyen de faciliter la circulation de l'information. France Connect a été mis en place au service des particuliers. Pour les entreprises, la stratégie TECH.GOUV prévoit de développer et déployer Pro Connect et l'écosystème associé. Les prochaines étapes seront de développer l'outil Pro Connect et de faire émerger un réseau d'acteurs concernés (fournisseurs d'identité, de services et de données, éditeurs).

La bonne exploitation du patrimoine des données de l'administration par celle-ci pour évaluer les politiques publiques, contribuer à leur élaboration, et, dans certains cas les conduire (comme par exemple la réalisation des analyses de risques nécessaires à l'élaboration de plans de contrôle) est aussi un enjeu majeur, qui demande la construction d'outils de connaissance de ce patrimoine et d'accès aux données garantissant le respect des règles de confidentialité, mais aussi la présence des compétences pour réaliser les traitements pertinents.

La Cour souligne à juste titre que répondre à ces défis demande de dégager des moyens. Cette ambition de valoriser au mieux ces données est fortement inscrite dans le plan de transformation numérique du MAA, qui a déjà lancé des actions pilotes comme le projet iAlim cité par la Cour.

2- L'encouragement à la réutilisation de la partie communicable de ce patrimoine par des acteurs privés est aussi un enjeu.

Celui-ci est plus difficile à appréhender car le principe même de l'open data est de faire confiance aux acteurs extérieurs pour inventer des utilisations auxquelles le détenteur initial de la donnée n'avait pas pensé. À ce titre, le MAA essaie, comme dans le cas d'Alim'confiance, de faciliter les réutilisations qu'il a identifiées comme pouvant appuyer les politiques dont il a la charge.

La publication de données sur le portail interministériel de mise à disposition de données publiques data.gouv.fr géré par Etalab permet leur réutilisation dans le cadre établi par le code des relations entre le public et l'administration (CRPA), notamment dans le cadre d'une licence ouverte, ce qui impose le respect de certaines obligations (mentionner la paternité de l'information et la date de dernière mise à jour).

Il n'y a pas d'obligation d'informer en retour le producteur de donnée sur les réutilisations réalisées, obligation qui pourrait être vécue par certains acteurs comme une surveillance de leur action et un frein à celle-ci. C'est pourquoi l'inventaire exhaustif des réutilisations des données publiques, que la Cour appelle de ses vœux, ne me semble pas opportune.

Au-delà du patrimoine de données détenu par l'administration, il apparaît que les données, y compris celles détenues par des acteurs privés, sont devenues un enjeu de politique publique. Comme l'indique la Cour à propos des données agricoles et de la société API Agro, le rôle de l'État est de donner une impulsion et de faciliter l'organisation par les acteurs concernés de l'utilisation de ces données dans un sens conforme à l'intérêt général, sans pour autant être l'opérateur de cette action.

RÉPONSE DU PRÉSIDENT DE LA COMMISSION D'ACCÈS AUX DOCUMENTS ADMINISTRATIFS (CADA)

Ce document, qui comprend un encadré sur l'application des règles de protection du secret par la Commission d'accès aux documents administratifs (CADA) n'appelle aucune réponse de la Commission.

Je me permets néanmoins de souligner deux points qui ne ressortent pas clairement de l'extrait que vous m'avez communiqué et qui pourraient y être précisés.

En premier lieu, en ce qui concerne l'information du public en matière environnementale, d'essence communautaire, il y a lieu de distinguer selon que l'information environnementale est, d'une part, liée ou pas à l'activité d'une personne morale, cette dernière ne pouvant se prévaloir de la protection de la divulgation d'un comportement susceptible de lui porter préjudice et, d'autre part, si cette information porte précisément sur des émissions de substance dans l'environnement, seules ces dernières bénéficiant d'un régime très libéral de communication partiellement dérogatoire aux secrets protégés.

En second lieu, il y a lieu de distinguer, en matière d'information du public, le mode de communication. En effet, si la protection des données à caractère personnel n'intervient qu'au titre de la protection de la vie privée pour la communication d'un document ou d'une information à une personne, les dispositions de l'article L. 312-1-2 du code des relations entre le public et l'administration prévoient que, sauf disposition contraire, ne peuvent faire l'objet d'une diffusion publique, généralement entendue comme une mise en ligne, que les documents qui ne comportent aucune mention relevant d'un secret protégé et qui ne comportent aucune donnée à caractère personnel, à l'exception de ceux qui figureraient dans la liste des documents définie par l'article D. 312-1-3 de ce code.
