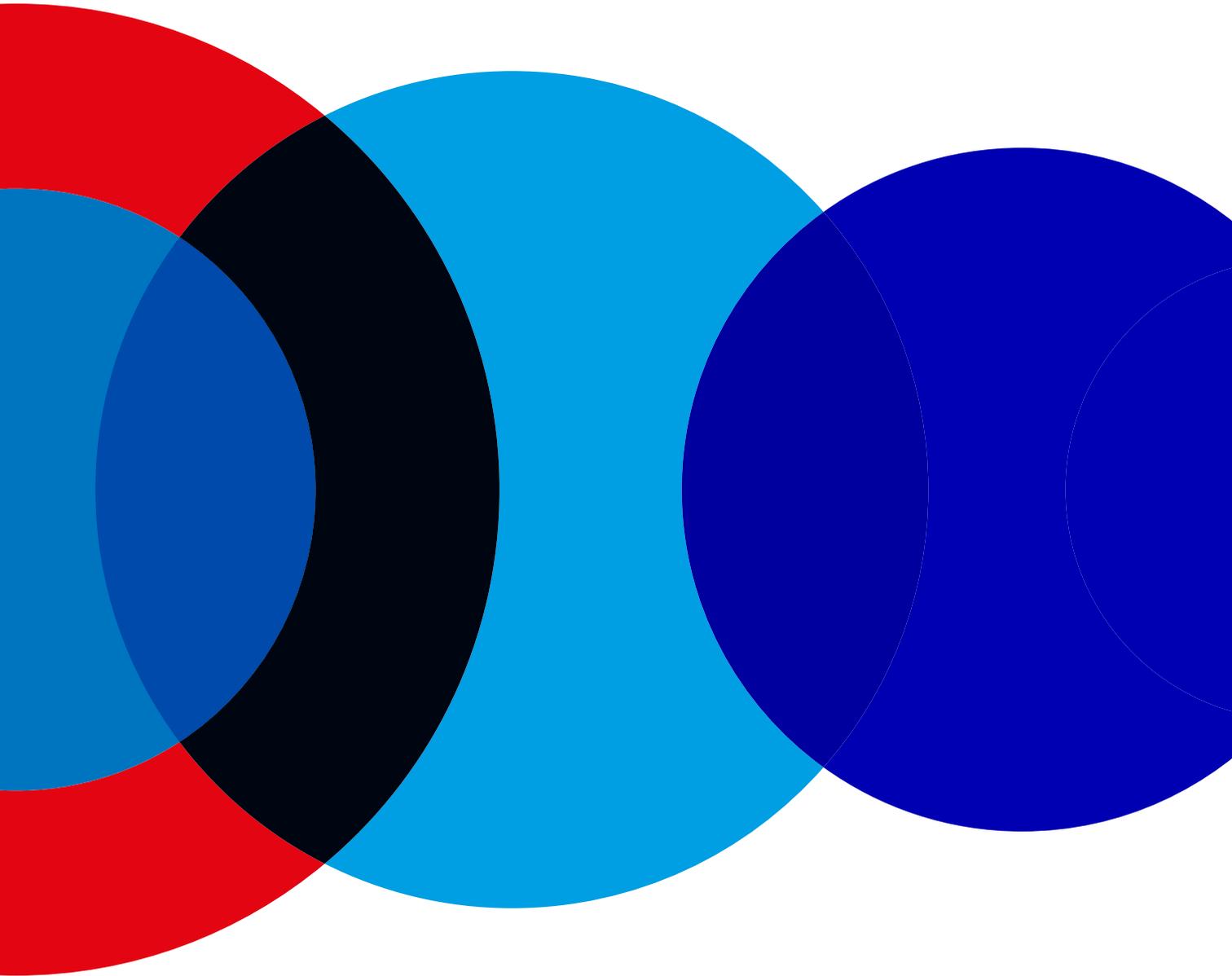


LIVRE BLANC

**COMMENT LES ADMINISTRATIONS FISCALES
ET DOUANIÈRES PEUVENT AMÉLIORER LEURS
PERFORMANCES GRÂCE À LA DONNÉE ?**

TABLE DES MATIÈRES

I. INTRODUCTION	5
II. EN QUOI LES DONNÉES SONT UN ATOUT ET QUEL EST LEUR POTENTIEL ?	6
III. QUELS SONT LES PROBLÈMES RENCONTRÉS PAR LES ADMINISTRATIONS ET COMMENT LES SURMONTER ?	11
IV. COMMENT DATAFID PEUT AIDER LES ADMINISTRATIONS FISCALES À AMÉLIORER LA GESTION DES DONNÉES ?	17



I. INTRODUCTION

Au cours des dernières années, les administrations fiscales et douanières ont digitalisé les procédures de collecte et de distribution des recettes fiscales : enregistrement des contribuables, déclarations de ces derniers, contrôle et recouvrement.

L'une des conséquences de cette digitalisation est la génération de grandes masses de données : les administrations fiscales et douanières disposent aujourd'hui d'un volume conséquent d'informations sur leurs contribuables. Il s'agit là d'une ressource considérable en vue d'améliorer leurs opérations. Les bénéfices retirés d'une meilleure utilisation des données fiscales et douanières sont légion : augmentation des recettes, lutte contre la fraude et réduction des risques fiscaux, gestion proactive des activités fiscales, plus grande efficacité, possibilité d'innover et de proposer de nouveaux services, etc.

Toutefois, la faculté à exploiter ces avantages est déterminée par la manière dont les données fiscales sont stockées, utilisées et partagées, et les prérequis dans ces domaines ne sont souvent pas atteints. Par ailleurs ces données génèrent aussi des risques : en raison de leur confidentialité et sensibilité, elles sont vulnérables aux attaques ou aux mauvaises manipulations.

Afin de créer des conditions propices à l'exploitation du potentiel des données tout en préservant leur sécurité, il est nécessaire que les administrations fiscales et douanières abordent une mutation pour devenir des organisations « orientées vers les données ». Il ne s'agit pas simplement d'améliorer ses systèmes d'informations : les efforts doivent porter sur l'adoption de cadres stratégique, juridique, administratif et organisationnel propices à une meilleure utilisation des données. C'est un chantier complexe et semé d'embûches qui nécessite de réfléchir collectivement aux options qui se présentent, de sensibiliser les équipes au potentiel des données, de développer des capacités d'analyse, de se doter d'équipements pertinents et d'adopter une gouvernance à la hauteur.

Le présent livre blanc, rédigé par le projet DATAFID, vise à accompagner les administrations fiscales et douanières dans cet exercice. En quoi les données sont-elles un atout ? Quel est leur potentiel ? Que faire pour maximiser ce dernier ? Ce document tente d'apporter des réponses à ces questions. Il identifie les bénéfices tirés d'une bonne utilisation des données, propose des recommandations sur les étapes à suivre pour y parvenir, et souligne les écueils à éviter.

II. EN QUOI LES DONNÉES SONT UN ATOUT ET QUEL EST LEUR POTENTIEL ?

Toutes les tendances et projections témoignent d'une hausse exponentielle des volumes de données produites. Les nouvelles technologies, l'intelligence artificielle, la science des données sont amenées à prendre une place croissante dans les organisations à mesure que la quatrième révolution industrielle avance, y compris dans les administrations fiscales et douanières. Elles pourront tirer profit de ces avancées, à condition d'opérer une révolution organisationnelle et d'adopter une culture axée sur les données. Ces dernières existent et sont déjà collectées, mais elles ne révèlent leur potentiel que lorsqu'elles sont croisées, échangées et analysées dans leur complexité et leurs multiples dimensions.

Toutes ces tâches doivent se faire dans un cadre structuré, issu d'une stratégie de valorisation des données inclusive et transversale. Plus tôt une administration se dotera d'une telle stratégie, mieux elle saura tirer profit du potentiel de ses données. A l'inverse, une administration qui ne se saisirait pas du sujet prendrait le risque de se priver de nouvelles ressources financières, mais aussi et surtout d'accumuler un retard difficilement rattrapable dans un domaine où les technologies évoluent rapidement.

Si les usages dérivés d'une bonne exploitation des données sont potentiellement infinis, tous convergent vers un objectif unique : permettre à l'administration de prendre des décisions avisées, justes et rapides, pour au final contribuer à mobiliser les ressources nécessaires au financement des besoins du pays.

Afin de rendre ces bénéfices plus concrets et intelligibles, nous les présentons ci-dessous regroupés en quatre catégories :



01

L'usage des données permet d'améliorer les performances fiscales

02

L'usage des données permet d'améliorer les performances administratives et financières

03

L'ouverture, le partage et l'accès aux données amplifient leur potentiel

04

L'analyse des données est à l'origine du développement de services nouveaux, jusqu'à présent non imaginés

1. L'usage des données améliore les performances fiscales de l'administration

● Les données peuvent être utilisées pour prévoir des tendances économiques et anticiper des manques à gagner fiscaux

Grâce aux données, il est possible de procéder à des modélisations pour tester des scénarios et effectuer des prévisions. Plus les données sont nombreuses et de qualité, plus les résultats des modélisations sont précis. Il devient plus aisé d'anticiper d'éventuels changements de comportements en réaction à une nouvelle réglementation fiscale, de quantifier les gains ou les pertes qu'elle est susceptible de générer, ou encore d'identifier de façon proactive d'éventuels problèmes socio-économiques impactant la mobilisation des ressources.

○ **L'exploitation des données permet une bonne gestion des risques**

L'utilisation des données historiques de fraudes permet d'identifier des tendances et schémas de comportements frauduleux : marchandises régulièrement sujettes à des déclarations erronées, types d'impôt généralement mal renseignés, etc. En explorant ces informations, les administrations peuvent identifier les risques et prendre les mesures adéquates pour éviter qu'ils ne surviennent : programmation des contrôles, communication particulière auprès des contribuables, etc.

○ **Le recoupement de données avec des sources externes permet de mieux connaître vos contribuables ou votre public**

De nombreuses administrations fiscales procèdent au recoupement de données externes dans le monde, pour des usages divers.

En termes de lutte contre la fraude fiscale, aux États-Unis, le Internal Revenue Services (IRS) utilise le big data : les données reçues des banques, des registres fonciers et des cartes de crédit sont utilisées pour détecter des trains de vie en inadéquation avec les dossiers fiscaux des contribuables, leur permettant de récupérer plus de 300 milliards de dollars d'impôts perdus chaque année.

A Singapour, l'administration a commencé à utiliser l'analyse prédictive pour pré-remplir les déclarations ou les évaluations individuelles avec les données personnelles des contribuables ou les données bancaires. Il en résulte une meilleure anticipation des recettes à collecter, et un suivi plus fin de leur recouvrement.

2. L'usage des données améliore les performances administratives et financières de l'administration

○ **Les données peuvent être utilisées pour mieux évaluer son action**

Les performances fiscales et douanières sont jugées à l'aune d'indicateurs spécifiques : taux de succès des contrôles, degré d'efficacité d'une politique économique, délai de traitement des contentieux, etc. Ces informations sont mesurées à échéances régulières et permettent de constater une évolution dans le temps. Le calcul de ces indicateurs se fait sur la base d'informations disponibles dans les services concernés. L'afflux massif de données permet d'accélérer la compilation des indicateurs et d'en améliorer la qualité. Les rapports de performance sont plus précis et reflètent mieux la vérité : les fonctionnaires appréhendent mieux les dynamiques constitutives de certains problèmes et l'administration peut rapidement opérer des changements si nécessaire.

○ **La science des données peut aider à automatiser les tâches et à traiter de grandes quantités de données**

La science des données peut aider les administrations fiscales à examiner des grands volumes de données, et à automatiser des tâches réalisées auparavant manuellement. Grâce à l'utilisation de langages de programmation, il ne faut que 2 minutes pour analyser 7 ans de données douanières en Côte d'Ivoire, contre plus de 2 heures auparavant pour analyser 1 an de données avec des outils plus traditionnels (tableur notamment). Ces gains de temps, considérables, permettent de redéployer les ressources humaines sur des tâches plus complexes, où leur action ne peut être remplacée.

3. L'ouverture, le partage et l'accès aux données amplifient leur potentiel

○ **L'échange de données avec d'autres pays permet d'avoir une meilleure vision des contribuables à l'étranger**

La mondialisation des affaires a engendré des schémas d'évasion fiscale internationaux difficile à appréhender. Ces schémas peuvent être tracés et démantelés grâce à un partage sûr et organisé des données. Des initiatives promues à l'échelle internationale, comme le Forum Mondial de l'OCDE sur la transparence et l'échange de renseignements, permettent aux pays d'adopter des standards reconnus pour faciliter l'échange de renseignements fiscaux sur demande, ou de façon automatique. Les informations partagées aident l'administration à mieux connaître les situations fiscales de leurs contribuables à l'étranger, et à prendre les mesures nécessaires pour les mettre en conformité.

○ **Des données en accès libre pour une transparence accrue et un meilleur civisme fiscal**

Le consentement à l'impôt est, entre autres, déterminé par le niveau de confiance des contribuables en leur administration fiscale. L'ouverture des données quantitatives et qualitatives, dans le respect des règles de confidentialité en vigueur, est un gage de transparence et de responsabilité de l'administration : elles doivent permettre aux organisations de la société civile, aux chercheurs et aux citoyens qui le souhaitent d'obtenir des informations sur l'efficacité de l'administration et sur le respect de son mandat. Une administration qui permettra l'accès à certaines de ses données inspirera la confiance de ses contribuables.

4. L'analyse des données est à l'origine du développement de services nouveaux, jusqu'à présent non imaginés

○ La modernisation des infrastructures est indispensable pour profiter de la révolution des données

Une grande partie du potentiel des données demeure inconnu. Les usages d'aujourd'hui étaient insoupçonnés, il y a quelques années encore. Les anciens systèmes peuvent sembler rentables pour gérer les exigences administratives actuelles, car ils sont déjà payés et ne nécessitent pas de dépenses supplémentaires au-delà des coûts d'assistance actuels. Cependant, pour bénéficier des perpétuelles avancées de la recherche, des investissements doivent être régulièrement effectués pour maintenir l'infrastructure à jour. Il est préférable de les réaliser le plus tôt, afin d'éviter de subir des coûts de rattrapage qui croissent à mesure que l'échéance est reculée.

○ Les données bien utilisées permettent de proposer des nouveaux services

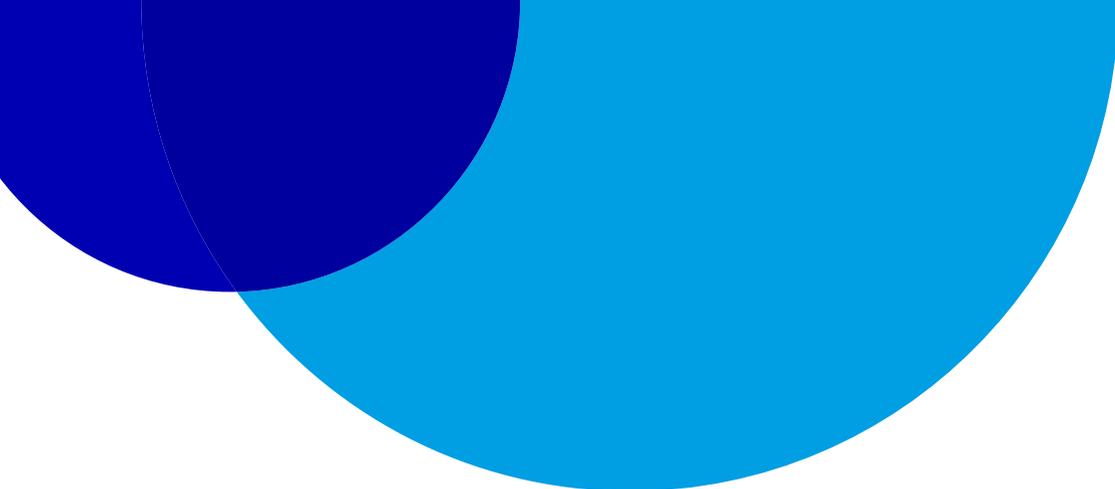
La disponibilité de grandes quantités de données permet à l'administration d'adopter une approche proactive, attitude relativement « nouvelle ». Le traitement de l'information disponible contribue à réaliser une segmentation améliorée des contribuables selon leurs comportements, leurs caractéristiques, leurs revenus, etc. Elle autorise des analyses fines, susceptibles d'aboutir à la compréhension de besoins singuliers, point de départ de la création de nouveaux services : identification de signaux faibles pour des entreprises en difficulté, ciblage de contribuables éligibles à des aides mais qui ne les sollicitent pas, anticipation de difficultés dans la mise en conformité d'une situation fiscale et proposition d'échéances de paiements, etc. Ces innovations dépendent de la faculté à exploiter les données disponibles et à en tirer des informations intéressantes.

III. QUELS SONT LES PROBLÈMES RENCONTRÉS PAR LES ADMINISTRATIONS ET COMMENT LES SURMONTER ?

« Data is power » : la donnée est un enjeu de pouvoir. La formulation de politiques publiques peut et doit bénéficier de l’afflux massif de données pour améliorer la conformité, lutter contre la fraude, augmenter le recouvrement des recettes et identifier de nouveaux gisements de ressources. En bref : les données sont là pour aider votre administration à mieux fonctionner. S’en priver, c’est renoncer à améliorer ses performances fiscales et maintenir de la distance avec vos contribuables.

La recherche a mis en évidence les principales difficultés rencontrées par les administrations fiscales et douanières dans leur relation à la donnée. Elles se situent à deux niveaux, stratégique et conceptuel d’un côté, et opérationnel et technique de l’autre :

Niveau stratégique et conceptuel	Une mauvaise vision stratégique en raison d’une faible acculturation à la donnée, une mauvaise compréhension de son potentiel
	Une gouvernance dysfonctionnelle
	Des difficultés à transformer les données disponibles en une information exploitable
	Un cadre juridique et réglementaire n’encourageant pas la collaboration entre services et administrations
Au niveau opérationnel et technique	Des systèmes informatiques faiblement équipés ou mal paramétrés, qui dépendent fortement de fournisseurs privés
	L’absence de plateforme intégrée unique pour l’analyse des données
	Une inadéquation entre les besoins en ressources humaines pour le traitement des données fiscales et les profils recrutés dans les administrations
	Une interconnexion déficiente entre administrations qui ne permet pas le recoupement de données
	Des données de mauvaise qualité
	L’inadéquation des processus de gestion du changement



La liste n'est pas exhaustive mais donne une idée de l'ampleur des défis à relever.

La réponse globale à ces difficultés se trouve dans **l'élaboration et la mise en œuvre d'une stratégie de données, par laquelle l'administration adopte une vision « orientée vers les données »**. Voici quelques conseils pour entamer cette démarche, dans ses dimensions stratégiques, organisationnelles, techniques et réglementaires. Il convient de préciser que plusieurs étapes peuvent être conduites en même temps, l'important étant que le calendrier d'élaboration puis de mise en œuvre de la stratégie soit intériorisé par les personnes concernées et qu'il fasse l'objet d'un suivi attentif.

Définition : Une stratégie de données est une planification à long terme qui définit la technologie, les processus, les personnes et les règles nécessaires à la gestion des actifs informationnels d'une organisation. Elle doit permettre de gérer les données de manière cohérente, afin d'exploiter leur potentiel au maximum, ainsi que présenté dans la section précédente.

La stratégie a pour objectif une transformation globale de l'organisation. Il convient dès lors d'obtenir d'un côté un **portage politique** (Direction Générale) donnant sa légitimité à l'exercice, mais aussi d'associer dès la genèse les Directions et services concernés par la production et l'utilisation de la donnée. La **constitution d'une équipe représentative** de ces Directions et/ou services est un préalable à la bonne compréhension et conditionnera le succès des opérations de conduite du changement.

Dans un second temps, il importe de **définir des objectifs en lien avec les priorités de l'organisation**, et donc en phase avec le contexte politico-économique dans lequel l'administration évolue. Ces priorités doivent naturellement aboutir à l'identification de besoins en matière de données : s'agit-il d'améliorer l'efficacité des contrôles pour diminuer la fraude ? D'améliorer le modèle de prévisibilité des recettes ? De réduire le temps alloué aux opérations de dédouanement ? De mieux gérer les contentieux ? Une priorisation des objectifs doit être opérée, qui pourra aider à déterminer comment procéder par la suite. De façon concrète, cette approche peut passer par l'organisation d'entretiens individuels ou collectifs pour comprendre les pratiques et attentes de chacun, des ateliers pour identifier un modèle de gouvernance pertinent, ou l'identification de cas d'usage pour analyser une problématique métier et trouver une solution y remédiant.

En parallèle de ces opérations, il convient de **procéder à un inventaire du patrimoine des données disponibles**. Il est possible pour ce faire d'employer la méthode des 5V :

LA MÉTHODE 5V

1

VOLUME

Quelles quantités sont stockées (périodes couvertes, champs couverts, etc.) ?

2

VARIÉTÉ

Données structurées, non structurées ? De quels systèmes d'information proviennent-elles ? Dans quel format sont-elles stockées ?

3

VITESSE

À quelle vitesse les données arrivent dans la base, et à quelle vitesse doivent-elles être traitées pour les besoins des opérations de l'administration ? La réponse à cette question peut aider à déterminer les besoins en infrastructure

4

VÉRACITÉ

Qu'il s'agisse d'analyse descriptive, prescriptive ou d'intelligence artificielle, la qualité des données détermine leur potentiel et donc la qualité des services que vous proposerez. Il importe d'identifier les sources d'erreurs potentielles (erreur de déclaration, malveillance et bug, mauvaise manipulation, etc.) pour proposer des mesures d'atténuation et fiabiliser vos données

5

VALEUR

De quelles données avez-vous réellement besoin pour vos opérations ?
Quelles sont celles qui apportent réellement de la valeur ?

Cet état des lieux de vos actifs en matière de données constitue un point de départ important de votre stratégie.

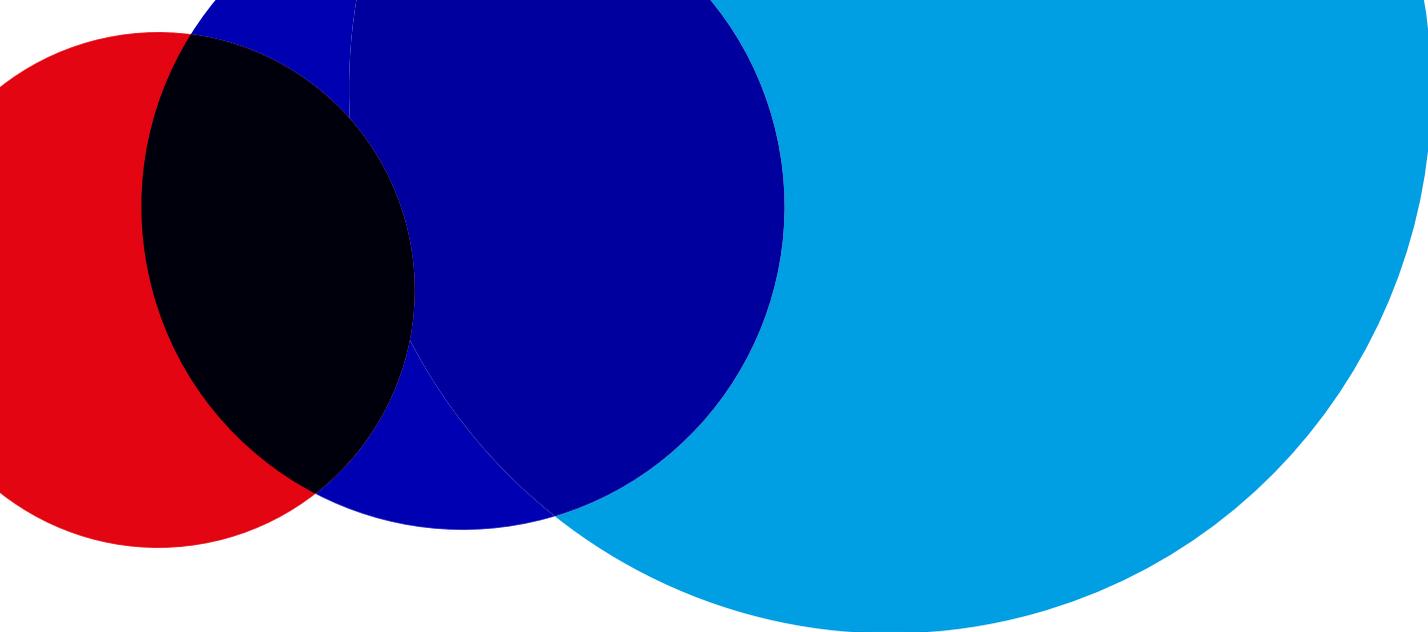
Par la suite, l'administration doit amorcer une **réflexion sur l'organisation des fonctions data**. Il n'y a pas d'organisation standard, mais plusieurs modèles existent :

- ◆ Concentration de l'ensemble des fonctions data dans une seule cellule vers laquelle se tournent les directions pour la formulation de cas d'usage
- ◆ Création d'une unité active sur un périmètre restreint d'activités nécessitant de la donnée (ex : ciblage des contrôles). Ce modèle permet une montée progressive en compétence et répond à des contraintes en termes de ressources, en limitant les champs d'application de la donnée sur un périmètre donné
- ◆ Formation d'agents à la science des données puis dissémination de ces derniers dans l'ensemble des directions concernées

La détermination des aspects organisationnels doit s'accompagner d'une **réflexion sur la gouvernance des données**, c'est-à-dire l'ensemble des processus, fonctions, règles, normes et paramètres qui garantissent que les informations sont utilisées de manière efficace et efficiente. Elle définit également qui peut effectuer quelle action sur quelle donnée et selon quelle méthode. Il n'existe pas de modèle standard de gouvernance, ce dernier étant défini par les spécificités de l'organisation concernée. Dans le cas des administrations fiscales et douanières, les données se caractérisent par leur confidentialité et leur implication en termes de conformité des contribuables à la loi. Dès lors, la gouvernance des données doit être orientée de façon à assurer la protection des données (anonymisation, sécurisation), ainsi que leur protection contre des altérations volontaires ou accidentelles (gestion des accès, traçabilité et horodatage des modifications, etc.). La gouvernance des données détermine par ailleurs comment les données doivent être stockées, comment elles peuvent être échangées, et par quels procédés elles sont transformées pour y ajouter de la valeur.

La gouvernance de la donnée peut influencer la **stratégie informatique** de l'administration : quels systèmes d'information doivent être mis en place pour collecter, échanger, stocker les données utiles à l'amélioration des performances de l'administration ? Là encore, il n'existe pas de modèle standard, mais certains points doivent être abordés pour mettre l'administration sur la bonne voie :

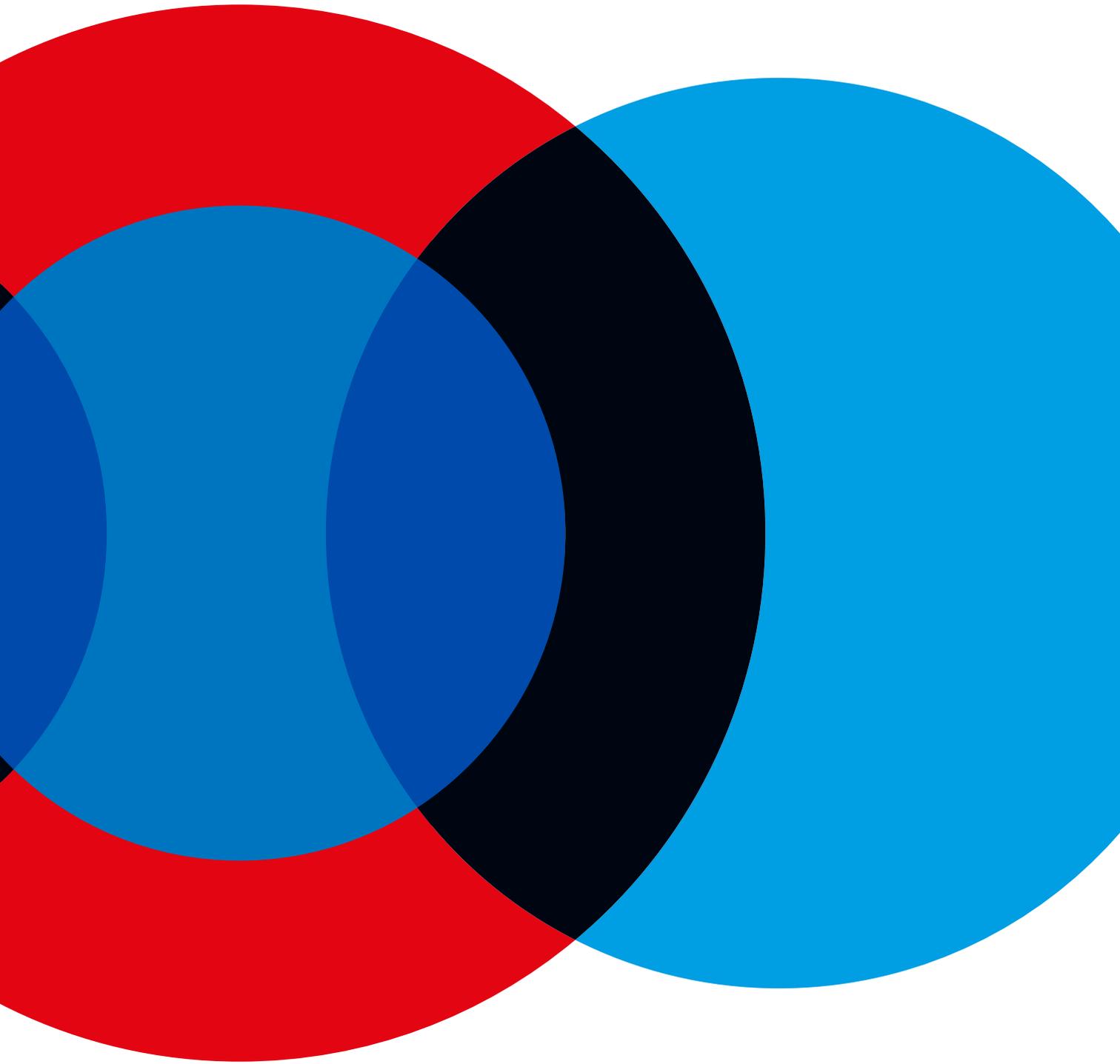
- ◆ Quelle infrastructure peut permettre d'atteindre la puissance de calcul requise pour la valorisation de la donnée ?
- ◆ L'infrastructure permet-elle aux données d'être protégées (existence de sites de redondances, maintien des éventuelles licences, etc.) ?
- ◆ Faut-il revoir le schéma directeur des systèmes d'information ?
- ◆ Quelle est la stratégie réseau, c'est-à-dire comment prévoit-on de connecter tous les centres et bureaux concernés sur l'ensemble du territoire ?



Une bonne stratégie data ne peut être mise en œuvre sans **ressources humaines**. Dès la genèse de l'exercice l'administration doit procéder à une cartographie des compétences requises pour opérationnaliser la stratégie data : analystes de données, data scientists, ingénieurs informaticiens, statisticiens, etc. Les compétences identifiées doivent être comparées aux effectifs en présence dans l'administration en vue d'identifier les besoins en formation et/ou en recrutement. Les métiers de la donnée sont relativement nouveaux, et dans de nombreux cas les postes proposés par l'administration ne prennent pas ces spécificités en considération ; une véritable réflexion doit être amorcée en étroite collaboration avec les directions en charge de la gestion des ressources humaines pour attirer ces talents dans l'administration, sous peine d'échouer à mettre en œuvre la stratégie. Par ailleurs, l'offre de formation continue sur les métiers de la donnée est abondante, et doit être intégrée dans la politique RH de formation des fonctionnaires déjà en place.

Enfin, la mise en œuvre d'une stratégie doit faire l'objet d'un **suivi fin, inclusif et régulier** pour remonter les difficultés et procéder à d'éventuels ajustements. Il importe par ailleurs de proposer des indicateurs de réussite consensuels auxquels toutes les personnes impliquées dans la mise en œuvre peuvent s'identifier. Des mécanismes de remontée d'information (réunions périodiques, questionnaires) doivent être proposés et considérés comme éléments à part entière de la conduite du changement.

Il existe un outil particulièrement intéressant pour accompagner une administration fiscale dans la construction de sa stratégie de données : il s'agit du **modèle de maturité analytique des administrations fiscales**, conçu par l'OCDE. Ce modèle est une sorte de grille d'auto-évaluation qui passe en revue plusieurs dimensions propres à l'administration fiscale aux niveaux stratégique (stratégie, gouvernance, culture) et opérationnel (systèmes d'information, gestion des données, compétences, RH, etc.). Pour chacune de ces dimensions, le modèle décrit comment fonctionne une administration selon son degré d'avancement (émergent, en progrès, établi, élevé ou idéal). L'administration qui s'autoévalue peut s'identifier au niveau qui lui correspond le plus et ainsi déterminer sa maturité analytique. L'exercice permet de mettre en avant les forces et faiblesses de la structure, et peut aider à établir une stratégie de données.



IV. COMMENT DATAFID ACCOMPAGNE LES ADMINISTRATIONS FISCALES ET DOUANIÈRES DANS LA VALORISATION DE LEURS DONNÉES ?

Le projet DATAFID, financé par la Direction Générale du Trésor français et mis en œuvre par Expertise France, a pour vocation d'accompagner les administrations fiscales et douanières dans la valorisation de leurs données. Le projet intervient aux deux niveaux clés identifiés ci-dessous : stratégique et opérationnel.

NIVEAU STRATÉGIQUE

- ◆ **Sensibilisation**

La stratégie passe par la **sensibilisation auprès des plus hautes instances**. Des rencontres régulières avec les Directions sont organisées pour rendre des comptes et promouvoir la valorisation des données.

- ◆ **Stratégies data**

Le projet va débiter des réflexions sur la constitution de **stratégies data** avec quelques administrations intéressées et « matures ». L'objectif affiché est de préparer des feuilles de route data qui permettront d'adopter une vision basée sur les données, pour une utilisation optimale du potentiel de ces dernières.

- ◆ **Modèles de maturité analytique**

DATAFID souhaite par ailleurs accompagner les administrations le souhaitant dans **l'utilisation des modèles de maturité analytique** pour les aider à se positionner dans le domaine, exercice préalable à la constitution d'une stratégie à part entière.

NIVEAU OPÉRATIONNEL

Trois grands axes de travail sont réalisés :

- ◆ **Construction de cas d'usage**

Le projet accompagne les administrations dans l'identification de cas d'usage, à savoir des problèmes rencontrés qui peuvent être résolus par une meilleure exploitation des données. Parmi les cas d'usage développés figurent un travail permettant la détection de déclarations suspectes grâce à des modèles statistiques (loi de Benford), un travail sur les modèles de sélectivité des contrôles, une réflexion sur l'automatisation de certaines tâches, et la création d'une base de données de valeur pour faciliter l'évaluation des marchandises en douanes.

- ◆ **Coaching et formation**

Les équipes ont rédigé un certain nombre d'exercices (travaux pratiques) et accompagnent les administrations le souhaitant dans leur mise en œuvre, à un rythme régulier, en présentiel ou à distance. Par ailleurs le projet travaille de concert avec ces mêmes administrations pour co-construire un programme de formation en ligne sur l'usage de la science des données dans les administrations fiscales et douanières : prise en main des langages de programmation, visualisation des données, construction d'analyses miroir, utilisation des données satellites, etc.

- ◆ **Déploiement d'une solution technique**

La valorisation des données passe par la bonne maîtrise des outils de la science des données. Le projet met à disposition un « datalab », une plateforme permettant aux partenaires de s'exercer et de manipuler leurs données pour en explorer la valeur. Certaines administrations ont demandé à ce que ce datalab soit pleinement intégré à leurs systèmes d'information pour une utilisation encore plus aboutie.

Le projet DATAFID est financé
par la Direction Générale du Trésor français



Clément Dupont
Chef de projet DATAFID
clement.dupont@expertisefrance.fr

