

**Rapport du groupe de travail  
franco-britannique sur  
l'économie de la donnée**

**La révolution de la donnée  
au service de la croissance**

**Innovation, Infrastructure,  
Compétences et "Pouvoir  
d'agir" à l'ère numérique**

Le 27 novembre 2015, Emmanuel Macron, ministre de l'économie, de l'industrie et du numérique français, George Osborne, chancelier de l'Echiquier et Axelle Lemaire, Secrétaire d'Etat au numérique, ont lancé la Taskforce franco-britannique sur la Donnée au service de la croissance chargée d'établir un rapport sur l'économie de la donnée et sur les axes de collaboration possibles entre nos deux pays.

Henri Verdier, Administrateur général des données, et Sir Nigel Shadbolt, directeur du Jesus College et co-fondateur de l'Open Data Institute, ont été nommés coprésidents de ce groupe.

Pour les assister, les coprésidents ont choisi ensemble un petit groupe d'experts français et britanniques, issus du secteur privé et du milieu universitaire :

- Patrick Cocquet (Cap Digital) ;
- Alban Schmutz (OVH) ;
- Mokrane Bouzeghoub (Allistene) ;
- Rand Hindi (Snips) ;
- Yann Lechelle (Snips) ;
- Clive Humby (DunnHumby) ;
- Doug Monro (Adzuna) ;
- Andrew Lawson (Salesforce) ;
- Stéphanie Finck (Salesforce).

Dans le cadre de ses travaux, la Taskforce a consulté de nombreuses parties prenantes, notamment des administrations publiques, des organisations internationales, des PME, de grandes sociétés et des startups. Ses recommandations tiennent compte de la grande variété des attitudes existantes - et qui continuent à évoluer - dans la manière dont nous utilisons et conservons les données, allant de ceux qui recommandent la plus grande prudence aux plus convaincus, en passant par ceux qui ne sont pas encore convaincus des bénéfices qu'elles pourraient apporter.

Les principaux objectifs des travaux de la Taskforce étaient les suivants :

- instaurer les conditions nécessaires à l'émergence d'un écosystème de l'économie de la donnée florissant, notamment en renforçant la confiance des citoyens et des entreprises ;
- approfondir la collaboration entre les entités publiques françaises et britanniques en matière de données ouvertes, et améliorer les services publics grâce aux données ;
- développer les compétences autour des données et de la science des données dans les secteurs publics et privés ;
- encourager les entreprises (en particulier les PME) à exploiter les données publiques et privées dans l'objectif de renforcer leur compétitivité ;
- acquérir une meilleure connaissance des modèles d'innovation par les données et des bénéfices économiques potentiels des données ;
- encourager le développement des prestataires de services liés aux données en France et au Royaume-Uni.

En particulier, il était demandé à la Taskforce d'examiner plus particulièrement les points suivants :

- les bénéfices économiques, sociaux et environnementaux pouvant être apportés par différents types de données ;
- le rôle de la technologie et la manière de créer un environnement favorisant la pleine exploitation du potentiel des données ;
- les comportements face aux données et les courants contraires de l'exploitation des données et de la crainte montante d'une utilisation abusive.

Le présent rapport est une synthèse des conclusions qui ont émergées de la consultation de nombreux experts, mais aussi de la grande quantité d'informations et d'exemples de réussites et d'échecs qui ont pu être recueillis dans les deux pays, dans d'autres pays d'Europe et dans le reste du monde.

La discussion sur tous ces sujets reste ouverte. Elle n'en est même qu'à ses débuts. C'est maintenant que le véritable travail commence : il faut établir une feuille de route pour soumettre ces recommandations à l'épreuve de la réalité, les appliquer lorsque cela se justifie et approfondir la collaboration que cette Taskforce a créée entre les administrations publiques, les communautés du secteur des technologies numériques, les entreprises et les représentants des consommateurs des deux côtés de la Manche.

*Nigel Shadbolt & Henri Verdier*

# TASKFORCE FRANCO-BRITANNIQUE SUR L'INNOVATION PAR LES DONNÉES

Dans ce rapport, nous proposons quatre grands axes d'action commune entre la France et le Royaume-Uni exploitant l'expertise des deux pays pour stimuler la croissance économique, gagner en efficacité, améliorer la productivité et accroître le bien-être social. Le premier de ces axes est l'encouragement de l'innovation par les données. Le deuxième est centré sur l'établissement d'infrastructures de données robustes comme préalable à des services et des opportunités

profitant à tous. Le troisième s'articule autour du besoin urgent d'amélioration des compétences et connaissances en matière de données dans nos sociétés. Le quatrième aborde la question essentielle de la confiance, dans les données elles-mêmes et dans la manière dont elles sont utilisées, et la façon de parvenir à une utilisation responsable et proportionnée des données. De nos jours, les données touchent tous les aspects de la société et de l'économie. Leurs bénéfices doivent être répartis de manière juste et équitable.

1

## encouragement de l'innovation par les données

Les plus grandes innovations naissent dans des écosystèmes où règnent la créativité, la coopération, l'interaction et une concurrence loyale. Ensemble, la France et le Royaume-Uni pourraient exploiter leurs organisations et projets existants dans le domaine des données pour construire l'un des écosystèmes de l'économie de la donnée les plus créatifs au monde.

2

## établissement d'infrastructures de données robustes

L'importance stratégique des données, infrastructures indispensables à nos économies, n'est pas encore reconnue par tous. De la même manière que nous investissons dans les routes, les chemins de fer et d'autres infrastructures publiques pour encourager l'activité, créer de nouvelles opportunités et améliorer la vie des citoyens, nous devons investir dans les infrastructures de données.

3

## amélioration des compétences et connaissances en matière de données dans nos sociétés

Nous sommes confrontés à un manque de compétences dans le domaine des données. En Europe, le nombre de postes de data scientists vacants est en constante augmentation. Il nous faut renforcer ces compétences dans les milieux professionnels où les données sont déjà utilisées sous diverses formes : épidémiologistes, statisticiens, médecins, manipulateurs radio, sociologues ou encore spécialistes du marketing. Chaque jour, des opportunités sont perdues car les sociétés et les administrations n'ont pas pleinement conscience de la valeur de leurs données. Les États-Unis et la Chine forment déjà des milliers de data scientists grâce à des programmes éducatifs innovants. La France et le Royaume-Uni doivent se saisir de ce nouveau défi.

4

## Confiance et accroissement du "pouvoir d'agir"

Nous devons réfléchir à la société que nous voulons. Le risque que les données soient utilisées pour opprimer et contrôler les individus, que des organisations aient la maîtrise exclusive des données et puissent les utiliser à des fins monopolistiques, et que nos données à caractère personnel soient utilisées de manière abusive suscite des inquiétudes. L'effondrement de la confiance dans les organisations collectant et utilisant les données, notamment nos données à caractère personnel, est un vrai risque qui pourrait nous empêcher, s'il se concrétisait, de profiter pleinement des bénéfices économiques potentiels de ces données.

## TABLE DES MATIÈRES

INTRODUCTION .....	2
PREMIÈRE PARTIE : LA REVOLUTION DES DONNEES .....	5
DEUXIÈME PARTIE : SYNTHÈSE DES RECOMMANDATIONS .....	9
1. INNOVATION .....	10
2. INFRASTRUCTURE DE DONNEES .....	11
3. COMPETENCES .....	12
4. CONFIANCE ET ACCROISSEMENT DU "POUVOIR D'AGIR" .....	13
TROISIÈME PARTIE : SAISIR LA REVOLUTION DE LA DONNEE .....	14
1. LA CHAÎNE DE VALEUR DES DONNÉES .....	14
Faire naître des entreprises innovantes .....	15
Créer de solides ressources en données .....	16
Utiliser les données pour prendre des décisions .....	17
Générer de plus grands bénéfices économiques .....	18
2. RECOMMANDATIONS .....	19
Innovation .....	19
Infrastructure de données .....	21
Compétences .....	23
Confiance et accroissement du "pouvoir d'agir" .....	25
REMERCIEMENTS .....	27

**1**

**La revolution  
des données**

## LA RÉVOLUTION DES DONNÉES

Nous vivons une « révolution des données ». Avec la dématérialisation d'un nombre croissant de processus et l'apparition de produits et services numériques totalement nouveaux, la quantité de données numériques créées augmente de manière exponentielle. Ces données prennent de plus en plus d'importance pour les entreprises, les consommateurs, les États et les citoyens. Leur traitement devient bon marché et les outils utilisés pour les traiter et les analyser font l'objet d'une innovation constante. De nouvelles politiques

d'utilisation des données voient le jour. La demande de personnel possédant des compétences dans le traitement des données connaît une hausse inexorable. La donnée est désormais le fer de lance de la révolution numérique. Comme les autres révolutions industrielles avant elle, la révolution des données transforme notre économie, notre société et nos politiques.

Au cours de l'histoire, les plus grandes transformations se sont produites lorsqu'une ressource jusque-là rare est devenue abondante : le passage d'une société de chasseur-cueilleur à une société agricole (plus de nourriture), l'invention de l'imprimerie (plus d'instruction), l'émergence de nouveaux processus de fabrication et la révolution industrielle (plus de produits fabriqués en grande série), ou encore le web, qui a donné naissance à notre monde interconnecté (plus de données).

Ces « révolutions » ont systématiquement généré de nouvelles structures sociales, de nouvelles formes de gouvernance, de nouvelles sources de richesse, de nouvelles opportunités et, il faut le souligner, de nouvelles inégalités. C'est ce qui s'est produit dans le passé et ce qui se produit aujourd'hui, avec la révolution des données.

La donnée est parfois comparée à un « nouveau pétrole ». Toutefois, cette image ne rend pas compte de la nature singulière des données et de ce qui rend cette révolution différente des précédentes. À la différence des gisements de pétrole, nos ressources en données ne sont pas limitées et ne diminuent pas avec le temps : plus la dématérialisation touche de services et de personnes, plus nos ressources en données grandissent. C'est un bien dont la jouissance n'implique aucune rivalité : le fait qu'il soit consommé par une personne n'empêche pas sa consommation simultanée par une autre personne. Il peut être produit et traité par de nombreux types d'organisations, à l'aide d'une grande variété d'outils, et plus il est utilisé, plus sa valeur augmente.

De nombreuses études ont cherché à quantifier la valeur de cette profusion de données. En 2015, la Commission européenne a estimé que la taille cumulée du marché direct des données publiques ouvertes pour la période 2016-2020 devrait s'élever à 325 milliards d'euros en Europe<sup>1</sup>. Un rapport du McKinsey Global Institute a conclu, en 2014, que le big data pourrait rapporter 3 000 milliards de dollars chaque année dans seulement sept secteurs<sup>2</sup>. Il s'est avéré difficile de quantifier la valeur des données, car elles n'ont pas une grande valeur intrinsèque : leur valeur réside dans la manière dont on les utilise. La valeur obtenue résulte des services créés grâce à ces données, des nouveaux produits et idées, qui donnent naissance à de nouveaux modèles économiques et à de nouvelles activités pour nous rendre la vie meilleure.

Dans le présent rapport, la Taskforce mentionne de nombreux types de données et d'utilisation des données. Les données que nous collectons, utilisons et partageons dans nos sociétés peuvent être des données « ouvertes », c'est-à-dire auxquelles tout le monde a accès et qui peuvent être utilisées et partagées par tous, des données partagées entre des entités publiques ou privées et dont l'utilisation est autorisée en contrepartie d'une redevance, ou des données « fermées », tenues confidentielles, qui ne sont accessibles qu'à quelques personnes<sup>3</sup>. Il peut s'agir de big data sous licence ouverte ou de big data partagées. L'économie de la donnée du Royaume-Uni et de la France reflète la diversité, et la complexité, de ce spectre des données.

Le Royaume-Uni et la France ont déjà lancé, avec succès, de nombreuses initiatives pour parvenir à une meilleure utilisation des données dans une grande variété de secteurs.

Au Royaume-Uni, plusieurs gouvernements successifs ont financé des organisations et des projets renforçant l'économie numérique et l'économie de la donnée du pays, notamment l'Open Data Institute (ODI), Tech City UK, la Digital Catapult, Tech North et l'Alan Turing Institute. 43 % des licornes européennes (sociétés non cotées, spécialisées dans les nouvelles technologies, valorisées plus d'un milliard de dollars) ont leur siège au Royaume-Uni. Les sociétés de gestion et d'analyse des données représentent 12 % des entreprises du secteur du numérique au Royaume-Uni et connaissent une croissance rapide<sup>4</sup>. De plus, pour la troisième fois consécutive, le Royaume-Uni arrive en tête de l'Open Data Barometer<sup>5</sup> au niveau mondial.

En France, des efforts similaires ont été déployés pour développer à la fois l'offre et la demande dans le secteur de la donnée. La « Nouvelle France Industrielle » a ainsi été lancée en 2013. Elle a permis de mettre en avant et de renforcer des projets existants (la French Tech, le pôle Cap Digital, MesInfos), d'en créer de nouveaux (notamment des appels à projets de R&D et le financement de Challenges Big Data) et de rapprocher l'ensemble des parties prenantes de l'écosystème.

1 <http://www.europeandataportal.eu/en/content/creating-value-through-open-data>.

2 <https://www.ced.org/blog/entry/big-datas-economic-impact>.

3 <https://theodi.org/data-spectrum>.

4 <http://www.techcityuk.com/investors/>.

5 <http://opendatabarometer.org/>.

En 2015, 61 % des sociétés françaises estimaient que le big data était devenu l'un des principaux moteurs de croissance, aussi important pour elles que leurs produits et services existants<sup>6</sup>. L'écosystème de l'économie de la donnée est en forte expansion et, en 2018, le marché français du big data devrait atteindre 652 millions d'euros, soit une hausse de 129 % par rapport à 2014, malgré une conjoncture économique européenne défavorable<sup>7</sup>. En 2014, la France a nommé son premier administrateur général des données et, en 2015, le pays était classé au deuxième rang mondial de l'Open Data Barometer.

Ce rapport arrive à un moment déterminant. Le Royaume-Uni comme la France cherchent à consolider leurs fondations en matière de données et à mieux exploiter les potentialités de l'utilisation des données, dans le secteur public, le secteur privé et le tiers secteur. Dans l'Union européenne, le règlement général sur la protection des données (RGPD) imposera de grands changements à beaucoup d'organisations collectant, conservant et traitant des données à caractère personnel, et nous devons y être préparés. La question de savoir si le Royaume-Uni devra ou non appliquer ce règlement reste indéterminée, mais la Taskforce suggère qu'à la fois le Royaume-Uni et d'autres pays européens pourraient bénéficier de l'adoption de règlements alignés qui permettront une interaction étroite, du commerce et de la collaboration.

Les citoyens britanniques et français sont de plus en plus mobiles et s'attendent à ce que les services dont ils ont besoin soient accessibles en tout lieu. De plus, nous souhaitons exploiter au mieux les données mais, pour cela, notre main-d'œuvre doit disposer des compétences adéquates.

Ce rapport est l'occasion pour le Royaume-Uni et la France d'approfondir leurs liens et d'établir de nouveaux modes de coopération et de partage d'expérience. Les données peuvent être un langage commun et une ressource commune. Les recommandations formulées dans ce rapport s'appuient sur les expériences les plus réussies des deux pays en matière de données.

L'une de ces expériences est celle de l'ouverture des données, qui est une priorité pour nos deux pays. Celle-ci favorise la transparence et la responsabilité, l'efficacité, l'innovation et la création de valeur au niveau économique et social. En France, les adresses sont maintenant disponibles sous la forme de données ouvertes. Ces données font partie de notre quotidien : elles indiquent où nous vivons et nous relient aux services publics et privés dont nous avons besoin. Désormais, tout le monde peut avoir accès à ces données, les interroger et construire des services les utilisant, et ce gratuitement. Ce qui n'était au départ qu'une partie de l'infrastructure postale française est maintenant accessible à tous via le web et fait partie de l'infrastructure de données essentielle du pays. Au Royaume-Uni, l'État a décidé de consacrer 5 millions de livres sterling de son budget 2016 à un projet de création de base de données ouverte contenant les adresses du pays.

Dans ce rapport, nous proposons quatre grands axes d'action commune entre la France et le Royaume-Uni exploitant l'expertise des deux pays pour stimuler la croissance économique, gagner en efficacité, améliorer la productivité et accroître le bien-être social. Le premier de ces axes est l'encouragement de l'innovation par les données. Le deuxième est centré sur l'établissement d'infrastructures de données robustes comme préalable à des services et des opportunités profitant à tous. Le troisième s'articule autour du besoin urgent d'amélioration des compétences et connaissances en matière de données dans nos sociétés. Le quatrième aborde la question essentielle de la confiance, dans les données elles-mêmes et dans la manière dont elles sont utilisées, et la façon de parvenir à une utilisation responsable et proportionnée des données. De nos jours, les données touchent tous les aspects de la société et de l'économie. Leurs bénéfices doivent être répartis de manière juste et équitable.

Pour tirer pleinement profit des données, il nous faut encourager l'innovation par les données. Les plus grandes innovations naissent dans des écosystèmes où règnent la créativité, la coopération, l'interaction et une concurrence loyale. Ensemble, la France et le Royaume-Uni pourraient exploiter leurs organisations et projets existants dans le domaine des données pour construire l'un des écosystèmes de l'économie de la donnée les plus créatifs au monde.

Cette ère d'abondance des données verra se produire des changements dans des domaines tels que l'emploi, la santé, les transports, la finance, l'alimentation et l'agriculture, l'énergie et l'environnement, l'éducation, la sécurité, la lutte contre la criminalité et la justice.

Au Royaume-Uni, l'utilisation de données ouvertes a permis de réduire les infections nosocomiales de 85 % en deux ans. Dans le secteur public, le ministère du travail et des retraites et le ministère de l'énergie et du changement climatique, en collaboration avec le gouvernement écossais et onze grands fournisseurs d'énergie, utilisent les données depuis plusieurs années pour lutter contre la précarité énergétique : celles-ci sont exploitées pour accorder à 1,4 million de ménages vulnérables des réductions sur leurs factures énergétiques. Par ailleurs, la mise à disposition de données sur les transports publics a entraîné la création d'applications qui modifient notre façon de nous déplacer à Londres.

En France, les données sur la consommation d'électricité ont permis à l'État d'économiser des centaines de millions d'euros. Au Royaume-Uni et en France, l'analyse des données, par exemple pour prévoir les achats des consommateurs, adapter les modes de déplacement proposés et réduire considérablement les délais de mise sur le marché de nouveaux produits, est une activité plus que rentable pour certaines sociétés. Nous voulons mettre en œuvre un programme collaboratif visant à rechercher et à encourager des solutions innovantes basées sur les données qui permettent non seulement de faire des économies mais aussi de dégager des bénéfices.

---

6 Markess, 2015.

7 IDC, 2015.

Nous devons considérer les données comme des infrastructures essentielles pour nos activités économiques émergentes. L'importance stratégique des données, infrastructures indispensables à nos économies, n'est pas encore reconnue par tous. De la même manière que nous investissons dans les routes, les chemins de fer et d'autres infrastructures publiques pour encourager l'activité, créer de nouvelles opportunités et améliorer la vie des citoyens, nous devons investir dans les infrastructures de données.

Deux caractéristiques des données révèlent la profonde différence qui existe entre cette révolution et les précédentes. Au XIXe siècle, l'extraction et l'utilisation massive de matières premières (charbon, pétrole, fer et acier) à grande échelle a conduit à la révolution industrielle. Au XXe siècle, la distribution généralisée d'énergie par le biais des réseaux d'alimentation électrique a contribué à entraîner une révolution technologique. Dans les deux cas, il s'agissait de biens dont la jouissance impliquait une rivalité : si vous utilisiez du combustible fossile, vous empêchiez toute autre personne d'utiliser le même combustible et, une fois utilisé, ce combustible ne pouvait plus l'être par qui que ce soit et n'avait plus de valeur.

Les données sont un bien fondamentalement différent. En principe, le fait que vous utilisiez des données n'empêche pas une autre personne de le faire. Par ailleurs, elles ne sont pas détruites lors de leur utilisation : leur utilisation répétée augmente même leur valeur et peut conduire à la création de nouvelles formes de la ressource (métadonnées, par exemple), qui lui donnent encore plus de valeur. C'est la raison pour laquelle les infrastructures de données devraient être aussi ouvertes et accessibles que possible.

Nous avons besoin de promouvoir la connaissance des données, de leur potentiel et de leurs limites dans toute la société, et de plus grandes compétences en sciences des données.

Nous sommes confrontés à un manque de compétences dans le domaine des données. En Europe, le nombre de postes de data scientists vacants est en constante augmentation. Il nous faut renforcer ces compétences dans les milieux professionnels où les données sont déjà utilisées sous diverses formes : épidémiologistes, statisticiens, médecins, manipulateurs radio, sociologues ou encore spécialistes du marketing. Chaque jour, des opportunités sont perdues car les sociétés et les administrations n'ont pas pleinement conscience de la valeur de leurs données. Les États-Unis et la Chine forment déjà des milliers de data scientists grâce à des programmes éducatifs innovants. La France et le Royaume-Uni doivent se saisir de ce nouveau défi.

Il sera essentiel de mettre en relation les professionnels des données et data scientists et les employeurs pour tirer profit de cette abondance de données. Toutefois, il est tout aussi important d'aider les décideurs du secteur privé et du secteur public à savoir détecter les véritables défis et opportunités liés aux données et, dans cette perspective, de les doter des compétences dont ils ont besoin pour faire exécuter des travaux de sciences des données, exploiter ces travaux et les encourager. Sans décideurs capables d'organiser et de planifier les transformations liées aux données, en détectant les opportunités et en anticipant les risques, nous ne pourrions pas exploiter pleinement le potentiel de la révolution des données.

En plus de la formation de nos décideurs, le développement des compétences nécessaires pour évaluer, exploiter et utiliser les données doit être au cœur de toute stratégie économique. Par ailleurs, des compétences, des connaissances et un savoir-faire de base dans le domaine des données doivent être enseignés dans les écoles en complément de l'éducation civique numérique, de la sécurité informatique et de la programmation.

Toute action liée aux données doit reposer sur la confiance et la responsabilité. Nous devons réfléchir à la société que nous voulons. Le risque que les données soient utilisées pour opprimer et contrôler les individus, que des organisations aient la maîtrise exclusive des données et puissent les utiliser à des fins monopolistiques, et que nos données à caractère personnel soient utilisées de manière abusive suscite des inquiétudes. L'effondrement de la confiance dans les organisations collectant et utilisant les données, notamment nos données à caractère personnel, est un vrai risque qui pourrait nous empêcher, s'il se concrétisait, de profiter pleinement des bénéfices économiques potentiels de ces données.

Le Royaume-Uni et la France sont à l'avant-garde pour ce qui est des politiques et programmes mis en place afin de donner aux citoyens une meilleure maîtrise de la manière dont leurs données sont collectées et utilisées. Ces politiques et programmes méritent un examen et une mise en commun afin d'établir des bonnes pratiques en la matière. Par ailleurs, la transparence doit être au cœur de la collecte et du traitement des données. Les dispositions du RGPD<sup>8</sup> en font une priorité. Nous avons besoin de codes de conduite et d'exemples de bonnes pratiques pour aider les organisations à respecter ce règlement.

Nous devons faciliter l'accès des citoyens aux services dont ils ont besoin, par voie électronique et sans limite de frontières, tout en préservant leur vie privée. En outre, nous devons anticiper les questions qui se présenteront à l'avenir et façonneront la manière dont nos données à caractère personnel seront utilisées pour prendre des décisions à notre sujet, comme la question de la transparence des algorithmes, qui se pose déjà dans certains secteurs. Par exemple, les agences de notation financière sont tenues de rendre compte de la manière dont elles attribuent leurs notes.

C'est dans ce contexte que la Taskforce a entrepris ses travaux. Nous avons identifié de grands axes potentiels d'action commune entre le Royaume-Uni et la France, qui exploiteraient les capacités considérables des deux pays dans le domaine des données. Nous pensons que nos recommandations, en plus de bénéficier à nos deux pays, pourraient être utilisées pour faire avancer les choses dans toute l'Europe et même au-delà. La révolution des données n'est pas circonscrite à une région ou un pays particuliers. Si nous saisissons ces opportunités, non seulement nos économies et nos sociétés en bénéficieront mais nous pourrions également servir de sources d'inspiration à d'autres.

---

8 [http://ec.europa.eu/justice/data-protection/reform/index\\_en.htm](http://ec.europa.eu/justice/data-protection/reform/index_en.htm).



**2**

## **Synthèse des recommandations**

## SYNTHÈSE DES RECOMMANDATIONS

Nous faisons ici la synthèse de nos recommandations, en les regroupant en fonction des grands enjeux en matière de données : comment innover par les données, mettre en place des infrastructures de données efficaces, développer les compétences en matière de données et veiller à ce que l'écosystème des données suscite la confiance et accroisse notre pouvoir d'agir. Une description plus détaillée de ces recommandations figure dans la section suivante, intitulée "Saisir la révolution de la donnée". Une feuille de route décrivant la manière dont elles pourraient être mises en œuvre sera présentée à l'automne, après échanges entre nos deux gouvernements. Par la suite, le Royaume-Uni et la France devraient mettre en place un sommet annuel sur les données pour présenter les initiatives concluantes, partager leurs expériences et suivre l'avancement de la mise en œuvre de ces recommandations.



### 1. INNOVATION

*Objectif : encourager la collaboration et concentrer nos efforts.*

**Recommandation : mettre en place un programme franco-britannique rassemblant des organisations de nos deux pays dans une recherche de solutions, par les données, à des problèmes communs.**

Lancer un nouveau programme conjoint axé chaque année sur deux problématiques majeures, avec l'objectif affirmé de présenter des améliorations claires et mesurables. Etalab, Cap Digital, l'Open Data Institute et la Digital Catapult pourraient être partenaires de ces travaux. Les objectifs pourraient être, par exemple, de tirer des avantages économiques, de parvenir à une plus grande efficacité, de réduire les goulets d'étranglement et de créer de nouvelles opportunités d'innovation dans des domaines comme l'emploi, la santé, les transports, la finance, l'alimentation et l'agriculture, l'énergie et l'environnement, l'éducation, la lutte contre la criminalité et la justice.

*Objectif : développer nos réseaux.*

**Recommandation : créer un « passeport pour l'innovation par les données » afin de partager les bonnes pratiques et d'étendre les réseaux de startups entre le Royaume-Uni et la France.**

Exploiter, soutenir et mettre en relation les programmes existants d'accompagnement de startups au Royaume-Uni et en France, par le biais d'un accélérateur. Les startups participant à ce programme pourraient passer d'une pépinière d'entreprises à une autre, telles que l'Open Data Institute, la Digital Catapult, Cap Digital, Etalab, Numa, Tech City UK et Scale Up!, et apprendre ainsi à connaître les écosystèmes et les marchés français et britanniques de manière à se développer plus facilement.

*Objectif : créer des échanges au niveau régional entre les innovateurs en matière de données et sur l'innovation par les données.*

**Recommandation : mettre en place un programme de jumelage dans le domaine des données entre des villes situées dans des régions dotées d'un tissu économique et de caractéristiques démographiques similaires, et de solides communautés de startups.**

Les acteurs de l'innovation par les données ne se trouvent pas seulement à Londres et à Paris. En région, des pôles d'activité et des pépinières d'entreprises soutiennent les entreprises britanniques et françaises hors des deux capitales, comme la French Tech à Lille, Avignon-Provence French Tech Culture, Tech North à Manchester et le Leeds Data Mill. La Mission French Tech et Tech City UK pourraient être les initiateurs de cette nouvelle collaboration entre les régions.



## 2. INFRASTRUCTURES DE DONNÉES

*Objectif : faciliter l'utilisation et l'échange de données.*

**Recommandation : développer l'utilisation d'interfaces de programmation (API) dans l'économie de la donnée, en s'appuyant sur l'expérience britannique d'open banking.**

Les API sont le système nerveux du XXI<sup>e</sup> siècle. Nous devons donc encourager le développement d'API ouvertes communes dans des secteurs importants pour les deux pays et collaborer en la matière. De la même façon que les échanges commerciaux dépassent les frontières, l'échange et l'accessibilité des données doivent également être fluides. Le Royaume-Uni a déjà commencé à travailler sur une norme de système bancaire ouvert pour le secteur des services destinés aux particuliers. Ces travaux devraient être étendus.

*Objectif : développer l'infrastructure de base des services interconnectés.*

**Recommandation : investir dans la création de registres de données de référence, en s'appuyant sur l'expérience française de la Base Adresse Nationale, et partager les expériences en la matière.**

Les registres constituent une source de données de référence faisant autorité qui permet d'interconnecter les services et de les améliorer. La France, en collaboration avec la communauté OpenStreetMap, a élaboré la Base Adresse Nationale, un bien commun basé sur la production participative de données, librement accessible et utilisable par tous. Le Royaume-Uni et la France doivent coopérer et partager les leçons tirées de cette expérience, alors que le Royaume-Uni envisage de créer sa propre base d'adresses ouverte.

*Objectif : faciliter l'accès des citoyens aux services dont ils ont besoin, où qu'ils se trouvent.*

**Recommandation : poursuivre les investissements dans des programmes d'identification électronique pour faciliter l'accès aux services publics et privés dans les deux pays, et partager les expériences en la matière.**

La France et le Royaume-Uni œuvrent au renforcement des mécanismes publics transversaux, comme celui de l'identification électronique, qui sont utilisés dans tous les services publics. La protection de la vie privée doit être intégrée au système public dès sa conception, en s'appuyant sur les enseignements tirés de projets tels que France Connect (France) et Verify (Royaume-Uni).



### 3. COMPETENCES

*Objectif : aider les dirigeants à prendre conscience du potentiel des données.*

**Recommandation : développer les programmes d'éducation aux données destinés aux dirigeants du public et du privé.**

Pour mettre en œuvre à grande échelle la transformation numérique, nous devons agir au sommet. Il pourrait être envisagé de développer les formations destinées aux « cadres à haut potentiel » (dirigeants d'entreprise, responsables politiques, hauts fonctionnaires, etc.) pour leur apporter et leur permettre de développer les connaissances, les compétences et les stratégies en matière de données qui leur seront nécessaires pour transformer leurs organisations. Ces dirigeants peuvent appartenir à des entités du secteur public ou du secteur privé.

*Objectif : surmonter la pénurie chronique de compétences en matière de données.*

**Recommandation : intégrer aux formations d'analystes existantes l'enseignement de nouvelles compétences en matière de données et de résolution de problèmes, et créer des programmes d'apprentissage et des stages de reconversion communs afin de fournir au Royaume-Uni et à la France un nombre suffisant de professionnels des données.**

Il serait profitable d'intégrer aux programmes d'enseignement des disciplines utilisant les données (comme la gestion, le marketing, les sciences de l'ingénierie, l'économie et la statistique) une formation de base sur les données, la résolution de problèmes, la programmation collaborative et l'éthique et la visualisation des données. En parallèle, il serait souhaitable de créer de nouvelles voies d'entrée pour les futurs professionnels des données. En France comme au Royaume-Uni, il existe des initiatives dont l'objet est d'enrichir les compétences des développeurs et des data scientists, mais elles sont généralement axées sur la sécurité et la recherche de pointe. Certains secteurs ont des difficultés à attirer et/ou exploiter les compétences en matière de données. Il est nécessaire d'investir dans des programmes d'apprentissage et des stages de reconversion destinés aux futurs professionnels des données. Le nombre de formations, stages et autres formes d'éducation aux données doit considérablement augmenter. C'est pourquoi il est important de « former les formateurs ». Ces initiatives pourraient être combinées à un programme de mise à disposition, qui permettrait à des professionnels des données de travailler dans des organisations qui souhaitent, sur un projet particulier, bénéficier d'une assistance en matière de données.

*Objectif : promouvoir l'acquisition de connaissances de base en matière de données.*

**Recommandation : mettre en œuvre des projets destinés à dispenser des connaissances de base sur les données aux élèves du primaire et du secondaire.**

De la même manière que les élèves reçoivent une éducation civique et sont sensibilisés aux risques liés à l'internet, ils doivent aussi être sensibilisés aux données si nous voulons faire des choix plus éclairés sur l'utilisation de nos données par les organisations. Plusieurs expériences en ce sens se sont révélées efficaces et pourraient facilement être soutenues et étendues. Ainsi, france-ioi.org propose gratuitement des activités et des concours sous une approche ludique, indépendante et participative. Une présentation générale du potentiel que représentent les données pourrait être intégrée au programme des cours d'éducation morale et civique.



## 4. CONFIANCE ET ACCROISSEMENT DU “POUVOIR D’AGIR”

*Objectif : renforcer le pouvoir d'action des citoyens sur leurs données.*

**Recommandation : améliorer et étendre les bonnes pratiques issues des projets MesInfos et MiData pour favoriser le renforcement du pouvoir d'action des citoyens sur leurs données.**

La propriété, la portabilité et la maîtrise des données ainsi que la responsabilité en la matière sont les principes fondamentaux inscrits dans le RGPD et au cœur des projets MesInfos et MiData. La manière dont les organisations se conformeront dans la pratique au futur RGPD demeure incertaine. Une initiative visant à accroître le pouvoir d'action des citoyens sur leurs données pourrait être lancée en s'appuyant sur les expériences MesInfos (France) et MiData (Royaume-Uni). La recherche dans les technologies et la réglementation susceptibles d'aider les entreprises à donner aux citoyens les moyens d'avoir une plus grande maîtrise sur leurs données à caractère personnel et de prendre des décisions en la matière doit être encouragée. De nouvelles expériences devraient également être menées pour faire de ces nouvelles règles une chance à saisir pour les entreprises.

*Objectif : anticiper les enjeux futurs en matière de politique des données.*

**Recommandation : mener des recherches sur la transparence des algorithmes et la responsabilité.**

Les décideurs et les organisations sont confrontés à de nouvelles questions portant sur la prise de décisions basée sur les données, par exemple au moyen d'algorithmes. Dans certains secteurs, cette pratique est déjà intégrée aux services que nous utilisons au quotidien. Comment s'assurer que notre processus décisionnel fondé sur les données est équitable et n'entraîne pas de discriminations ? La France et le Royaume-Uni devraient lancer des travaux de recherche sur les moyens de rendre les organisations plus transparentes par rapport aux algorithmes, aux sources de données et au processus d'analyse des données qu'elles utilisent pour prendre des décisions qui touchent les consommateurs et les citoyens.

# 3

## Saisir la révolution de la donnée

### La chaîne de valeur des données

La révolution des données peut être mise au service de la croissance économique et transformer les entreprises, le vécu des consommateurs et les services publics. Si elles représentent des opportunités, la disponibilité et l'utilisation croissantes des données pour créer de la valeur représente également d'importants défis.

Il nous faut comprendre les mécanismes de création de valeur dans ce domaine en pleine évolution si nous voulons que toutes les transformations bénéfiques potentielles de la révolution des données se concrétisent.



## FAIRE NAÎTRE DES ENTREPRISES INNOVANTES

L'accès aux données donne l'occasion de créer de la valeur par l'innovation. Les données peuvent être exploitées par les organisations comme un actif ou comme une source d'externalités positives. Leur utilisation en tant qu'actif entraîne des changements positifs au sein de l'organisation tandis que leur utilisation comme source d'externalités positives a un impact plus large sur l'économie et la société dans son ensemble. Pour générer ces externalités, il convient de fluidifier la circulation des données, en les partageant ou en les rendant ouvertes.

Les données deviennent un actif :

- lorsqu'elles sont utilisées pour améliorer l'allocation des ressources ou les processus et pour réduire les coûts de transaction dans un objectif de gains d'efficacité ;
- lorsqu'elles sont utilisées pour améliorer et personnaliser les services afin d'en accroître la qualité ;
- lorsqu'elles sont utilisées comme élément de l'analyse prédictive visant à étayer la prise de décision et à améliorer la pertinence de cette dernière ;
- lorsqu'elles sont utilisées pour mener des recherches en sciences sociales et établir des modèles comportementaux.

La vente de l'accès aux données crée de la valeur sous forme de revenus directs pour les sociétés ou administrations qui les possèdent mais s'est révélée globalement inefficace, en particulier pour les informations du secteur public. En Australie, en 2002, les données géographiques ont été ouvertes. Après prise en compte de la perte de revenus directs, le gain annuel de cette ouverture a été estimé à 1,7 million de dollars<sup>1</sup>. S'agissant de la tarification des données entre administrations, les coûts de transaction sont proportionnellement plus élevés et les échanges existants sont contraires à la rationalité économique<sup>2</sup>.

Les données créent aussi de la valeur en aidant la prise de décision, ce qui permet d'adopter de meilleures stratégies et de dégager des gains financiers concrets. À titre d'exemple, en France, l'administrateur général des données a réalisé une analyse des consommations d'électricité des bâtiments publics<sup>3</sup>, dans le cadre d'une collaboration avec le service des achats de l'État, et a déterminé sur cette base les meilleures stratégies d'approvisionnement, au moment des discussions du gouvernement au sujet de l'ouverture à la concurrence du marché de l'électricité. La stratégie ainsi dégagée a permis d'économiser plusieurs centaines de millions d'euros.

---

1 « Re-use of public sector information. Report for Danish Ministry for Housing, Urban and Rural Affairs », Marc de Vries, 2012.

2 Rapport Fouilleron.

3 <https://agd.data.gouv.fr/2015/05/17/analyser-les-consommations-energetiques-des-batiments-publics/>.



## CRÉER DE SOLIDES RESSOURCES EN DONNÉES

Lorsque l'on analyse la création de valeur par les données, la première chose à faire est d'évaluer les infrastructures de données existantes. Les données essentielles sont-elles présentes ? Les données disponibles répondent-elles aux usages souhaités ? Certains jeux de données sont essentiels car ils servent de référence lors du traitement et de la combinaison de données provenant de différentes sources. Le registre des sociétés, le cadastre et la base de données nationale des adresses sont des exemples notables d'infrastructures de données existantes. La base de données nationale des adresses permet la géolocalisation d'adresses officielles faisant autorité, sur laquelle s'appuient de très nombreuses applications.

La disponibilité, la visibilité et la qualité des données sont des aspects qui influent directement sur l'efficacité de notre infrastructure de données. La disponibilité des données (par ouverture ou partage) dépend de l'entité qui les possède, les produit et les finance. Par ailleurs, il peut être difficile de répertorier les ressources en données disponibles lorsque leur visibilité est insuffisante. Enfin, même lorsqu'elles sont disponibles et faciles à trouver, les données peuvent être de mauvaise qualité, c'est-à-dire inexactes, mal documentées ou trop rarement mises à jour, ce qui les rend peut-être aptes à être réutilisées de manière efficace.

Après avoir traité les questions relatives aux infrastructures de données, il convient de s'intéresser à la manière de rendre les données exploitables. L'accessibilité réelle requiert des formats ouverts et lisibles par la machine. La mise à disposition des données par le biais d'API peut augmenter considérablement la valeur et les bénéfices générés grâce aux données. En outre, l'interopérabilité, obtenue par une normalisation des données facilitant leur transfert et leur comparaison, est une condition préalable à une vaste réutilisation des données et à la réalisation d'économies d'échelle. Dans certains cas, les données doivent être anonymisées avant d'être utilisées ou partagées, de manière à conserver les données utiles tout en protégeant les données à caractère personnel et en limitant les possibilités d'identification des personnes.







## UTILISER LES DONNÉES POUR PRENDRE DES DÉCISIONS

API, interopérabilité, normalisation, anonymisation, ... On le voit : parvenir à une bonne utilisation des données nécessite des compétences très spécialisées. C'est la raison pour laquelle les data scientists, statisticiens innovants qui utilisent les mathématiques et l'informatique aux fins du traitement des données, sont devenus des acteurs clés de la conversion des données brutes en actions concrètes. Ils sont de plus en plus amenés à assumer des fonctions de (quasi) décideurs, comme en témoigne le nombre croissant d'administrateurs des données dans les organisations publiques et privées.

Les compétences en matière de données sont nécessaires non seulement au sein des organisations mais aussi à l'extérieur. Les citoyens doivent être capables d'utiliser les données ouvertes et de tirer profit de la transparence des algorithmes et doivent disposer d'une culture générale des données si l'on veut garantir une responsabilité effective, la protection des consommateurs et une citoyenneté active.





## GÉNÉRER DE PLUS GRANDS BÉNÉFICES ÉCONOMIQUES

Les données deviennent une source d'externalités positives :

- lorsqu'elles sont utilisées pour encourager l'innovation par la création de nouveaux services et de nouveaux modèles économiques ;
- lorsqu'elles créent de nouveaux marchés et permettent leur régulation ;
- lorsqu'elles constituent une source de développement en produisant des effets sociaux positifs et en améliorant la rationalité économique des agents ;
- lorsqu'en réduisant l'asymétrie d'information elles constituent une source de transparence ;
- lorsqu'elles sont utilisées pour promouvoir la coopération par la création de communs contributifs et l'exploitation d'économies d'échelle.

L'ouverture des données a un effet positif sur l'innovation. Au Royaume-Uni, le rapport « Open data means business »<sup>1</sup> de l'Open Data Institute a recensé et analysé 270 sociétés qui utilisent et produisent des données ouvertes ou investissent dans ces données dans le cadre de leur activité. Ces sociétés représentent un chiffre d'affaires annuel global de 92 milliards de livres sterling et plus de cinq cent mille salariés. En France, le concours Dataconnexions<sup>2</sup>, organisé par Etalab depuis 2012, a permis de recenser 200 startups à haut potentiel. En Espagne, une étude a montré que le secteur des « infomédiaires » (sociétés qui vendent des services fondés sur les données ouvertes) comprenait au moins 150 entreprises, employait environ 4 000 personnes et générait annuellement entre 330 et 350 millions d'euros de chiffre d'affaires<sup>3</sup>.

Les données peuvent donner naissance à des marchés nouveaux et permettre leur régulation. À titre d'exemple, dans le secteur de la mobilité, avec Uber et Blablacar, et dans celui de l'hébergement de courte durée, avec Airbnb et Bedycasa, les données ont permis de créer de nouveaux modèles économiques. Dans ce contexte, les données ouvertes peuvent redéfinir le périmètre d'un marché. En France, le projet « Le Taxi »<sup>4</sup> a permis aux taxis de se battre à armes égales contre Uber grâce aux données ouvertes sur la géolocalisation des taxis et aux API ouvertes permettant de réserver un véhicule, créant ainsi une nouvelle stratégie de régulation du secteur.

Les données peuvent devenir un facteur du développement international, comme l'illustre nettement la « révolution des données »<sup>5</sup> lancée par les Nations Unies. En particulier, l'existence de données sur les parcelles cadastrales peut créer de la richesse. Elle favorise l'émergence d'une classe moyenne dans les pays en développement, car elle permet la collecte de l'impôt et un meilleur accès aux prêts et aux investissements grâce aux titres de propriété.

Enfin, la très dynamique communauté OpenStreetMap (OSM) constitue un parfait exemple de création de communs contributifs et d'économies d'échelle grâce à la collaboration à l'échelon national comme à l'échelon mondial. Les données issues d'une production collaborative se sont avérées être d'une qualité au moins égale à celle des statistiques officielles. La collaboration permet de réduire au minimum les activités redondantes entre les secteurs public et privé. OSM a conduit des projets très différents, de la construction de biens publics comme la Base Adresse Nationale<sup>6</sup> en France jusqu'au soutien logistique de gestion de crise grâce à des cartes lors des tremblements de terre en Afghanistan et au Népal ou de l'épidémie Ébola en Afrique de l'Ouest<sup>7</sup>.

Les mécanismes de création de valeur par les données sont très divers et peuvent être mis en œuvre avec plus ou moins de facilité en fonction du contexte, des acteurs concernés et des ressources disponibles. Néanmoins, pour tirer parti de la révolution des données, certaines méthodes, certaines compétences et un certain état d'esprit sont essentiels. L'objectif de la Taskforce est de recenser les aspects sur lesquels le Royaume-Uni et la France pourraient collaborer concrètement pour mobiliser et renforcer ces leviers essentiels, et de mettre en œuvre une telle collaboration.

1 <http://theodi.org/open-data-means-business>.

2 <https://www.data.gouv.fr/fr/dataconnexions-6>.

3 Rapport annuel du portail espagnol des données ouvertes, « Characterization Study of the Infomediary Sector », juillet 2012.

4 <http://le.taxi/>.

5 <http://www.undatarevolution.org/>

6 <https://adresse.data.gouv.fr/>.

7 <https://hotosm.org/projects/ongoing>.



## 1. INNOVATION

**Ensemble, le Royaume-Uni et la France peuvent développer l'un des écosystèmes de l'économie de la donnée les plus créatifs et dynamiques au monde.**

Au Royaume-Uni, les entreprises spécialisées dans la gestion et l'analyse des données représentent déjà 12 % des entreprises du numérique et cette proportion croît rapidement<sup>1</sup>. En France, le big data, associé à l'émergence d'un internet des objets, pourrait représenter 3,6 % du PIB en 2020 (et près de 7 % en 2025)<sup>2</sup>. Les sociétés qui proposent des solutions, des outils et des services liés aux données connaissent un développement continu qui est voué à avoir un effet transformateur sur chaque secteur d'activité. Les sociétés faisant un usage efficace des données sont jusqu'à 13 % plus productives et ont jusqu'à 40 % plus de chances de lancer de nouveaux produits que leurs concurrentes<sup>3</sup>.

Les startups qui appliquent de nouvelles formes de gestion et d'analyse des données tirent l'innovation dans ce domaine. Au Royaume-Uni et en France, toute une série de programmes et d'initiatives apportent un soutien aux startups et aux PME du secteur de la donnée, dans les capitales, avec des organisations comme Tech City UK, l'Open Data Institute et la Digital Catapult à Londres et Etalab, Cap Digital et NUMA à Paris, mais aussi en région (Tech North et le Leeds Data Mill au Royaume-Uni et toutes les Métropoles French Tech, notamment Aix-Marseille et Digital Grenoble, en France).

Aujourd'hui, les sociétés en réseau s'étendent au-delà des frontières. Les équipes de spécialistes des technologies numériques sont de plus en plus réparties dans différentes villes et différents pays et ont recours à des développeurs de talent partout dans le monde. Le Royaume-Uni et la France doivent continuer à réduire les difficultés que rencontrent les entreprises pour trouver les compétences nécessaires en matière de données et pour étendre leur activité à de nouvelles régions et de nouveaux secteurs. Les deux pays doivent leur apporter le soutien et les aides financières indispensables à l'innovation.

Si nous voulons encourager des économies de la donnée plus ouvertes et collaboratives, il faut que nous nous fixions des objectifs communs et que nous instaurions des modes de communication et de partage d'expérience. La Taskforce formule donc les recommandations suivantes :

*Objectif : encourager la collaboration et concentrer nos efforts.*

**Recommandation : mettre en place un programme franco-britannique rassemblant des organisations de nos deux pays dans une recherche de solutions, par les données, à des problèmes communs.**

Lancer un nouveau programme conjoint axé chaque année sur deux problématiques majeures, avec l'objectif affirmé de présenter des améliorations claires et mesurables. Etalab, Cap Digital, l'Open Data Institute et la Digital Catapult pourraient être partenaires de ces travaux. Les objectifs pourraient être, par exemple, de tirer des avantages économiques, de parvenir à une plus grande efficacité, de réduire les goulets d'étranglement et de créer de nouvelles opportunités d'innovation dans des domaines comme l'emploi, la santé, les transports, la finance, l'alimentation et l'agriculture, l'énergie et l'environnement, l'éducation, la lutte contre la criminalité et la justice.

1 Tech Nation 2016.

2 Big Data et Objets Connectés, Institut Montaigne, 2015.

3 Nesta, Rise of the Datavores, 12.

En 2013, le ministère britannique des entreprises, de l'innovation et des compétences (BIS) a investi 1,2 million de livres dans un programme, axé sur l'utilisation des données ouvertes, destiné à faire émerger des solutions innovantes et durables à sept problématiques sociales (Open Data Challenge Series). Une étude menée en 2015 a montré que ce programme avait rapporté à l'économie britannique de cinq à dix fois plus qu'il n'avait coûté<sup>4</sup>.

En France, deux grands programmes ont été mis en œuvre pour promouvoir la culture des données : le concours Dataconnexions (créé par Etalab) et les Challenges Big Data (créés dans le cadre de la Nouvelle France Industrielle). Plus de vingt-cinq Challenges Big Data ont été organisés par cinq entités (Cap Digital, Numa, Images & Réseaux, Euratechnologies et TUBÀ Lyon) depuis novembre 2014 pour aider les PME à répondre aux défis de parrainages publics et privés. L'État français investira 2,8 millions d'euros ces deux prochaines années pour soutenir ces Challenges Big Data. La France a également lancé des Challenges Numériques, dans le cadre desquels près de cent concours seront organisés par douze organisations, pour un budget de 7,7 millions d'euros, jusqu'en 2019.

On pourrait imaginer l'organisation d'une initiative similaire entre le Royaume-Uni et la France utilisant à la fois des données fermées, des données partagées et des données ouvertes pour apporter des solutions à des problèmes communs au Royaume-Uni et à la France, notamment :

- **la gestion des transports publics et la régulation du trafic dans les villes en pleine expansion;**
- **l'amélioration des résultats sanitaires pour les usagers utilisant les données;**
- **la transformation des établissements bancaires proposant des services aux particuliers.**

L'organisation de concours sur les données rassemblant des startups et des entrepreneurs britanniques et français serait un catalyseur de changement dans nos modes de coopération, mais cela suppose de forger une identité partagée, transfrontière.

*Objectif : développer nos réseaux.*

**Recommandation : créer un « passeport pour l'innovation par les données » afin de partager les bonnes pratiques et d'étendre les réseaux de startups entre le Royaume-Uni et la France.**

Exploiter, soutenir et mettre en relation les programmes existants d'accompagnement de startups au Royaume-Uni et en France, par le biais d'un accélérateur. Les startups participant à ce programme pourraient passer d'une pépinière d'entreprises à une autre, telles que l'Open Data Institute, la Digital Catapult, Cap Digital, Etalab, Numa, Tech City UK et Scale Up!, et apprendre ainsi à connaître les écosystèmes et les marchés français et britanniques de manière à se développer plus facilement.

Ce programme vise à tirer parti du soutien entre pairs et à mettre en place un réseau virtuel. Toute pépinière d'entreprises ou startup qui mène des activités dans le domaine des données pourra prétendre à la remise du passeport. Cette idée pourra donner lieu à la création d'un réseau virtuel des startups faisant partie de l'écosystème de l'économie de la donnée.

*Objectif : créer des échanges au niveau régional entre les innovateurs en matière de données et sur l'innovation par les données.*

**Recommandation : mettre en place un programme de jumelage dans le domaine des données entre des villes situées dans des régions dotées d'un tissu économique et de caractéristiques démographiques similaires, et de solides communautés de startups.**

Les acteurs de l'innovation par les données ne se trouvent pas seulement à Londres et à Paris. En région, des pôles d'activité et des pépinières d'entreprises soutiennent les entreprises britanniques et françaises hors des deux capitales, comme la French Tech à Lille, Avignon-Provence French Tech Culture, Tech North à Manchester et le Leeds Data Mill. La Mission French Tech et Tech City UK pourraient être les initiateurs de cette nouvelle collaboration entre les régions.

La Mission French Tech et Tech City UK pourraient travailler avec les villes souhaitant partager les ressources et les atouts de leur écosystème dans des domaines particuliers (transport, tourisme, culture, agriculture, etc.).

<sup>4</sup> <https://theodi.org/news/investment-in-open-data-challenge-series-could-see-10fold-return-to-uk-economy-over-3-years>.



## 2. INFRASTRUCTURES DE DONNÉES

**Les données sont des infrastructures.** Les infrastructures de données deviendront de plus en plus incontournables à mesure que la population de nos deux pays augmentera et que nos économies et nos sociétés seront de plus en plus dépendantes de la valeur qui peut être tirée des données.

Selon l'Open Data Institute, une infrastructure de données « est composée de données en tant que telles, d'organisations qui les exploitent et en assurent la maintenance, et de guides qui décrivent comment utiliser et gérer ces données. Pour être fiable, une infrastructure de données doit être financée de manière soutenable et faire l'objet d'une surveillance permettant d'optimiser l'utilisation et la valeur des données en répondant aux besoins de la société. **Une infrastructure de données, c'est à la fois de la technologie, des processus et une organisation** ».

Il existe des infrastructures de données qui sont détenues et gérées par des organisations du secteur public. De la même façon, il existe des données de grande valeur qui sont détenues et gérées par le secteur privé. Comme les infrastructures physiques, telles que les routes, les chemins de fer ou les infrastructures de transport aérien, le modèle de propriété des infrastructures de données est mixte.

Les données doivent être **fluides et doivent pouvoir circuler librement là où il existe des besoins en services. C'est l'une des conditions préalables pour que l'économie de la donnée puisse produire de la croissance, renforcer la coopération et créer des liens entre les services.** Pour peu qu'elle soit robuste, l'infrastructure de données est un facteur de création de valeur (utilisation et réutilisation de données provenant de tous les secteurs), vient en appui à la communauté existante des acteurs de la donnée et contribue à son développement.

**La propriété et le stockage local des données** sont des éléments importants à prendre en compte pour déterminer la façon dont doit être construite et maintenue une infrastructure de données dont les couches sous-jacentes sont bien sécurisées. Nous devons pouvoir faire des choix quant au lieu de stockage des données et aux personnes y ayant accès qui facilitent la fourniture de services tout en garantissant le respect de la vie privée et un contrôle adéquat, sans pour autant entraver la flexibilité et l'innovation propres aux services d'informatique en nuage.

L'augmentation de la demande en données, réutilisables de surcroît, fait naître en permanence de nouvelles exigences en termes de qualité, d'accessibilité et de maintenance. Une infrastructure de données doit prioritairement répondre à la nécessité de répertorier et cartographier les données existantes, d'en améliorer l'accès et d'en faciliter l'échange. Le recensement et la maintenance constante d'un ensemble de données de référence, qui doivent être aussi accessibles et réutilisables que possible, doivent être considérées comme un bien public fondamental. En 2013, dans le cadre de son projet d'Infrastructure nationale d'information (NII), le Royaume-Uni a commencé à faire l'inventaire des données (publiques) de référence. En France, la loi pour une République numérique, adoptée en juin 2016, réaffirme l'existence d'un « service public de la donnée » et garantit l'accessibilité, la qualité et la maintenance de banques de données de référence considérées comme ayant un grand potentiel social et économique.

**Le secteur public a l'obligation de fournir aux citoyens des infrastructures de données de référence : il s'agit d'un bien commun, d'un service public fondamental. Il se peut que certaines données de référence doivent être élaborées par le secteur public, ce qui nécessitera des investissements. Les quantités importantes de données issues du secteur privé devront être rendues aussi accessibles que possible, de manière à optimiser la valeur des infrastructures de données pour tous.**

Les obstacles au développement d'infrastructures de données de référence capables de fournir des services essentiels sont souvent liés à une approche fragmentée de la gouvernance des données. Aujourd'hui, les décideurs britanniques et français sont notamment confrontés aux défis suivants :

- une connaissance et une visibilité insuffisantes des données de référence existantes ;
- des systèmes d'information qui, pour la majorité des acteurs, ne sont pas optimisés pour des approches axées sur les données et basées sur la machine ;
- les barrières culturelles et légales qui entravent l'échange de données ;
- le manque de financements et de ressources qui permettraient de combler les écarts et d'investir dans de nouvelles données et dans de nouveaux organismes lorsque cela est nécessaire ; du fait de son caractère stratégique, une infrastructure de données exige des investissements.

Compte tenu de ces éléments, la France et le Royaume-Uni devraient se fixer les objectifs suivants :

*Objectif : faciliter l'utilisation et l'échange de données.*

**Recommandation : développer l'utilisation d'interfaces de programmation (API) dans l'économie de la donnée, en s'appuyant sur l'expérience britannique d'open banking.**

Les API sont le système nerveux du XXI<sup>e</sup> siècle. Nous devons donc encourager le développement d'API ouvertes communes dans des secteurs importants pour les deux pays et collaborer en la matière. De la même façon que les échanges commerciaux dépassent les frontières, l'échange et l'accessibilité des données doivent également être fluides. Le Royaume-Uni a déjà commencé à travailler sur une norme de système bancaire ouvert pour le secteur des services destinés aux particuliers. Ces travaux devraient être étendus.

Il pourrait être intéressant de développer des API ouvertes communes dans d'autres secteurs réglementés, comme l'énergie, les transports, l'eau, le haut débit ou les télécommunications, dans des domaines dans lesquels il serait profitable pour les citoyens de permettre à des tiers d'accéder à leurs données, notamment leurs registres médicaux ou fiscaux, ou encore dans des secteurs dans lesquels collaborent de multiples parties prenantes, comme la filière agroalimentaire.

*Objectif : développer l'infrastructure de base des services interconnectés.*

**Recommandation : investir dans la création de registres de données de référence, en s'appuyant sur l'expérience française de la Base Adresse Nationale, et partager les expériences en la matière.**

Les registres constituent une source de données de référence faisant autorité qui permet d'interconnecter les services et de les améliorer. La France, en collaboration avec la communauté OpenStreetMap, a élaboré la Base Adresse Nationale, un bien commun basé sur la production participative de données, librement accessible et utilisable par tous. Le Royaume-Uni et la France doivent coopérer et partager les leçons tirées de cette expérience, alors que le Royaume-Uni envisage de créer sa propre base d'adresses ouverte.

*Objectif : faciliter l'accès des citoyens aux services dont ils ont besoin, où qu'ils se trouvent.*

**Recommandation : poursuivre les investissements dans des programmes d'identification électronique pour faciliter l'accès aux services publics et privés dans les deux pays, et partager les expériences en la matière.**

La France et le Royaume-Uni œuvrent au renforcement des mécanismes publics transversaux, comme celui de l'identification électronique, qui sont utilisés dans tous les services publics. La protection de la vie privée doit être intégrée au système public dès sa conception, en s'appuyant sur les enseignements tirés de projets tels que France Connect (France) et Verify (Royaume-Uni).



## 3. COMPETENCES

À mesure que l'économie numérique mondiale se développe, toutes les grandes puissances économiques, dont le Royaume-Uni et la France, connaissent une forte augmentation de la demande en compétences en matière de données. Accéder aux talents et aux compétences en la matière est un défi permanent tant pour le secteur public que pour le secteur privé. S'attaquer au problème de la pénurie des compétences dans ce domaine doit devenir une priorité stratégique compte tenu de ses importantes incidences économiques directes. D'importants investissements seront donc nécessaires.

En France, l'objectif inscrit dans la feuille de route du plan Big Data publiée en décembre 2014<sup>1</sup> était de créer ou consolider, à l'horizon 2020, de l'ordre de 137 000 emplois, soit directement dans l'industrie informatique, soit dans les fonctions technologiques au sein des entreprises, soit dans les fonctions métiers (par exemple les fonctions de vente) des entreprises. Au départ, les chefs de projet de ce plan industriel, François Bourdoncle (FB&Cie) et Paul Hermelin (Cap Gemini), ont relevé cinq secteurs clés exposés au risque lié aux transformations induites par la révolution des données : l'assurance, l'énergie, les transports, la santé et l'emploi, tous liés à des projets phares. Un observatoire des usages du Big Data<sup>2</sup> a été créé en 2016 dans le cadre de la Nouvelle France Industrielle. Il analysera quarante secteurs et leur potentiel de transformation par le big data, et recensera également les compétences dont les sociétés du privé ont besoin.

De nouveaux diplômes d'enseignement supérieur dans les domaines, notamment, des données et de la connaissance, de l'analyse par le big data et de l'apprentissage automatique sont créés régulièrement en France et au Royaume-Uni pour répondre aux besoins qui se présenteront dans les années à venir mais ils ne suffiront certainement pas à faire face aux trois grands enjeux qui nous attendent.

Premièrement, nous devons aider nos dirigeants actuels à imaginer des moyens novateurs et pragmatiques d'utiliser les données pour transformer leurs organisations.

Deuxièmement, nous assistons dans nos deux pays à une augmentation de la demande en analystes de données, qui présentent à la fois des compétences techniques et analytiques et des compétences générales, ainsi qu'une connaissance du secteur, de manière à transformer les données en valeur réelle pour leur organisation. Alors que les compétences en science des données sont de plus en plus intégrées à la formation formelle, l'apprentissage de compétences générales telles la capacité de travailler en équipe, l'apprentissage par problèmes ou le travail en collaboration avec des personnes non formées aux données avant d'entrer sur le marché du travail, sont rarement possibles. Le Royaume-Uni et la France investissent fortement dans les compétences et la recherche dans le domaine de la science des données, comme l'illustrent, par exemple, l'Alan Turing Institute et le projet d'Institute of Coding au Royaume-Uni. Si ces projets sont importants, ils n'entraîneront pas pour autant les profondes transformations nécessaires dans chaque secteur de l'économie. Davantage devrait être fait pour mettre en relation les professionnels des données et les secteurs qui ont des difficultés à utiliser les données.

Troisièmement, compte tenu des bouleversements qu'entraînent les technologies numériques pour un nombre croissant de secteurs, il va falloir prévoir de renforcer les connaissances de base et la culture des salariés dans le domaine des données, au Royaume-Uni comme en France, à mesure que nos concitoyens sont encouragés à faire des choix plus éclairés quant à la façon dont leurs données sont collectées et utilisées, et quant aux entités autorisées à y accéder.

1 <http://www.economie.gouv.fr/big-data-feuille-route-en-action>.

2 <http://observatoirebigdata.fr/>.

La France et le Royaume-Uni devraient se fixer les objectifs suivants :

*Objectif : aider les dirigeants à prendre conscience du potentiel des données.*

**Recommandation : développer les programmes d'éducation aux données destinés aux dirigeants du public et du privé.**

Pour mettre en œuvre à grande échelle la transformation numérique, nous devons agir au sommet. Il pourrait être envisagé de développer les formations destinées aux « cadres à haut potentiel » (dirigeants d'entreprise, responsables politiques, hauts fonctionnaires, etc.) pour leur apporter et leur permettre de développer les connaissances, les compétences et les stratégies en matière de données qui leur seront nécessaires pour transformer leurs organisations. Ces dirigeants peuvent appartenir à des entités du secteur public ou du secteur privé.

Ces programmes devraient comprendre à la fois des ateliers et un réseau de pairs, et être axés sur la réponse à des problèmes ou enjeux particuliers. Les structures publiques et privées existantes créent de nouvelles formations et la Taskforce soutiendra les programmes conçus pour les « cadres à haut potentiel ».

*Objectif : surmonter la pénurie chronique de compétences en matière de données.*

**Recommandation : intégrer aux formations d'analystes existantes l'enseignement de nouvelles compétences en matière de données et de résolution de problèmes, et créer des programmes d'apprentissage et des stages de reconversion communs afin de fournir au Royaume-Uni et à la France un nombre suffisant de professionnels des données.**

Il serait profitable d'intégrer aux programmes d'enseignement des disciplines utilisant les données (comme la gestion, le marketing, les sciences de l'ingénierie, l'économie et la statistique) une formation de base sur les données, la résolution de problèmes, la programmation collaborative et l'éthique et la visualisation des données. En parallèle, il serait souhaitable de créer de nouvelles voies d'entrée pour les futurs professionnels des données. En France comme au Royaume-Uni, il existe des initiatives dont l'objet est d'enrichir les compétences des développeurs et des data scientists, mais elles sont généralement axées sur la sécurité et la recherche de pointe. Certains secteurs ont des difficultés à attirer et/ou exploiter les compétences en matière de données. Il est nécessaire d'investir dans des programmes d'apprentissage et des stages de reconversion destinés aux futurs professionnels des données. Le nombre de formations, stages et autres formes d'éducation aux données doit considérablement augmenter. C'est pourquoi il est important de « former les formateurs ». Ces initiatives pourraient être combinées à un programme de mise à disposition, qui permettrait à des professionnels des données de travailler dans des organisations qui souhaitent, sur un projet particulier, bénéficier d'une assistance en matière de données.

En France, le projet EdFab3, lancé par Cap Digital en 2016, est consacré à l'innovation dans les secteurs de l'enseignement, de la formation et de la transformation des métiers. Au Royaume-Uni, plus de cinquante formateurs ont été formés en 2015 grâce au programme « Train the Trainer » de l'Open Data Institute.

*Objectif : promouvoir l'acquisition de connaissances de base en matière de données.*

**Recommandation : mettre en œuvre des projets destinés à dispenser des connaissances de base sur les données aux élèves du primaire et du secondaire.**

De la même manière que les élèves reçoivent une éducation civique et sont sensibilisés aux risques liés à l'internet, ils doivent aussi être sensibilisés aux données si nous voulons faire des choix plus éclairés sur l'utilisation de nos données par les organisations. Plusieurs expériences en ce sens se sont révélées efficaces et pourraient facilement être soutenues et étendues. Ainsi, france-ioi.org propose gratuitement des activités et des concours sous une approche ludique, indépendante et participative. Une présentation générale du potentiel que représentent les données pourrait être intégrée au programme des cours d'éducation morale et civique.



#### 4. CONFIANCE ET ACCROISSEMENT DU “POUVOIR D’AGIR”

L’exploitation des opportunités offertes par l’économie de la donnée suppose la confiance et le renforcement du pouvoir d’agir.

Pour gagner et conserver la confiance de nos citoyens et consommateurs, nous avons besoin d’un réel engagement des gouvernements européens en faveur de la protection des données des citoyens et des entreprises. La sécurité des réseaux et la souveraineté numérique pour les citoyens, les organisations et les gouvernements sont des éléments essentiels de notre souveraineté économique, mais aussi de la confiance globale des citoyens dans la révolution numérique, voire dans les gouvernements.

En 2014, une étude réalisée par Ipsos MORI à la demande de la Royal Statistical Society a montré que c’est aux médias, et aux entreprises des secteurs de l’internet, des télécommunication et des assurances que les citoyens font le moins confiance pour ce qui est de l’utilisation qui est faite de leurs données. Ainsi, seulement 4 à 7 % des citoyens interrogés ont déclaré faire tout à fait confiance à ces organisations pour utiliser les données de manière appropriée, à comparer au pourcentage de confiance de 36 % et de 41 % dont sont crédités, respectivement, le système de soins britannique (NHS) et les médecins traitants<sup>1</sup>.

En France, les autorités ont travaillé avec des parties prenantes publiques et privées sur un label « Secure Cloud »<sup>2</sup> afin de renforcer la confiance dans les services d’informatique en nuage tout en garantissant un certain niveau de sécurité des services d’infrastructure. Il est important de bâtir une chaîne de la confiance qui aille de l’infrastructure sous-jacente aux applications et à leur utilisation des données : aujourd’hui encore, 74 % des Français disent ne pas avoir confiance en la manière dont les applications pour mobile utilisent leurs données<sup>3</sup>.

Pour que la révolution des données soit un succès, les organisations doivent faire montre à l’égard de leurs usagers d’une plus grande transparence sur la façon dont elles traitent et protègent leurs données, et notamment préciser avec quelles organisations elles les partagent et quelles données elles peuvent conserver par-devers elles. Ces principes sont inscrits dans le règlement général sur la protection des données (RGPD).

De nombreux services de données proposés aujourd’hui sont de véritables « boîtes noires », qui fournissent des résultats ou des décisions sur la base d’algorithmes et de sources de données dont les mécanismes et le contenu sont opaques pour l’utilisateur. Les utilisateurs de ces systèmes ne savent pas comment ces derniers sont conçus, comment leurs critères sont établis ni tout ce qu’ils permettent d’effectuer. Lorsque des explications sont données, c’est souvent dans un langage technique que l’utilisateur lambda ne peut comprendre. Les citoyens, les entreprises, les acteurs publics et les gouvernements sont de plus en plus amenés à se saisir de cette situation.

Il peut arriver que le partage de nos données à caractère personnel soit nécessaire pour nous permettre de bénéficier des services dont nous avons besoin (en matière de santé par exemple), ou encore qu’il soit impossible d’obtenir un consentement préalable à l’utilisation des données (par exemple des données implicites issues des actions de l’utilisateur lors d’une session de navigation sur le web). Néanmoins, la gouvernance des données centrées sur le citoyen est de plus en plus placée au cœur de l’économie de la donnée. Le citoyen doit devenir un acteur capable de prendre les décisions qui concernent ses propres données. **Qui dit renforcement du pouvoir d’agir des citoyens dit possibilité pour eux d’avoir la maîtrise de leurs propres données et de décider quelle utilisation en faire, par exemple les partager avec d’autres citoyens, avec des associations, des organismes de recherche ou des entreprises du privé.**

L’État français (Conseil Général de l’Économie) étudie avec différents acteurs des secteurs public et privé l’éventualité d’une étude sur la réglementation des algorithmes traitant du contenu. Cette étude permettrait de décider de l’opportunité de mettre en place, dans nos économies de la donnée, de nouvelles politiques favorisant la confiance et l’accroissement du pouvoir d’agir des citoyens.

1 <https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-les-sons-for-policy-makers.aspx>.  
2 <http://www.ssi.gouv.fr/actualite/appele-public-a-commentaires-sur-le-referentiel-dexigences-applicables-aux-prestataires-de-services-securises-dinformatique-en-nuage/>.  
3 Baromètre de l’innovation (Odoxa, Microsoft, L’Usine Digitale), 2016.

Compte tenu de l'importance de la confiance et de l'accroissement du pouvoir d'agir, la France et le Royaume-Uni devraient se fixer les objectifs suivants :

*Objectif : renforcer le pouvoir d'action des citoyens sur leurs données.*

**Recommandation : améliorer et étendre les bonnes pratiques issues des projets MesInfos et MiData pour favoriser le renforcement du pouvoir d'action des citoyens sur leurs données.**

La propriété, la portabilité et la maîtrise des données ainsi que la responsabilité en la matière sont les principes fondamentaux inscrits dans le RGPD et au cœur des projets MesInfos et MiData. La manière dont les organisations se conformeront dans la pratique au futur RGPD demeure incertaine. Une initiative visant à accroître le pouvoir d'action des citoyens sur leurs données pourrait être lancée en s'appuyant sur les expériences MesInfos (France) et MiData (Royaume-Uni). La recherche dans les technologies et la réglementation susceptibles d'aider les entreprises à donner aux citoyens les moyens d'avoir une plus grande maîtrise sur leurs données à caractère personnel et de prendre des décisions en la matière doit être encouragée. De nouvelles expériences devraient également être menées pour faire de ces nouvelles règles une chance à saisir pour les entreprises.

Les parties prenantes françaises et britanniques (MesInfos, MiData, Cozycloud, Imperial College, Digi.me et d'autres) auraient beaucoup à gagner à travailler ensemble en faveur du développement des technologies et des services liés aux données à caractère personnel : plates-formes de gestion des données à caractère personnel, détection des fuites de données grâce à des systèmes spécialisés, etc.

*Objectif : anticiper les enjeux futurs en matière de politique des données.*

**Recommandation : mener des recherches sur la transparence des algorithmes et la responsabilité.**

Les décideurs et les organisations sont confrontés à de nouvelles questions portant sur la prise de décisions basée sur les données, par exemple au moyen d'algorithmes. Dans certains secteurs, cette pratique est déjà intégrée aux services que nous utilisons au quotidien. Comment s'assurer que notre processus décisionnel fondé sur les données est équitable et n'entraîne pas de discriminations ? La France et le Royaume-Uni devraient lancer des travaux de recherche sur les moyens de rendre les organisations plus transparentes par rapport aux algorithmes, aux sources de données et au processus d'analyse des données qu'elles utilisent pour prendre des décisions qui touchent les consommateurs et les citoyens.

Plusieurs initiatives internationales ont été lancées autour de la transparence. L'INRIA s'associera au Data Transparency Lab<sup>4</sup> cet été pour définir et mettre en œuvre de nouvelles activités. L'Open Data Institute est l'un des partenaires à l'origine de cette initiative. Les travaux du Data Transparency Lab pourraient être exploités.

---

4 <http://www.datatransparencylab.org/>.

## REMERCIEMENTS

Ce rapport n'aurait pu voir le jour sans le soutien et les conseils de nombreuses personnes et organisations. La Taskforce aimerait remercier tous ceux qui ont contribué à cette entreprise.

Nous remercions en particulier Cap Digital, l'Open Data Institute et NUMA, qui ont gentiment accepté d'accueillir des réunions de la Taskforce entre novembre 2015 et avril 2016 et de mobiliser leurs équipes pour consulter leurs écosystèmes.

Nous sommes infiniment reconnaissants à Salesforce, Access Partnership, techUK, IRT SystemX, et OVH de nous avoir accordé leur temps et leur assistance pour organiser des consultations auprès de nombreuses entreprises du secteur des technologies numériques afin de nous aider à élaborer les recommandations figurant dans ce rapport.

Enfin, nous exprimons notre sincère gratitude à Cédric Mora (DGE), Fabien Terrailot (DGE), Achille Lerpinière (DGE), Paula Forteza (Etalab), Vicki Brown (DCMS), Paul Driver (DCMS) et Ellen Broad (ODI) pour l'implication et l'énergie dont ils ont fait preuve dans la gestion du secrétariat de la Taskforce.

