

LE DATA ALTRUISME : UNE INITIATIVE EUROPÉENNE

LES DONNÉES AU SERVICE
DE L'INTÉRÊT GÉNÉRAL



HUMAN TECHNOLOGY
FOUNDATION

l'exploratoire
sopra  steria
next

SOMMAIRE

P. 6 — INTRODUCTION

P. 8 — LE DATA ALTRUISME,
ACTEURS ET MÉCANISMES

P. 12 LES ACTEURS DE L'ALTRUISME

P. 16 LE DATA ALTRUISME : UN PAS DE CÔTÉ
EUROPÉEN DANS LE CHAMP NUMÉRIQUE

P. 16 LES DÉFIS RENCONTRÉS
PAR LE DATA ALTRUISME

P. 24 — UN PAYSAGE DES NOUVEAUX
MODES DE PARTAGE DE
DONNÉES RICHE ET MOUVANT

P. 26 DYNAMIQUES DU SECTEUR DE LA SANTÉ

P. 29 DYNAMIQUES DANS LE SECTEUR DE LA MOBILITÉ

P. 32 DYNAMIQUES DANS LE SECTEUR DE L'ENVIRONNEMENT

P. 35 FIDUCIE DES DONNÉES :
UNE PERSPECTIVE CANADIENNE

P. 37 FIGURES INSPIRANTES ET RISQUES
DE DÉRIVES DU DATA ALTRUISME

P. 44 — VADE-MECUM POUR UN
DATA ALTRUISME EFFICACE
DANS LA CONFIANCE

P. 45 **ÉTAPE N°1**
CRÉER UNE ORGANISATION DATA ALTRUISTE

P. 48 **ÉTAPE N°2**
DÉFINIR LA GOUVERNANCE DE
L'ORGANISATION DATA ALTRUISTE

P. 51 **ÉTAPE N°3**
COLLECTER DES DONNÉES

P. 54 ENCOURAGEMENT AU PARTAGE
PAR LES PERSONNES PHYSIQUES

P. 56 ENCOURAGEMENT AU PARTAGE
PAR LES PERSONNES MORALES

P. 59 **ÉTAPE N°4**
ASSURER LES MODALITÉS
DE PARTAGE DES DONNÉES

P. 61 **ÉTAPE N°5**
ASSURER LE FINANCEMENT DE
L'ORGANISATION DATA ALTRUISTE

P. 66 — CONCLUSION

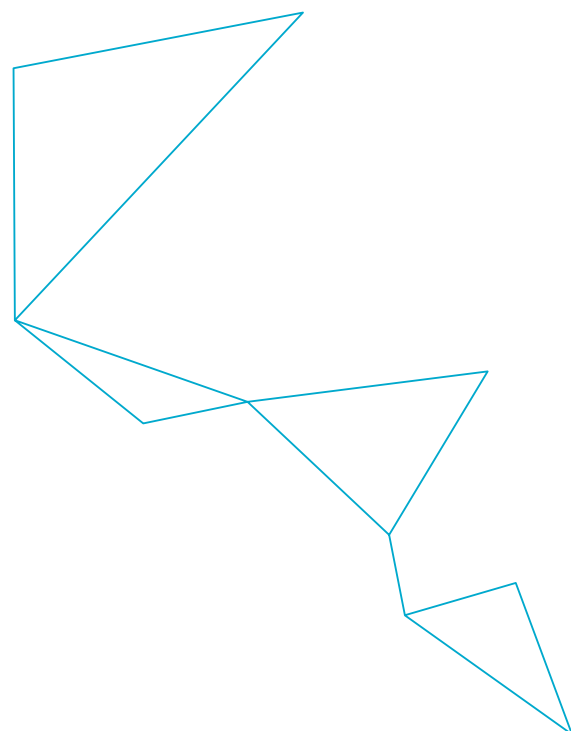
Au nom de la Human Technology Foundation, je tiens à remercier tous ceux qui se sont impliqués dans cette étude et, en premier lieu, L'Exploratoire Sopra Steria Next, partenaire avec lequel nous avons établi une collaboration fructueuse.

Mes remerciements vont également aux rapporteurs, Jean Leviste, Grimaud Valat et Victor de Salins, ainsi qu'à tous les membres du groupe de travail, dont la compétence s'est avérée déterminante dans la rédaction de ce rapport. Parmi eux, je tiens à saluer l'expertise apportée par les équipes de la CNIL et du MEDEF.

Merci également à l'ensemble des acteurs auditionnés, qui nous ont consacré de leur temps et de leur énergie.

Éric SALOBIR

Président du comité exécutif
Human Technology Foundation



MEMBRES DU GROUPE DE TRAVAIL

Jawaher ALLALA
CEO/Administratrice
SystnapsNumeum

Éric BOTHOREL
Député des Côtes-d'Armor

Bertrand BRAUNSCHWEIG
Coordinateur scientifique
Confiance.ai

Régis CHATELLIER
Responsable d'études prospectives
CNIL (LINC)

Emmanuel CRAIPEAU
Partner
Sopra Steria Next

Victor de SALINS
Rapporteur
Human Technology Foundation

Olivier DION
Cofondateur
aNewGovernance

Clément EMINE
Chargé de mission
Economie numérique
MEDEF

Clémentine FURIGO-CHRISTIN
Juriste Sherpa du Comité
Régulations du numérique
MEDEF

Yacine JERNITE
Chercheur
BigScience/Hugging Face

Jean LEVISTE
Consultant / Rapporteur
Sopra Steria Next

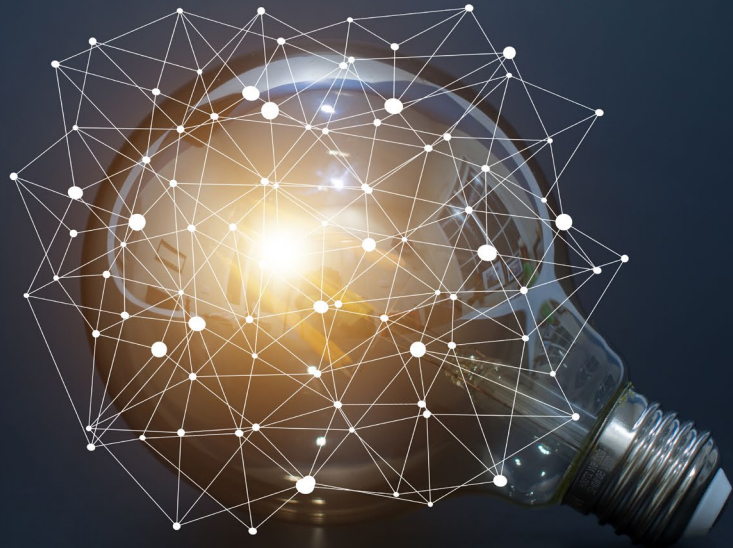
Charles MORGAN
Avocat associé, coleader national du groupe Cyber/Données
McCarthy Tétrault

Éric PETERS
Senior Adviser to the Director General
DG Connect

Éric SALOBIR
Président
Human Technology Foundation

Floran VADILLO
Directeur Conseil éthique et souveraineté
Sopra Steria Next

Grimaud VALAT
Avocat associé / Rapporteur
DTMV



ACTEURS AUDITIONNÉS

Yoshua BENGIO

Fondateur et directeur scientifique

Mila - Institut québécois d'intelligence artificielle

Julien BOVET

Head of Data

Lepermislibre

Javier CREUS

Fondateur/Membre

Ideas for change/
SalusCoop

Jacopo DOMENICUCCI

Professeur agrégé de philosophie

Milad DOUEIHI

Professeur des universités Titulaire de la chaire d'humanisme numérique

Université de Paris-Sorbonne

Romain DURAND

Directeur général

Lepermislibre

Marco FIORINI

Délégué général

Alliance pour la recherche et l'innovation des industries de santé

Marion GLATRON

Directrice de la donnée, de la transition écologique et des ressources

Rennes Ville et Métropole

Gloria GONZALEZ FUSTER

Professeure

Vrije Universiteit Brussel

Jean-François HULLO

Head of Data, Analytics and Reporting

EDF

Florian LAINEZ

CEO

Jungle Bus

Alexandre LE BOUTHILLIER

Co-fondateur et Chef des affaires corporatives

Imagia

Erwan MEDY

Data and digital strategy

Groupe VVV

Carlos MORENO

Professeur, co-fondateur de la Chaire eTI (Entrepreneuriat Territoire Innovation)

Université Paris I Panthéon-Sorbonne

Valérie PISANO

Président & CEO

Mila - Institut québécois d'intelligence artificielle

Simon SAINT-GEORGES

Chef de projet européen Rudi Rennes Ville et Métropole



INTRODUCTION



In Europe, basically we do have all what we need, but we have to unleash this potential.¹



¹ Press remarks by President von der Leyen on the Commission's new strategy: Shaping Europe's Digital Future Brussels, 19 February 2020.

En 2025, le volume de données générées dans le monde devrait dépasser 180 zettaoctets contre 64 zettaoctets² en 2020 et « seulement » 2 zettaoctets en 2010³. Le marché mondial des services de traitement de données massives est en constante augmentation. D'après la Commission européenne il représenterait plus de 200 Mds\$ en 2020 alors que le volume mondial des données devrait augmenter de 530 % d'ici 2025⁴.

Alors que le big data recèle de très fortes perspectives de développement scientifique, culturel ou économique, les scandales liés à l'exploitation de données ont profondément terni son image, dans l'esprit du grand public comme des entreprises.

De fait, la confiance dans le domaine numérique est loin d'être installée. Les grands acteurs du numérique ont déployé des investissements massifs pour capter ces données et alimenter de puissants outils d'analyse. Dans le domaine si sensible des données personnelles, Shoshana Zuboff compare même la vie intérieure des gens à un « *continent précolonial*⁵ », en lutte permanente avec les dispositifs numériques qui s'évertuent à en exploiter les ressources.

C'est ainsi que, paradoxalement, les grandes plateformes numériques parviennent à collecter les données personnelles de leurs utilisateurs à des fins commerciales, tandis que, selon les chiffres de la Commission européenne, « 80 % des données industrielles sont collectées mais ne sont jamais utilisées⁶ ».

Amélioration des services publics, transparence de l'information, stratégies climatique ou sanitaire, développement économique, les thématiques d'intérêt général sont nombreuses, et toutes pourraient bénéficier d'une optimisation des traitements de données. S'est ainsi développée une politique de fluidification et d'ouverture des données, notamment via la Directive Open Data, qui reste pourtant essentiellement attachée à la sphère publique.

L'ambition de l'Union européenne est donc de libérer le potentiel de ces données « immobilisées », et de les mettre au service de l'intérêt général.

La donnée n'est pourtant pas un objet juridique comme un autre. Elle revêt une nature plurielle : les données personnelles, techniques, territoriales ou d'intérêt général ont toutes des caractéristiques et des usages différents.

En termes économiques, la donnée est dite « non rivale », en ce sens qu'elle peut être utilisée plusieurs fois par des acteurs différents sans pour autant être consommée comme une matière première. Elle est aussi dite « excluible » car il est possible d'en faire payer l'accès. Enfin, elle nécessite d'être traitée pour acquérir de la valeur. Cette nature devrait faire de la donnée un objet circulant activement et pour des finalités multiples.

Pourtant, l'essentiel des données disponibles sont détenues par des personnes physiques ou morales et ne sont pas partagées. Elles sont donc sous exploitées, alors même qu'elles pourraient être utilisées pour l'intérêt général et sans nuire aux avantages concurrentiels ou à la vie privée de leurs détenteurs initiaux.

Ainsi, outre les données industrielles collectées qui ne sont pas exploitées, les autres circulent majoritairement à travers deux voies, qui ont montré leurs limites :

- **La voie commerciale** : les données sont achetées ou collectées puis traitées dans un but exclusivement marchand. Très souvent, elles « *alimentent les patterns produits par les plateformes qui les revendent aux marques*⁷ ». L'exploitation commerciale des données obéit aux règles du marché concurrentiel, très opposées au principe d'ouverture et de libre circulation ;

- **La voie de l'open data légal** : le partage des données est imposé notamment aux services publics. Ce mécanisme, basé sur la contrainte, aboutit à des résultats très mitigés notamment au regard de la qualité et de l'utilisabilité des données mises à disposition.

Aucun de ces deux modes de partage n'est cependant à exclure, puisqu'ils participent à la circulation des données et à la croissance économique. De la même manière, la création de mécanismes d'obligation de partage de données, dans des circonstances spécifiques, paraît nécessaire pour répondre à certains enjeux, et ne doit pas être abandonnée⁸.

Le partage par la contrainte est toutefois complexe à mettre en œuvre et se heurte nécessairement à des barrières légales, au premier rang desquelles la territorialité des normes juridiques qui rend difficile, voire impossible, la réquisition de données auprès de détenteurs internationaux. La Commission européenne a ainsi proposé un outil réglementaire relatif à la gouvernance des données dit Data Governance Act (DGA) dont l'objectif est de : « *Favoriser la disponibilité de données en vue de leur utilisation, en augmentant la confiance dans les intermédiaires de données et en renforçant les mécanismes de partage de données dans l'ensemble de l'UE*⁹ ».

Est ainsi introduite dans le Chapitre IV la notion de data altruisme, une potentielle troisième voie de partage de données, basée sur l'initiative et l'altruisme à des fins d'intérêt général. Elle repose sur la création d'organisations altruistes en matière de données, des tiers de confiance, indépendants, à but non lucratif, chargés d'assurer dans la plus totale transparence le recueil de données transmises par des tiers à des fins d'intérêt général.

Le data altruisme est, en l'état, un concept encore plastique et désincarné. L'ambition du présent rapport est de permettre de participer à sa définition, d'en mieux saisir les contours, mais aussi de mettre en avant les opportunités et les limites inhérentes à ce nouvel outil (PARTIE 1). Il propose des possibilités réglementaires et d'actions pratiques ce que la notion de data altruisme peut apporter à un partage et une exploitation éthique des données. Le rapport comprend ainsi une analyse des modes de partage de données actuels et inspirants en vue de l'élaboration de modèles data altruistes (PARTIE 2). Des recommandations sont ainsi détaillées pour permettre la mise en œuvre et le développement de ces modèles et d'en assurer véritablement le caractère altruiste (PARTIE 3).

² 1 zettaoctet fait 10 exposant 21 octets (1 000 000 000 000 000 000 octets).

³ [Le Big Bang du Big Data, Statista, 2021.](#)

⁴ Stratégie européenne pour les données, communication de la Commission européenne du 19 février 2020, projections pour 2025.

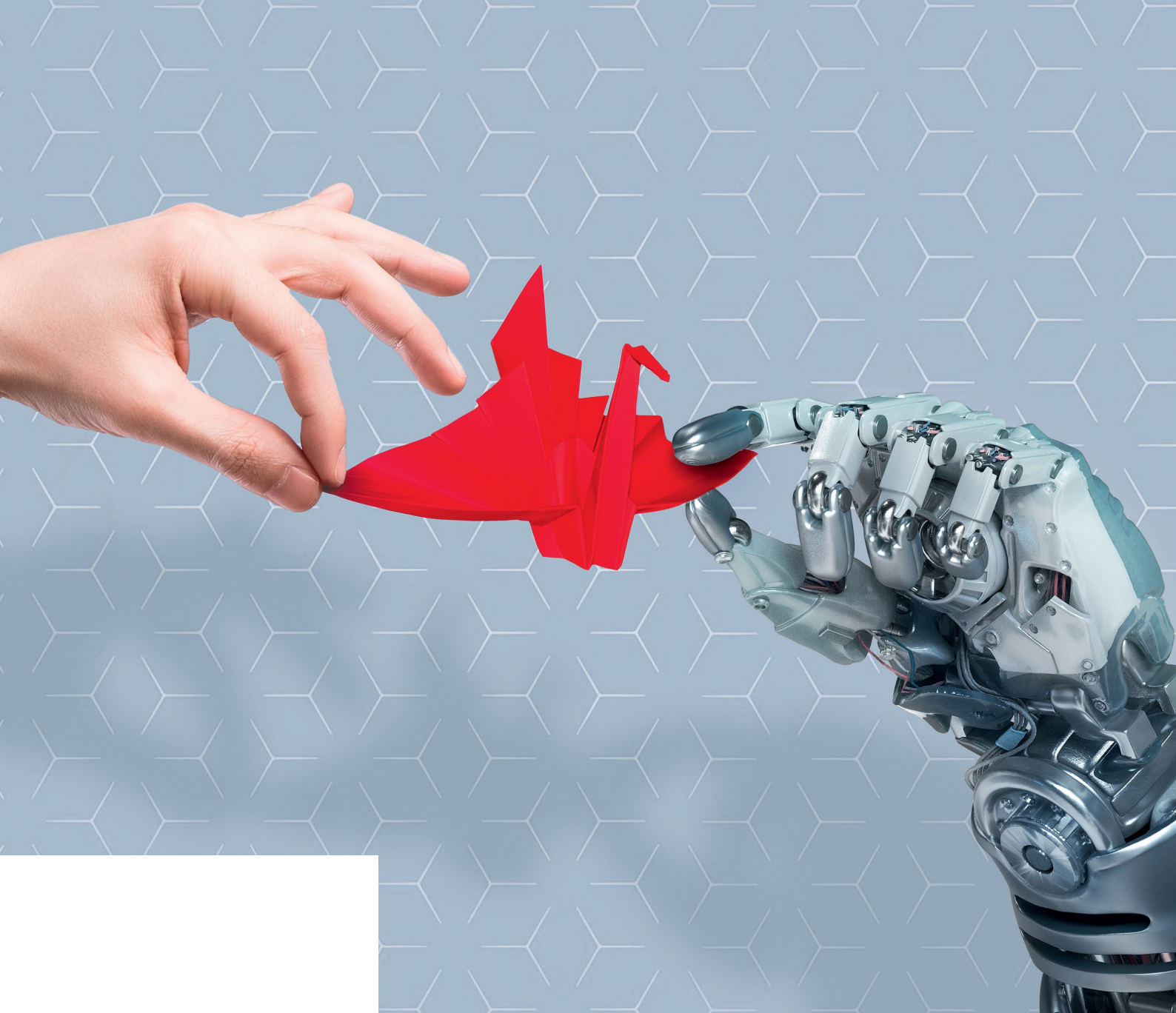
⁵ ZUBOFF Shoshana, *Age Of Surveillance Capitalism: The Fight For A Human Future At The New Frontier Of Power*, 2019 cité in VILJOEN Salomé, *A Relational Theory Of Data Governance Forthcoming*, Yale Law Journal, Vol. 131, 2021.

⁶ Discours sur l'état de l'Union de la présidente von der Leyen en session plénière du Parlement européen 16 septembre 2020.

⁷ [BOULLIER Dominique, Puissance des plateformes numériques, territoires et souverainetés – Sciences Po Chaire digital, gouvernance et souveraineté, 2021.](#)

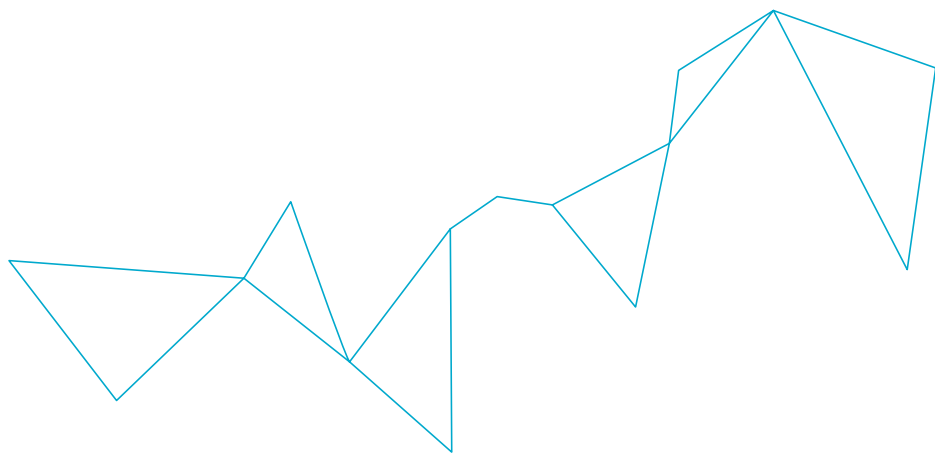
⁸ Les mécanismes de réquisition restent flous, un point soulevé par la mission Bothorel sur *La Politique publique de la donnée, des algorithmes et des codes sources* qui recommande de « *clarifier le régime juridique de la réquisition pour permettre à la puissance publique d'accéder à des données du secteur privé en cas de motif impérieux d'intérêt général et d'urgence* » (recommandation n°34).

⁹ Proposition de Règlement du Parlement européen et du Conseil sur la gouvernance européenne des données, Exposé des motifs.



1.

LE DATA ALTRUISME, ACTEURS ET MÉCANISMES



Le DGA définit l'altruisme des données comme : « *le partage volontaire de données basé sur le consentement donné par les personnes concernées au traitement de données à caractère personnel les concernant, ou les autorisations accordées par d'autres titulaires de données pour l'utilisation de leurs données à caractère non personnel sans demander de contrepartie, à des fins d'intérêt général, telles que la santé, la lutte contre le changement climatique, l'amélioration de la mobilité, la facilitation de l'établissement de statistiques officielles, l'amélioration de l'offre de services publics, l'élaboration des politiques publiques* ».

Toutefois, comme le rappelle le Comité européen de la protection des données¹⁰, le DGA ne crée pas mais formalise la possibilité de transmettre ses données à un tiers, dans un cadre propice à la confiance.

À cette notion, le présent rapport apporte les précisions suivantes :

LES MODALITÉS DU DATA ALTRUISME

Le data altruisme rassemble les modes de partage librement consentis des données, générées pour l'occasion ou préexistantes, en vue d'un traitement ultérieur au service de l'intérêt général et des individus. Un modèle data altruiste peut inclure :

- des mécanismes de financement afin d'assurer la mise en place et la maintenance des infrastructures de partage ;
- des incitations à participer aux modèles, qu'elles soient directes ou indirectes.

Cette approche précise que des modèles data altruistes peuvent incorporer des modalités de financements ainsi que des incitations à agir pour les différentes parties prenantes. Les travaux du groupe de travail ont abouti à considérer que si le DGA prévoit d'ores et déjà que les tiers de confiance peuvent instaurer des systèmes d'accès aux données qu'ils détiennent en contrepartie de redevances, il convient de penser à l'instauration de systèmes d'incitations directes ou indirectes au bénéfice des contributeurs de données. L'ambition affichée est double :

- contribuer à répondre aux enjeux de ce siècle en sollicitant des gisements de données peu ou pas exploités ;
- refonder une mécanique d'échange dans le monde numérique jusqu'alors dominée par l'extractivisme, la prédation ou la contrainte.

¹⁰ European Data Protection Board (EDPB) and European Data Protection Supervisor (EDPS), *Joint opinion on the on the Proposal for a Regulation of the European Parliament and of the Council on European Data governance (Data Governance Act)*, 10 March 2021.

OPEN DATA ET ALTRUISME DES DONNÉES

Quoique proches en apparence l'*open data* et l'altruisme des données diffèrent l'un de l'autre sur au moins deux points :

- **Dans leur nature** : l'*open data* consiste en une mise à disposition de données, afin de favoriser l'émergence des finalités, tandis que le data altruisme vise à permettre la transmission de données pour les besoins de finalités prédéfinies.
- **Dans leur contexte historique de mise en place** : l'*open data* a été marqué assez tôt par une obligation d'ouverture imposée aux pouvoirs publics. Cette obligation, bien que non coercitive, s'oppose au principe d'altruisme. L'*open data* ne revêt toutefois pas systématiquement ce caractère obligatoire.

Anne Lehmans dans son analyse des politiques d'*open data* y voit un enjeu de culture organisationnelle « *Il ne s'agit pas seulement d'une culture informatique, mais d'une convergence entre la capacité à produire et traiter des données informatiquement, [...] et la compréhension des enjeux sociaux, politiques, économiques de la production de données ouvertes* »¹¹. Pourtant cet esprit d'ouverture de la donnée ne se décreète pas. Il nécessite des ajustements organisationnels, une revue des valeurs et des principes d'organisation et génère donc des tensions. Lorsque le Secrétariat général pour la modernisation de l'action publique remet, en 2015, un rapport sur *La gouvernance de la donnée*¹², il met en avant le fait que la culture administrative est fort éloignée de la logique de coopération et de partage de données. Cette culture de la donnée et de son partage s'articule autour de trois valeurs¹³ :

- **La pluralité** : cette stratégie d'ouverture de la donnée doit naviguer au sein d'un patchwork de compétences, de métiers, tous ayant leurs routines, symboles, références propres. Travailler différemment impose une revue de ces routines, leur ajustement voire leur suppression. C'est pour cela que « *la culture de la donnée relève donc d'un processus de négociation* »¹⁴.
- **La transversalité**, pour conjuguer ces différentes références professionnelles. Il faut convoquer une notion clé, *data literacy* employée dans le monde anglo-saxon. Elle renvoie au « *désir et la capacité de s'engager de façon constructive dans la société à travers et avec les données* »¹⁵ est alors une notion centrale.
- **La technicité** qui sous-tend et rend possible les pratiques de partage à travers une infrastructure qui est maintenue, étendue et améliorée.

L'absence de l'un de ces trois items, notamment lors d'événements *ad hoc* (hackathon par exemple) génère souvent l'impression préjudiciable d'*openwashing*¹⁶. Ainsi, en *open data* comme en data altruisme, si une culture du partage n'est pas en place, ou si l'initiative est issue d'un acteur extérieur, « *la mobilisation des acteurs est très faible et les procédures mises en place peu efficaces pour atteindre cette nécessaire convergence* »¹⁷.

Le trilogue¹⁸ entre la Commission, le Parlement et le Conseil a été riche et mouvant. Un point débattu en particulier a symbolisé les questionnements autour de la nature profonde du data altruisme : le recours au terme de « *bien commun* » ou « *d'intérêt général* » pour qualifier l'objet recherché par les projets se réclamant du data altruisme. Bien que subtile, la différence sémantique est réelle. Si le second a fini par s'imposer¹⁹, définir clairement ces deux notions permet de mettre l'accent sur les particularités du data altruisme.

¹¹ LEHMANS Anne, « Données ouvertes et redéfinition de la culture de l'information dans les organisations », Communication et organisation, n°51, 2017.

¹² Secrétariat général pour la modernisation de l'action publique, Rapport de l'administrateur général des données. Les données au service de la transformation de l'action publique, 2015.

¹³ LEHMANS Anne, 2017.

¹⁴ LEHMANS Anne, 2017.

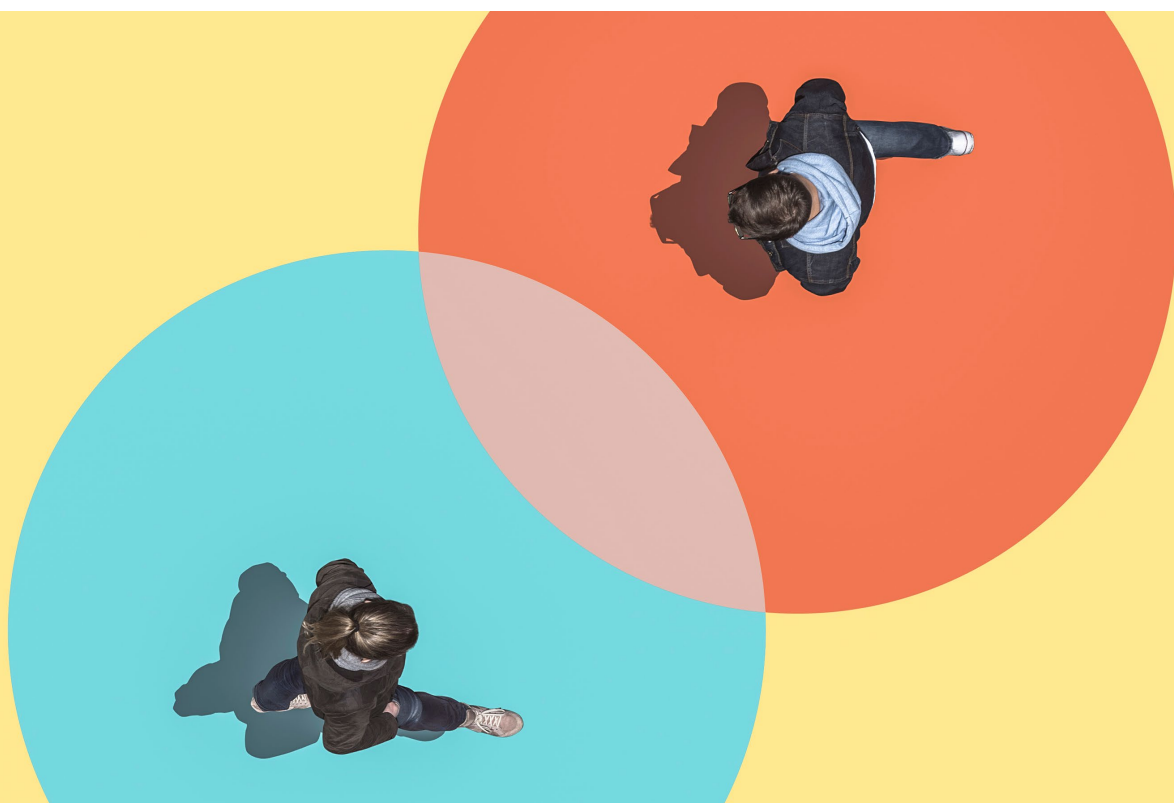
¹⁵ Traduction in LEHMANS Anne, 2017.

¹⁶ DILARA Vanessa Trupia, "Open transport data et développement d'applications de mobilité : Un travail d'équipement à la frontière de mondes variés", Réseaux, pp. 95-129 n°228, 2021.

¹⁷ LEHMANS Anne, 2017.

¹⁸ Un trilogue est un type de réunion utilisé dans le processus législatif de l'Union européenne. Il est composé de la Commission européenne, du Conseil de l'Union européenne et du Parlement européen. La Commission assume la fonction de médiation.

¹⁹ Une référence au « *common good* » est faite dans L'*Explanatory Memorandum* du DGA.



INTÉRÊT GÉNÉRAL OU BIEN COMMUN ?

Il n'existe pas de définition légale officielle et unifiée de l'intérêt général, mais il est possible d'en identifier certaines caractéristiques dans plusieurs textes.

Ainsi, en droit fiscal français²⁰, l'intérêt général consiste en une activité non lucrative à caractère philanthropique, scientifique, social, éducatif, sportif, culturel ou humanitaire, qui n'est pas mise en œuvre au profit d'un cercle restreint de personnes, et dont la gestion est désintéressée. De par sa finalité transcendant les intérêts individuels, il est le corollaire d'une mission définie par la loi dans le cadre d'un contrat social. Il suggère un rééquilibrage du rapport de force entre les individus et l'Etat, renforçant le contrôle des citoyens dans les mécanismes de vie en communauté et de cohésion sociale²¹. Chacun est ainsi incité à participer à la société en mettant à disposition son intelligence, son temps, ses compétences, ou ses données dans le cas du data altruisme. La poursuite de l'intérêt général légitime la construction d'un projet commun par des méthodes participatives et façonne une destinée collective incarnée par un intérêt supérieur.

Au niveau européen, un cadre de qualité pour les services d'intérêt général (SIG) a été adopté en 2011 pour harmoniser la définition des activités concernées. Fournies par l'Etat ou le secteur privé, ces dernières font l'objet « d'obligations de service public spécifique » indissociables²², à l'instar des transports publics, des services postaux ou des soins de santé.

L'idée de bien commun a été théorisée par le philosophe dominicain Thomas d'Aquin au XII^e siècle, inspiré par la littérature antique. Elle offre une vision morale de la société. Chaque membre du corps social est invité à rechercher le bien de la communauté, en même temps que le sien, en exerçant sa liberté dans une démarche spontanée. L'intérêt général évacue donc la dimension morale individuelle portée par le bien commun. Son caractère participatif et formalisé et sa promotion d'un projet politique commun en rend son utilisation plus pertinente dans le cadre du data altruisme.

²⁰ Code général des impôts, article 200 définissant les activités d'intérêt générale éligibles à des réductions d'impôts au titre des dons faits par les particuliers.

²¹ *Traité de droit constitutionnel*, cité n°5, p. 61.

²² ESPLUGAS-LABATUT Pierre, *Fasc. 149 : notion de service public – Droit interne et droit de l'Union européenne*, LexisNexis, p. 17, 2019.

Le data altruisme ne discriminant pas la nature, la destination ou la typologie des données transmises implique d'avoir une approche multifactorielle lors de la définition de modèle de partage altruiste. Si le présent rapport s'efforcera de saisir un maximum de cas de figure, il ne pourra prétendre traiter de tous les types de données pouvant entrer dans un modèle data altruiste étant donné le foisonnement des cas de figures.

LE SYSTÈME ORANGE JUDD : LE DATA ALTRUISME AGRICOLE DU XIX^e ?

Dans les années 1860, le système agricole américain connaît une asymétrie informationnelle particulièrement forte : les producteurs ne connaissent que les données de leur production, et seuls les acheteurs itinérants ont une vision globale du marché. Libre à eux de fixer des prix qui leurs sont avantageux. Orange Judd, journaliste et éditeur à la tête de *l'American Agriculturist*²³, décide de lancer des sondages demandant à ses abonnés de remplir des questionnaires portant sur l'état de leurs récoltes. Le fermier envoie les données de son activité individuelle en échange de données agrégées à même de réduire l'asymétrie entre producteurs et acheteurs. Le gouvernement fédéral demande à avoir accès aux données afin de nourrir sa politique agricole en pleine construction. Les producteurs ont toutefois généré des effets d'anticipation : chaque fermier avait intérêt à sous-évaluer sa production pour que les prix soient au plus haut²⁴. Il revient alors au système de corriger ces biais. La gouvernance adoptée est riche : les statisticiens les plus proches du terrain corrigent les biais favorables aux paysans, les autres correctifs sont effectués par l'administration fédérale à Washington D.C.

Plus récemment, cette méthode d'action a pris un nouveau tournant grâce à la révolution de la téléphonie en Afrique subsaharienne entre les années 2000 et 2010. Les fermiers africains des zones rurales ont pu contourner les intermédiaires et combler l'asymétrie d'information en échangeant directement les prix de marché de leurs produits via leurs téléphones mobiles²⁵.

LES ACTEURS DE L'ALTRUISME

Pour mettre en œuvre le data altruisme, il importe de comprendre quels acteurs sont en jeu et en quoi cela diffère du schéma classique de circulation des données.

De façon générale, le mécanisme du partage de données regroupe quatre types d'acteurs :

- **Le sujet des données** - par exemple l'individu qui porte une montre connectée et génère de la donnée en l'utilisant avec une application de sport ;
- **Le détenteur des jeux de données** - par exemple la société éditrice de l'application de sport intégrée à la montre connectée, qui rassemble les données de tous ses utilisateurs. Il faut noter que le sujet peut parfois être également directement détenteur de ses propres données ;
- **L'utilisateur des jeux de données** - qui peut être le détenteur mais aussi, par exemple, des équipementiers sportifs qui achètent au détenteur ses jeux de données pour faire des études destinées à créer de nouveaux produits ou à lancer une campagne de publicité ciblée ;
- **L'intermédiaire de données**²⁶ - il opère techniquement le transfert de données d'un sujet/détenteur vers un autre acteur tel que l'utilisateur ou l'organisation altruiste (recueil du consentement, traitement pour assurer l'interopérabilité, gestion des API). Il est considéré comme neutre au sens où il ne réalise aucune opération d'exploitation des données et se contente de réaliser des prestations de services liées à la circulation des données.

²³ HENDERSON, W. Harry, «The Early Poll », *The Public Opinion Quarterly*, 1942.

²⁴ MIEN Edouard, PAPP Alizé, "L'afflux massif de données est-il nouveau ? Entretien avec Emmanuel Didier", *Regards croisés sur l'économie*, n°23, pp. 14-26, 2018.

²⁵ SUBERVIE Julie, GALTIER Franck, *L'information sur les prix agricoles par la téléphonie mobile : le cas du Ghana*, AFD, 2014.

²⁶ Termes utilisés au sein du DGA, également appelés *data sharing intermediaries*.

La nouveauté apportée par le DGA réside notamment dans la création d'un cinquième type d'acteurs : un tiers de confiance appelé « *organisation altruiste en matière de données* », qui collecte la donnée auprès des détenteurs, pour la partager ultérieurement avec des utilisateurs.

DÉFINITION DE L'ORGANISATION ALTRUISTE EN MATIÈRE DE DONNÉES

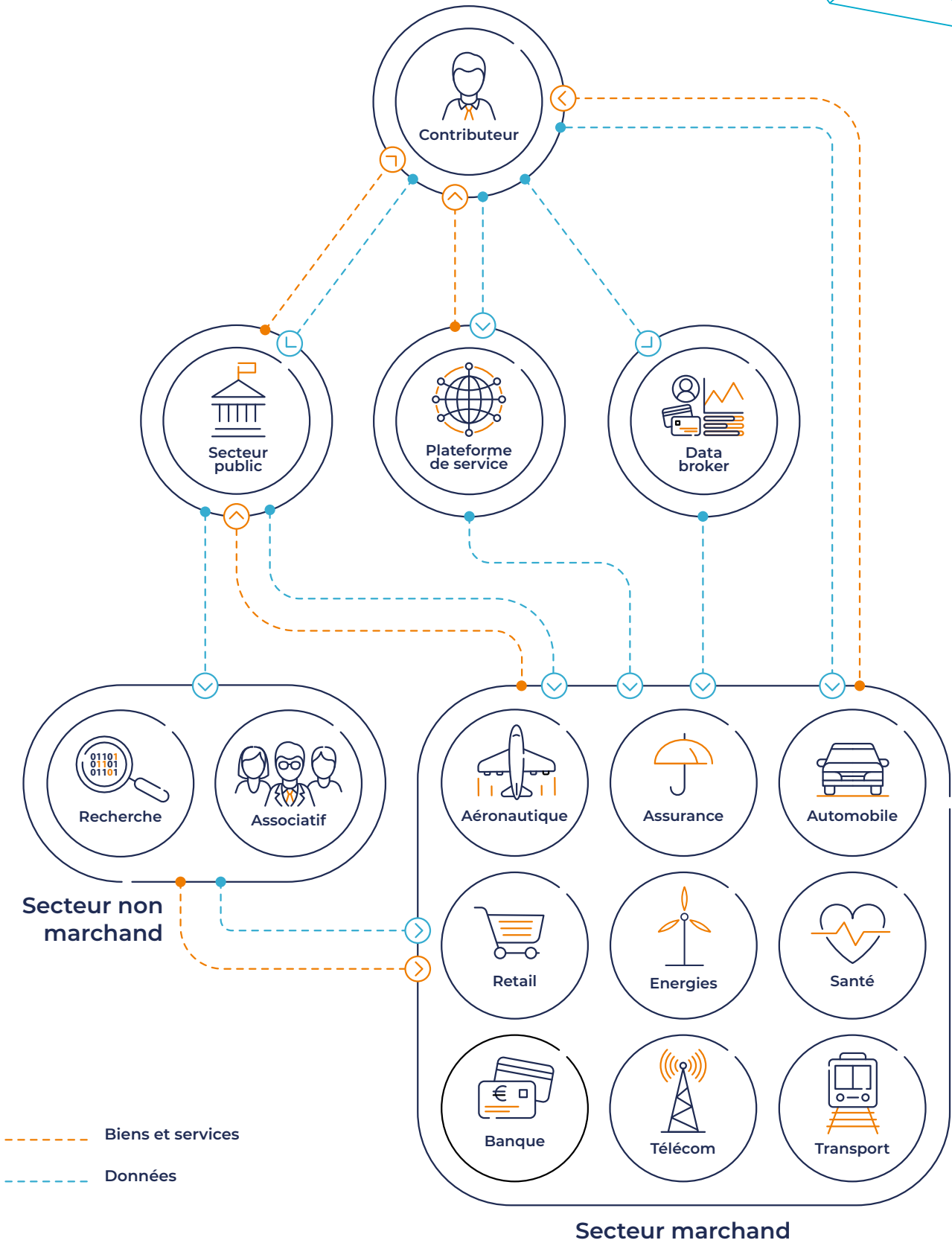
Pour être enregistrée comme étant altruiste en matière de données une organisation doit :

- Effectuer des activités de data altruisme ;
- Être une personne morale constituée conformément au droit national pour répondre à des objectifs d'intérêt général, telle que la notion a été définie dans le droit national ;
- Opérer dans un but non lucratif et être légalement indépendante de toute entité poursuivant un but lucratif ;
- Mener les activités liées à l'altruisme en matière de données par l'intermédiaire d'une structure fonctionnellement séparée d'autres activités.

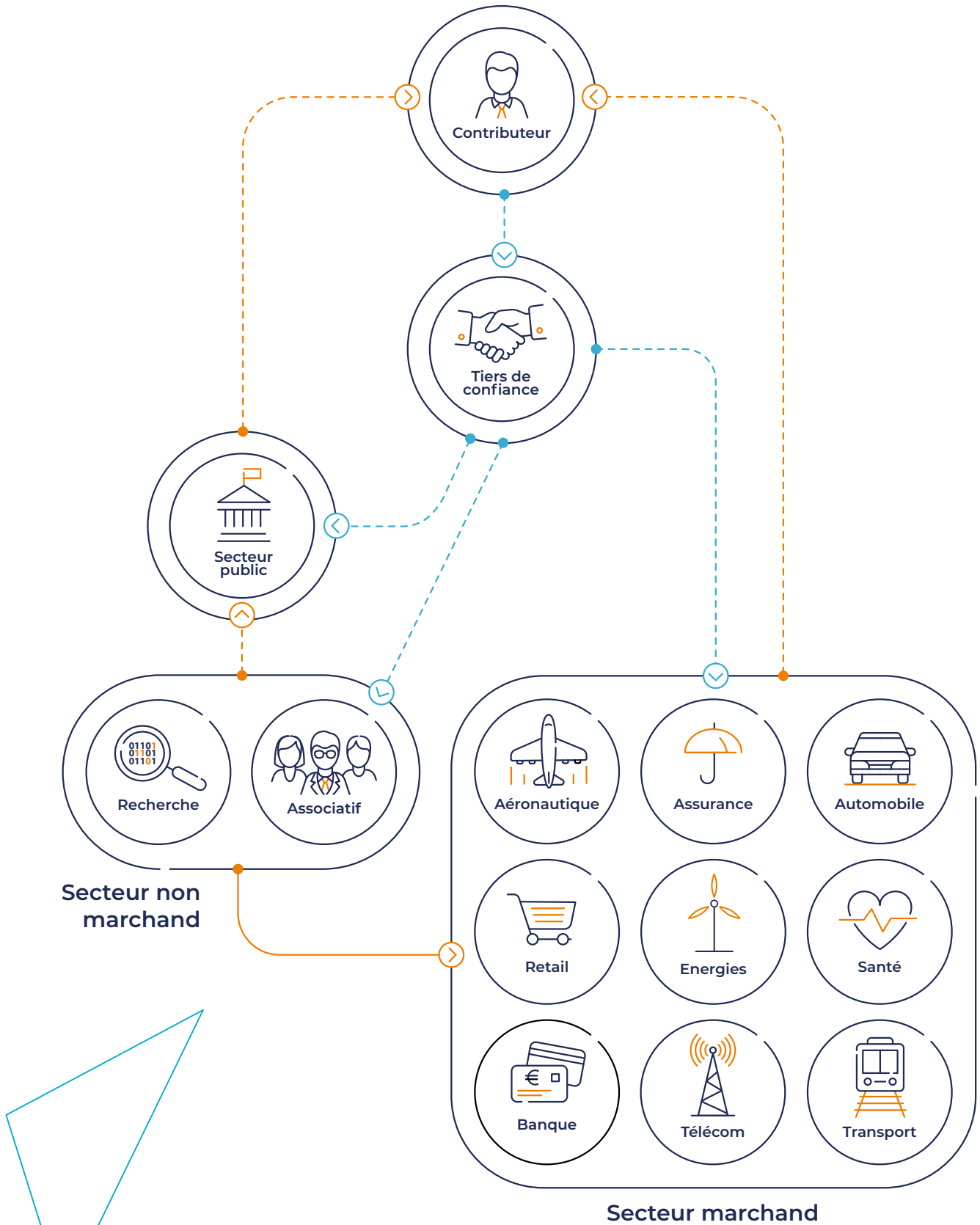
De façon concrète, le DGA n'identifie que trois types d'acteurs jouant un rôle dans l'altruisme des données : « l'agent altruiste » qui accepte le transfert de ses données - le détenteur des jeux de données (article 2 définition) ; l'intermédiaire de données qui assure l'infrastructure de partage et « l'organisation altruiste en matière de données » - le tiers de confiance chargé de traiter les données (article 16). Cependant, ce sont bien le sujet, le détenteur, le tiers de confiance et l'utilisateur qui interagissent au sein de la démarche altruiste, selon le schéma suivant :



PARTAGE DE DONNÉES DANS UN CADRE MARCHAND



PARTAGE DE DONNÉES DANS UN CADRE DATA ALTRUISTE



LE DATA ALTRUISME : UN PAS DE CÔTÉ EUROPÉEN DANS LE CHAMP NUMÉRIQUE

Le marché unique européen est marqué par la distorsion due à la présence de géants de la donnée à la position quasiment monopolistique, désormais les gardiens du seuil de l'économie numérique mondiale. Cela appelle à une réaction de l'UE et des États membres pour construire des capacités au niveau de l'état de l'art mondial. Dans ce contexte international, le data altruisme européen représentant une approche spécifiquement européenne.

Les États-Unis s'appuient sur ces multinationales de la Silicon Valley et la portée extraterritoriale de leur réglementation, afin d'édicter des normes applicables à toute personne morale ou physique de nationalité étrangère. Ce cadre juridique offre l'opportunité aux autorités américaines de collecter des données auprès d'entreprises étrangères positionnées dans des secteurs stratégiques (la banque, les transports, l'énergie et la défense).

Concurrente de taille, la Chine a fondé son modèle sur sa puissance économique et un contrôle sans faille des données. Depuis 2021, les données des entreprises et

des citoyens chinois font l'objet de transactions sur le Shanghai Data Exchange (SDE). Cette monétisation des données est une stratégie assumée des autorités chinoises dans leur quête de puissance, alors que les données sont considérées comme « *un cinquième facteur de production*²⁷ ». Les données font également l'objet d'un contrôle étroit de la part des citoyens. Ils disposent ainsi d'une copie de leurs données afin de les corriger, d'en contrôler l'utilisation et de les détruire. Ce dispositif offre aussi la possibilité aux autorités de superviser les transferts de données à l'étranger lui assurant une position de force sur le marché mondial des données.

Qu'il s'agisse de l'exemple américain, chinois ou européen, tous observent une finalité similaire : la diffusion de leur modèle et le développement d'activités économiques alimentées par les données. Pour l'Europe, le développement d'un mécanisme de captation et d'exploitation des gisements des données est une capacité clé sur laquelle les entreprises européennes accumulent du retard par rapport à leurs concurrentes chinoises et américaines.

LES DÉFIS RENCONTRÉS PAR LE DATA ALTRUISME

Pour se développer, le data altruisme doit affronter trois défis majeurs :

- gagner la confiance des parties prenantes ;
- identifier et entretenir leurs motivations à participer à de tels modèles ;
- susciter des mécanismes de financement assurant la pérennité des organisations altruistes.

Ces obstacles s'articulent de différente manière selon les spécificités des acteurs concernés : entreprises, individus, et pouvoirs publics.

DÉFI N°1 : ASSURER UN HAUT NIVEAU DE CONFIANCE DANS LES MODÈLES DATA ALTRUISTES

La confiance est nécessaire pour initier la relation de partage et pour la faire perdurer dans le temps. Or le niveau de confiance actuellement existant a été gravement obéré par la révélation de scandales aux répercussions considérables liés à l'exploitation de données :

- Le scandale dit « Cambridge Analytica » a révélé que l'exploitation déloyale de données pouvait avoir une influence considérable sur les scrutins démocratiques ;
- Les publications successives de fichiers contenant des milliards de combinaisons d'adresses emails et de mots de passe - récemment ROCKYOU 2021 qui a rendu publique une liste de plus de 8 milliards de mots de passe - rappellent à quel point la sécurité informatique est fragile ;
- La récente fuite de données de l'Assistance Publique-Hôpitaux de Paris (AP-HP) concernant 1,4 millions de personnes testées contre la COVID-19 mi-2020.

Des systèmes de contrôle étatiques au « *capitalisme de surveillance*²⁸ », plusieurs facteurs jouent ainsi contre le partage des données. Par ailleurs, les entreprises manifestent également une réelle défiance face à toute possibilité de partage de ce qui constitue bien souvent pour elles un actif stratégique.

²⁷ Après le capital, le travail, la terre et la technologie selon XIX^e Congrès du Parti communiste chinois.

²⁸ ZUBOFF, 2019 cité in VILJOEN, 2021.

La défiance qui en résulte fait de la construction d'un système inspirant la confiance la clef de voûte d'un mécanisme nouveau et efficace de partage de données. Le lien de confiance, au cœur du data altruisme, se situe à deux niveaux :

- **Entre les contributeurs et l'organisation altruiste** : ce lien est sans doute le plus difficile à créer, l'enjeu étant évidemment d'être en mesure de rassurer les contributeurs potentiels pour les amener à partager sereinement les données qu'ils détiennent ;

- **Entre l'organisation altruiste et les utilisateurs des données** : ce lien est sans doute plus aisé à construire, et repose sur la création d'une relation permettant un contrôle de l'utilisation des données qui ne bride toutefois pas inutilement cette dernière.

Les ressorts de la confiance sont complexes à appréhender car relevant d'une « *famille conceptuelle aux formes variées et aux frontières floues*²⁹ ». Ainsi, la confiance apparaît :

- **Polysémique**, car « faire confiance » implique des modalités d'action qui diffèrent selon l'acception que l'on a du mot. Sans épuiser le débat, il est possible de mentionner deux types de confiance de nature diverse : la confiance systémique et la confiance épistémique³⁰. La première vise les conditions pratiques de l'échange (la structure qui le supporte, les modalités de contractualisation et de révision, etc.), tandis que l'autre s'attache aux représentations symboliques pendant l'échange (l'institution garante de son bon déroulement, les rituels, etc.).

- **Multifactorielle**, les travaux de confiance dans le numérique par l'ISO considérant, par exemple, une vingtaine de facteurs principaux³¹. Dans le cas du data altruisme, le tiers de confiance, les autres contributeurs, les acteurs impliqués dans la maintenance de l'infrastructure, l'autorité de régulation peuvent tous avoir une acception différente de la notion et de ses leviers. Suivant les dispositions de chacun, la confiance repose sur le caractère altruiste de la finalité, les modalités de sécurisation des données ou les possibilités d'évolution de la gouvernance.

- **Contextuellement** changeante car, selon les pays, les époques, les dynamiques, les sujets, la capacité à faire confiance évolue considérablement. Une vision culturaliste est parfois appliquée à la confiance dans le monde numérique, les États du nord de l'Europe étant supposés plus susceptibles de faire confiance dans les dispositifs techniques que dans ceux du sud. Si une telle analyse peut être nuancée, on observe de nettes disparités dans le recours à des services numériques similaires en fonction de leur contexte d'utilisation³².

Le numérique introduit-il des logiques de confiances inédites ? Son caractère immatériel, instantané et parfois obscur, peut exacerber une certaine méfiance. À l'opposé, son rôle de facilitateur des échanges ouvre de nouvelles possibilités de partage. Les démarches de développement en source ouverte, au cœur des premiers projets numériques, ou plus récemment le financement participatif, en sont des témoins. Le numérique n'est ainsi pas qu'une matière désincarnée, il « *n'est pas un objet technique, c'est un ensemble de structures symbolique matériel culturel et social*³³ ».

LE CAS FINLANDAIS DU BIOBANK ACT : LES LIMITES DE L'EXPLICATION CULTURALISTE

Si la Finlande a fortement développé les collectes de données génétiques auprès d'une part importante de sa population, notamment suite au Biobank Act en 2013, la confiance reste « *fortement influencée par la façon dont un participant potentiel interprète les contextes de la demande de don*³⁴ ». Les premières collectes se sont ainsi déroulées avec grandes difficultés, des sondages montrant que seuls 40 % des Finlandais avaient entendu parler de telles collectes et, parmi eux, 63 % sont considérés comme ne les comprenant pas suffisamment pour prendre une décision éclairée³⁵. Le mode de recueil du consentement est ainsi capital. Un recueil par un ambassadeur de la biobanque en face à face ou pour des maladies spécifiques génère d'importants taux de réponses positives, contrairement à un simple mail. En conclusion, « *des attitudes favorables et positives ou des niveaux élevés de confiance générale dans une société donnée ne sont pas en eux-mêmes des indications directes de la volonté et de la préparation réelles des personnes à participer aux biobanques*³⁶ ».

²⁹ Milad DOUEIHI et Jacopo DOMENICUCCI (dir.), La confiance à l'ère numérique, Éditions rue d'Ulm 288 p., 2018.

³⁰ DOUEIHI et DOMENICUCCI, 2018.

³¹ ISO 16363:2012 - Systèmes de transfert des informations et données spatiales - Audit et certification des référentiels numériques de confiance.

³² Début 2021, l'application de suivi de contact utilisée dans la politique anti-covid avait été téléchargée par 21 % de la population en Allemagne, 14 % en Italie, 15 % en France et jusqu'à 40 % en Islande et Singapour, les pays avec les taux les plus élevés in « Pourquoi l'adoption des applications de traçage anti-Covid n'a pas été au rendez-vous en Europe », *La Tribune*, février 2021.

³³ DOUEIHI et DOMENICUCCI, 2018

³⁴ RAIVOLA Vera, SNELL Karoliina, HELEN Ilpo, PARTANEN Jukka, « Attitude of blood donors to their sample and data donation for biobanking », *European Journal of Human Genetics*, 2019.

³⁵ SNELL K, STARKBAUM J, LAUB G, VERMEER A, HELENI, « From protection of privacy to control of data streams: a focus group study on biobanks in the information society », *Public Health Genomics*, 2012.

³⁶ RAIVOLA, SNELL, HELEN, PARTANEN, 2019.

DÉFI N°2 : SAISIR LES SPÉCIFICITÉS, CRAINTES ET INTÉRÊTS DE CHAQUE PARTIE PRENANTE VIS-À-VIS DU DATA ALTRUISME

Si la création d'un lien de confiance est la condition *sine qua non* du succès de l'altruisme des données, les clefs de sa création auprès de ses différents acteurs ne sont pas les mêmes, les raisons de la défiance n'étant, elles-mêmes, pas identiques.

DÉMONTRER AUX ENTREPRISES L'INTÉRÊT DU DATA ALTRUISME Y COMPRIS DANS UN CONTEXTE DE FORTE PRESSION CONCURRENTIELLE

L'entreprise a pour particularité d'évoluer dans un milieu centré sur la concurrence et les lois de marché. Pour émerger, croître et se pérenniser, chaque entreprise doit trouver, créer, conserver un ou des avantages concurrentiels qui lui permettent de se distinguer des autres.

Ces avantages concurrentiels peuvent prendre des formes diverses, certains étant d'ailleurs conditionnés à leur caractère public, comme c'est le cas des marques ou des brevets d'invention. D'autres, en revanche, n'ont de valeur que parce qu'ils demeurent secrets. Tous partagent toutefois la recherche d'une exclusivité maximale. Cette recherche de l'exclusivité, voire du secret, s'accommode mal avec l'idée d'un partage altruiste.

Or, dans un contexte de numérisation presque totale de la société, et donc de l'économie, les données brutes ou traitées s'avèrent non seulement omniprésentes dans les entreprises mais également :

- essentielles à la conduite de leurs activités ;
- hautement convoitées par leurs concurrents.

Si leur exploitation efficace peut être une source considérable de profits, une telle exploitation est également génératrice de coûts potentiellement hors de portée de beaucoup de PME. Notons en effet que le secteur privé ne constitue pas un groupe homogène. Les capacités d'investissement dans la mise à disposition de données, la propension à partager des données sans exploitation commerciale, divergent considérablement entre un grand groupe et une PME. Pour le premier, il y a une plus grande facilité à mettre en œuvre un partage de données sans retour sur investissement direct. Pour la seconde, en revanche, le souci de rentabilité à court terme, la capacité à se projeter dans le temps et la difficulté à mettre en œuvre techniquement le partage de données peuvent limiter la portée de la démarche altruiste.

À charge pour l'organisation data altruiste de créer et déployer un système de collecte de données ou de prévoir les modalités de réutilisation d'un système de collecte existant. L'ensemble doit être sécurisé pour se prémunir de tout acte malveillant, tout du moins pour en limiter les impacts. C'est ainsi que, bien que stratégiques, les données des entreprises restent souvent secrètes et inusitées, conduisant à la déclaration faite par la présidente von der Leyen selon laquelle 80 % des

données industrielles collectées ne sont pas exploitées. Et si ce constat est qualifié de gaspillage, c'est parce que les entreprises sont des collecteurs naturels de données, qui proviennent de sources multiples :

- des clients, des usages observés, du suivi de la consommation, des expérimentations ;
- de l'interne, à travers des capteurs, des machines/robots dans l'amélioration des process ;
- de tiers, à travers des partenariats avec d'autres entreprises, des associations, etc. ou à travers la monétisation ou l'achat de fichiers.

Il va sans dire que le traitement de tels jeux de données, considérables dans leur volume, précieux dans leur nature, pourrait permettre des avancées considérables dans les secteurs d'intérêt général. Pour y accéder, il faut donc tout d'abord savoir susciter la confiance des entreprises, en répondant à leurs préoccupations principales.

Les organisations data altruistes devront ainsi être capables de démontrer que les données qui leur seront transmises ne seront pas mises à la disposition du marché concurrentiel. Une fois cet obstacle franchi, nous soulignons que, si le partage de données dans l'intérêt général n'est pas encore entré dans les mœurs, il en va différemment d'autres formes d'altruisme.

Le milieu de l'entreprise est de longue date accoutumé aux œuvres de charité, financement de fondations et autres donations, qui ont un impact significatif tant sur les projets soutenus que sur l'image des mécènes. Il faut donc parvenir à connecter une pratique traditionnelle à une modalité novatrice pour voir l'altruisme des données devenir une norme sociale, garante de son succès.

Ce, d'autant plus que la réquisition de données par les pouvoirs publics sur un certain nombre de secteurs clefs est une possibilité amenée à s'affermir. La solution portée par le DGA de créer des tiers de confiance indépendants, non gouvernementaux et à but non lucratif, apparaît comme réunissant nombre de conditions à même de susciter la confiance des entreprises. Il conviendra toutefois de créer au sein de ces organisations des modes de gouvernance suffisamment structurés et transparents pour rassurer les entreprises sur le contrôle des flux de données.

ASSURER LA CONFIANCE DES INDIVIDUS ALORS QUE LEURS DONNÉES SONT DÉJÀ SUR-SOLLICITÉES

Les individus ont un rapport particulier à leurs données, qui fluctue entre une très grande propension à les partager et une approche protectrice, voire défensive, parfois farouche. C'est ainsi que l'on constate très paradoxalement une volonté de protection en apparence moins exacerbée dans le cadre du partage de données en contrepartie de la fourniture de services que face à des demandes d'accès formulées par des services publics, pourtant supposés agir dans l'intérêt de la collectivité. En témoigne le taux de pénétration des réseaux sociaux dans la même population qui s'est pourtant fortement mobilisée, parfois d'ailleurs par l'intermédiaire desdits réseaux sociaux, contre le

déploiement d'applications dont le but était d'endiguer la propagation du Covid-19 mais qui étaient jugées attentatoires à la vie privée. Toutefois, les individus, personnes physiques, sujets des données personnelles qui alimentent les modèles économiques des grandes entreprises du numérique, ont de plus en plus conscience de la valeur économique que représentent leurs données.

Et, comme nous le soulignons précédemment, s'ils « consentent » encore à leur partager leurs données, un mouvement de fond est apparu qui milite pour la patrimonialisation de celles-ci et la rémunération des sujets de données.

LA QUESTION DE LA PATRIMONIALISATION DES DONNÉES

Il existe deux grandes approches permettant de décrire la relation de l'individu à ses données personnelles³⁷ : les approches patrimonialistes et les approches dignitaires.

Les premières renvoient à la capacité d'utiliser ses données comme des choses, et donc d'avoir sur elles l'*usus* (le droit d'user), le *fructus* (le droit de jouir, tirer les fruits) et l'*abusus* (le droit de disposer).

Les secondes conçoivent les données comme un attribut de la personne, une extension de son ipséité, ce qui fait d'elle une personne unique. Elles rappellent que « *les données personnelles existent à la frontière de l'être et de l'avoir*³⁸ ».

La construction de la réglementation européenne s'inscrit fondamentalement dans une approche dignitaire du statut des données personnelles, rejetant presque totalement l'idée d'une patrimonialisation de celles-ci. Le Contrôleur européen de la protection des données rappelle ainsi que la vie privée n'est pas « *qu'un droit individuel mais une valeur sociale*³⁹ ».

La patrimonialisation des données est un argument parfois avancé dans le but de redonner aux individus le contrôle sur leurs données tout en leur permettant d'en retirer les fruits en cas d'exploitation. Le Sénat français en 2009⁴⁰, le Conseil d'État en 2014⁴¹, ou la Commission nationale de l'Informatique et des Libertés (CNIL) en 2021, ont pris position contre cette patrimonialisation. La rémunération potentielle des transferts de données personnelles réalisés par des individus paraît être sans commune mesure avec :

- la valeur personnelle du contrôle de ces données (considérable) ;
- la valeur économique initiale desdites données (nulle) ;
- la valeur finale de la base de données constituée (colossale).

Toutefois, en l'état des pratiques, force est de constater qu'existent déjà des formes nombreuses de monétisation indirecte des données personnelles (l'accès gratuit à des services numériques est réalisé en contrepartie d'un accès aux données des utilisateurs, lesquelles sont revendues à des tiers, souvent à des fins de publicité ciblée). Il est par ailleurs très probable que des pratiques de monétisation directe soient de plus en plus fréquentes, certaines étant déjà en place actuellement⁴². Néanmoins, l'approche patrimonialiste est à l'opposé, non seulement de la réglementation en vigueur, mais aussi de la proposition faite par le DGA de créer un mouvement d'altruisme en matière de données.

³⁷ VILJOEN, 2021.

³⁸ [Nos données n'appartiennent qu'à notre humanité, La Quadrature du Net, 2021.](#)

³⁹ [Data Protection, EUR. DATA PROT. SUPERVISOR, in VILJOEN, 2021.](#)

⁴⁰ « Notre conception de la vie privée place sa protection sur le terrain de la dignité humaine. En se référant au concept de propriété, le risque de marchandisation de ses données personnelles est évident. La propriété comprend aussi le droit de céder la chose. Or, comment pourrait-on céder une donnée qui peut aussi être un attribut de sa personnalité ? » p. 106 in DETRAIGNE Yves, ESCOFFIER Anne-Marie, Rapport d'information relatif à la vie privée à l'heure des mémoires numériques, n°441, p.106, 2009.

⁴¹ [Le Numérique et les droits fondamentaux](#), Conseil d'État, 2014.

⁴² [Amazon propose de payer 10 dollars pour l'empreinte palmaire de ses clients, L'Usine digitale, 2021.](#)

L'EXEMPLE FICTIF DE TATTOOVIEW IA

Dans un article analysant les limites des approches propriétaires et dignitaires, la juriste américaine, Salomé Viljoen⁴⁴, propose un exemple inspiré de développements réels au sein des forces de police étatsuniennes. Elle imagine la signature d'un partenariat entre le FBI et l'entreprise TattoView IA dans le but de fournir des solutions de reconnaissance d'un individu grâce à son tatouage, mais aussi d'outils d'évaluation de l'appartenance à un gang en fonction des tatouages spécifiques qu'il utilise.

Deux personnages entrent alors en scène : Adam, ancien membre d'un gang et tatoué et Ben, également tatoué et suspecté d'être membre d'un gang (à tort ou à raison). Selon les scénarios explorés, Adam pourrait fournir l'image de son tatouage et la signification associée de plusieurs façons :

- de manière coercitive car arrêté par la police dans le cadre d'une enquête (obligation légale) ;
- en le partageant sur un réseau social d'amateurs de tatouage racheté par TattoView IA qui en exploite les données (contrat général d'utilisation d'un service) ;
- en le vendant à un courtier de données proposant des sommes modiques en l'échange d'informations corporelles (contrat de patrimonialisation) ;
- dans le cadre d'un programme d'anciens membres de gang repentis souhaitant soutenir la stratégie fédérale de lutte contre la criminalité (data altruisme).

D'après Viljoen, ce qui rend ce flux de données sur les tatouages potentiellement injuste ou dommageable, ce n'est pas le niveau de rémunération reçu par Adam, ou la possibilité qu'il a eu de faire prévaloir ses droits lors de la collecte de la donnée. Le risque réel est que l'utilisation de ces informations personnelles ait un impact négatif sur un groupe plus large (en facilitant la mise en place d'une politique répressive dans ce cas).

Ainsi, une gouvernance de la donnée doit prendre en compte, non seulement la relation verticale d'Adam avec ses données, mais aussi la relation horizontale entre Ben et les données d'Adam.

En l'état, il paraît qu'en dehors d'une incitation forte, voire contrainte, à partager des données avec un acteur économique (tel que l'accès à un réseau social devenu incontournable ou à un service devenu indispensable) il sera de plus en plus difficile d'obtenir des partages de données sans proposer un « retour sur investissement ».

Une telle démarche étant à l'opposé de la politique européenne en la matière (voir encadré sur la patrimonialisation des données), il est nécessaire de trouver des incitations non financières au partage.

Or, il existe chez les individus une forme de méconnaissance des bénéfices que pourrait retirer la collectivité de l'exploitation correctement encadrée de leurs données, ce qui ne contribue pas à créer une dynamique de partage.

S'il est en effet démontré et connu de tous que la collecte puis le traitement à grande échelle de données personnelles permettent la création de biens et de services qui ont des effets sociaux, les illustrations médiatisées de ce phénomène sont majoritairement négatives (voir notamment Cambridge Analytica, ROCKYOU 2021 cités plus haut).

Les individus doivent donc être, plus que sensibilisés, pleinement informés des finalités pour lesquelles leurs données pourraient être récoltées, et des bénéfices qui pourraient en résulter pour la société et, indirectement donc, pour eux individuellement.

Pour surmonter ces obstacles, susciter la confiance des individus et les inciter à partager leurs données, la création de tiers de confiance indépendants, non gouvernementaux, à but non lucratif, tenus par un code de conduite⁴³, apparaît comme une solution particulièrement adéquate. Ce code aurait notamment pour objectif de préciser les activités de l'organisation data altruiste et les principes qu'elle se doit de respecter. Si ce code reste à construire et se fera avec un dialogue avec les États membres, la société civile ou encore l'*European data innovation board* (groupe d'experts), les deux propositions d'actions ci-dessous en propose une ébauche.

La définition des données recherchées et la gouvernance qui doit être mise en place au regard de la nature et de la sensibilité desdites données, dépendent de la finalité. La première étape de la création d'une organisation data altruiste est donc la définition de la ou les finalités - nécessairement d'intérêt général - qu'elle entend servir.

Ces organisations, n'ayant pour seule finalité que de promouvoir des finalités d'intérêt général, informeront les individus desdites finalités, des projets auxquels elles contribuent, et donc des bénéfices sociaux en résultant.

L'approche non lucrative devrait permettre de désamorcer les velléités patrimoniales et les inquiétudes liées à création de valeurs non redistribuées aux contributeurs et sujets des données.

Enfin, l'indépendance des organisations vis-à-vis des pouvoirs publics permettra de garantir une « non-ingérence » de ces derniers dans l'exploitation des données, au bénéfice d'un renforcement de la confiance.

⁴³ En préparation par la Commission.

⁴⁴ VILJOEN, 2021.



DÉFI N°3 : IDENTIFIER DES SOURCES DE FINANCEMENTS PÉRENNES POUR LES MODÈLES DATA ALTRUISTES TOUT EN ASSURANT LEUR INDÉPENDANCE

La création d'une organisation est, en soit, synonyme de charges et de coûts liés notamment à son organisation ou à sa gestion administrative.

Ces coûts sont d'autant plus élevés pour les organisations altruistes en matière de données que leur activité, éminemment technique, implique des processus coûteux qui vont au-delà de la collecte de données et de leur stockage et impliquent⁴⁵ :

- L'enrichissement des données par l'ajout d'informations supplémentaires dans le jeu de données ;
- La standardisation pour homogénéiser les données collectées par différents acteurs ;
- L'insertion « d'opérateurs d'articulation » venant combler le manque d'informations contenues dans le jeu de données par l'ajout de données pivot⁴⁶, et favoriser la polyvalence.

Le DGA prévoit la possibilité pour les organisations altruistes de mettre en place des systèmes d'accès aux données en contrepartie de redevances ou de frais administratifs. La mise en place de licences d'utilisations des données impliquant le paiement de redevances pourra ainsi permettre aux organisations altruistes de créer un modèle économique qui, bien que n'ayant pas une finalité lucrative, permettra le fonctionnement de l'institution et le financement d'investissements.

Il convient toutefois de veiller à l'instauration d'un système de redevances proportionné et équitable, adapté aux caractéristiques des différents utilisateurs et propice à l'exploitation des données dans l'intérêt général.

Nous notons, par ailleurs, que les dispositions du DGA, qui prévoient que l'organisation altruiste doit assurer son indépendance juridique et fonctionnelle, n'interdisent pas la mise en place d'une politique de financement par mécénat, à condition toutefois que des mesures claires soient mises en place pour préserver l'indépendance de l'organisation.

Enfin, si l'indépendance des organisations altruistes est un élément clé de la confiance qui leur est accordée, elle ne paraît pas incompatible avec le financement d'une partie de leur fonctionnement par les pouvoirs publics, notamment au stade de leurs premiers développements (subvention, crédit d'impôt, soutien technique, appel d'offre ou appel à projet, cf. actions n°20 et 21).

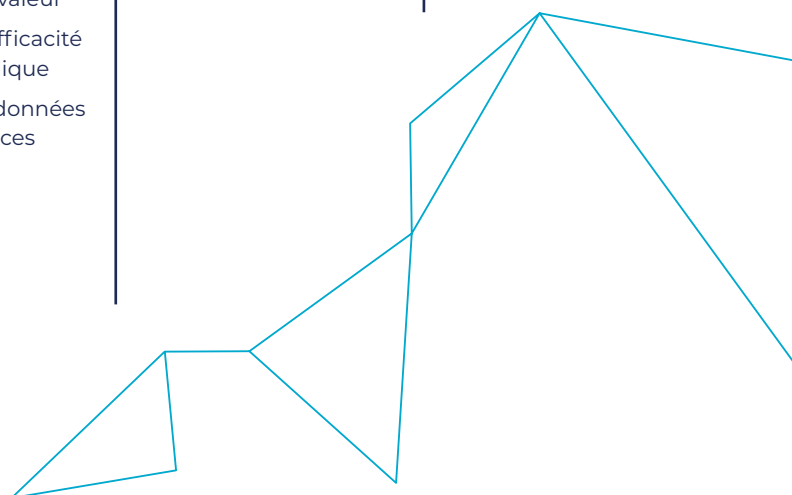
⁴⁵ COURMONT Antoine, « Open data et recomposition du gouvernement urbain : de la donnée comme instrument à la donnée comme enjeu politique », *Informations sociales*, n°191, 2015.

⁴⁶ Code Insee, Iris, Siren, Siret, numéro de sécurité sociale, etc.



SYNTHÈSE DES INCITATIONS À PARTICIPER À UN MODÈLE DATA ALTRUISTE

Individus	Pouvoirs publics	Entreprises	
		qui donnent leurs données	qui reçoivent des données
<ol style="list-style-type: none"> 1. Obtenir des retombées concrètes à l'utilisation de ses données 2. <i>A minima</i> sans détérioration de service 3. Se sentir acteur d'un système de production de valeur, que cela soit avec <ol style="list-style-type: none"> a. une agentivité forte (portabilité citoyenne des données, implication importante dans le modèle data altruiste) b. une agentivité faible (gain moral à participer à une bonne cause) 	<ol style="list-style-type: none"> 1. Appliquer les valeurs partagées en société dans le monde numérique 2. Disposer de solutions d'exploitation de données en-dehors du marché classique (gage d'indépendance) 3. Stimuler la création d'un écosystème producteur de valeur 4. Améliorer l'efficacité de l'action publique 5. Valoriser les données issues des services publics 	<ol style="list-style-type: none"> 1. Disposer d'un levier d'engagement social 2. Valoriser son image auprès de parties prenantes 	<ol style="list-style-type: none"> 1. Accéder à des données non présentes sur le marché



SYNTHÈSE DES FREINS À PARTICIPER À UN MODÈLE DATA ALTRUISTE

	Individus	Pouvoirs publics	Entreprises	
			qui donnent leurs données	qui reçoivent des données
Confiance	<ul style="list-style-type: none"> - Méfiance à l'égard des acteurs économiques et étatiques - Opacité de la gouvernance du système 		<p>Crainte d'une exploitation des données fournies par des concurrents</p>	<p>Opacité sur l'origine des données engendrant la crainte d'un acte malveillant</p>
Motivation	<ul style="list-style-type: none"> - Absence de considération pour les causes altruistes proposées - Absence de bénéfice direct à partager ses données - Forte sollicitation dans le champ de la philanthropie classique 	<ul style="list-style-type: none"> - Absence d'une culture organisationnelle favorable au développement de ce genre de projets - Numérisation hétérogène du secteur public - Difficulté à identifier les acteurs adéquats parmi les différents champs de compétence 	<ul style="list-style-type: none"> - Refus de capter et de fournir des données qui ne seraient pas directement utiles à l'activité de l'entreprise - Absence de considération pour les causes altruistes proposées 	<p>Abondance de sources de données auxquelles les entreprises ont accès habituellement</p>
Financement	<p>Faible capacité de financement (hors philanthropie)</p>	<p>Investissement de départ significatif</p>	<p>Investissement potentiellement significatif</p>	<p>Difficulté à mettre en place un modèle d'affaire</p>
Technique	<ul style="list-style-type: none"> - Maîtrise de la technologie hétérogène - Difficulté à recueillir le consentement éclairé - Fracture numérique 	<p>Numérisation en cours du service public pouvant conduire à une priorisation moindre des projets data altruistes</p>	<ul style="list-style-type: none"> - Investissement sur la qualité des données en amont de leur transmission - Investissement de moyens (humains, techniques, financiers) 	<ul style="list-style-type: none"> - Qualité des données insuffisante pour répondre au besoin - Mobilisation de moyens (humains, techniques, financiers)

ÉLÉMENTS CLEFS PARTIE 1

- Le data altruisme a pour objectif de stimuler les échanges de données dans le cadre d'une compétition internationale exacerbée pour la maîtrise des outils d'analyses (IA, machine learning, etc.).
- Il n'invente pas toutefois les mécanismes de partage de données hors-marché, il les complète. Sur certains aspects (flux BtoG), il devance des mécanismes coercitifs (réquisition) qui pourraient se développer.
- Le data altruisme est donc une vision technique et politique du numérique. Elle propose un rapport à la création de valeur qui ne passe pas par l'économie marchande.



2.

UN PAYSAGE DES
NOUVEAUX MODES DE
PARTAGE DE DONNÉES
RICHE ET MOUVANT

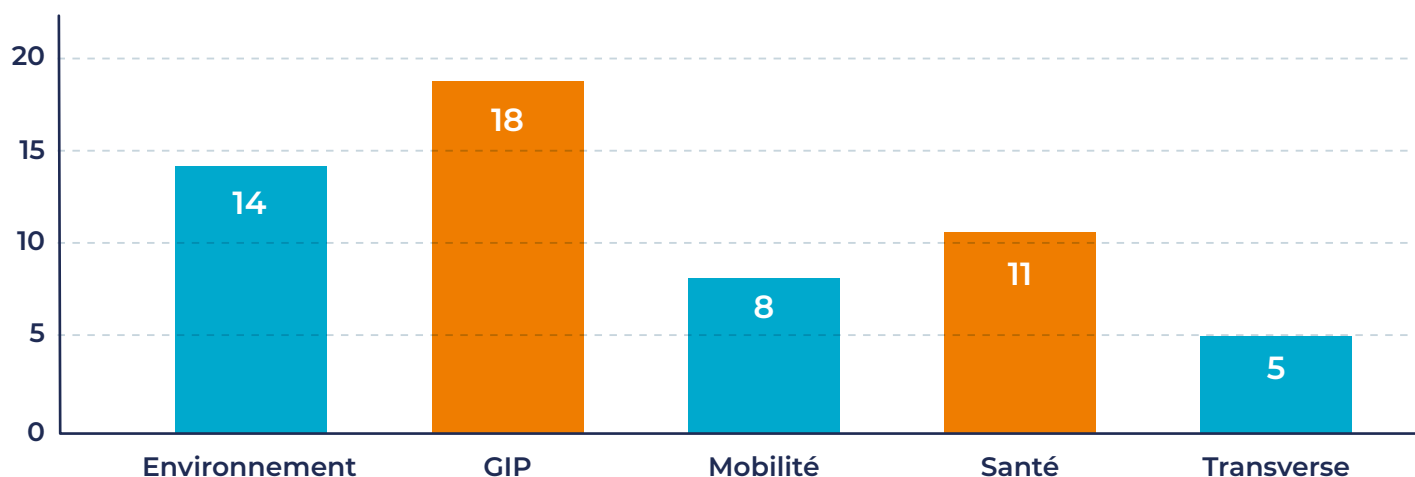
L'institution par le DGA d'une démarche altruiste structurée intervient dans un contexte où les modes de partage de données novateurs et collaboratifs, servant l'intérêt général sont déjà nombreux. Il convient de penser les modalités d'insertion de la notion de data altruisme dans cet écosystème riche et mouvant.

Un travail de recherche a permis l'identification de 56 initiatives existantes. Elles ont été réparties dans une matrice dont l'objectif est d'identifier les invariants, les dynamiques sectorielles et les risques ; autant de références que des modèles data altruistes devront investir et d'écueils qu'ils devront éviter.

Cette matrice est divisée en quatre secteurs : Environnement, Santé, Gestion des informations personnelles (GIP ou PIM⁴⁷) et Transport. Une catégorie désignée comme Transverse a également été ajoutée. À noter que certains cas pratiques sont concernés par différents secteurs d'activité.

L'ensemble des cas étudiés figure en annexe de ce rapport et les plus emblématiques seront analysés en détail ci-après, via des fiches projet.

NOMBRE DE CAS PAR SECTEUR ANALYSÉ



Les initiatives étudiées ne correspondent pas forcément à une pratique que l'on pourrait pleinement qualifier d'altruiste, notamment au sens du DGA. Elles permettent toutefois de mieux saisir l'écosystème au sein duquel des modèles altruistes pourraient émerger.

Nous avons cherché, à travers cette analyse, à identifier des dynamiques de partage de données propres à certains secteurs et d'autres en tout ou partie généralisables.

Cela nous a permis d'identifier de bonnes pratiques potentiellement transposables dans le cadre de l'altruisme des données, ainsi que des risques de dérives, qui pourraient être amenées à prospérer concomitamment.

LE *PERSONAL INFORMATION MANAGEMENT* ET LA PATRIMONIALISATION DES DONNÉES

Le PIM, également appelé Self data, est un mouvement qui a pour objectif de redonner aux individus une forme de contrôle sur leurs données. Le plus souvent, ces modèles prennent la forme d'un coffre-fort où est compilé l'ensemble des données auxquelles les services marchands ont accès. Le PIM permet donc une information éclairée sur les données personnelles utilisées et des possibilités de réallocation de ces données vers d'autres services marchands ou non.

La finalité de ces réallocations distingue le PIM altruiste de la patrimonialisation. Par exemple, de nombreux cas de PIM sont orientés vers une logique de revente des données personnelles (BitsAboutMe), alors que d'autres incluent explicitement la cession de données vers le domaine scientifique ou des œuvres caritatives, etc. (DataForGood Foundation).

⁴⁷ *Personal Information Management* ou Gestion des informations personnelles, voir PARTIE 1.

DYNAMIQUES DU SECTEUR DE LA SANTÉ

Dans le secteur de la santé, nous avons étudié 11 initiatives concernant essentiellement des données personnelles et le plus souvent sensibles.

UNE PROFUSION DE DONNÉES, MAIS DE QUALITÉ ET DE SENSIBILITÉ VARIABLES

Du fait d'un héritage historique, les pays européens sont particulièrement riches en données de santé. La création de systèmes de sécurité sociale après-guerre a produit d'importantes bases de données centralisées. Il existe donc, dans le domaine de la santé, de très importantes ressources disponibles et représentatives d'échantillons de populations vastes et significatifs puisque répartis sur plusieurs générations. Ces données historiques sont toutefois de nature et de qualité très hétérogènes.

Or la question ne se résume pas à la volumétrie des données. Leur qualité et diversité sont essentielles et ne sont pas toujours suffisantes dans le secteur de la santé pour en assurer une exploitation pertinente, soit parce qu'elles sont trop anciennes, soit parce qu'elles sont difficilement rattachables à un individu.

De par leur activité de complémentaire santé, les entités composant [le groupe VYV](#) possèdent un important stock de données liés à des prestations de santé. Historiquement, ces données avaient une vocation en termes de finalité de traitement plutôt portée vers la liquidation et le remboursement de prestations rattachées à des contrats et non à des personnes. Or plusieurs personnes peuvent bénéficier d'un même contrat, dit de "mutuelle" par exemple. L'exploitation de ces données, pour être pertinente et pouvoir bénéficier à une politique, sanitaire, par exemple, nécessite ainsi un retraitement pour transformer des données contractuelles en données individuelles.

Une autre dimension de la qualité des données est leur représentativité. Auditionné, Marco FIORINI, délégué général de l'Alliance pour la recherche et l'innovation des industries de santé ([ARIIS](#)), indique que plus la recherche dispose de typologies de données importantes provenant de sources variées :

- plus les possibilités de R&D sont nombreuses
- et plus les risques de biais sont minorés.

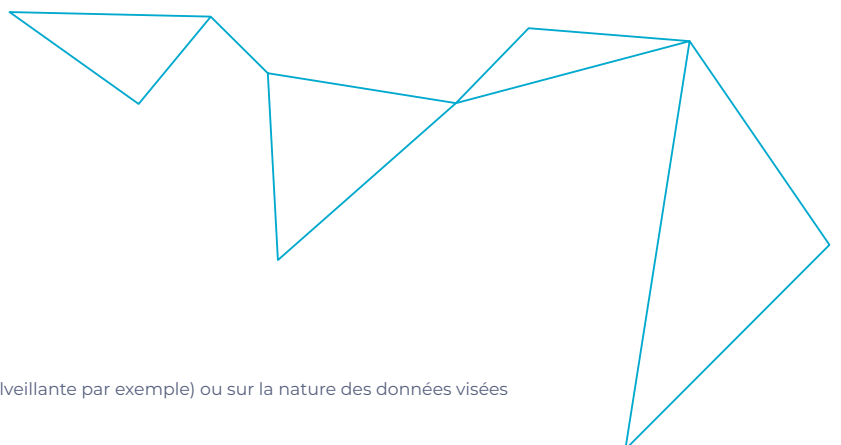
[Imagia](#) avance le fait que son modèle multicentrique (c'est à dire basé sur l'exploitation de données provenant de plusieurs centres hospitaliers) renforce la diversité et la représentativité des données sur lesquelles elle entraîne ses modèles.

Le groupe VYV souligne son souci d'appréhender les notions d'éthique de l'IA et plus généralement d'éthique du numérique. Cette approche illustre les problématiques liées aux biais des données auxquels les modèles peuvent faire face. Pour ce faire les *datascientists* du groupe sont impliqués dans des initiatives d'explicabilité de l'IA, mais aussi d'appréhension des considérations sociales que les technologies engendrent.

Impliquant de saisir l'intégralité du cycle de vie de la donnée, les enjeux de qualité sont corrélés avec ceux de sécurité. La sensibilité des données de santé est évidente. S'il existe des mécanismes pour réduire les risques, notamment la (pseudo)anonymisation, la possibilité de réidentification de données est réelle.

À noter que le projet COVI APP optait pour un parti pris pertinent : le refus d'un discours basé sur le risque zéro. Plutôt que d'assurer que son modèle est inviolable, le livre blanc du projet expose l'ensemble des risques et actes malveillants, susceptibles d'advenir et les mesures d'atténuation envisagées. Les possibilités de *hackage*⁴⁸ de l'application sont détaillées et exemplifiées.

Une politique de transparence, en ce qui concerne le risque de cybersécurité, couplée à la mise en place de mesures de protection aussi fortes que possible, paraissent créer un juste équilibre entre incitation et création d'un lien de confiance.



⁴⁸ Sur le type d'attaque (dite du « justicier », ou par une autorité malveillante par exemple) ou sur la nature des données visées (données médicales, antécédents de localisation de l'utilisateur).

L'ÉPINEUSE QUESTION DE LA PRÉCISION DE LA FINALITÉ

Si le secteur de la santé et de la recherche en général ont des besoins identifiés en données pour une certaine finalité, le principe de la recherche implique l'existence de besoins non identifiés dont l'apport ne pourrait être découvert que de manière fortuite ou à travers des processus itératifs.

Il convient donc que les collecteurs de données puissent déterminer le juste niveau de précision pour convaincre les contributeurs en données de participer à une finalité suffisamment précise pour être comprise. C'est par ailleurs une obligation dans le RGPD. La question de son respect au sein de modèles data altruistes a été soulevée par l'EDPB dans son avis sur le DGA⁴⁹. Toutefois, cette finalité doit rester suffisamment souple dans sa définition pour faciliter le travail de la recherche.

Cette tension est évidente dans le cas particulier de la donation posthume. Déjà commune dans le champ médical via le « don du corps à la science », elle se pose en des termes renouvelés dans le cadre du data altruisme. Il importe pour les organes de gouvernance en charge d'exploiter ou de transmettre ces données posthumes de s'assurer qu'elles le soient dans le cadre des volontés de la personne⁵⁰.

Toutefois, la recherche médicale est marquée par la praticité : les praticiens, notamment dans le milieu hospitalier, s'avèrent peu enclins à participer à un modèle de partage de données s'ils ne voient pas d'application dans leur quotidien. En outre, « *plus on éloigne les données de leur lieu de collecte, plus on les décontextualise en prenant le risque de mal les interpréter. En retour, c'est l'application des algorithmes au soin qui devient moins précise. En effet, en éloignant géographiquement le lieu de collecte du lieu de traitement, on perd le bénéfice d'un aller-retour correctif entre les algorithmes et la pratique réelle des soins sur de vrais patients*⁵¹».

Cette problématique se pose aussi dans le cadre d'initiatives de mise en commun de données par divers acteurs. Comme l'évoque Erwan MEDY du groupe VYV, les acteurs du monde de la santé, assurantiels ou pharmaceutiques n'auraient *a priori* pas d'opposition de principe à la constitution de modèle de partage de données de recherche, ce qu'ils peuvent d'ailleurs faire dans d'autres cadres (*open data*, collaboration sur des projets spécifiques, etc.). En revanche, il y aura des exigences pratiques multiples : cadre de gouvernance, modalités de partage et d'hébergement.

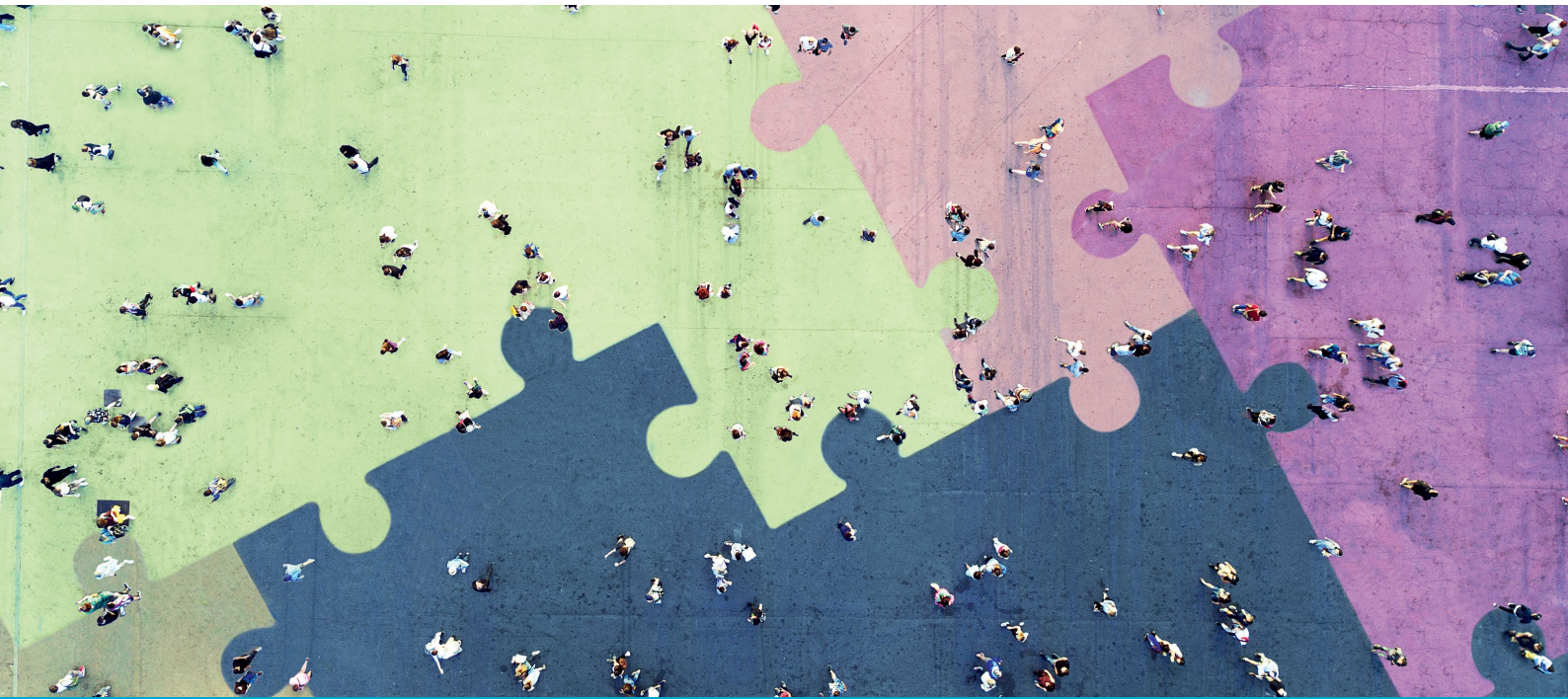
SYNTHÈSE ET ILLUSTRATIONS

Dynamique sectorielle	<ul style="list-style-type: none">- Le secteur de la recherche dirige de plus en plus son attention sur des données dites centrées sur l'individu (le patient) de la collecte à l'utilisation finale de la donnée.- Le secteur de la santé est marqué par une connaissance limitée du besoin précis en donnée (nature, typologie) en amont du projet de recherche. Cela implique que l'option pour des démarches très prospectives est difficilement compatible avec l'annonce d'une finalité précise en début de projet.
Dynamique générale	<ul style="list-style-type: none">- Les secteurs à composante technologique ont des besoins massifs en terme de données.- Les secteurs générateurs de données voient s'opérer des mouvements de consolidation (ex : Health Data Hubs).
Figures	<ul style="list-style-type: none">- L'acte citoyen : transmission de données d'un individu vers la communauté dans une finalité d'intérêt général. La constitution de ce modèle peut s'inscrire dans un programme politique.- Le commun : mutualisation des données des acteurs d'un même secteur.

⁴⁹ [Adopted EDPB-EDPS Joint Opinion 03/2021 on the Proposal for a regulation of the European Parliament and of the Council on European data governance \(Data Governance Act\)](#).

⁵⁰ [Milestone M8.4 Presentation of a first set of data altruism definitions, use cases and findings, Towards European Health Data Space, 2021.](#)

⁵¹ [BOULARD Audrey, FAVIER-BARON Eugène, WOILLET Simon, Le Health Data Hub ou le risque d'une santé marchandisée, 2020.](#)



FOCUS CAS SANTÉ LIMITE : CORONA DATA DONATION APP

Face à la propagation de la Covid-19 en Europe, l'Institut Robert Koch a mis en place le projet Corona-data-donation-app en avril 2020, à l'échelle du territoire allemand. La population est ainsi invitée à partager ses données de santé, notamment les symptômes liés à la contraction du virus. L'anonymat des contributeurs est préservé grâce au recours à des pseudonymes, sans pour autant compromettre la précision des données collectées et associées à chaque usager. En ce sens, ce projet est une démonstration de la compatibilité de la démarche data altruiste avec le RGPD.

Cependant, un an après la mise en ligne de l'application, cette dernière comptait moins d'un million d'utilisateurs, sur une population nationale de 83 millions, en dépit des perspectives d'accélération de la recherche médicale et d'amélioration des services de santé. Le Commissaire fédéral allemand à la protection et la liberté de l'information a expliqué ce revers par l'ambiguïté de la notion de « don de données » (*datenspende*) et le choix fait de la pseudonymisation, au lieu de l'anonymat total.

Pourtant, une étude, réalisée par Data4Life⁵² en avril 2020 sur l'acceptabilité sociale d'un tel projet semblait prometteuse :

- 73 % des Allemands se déclarent prêts à partager leurs données de santé afin de permettre d'améliorer la fiabilité des diagnostics ;
- Mais seulement 65 % dans le cadre du Covid 19 ;
- Avec une défiance particulière des citoyens résidant en Allemagne de l'Est (58 %) et des foyers avec des enfants (63 %).

L'exemple de Corona-data-donation-app est donc symptomatique des difficultés rencontrées à mettre en pratique le data altruisme, en dépit de la conscience des bénéfices qui en découlent, et souligne le haut niveau de volontarisme requis de la part des participants.

⁵² L'étude a été réalisée sur un échantillon de 5 002 participants représentatifs des différents corps sociaux (Corona data donation survey | Data4Life).

DYNAMIQUES DANS LE SECTEUR DE LA MOBILITÉ

La *smart city*, le phénomène de numérisation du tissu urbain et de ses flux (transport, économique, etc.), est un concept qui a largement occupé le débat autour de la mobilité. On peut toutefois en relever les injonctions contradictoires⁵³ :

- Une territorialisation de la notion et un mille-feuille administratif national qui compliquent l'émergence d'un modèle et la coordination des politiques publiques en matière de données ;
- Une personnalisation poussée des services de la ville, tout en affirmant assurer le respect de la vie privée ;
- Une optimisation de la gestion des données pour répondre à des problèmes pratiques, tout en cherchant à en capter le plus possible sans finalité précisément définie ;
- Une opposition entre les nouveaux entrants sur le marché de la donnée urbaine, notamment les acteurs privés pas toujours soumis aux règles des marchés publics et les acteurs traditionnels notamment publics.

Pour ces raisons, le secteur de la mobilité apparaît comme particulièrement propice à l'émergence de modèles data altruistes.

UNE ÉVOLUTION DES MODES DE MOBILITÉ GÉNÉRATEURS ET CONSOMMATEURS DE DONNÉES

La mobilité, notamment urbaine, connaît des changements importants, alors que l'élargissement des infrastructures se heurte à un double obstacle budgétaire et technique. La concentration des populations dans les métropoles et l'augmentation des déplacements pendulaires provoquent des congestions que les autorités de transports s'efforcent de réduire. Une gestion plus fine des déplacements des personnes, comme des marchandises, les trajets « porte-à-porte » pour réduire les frictions de l'intermodalité, l'efficacité de la logistique notamment dans « le dernier kilomètre » sont au centre des préoccupations des acteurs de la mobilité.

L'émergence de grands services numériques (Uber, Waze, Airbnb) a généré un impact fort sur le tissu urbain⁵⁴ grâce à une maîtrise fine de la donnée. Ces nouveaux acteurs se sont fortement développés en s'appuyant sur le système de l'auto-entrepreneuriat peu protecteur pour les employés, tout en conservant une autorité hiérarchique semblable à celle du salariat. S'est alors posée alors la question de la reprise de contrôle de la donnée par les pouvoirs publics. Jusqu'alors, la seule alternative envisagée à la propriété privée de la donnée était l'open data. Apparue récemment dans un rapport parlementaire⁵⁵, la notion de donnée d'intérêt territorial a opéré un changement de paradigme : comment penser la transmission de données générées dans le cadre privé vers la sphère publique, pour des finalités d'intérêt général ? De nouveaux modes de régulation de la donnée urbaine ont donc été imaginés afin d'assurer le respect des individus et de leurs libertés, s'appuyant sur les règles de gouvernance des communs.

À l'image du secteur de la santé, les acteurs de la mobilité rencontrent des difficultés quant à la qualité des données. L'entreprise [Jungle Bus](#) évoque ainsi un changement de doctrine au sein de la communauté *open data* sur la mobilité. Depuis que Tim Berners-Lee⁵⁶ a demandé en 2012 « *raw data now!*⁵⁷ », la société civile demande aux pouvoirs publics de publier en *open data* les données publiques. Aujourd'hui, la norme est de publier des jeux de données de qualité et de structuration très hétérogènes. La compilation de ces données n'est possible que par un acteur disposant de ressources et de compétences importantes. De fait, les grands acteurs du numérique sont donc favorisés dans l'exploitation de données *open data*. La co-création de communs numériques avec les pouvoirs publics et société civile pourrait permettre de sortir de cette impasse.

Le cas Jungle Bus montre l'importance de la dimension sociologique et communicationnelle dans la mise en place d'un flux de données qualitatif. L'entreprise ainsi que l'une de ses principales plus-values est sa capacité à faire le lien entre la communauté bénévole (souvent experte) et les professionnels des collectivités territoriales. Ces deux mondes sont marqués par d'importantes différences (horaires de travail, culture, outils, etc.). Cette interface entre la mobilisation des ressources bénévoles d'un côté et les besoins spécifiques d'acteurs est un créneau que les organisations data altruiste pourraient être amenées à occuper.

⁵³ La plateforme d'une ville », Cahier IP Innovations & Prospective N°05, CNIL - LINC, 2017.

⁵⁴ Modification des flux de déplacement, des règles d'aménagement, hausse des prix du foncier, etc.

⁵⁵ BELOT Luc, *De la « smart city » au territoire d'intelligence(s) - L'avenir de la « smart city »*, 2017.

⁵⁶ Informaticien britannique, principal inventeur du World Wide Web.

⁵⁷ Publiez les données brutes dès maintenant.

UN SECTEUR PARTICULIÈREMENT SUJET AU PHÉNOMÈNE DE LA PORTABILITÉ CITOYENNE

Le secteur de la mobilité est un terreau particulièrement favorable à l'émergence d'une portabilité citoyenne des données, certes encore limitée. Elle a une parenté avec la démarche altruiste promue par le DGA, puisqu'elle est basée sur une initiative volontaire et vise l'intérêt général.

LA QUESTION DE LA PORTABILITÉ CITOYENNE

Introduit par le RGPD, ce mécanisme permet à un acteur de demander la transmission des données générées dans le cadre de son utilisation d'un acteur vers un autre. L'objectif est de favoriser la concurrence en évitant les effets de *lock in*, d'enfermement dans un système.

Ce concept s'est enrichi pour générer la portabilité citoyenne : des individus demandent la portabilité de leurs données vers une organisation servant l'intérêt général. Ces organisations réceptrices de données peuvent être gouvernées par des communs⁵⁸.

Le rapport Villani⁵⁹ précise que la constitution de tels modèles de portabilité présente trois intérêts pour la puissance publique :

- La création de nouvelles bases de données pour le service public ;
- Une meilleure circulation des données sous le contrôle exclusif des citoyens ;
- Ils opèrent à droit constant, ce qui n'impose pas de nouvelles contraintes aux acteurs privés.

L'exercice plein et entier de ce droit nécessite toutefois un outillage en amont, notamment via des personal information management systems.

Des projets tels que [Tracemob](#) proposent aux individus de suivre automatiquement leurs déplacements en détectant le mode de transport utilisé via une application installée sur *smartphone* et développée en *open-source*. Cette application permet à l'utilisateur de visualiser ses déplacements et de calculer son empreinte carbone, étant précisé que les données (brutes et traitées) ne sont accessibles que par l'utilisateur et font l'objet d'un effacement rapide (30 jours). Outre cela, Tracemob permet à l'utilisateur de transférer ses données de déplacement moyennées et anonymisées à la collectivité dans laquelle il habite pour aider à l'amélioration des services de transport et de l'infrastructure publique.

Dans une approche similaire, STRAVA METRO utilise les données agrégées et anonymisées des utilisateurs de STRAVA (application de suivi GPS de course à pied ou vélo) pour aider les planificateurs des transports, les administrations municipales et les acteurs du domaine de la mobilité à améliorer la conception et la localisation des infrastructures dédiées. Il s'agit là d'une démarche

de STRAVA METRO, et non de l'individu. On ne peut donc pas parler de portabilité citoyenne. Néanmoins, la transmission des données est réalisée de manière altruiste et dans une finalité d'intérêt général.

Les données de mobilité peuvent être sensibles elles aussi. Cela implique une attention particulière des acteurs dans leur gestion et exploitation. C'est le cas du projet [WeFlo Afrique](#) de Jungle Bus qui collecte les traces GPS d'individus via une application et cartographie ainsi des réseaux de bus. Ces données de déplacement individuel sont agrégées avant d'être transmises pour analyse. De même, le projet Commute au sein de la Métropole de Toulouse collecte des données RH des grandes entreprises de la zone aéronautique⁶⁰ pour améliorer les flux de déplacement. Ces données sont donc particulièrement sensibles. Pour pouvoir être traitées, elles sont partitionnées⁶¹, c'est-à-dire rassemblées au sein de catégories plus larges ne permettant pas les réidentification des individus à l'origine des données.

⁵⁸ CNIL - LINC, 2017.

⁵⁹ VILLANI Cédric, *Donner un sens à l'Intelligence artificielle : pour une stratégie nationale et européenne*, Conseil national du numérique, 2018.

⁶⁰ Lieu de production industrielle rassemblant de grands acteurs du secteur : Airbus, Safran, etc.

⁶¹ Procédé également nommé *data clustering*.

SYNTHÈSE ET ILLUSTRATIONS

Dynamique sectorielle	<ul style="list-style-type: none">- Les acteurs de la mobilité ont besoin de données territoriales à un niveau relativement fin et actualisé pour répondre aux défis de la congestion urbaine et de la décarbonation des transports.- Le phénomène de la portabilité citoyenne, notamment à l'échelon municipal, se développe notamment en réaction aux dynamiques de numérisation marchande de la ville (<i>smart city</i>).
Dynamique générale	De nouveaux acteurs qualifiés de tangentiels peuvent, par le partage de données, contribuer à un secteur auquel ils ne sont pas directement rattachés. Les jeux de données sont perméables d'un secteur à l'autre (ex : Strava).
Figures	<ul style="list-style-type: none">- Le pas de côté altruiste : transmission de données d'un acteur économique vers un acteur non concurrent dans une finalité d'intérêt général.- L'acte citoyen : transmission de données d'un individu vers la communauté dans une finalité d'intérêt général. La constitution de ce modèle peut s'inscrire dans un programme politique.

FOCUS CAS MOBILITÉ EMBLÉMATIQUE : WORKER INFO EXCHANGE

Worker Info Exchange est une organisation à but non lucratif britannique qui a pour objectif de réduire l'asymétrie d'information entre des auto-entrepreneurs et les plateformes de mise en relation, telles qu'Uber et Deliveroo. Elle favorise ainsi la transparence en matière de conditions salariales et lutte contre l'opacité et la discrimination induites par les systèmes de gestion automatisés. L'enjeu est d'aboutir à une amélioration des conditions de travail grâce à une stratégie en trois temps :

- La récolte et la facilitation de l'accès aux données générées par les différents livreurs et chauffeurs travaillant pour ces plateformes, afin d'établir la répartition de leur temps de travail, le calcul de leur rémunération et de leur performance ;
- La réalisation d'enquête et l'analyse des données, permettant aux chauffeurs/livreurs de contester les licenciements automatiques illégaux et les décisions arbitraires de la direction ;
- L'ouverture de la base de données aux syndicats et aux organisations de chauffeurs/livreurs pour appuyer les négociations et les actions collectives.

En mettant en commun les données des travailleurs, Worker Info Exchange est un exemple de confiance des données avec une finalité d'acte citoyen visant à améliorer les conditions de travail de chauffeurs et livreurs dont le statut d'auto-entrepreneur est ambigu au regard de la relation contractuelle qui les lie aux plateformes.

FOCUS CAS MOBILITÉ : UBER MOVEMENT

Cette plateforme créée par Uber en 2017 rassemble des données en matière de mobilité, calculées à partir des temps de trajet et des parcours effectués par les conducteurs de VTC. Ce projet est symptomatique de l'émergence du big data urbain soutenant le développement des *smart cities*. Il ouvre la voie à une nouvelle forme de collaboration entre les collectivités territoriales et des multinationales, proposant une rénovation du mode de gouvernance urbaine visant à l'amélioration des services.

Les données mises à disposition par les chauffeurs permettent aux organismes en charge de l'urbanisme de mesurer la congestion du trafic, l'efficacité des politiques publiques, de mettre à l'épreuve les futurs projets et investissements à l'étude, ou encore de réduire la vulnérabilité des cyclistes et des piétons lors de leurs déplacements. Les chercheurs universitaires bénéficient également des données collectées par Uber pour améliorer leurs modèles de recherches et mener des simulations plus précises.

Uber Movement propose une vision du don altruiste de données pour améliorer la mobilité, cependant entachée par la volonté sous-jacente d'améliorer une image de marque dégradée, en partageant gratuitement des données déjà exploitées commercialement.

DYNAMIQUES DANS LE SECTEUR DE L'ENVIRONNEMENT

Le secteur de l'environnement a la particularité d'être transversal, tant par les finalités d'intérêt général qui en découlent que par les données qui peuvent l'alimenter. Dans son rapport⁶², le Conseil national du numérique distingue ainsi les données environnementales par nature, directement liées à l'environnement (données géographiques), et les données environnementales par destination (mobilité, consommation d'eau ou d'énergie), qui peuvent servir la transition écologique lorsqu'elles sont croisées avec d'autres données. Deux thèmes en particulier ont été étudiés au sein de ce vaste ensemble : l'énergie et la biodiversité.

LE CAS DE L'ÉNERGIE

Les mix énergétiques sont propres à chaque pays. Toutefois, les politiques de réduction des énergies fossiles et de développement des énergies renouvelables provoquent des évolutions substantielles dans la manière dont l'énergie est pensée dans les pays de l'UE. Une transition d'un modèle de production centralisé basé sur des moyens pilotables (centrale à gaz, charbon ou nucléaire) à un modèle décentralisé reposant sur la multiplication de moyens de productions non pilotables (éolien, solaire) s'est amorcée de manière plus ou moins soutenue partout sur le continent. Dans le présent rapport, l'utilisation de l'expression « secteur de l'énergie » renvoie principalement à la production et la distribution d'électricité, même si les autres énergies ne sont pas étrangères aux problématiques émergentes en matière de partage de données.

Le besoin en données, déjà fort, se fait de plus en plus pressant pour les acteurs du monde de l'énergie. La rencontre entre la demande et l'offre d'énergie s'opère désormais à un niveau de plus en plus fin. Si les acteurs importants (producteurs ou distributeurs) sont déjà largement impliqués dans des démarches d'innovation autour du partage de la donnée, l'échelon local (métropole, département) prend désormais une place prépondérante. L'audition du projet [RUDI](#) à Rennes,

qui comprend certaines activités liées à la distribution d'énergie, a permis de mettre en avant un certain nombre de problématiques concrètes rencontrées par ces acteurs. Le développement de réseaux de distribution connectés, ou *smart grids*, nécessite ainsi d'obtenir une donnée actualisée en temps réel. Dans une autre perspective de partage de données énergétiques respectant la confidentialité, les îlots morphologiques urbains apparaissent comme une maille pertinente, mais nécessitant un travail autour de la gouvernance de la donnée⁶³.

Alors que les problématiques liées au changement climatique se développent, la question du partage et de l'appropriation de ces données au sein d'une population large se pose. La connaissance portant sur la manière dont on produit et consomme l'énergie serait un moyen pour les citoyens d'être plus impliqués dans des démarches de sobriété et d'économie d'énergie. Toutefois, le projet RUDI montre qu'associer des habitants, des personnes non-expertes à un débat sur les données elles-mêmes, leur utilisation, leur production, leur partage reste complexe.

⁶² Feuille de route sur l'environnement et le numérique – 50 mesures pour un agenda national et européen sur un numérique responsable, c'est-à-dire sobre et au service de la transition écologique et solidaire et des objectifs de développement durable, Conseil national du numérique, 2020.

⁶³ Échelle intermédiaire entre la parcelle cadastrale et la commune qui permet d'agréger, structurer et diffuser certaines informations. [Story Map Journal](#).

LE CAS DE LA BIODIVERSITÉ

La sauvegarde de la biodiversité constitue une autre thématique d'intérêt général pouvant tirer parti du data altruisme. Dans ce domaine, un dispositif de recherche est particulièrement susceptible de participer à des modèles data altruistes : le tiers secteur de la recherche. Il se définit comme regroupant « les activités de production de savoirs du secteur non marchand (associations, syndicats, collectivités territoriales, etc.), du secteur marchand à but non lucratif (économie sociale et solidaire, groupements professionnels, etc.) et des organisations à but lucratif de petite taille (auto-entrepreneur, groupements agricoles ou artisanaux, etc.)⁶⁴».

Ce tiers secteur de la recherche apparaît comme une interface de coopérations entre le monde académique et les autres contributeurs aux thématiques d'intérêt général. Si cette approche n'est pas limitée au domaine de la biodiversité, ce secteur nécessitant la collecte de données sur des surfaces et temporalités relativement longues s'y prête particulièrement. Le développement de l'information géographique volontaire, la collecte de données par un ensemble d'acteurs (souvent non-scientifiques) est un autre exemple d'application du tiers secteur de la recherche.



SYNTHÈSE ET ILLUSTRATIONS

Dynamique sectorielle	<p>Biodiversité : le tiers secteur de la recherche se développe en appui de thème de recherche (ex : Fishbrain, NoisePlanet).</p> <p>Énergie : le pilotage des activités dans un cadre de développement des moyens de production non pilotables nécessite un besoin croissant de données.(Enedis Data connect).</p>
Dynamique générale	<ul style="list-style-type: none">- Extension du domaine scientifique avec le développement du tiers secteur de la recherche.- Développement d'une démocratie technique, notamment via la constitution d'une culture de l'énergie.
Figures	<ul style="list-style-type: none">- Le collecteur : l'individu fournit des données qu'il collecte au cours de son cheminement.- Le pas de côté altruiste : transmission de données vers un acteur étranger à mon secteur dans un but d'optimisation de leur valorisation.- Le commun : mutualisation des données des grands acteurs d'un même secteur.

⁶⁴ Assises du tiers secteur de la recherche.

FOCUS CAS ENVIRONNEMENT EMBLÉMATIQUE : *THE WORLD BEE PROJECT*

Fondée par l'humanitaire britannique Sabiha Rumani Malik, The World Bee Project est une société d'intérêt général (CIC – Community Interest Company) engagée en faveur de la préservation des abeilles. Le projet a noué de nombreux partenariats de recherche avec des universités, à l'instar de Reading. Il fournit, sans contrepartie, de nombreuses données relatives à l'environnement et au comportement des abeilles, en s'appuyant sur des techniques d'IA et d'IoT élaborées par Oracle Cloud, et des capteurs fournis par Bee Hero. Le projet repose sur un réseau de 50 000 ruches dites intelligentes réparties dans le monde. The World Bee Project offre ainsi gratuitement aux chercheurs les données récoltées en partenariat avec des entreprises du secteur technologique.

FOCUS CAS ENVIRONNEMENT : LE COMPTEUR ÉLECTRIQUE LINKY

Le dispositif des compteurs « intelligents » Linky répond à une initiative européenne inscrite dans le Clean Energy Package (aussi connue sous le nom de directive 2009/72/CE), transposée, dans la législation française, par la loi du 17 août 2015 relative à la transition énergétique pour la croissance verte. L'installation de ces compteurs électriques de nouvelle génération a pour objectif d'aboutir à un marché européen de l'électricité intégré servant de pierre angulaire à la transition énergétique du continent. Le meilleur suivi de la consommation électrique permis par les compteurs Linky doit inciter les usagers à diminuer leur consommation d'énergie et apporter davantage de flexibilité à la demande, tout en favorisant le recours aux énergies renouvelables⁶⁵.

Ce dispositif doit améliorer la qualité du service pour les clients : tarification horaire, possibilité de d'éteindre ou d'allumer à distance des appareils électroménagers afin d'éviter les pics de consommation, introduction de nouveaux moyens d'autoproduction comme les panneaux photovoltaïques, facturation sur données réelles, raccourcissement des délais d'intervention, etc.

Malgré les externalités positives annoncées, le Baromètre Energie-infos publié en 2016 indique que 32 % des Français sont défavorables à l'installation de cet équipement. Les motifs d'opposition invoqués sont d'une grande diversité : 9 % craignent les ondes émises par l'équipement et un potentiel effet néfaste sur la santé, 10 % s'inquiètent du risque pour les données personnelles, d'autres encore doutent des nouveaux services proposés et de leur impact annoncé sur la transition énergétique. De plus, certaines options des compteurs Linky ont aussi pu offrir une vision répressive et de surveillance, notamment le relèvement automatique de la consommation et la limitation, voire la coupure, du service à distance, suscitant l'inquiétude quant à la protection du consommateur.

Cette agrégation d'arguments hétéroclites souligne que l'absence de finalité publiquement identifiée a rendu le processus d'adoption plus difficile et a constitué l'un des principaux écueils du projet Linky. En outre, les difficultés rencontrées par ce projet illustrent la dimension clé du consentement. Il a notamment pu être reproché que le déploiement des compteurs nouvelle génération dans l'Union européenne ne s'appuie pas sur le choix éclairé des consommateurs. En effet, le client est réputé accepter par défaut l'installation de cet équipement. Si ce n'est pas le cas, il doit lui-même manifester son opposition auprès du gestionnaire de réseau de distribution. Ce choix « passif » a ainsi pu entacher la confiance déjà fragile des consommateurs.

⁶⁵ La flexibilité de la demande devrait principalement prendre la forme d'une tarification horaire déterminée à partir des données de consommation.

FIDUCIE DES DONNÉES : UNE PERSPECTIVE CANADIENNE

La gouvernance des données est devenue un enjeu majeur. Diverses solutions sont proposées pour protéger les droits et libertés des citoyens tout en favorisant le développement de nouvelles technologies bénéfiques. La fiducie de données représente une solution potentielle à ces enjeux. Au Canada, elle y a connu un terreau favorable.

GOVERNEMENT CANADIEN

Une année après la publication de l'influent rapport *Growing the Artificial Intelligence Industry in the UK* (2017) qui a popularisé la notion de fiducie de données, le Gouvernement du Canada a signifié son intérêt pour ce concept dans son Rapport des Tables de stratégies économiques du Canada publié en septembre 2018. Fruit de discussions entre près d'une centaine de PDG d'entreprises et de fonctionnaires fédéraux, ce rapport portait notamment sur les industries numériques, identifiées comme un des six secteurs-clés pour la croissance future du Canada⁶⁶. Le rapport soulignait notamment l'importance des mégadonnées (« les marchandises les plus lucratives de la nouvelle économie mondiale⁶⁷ ») et de l'IA qu'elles alimentent, tout en soulevant les enjeux de protection des données accompagnant ces technologies⁶⁸. Pour maintenir la confiance des Canadiens, le rapport suggérait de mettre à jour les lois canadiennes, mais aussi de développer une stratégie pancanadienne en matière d'IA incluant le financement de « fiducie de données ». À notre connaissance, il s'agit de la première mention officielle du concept de « fiducie de données » au Canada, défini alors comme « un organisme formé pour gérer des données pour le compte de ses membres. Ces derniers mettent en commun leurs données et conviennent expressément des conditions afférentes à leur partage ». Le concept n'était cependant pas plus développé et la stratégie ne semble pas avoir mené à la création de fiducie de données.

L'idée d'encourager la création de fiducies de données apparaît également brièvement dans le rapport La Charte numérique du Canada en action : un plan par des Canadiens, pour les Canadiens, publié en 2019, par le ministère de l'Innovation, des Sciences et du Développement économique (le « MISDE »)⁶⁹. De vastes consultations publiques menées de juin à octobre 2018 sont à l'origine de cette Charte qui vise à favoriser la confiance des Canadiens dans le monde numérique et encourager la croissance économique. Comme le Rapport des Tables de stratégies économiques, la Charte numérique elle-même ne mentionne les fiducies de données qu'au passage. Cependant, dans le cadre des actions de mise en œuvre de la Charte, le Gouvernement canadien soutient le collectif canadien de normalisation

en matière de données (le « Collectif ») créé par le Conseil canadien des normes pour accélérer le développement de normes de gouvernance des données. Le Collectif a publié une feuille de route en juin 2021 listant 35 enjeux d'importance en matière de gouvernance des données devant être étudiés. Ceux-ci incluent la fiabilité des intermédiaires du traitement des données, dont les fiducies de données, et le Collectif souhaite explorer les principes de gouvernance applicables à ces entités, tout particulièrement les normes d'accréditation et d'attestation qui pourraient s'appliquer⁷⁰.

Peu après le dévoilement de la Charte numérique, le MISDE a publié ces Propositions pour moderniser la loi sur la protection des renseignements personnels et des documents électroniques, soit la loi fédérale encadrant la protection de la vie privée (« PIPEDA »). Selon le MISDE, « la confiance est le pilier de l'économie numérique et axée sur les données⁷¹ ». Parmi ses propositions de réforme, le MISDE suggérait de modifier la loi pour ajouter une exception à l'obligation d'obtenir un consentement (un des principes fondamentaux de la loi) lorsque des données désidentifiées seraient traitées par une fiducie de données, soit une tierce partie fiable avec des obligations de gestion responsable des données et encadrée par un régime juridique spécifique.

Le législateur n'a cependant pas donné suite à ces recommandations puisque le Projet de loi C-11, introduit en novembre 2018 au parlement Canadien pour remplacer PIPEDA, ne contient aucune disposition spécifique sur les fiducies de données. À la place, le Projet de loi C-11 permettrait à une organisation de communiquer des renseignements personnels sans consentement si (1) ces renseignements sont dépersonnalisés, (2) la communication est faite notamment à une institution gouvernementale, une entité mandatée par une loi fédérale ou provinciale ou toute autre entité réglementaire et (3) la communication est faite pour une fin socialement bénéfique⁷². Le Projet de loi C-11 est devenu lettre morte, vu les élections fédérales qui ont eu lieu en 2020 avant son adoption.

⁶⁶ Les cinq autres secteurs étant : la fabrication de pointe, le secteur agroalimentaire, les technologies propres, le secteur des sciences biologiques et de la santé ainsi que les ressources de l'avenir. L'impératif de l'innovation et de la compétitivité - Tables sectorielles de stratégies économiques.

⁶⁷ Rapport des Tables de stratégies économiques du Canada : Industries numériques, à la p. 12.

⁶⁸ Pour une discussion plus approfondie des enjeux de vie privée soulevés par l'IA, voir le chapitre Privacy du rapport Responsible AI : A Global privacy framework publié par ITechLaw en 2019 et mis à jour en 2021.

⁶⁹ La Charte numérique du Canada en action : un plan par des Canadiens, pour les Canadiens, à la p. 24.

⁷⁰ Feuille de route du Collectif canadien de normalisation en matière de gouvernance des données, à la p. 64.

⁷¹ Renforcer la protection de la vie privée dans l'ère numérique - Innover pour un meilleur Canada.

⁷² Projet de loi C-11.

En revanche, au Québec, un régime semblable a été formellement mis de l'avant avec l'adoption du Projet de loi 64 en septembre 2021, qui a modernisé le régime de protection des renseignements personnels au Québec. Celui-ci prévoit des exceptions à l'obligation de consentement applicable à la communication des données personnelles lorsque ces données sont utilisées à des fins de recherche dans l'intérêt public. De plus, le Projet de loi 19 (aussi du Québec) portant sur la protection des renseignements de santé contient également une exception permettant aux chercheurs d'utiliser les renseignements personnels recueillis par les organismes de santé publique et de services sociaux à des fins autres que celles pour lesquelles un consentement a été obtenu, à condition qu'une évaluation éthique soit effectuée pour le projet de recherche. En somme, au lieu de soutenir spécifiquement la création de fiducies de

données ou d'intermédiaire fiable en droit canadien, ces dispositions facilitent l'utilisation des données personnelles pour des fins « altruistes » à travers un mécanisme d'exception au fardeau du consentement.

En définitive, malgré l'intérêt porté à l'idée de favoriser la création de fiducies de données, le concept n'a pas encore de support dans les lois canadiennes et il n'y a pas d'équivalent aux intermédiaires de données proposés par la Commission européenne⁷³. Ce silence crée certaines ambiguïtés relativement à l'application du concept de fiducie de données au Canada, y compris les responsabilités juridiques qu'impliquent ces entités et leur interaction avec les lois sur la protection de la vie privée. Comme le montrent l'exemple suivant, cela complexifie la tâche de ceux qui souhaitent créer de telles entités au Canada.

EXEMPLE : L'URBAN DATA TRUST DE SIDEWALK LABS

C'est au Canada que Sidewalk Labs, une société affiliée à Alphabet inc., a proposé le projet de fiducies de données le plus connu, le Urban Data Trust (le « Trust »), dans le cadre de sa participation au projet de revitalisation urbaine du district Quayside, à Toronto⁷⁴. Waterfront Toronto, organisme responsable du développement du district, proposait un partenariat avec Sidewalk Labs pour créer une ville intelligente où l'installation de différents capteurs permettraient de collecter une vaste gamme de données sur les activités urbaines, incluant des renseignements personnels. Le Trust devait prendre la forme d'une organisation indépendante à but non lucratif qui aurait approuvé l'utilisation des données collectées dans le district Quayside, avec un niveau de surveillance plus serré en fonction des enjeux de protection des renseignements personnels liés à certaines catégories de données. Le Trust devait approuver au préalable les contrats conclus avec des entreprises privées voulant collecter et utiliser ces données et ainsi servir d'instance de gouvernance des données liées au projet⁷⁵.

Toutefois, Sidewalk Labs n'est pas parvenue à convaincre Waterfront Toronto que son Urban Data Trust était un modèle de gouvernance des données en lequel ils pouvaient avoir confiance. En effet, dès l'annonce de la participation de Sidewalk Labs dans le projet Quayside, de nombreux Torontois ont fait entendre leurs inquiétudes. Les difficultés rencontrées par Sidewalk Labs proviennent en partie des spécificités du projet Quayside puisque celui-ci concernait des données collectées directement dans l'espace public. Mais il semble y avoir eu aussi un malaise directement lié à l'implication de Google (via sa filiale) dans le projet. Dans une lettre d'opinion publiée par le Toronto Star en novembre 2019, Mark Surman, directeur général de la Fondation Mozilla, résumait le sentiment d'une partie de l'opinion publique comme suit : « *Comme beaucoup,*

*j'ai craint dès le début que l'approbation de l'entente avec Sidewalk ne signifie transmettre toutes nos données à Google*⁷⁶ ».

Également, le Trust devait être établi par des moyens de droit privé, ce qui a compliqué son intégration dans les régimes municipaux et fédéraux de protection des données existantes. Combiné avec l'aspect novateur du concept et les ambiguïtés dans la structure, la gouvernance et le fonctionnement du Trust, l'absence de cadre légal clair entourant les fiducies de données a fait en sorte que Sidewalk Labs n'a pas pu bâtir la confiance nécessaire et faire aboutir son projet. En novembre 2019, Waterfront Toronto a décidé d'abandonner l'idée de créer le Trust pour plutôt favoriser un contrôle des données par les instances gouvernementales. Finalement, en mai 2020, Sidewalk Labs a annoncé son retrait du projet, expliquant que la pandémie rendait le projet économiquement infaisable⁷⁷.

Une des principales leçons à tirer de l'échec de Sidewalk Labs est que les fiducies de données évoluent au travers d'enjeux sensibles, susceptibles de faire l'objet de débats publics. Une fiducie de données ne soulève pas uniquement des enjeux technologiques et de sécurité, car le modèle vise à bâtir une confiance réciproque entre les participants.

La fiducie de donnée reste ainsi une idée embryonnaire au Canada et les tentatives concrètes de créer de telles entités n'ont pas rencontré, jusqu'à présent, les succès escomptés. Clarifier le cadre réglementaire et légal entourant les fiducies de données, tout particulièrement leurs responsabilités eu égard aux lois sur la protection de la vie privée, faciliterait l'adoption de la fiducie de données comme mode de gouvernance fiable des renseignements personnels.

⁷³ Commission européenne Bruxelles, Le 25.11.2020 Com(2020) 767 Final 2020/0340 (Cod) Proposition de Règlement du Parlement Européen.

⁷⁴ La Charte numérique du Canada en action : un plan par des Canadiens, pour les Canadiens, à la p. 11.

⁷⁵ Digital Governance Proposals for DSAP Consultation, à la p. 12-16.

⁷⁶ The Sidewalk Labs Dividend: Can this debate usher in a new era of digital urban planning? | The Star.

⁷⁷ Why we're no longer pursuing the Quayside project — and what's next for Sidewalk Labs | by Daniel L. Doctoroff | Sidewalk Talk | Medium.

FIGURES INSPIRANTES ET RISQUES DE DÉRIVES DU DATA ALTRUISME

La recension de cas pratiques et la conduite d'auditions auprès de praticiens ont permis d'identifier un ensemble de besoins en matière de données, de conditions nécessaires au succès de modèles de partage de données actuels ou futurs. La synthèse de ces éléments s'incarne dans des figures, des positions dans lesquelles un acteur (individu, entreprise, organisation) pourrait s'identifier dans la définition de sa démarche altruiste. Chaque figure comprend une description générale, un intérêt de l'adopter, ses limites et des exemples pouvant s'en approcher.

FIGURES

LE PAS DE CÔTÉ ALTRUISTE

L'un des principaux obstacles au partage de données par des acteurs économiques réside dans la crainte de divulguer des connaissances à des concurrents ou de leur permettre de les reconstituer et de prendre ainsi l'avantage. Si la démarche altruiste est suspectée de permettre cela, le tiers de confiance perd son objet.

Le processus mis en place gagnerait donc à favoriser le partage de données entre acteurs non concurrentiels au sein d'un même secteur d'activité. Cette disposition n'est pas aisée à mettre en œuvre, tant la concurrence peut prendre des formes diverses, mais elle peut passer par une sélection stricte de la typologie des utilisateurs de données et des projets dans le cadre de la démarche altruiste.

En outre, une grande partie de la valeur des données réside dans une fertilisation croisée entre secteurs d'activité. Par exemple, l'utilisation de données de mobilité ou d'habitat peut avoir du sens dans le cadre de la recherche médicale. Ainsi, **une promotion des échanges intersectoriels rendrait l'altruisme de la donnée plus fructueux.**

Un acteur d'un secteur transmet ses données vers un acteur ou un ensemble d'acteurs non concurrents.	Intérêt(s)
	Cette approche permet de contourner la problématique concurrentielle inhérente au partage de données.
	Limite(s)
	Il est difficile d'identifier des secteurs dans lesquels l'usage des données transmises serait pertinent. Cela implique d'avoir une connaissance fine des spécificités de chaque secteur. Le risque de voir la concurrence exploiter les données utilisées dans un autre secteur est non nul.
	Potentialités
	Développement d'autres applications concrètes liées aux données générées, peuvent être cités les cas : <ul style="list-style-type: none">- De l'imagerie spatiale- Des producteurs de contenus (presse, culture) pour l'entraînement d'algorithmes- Des plateformes de services à forte fréquentation pour affiner les plans de mobilités
Exemples	
Strava, Fishbrain, BigScience	



L'ACTE CITOYEN

En sa qualité de citoyen, l'individu fait preuve d'une propension à partager ses données dans l'intérêt de la communauté dont il fait partie. C'est particulièrement vrai au niveau local, mais cela demeure transposable à des projets de plus grande ampleur.

La confiance est, en effet, générée par la démarche participative et le sentiment d'agir directement dans un objectif déterminé. Partant, **l'organisation de systèmes replaçant le citoyen au cœur d'un programme de collecte de données dans un intérêt général nous apparaît comme étant un facteur essentiel.**

En outre, le data altruisme soulève des craintes légitimes en matière de protection des droits fondamentaux. **Des corps intermédiaires, notamment au niveau des collectivités locales, peuvent être une interface d'information et de renforcement de la capacité des individus à faire valoir leurs droits dans le cadre du data altruisme** (droit de retrait, droit à l'oubli, droit d'information, etc.).

<p>Un individu transmet ses données à la communauté, directement ou via un tiers de confiance. Par là même, il s'engage dans un projet qui vise à servir une cause d'intérêt général.</p>	<p>Intérêt(s)</p>
	<ul style="list-style-type: none"> - Renforcement du rôle du citoyen. - Permet de créer des groupes citoyens et de confier à un tiers la préservation de leurs droits. - Appuis des pouvoirs publics et/ou du milieu associatif comme tiers de confiance ou financeur.
	<p>Limite(s)</p>
	<ul style="list-style-type: none"> - Un tel modèle nécessite un haut niveau d'engagement de la part des individus, voire une démarche militante. - Ce niveau d'engagement peut générer une réduction de la population participante, et donc nuire à la représentativité des données fournies.
	<p>Potentialités</p>
	<p>Développement de communs citoyens autour de problématiques plus ou moins larges, notamment via des coopératives.</p>
<p>Exemples</p>	
<p>BitsAboutMe, CitizenMe, DataForGood Foundation, Ethi.me.</p>	



LA COLLECTE DE DONNÉES SCIENTIFIQUES PAR LES PARTICULIERS

L'expansion et la miniaturisation des dispositifs de captations de données (*smartphone*, montre connectée) permet d'étendre et de démocratiser les programmes de recherche. Un double mouvement est venu renforcer ces possibilités de génération de données : la numérisation de la sphère des loisirs et la ludification de certaines collectes scientifiques. Dans certains cadres, et avec un protocole de recherche adapté et contrôlé par une autorité compétente (institut de recherche), les individus et entreprises peuvent ainsi collecter des données concomitamment à leur activité initiale : déplacements pendulaires, activités de loisir, tournées de sociétés de service, etc.

<p>Un individu fournit des données qu'il collecte au cours de son cheminement de manière régulière ou ponctuelle.</p>	<p>Intérêt(s)</p>
	<ul style="list-style-type: none"> - La génération de données peut être effectuée en même temps qu'une activité autre (randonnée, tournée de facteur) - Cette approche ajoute une dimension ludique au data altruisme.
	<p>Limite(s)</p>
	<ul style="list-style-type: none"> - Pour que les données soient qualitatives, doivent être collectées, soit avec une dimension experte (nombre important de données et à un niveau fin), soit par une masse critique d'utilisateur. - Le caractère ludique peut aisément s'estomper : <ul style="list-style-type: none"> · Dans le cadre d'une activité de loisir ponctuelle (lassitude). Des associations (sportives, culturelles) organisées peuvent être un relais pertinent. · Dans le cadre d'une activité professionnelle, si elle est imposée et/ou parasite les actions de l'employé).
	<p>Potentialités</p>
<ul style="list-style-type: none"> - Extension de cette dimension de la collecte à des applications plus grand public. - Développement auprès de réseaux d'acteurs professionnels hiérarchisés et répandus sur le territoire (facteur, opérateur téléphonique, etc.). 	
<p>Exemples</p>	
<p>Noise Planet, Pollutrack, Jungle Bus, Aubudon, Cros, EDD maps, Montre Verte, Ushahidi, PlantNet.</p>	

LE COMMUN

Si l'altruisme peut renvoyer à une question d'éthique individuelle⁷⁸ le transfert de données, notamment personnelles, est souvent l'occasion de créer une organisation collective. Parmi ces projets qui ambitionnent de gérer collectivement les données, la notion de commun est régulièrement convoquée. Ils se définissent comme une ressource numérique avec une gouvernance dédiée et la capacité à en faire évoluer collectivement les règles tout en respectant le droit des personnes. Empruntant à la fois aux logiques de gouvernance des biens publics et privés, ils ne sont pas organisés par une puissance publique, ni ne sont la propriété exclusive d'un groupe d'individus. Le commun a ainsi pour spécificité « *que la communauté qui veille sur lui a, non pas des droits, mais une autorité particulière dans sa gouvernance*⁷⁹ ». Les principes pour la gestion des communs ont été définis dans la littérature⁸⁰, et certains sont particulièrement pertinents dans le cadre du data altruisme :

- Définition claire de l'objet de la communauté et de ses membres⁸¹ ;
- Assurer l'adaptation des règles opérationnelles dans le temps et accès rapide à des instances locales de résolution de conflits⁸² ;
- Reconnaissance du commun par les autorités gouvernementales externes.

Ils se différencient de l'*open data* et ils n'impliquent pas une gratuité des services.

<p>Des acteurs mutualisent leurs données à travers une gouvernance où chaque participant a la possibilité d'en modifier les règles. Le commun peut être pensé dans une logique sectorielle ou transverse.</p>	<p>Intérêt(s)</p>
	<ul style="list-style-type: none"> - La taille du gisement de données à traiter augmente. - Le modèle est résilient car il peut évoluer et prévoit des instances de résolution des conflits.
	<p>Limite(s)</p>
	<ul style="list-style-type: none"> - Le commun nécessite un niveau d'organisation relativement important. - Il implique d'accepter le fait que l'organisation est souple et sujette à des modifications de la part de ses adhérents. Cette flexibilité peut être un frein pour des organisations structurées comme des entreprises.
	<p>Potentialités</p>
	<p>Constitution de fiducies sectorielles.</p>
<p>Exemples</p>	
<p>SalusCoop, Worker Info Exchange, ODRé.</p>	

⁷⁸ À noter que lors de leur audition messieurs DOUEIHI et DOMENICUCCI sont revenus sur les origines de l'utilisation de l'altruisme dans le champ économique et en particulier de l'économie numérique et sur celle de commun. Le commun est avant tout une réponse à un problème de gestion collective, alors que l'altruisme est plus une question individuelle : dans quelle mesure suis-je prêt à sacrifier mon intérêt privé pour le bénéfice d'autres ?

⁷⁹ CARDON Dominique, *Culture numérique*, Presses de Sciences Po, 2019.

⁸⁰ OSTROM Elinor, *Governing the Commons: The Evolution of Institutions for Collective Action*, De Boeck, 2010.

⁸¹ Cf. recommandations sur la création de l'organisation data altruiste.

⁸² Cf. recommandations sur la gouvernance de l'organisation data altruiste.



LES RISQUES

LA RÉPARATION

Le data altruisme est une activité qui peut bénéficier à l'image d'un acteur. Cela en est une des incitations à participer, notamment pour les entreprises, afin de contrebalancer le coût (réel et d'opportunité) généré par la mise à disposition de données de manière altruiste. Le degré d'implication d'une entreprise reste variable et basé sur le volontariat, allant de l'investissement fort et durable à l'action mineure et ponctuelle. De même, libre à elles de fixer le niveau de communication qu'elles souhaitent mettre en place.

La réparation est une tentation puissante pour les sociétés dont l'activité peut être particulièrement dommageable ou pointée du doigt pour ses dérives dans les médias. En mettant à disposition les données générées par son activité l'acteur fait taire les critiques et redore son image auprès du public.

À l'image des actions menées dans le développement durable, le data altruisme pourrait être sujet à du *greenwashing*, appelé aussi *bluwashing*, pour les actions relatives à l'éthique. Des entreprises ayant mené des actions particulièrement néfastes pourraient adopter une démarche data altruiste *a minima* pour camoufler les dommages qu'elles ont créés en s'appuyant sur une organisation *ad hoc* selon les règles établies dans le DGA.

LE CHEVAL DE TROIE

Comme exposé dans la PARTIE 1, l'instrument juridique peut être utilisé dans le cadre de la compétition internationale pour amener des entreprises à dévoiler leurs secrets industriels, ou à payer de fortes amendes. Le data altruisme, dans la mesure où il cherche à partager des données, qui, peuvent être sensibles, représente un risque pour le contributeur. Des acteurs malveillants pourraient mettre en place des mécanismes qui, sous couvert de data altruisme, viseraient la captation de données stratégiques d'un concurrent. Ces mécanismes peuvent être directs (piratage du serveur d'un tiers de confiance) ou indirect (hameçonnage en utilisant la couverture d'un projet data altruiste).

Le Cheval de Troie est un stratagème qui pourrait être utilisé par des acteurs malveillants pour s'introduire chez des concurrents, à l'insu de ces derniers, pour les piéger, sous couvert de participer à une initiative altruiste. Ces agissements, relevant de l'espionnage, soulèvent notamment les problématiques de la sécurité économique et de l'anonymisation des données partagées. En effet, les informations rassemblées visent à réduire l'asymétrie de l'information et à conquérir les parts de marchés détenues par le concurrent ciblé. Dans la mesure où les modes de partage de données d'entreprises nécessitent une importante formalisation en amont, le risque de captation indirecte reste faible. Néanmoins, des tiers de confiance moins expérimentés ou dont le processus d'anonymisation n'est pas assez rigoureux pourraient faire l'objet d'atteintes malveillantes.

LA TROMPERIE

Alors que le projet DGA a marqué une étape importante de son développement avec la clôture du trilogue européenne, le Bureau européen des unions de consommateurs rappelait son inquiétude sur « *la façon dont une définition floue de l'altruisme [...] pourrait permettre aux entreprises d'abuser de motifs vagues et altruistes pour pousser les consommateurs à partager leurs données*⁸³ ». Le projet COVI APP du [Mila](#) (Institut québécois d'intelligence artificielle) détaillait dans son livre blanc l'ensemble des possibilités de hameçonnage susceptibles d'advenir suite au déploiement de l'application⁸⁴. À l'image du Cheval de Troie pour les entreprises, l'anti-figure de la tromperie pourrait être une des formes néfastes prises par le data altruisme afin de capter des données personnelles sous couvert d'une démarche data altruiste.

⁸³ Data governance: new EU law for data-sharing adopted, Euratic, 1^{er} décembre 2021.

⁸⁴ Accès à un faux lien URL lors de la mise en œuvre de l'application, installation d'une fausse application, faux messages texte de diagnostic COVI, divulgation de renseignements protégés provenant d'une application.

SYNTHÈSE

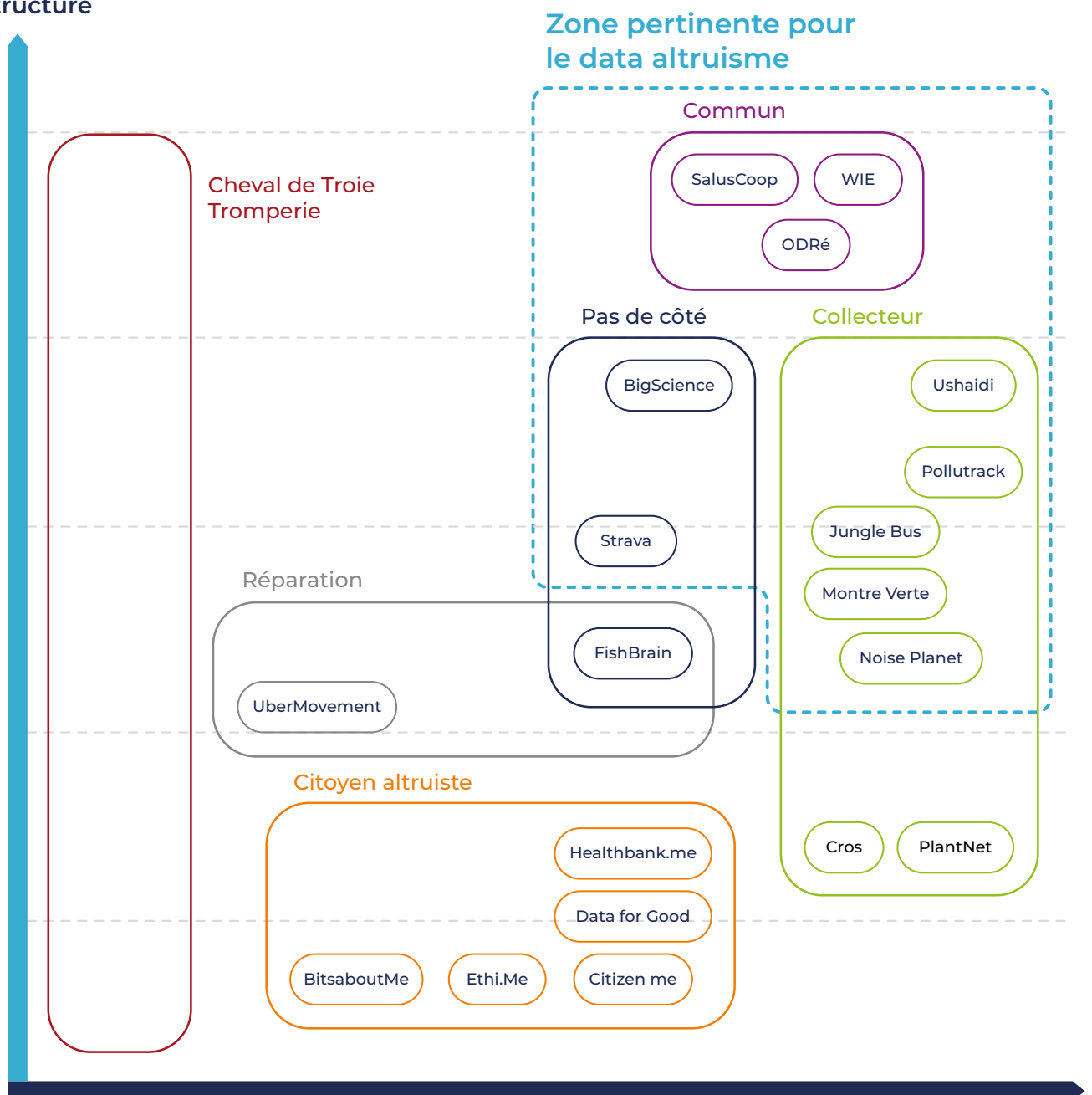
Rassembler ces figures et risques de dérives au sein d'un même graphique permet d'envisager deux questions essentielles avant la mise en place de modèles data altruistes :

- La question de **la finalité** du modèle : son objectif concerne-t-il une population vaste ou réduite ? Les potentiels bénéfiques qui en découleront seront-ils importants ?
- La question de **la complexité** du modèle : cette approche nécessite-t-elle un haut niveau de sophistication ? Une gouvernance multipartite ? Une planification à long terme ? Un investissement matériel et temporel conséquent ?

Envisager ces deux questions à travers les figures présentées ci-dessous pourrait permettre d'orienter les acteurs qui souhaiteraient s'impliquer dans un modèle data altruiste.

MATRICE DES CAS PRATIQUES ÉTUDIÉS

-/+ structuré



-/+ collectif

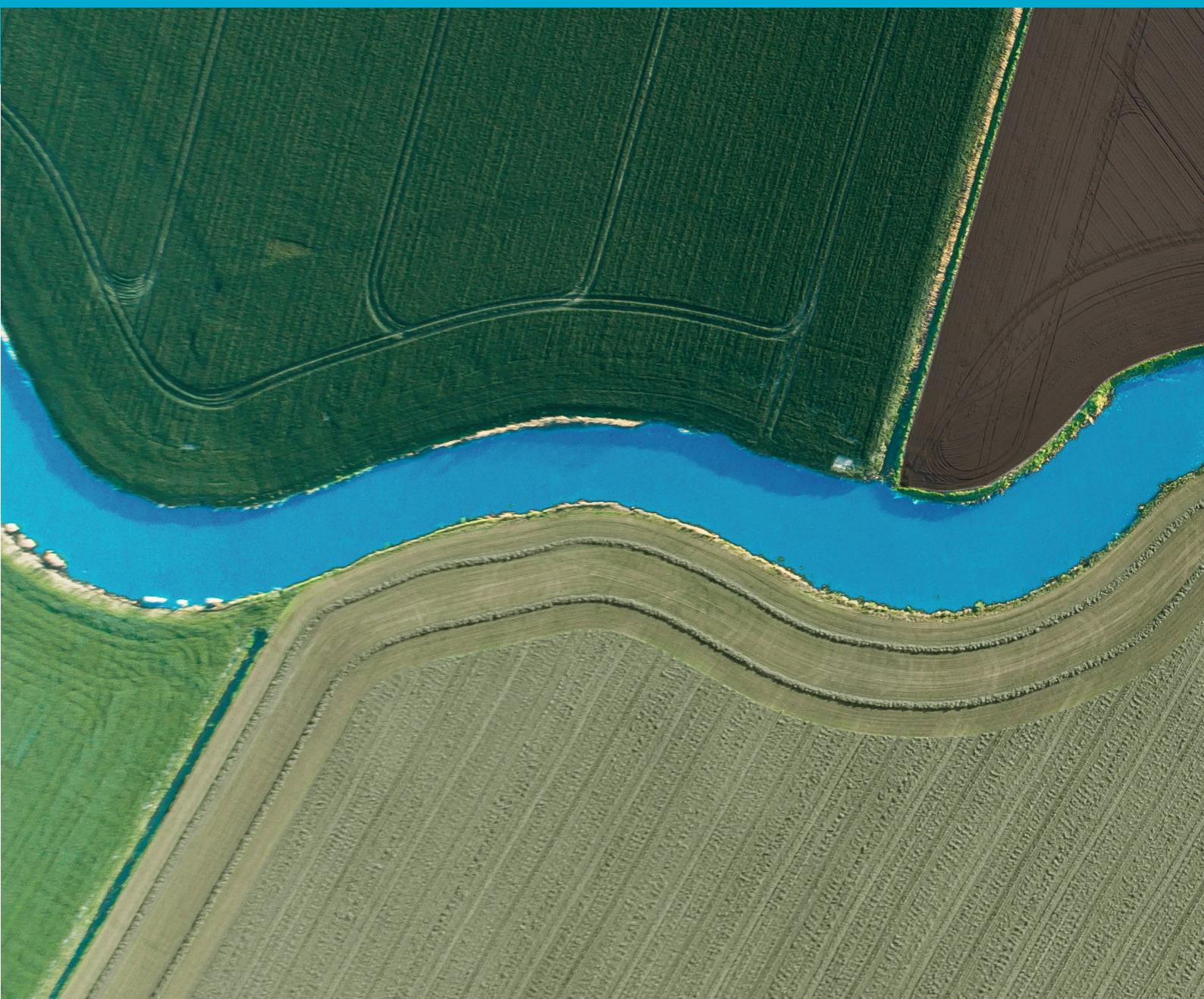
ÉLÉMENTS CLEFS PARTIE 2

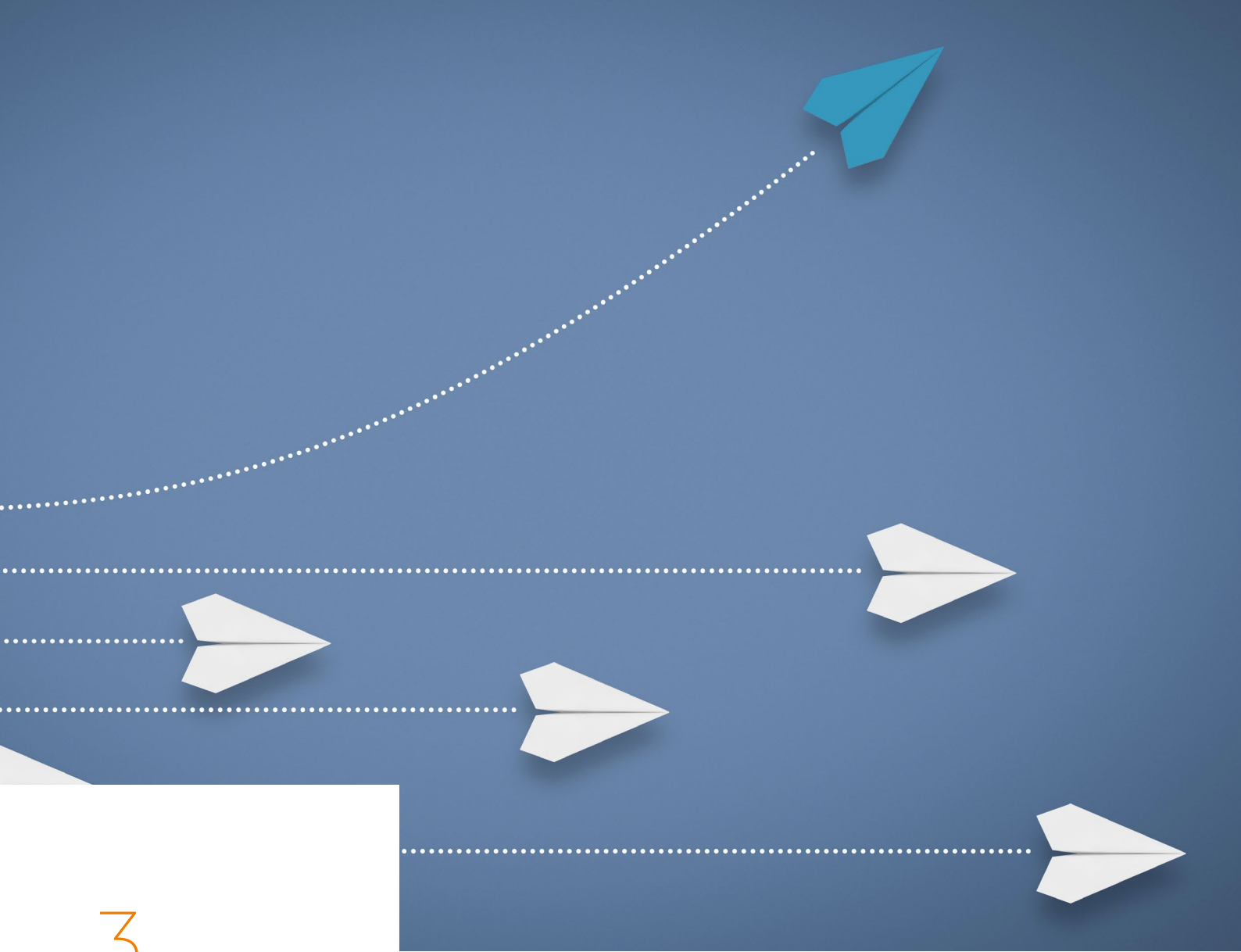
Les modes de partage novateurs de données nous montrent qu'il est possible d'imaginer plusieurs formes de data altruisme. Ces figures sont des orientations que des individus ou les entreprises pourraient prendre : le pas de côté, l'acte citoyen, le collecteur, le commun. Cependant, le data altruisme peut conduire à certains comportements non vertueux qu'il convient de combattre : la réparation, le Cheval de Troie et la tromperie.

De telles attitudes ne peuvent être considérées comme des actions data altruistes. Elles dévoient la finalité d'intérêt général énoncée dans le DGA et ouvrent une brèche dans la confiance à la base du dispositif.

La vision globale de ces figures et risques de dérives nous montre que :

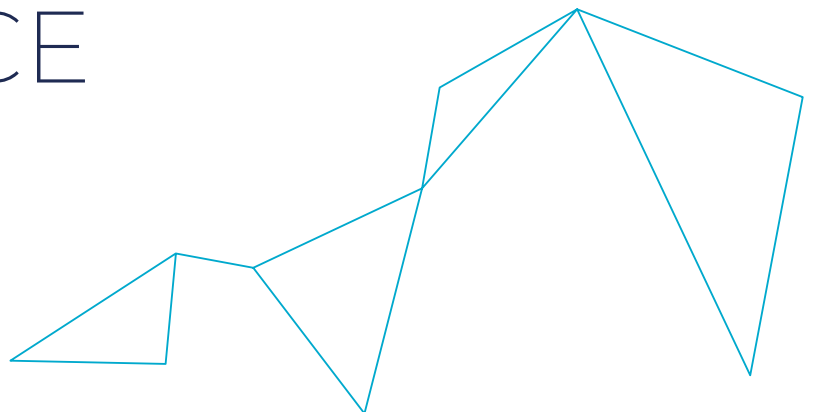
- Les projets de partage de données prennent des formes très diverses ;
- Les modèles data altruistes gagneraient à se développer au sein des modèles les plus ambitieux en termes d'impacts ;
- Cela nécessite de prévoir des niveaux d'organisation relativement complexes.





3.

VADE-MECUM POUR UN DATA ALTRUISME EFFICACE DANS LA CONFIANCE



Dans le cadre des travaux menés, nous avons adopté une démarche de recherche itérative reposant sur des études de cas d'exploitation de données. Nous avons, à cet effet, développé une grille d'analyse standardisée, permettant de comparer les différents projets et de faire émerger des dynamiques. Cette méthode repose sur cinq axes :

- L'identification du besoin initial des données, des contraintes techniques et des pratiques sectorielles qui influencent la conception et l'émergence du projet ;
- Le mode de financement initial (subvention, financement participatif) et le modèle économique retenu, ainsi que la finalité affichée ;
- La structuration juridique et la gouvernance de l'organisation, ses éléments de conformité ou non avec le cadre du DGA ;
- Les principes visant à instaurer une relation de confiance avec les contributeurs de données ;
- Les incitations mises en œuvre pour faciliter le partage des données.

Ce cadre d'enquête définit également une ligne de vie d'un projet data altruiste : les moments clés qu'il doit traverser. À chaque phase, sont exposés les questionnements des différents acteurs, les principaux enseignements issus des acteurs auditionnés, les meilleures pratiques rencontrées, ainsi que des recommandations en vue de l'établissement de modèles data altruistes.

ÉTAPE N°1 CRÉER UNE ORGANISATION DATA ALTRUISTE

Pour diminuer la défiance et, corrélativement, augmenter la confiance, il est nécessaire d'inclure *by design* des mesures de nature à susciter la confiance en l'organisation altruiste. Les organisations altruistes doivent afficher une transparence optimisée afin de mettre à la disposition des contributeurs potentiels, puis aux contributeurs effectifs, les informations nécessaires et proportionnées à leur implication. Cela implique d'identifier lors de la création de l'organisation altruiste les informations devant être ouvertes, mais également les règles d'accès aux dites informations.

ACTION N°1	DÉFINITION DES FINALITÉS D'INTÉRÊT GÉNÉRAL DE L'ORGANISATION DATA ALTRUISTE
Objectif	
Susciter la confiance en faisant adhérer les contributeurs potentiels au projet porté par l'organisation altruiste.	
Recommandation	
L'organisation data altruiste doit définir les finalités qu'elle entend promouvoir, d'une manière précise pour être comprise, mais aussi d'une manière extensible pour permettre l'exploitation des données pour des cas d'usages non envisagés de manière prospective.	
Mise en œuvre⁸⁵	
<ul style="list-style-type: none"> - Définir le domaine d'intérêt général dans lequel l'organisation entend s'inscrire. Ex : poursuite d'un objectif de développement durable des Nations Unies (écologie, santé, éducation, etc.). - Déterminer une modalité d'action précise Ex : lutte contre les émissions de CO2, recherche de traitements contre le cancer. 	

⁸⁵ Référence complémentaire : avis conjoint du Comité européen de la protection des données et du contrôleur européen de la protection des données (CEPD) concernant l'acte sur la gouvernance des données.

ACTION N°2

CRÉATION D'UNE CHARTE ÉTHIQUE DE L'ORGANISATION

Objectif

Fournir aux contributeurs et utilisateurs, ainsi qu'aux organes de gouvernance, un référentiel éthique et pratique sur l'utilisation des données et l'acceptation des projets.

Recommandation

L'organisation doit publier une charte éthique contraignante définissant les qualités requises des utilisateurs de données et les critères d'accès aux données. Cette charte peut réitérer des obligations légales ou réglementaires en leur donnant une force contractuelle entre les parties, ce qui en facilite la sanction directe en cas de non-respect.

Mise en œuvre

La charte peut étendre aux données non personnelles, sous la forme d'obligations contractuelles, des obligations prévues par le RGPD pour les données personnelles, ainsi que d'autres principes adaptés à l'organisation et aux données qu'elle traite. À titre d'illustrations :

- **Principe de bienfaisance** : le moyen mis en place doit avoir, ou à tout le moins viser, une portée bénéfique et, *a contrario*, ne doit pas contribuer directement à faire subir un dommage à une personne physique ou à dégrader sa situation. Ce principe s'applique aux moyens mis en place par l'utilisateur pour exploiter les données et à l'organisation pour les collecter.
- **Principe de transparence** : fournir une information claire, accessible, explicable et auditable depuis l'acquisition des données jusqu'à la fin de leur cycle de vie.
- **Principe de proportionnalité** : la collecte de données doit être adaptée à la finalité recherchée.
- **Principe de protection** : les choix et actes opérés doivent tenir compte d'un impératif de sécurité et de protection des données.
- **Principe de responsabilité** : l'organisation doit assumer la responsabilité de ses choix et réalisations.
- **Principe d'interopérabilité et de portabilité des données** : les choix technologiques de l'organisation favorisent l'échange et l'utilisation des données.
- **Principes de finalité et de loyauté** : l'utilisateur ne doit accéder aux données que pour le projet présenté et accepté par l'organisation.

ACTION N°3

PUBLICATION D'INFORMATIONS SUR LA STRUCTURE INTERNE DE L'ORGANISATION

Objectif

Limiter la défiance liée à l'opacité de la structuration de l'organisation.

Recommandation

Publication d'informations sur la structure interne de l'organisation.

Mise en œuvre

L'organisation peut mettre à la disposition des contributeurs potentiels et des utilisateurs des informations dont la publication n'est pas obligatoire, qui sont aisées à récolter et publier et qui sont de nature à instaurer un climat de confiance.

Nous préconisons notamment la publication de l'organigramme des différents organes de gouvernance de l'organisation.

**ACTION
N°4****PUBLICATION DES INFORMATIONS TRANSMISES
ANNUELLEMENT À L'AUTORITÉ NATIONALE DE CONTRÔLE****Objectif**

Susciter la confiance en instaurant une transparence *by design*.

Recommandation

Rendre publiques les informations transmises annuellement à l'autorité nationale de contrôle.

Mise en œuvre

Partage des informations dont le DGA prévoit qu'elles doivent être transmises annuellement à l'autorité nationale de contrôle :

- Les statuts de l'organisation (article 17 DGA) : disponibles librement pour les contributeurs potentiels ;
- Une description des finalités d'intérêt général pour lesquelles des données sont collectées par l'organisation, et des moyens employés pour les atteindre (article 18 DGA) : disponible librement pour les contributeurs potentiels ;
- Une liste des utilisateurs qui ont eu accès aux données qu'elle détient, assortie d'une description sommaire des finalités d'intérêt général poursuivies par cette utilisation de données (voir développements ci-après sur la précision de la description des finalités) et de la description des moyens techniques employés en vue de cette utilisation, y compris une description des techniques appliquées pour préserver la vie privée et la protection des données (article 18 DGA) : disponible sur demande aux contributeurs potentiels démontrant une capacité contributive réelle - des règles plus contraignantes peuvent être mises en place par la gouvernance de l'organisation notamment en anonymisant la liste et en préservant les informations devant rester confidentielles ;
- Une synthèse des résultats des utilisations de données autorisées par l'organisation, s'il y a lieu (article 18 DGA) : disponible sur demande aux contributeurs potentiels démontrant une capacité contributive réelle ;
- Des informations de synthèse sur les sources de recettes de l'organisation, en particulier toutes les recettes générées par le fait d'avoir autorisé l'accès aux données, et sur les dépenses (article 18 DGA) : disponible librement pour les contributeurs potentiels.

**ACTION
N°5****FAVORISER LES CONDITIONS DE L'ÉMERGENCE
D'ORGANISATIONS DATA ALTRUISTES AU NIVEAU EUROPÉEN****Objectif**

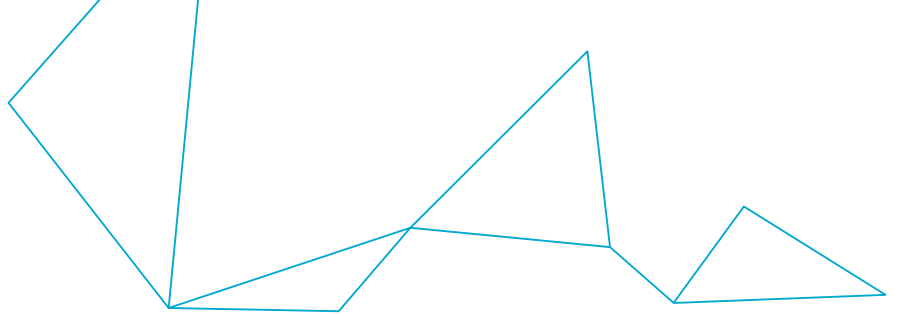
Initier le processus dans les premiers mois suivant le vote du DGA et partager sur les usages et les projets afin de permettre la répliquabilité.

Recommandation

Mener des actions concrètes pour solliciter des acteurs susceptibles de participer à des modèles data altruistes, améliorer la communication autour de l'initiative.

Mise en œuvre

- Répertoire au sein d'un catalogue en sources ouvertes les initiatives proches du data altruisme déjà actives ;
- Mettre en place un groupe public privé sous l'égide de la Commission européenne pour rassembler les acteurs susceptibles de participer à un modèle data altruiste ;
- Organiser une conférence annuelle afin d'acculturer les parties prenantes au data altruisme ;
- Préparer l'organisation de la fédération des organisations data altruiste (cf. action N°9) ;
- Prévoir dans les financements Digital Europe Programme/Horizon Europe un cadre spécifique au data altruisme ;
- Organiser des hackathons pour identifier les solutions techniques à même de servir des modèles data altruistes ;
- Créer un site internet centralisant les informations pertinentes sur le data altruisme.



ÉTAPE N°2

DÉFINIR LA GOUVERNANCE DE L'ORGANISATION DATA ALTRUISTE

Le modèle de gouvernance doit traduire les mesures mises en œuvre pour assurer l'atteinte des objectifs de l'organisation tout en offrant des garanties solides aux contributeurs quant aux modalités d'exploitation des données par les utilisateurs. Il apparaît qu'un modèle de gouvernance adapté devrait scinder les fonctions de gestion de données et celles de gestion administratives et financières de l'organisation.

ACTION N°6	GOUVERNANCE MULTIPARTITE DE L'ORGANISATION DATA ALTRUISTE
Objectif	
S'assurer que l'organisation data altruiste est pérenne et que son caractère altruiste est garanti par une séparation des fonctions au sein de la gouvernance.	
Recommandation	
Application du principe de spécialisation des organes de gouvernance par périmètre d'activité.	
Mise en œuvre	
<p>L'organisation data altruiste doit répartir ses compétences entre différents rôles :</p> <ul style="list-style-type: none">• La gestion administrative et financière de l'organisation :<ul style="list-style-type: none">- Organe de gestion administrative et opérationnelle de l'organisation, comprenant le Président, la Présidente, ou mandataire chargé de représenter l'organisation vis-à-vis des tiers. Cet organe doit être chargé de la définition et de l'instauration de la compliance juridique relative à la collecte et à l'exploitation de données.- Organe de gestion comptable et financière.• La gouvernance des données :<ul style="list-style-type: none">- Organe de gouvernance des données, mettant en œuvre les mesures édictées par l'organe de gestion administrative :<ul style="list-style-type: none">> Suivi et contrôle de la compliance juridique sur la collecte et l'exploitation des données ;> Gestion et suivi des mesures techniques de protection des données collectées ;> Gestion des relations avec les contributeurs (information, exercice des droits sur les données, gestion des différends) ;> Élaboration et mise à jour des critères d'accès aux données par les utilisateurs ;> Mise en œuvre d'analyses d'impact relatives à la protection des données (AIPD) ;> Instruction des demandes d'accès aux données conformément aux critères posés. <p>Il incombe à l'organe de gouvernance chargé de la gestion des données d'élaborer une méthodologie d'évaluation de l'exploitation des données par les utilisateurs. Cette méthodologie doit tenir compte de la nature et de la sensibilité des données.</p>	

**ACTION
N°7****GARANTIR LA COMPÉTENCE ET L'AUTONOMIE
DE L'ORGANE DE GOUVERNANCE DES DONNÉES****Objectif**

S'assurer que l'organe de gouvernance des données au sein de l'organisation data altruiste est structuré de manière à pouvoir réaliser un contrôle effectif et indépendant.

Recommandation

Organisation de la gouvernance selon les règles similaires à celles édictées par le RGPD applicables au Délégué à la protection des données (DPO)⁸⁶.

Mise en œuvre

L'organe de gouvernance en charge des données devrait être structuré selon les principes suivants :

- Avoir une expertise juridique et technique en matière de protection des données personnelles et non personnelles ;
- Avoir une bonne connaissance du secteur d'activité : finalités, nature des données collectées, sensibilité, contraintes sectorielles ;
- Avoir une parfaite connaissance de l'organisation interne, en particulier des opérations de traitements, des systèmes d'information, des besoins en matière de protection et de sécurité des données ;
- Disposer de moyens suffisants ;
- Pouvoir librement accéder aux informations utiles ;
- Être associé en amont des projets impliquant des données ;
- Pouvoir agir en toute indépendance, c'est à dire :
 - > Ne pas être en situation de conflit d'intérêt (impliquant une séparation de la gestion financière de l'organisation et une indépendance des membres de cet organe vis-à-vis des utilisateurs de données) ;
 - > Pouvoir rendre compte de son action au plus haut niveau de la direction de l'organisme ;
 - > Ne pas être sanctionné pour l'exercice de ses missions ;
 - > Ne pas recevoir d'instruction dans le cadre de l'exercice de ses missions.

L'instauration d'un filtre de sélection des utilisateurs pouvant avoir accès aux données, fut-elle réalisée sur la base de critères aussi objectifs que possibles, entraîne nécessairement des contestations tant de la part des utilisateurs que des contributeurs.

**ACTION
N°8****CRÉATION D'UNE INSTANCE DE RÉOLUTION
DES CONFLITS INTERNES****Objectif**

Permettre la gestion des différends entre les organes de gouvernance.

Recommandation

Prévoir la création d'une instance, qui peut être non permanente, chargée d'arbitrer les conflits internes portant, par exemple, sur l'autorisation d'un utilisateur.

⁸⁶ [Le Délégué à la protection des données \(DPO\) est chargé de mettre en œuvre la conformité au règlement européen sur la protection des données au sein de l'organisme qui l'a désigné s'agissant de l'ensemble des traitements mis en œuvre par cet organisme. Délégué à la protection des données \(DPO\) | CNIL.](#)

**ACTION
N°9**

**CRÉATION D'UNE FÉDÉRATION DES
ORGANISATIONS ALTRUISTES**

Objectif

Rassembler les organisations de l'altruisme des données et promouvoir leur activité.

Recommandation

Création par les organisations altruistes d'une fédération.

Mise en œuvre⁸⁷

- Création d'une fédération dont l'objet est de regrouper des organisations altruistes en matière de données afin de :
 - Mettre en commun des connaissances ;
 - Proposer des soutiens techniques ou administratifs aux organisations ;
 - Partager des coûts, par exemple liés à la mise en place de contrats d'assurances mutualisés ou à la création d'outils techniques communs ;
 - Réaliser la promotion de l'altruisme en matière de données ;
 - Travailler à l'évolution du cadre de l'altruisme des données ;
 - Éventuellement, avoir un rôle de certificateur de démarches altruistes (cf. marque collective de certification).
- La fédération pourrait également avoir un rôle dans le règlement de conflits entre organisations et utilisateurs ou contributeurs.
- La fédération devra veiller et mettre en œuvre des campagnes de sensibilisation contre les actes malveillants à l'encontre des contributeurs (arnaques par de fausses organisations altruistes).

⁸⁷ Référence complémentaire : Recommandation n°23 mission BOTHOREL : Créer un label de service producteur de la donnée pour reconnaître les efforts investis dans la donnée, par exemple dans le cadre du service public de la donnée.



ÉTAPE N°3

COLLECTER DES DONNÉES

La donnée a une valeur d'usage complexe à définir et hypothétique. Dans un rapport sur l'économie des données personnelles⁸⁸, le CIGREF rappelle qu'un traitement de ces données (lisibilité, format, etc.) est un prérequis avant toute utilisation génératrice de valeur. La qualité d'un jeu de données reste une notion floue, sujette à discussion. Il est possible de la décomposer en deux dimensions⁸⁹ : des attributs objectifs (complétude, précision, cohérence) et subjectifs (pertinence, crédibilité, accessibilité). Si, avec le data altruisme, il est envisageable d'espérer un afflux de données, il n'y a aucune garantie de leur « aptitude à l'emploi⁹⁰ ». Une donnée de faible qualité peut être traitée, mais nécessite un investissement en temps et compétences important, que des organisations altruistes en matière de données fraîchement constituées n'auront peut-être pas à disposition. Pour minorer une phase de retraitement coûteuse, il importe de mener une politique de données qualitatives en amont du processus.

ACTION N°10	CRÉATION D'UN CAHIER DES CHARGES SUR LA QUALITÉ ET L'EXPLOITABILITÉ DE LA DONNÉE
Objectif	
Obtenir et mettre à disposition des jeux de données dont l'aptitude à l'emploi est la plus élevée possible et éviter des coûts supplémentaires liés au retraitement.	
Recommandation	
L'organisation data altruiste doit s'assurer que les données qu'elle collecte et met à disposition des utilisateurs respectent un niveau suffisant de qualité. Pour cela, elle doit établir les règles qui définissent les standards de qualité des données qu'elle collecte et met à disposition. Ces règles doivent permettre de se conformer et maintenir les standards de qualité.	
Mise en œuvre⁹¹	
<p>L'organisation doit donc établir et publier un cahier des charges de la qualité des données. Cela implique :</p> <ul style="list-style-type: none"> - D'identifier les besoins qu'elle-même ou les potentiels utilisateurs de données en termes de qualité de données (format, fréquence de renouvellement, etc.) ; - D'analyser la capacité des contributeurs à fournir des données atteignant un certain standard ; - D'établir des indicateurs de niveau de qualité du jeu de données et des métadonnées ; - D'assurer une médiation et, lorsque cela est nécessaire et possible, de corriger l'écart entre les niveaux de qualité possibles et les attendus. <p>Dans une démarche de qualité, nous préconisons :</p> <ul style="list-style-type: none"> - Une recherche de diversité des sources de données ; - Une recherche de diversité des chargés de traitement des données pour éviter les biais (genre, âge, origine, formation etc.) ; - Une politique de recherche et d'identification des biais, puis de publication des biais identifiés, de leurs causes et des solutions de remédiation. <p>L'organisation peut soumettre l'obtention du droit d'utiliser une marque collective déposée par elle (cf. Action N°19) au respect de ce cahier des charges.</p>	

⁸⁸ Rapport CIGREF « Economie des données personnelles, 2015.

⁸⁹ ARRUBARRENA Béa, *Big data & éthique : la qualité des données en débat*. Cité dans BALICCO Laurence et al., *L'éthique en contexte info-communicationnel numérique*, pp. 23-38, Collection Information et stratégie, De Boeck Supérieur 2018.

⁹⁰ WANGRY., STRONG D.M., « Beyond Accuracy: What Data Quality Means to Data Consumer », *Journal of Management Information Systems*, vol. 12, n°4, 1996, pp. 5-33 cité in ARRUBARRENA, 2018.

⁹¹ Référence complémentaire : Recommandation n°25 Mission BOTHOREL : Encourager les écosystèmes à définir des principes de gouvernance de la qualité, en désignant un référent qualité et en créant des communautés de réutilisation avec participation active des producteurs de la donnée.



Le discours sur le data altruisme est principalement soutenu par deux éléments : la dimension symbolique de l'organisation (recherche publique, mutuelle, etc.) et le volontarisme, voire le charisme, affiché par les dirigeants d'organisation tenant lieu de tiers de confiance. Si la première composante est nécessaire, elle ne peut être suffisante. Cette dimension symbolique doit être accompagnée de mécanismes concrets permettant d'assurer dans les faits la réputation de l'organisation. Le second point est, lui aussi, moteur de la promotion de nouveaux modes de partage de la donnée. Toutefois, faire reposer le développement d'un projet altruiste sur la seule personnalité de sa direction ne peut être pérenne et nécessite là aussi de l'accompagner d'une praticité dans les bénéfices attendus de modèles data altruistes.

ACTION N°11	MISE EN PLACE D'UN DROIT DE RÉTRACTATION EFFECTIF
Objectif	
Permettre aux contributeurs, personnes physiques et morales, de retirer leur accord à l'utilisation des données transférées.	
Recommandation	
Créer un véritable droit de rétractation au profit des contributeurs en données.	
Mise en œuvre	
<p>L'exercice du droit de rétractation doit être :</p> <ul style="list-style-type: none"> - Simple à mettre en œuvre, via la gouvernance des données ; - Effectif en ce que les données doivent pouvoir être retirées : <ul style="list-style-type: none"> > A <i>minima</i> des données disponibles⁹² ; > Si possible des systèmes auxquels elles ont été intégrées. 	

⁹² La mesure est rendue obligatoire par le RGPD en cas de retrait de consentement à l'usage de données personnelles.

Objectif

Sécuriser l'obtention de l'accord à la collecte et encadrer traitement des données.

Recommandation

Pour les individus, nous recommandons l'utilisation de formulaires de consentement standardisés. Pour les personnes morales, nous recommandons le recours à des contrats de contributeur et d'utilisateur.

Mise en œuvre

Le DGA prévoit qu'un formulaire de consentement européen à l'altruisme en matière de données sera créé par la Commission.

Dans l'attente de la mise en place de cette mesure, l'organisation peut créer son propre formulaire destiné à :

- Informer le contributeur de la ou des finalités pour lesquelles la collecte est organisée (s'il y a plusieurs finalités, elles doivent être exposées distinctement et indépendamment) ;
- Informer le plus clairement possible le contributeur des modalités de mise à disposition des données transférées ;
- Informer le contributeur de son droit de retrait et des modalités d'exercice de ce droit ;
- Présenter les garanties offertes par l'organisation sur la sécurité des données ;
- Permettre l'expression univoque du consentement dans ce contexte, notamment en empêchant la mise en place de *dark patterns*⁹³ rendant le refus plus difficile que l'acceptation.

Nous préconisons l'inclusion des stipulations suivantes pour les contrats standardisés :

• **Contrat contributeurs :**

- Modalités de collecte des données ;
- Classification du niveau de sensibilité des données ;
- Mesures et durée de stockage et de protection adaptée desdites données ;
- Conditions dans lesquelles ces données seront rendues accessibles à des utilisateurs ;
- Droits du contributeur sur ces données une fois transférées à l'organisation :
 - > Retrait de l'autorisation d'exploitation ;
 - > Portabilité ;
 - > Droits de suivi sur le long terme des projets les utilisant ;
- Droits de l'utilisateur sur lesdites données ;
- Modalités de contrôle de l'utilisateur par l'organisation.

• **Contrat utilisateurs :**

- Leurs droits et obligations relatifs à l'exploitation des données ;
- La durée de conservation et le sort final des données après traitement ;
- Les conditions financières de l'accès aux données ;
- Les conditions d'exploitation des résultats issus de l'exploitation de ces données ;
- Les modalités de contrôle de leur activité par l'organisation ;
- Les potentielles sanctions applicables en cas de manquement ;
- Les droits du contributeur et notamment les conséquences de l'exercice d'un droit tel que le droit de retrait.

Il est essentiel que ces contrats standardisés remplissent les conditions suivantes :

- **Lisibilité et intelligibilité** : ils doivent être focalisés sur l'essentiel, présentés sous un format facilement navigable et être rédigés en des termes précis et facilement compréhensibles ;
- **Adaptabilité** : ils doivent pouvoir être négociés et adaptés dans la limite de la cohérence du modèle de l'organisation. Les contrats doivent également permettre au contributeur un contrôle suffisant sur ses données, tout en permettant à l'organisation une souplesse suffisante pour exploiter les données conformément aux finalités affichées ;
- **Accessibilité** : ils doivent pouvoir être consultés facilement par tout intéressé ;
- **Evolutivité** : le cadre contractuel doit être régulièrement réévalué et mis à jour pour faire face aux contraintes rencontrées dans la pratique et aux évolutions législatives.

⁹³ Un *dark pattern* est une interface utilisateur qui a été spécialement conçue pour tromper ou manipuler un utilisateur.

ENCOURAGEMENT AU PARTAGE PAR LES PERSONNES PHYSIQUES

Le grand public reste un acteur difficile à mobiliser. Il y a un arbitrage à faire entre le choix d'entraîner une population large dans le modèle, au risque de déperdition, et de se concentrer sur une population plus restreinte, experte et motivée, au risque de ne s'adresser qu'aux *happy few*.

ACTION N°13	CRÉER UN GUICHET D'ÉCHANGE AVEC LES INDIVIDUS
Objectif	
Renforcer la confiance en améliorant les échanges et l'accessibilité des informations pour les individus.	
Recommandation	
Mettre en place une interface d'information spécifique pour les contributeurs individuels, satisfaisant au devoir de conseil auprès des individus qui en feront la demande.	
Mise en œuvre	
Ce guichet peut être constitué d'une ligne téléphonique, d'une plateforme de chat ou d'un site web permettant de joindre la gouvernance dédiée aux données et offrant l'accès à : <ul style="list-style-type: none">- une aide technique ;- une aide juridique.	

ACTION N°14	CRÉER UNE INTERFACE DE COMMUNICATION ENTRE LES DIFFÉRENTES PARTIES PRENANTES D'UN MODÈLE DATA ALTRUISTE
Objectif	
Renforcer la confiance en renforçant les échanges et l'accessibilité des informations.	
Recommandation	
Mettre en place une interface de communication pour faciliter les échanges et consultation d'informations entre l'organisation, les contributeurs et utilisateurs.	
Mise en œuvre	
L'interface de communication peut prendre la forme d'une plateforme numérique permettant : <ul style="list-style-type: none">- Le suivi des programmes ayant eu accès à la base de données de l'organisation ;- La mise en avant des programmes (anonymisés ou non selon leur accord), s'étant vu accorder l'accès aux données assorties d'une présentation d'éléments pertinents ayant conduit à cet accord (bonne pratiques).- L'accès au catalogue des jeux de données (sans accès aux données) montrant de manière figurée et statistique l'importance des contributions et permettant de comprendre l'intérêt de la base de données.	

**ACTION
N°15****ORGANISER UNE CAMPAGNE DE SENSIBILISATION****Objectif**

Encourager le partage de la donnée.

Recommandation

Mettre en place des démarches proactives de sensibilisation des individus *in situ*.

Mise en œuvre

- Organisation d'une ou plusieurs conférences de consensus en amont du lancement d'un projet data altruiste, moment de réunion physique, présentant les objectifs du projet, répondant aux interrogations de l'assistance.
- Mise en place un programme d'ambassadorat dans le cadre d'une collecte de données personnelles.

**ACTION
N°16****UTILISER DES MODES DE COMMUNICATION LOYAUX****Objectif**

Conserver un haut niveau de confiance.

Recommandation

S'assurer que les éléments de proposition et de communication sont loyaux, complets et réalistes. Limiter le recours à tout artifice de communication de nature à biaiser le jugement des contributeurs et utilisateurs (rejet de *dark patterns* et de designs trompeurs)⁹⁴.

**ACTION
N°17****DOCTRINE D'ANONYMISATION AU SEIN
DU PROJET DATA ALTRUISTE****Objectif**

Sécuriser les données.

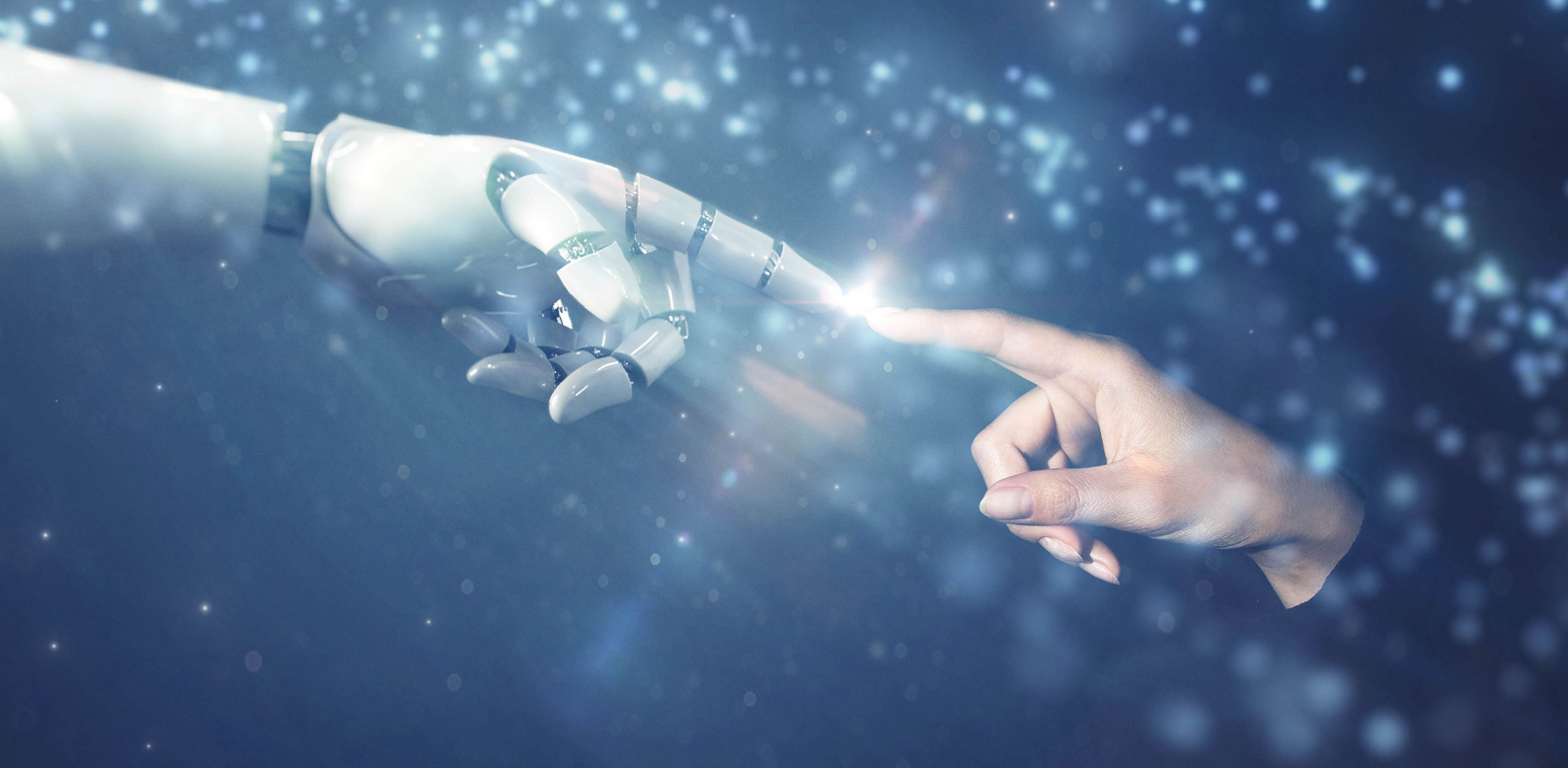
Recommandation

Mise en place d'une doctrine de pseudonymisation ou d'anonymisation des données publiées et consultables *a minima* pour les données personnelles.

Mise en œuvre

Mise en place par l'organisation d'un système de pseudonymisation ou d'anonymisation adapté à la nature des données et à leur sensibilité, garantissant la protection des données des individus tout en permettant une exploitation fructueuse desdites données.

⁹⁴ Directive 2005/29/CE du Parlement européen et du Conseil du 11 mai 2005 relative aux pratiques commerciales déloyales des entreprises vis-à-vis des consommateurs dans le marché intérieur.



ENCOURAGEMENT AU PARTAGE PAR LES PERSONNES MORALES

L'altruisme des données dépend directement :

- de la mobilisation des contributeurs en données ;
- de l'appropriation du système par les utilisateurs.

Le DGA prévoit que le contributeur altruiste doit donner son autorisation à l'exploitation de ses données sans demander de contrepartie. Il est néanmoins permis de mettre en place un système incitatif optimisant l'élan contributif ainsi que l'intérêt des utilisateurs pour le système altruiste. L'appartenance des acteurs de l'altruisme à un réseau concret et organisé et la possibilité de s'identifier comme altruiste en matière de données par l'usage de signes distinctifs tels que les marques collectives ou marques de certification sont des mesures considérées comme incitatives.

**ACTION
N°18**

CRÉER UNE COMMUNAUTÉ DES CONTRIBUTEURS

Objectif

Favoriser l'amplification du mouvement de partage altruiste des données.

Recommandation

Nous recommandons la création d'une communauté des contributeurs altruistes visant :

- À mettre en relation des entreprises partageant des valeurs éthiques ;
- À promouvoir le mode altruiste de partage des données ;
- À permettre aux entreprises partageant des données de communiquer sur leur démarche altruiste ;
- La communauté devra veiller et mettre en œuvre des campagnes de sensibilisation contre les actes malveillants tels que les démarchages par de fausses organisations altruistes en vue d'obtenir des accès illégitimes à des données.

L'accès à cette communauté pourrait être soumis à l'obtention d'une certification d'acteur altruiste, soit émise par le régulateur, soit, le cas échéant, par une marque de certification créée par la fédération.

**ACTION
N°19**

**CRÉATION D'UNE MARQUE DE CERTIFICATION (LABEL)
DES ACTEURS ALTRUISTES**

Objectif

Création d'une certification contrôlée de la qualité des acteurs de l'altruisme des données.

Recommandation

Création d'une marque de certification par une entité capable d'établir des critères de labellisation des acteurs de l'altruisme. La fédération pourrait remplir ce rôle.

Mise en œuvre

- Dépôt d'une marque collective de certification.
- Rédaction et publication d'un règlement d'utilisation de la marque détaillant :
 - > Les critères à remplir pour avoir le droit d'afficher la marque ;
 - > Les conditions d'affichage de cette dernière (supports, destination des supports, taille et forme de l'application de la marque, etc.) ;
 - > Les sanctions du non-respect des conditions ;
 - > Les conditions du retrait de l'autorisation d'exploitation de la marque.
- Exemples de critères et conditions d'utilisation de la marque pour les contributeurs :
 - > Avoir effectivement donné accès à des données à au moins une organisation membre de la fédération, ce qui pourrait éventuellement être mesuré par un indicateur et une déclaration de la part de l'organisation membre ;
 - > Maintenir cet accès pendant toute la durée d'utilisation de la marque dans une perspective de collaboration sur le long terme (sans contradiction avec le droit de rétraction effectif, cf. action N°11) ;
 - > Adhérer à la charte éthique de la fédération.
- **Avantages :**
 - > Permet d'afficher la marque sur les supports de communication ;
 - > Permet de communiquer autour de la démarche altruiste ;
 - > Interdit d'utiliser la marque pour faire la promotion de toute activité incompatible avec la charte éthique de la fédération.

**ACTION
N°20**

**CRÉATION D'UN CRITÈRE D'ÉVALUATION
DANS LES APPELS D'OFFRES**

Objectif

Augmenter la participation à l'altruisme de la donnée.

Recommandation

Inclusion dans les appels d'offres d'un critère d'évaluation et de sélection des offres basé sur la participation ou l'engagement à participer à l'altruisme des données (ex : Programmes d'investissement nationaux, Agence nationale de la recherche, Digital Europe Programme/Horizon Europe, etc.).

Mise en œuvre

- La démonstration par les candidats de leur participation à l'altruisme des données peut découler de leur obtention du droit d'afficher :
- La marque de certification créée par une instance de contrôle (comme une fédération) ;
 - Ou la marque collective créée par une organisation altruiste en particulier.
- L'appel d'offre peut également comprendre un engagement des candidats à partager des données avec une ou des organisations altruistes prédéfinies.

Objectif

Compenser les coûts éventuellement liés à la démarche altruiste du contributeur.

Recommandation

Création d'un crédit d'impôt « data altruiste » compensant les coûts liés à la mise à disposition des données (coûts salariaux, frais d'infrastructures) sur les modèles des crédits d'impôts récompensant les dons aux associations et des crédits d'impôt recherche.

Mise en œuvre

La négociation de la mise en place de cette mesure peut être réalisée, par exemple, par une fédération des organisations altruistes.

- Le transfert de données à une organisation altruiste peut engendrer des coûts de mise en œuvre pour le contributeur potentiel.

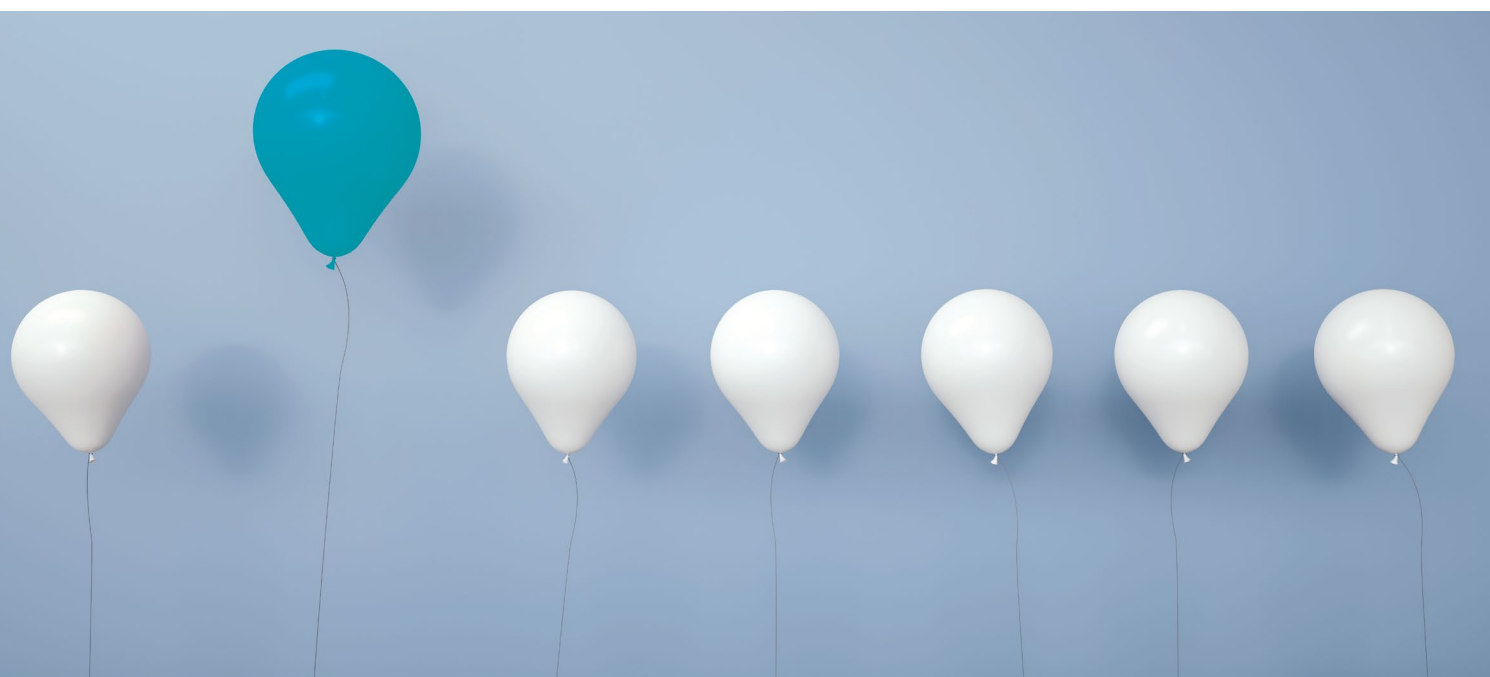
- Il existe des mécanismes d'incitation aux dons : réduction d'impôt sur le revenu de 66 % à 75 % du montant versé, selon l'association choisie, dans la limite de 20 % du revenu imposable.

- Il existe des mécanismes d'incitation à la recherche et au développement : le crédit d'impôt recherche, par exemple, permet de déduire, dans certaines conditions, les dépenses de recherche fondamentale et de développement expérimental des impôts.

- La mise en place d'une mesure d'incitation fiscale permettant au contributeur de déduire de ses impôts les frais liés au transfert de ses données à une organisation altruiste lèverait un obstacle important.

- Le mécanisme pourrait comporter deux options :

- > Une option forfaitaire, permettant une déduction fixe sans démonstration d'engagement effectif des sommes mais sous condition de démonstration d'une contribution effective, certifiée par une organisation altruiste ;
- > Une option « au réel », permettant une déduction fiscale des efforts matériels réellement engagés et justifiés, soumis à un plafond, et sous condition de la démonstration d'un transfert effectif de données, certifié par une organisation altruiste.



ÉTAPE N°4

ASSURER LES MODALITÉS DE PARTAGE DES DONNÉES

La question de l'hébergement des données est centrale, non tant d'un point de vue de faisabilité technique qu'afin d'assurer la confiance des parties prenantes dans le modèle.

La centralisation de données (personnelles ou stratégiques pour des entreprises) au sein d'un unique hébergement de l'organisation altruiste en matière de données suscite, à juste titre, une certaine méfiance.

Le risque de fuite, du fait d'erreur humaine, de dysfonctionnement ou d'acte malveillant, en dehors du modèle, ou entre des participants concurrents, est ainsi constamment relevé par les acteurs interrogés. De ce fait, il est de plus en plus fréquent de rencontrer des systèmes s'efforçant de contourner ce problème. Ces approches novatrices sont autant de possibilités pour les modèles data altruistes qui devront s'en saisir pour assurer un partage de données dans la confiance.

ACCÈS À LA DONNÉE

ACTION N°22	CONTRÔLE DES DEMANDES D'ACCÈS PAR LES UTILISATEURS DES DONNÉES
Objectif	
Assurer le respect et la poursuite des finalités déclarées par l'organisation altruiste.	
Recommandation	
L'organisation doit mettre en place une procédure de contrôle des candidatures à l'accès aux données par les contributeurs, mise en œuvre par la gouvernance des données, pour garantir aux contributeurs une utilisation conforme de leurs données. Nous recommandons la création d'un dossier type de candidature comprenant la liste des pièces et informations obligatoires à fournir.	
Mise en œuvre	
La détermination des critères de sélection des candidats dépend de la ou les finalités promues par l'organisation et de la sensibilité des données qui lui sont confiées. Nous recommandons que toute candidature comprenne notamment : <ul style="list-style-type: none">- Une description précise du projet et de ses objectifs ;- Une démonstration proposée par l'utilisateur de l'adéquation du projet avec les finalités promues par l'organisation ;- Un engagement d'adhérer à la charte éthique de l'organisation ;- Une déclaration des mesures qui seront mises en œuvre pour garantir la sécurité des données et leur utilisation uniquement dans le cadre du projet décrit. Outre l'archivage de ces déclarations faites par les utilisateurs aux fins d'établissement du rapport annuel prévu à l'article 18 du DGA, nous recommandons une ouverture au public de ces informations (cf. actions n°5 et 6). Nous attirons l'attention des organisations altruistes sur le risque lié aux fausses candidatures présentées par des utilisateurs malveillants.	

**ACTION
N°23****CRÉATION D'UNE PROCÉDURE DE RECOURS DES
CONTRIBUTEURS ET UTILISATEURS AUPRÈS
DE L'ORGANISATION DATA ALTRUISTE****Objectif**

Assurer un traitement équilibré des litiges sur les accès aux données.

Recommandation

L'organisation doit mettre en place un droit de recours, *a priori* géré par la gouvernance des données, à disposition des contributeurs et utilisateurs concernant les ouvertures d'accès aux données. La procédure de traitement des contestations et des litiges doit être transparente et publique.

Mise en œuvre

La procédure doit prévoir :

- Modalités de saisine :

- > Par les utilisateurs en cas de rejet de leur demande d'accès ;
- > Par les contributeurs en cas de contestation sur une ouverture d'accès.

- Délai de traitement ;

- Recours éventuel contre la décision de l'instance de gestion des litiges.

Composition de l'instance de gestion des litiges :

- Il n'est pas nécessaire de prévoir la création d'un organe permanent chargé de la gestion des litiges, surtout à la création de l'organisation ;

- Néanmoins, il faut prévoir, dès l'origine, qui sera amené à statuer en cas de litige et donc comment sera composée l'instance de gestion des litiges réunie spécialement le cas échéant ;

- Les responsables de traitement d'un recours doivent dans la mesure du possible être différents des décisionnaires initiaux ;

- Il peut être envisagé de prévoir deux formations en fonction de la sensibilité des données :

- > Collégiale ;
- > Unipersonnelle.

Dans le cadre de l'évolution de l'organisation altruiste, il pourra être créé un organe *ad hoc* permanent de gestion des litiges.

ÉLÉMENTS TECHNIQUES

**ACTION
N°24****PROTÉGER LES DONNÉES EN LIMITANT
TECHNIQUEMENT L'ACCÈS AUX DONNÉES BRUTES****Objectif**

Permettre un accès efficace tout en préservant la sécurité de celles-ci, tout le long de son cycle de vie.

Recommandation

L'organisation doit s'efforcer, à chaque fois que cela est possible, de limiter l'accès aux données brutes qui lui sont transmises, grâce à l'utilisation d'outils et de méthodes techniques appropriés et souverains lorsque cela est rendu possible.

Mise en œuvre

L'organisation doit déterminer le mode d'hébergement des données le plus adéquat et sécurisé en fonction de la nature des données et des risques afférents. Le recours à un hébergement centralisé avec transmission directe des données brutes à l'utilisateur doit être évité dans la mesure du possible. Des mécanismes plus protecteurs des données brutes doivent être investigués notamment :

- Via un système centralisé mais ne traitant que des données agrégées et/ou chiffrées ;
- Via un système décentralisé.

Les expérimentations en matière de *federated learning* sont prometteuses et peuvent constituer un modèle de choix lorsqu'il est compatible avec les données collectées et les finalités promues. Indépendamment de ce choix de modèles d'autres mécanismes peuvent être ajoutés : destruction automatique des données transférées après un certain laps de temps.



ÉTAPE N°5

ASSURER LE FINANCEMENT DE L'ORGANISATION DATA ALTRUISTE

Un modèle data altruiste s'accompagne d'outils de collecte de données, de centralisation via une plateforme, de flux de partage, autant de briques technologiques qui doivent être conçues, déployées et maintenues. Cela implique la définition de règles de partage sous forme de licences (implicites ou explicites) et d'une méthodologie de standardisation et d'homogénéisation (respect de normes, d'un format, d'une structure et d'un niveau de précision). Ces développements nécessitent du temps, des ressources, des personnels compétents, et donc un financement.

ACTION N°25	ENCADRER LES FINANCEMENTS PAR DONATIONS
Objectif	
Permettre un financement respectant les conditions d'existence et d'enregistrement de l'organisation altruiste.	
Recommandation	
L'organisation doit organiser ses campagnes de collectes de fonds en ne proposant aucune contrepartie aux donateurs de nature à compromettre son indépendance, notamment juridique, vis-à-vis des donateurs.	
Mise en œuvre	
<ul style="list-style-type: none"> - La donation financière au profit de l'organisation doit se faire sans contrepartie autre que les incitations fiscales (mises en œuvre par l'État) ou la « labellisation » par la marque collective ; - La dénomination de l'organisation ne doit pas être identique ou assimilée à celle de son ou ses donateurs ; - La nature et le volume des financements par donations doivent être rendus publics ; - Les donateurs ne doivent d'aucune manière disposer de pouvoirs directs ou indirects au sein de la gouvernance de l'organisation. 	

Objectif

Organiser de manière prévisible et transparente les conditions financières d'accès et d'exploitation des données de l'organisation.

Recommandation

Création par les organisations de conditions financières d'accès aux données distinguant de manière :

- objective ;
- non discriminante ;
- proportionnée.

Les conditions financières d'accès aux données en fonction notamment :

- de la forme juridique de l'utilisateur ;
- de ses capacités financières ;
- de l'utilisation qu'il veut faire des données.

Mise en œuvre

Par exemple, peuvent être mises en place les licences suivantes :

- **Licence gratuite et contaminante** : destinée aux acteurs de type universités, laboratoires publics, ne disposant pas de fonds importants mais travaillant sur un programme de recherche déterminé.

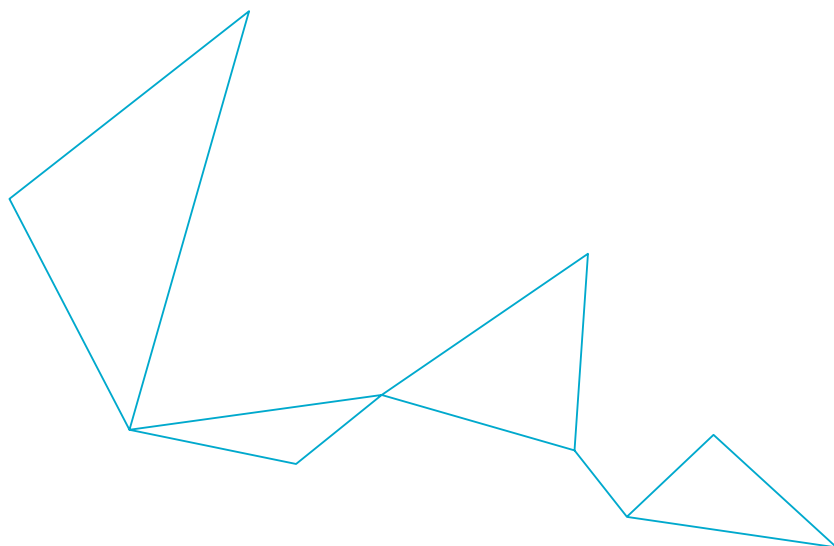
La gouvernance peut ouvrir un accès gratuit aux données de l'organisation, à condition de recevoir en contrepartie les données retraitées dans le cadre du programme de recherche ou, *a minima*, un rapport sur l'utilisation des données qui puisse bénéficier à des tiers et à la finalité de l'organisation.

- **Licence payante *in fine*** : destinée aux acteurs privés ne disposant pas de fonds importants mais travaillant sur un programme de recherche déterminé.

La gouvernance peut octroyer un accès initialement gratuit aux données disponibles, à condition de verser à l'organisation des redevances proportionnelles au chiffre d'affaires réalisé en cas de commercialisation d'un produit créé à l'issue du programme de recherche, que les données transmises par l'organisation aient contribué ou non au résultat final.

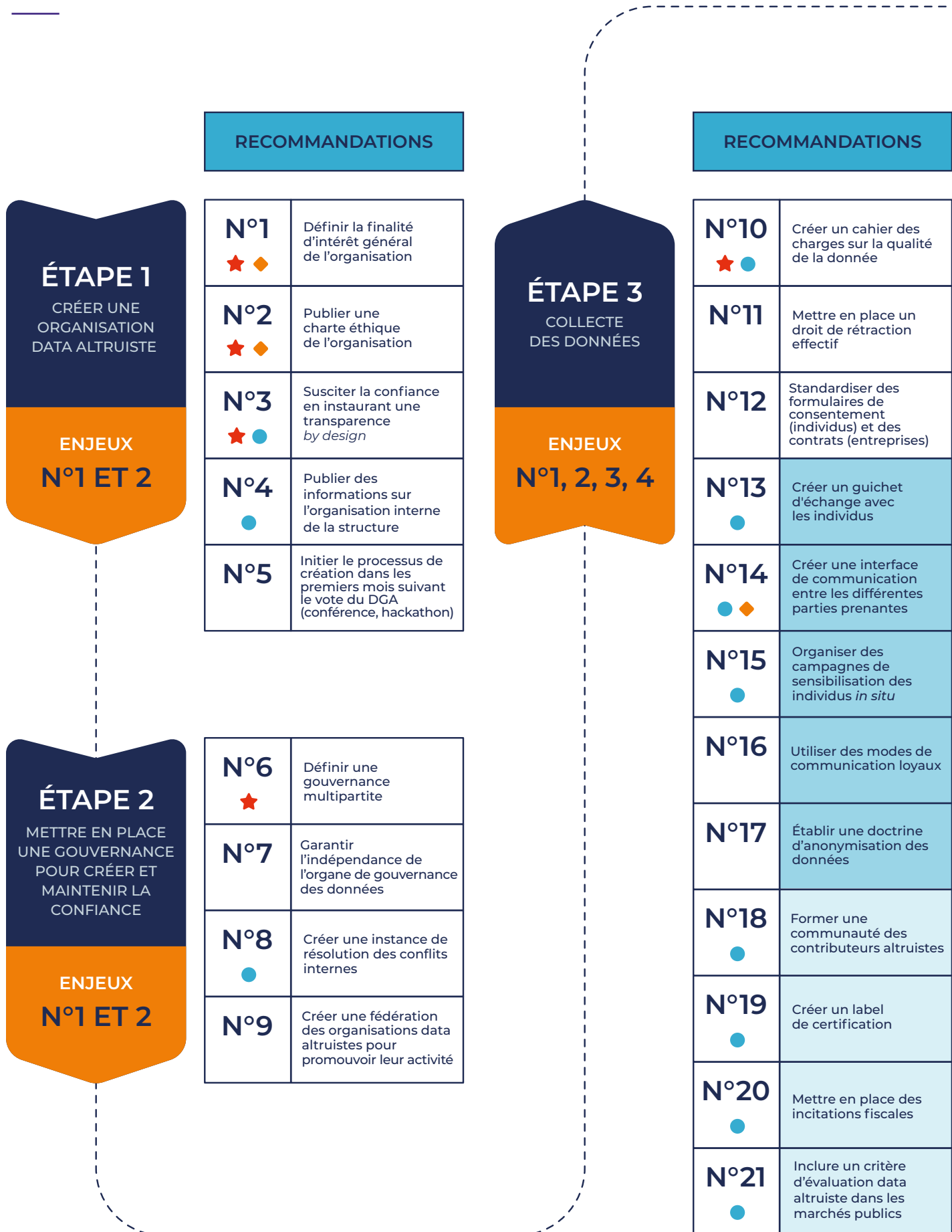
- **Licence payante** : destinée aux acteurs privés ayant une activité commerciale.

La gouvernance peut donner accès aux données détenues par l'organisation pour mener des travaux de recherche visant à créer un produit commercialisable en contrepartie d'une redevance calculée proportionnellement au chiffre d'affaires annuel de l'utilisateur.





RECOMMANDATIONS PAR ÉTAPE



ÉTAPE 4

ASSURER LES MODALITÉS DE PARTAGE

ENJEUX

N°1, 2, 3, 6

ÉTAPE 5

ASSURER DES MODALITÉS DE FINANCEMENT DU MODÈLE

ENJEUX

N°1 ET 5

RECOMMANDATIONS

N°22 ★ ●	Mettre en place une procédure de contrôle des candidatures à l'accès aux données
N°23 ◆	Créer une procédure de recours
N°24 ★ ◆	Protéger les données en limitant l'accès aux données brutes

N°25	Encadrer les financements par donations
N°26 ★ ◆	Définir les modalités financières d'accès aux données (ex : licences)

ENJEUX

N°1	Assurer un haut niveau de confiance dans le modèle data altruiste et le tiers de confiance sur le long terme
N°2	Saisir les spécificités, craintes et intérêts de chaque parties prenantes vis-à-vis du data altruisme
N°3	Démontrer aux entreprises l'intérêt du data altruisme, y compris dans un contexte de forte pression concurrentielle
N°4	Assurer la confiance des individus alors que leurs données sont sur-sollicitées dans l'espace numérique afin d'alimenter le système marchand
N°5	Trouver l'équilibre entre la nécessaire implication des pouvoirs publics pour accompagner l'émergence de modèles data altruistes et le risque de nuire à la confiance des autres parties prenantes, notamment des entreprises et des individus
N°6	Identifier des sources de financements pérennes pour les modèles data altruistes autres que les subventions publiques, tout en assurant leur indépendance

Encouragement des personnes physiques

Encouragement des personnes morales

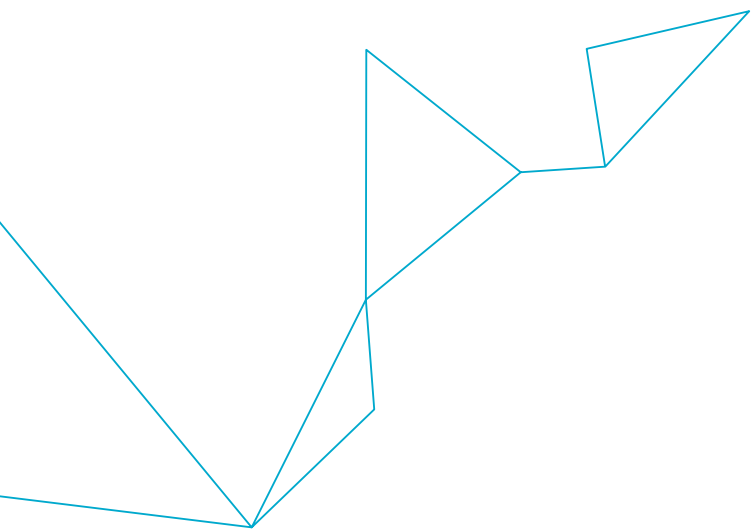
★ Recommandation prioritaire

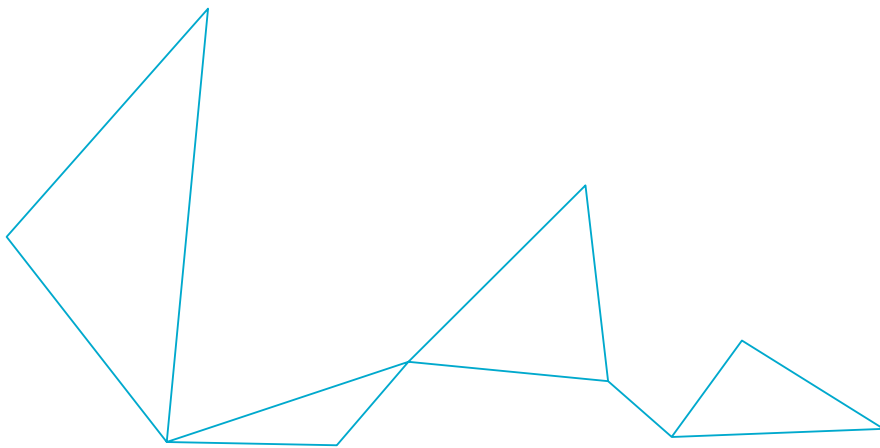
● Contributeur

◆ Utilisateur



CONCLUSION





Les données constituent une ressource disponible et sous-exploitée, dont le partage est nécessaire à notre développement économique et social. L'Union européenne ambitionne de mettre en place un cadre à leur bonne utilisation. Un cadre dont le DGA et le data altruisme sont des briques constitutives.

Plus qu'une stimulation des flux des données, le data altruisme ambitionne d'ouvrir une voie nouvelle pour des modes d'échanges dans le numérique jusqu'alors dominé par l'extractivisme ou la contrainte. Il vise à créer une dynamique unique, reflétant profondément les valeurs européennes, fondée sur le décloisonnement des données et leur partage, en vue de l'intérêt général. Cette opportunité s'articule autour d'un nouvel acteur : une organisation data altruiste, tiers de confiance garant de la finalité du partage des données.

La mise en œuvre de cette démarche nécessite cependant de relever plusieurs défis. Pour convaincre de leur bien-fondé et de leur efficacité, les organisations data altruistes auront donc fort à faire. Elles auront tout intérêt à s'inspirer d'expériences actuelles déjà riches d'enseignements. Car la confiance dans le monde numérique, qui se trouve au cœur du data altruisme, se doit de reposer sur des mécanismes concrets toujours questionnés et actualisés.

Les actions suggérées dans le présent rapport n'ont pour objectif que de faciliter la mise en pratique d'un concept qui nous semble riche de promesses. Une fois établis ces cadres pratiques, il appartient aux citoyens européens et aux entreprises développant leur activité au sein de l'Union de s'emparer de l'opportunité que constitue la démarche de data altruisme. Seul l'engagement de chacun à partager ses données dans l'intérêt général permettra de saisir les possibilités offertes par la transformation numérique.





Scanner le QR code pour accéder à l'annexe du rapport :

- Des fiches sur des modèles novateurs de partage de données
- Une cartographie juridique du data altruisme
- La liste des cas pratiques étudiés