

# Pour un partage efficace et éthique des données à grande échelle

Un livre de recettes pour les producteurs de données, les donateurs, les décideurs et autres praticiens du développement



Avec le soutien de  
**Google.org**

# À propos de ce livre de recettes

## Qu'est-ce qu'un partage de données multipartite efficace ?

Le **partage de données multipartite** dans le secteur du développement fait référence à des situations dans lesquelles deux ou plusieurs organisations collaborent pour collecter, partager et/ou analyser des données afin de relever des défis sociétaux.

Le partage de données multipartite se produit souvent entre des parties prenantes de différents secteurs, notamment des agences gouvernementales, des entreprises privées et des organisations non gouvernementales, qui se réunissent pour établir des partenariats de collaboration. Ce document se concentre sur les initiatives de partage de données qui impliquent des partenaires de différents secteurs, et ne traite pas du partage de données entre entreprises (B2B) et entre gouvernements (G2G).

Il n'existe pas de consensus formel ou largement accepté sur ce qui rend un partenariat de partage de données efficace en général, et encore moins dans le secteur du développement. Dans ce livre de recettes, nous nous appuyons sur les contributions des praticiens du développement pour caractériser le partage efficace des données comme une **réduction des frictions** dans le partage des données sur le long terme qui **répond à un défi sociétal** ou cherche à améliorer le bien-être public sans produire d'externalités négatives, et qui est **responsable et transparent** envers les principales parties prenantes et le grand public.



Pour un partage efficace des données, vous aurez besoin de :

Une **réduction des frictions** dans le partage des données sur le long terme...



qui **répond à un défi sociétal** ou cherche à améliorer le bien-être public sans produire d'externalités négatives...

et qui est **responsable et transparent** envers les principales parties prenantes et le grand public.



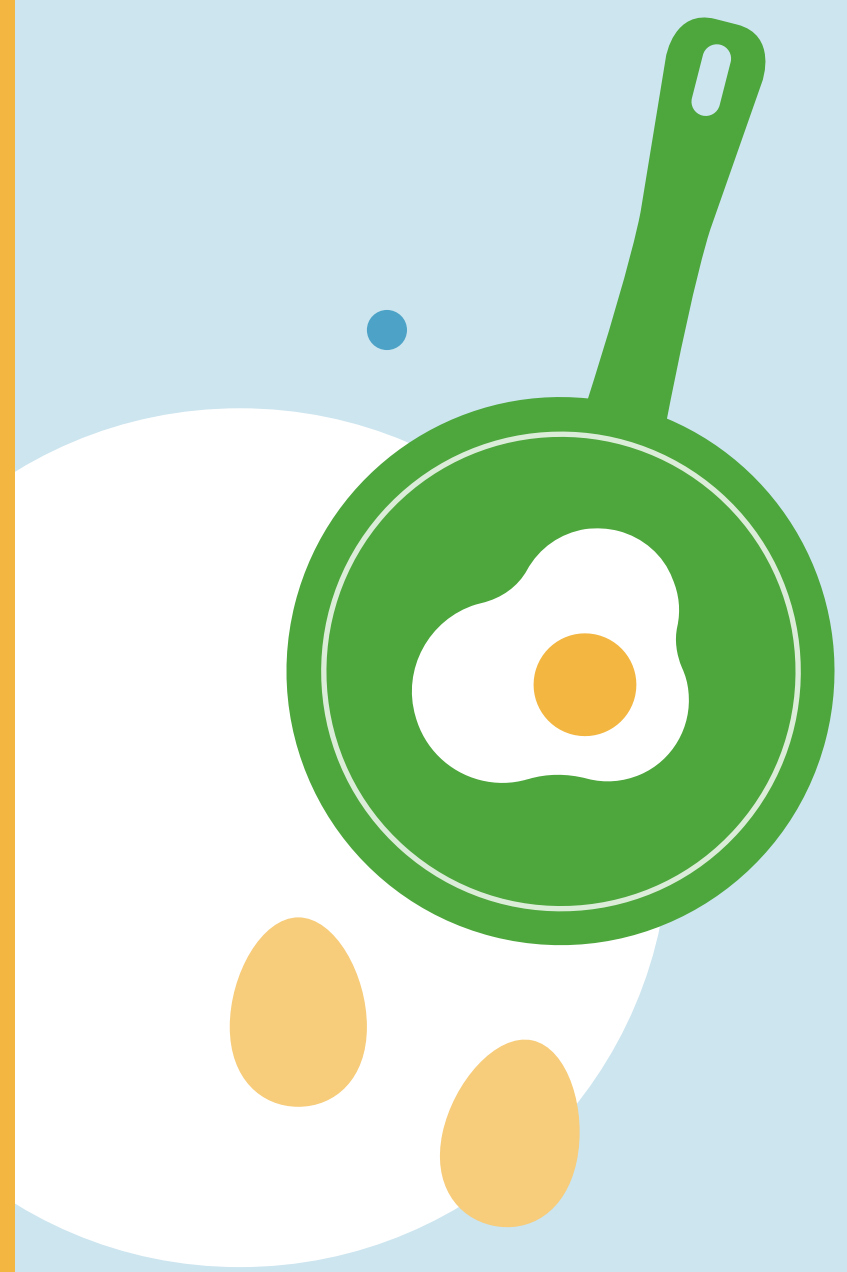
## Pourquoi avons-nous besoin d'un livre de recettes sur le partage de données multipartite ?

Les initiatives de partage de données multipartites dans le secteur du développement ont proliféré au cours de la dernière décennie, à mesure que s'intensifient les efforts visant à mieux exploiter les données pour atteindre les objectifs de développement durable.

Cependant, **il existe peu d'espaces pour échanger des connaissances** sur les meilleures pratiques ou pour permettre aux partenaires engagés dans des initiatives (plus ou moins efficaces) de consolider l'apprentissage et de dégager des idées communes. Ce déficit persiste malgré les demandes des membres de la communauté des données pour le développement pour des recommandations et des idées concrètes sur ce qui permet un partage efficace des données.

L'objectif de ce livre de recettes est de fournir aux praticiens et aux organisations de développement **des idées, des recommandations et des exemples concrets et fondés sur des preuves, pour créer des initiatives de partage de données réussies**. Ce livre de recettes aspire à être un outil facile à utiliser, pragmatique, concret et concis pour notre communauté.





## A propos du succès du partage des données dans le secteur du développement

Si le concept de partage des données est simple, mesurer le succès des initiatives de partage des données est plus compliqué. Lors de conversations avec des praticiens du développement, trois concepts généraux sont apparus comme des indicateurs de réussite :



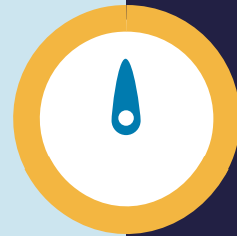
### Atteindre les objectifs des indicateurs clés de performance

Ces définitions du succès se concentrent sur les résultats des accords de partage de données, tels que le nombre d'utilisateurs de données, les applications ou services développés, ou les rapports produits.



### Résoudre les défis sociétaux

Pour certains experts, le succès des initiatives de partage des données est représenté par leur efficacité à résoudre les problèmes du monde réel sans créer d'externalités négatives ou de préjudices (comme l'exclusion de groupes marginalisés). Ces mesures portent sur les résultats à long terme plutôt que sur les résultats immédiats.



### Devenir autonome

Certains experts considèrent que le succès se mesure à la capacité d'une initiative de partage de données à devenir autonome au fil du temps. Cela englobe des aspects de stabilité financière ainsi que les relations entre les principales parties prenantes, en plus du maintien des opérations et de la qualité des données.

## Comment ce livre de cuisine a-t-il été créé ?

Ce livre de recettes est le résultat d'un projet d'un an entrepris par le [Global Partnership for Sustainable Development Data](#) (Global Partnership) et parrainé par [Google.org](#), visant à identifier les moteurs d'un partage efficace des données dans le secteur du développement.

Il s'appuie sur les informations recueillies dans le cadre d'une analyse du paysage des initiatives de partage des données et sur les commentaires d'un groupe de travail diversifié composé de plus de 20 experts qui se sont réunis à quatre reprises entre janvier et septembre 2022. Il s'appuie également sur l'expérience du Global Partnership, qui aide les pays et les organisations à combler leurs besoins et leurs lacunes en matière de données en collaborant avec des partenaires. Le Global Partnership a facilité plus de [120 collaborations](#) de partage de données depuis 2015.

Si le partage des données peut être comparé à notre régime alimentaire, alors le **partage efficace des données est comme une alimentation saine**. Pour avoir un régime alimentaire sain, il faut savoir quels aliments sont bons pour la santé. Cuisiner des

recettes saines nécessite de savoir quels ingrédients sont importants et comment les combiner pour chaque plat. Tout comme une alimentation saine et équilibrée, le partage efficace des données dans le secteur du développement est essentiel pour **garantir que les activités liées aux données bénéficient aux personnes** sur le terrain.

Les similitudes entre une alimentation saine et un partage efficace des données ne s'arrêtent pas là. Bien sûr, aucun d'entre nous ne mange exactement comme ses ancêtres il y a quelques siècles, car les aliments et les goûts ne sont pas immuables. Non seulement les ingrédients, les recettes et les méthodes varient considérablement à travers le monde, chaque pays et région développant sa propre cuisine, mais ils sont également en constante évolution. De même, ce livre de cuisine n'est qu'un instantané des connaissances actuelles et des expériences de partage de données des "chefs" du monde entier. Comme pour tout livre de cuisine, des recettes nouvelles et actualisées devraient être ajoutées régulièrement pour que cet outil reste pertinent et pour accroître la diversité des secteurs, des régions et des initiatives qui y sont représentés.

Enfin, il faut reconnaître que, malgré tous les efforts déployés pour rechercher et consulter un éventail diversifié d'experts et d'initiatives en matière de partage des données, ainsi que pour tirer parti de l'expérience du Global Partnership en matière de collaboration avec des organisations du Sud, ce document s'appuie essentiellement sur des connaissances théoriques et des recettes qui ont été élaborées dans le Nord du monde. Des efforts supplémentaires seront nécessaires pour rééquilibrer nos connaissances culinaires afin d'y ajouter davantage d'ingrédients et d'astuces provenant d'autres contextes.



**Dans les recettes, cette icône indique qu'il faut poursuivre la recherche, l'expérimentation et/ou l'échange de connaissances.**

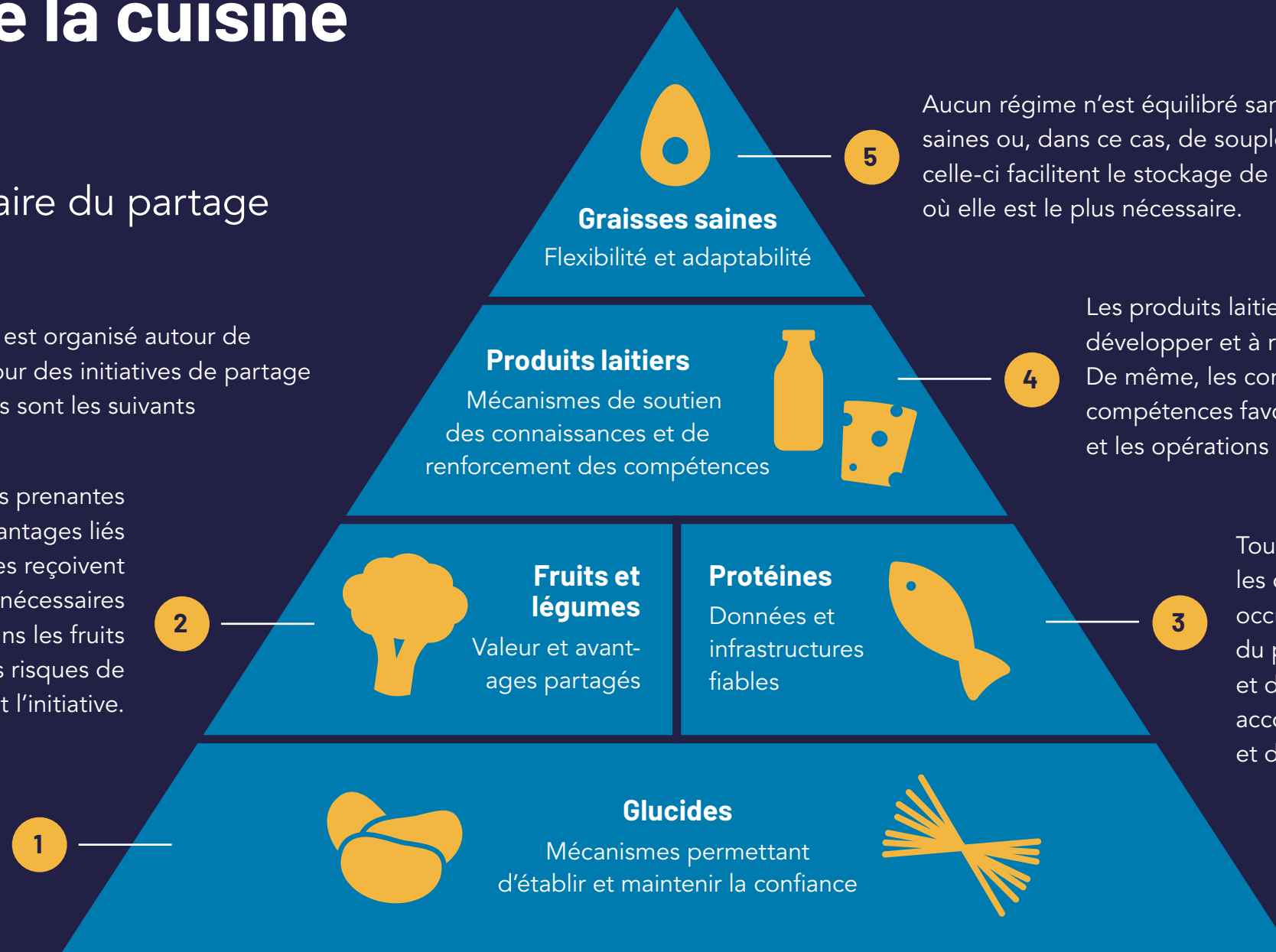
# Comprendre la cuisine de base

## La pyramide alimentaire du partage des données

Le contenu de ce livre de recettes est organisé autour de cinq groupes d'ingrédients clés pour des initiatives de partage de données efficaces. Ces groupes sont les suivants

Lorsque toutes les parties prenantes partagent la valeur et les avantages liés au partage des données, elles reçoivent les vitamines et les minéraux nécessaires (comme ceux que l'on trouve dans les fruits et légumes) pour réduire les risques de "maladies" menaçant l'initiative.

Tout régime sain commence par une base de confiance entre les partenaires de partage des données, les utilisateurs des données et les autres parties prenantes.



Aucun régime n'est équilibré sans une dose de graisses saines ou, dans ce cas, de souplesse et d'adaptabilité: celle-ci facilitent le stockage de l'énergie pour le moment où elle est le plus nécessaire.

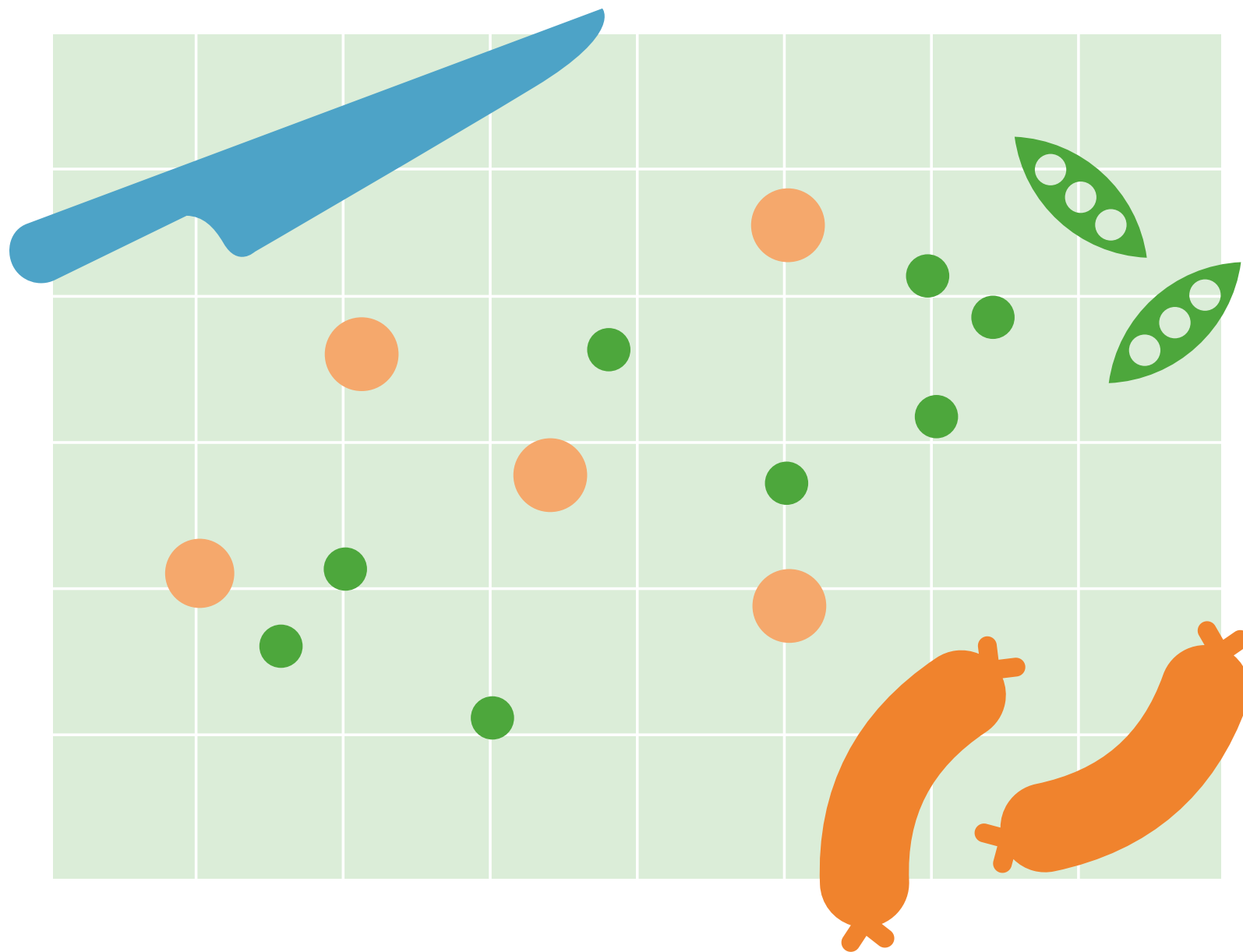
Les produits laitiers contribuent à développer et à renforcer les os humains. De même, les connaissances et les compétences favorisent le fonctionnement et les opérations de partage des données.

Tout comme les protéines, les données et l'infrastructure occupent le centre de l'assiette du partage des données et doivent toujours être accompagnées d'une valeur et d'avantages partagés.

Ces cinq groupes alimentaires se retrouvent dans toutes les initiatives de partage de données réussies. Ils constituent la base d'un régime varié et équilibré de partage des données. La quantité (ou l'importance) de chacun d'entre eux varie en fonction des initiatives de partage de données individuelles, tout comme les recettes nécessaires pour les créer.

Chacune des cinq sections de ce livre de cuisine contient donc des "recettes", c'est-à-dire des méthodes et des outils testés qui peuvent être reproduits par d'autres.

Le partage des données dans le secteur du développement est un domaine relativement nouveau, et il existe des lacunes évidentes en matière de connaissances et d'expérience. C'est pourquoi le livre indique les endroits où il manque des recettes ou des informations sur les ingrédients. La mise en évidence de ces lacunes va aider les chefs en herbe à comprendre ce qu'il reste à faire et va ouvrir la voie à davantage de réflexion et d'expérimentation.



## Facteurs influençant le choix et la quantité des ingrédients

Il n'existe pas d'approche unique pour un partage de données multipartite efficace. Ce livre de cuisine propose plutôt des suggestions d'ingrédients clés et des recettes pour inspirer des approches créatives.

L'importance (ou la quantité) de chaque groupe d'aliments, ainsi que les recettes à utiliser, dépendent de deux grandes catégories de facteurs liés à :

► **Les caractéristiques de l'initiative spécifique de partage des données**, telles que les parties prenantes ou les secteurs impliqués, les types de données (c'est-à-dire personnelles, non personnelles) utilisées, le stade auquel se trouve l'initiative, les objectifs du partenariat, le nombre de parties prenantes impliquées et l'ouverture des données ou de l'initiative.

► **Le contexte dans lequel l'initiative de partage des données fonctionne**, y compris, par exemple, l'environnement réglementaire et politique, les autres parties prenantes au sein d'un écosystème de données particulier, et d'autres facteurs contextuels tels que le fait que le partage des données ait lieu pendant une situation d'urgence (par exemple, une pandémie ou une catastrophe naturelle) ou pendant les "activités habituelles".





1

# Mécanismes permettant d'établir et de maintenir la confiance

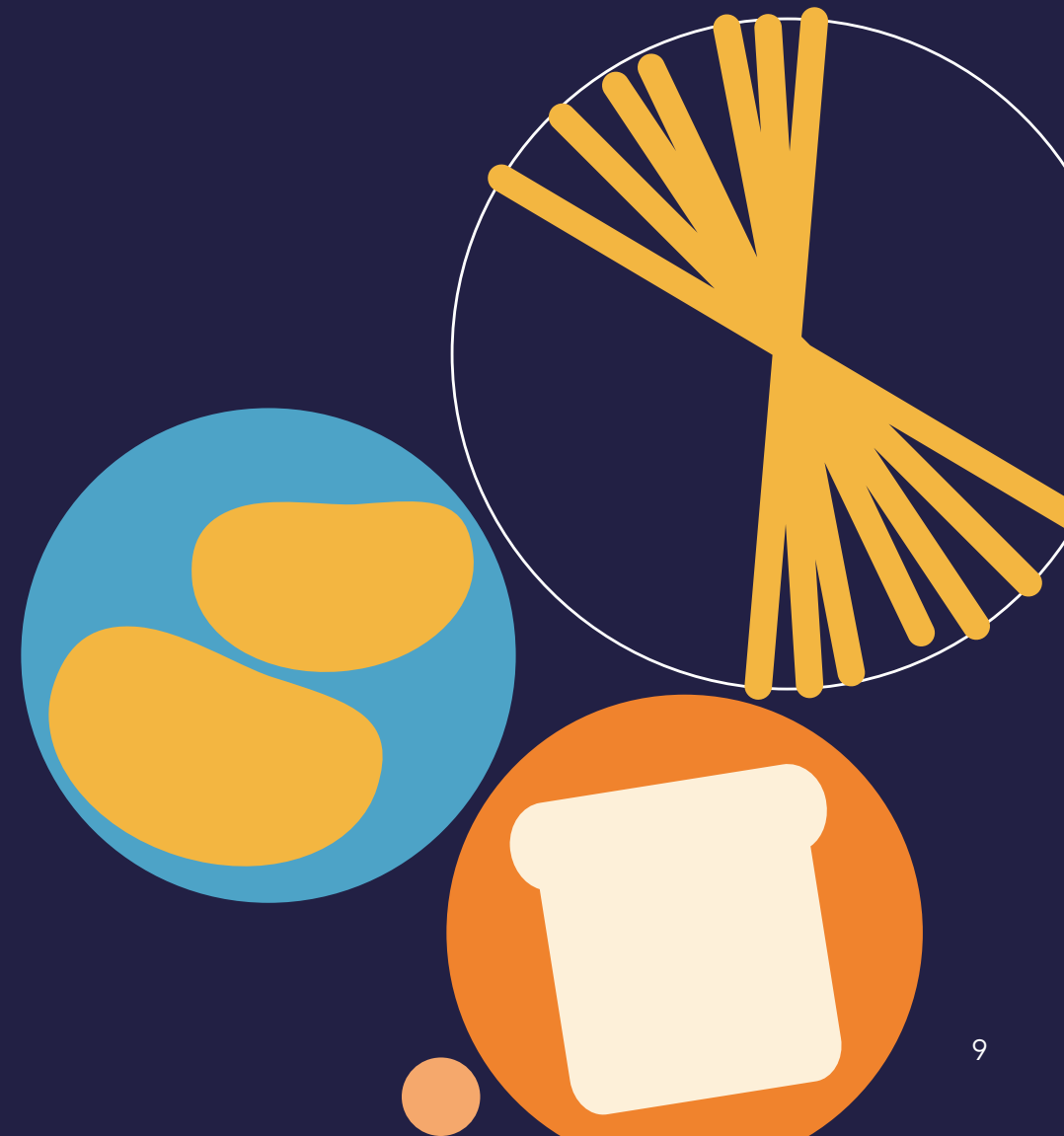
## Carburateurs

Recettes de cette section

Des modèles de gouvernance adaptés aux besoins

Communiquer avec le public et les parties prenantes

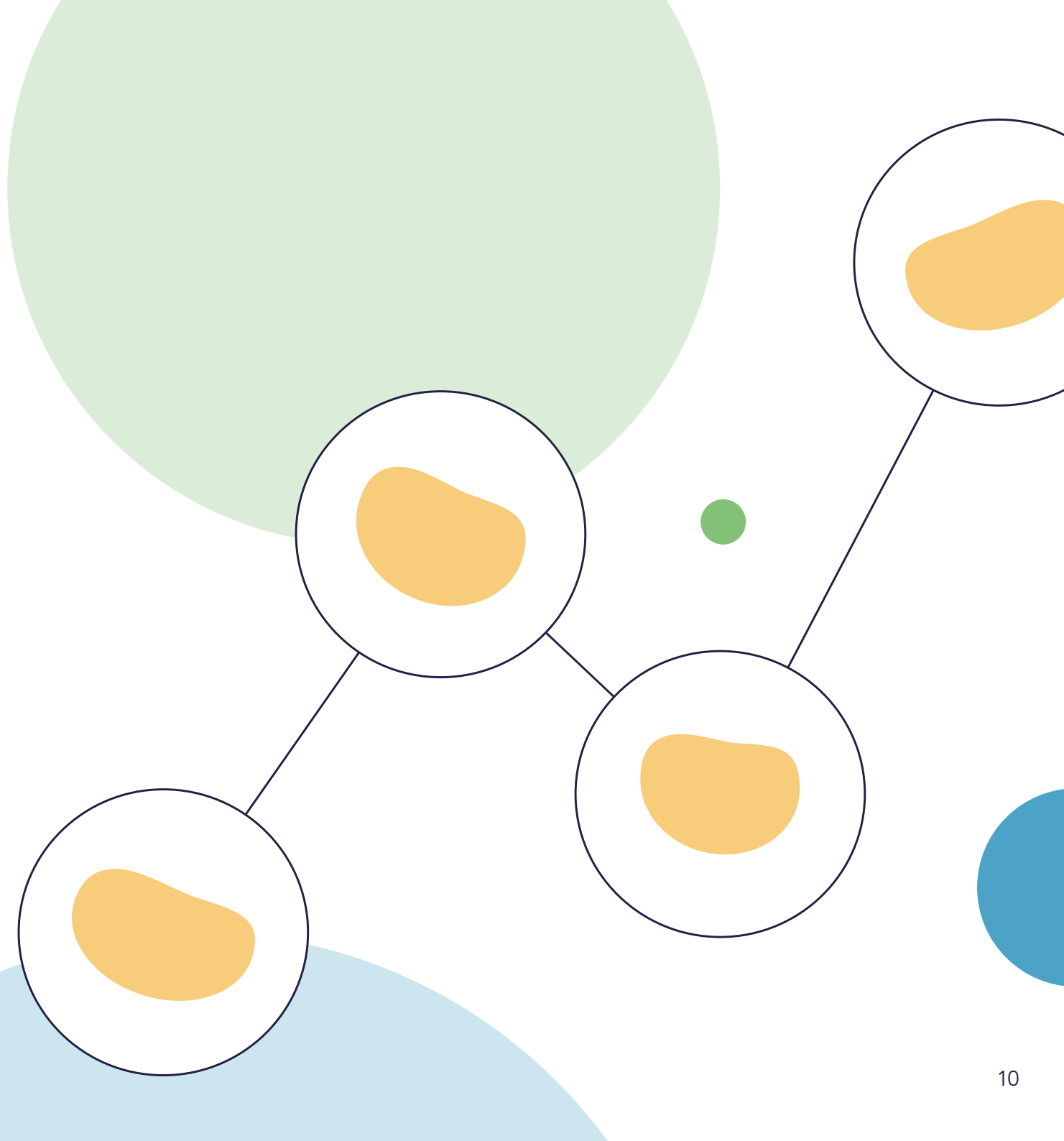
Responsabilisation



## Introduction

La **confiance** est l'ingrédient le plus important pour un partage de données multipartite efficace et, précisément comme les glucides, elle **fournit aux initiatives l'énergie dont elles ont besoin pour fonctionner et atteindre leurs objectifs**. Des recherches récentes mettent en évidence les preuves quantitatives solides selon lesquelles une plus grande confiance est associée à un partage accru des données et soulignent que l'impact de la confiance est particulièrement significatif lorsque les niveaux initiaux de confiance sont faibles. Toutefois, ces mêmes recherches révèlent également que, pour atteindre des niveaux optimaux de partage des données, l'augmentation de la confiance doit être associée à d'autres stratégies ou actions, c'est pourquoi la confiance doit être combinée à d'autres ingrédients.

Le simple fait de dire que la confiance est importante n'aide pas les organisations ou les individus travaillant sur des initiatives de partage de données à comprendre clairement comment ou ce qu'ils doivent faire. Pour cette raison, et afin de favoriser un partage des données plus efficace, cette section se concentre plutôt sur la compréhension des mécanismes efficaces pour construire et maintenir la confiance dans les initiatives de partage des données.



## Qu'est ce qu'on cuisine ?

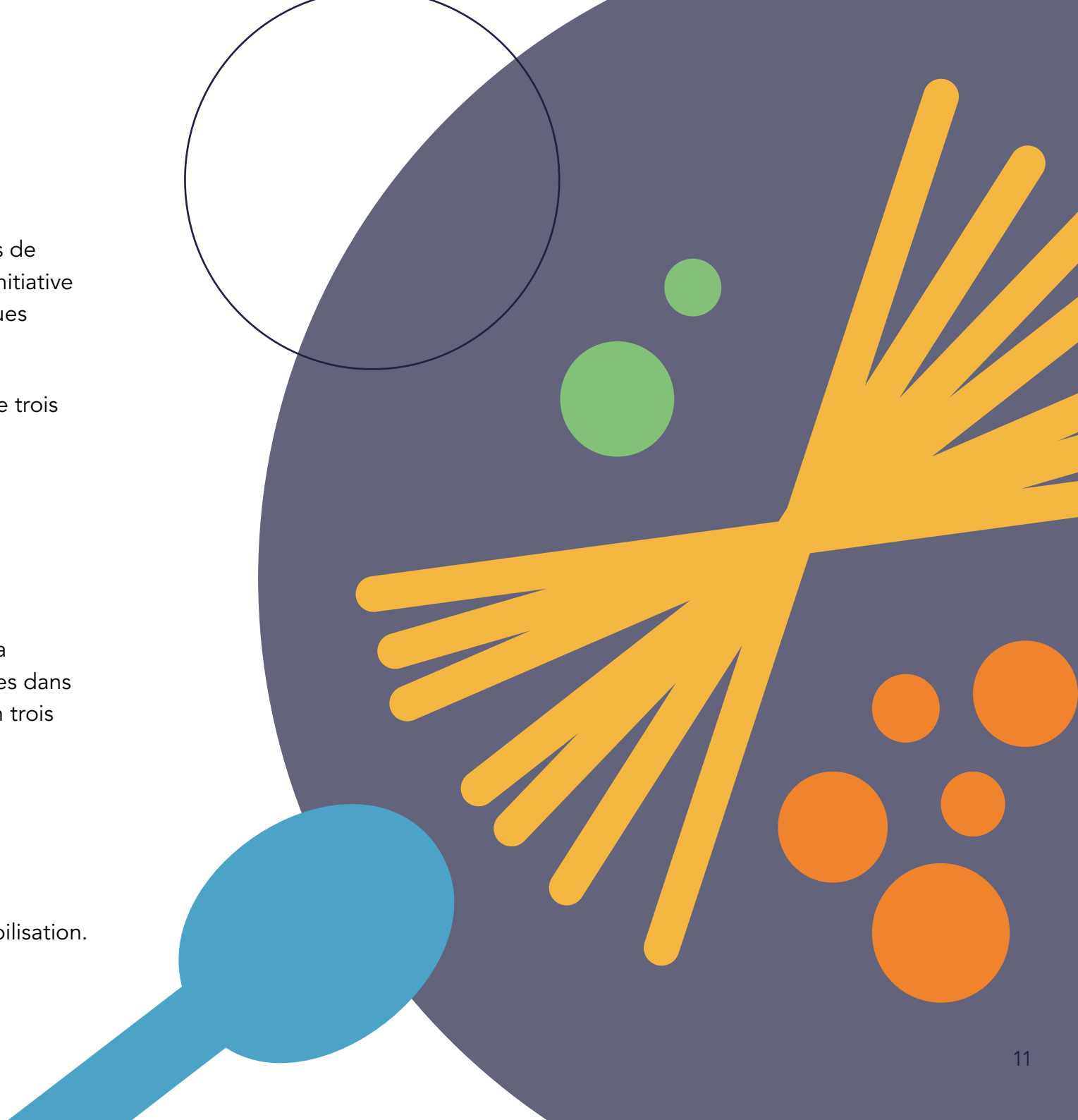
Pour que les initiatives de partage des données soient couronnées de succès, les partenaires doivent établir et maintenir la fiabilité de l'initiative et intégrer des valeurs fondamentales et des considérations éthiques dans son fonctionnement.

La fiabilité d'une initiative de partage de données est le résultat de trois éléments clés :

- ▶ La confiance entre les partenaires,
- ▶ La confiance de la communauté en général, et
- ▶ Mesures de responsabilisation.

Il s'ensuit que les recettes permettant d'instaurer et de maintenir la confiance et d'intégrer des valeurs et des lignes directrices éthiques dans les initiatives de partage des données peuvent être regroupées en trois catégories, qui sont décrites dans la section suivante :

- ▶ Modèles de gouvernance adaptés aux besoins
- ▶ Méthodes de communication efficaces à l'égard du public et des principales parties prenantes
- ▶ Audits, cadres de référence et autres mécanismes de responsabilisation.



## Des recettes pour des modèles de gouvernance adaptés aux besoins

L'instauration et le maintien de la confiance est un processus continu et itératif qui reste au centre des préoccupations tout au long du cycle de vie d'un partenariat de partage de données.

Bien entendu, il est plus facile d'instaurer la confiance lorsque les partenaires de données ont déjà travaillé ensemble. Une collaboration antérieure réussie améliore les niveaux de base de confiance et de confort. Grâce aux expériences positives antérieures de travail en commun, les canaux de communication et les relations établies facilitent le travail en commun sur de nouvelles initiatives.

Mais les expériences antérieures ne sont pas nécessairement des précurseurs de la confiance dans les nouveaux partenariats. La mise en place de mécanismes de gouvernance pertinents peut contribuer à instaurer et à maintenir la confiance, même lorsque les partenaires n'ont jamais travaillé ensemble auparavant.

**Les modèles de gouvernance participative** tels que les **"data trusts"** ou les **collaborations** en matière de données, les cadres de partage des **données détaillés et standardisés** ou les protocoles d'accord entre partenaires, ainsi que les **mécanismes de résolution des conflits**, comptent parmi les mécanismes les plus répandus et les plus efficaces pour y parvenir.

### Modèles de gouvernance participative – Hong Kong Data Trust 1.0

Le système de transport public de Hong Kong se compose de taxis, du Mass Transit Railway (MTR), de bus, de minibus et de tramways. Chaque prestataire de transport est privé et le partage des données entre les fournisseurs a toujours été limité. Le principal obstacle au partage des données a été le manque de confiance entre les parties prenantes et l'absence d'un cadre pour régir le partage des données.

Un [Data Trust](#) à l'Université de Hong Kong (HKU) visait à surmonter cet obstacle grâce au partage des données en temps réel. Ce Data Trust était une preuve de concept pour un modèle de tiers de confiance qui encourage les contrôleurs de données (opérateurs de transport public) à partager leurs données en toute sécurité avec le Data Trust. Le Trust lui-même était une entité hébergée par HKU (un tiers neutre) avec une responsabilité fiduciaire envers les contrôleurs de données et la capacité technique d'analyser et de traiter les données.

Les relations entre les opérateurs de transport public et HKU ont été formellement établies par des protocoles d'accord. Le partage des données via le Data Trust a également été facilité par un cadre juridique clair, une communication constante avec les parties prenantes et des canaux ouverts pour le retour d'informations.

## Collaborations en matière de données – un exemple en Californie

Le [California Data Collaborative](#) (CaDC) est un organisme indépendant à but non lucratif qui travaille en tant que coalition d'agences d'approvisionnement en eau en Californie pour faciliter les politiques et les opérations relatives à l'eau basées sur les données. Selon le GovLab, "les [Data Collaboratives](#) sont une nouvelle forme de collaboration, au-delà du modèle de partenariat public-privé, dans laquelle les participants de différents secteurs – en particulier les entreprises – échangent leurs données pour créer une valeur publique."

Le CaDC cherche à faire de l'analyse des données sur l'eau un élément central de l'amélioration de la gestion de l'eau. Il maintient une base de données sécurisée, exactes et normalisée sur l'utilisation de l'eau pour chaque organisme membre. À l'aide de ces données, le CaDC crée des outils d'analyse de données et développe des idées directement exploitables.

Le CaDC est régi par un comité directeur composé de membres de toutes les agences d'approvisionnement en eau participantes. Ce comité fixe les priorités de recherche et de politique du CaDC et identifie les projets pilotes pour la collaboration. Lors de ses premiers ateliers, le CaDC a élaboré un "cadre de confiance" comprenant des accords standards de partage et de transfert de données accessibles à tous. Les organisateurs espéraient que cet accord standardisé favoriserait la transparence et garantirait que chacun ait accès aux mêmes règles de partage des données. Ce mode de gouvernance équitable et transparent a contribué au bon fonctionnement du CaDC.

## Cadres détaillés et standardisés de partage des données

Development Data Partnership est une initiative qui vise à répondre aux besoins des organisations internationales de développement. Elle a créé une plateforme pour les mettre en relation avec des entreprises technologiques ou des partenaires de données. Étant donné que l'instauration d'un climat de confiance entre les partenaires, la négociation des conditions et la conclusion d'un accord de partage des données demandent du temps, Development Data Partnership a créé un accord-cadre de licence très détaillé qui définit les procédures relatives aux aspects cruciaux de la gestion des données, tels que la propriété et l'accès. Tous les partenaires signent le même accord, et celui-ci fournit le cadre dans lequel le Development Data Partnership doit rassembler, harmoniser, traiter et stocker les données. L'accord rend les procédures prévisibles et dignes de confiance, ce qui réduit les obstacles et permet à un partenaire de partager facilement des données.

Les accords standardisés ne sont toutefois pas toujours la meilleure solution. Par exemple, Global Fishing Watch travaille directement avec les gouvernements nationaux pour suivre la pêche illégale et prendre des mesures contre les acteurs malhonnêtes. Les gouvernements nationaux sont à la fois fournisseurs de données (aux côtés d'autres fournisseurs publics et privés de systèmes d'identification automatique et de systèmes de surveillance des navires) et utilisateurs au sein de l'initiative. Étant donné l'engagement direct avec les gouvernements et les demandes réglementaires spécifiques de chaque partenaire, il n'est pas toujours possible de créer des conditions d'engagement standard.

## Résolution des conflits



Même si le conflit est un aspect inévitable des partenariats inter-organisationnels et intersectoriels, cet examen des partenariats de partage de données a trouvé peu de mentions de mécanismes formels de résolution des conflits. Les experts consultés ont suggéré que les désaccords tendent à se concentrer sur l'établissement de l'initiative plutôt que sur sa mise en œuvre ou son fonctionnement.

Des pratiques telles que la tenue de réunions régulières avec les partenaires et la mise en place de processus décisionnels et d'organes de gouvernance clairs contribuent à la mise en place d'outils de gestion des conflits et garantissent que les problèmes qui surviennent au sein de l'initiative sont réglés dès leur apparition. Cependant, davantage de recherches et d'expérimentations sont nécessaires pour comprendre quels mécanismes de résolution des conflits sont optimaux pour les initiatives de partage des données dans le secteur du développement.

## Recettes pour communiquer avec le public et les parties prenantes

La communication peut faire ou défaire une initiative de partage de données. Pour que le partage des données soit digne de confiance, une communication réussie entre les partenaires de données et avec le public est primordiale. Les leçons tirées d'initiatives antérieures, telles que [Sidewalk Toronto](#), suggèrent que l'absence de communication claire et efficace avec le public peut conduire à la fin des efforts de partage des données, en raison d'un manque de confiance et d'une forte résistance des citoyens.

C'est pourquoi les nouvelles initiatives de partage des données doivent investir considérablement dans la mise en place des bonnes approches de communication avec les parties prenantes.

Tout comme l'instauration et le maintien de la confiance, la communication entre les partenaires est facilitée par les exemples de collaboration antérieure réussie. Lorsque de nouveaux canaux de communication doivent être établis,

des mécanismes tels que des consultations ou des ateliers de parties prenantes peuvent être bénéfiques au stade du lancement pour amorcer le dialogue. Les investissements dans la communication entre les partenaires ne peuvent toutefois pas être relégués à la phase de lancement et doivent se poursuivre tout au long de l'initiative.

Pour garantir une communication efficace, il convient de s'appuyer sur des **champions organisationnels** ou des gestionnaires de données ("data stewards" en Anglais), d'investir dans la **communication** tout au long de la durée de vie de l'initiative et d'établir un **dialogue avec la société civile** et les citoyens dès le départ et de manière récurrente.

## Champions organisationnels – responsables des données (“Data Stewards”)

Le partage des données nécessite souvent un changement d'état d'esprit culturel de la part des dirigeants. Il peut être difficile d'obtenir l'adhésion des organisations étant donné la sensibilité avec laquelle elles conservent leurs données, ainsi que le coût du partage des données en termes d'investissement en personnel et de financement. Le fait d'avoir un champion parmi les partenaires clés de l'initiative ou au sein du secteur, qui parle ouvertement des avantages potentiels du partage des données, peut aider à établir la crédibilité et le soutien.

Par exemple, le [California Data Collaborative](#), dont il est question dans la section précédente, a trouvé son champion en la personne du directeur général d'un district de distribution d'eau en Californie, qui a soutenu le projet et défendu la valeur des collaborations en matière de données dès le début. Son travail de sensibilisation auprès de ses pairs a permis à l'initiative de rallier d'autres organisations du secteur.

Le rôle de communiquer les avantages du partage des données peut également être confié aux [gestionnaires de données](#) (ou “Data Stewards” en Anglais), des personnes ou des équipes au sein des organisations détentrices de données qui sont habilitées à lancer, faciliter et coordonner de manière proactive les collaborations en matière de données dans l'intérêt du public. L'existence de gestionnaires de données est considérée comme [un facteur clé de succès](#) dans la mise en place de partenariats durables et responsables entre entreprises et gouvernements dans l'Union européenne. Ces conclusions sont transférables au partage des données dans le secteur du développement.





## Investissements dans la communication tout au long de la durée de vie de l'initiative

---

Le Hong Kong Data Trust 1.0, dont il a été question précédemment, a été créé pour promouvoir le partage des données entre les fournisseurs de transport afin d'informer les politiques. Pour répondre à la réticence des fournisseurs de transport à partager des données à valeur commerciale, le Data Trust a communiqué avec les parties prenantes pendant 18 mois avant le lancement de l'initiative par le biais de consultations, d'ateliers, de tables rondes, de cafés de la connaissance et de bulletins d'information réguliers.

Ces consultations ont permis aux acteurs de s'engager les uns envers les autres et de comprendre les cas d'utilisation autorisés du partage des données et l'échange de valeur qui en résulte. Elles ont également fourni aux parties prenantes un forum pour échanger et discuter de leurs préoccupations, et identifier clairement les cas d'utilisation autorisés pour l'initiative.

La plupart de ces canaux de communication ont perduré après la phase de lancement, garantissant que les partenaires poursuivent leurs échanges formels et informels.

## Établir un dialogue avec la société civile et les citoyens



L'analyse des initiatives de partage de données existantes dans le secteur du développement a montré que très peu d'entre elles investissent dans des activités de communication destinées aux communautés concernées ou au grand public. Dans de nombreux cas, les citoyens ne sont pas informés de l'utilisation de leurs données (même sous forme anonyme ou agrégée) dans le cadre de partenariats public-privé de partage de données. Cela a entraîné des réactions négatives de la part du public une fois les partenariats établis, comme le montre [le scandale](#) lié à un accord de partage de données entre les services de santé nationaux du Royaume-Uni et la société américaine Palantir.

Une communication efficace, claire et continue avec le public est souvent coûteuse et difficile à mettre en œuvre. Cependant, les enjeux de l'absence de communication sont élevés et les initiatives de partage des données doivent entreprendre cet effort, à la fois parce qu'il fait partie des processus de renforcement de la confiance et parce que

[l'information du public](#) est une première étape nécessaire pour permettre la responsabilisation et l'examen public.

Il existe peu d'exemples d'initiatives investissant du temps et des efforts dans la communication avec des publics plus larges, bien que certaines aient commencé à intégrer cette approche dans leurs activités. Par exemple, au cours de la phase de lancement d'un projet de partage de données en Gambie, Spatial Data Commons, en collaboration avec le Bureau gambien des statistiques et l'Autorité de régulation des services publics, a organisé [un atelier multipartite](#) pour le secteur privé, les ministères et agences gouvernementales, les universités, la société civile et les partenaires du développement. L'atelier avait pour but d'informer les parties prenantes sur le projet potentiel, de présenter le cas d'utilisation de l'accès aux données des opérateurs de réseaux mobiles et de permettre aux parties prenantes d'exprimer leurs préoccupations potentielles aux premiers stades du projet.

## Recettes pour la responsabilisation

Pour être efficace et digne de confiance, le partage des données doit être clairement ancré dans les valeurs sociétales et répondre aux normes éthiques et opérationnelles les plus élevées.

Au cours des deux dernières années, plus de 350 personnes et organisations de plus de 60 pays ont participé à un processus de consultation publique qui a conduit au lancement d'un [Manifeste #DataValues](#) appelant les gouvernements, les entreprises, les organisations de la société civile, les donateurs et d'autres acteurs à apporter des changements positifs dans la manière dont les données sont financées, conçues, gérées et utilisées. Ses cinq appels à l'action et les éléments recueillis pour le livre blanc qui le soutient, "[Repenser la relation entre les données et le pouvoir](#)", montrent que les praticiens du développement souhaitent réellement instaurer un partage des données éthique et fondé sur les valeurs partagées.

Les valeurs humaines et l'éthique peuvent être intégrées dans la gouvernance des initiatives et de leurs opérations de nombreuses manières différentes, notamment par le biais de **cadres éthiques** ou de **structures de gouvernance** spécifiques.

## Cadres et lignes directrices éthiques

Divers cadres et lignes directrices peuvent aider les initiatives de partage des données à s'interroger sur l'éthique de leur gouvernance et de leurs opérations. Certains sont d'application générale, tandis que d'autres sont adaptés à des types de données ou des contextes spécifiques.

L'Open Data Institute, par exemple, a développé le [Data Ethics Canvas](#), un outil permettant d'identifier et de gérer les questions éthiques dans les projets de données. Cet outil fournit un cadre pour élaborer des conseils éthiques adaptés aux initiatives de partage de données, quels que soient leur portée, leur taille ou leur secteur.

En 2021, la collaboration "Data for Children" a publié son [cadre d'évaluation éthique](#). Ce document reconnaît que les enfants présentent des vulnérabilités spécifiques lorsqu'il s'agit de données et que, dans ce contexte, "la réalisation d'un projet Data for Children éthique exige davantage qu'une simple conformité juridique." Le cadre fournit une feuille de route, comprenant une série de questions auxquelles il faut répondre au stade initial, pendant le projet et au stade final, dans le but de naviguer de manière éthique dans toutes les activités.

Le Centre de données humanitaires du Bureau des Nations Unies pour la coordination des affaires humanitaires (OCHA) a récemment élaboré des [lignes directrices sur la responsabilité des données](#), un ensemble de principes, de processus et d'outils pour soutenir la responsabilité des données dans le travail humanitaire d'OCHA. Ce document offre la norme minimale à suivre en matière de responsabilité des données au sein d'OCHA et reconnaît les grands risques qu'une gestion irresponsable des données fait courir aux communautés déjà vulnérables.

## Conseils et comités d'évaluation éthique

Depuis 2020, l'Alliance Data-Pop a inclus un [Conseil d'orientation du développement et de l'éthique](#), ou CODE, dans la plupart de ses projets. Un CODE est un groupe de parties prenantes indépendantes qui partagent volontairement leur expertise dans des domaines ayant une pertinence directe pour un projet. En tant que groupe consultatif, il assure une surveillance pour garantir qu'un projet respecte les principes éthiques clés, notamment l'utilisation équitable et sûre des données et les préoccupations locales et contextuelles. Les CODE sont généralement composés d'une dizaine d'experts locaux issus du monde universitaire, de la société civile, du gouvernement et de l'entreprise privée. Le rôle consultatif des membres du CODE n'est pas juridiquement contraignant. Néanmoins, les équipes de mise en œuvre des projets accordent un poids considérable à leurs commentaires et intègrent les recommandations des CODE chaque fois que cela est possible.

Lorsque Dalberg Data Insights a commencé à travailler avec le gouvernement belge sur l'utilisation de données de télécommunications anonymes et de données d'incidence localisées pour surveiller l'impact des mesures de confinement pendant la pandémie de COVID-19, elle a mis en place un [comité d'éthique](#) composé d'experts de différentes disciplines et de différents horizons. Ce groupe d'experts a suivi le projet tout au long de son cycle de vie et a été chargé de soulever toute question d'ordre éthique. Le rôle du comité était particulièrement important compte tenu à la fois de l'urgence de la situation et des contraintes qu'imposaient les restrictions du COVID-19 pour recueillir les réactions du public.

2

# Valeur et avantages partagés

## Fruits et légumes

Recettes de cette section

Réduire les coûts

Réaliser des produits d'innovation et de données

Fourniture de services sur mesure



## Introduction

Les fruits et les légumes contiennent des nutriments essentiels à l'épanouissement des personnes. De même, **le sentiment que tous les partenaires en bénéficient** proportionnellement est essentiel pour créer des partenariats efficaces et réduire le risque de menaces pour l'initiative. Si les humains peuvent vivre un certain temps sans fruits et légumes, leur absence alimentaire entraîne à long terme d'importants problèmes de santé. Si les organisations ne perçoivent pas la valeur d'un partenariat pour le partage des données, elles ne resteront pas longtemps à la table des négociations. **Le partage des données échoue lorsque la valeur et les avantages perçus sont inégalement répartis pendant de longues périodes.** Lorsque les avantages profitent systématiquement à un petit nombre plutôt qu'à un grand nombre, les incitations à participer diminuent.

Les avantages **vont au-delà de la simple valeur économique.** Les partenaires de partage des données peuvent tirer profit d'une initiative en termes de réduction des coûts, de réputation, de connaissances, de compétences ou même d'accès aux données générées par d'autres. Par conséquent, l'accent doit être mis non seulement sur la valeur monétaire du partage des données, mais aussi sur la compréhension de la portée plus large de la valeur ajoutée de l'initiative pour les partenaires de données. Des outils tels que [l'outil et le guide de cartographie des écosystèmes de données](#) développés par l'Open Data Institute dans le cadre de la [campagne Microsoft Open Data](#) peuvent aider à analyser la valeur potentielle qui peut être échangée entre les acteurs d'un écosystème de données.

Le partage équitable de la valeur est déjà compliqué dans le contexte du partage bilatéral des données entre le secteur public et le secteur privé, comme le [suggère la littérature](#), car il peut être difficile de concilier des intérêts et des attentes divergents, quel que soit le nombre de parties prenantes concernées. Dans le contexte du partage de données multipartite, il peut devenir encore plus complexe d'équilibrer des intérêts divergents.

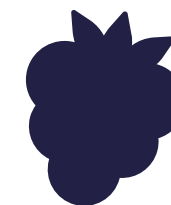
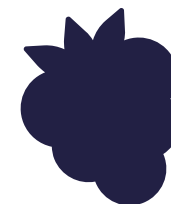


## Comprendre la valeur du partage des données



En termes économiques, le terme “valeur” est généralement utilisé pour désigner la “valeur ajoutée”, c’est-à-dire la différence entre les intrants et les extrants d’un certain produit. Dans le contexte du partage de données multipartite, cela se traduit souvent par un accent sur les avantages financiers qui sont relativement faciles à quantifier. Au niveau macroéconomique, l’Organisation de coopération et de développement économiques (OCDE), par exemple, suggère que l’amélioration de “l’accès et du partage des données publiques et privées peut contribuer à générer des avantages sociaux et économiques représentant entre 1 et 2,5 % du produit intérieur brut (PIB) (dans quelques études, jusqu’à 4 % du PIB)”. Au niveau microéconomique, une étude bien connue de Deloitte sur les impacts de la pratique du Transport of London (TfL) consistant à partager ouvertement des données non personnelles a montré que les entreprises utilisant les données de TfL ont généré une valeur ajoutée brute comprise entre 12 et 15 millions de livres sterling par an, y compris le soutien direct d’environ 500 emplois.

Cependant, de nombreux avantages du partage des données ne sont pas aussi faciles à mesurer. Pour le secteur public, des avantages sociaux divers et nombreux peuvent être obtenus grâce à des partenariats de partage de données. L’OCDE souligne en particulier les effets positifs sur la transparence et la responsabilité des gouvernants, ainsi que la responsabilisation accrue des utilisateurs, associés à un plus grand partage des données. Pour les entreprises, les avantages en termes de réputation ou de connaissances, comme nous l’avons vu plus haut, peuvent également servir de motivation pour s’engager dans des partenariats de données. L’OCDE souligne également la possibilité pour les fournisseurs de données du secteur privé d’obtenir de nouvelles idées et d’exploiter les innovations des utilisateurs liées à l’émergence d’une communauté qui crée une valeur supplémentaire qu’une organisation seule ne serait pas en mesure de créer.





## Qu'est ce qu'on cuisine?

**Le succès du partage des données dépend de l'établissement et du maintien d'une valeur partagée pour tous les partenaires.** Cela signifie que tous les partenaires bénéficient du partage des données dans une certaine mesure et qu'aucun partenaire n'en bénéficie de manière disproportionnée par rapport aux autres.

Cela dit, la répartition de la valeur n'est pas figée et peut changer au fil du temps, soit parce que les besoins et les attentes des partenaires de données changent, soit parce que l'initiative de partage des données évolue en termes d'objectif, d'activités et de niveau de contribution nécessaire des différents partenaires. Tout comme il peut être nécessaire d'ajuster la répartition de la valeur, il est également important que les initiatives soient claires quant à leur approche initiale et qu'elles soient en mesure de vérifier si les avantages promis se concrétisent pour les différents partenaires et dans quelle mesure. Cela leur permet de corriger les déséquilibres et de modifier les propositions de valeur en fonction des besoins.

Comme pour d'autres aspects du partage des données, il n'existe pas d'approche unique pour créer et distribuer la création de valeur. Les recettes visant à garantir que les avantages et la valeur sont distribués aussi équitablement que possible peuvent prendre de nombreuses formes, souvent adaptées aux besoins de partenaires spécifiques ou à des contextes et écosystèmes différents. Trois d'entre elles, en particulier, sont ressorties de l'analyse du paysage effectuée en vue de la préparation de ce livre de recettes : **la réduction des coûts, l'innovation et la fourniture de produits de données, et la prestation de services sur mesure.**



## Des recettes pour réduire les coûts

Réduire les frictions qui entravent le partage des données est un objectif clé pour de nombreuses initiatives multipartites. Le partage de données entre partenaires multiples se caractérise par des coûts de transaction importants, liés notamment au temps et à l'énergie nécessaires pour parvenir à un accord, le formaliser et le traduire en infrastructures appropriées et en mécanismes de coopération efficaces.

Les initiatives peuvent apporter de la valeur à leurs membres en réduisant ces coûts, ce qui permet également de réduire le temps nécessaire au partage effectif des données. Cela nécessite le plus souvent de développer des actifs ou des outils réutilisables qui peuvent être exploités pour accélérer la conclusion d'un accord et la mise en place de pipelines de données. Par exemple, le [Development Data Partnership](#) réduit le temps nécessaire à la conclusion d'accords et les coûts de transaction, tant pour les fournisseurs que pour les utilisateurs de données, en fournissant un modèle d'accord juridique commun (qui ne doit pas être négocié au cas par cas) et une infrastructure de partage et de stockage des données facilement accessible. Ces atouts sont prêts à l'emploi et peuvent être adoptés par de nouveaux partenaires potentiels sans longues négociations.

La réduction des coûts en tant que stratégie de création de valeur pour les partenaires peut également concerner les étapes du cycle de vie des données au-delà de la phase de partage. À titre d'exemple, la California Data Collaborative épargne à ses membres les coûts associés au recrutement de grandes équipes de scientifiques et de chercheurs en données, car ceux-ci sont embauchés et mis en commun par la Collaborative à un tarif inférieur pour toutes les agences de gestion de l'eau. La réduction des coûts est une stratégie visible et quantifiable qui permet aux initiatives de partage des données d'apporter de la valeur aux membres.

## Recettes pour réaliser des produits d'innovation et de données

Pour que le partage des données soit efficace, il faut que chaque partenaire ait une idée claire des objectifs, des cas d'utilisation, des données requises et des résultats attendus de l'initiative. Ceci est primordial pour définir la répartition de la valeur. Cependant, de nombreuses initiatives réussies innovent continuellement afin non seulement d'atteindre, mais aussi de faire évoluer et de renforcer leur proposition de valeur au fil du temps. L'expérience de ces initiatives suggère que (1) le partage des données peut être un catalyseur pour l'innovation mais aussi que (2) se concentrer sur l'innovation est souvent une incitation qui maintient l'engagement des partenaires et encourage de nouveaux partenaires à se joindre.

L'innovation en matière de données et les efforts de recherche dans le contexte du partage des données peuvent conduire au développement de **produits de données** (ou "data goods"), entendus comme des méthodes, des solutions et des outils produits par une organisation et

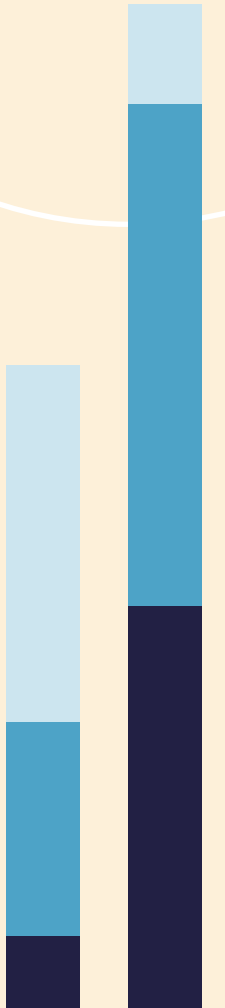
pouvant être utilisés avec succès par d'autres. Les initiatives créent souvent des produits de données qui sont accessibles à différents degrés, c'est-à-dire qu'elles peuvent donner un accès public gratuit à certains biens de données tout en restreignant l'accès à d'autres afin que seuls les partenaires de données puissent en bénéficier. Dans certains cas, en fait, la monétisation des produits de données est considérée comme une stratégie de viabilité financière, et il est donc dans l'intérêt de l'initiative de placer les produits de données derrière un mur payant. C'est l'approche adoptée par la California Data Collaborative : seuls les membres ont accès aux produits de données. Dans d'autres cas, lorsque d'autres mécanismes de viabilité financière sont en place, les produits de données sont considérés comme des biens publics et sont largement disponibles. En général, les produits de données font partie de la répartition de la valeur entre les partenaires lorsque leur accès est restreint, mais pas nécessairement lorsqu'ils sont en accès libre.



## Production et accès aux données

Le [Development Data Partnership](#) a créé et gère une place de marché pour différents types de produits de données. Cette place de marché comprend des dépôts de codes accessibles pour les produits de données et les algorithmes dérivés. Par exemple, des cartes temporelles (au lieu d'enquêtes auprès des ménages) ont été utilisées par la Banque mondiale en 2020 dans le cadre d'un projet visant à comprendre la [fracture numérique en Ukraine](#). Cette approche a été partagée en tant que produit de données sur la place de marché, puis reproduite par d'autres partenaires de l'initiative dans des projets en Amérique du Sud et en Afrique. Dans le cas du Development Data Partnership, les données sont partagées avec tous les partenaires mais ne sont pas mises à la disposition de la communauté du développement.

D'autres initiatives, conformément à leur objectif premier, accordent un accès universel à leurs données et produits. Par exemple, [Global Forest Watch](#), une plateforme en ligne qui fournit des données et des outils pour surveiller les forêts, offre un accès ouvert à divers outils et biens de données, tels que le nouveau Radar pour la détection de la déforestation, qui détecte les perturbations forestières à l'aide de données radar par satellite. Global Forest Watch a besoin d'une diffusion maximale des données pour atteindre son objectif de mettre fin à la déforestation illégale et aux activités non durables dans les forêts en donnant aux décideurs politiques et aux communautés concernées les moyens de surveiller l'état de leurs forêts et de développer des réponses politiques appropriées aux menaces. Restreindre l'accès aux biens de données serait contre-productif pour atteindre cet objectif ambitieux.



## Recettes pour la fourniture de services sur mesure

Si les produits ou biens de données décrits dans la section ci-dessus constituent des ressources utiles pour toutes les parties prenantes et les partenaires intéressés, les services sur mesure répondent aux besoins et exigences spécifiques des organisations.

Les services sur mesure comprennent la formation personnalisée, la réponse aux besoins spécifiques des partenaires en matière d'analyse ou de recherche, ou le développement de logiciels sur mesure. Ces services ont tendance à exiger beaucoup de ressources et peu d'initiatives de partage de données dans le secteur du développement les proposent aux partenaires. La recherche suggère que la fourniture de services personnalisés est plus typique dans les situations où le secteur privé est fortement impliqué et/ou dans les contextes où les partenariats de données prennent la forme de collaborations de données, comme dans le cas de la California Data Collaborative (voir ci-dessous).

Dans le secteur du développement, les **services sur mesure** constituent une stratégie visant à fournir de la valeur aux organisations de partage de données qui ne bénéficieraient pas autrement de la fourniture de biens de données ou d'autres mécanismes de distribution, comme les partenaires qui sont les plus gros contributeurs dans un partenariat de partage de données. Les services personnalisés peuvent inciter ces partenaires à rester engagés à long terme. Par exemple, [Global Fishing Watch](#), conscient du haut niveau de ressources requis par les gouvernements pour partager leurs données avec la plateforme, leur fournit des conseils personnalisés et répond à leurs questions politiques spécifiques afin de s'assurer que les partenaires de partage de données reçoivent un retour équitable.

## Fourniture de services de recherche sur mesure

Le [California Data Collaborative](#) offre une proposition de valeur claire à tous ses partenaires, en réduisant les coûts de transaction liés au partage des données et en réduisant le coût que les agences d'approvisionnement en eau individuelles encourent en engageant des talents indépendants en matière de données pour répondre aux exigences réglementaires que l'État de Californie a créées en 2015. En plus des services standard de partage de données et d'analyse qu'il fournit, le CaDC a travaillé à la mise au point d'un logiciel qui peut faciliter la gestion et l'analyse des données sur l'eau.

Le CaDC se caractérise également par l'accent mis sur les services de recherche personnalisés pour ses partenaires. Les partenaires peuvent venir au CaDC avec leurs propres questions de recherche de données, et l'équipe du CaDC les aidera à trouver la réponse. Par exemple, l'un des partenaires de la collaboration se demandait s'il devait investir dans des installations de stockage. L'équipe du CaDC a examiné l'utilisation actuelle de l'eau par l'agence d'approvisionnement en eau et a effectué des analyses de séries chronologiques pour prédire l'utilisation future de l'eau avant de conclure qu'un stockage supplémentaire n'était pas nécessaire. Cette décision a permis à l'agence d'économiser 20 millions de dollars.



3

# Données et infrastructures fiables

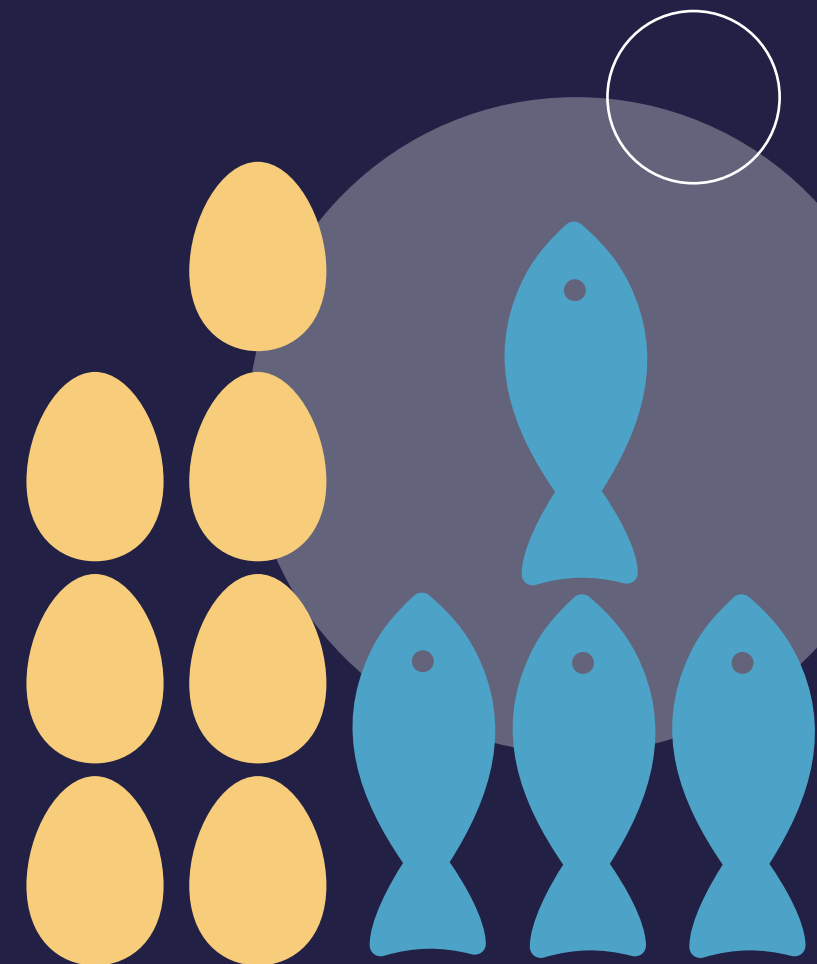
## Protéines

Recettes de cette section

Garantir la mise en place de protections techniques

Intégrer l'assurance qualité dans la collecte et l'analyse des données

Interopérabilité et le format des données



## Introduction

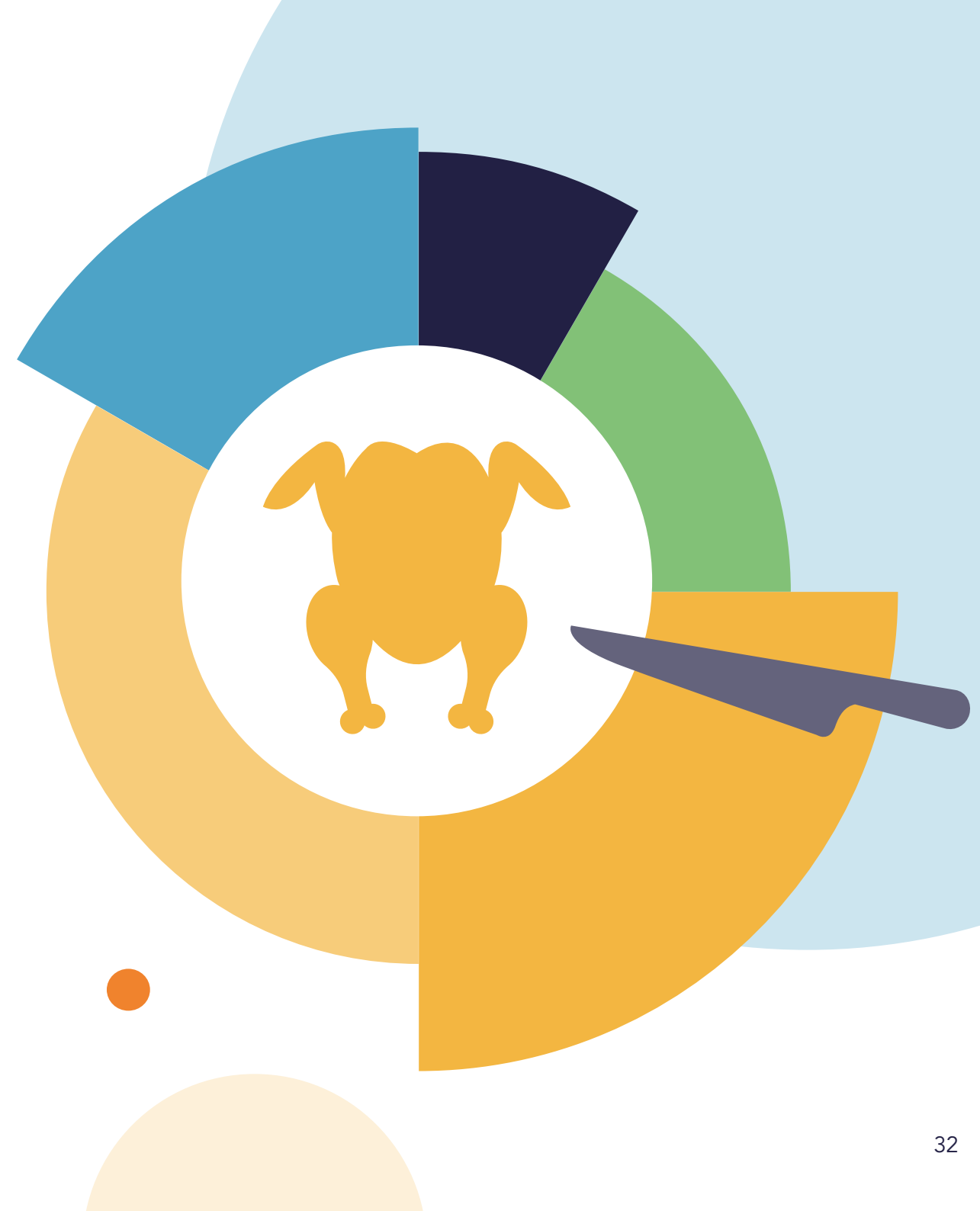
**Des données et une infrastructure de haute qualité (définies comme les logiciels et le matériel qui sous-tendent l'échange de données) sont au cœur d'un partage de données réussi.** On peut les comparer aux protéines, car elles sont généralement placées au centre de l'assiette et considérées comme le plat principal d'un repas.

Ce qui est moins évident, c'est que ce n'est pas la quantité de données partagées ni les caractéristiques technologiques de l'infrastructure qui conduisent finalement au succès des initiatives de partage de données, mais plutôt leur sûreté de fonctionnement.

La sûreté de fonctionnement englobe deux aspects des données et des infrastructures dans le partage des données : la fiabilité de ces systèmes et la qualité et l'adéquation des informations créées par rapport aux besoins des utilisateurs.

- ▶ **Les données** et les infrastructures fiables sont impartiales et régulièrement contrôlées.
- ▶ **La qualité et l'adéquation** sont liées à la sécurité et à l'interopérabilité des ensembles de données.

Ces deux concepts se renforcent mutuellement et garantissent que les partenariats de données débouchent sur des produits de grande qualité que les organisations peuvent utiliser pour développer de nouvelles applications, de nouveaux services, de nouvelles politiques et de nouvelles opérations.





## Qu'est ce qu'on cuisine ?

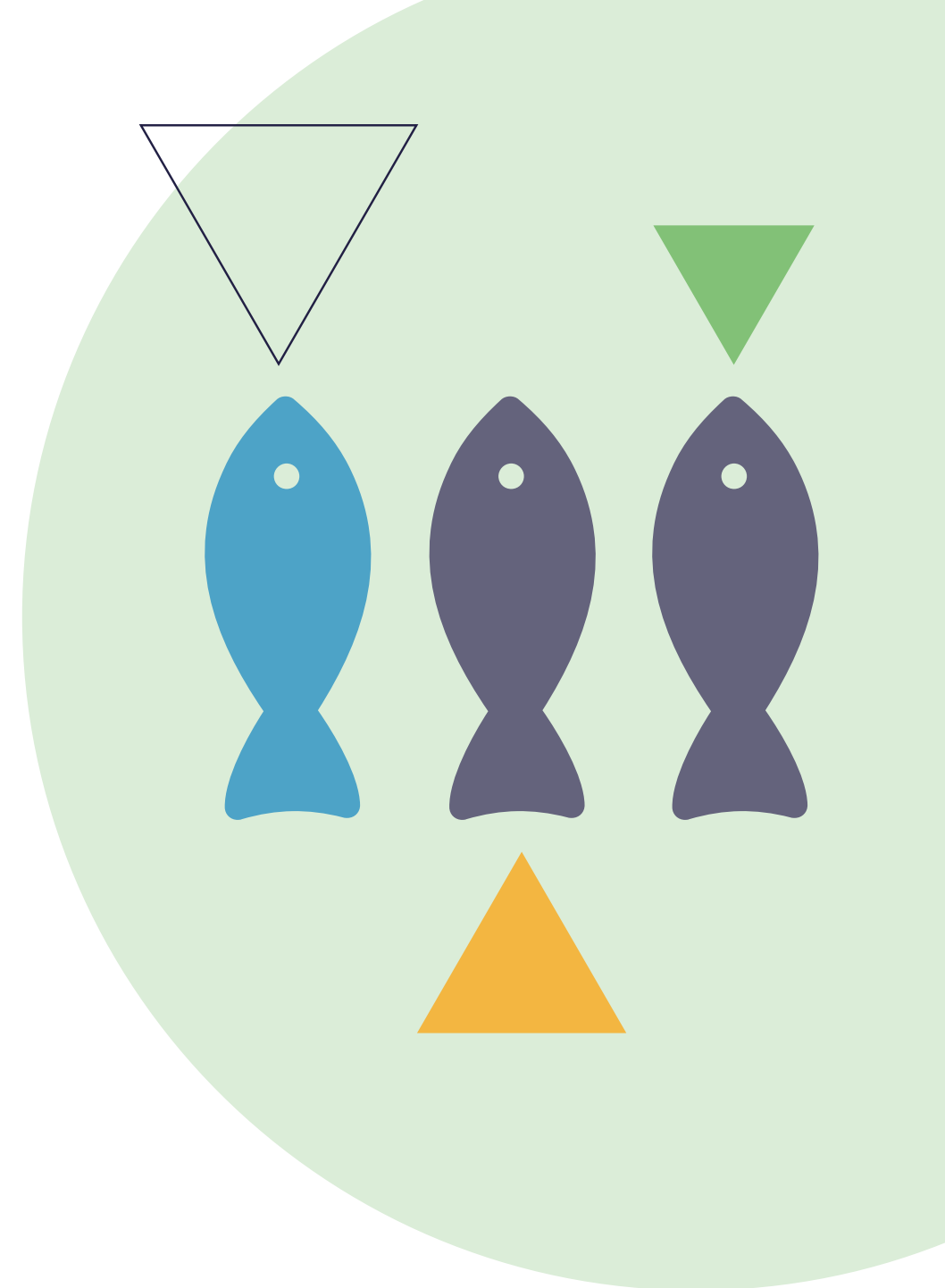
Les initiatives de partage de données peuvent adopter différentes approches pour garantir la sûreté de fonctionnement de leurs données et de leur infrastructure, mais généralement elles mettent l'accent sur les points suivants :

- ▶ Garantir la mise en place de protections techniques.
- ▶ Intégrer des mécanismes d'assurance qualité dans la collecte et l'analyse des données.
- ▶ Adopter des approches appropriées en matière d'interopérabilité et de format des données.

Le fait de mettre davantage l'accent sur l'un ou l'autre de ces éléments dépend principalement de la sensibilité des données en question, des caractéristiques des données elles-mêmes (c'est-à-dire en termes de variété des sources ou de granularité), des cas d'utilisation ciblés et des objectifs des initiatives, ainsi que des

ressources disponibles. **Les niveaux de confiance entre les partenaires influencent également l'attention portée aux garanties techniques et à l'assurance qualité en particulier.**

Par exemple, les agrégateurs de données tels que le [Humanitarian Data Exchange](#), qui est une plateforme web à code source ouvert et à accès libre permettant aux organisations humanitaires de partager des données, consacrent plus de temps et de ressources à assurer l'interopérabilité des ensembles de données, alors que des initiatives telles que [Global Fishing Watch](#), qui cherchent à informer les gouvernements, concentrent leurs efforts sur la qualité des données. En effet, la première collecte une grande variété de données provenant de différentes sources dont l'interopérabilité est essentielle pour rendre les données utilisables, tandis que la seconde a besoin de données très précises pour créer des cartes exactes des activités de pêche illégale.



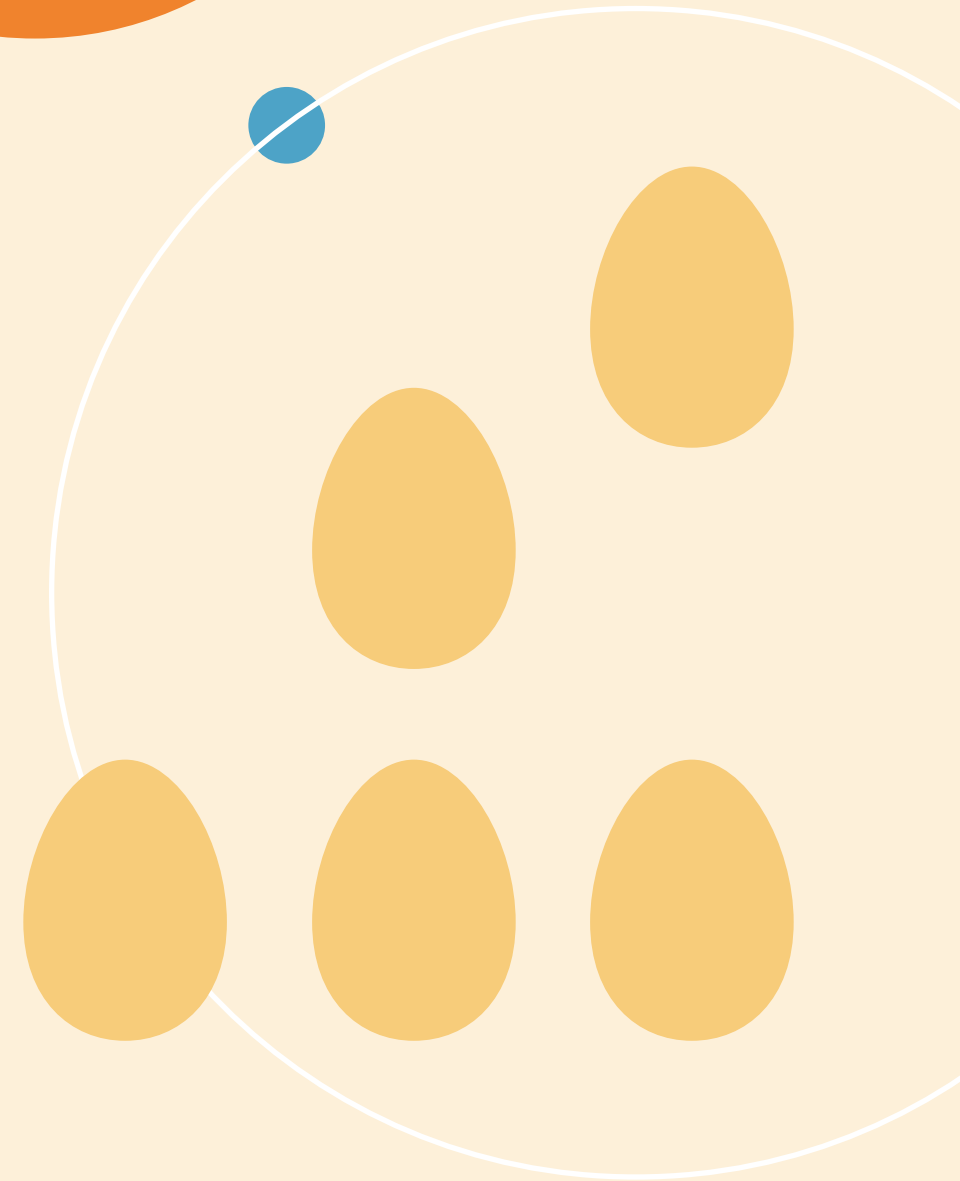
## Recettes pour garantir la mise en place de protections techniques

Les partenariats de partage de données se caractérisent par l'ouverture de l'accès aux informations et leur partage. Cependant, cela soulève des problèmes de sécurité, qui peuvent être particulièrement graves dans le contexte des données personnelles, définies comme **toute information qui se rapporte à une personne vivante identifiée** ou identifiable. Même si la plupart des initiatives réussies de partage de données dans le secteur du développement n'utilisent que des données anonymes, agrégées ou non personnelles, les fournisseurs de données restent préoccupés par la faillibilité des méthodes d'anonymisation et le risque de ré-identification.

Les politiques de sécurité et de stockage visent à répondre à ces préoccupations et constituent les principaux outils permettant de garantir la mise en place de protections techniques.

Les initiatives de partage de données peuvent se concentrer sur la sécurité des données en mettant en place les personnes, les processus et les outils nécessaires pour protéger la confidentialité et l'intégrité des données contre les attaques malveillantes ou les accidents involontaires tout au long du cycle de vie des données.

Les approches de **stockage sécurisé des données** peuvent être centralisées, fédérées ou distribuées, chacune utilisant des moyens différents pour assurer la sécurité des données concernées. Si la plupart des initiatives analysées s'appuient sur des systèmes de stockage centralisés, il n'existe pas d'approche unique et, dans le secteur du développement et au-delà, une tendance croissante à la décentralisation du stockage est liée aux préoccupations grandissantes concernant les déséquilibres de pouvoir et la thésaurisation des données.



## Création d'environnements de données contrôlés pour garantir la confidentialité des données

---

Depuis 2014, le défi [LinkedIn Economic Graph](#) invite les partenariats de recherche et d'analyse dans lesquels les données de LinkedIn peuvent être exploitées pour identifier les tendances macroéconomiques du travail et de l'économie.

Pour garantir la sécurité des données de ses membres, le défi fournit une formation approfondie en matière de sécurité à chaque équipe participante et exige que le travail soit effectué uniquement sur des ordinateurs portables fournis par LinkedIn sur le réseau LinkedIn dans un environnement de bac à sable surveillé. Les données téléchargées en dehors de ce réseau font l'objet de restrictions importantes et un collaborateur de LinkedIn supervise l'accès aux données et leur utilisation. L'utilisation des données est limitée à l'objectif spécifique identifié dans le partenariat de recherche. En outre, un comité d'examen interne évalue tous les produits de recherche créés dans le cadre du partenariat.

Des recherches récentes menées sur les données de LinkedIn, par exemple, montrent que les compétences vertes sont de plus en plus demandées sur le marché du travail, au moins [10 % des offres d'emploi](#) de l'année dernière les exigeant. Elles suggèrent également qu'un plus grand nombre de travailleurs acquièrent des compétences vertes et passent à des emplois verts ou dans des secteurs écologiques, ce qui entraîne des transitions nettes positives vers ces emplois.

## Passage d'un stockage centralisé à un stockage décentralisé des données

---

Le réseau [INSPIRE](#) (Implementation Network for Sharing Population Information from Research Entities) est un partenariat visant à partager les données des sites de surveillance sanitaire et démographique de cinq pays d'Afrique de l'Est. Au départ, INSPIRE a été mis en place pour créer un grand dépôt de données sur la santé, et les données ont été stockées de manière centralisée dans une installation Cloud, l'initiative, dirigée par le Centre africain de recherche sur la population et la santé, étant le dépositaire des données.

Toutefois, l'équipe INSPIRE, conformément à sa stratégie visant à accroître la capacité des fournisseurs de données et à renforcer les compétences des instituts de recherche participant à l'initiative, a l'intention de passer du statut de dépôt de données à celui de plateforme de services. Cela implique une évolution vers un système de stockage fédéré dans lequel les fournisseurs de données restent les gardiens de leurs propres données. La plateforme INSPIRE exploitera ensuite les données à distance pour des cas d'utilisation particuliers.

Comme le dit l'une des principales parties prenantes de l'initiative, l'objectif de cette transition est d'explorer une approche dans laquelle aucun partenaire ne détient toutes les données.

## Recettes pour intégrer l'assurance qualité dans la collecte et l'analyse des données

Pour être utilisables, les données doivent respecter des normes de qualité convenues. Les normes de qualité des données font référence au niveau d'exactitude, d'actualité, de précision et de fiabilité des performances. Chaque initiative de partage de données fixe ses propres normes de qualité et définit ce qui est acceptable et ce qui ne l'est pas en fonction des objectifs qu'elle se propose d'atteindre. Il est clair que certains cas d'utilisation (par exemple, dans le contexte des soins de santé ou de l'aide humanitaire) exigent des normes de qualité des données plus élevées que d'autres. Il est donc impossible d'établir une définition universelle des données "suffisamment bonnes".

Investir pour garantir le respect des normes de qualité des données convenues peut consommer du temps et des ressources. Toutefois, la mise en place d'un système à cet effet permet d'économiser des ressources à des stades ultérieurs des projets, car les erreurs et les biais peuvent être repérés avant que les données ne soient utilisées, ce qui évite des efforts coûteux pour corriger les erreurs une fois l'initiative ou la plate-forme établie.

L'adoption de **cadres de qualité** appropriés lors de la phase de collecte des données et la mise en place d'**approches transparentes pour limiter et atténuer les biais** lors de la phase d'analyse des données sont des étapes utiles pour améliorer la qualité des données.



## Définir les approches appropriées en matière de qualité des données

---

Les approches de la qualité et de la véracité des données varient selon les initiatives. Par exemple, les plateformes d'échange de données comme le [Humanitarian Data Exchange](#) (HDX) ne vérifient pas la qualité des données qu'elles reçoivent des partenaires. L'architecture de données du HDX n'est pas conçue pour nettoyer les données soumises. Cette initiative adopte une approche "d'acheteur averti", où la véracité des données est évaluée par l'utilisateur.

D'autres initiatives travaillent intensivement avec des partenaires de données pour garantir la qualité des données partagées. Les organisateurs de la [Global Fishing Watch](#) (GFW), par exemple, peuvent passer des mois à effectuer des contrôles de qualité des données reçues des gouvernements, car chaque pays communique ses données différemment. L'équipe du GFW standardise le format des données et vérifie l'absence d'erreurs. Les divergences telles que les champs de données manquants ou les fuseaux horaires erronés sont courantes, et l'équipe GFW travaille avec les gouvernements pour les corriger. Ce n'est que lorsque l'équipe GFW est convaincue de la qualité des données qu'elle passe à l'étape de l'analyse.

Une autre approche consiste à confier le nettoyage et le contrôle de la qualité des données aux fournisseurs de données. Ce point est généralement discuté au début de l'initiative, et les partenaires de données acceptent les exigences de format de données de l'initiative. Par exemple, [INSPIRE](#) exige que les partenaires effectuent le nettoyage des données, les contrôles de qualité et les mesures d'assurance qualité nécessaires avant le partage.

## Transparence pour l'atténuation et la limitation des distorsions liées aux données

En avril 2022, Haïti a été le théâtre de violences généralisées dues à des bagarres entre deux gangs. Le conflit a entraîné le déplacement d'environ 35 000 personnes de la zone touchée. [Flowminder](#) est une fondation à but non lucratif spécialisée dans l'analyse des Big Data, telles que les enregistrements des détails des appels, l'imagerie satellite et les enquêtes sur les ménages, afin de résoudre les problèmes de développement. Pour fournir plus de preuves et de détails sur le déplacement, elle a formé un partenariat de partage de données avec un opérateur de télécommunications représentant 74 % de la part de marché nationale.

L'objectif de ce partenariat était de générer des preuves pour mieux comprendre les mouvements à grande échelle de la population déplacée et soutenir la mise en place d'une réponse politique appropriée. Dans son [rapport final](#), Flowminder a fourni un avertissement détaillé sur les limites des données provenant des opérateurs de réseaux mobiles, qui ne sont pas statistiquement représentatives car l'accès aux téléphones n'est pas universel. Le rapport invite les lecteurs à tenir compte des limites des données lorsqu'ils tirent des conclusions à partir de l'étude.

## Recettes pour l'interopérabilité et le format des données

Le partage des données par différents partenaires et l'agrégation de différentes sources de données sur des plateformes communes nécessitent des investissements importants en termes d'**interopérabilité** et de format des données. Les données sont interopérables lorsqu'elles suivent des standards/**normes techniques** communément admises et qu'elles comportent suffisamment de **métadonnées** pour permettre à un analyste d'identifier la provenance, les concepts, les méthodes et les conseils d'utilisation.

Les standards et les métadonnées sont toutes deux nécessaires à l'interopérabilité. Sans standards, même les données bien décrites (c'est-à-dire avec des métadonnées) ne sont pas prêtes à être utilisées et, sans métadonnées, il est extrêmement long de comprendre les données ou de les utiliser.

Même au sein d'une même initiative, différents types de données peuvent nécessiter différentes approches en matière de standardisation et d'interopérabilité.

Par exemple, la [California Data Collaborative](#) reçoit deux types de données : les données de facturation et les données de participation aux programmes. Les premières nécessitent peu de standardisation, car le format de facturation est uniforme dans toutes les agences. En revanche, les formats des données de participation peuvent varier. Le CaDC a donc développé un logiciel que les agences partenaires peuvent utiliser pour standardiser les données de participation avant de les partager.

Si l'accent est souvent mis sur les standards, il est important de reconnaître que les initiatives réussies investissent fortement dans la production de métadonnées de haute qualité. De bonnes métadonnées peuvent faciliter le processus de transfert et d'utilisation des données. Afin de guider les fournisseurs de données, le [Humanitarian Data Exchange](#) a élaboré et mis à disposition des [instructions claires](#) sur les champs de ressources et de métadonnées requis et fournit un soutien aux nouveaux partenaires pour les appliquer.

## Assurer l'interopérabilité entre des ensembles de données disparates

Toute organisation humanitaire peut télécharger ses données sur le [Humanitarian Data Exchange](#). Afin d'assurer que le partage des données sur la plateforme soit transparent, HDX accepte les contributions dans n'importe quel format de données standard. Cependant, dès le début, HDX a reconnu qu'une certaine standardisation était nécessaire pour réduire les difficultés liées au partage des données.

Son approche originale de la standardisation des données nécessitait du temps et de l'expertise technique de la part des organisations humanitaires contributrices. Mais, en 2014, HDX a relancé cette approche pour développer le [langage d'échange humanitaire \(HXL\)](#). Comme la plupart des données sur HDX étaient stockées dans des feuilles de calcul, la nouvelle approche exigeait simplement que les données incluent un identifiant (un symbole utilisant un hashtag) dans le titre de la colonne, ce qui permet de fusionner les mêmes champs dans plusieurs ensembles de données. Si l'adoption de HXL pose encore des problèmes, les études menées par l'initiative de partage des données montrent qu'elle a considérablement amélioré l'utilisation des données et permis de gagner du temps dans leur traitement.



4

# Mécanismes d'appui au renforcement des connaissances et des compétences

## Produits laitiers

Recettes de cette section

Développement de la formation et du renforcement des capacités officielles

Développement informel des compétences et transfert de connaissances



## Introduction

Une caractéristique commune des initiatives de partage de données réussies est l'**accent mis sur les connaissances, les compétences et le renforcement des capacités**. Le partage des connaissances et des compétences renforce les accords de partage de données, tout comme le calcium contenu dans les produits laitiers aide à construire et à maintenir des os solides.

Il est largement reconnu au sein de la communauté du développement qu'il ne suffit pas de mettre les données à disposition et d'attendre que les organisations y accèdent ou les utilisent. Au contraire, il existe un consensus sur la nécessité de développer les compétences en matière d'utilisation des données au sein des communautés d'utilisateurs et sur l'importance de renforcer la culture des données pour parvenir à une utilisation durable des données.

Souvent, cependant, les utilisateurs ne sont pas les seuls à manquer de compétences en matière de données. L'analyse du paysage, par exemple, a souligné que les partenaires d'initiatives de partage de données réussies sont préoccupés par le manque de compétences en matière de données de leur propre organisation et par la difficulté de trouver et d'embaucher les bonnes personnes. En outre, le travail du Global Partnership a mis en évidence que les compétences

en matière de gouvernance des données, par exemple, font partie des principaux obstacles à l'établissement de partenariats public-privé.

Enfin, il ressort de ce travail qu'une attention limitée est accordée au renforcement des capacités au niveau communautaire. L'étude paysagère indique que presque aucun des partenariats de partage de données examinés ne s'engage directement avec les communautés d'où proviennent les données. Certains modèles avec des mécanismes d'engagement limités incluaient la participation d'organisations de la société civile et de représentants nationaux. Cependant, ils ne visaient pas explicitement à renforcer les capacités de ces communautés en matière de données.

## Qu'est ce qu'on cuisine ?

Pour réussir le partage des données, il faut **renforcer les compétences de la communauté des utilisateurs, ainsi que celles des principales parties prenantes et des partenaires.**

Pour cela, il est essentiel d'identifier l'ensemble des compétences nécessaires pour qu'un partenariat de partage des données atteigne ses objectifs et de le doter de personnes qui possèdent ces compétences.

En outre, les partenariats de partage de données doivent très bien connaître leurs utilisateurs et comprendre leurs obstacles et leurs exigences en matière d'accès aux données. Dans ce cadre, les partenariats doivent se concentrer sur la capacité des parties prenantes externes qui sont censées utiliser les données.

Enfin, pour rendre le partage des données plus équitable et plus durable, il faut des stratégies visant à accroître la capacité des communautés dont les données sont utilisées. Cependant, les exemples de la façon dont cela peut être fait dans la pratique sont limités, en partie en raison des coûts élevés des grands programmes d'alphabétisation des données qui sont perçus comme ne relevant pas de la portée des initiatives spécifiques de partage des données.

Pour accroître la capacité des parties prenantes externes et internes, les approches peuvent aller de la production d'outils de formation formels à la mise en place de mécanismes plus informels de transfert de connaissances. Quelques recettes réussies pour **renforcer les capacités institutionnelles** sont présentées ci-dessous.



## Recettes pour le développement de la formation et du renforcement des capacités officielles

La mise en place d'un renforcement des capacités et d'une formation institutionnalisés représente une activité centrale pour de nombreuses plateformes de partage de données réussies dans le secteur du développement. Les programmes de renforcement des capacités et de formation peuvent prendre différentes formes, de la formation virtuelle, comme dans le cas du [Humanitarian Data Exchange](#), aux bourses de données, comme le [Data Science Fellowship Program](#) du Global Partnership, en passant par des cadres et des documents d'orientation comme le [Open Data for Social Impact Framework](#) développé par la Microsoft Open Data Campaign.

Si l'accent mis sur le renforcement des capacités de partage des données est un facteur de réussite évident, son importance varie en fonction du type de parties prenantes impliquées dans les initiatives.

Par exemple, les initiatives de partage de données qui travaillent directement avec des organisations nationales et locales du secteur public pour éclairer la prise de décision sont plus susceptibles d'investir dans des activités

de renforcement des capacités de ces partenaires. Cela s'explique par le fait que l'objectif de l'initiative exige que la capacité à utiliser les données soit intégrée de manière permanente dans le fonctionnement de l'organisation du secteur public.

En outre, l'analyse du paysage suggère que les acteurs basés dans les pays du Sud sont plus amenés à se concentrer sur le renforcement des capacités locales. Dans l'analyse du paysage, les initiatives menées par les acteurs du Nord sont moins susceptibles de prendre en compte les avantages directs et à long terme pour les acteurs du Sud. Les entretiens indiquent que les acteurs basés dans les pays du Sud sont plus soucieux de remédier à ce déséquilibre des compétences en renforçant les capacités internes.

## Renforcement des capacités des organisations gouvernementales

---

Dans le contexte de la plateforme [Global Fishing Watch](#), les gouvernements sont à la fois fournisseurs de données et utilisateurs d'informations.

Cependant, tous les gouvernements ne sont pas au même niveau en termes de capacité à collecter et partager des données ou à utiliser les données disponibles sur la plateforme pour éclairer la prise de décision. C'est pourquoi GFW a adopté une approche consciente et structurée du renforcement des capacités.

GFW travaille avec les gouvernements nationaux pour s'assurer de leur capacité de base à collecter et à utiliser les données concernant l'activité humaine basée sur les navires. L'équipe de GFW collabore avec les responsables des données concernés pour nettoyer et normaliser les ensembles de données. En outre, une fois les données analysées, l'équipe travaille avec le service gouvernemental pour traduire les tendances des données afin d'identifier les mauvais acteurs et d'informer les mesures d'application.

## Renforcement des capacités dans le contexte des partenariats Sud-Sud en matière de données

---

Le réseau [INSPIRE](#) (Implementation Network for Sharing Population Information from Research Entities) vise à renforcer les capacités de ses partenaires et à les récompenser pour leur travail. Pour atteindre cet objectif, il fournit aux producteurs de données des outils et des services qui améliorent le partage des données, la reproductibilité et les processus de production de données auxiliaires. Ces outils aident également les producteurs de données à découvrir, accéder et utiliser de manière appropriée les données et métadonnées partagées. Dans les années à venir, INSPIRE vise à renforcer les compétences techniques internes des producteurs de données, ce qui leur permettra de participer à l'extraction, à la transformation et au chargement des données dans des modèles de données communs.



## Recettes pour le développement informel des compétences et transfert de connaissances

Le transfert de connaissances et le développement des compétences ne doivent pas nécessairement s'inscrire dans des programmes formels de formation ou de renforcement des capacités. Très souvent, l'apprentissage entre pairs ou les mécanismes informels peuvent accroître la capacité des partenaires ou des utilisateurs de données dans le contexte des initiatives de partage des données. En outre, l'adoption d'une approche institutionnalisée de la formation n'exclut pas d'autres mécanismes informels de transfert de connaissances. Au contraire, ces deux approches se complètent.

Il est ressorti de l'analyse du paysage que, même dans le contexte d'initiatives de partage de données réussies, chaque partenaire ne possédait pas toute la gamme de compétences nécessaires pour établir et maintenir un partage de données efficace au début de l'initiative. Au contraire, les partenaires possédaient souvent des compétences complémentaires qui, ensemble, ont conduit à un partage de données durable et réussi.

Si la diversité des compétences des partenaires constitue un bon point de départ pour le partage des données, à long terme, une initiative gagne à accroître la capacité des partenaires en partageant entre eux leurs connaissances et leurs forces uniques (par exemple leurs compétences techniques ou leur compréhension du droit).

Les mécanismes informels de développement des compétences et de transfert des connaissances peuvent prendre une multitude de formes différentes, allant des **réseaux d'apprentissage par les pairs** aux **centres de connaissances** et aux espaces de travail collaboratif comme GitHub et les plateformes de partage des ressources.

## Cafés des connaissances

---

Depuis sa création, le Hong Kong Data Trust 1.0 organise périodiquement des cafés des connaissances pour faciliter l'échange entre les partenaires de données et explorer des sujets d'intérêt dans un cadre informel. Ces forums sont ouverts à la communauté des parties prenantes de l'initiative de partage des données (composée de plus de 500 organisations du secteur des transports locaux) et au public. Les cafés des connaissances comprennent souvent des présentations et des conversations avec des experts externes.

Par exemple, en 2021, Data Trust 1.0 a organisé un café du savoir sur le thème des "[données sur la marche](#)" et a invité des experts du groupe de travail sur les normes internationales relatives aux données sur la marche et du réseau Walk21 à partager leurs idées sur les développements internationaux visant à accroître la collecte, l'accès et l'utilisation des données sur la marche.

Ces événements profitent à la communauté des parties prenantes et offrent une occasion d'échanger des connaissances et des informations entre partenaires de manière plus décontractée et en dehors du contexte des structures de gouvernance formelles. Cela contribue au renforcement des capacités et améliore la confiance entre les partenaires.

## Référentiels de connaissances

---

Les principaux objectifs de la [campagne Microsoft Open Data](#) étaient de faciliter les collaborations en matière de données et de créer des ressources communes pour soutenir le partage des données entre les organisations. Dès le départ, l'initiative a tenu à développer un référentiel de connaissances contenant des outils et des ressources pour favoriser le partage des données de manière plus générale.

L'initiative s'est associée à l'Open Data Institute et au GovLab de l'université de New York pour développer des ressources, des cours et des réseaux afin d'aider les organisations et les dirigeants à effectuer la transition vers des approches de données plus ouvertes et collaboratives. L'initiative a travaillé avec ses partenaires du savoir pour développer des ressources afin de fournir des conseils sur la gestion des données aux acteurs du secteur public et du secteur privé. Cet effort comprend la création d'une nouvelle académie de [gestion des données](#), un [projet de villes ouvertes](#) visant à créer une communauté et à partager des idées entre les villes qui utilisent des données pour stimuler le changement, et un [réseau d'apprentissage par les pairs](#), qui servira de communauté de pratique et d'apprentissage.

5

# Flexibilité et adaptabilité

## Graisses saines

Recettes de cette section

Intégrer la flexibilité dans la gouvernance,  
les opérations et l'architecture

Modèles financiers durables





## Introduction

Les écosystèmes de données dans lesquels opèrent les initiatives de partage de données ne sont jamais statiques, mais évoluent en permanence, notamment du point de vue de la technologie, des parties prenantes et de la culture. Il s'ensuit que les choix en matière de gouvernance, d'architecture et d'exploitation ne peuvent être gravés dans le marbre, et que **la possibilité de s'adapter ou de changer doit être envisagée dès le départ.**

Les approches financières et de financement ne peuvent pas non plus être immuables. La plupart des initiatives de partage de données commencent par le financement d'un projet spécifique, mais doivent passer à autre chose à un stade ultérieur, lorsque le financement du projet prend fin. Changer de modèle de financement peut signifier diversifier les sources de financement ou adopter des modèles d'abonnement ou de redevance. Quelle que soit l'approche choisie, cette transition d'un mode de financement à un autre reste un moment délicat qui peut briser même des initiatives très réussies (du point de vue des utilisateurs).

Des doses saines de flexibilité et d'adaptabilité sont donc aussi importantes dans le partage des données que des graisses saines dans l'alimentation humaine. Des doses raisonnables de flexibilité et d'adaptabilité sont essentielles pour que le partage des données réussisse à long terme et pour que les partenaires restent impliqués.

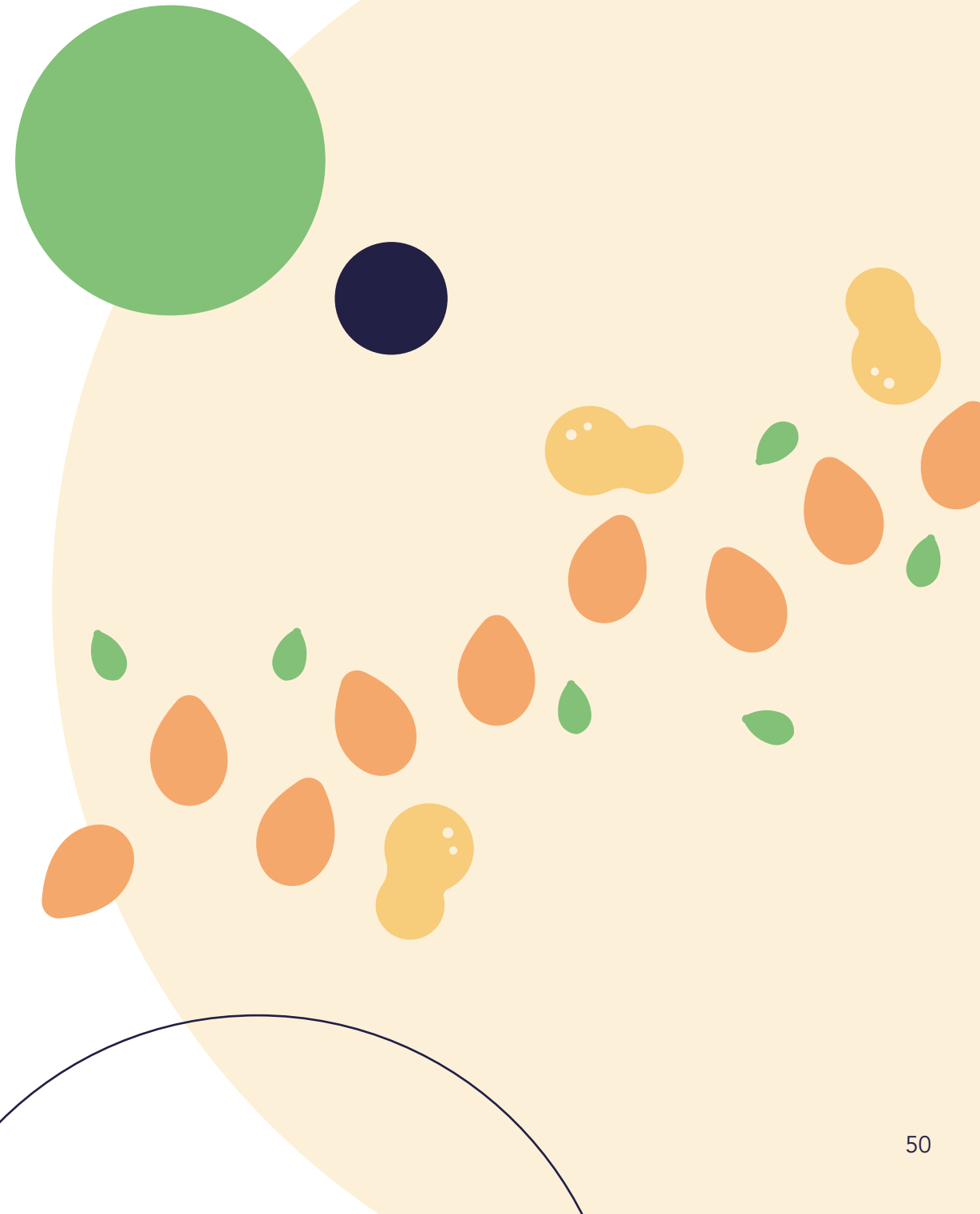


## Qu'est ce qu'on cuisine ?

Tout comme la confiance, la flexibilité et l'adaptabilité sont des termes génériques qui n'aident pas les organisations à déterminer ce qu'elles doivent faire exactement pour assurer le succès du partage des données. Plus concrètement, ce à quoi les initiatives de partage des données devraient prêter attention, c'est à la manière **d'intégrer la flexibilité dans leur gouvernance, leurs opérations et leur architecture de données, ainsi qu'à la manière d'identifier des modèles financiers durables** appropriés qui leur permettent d'évoluer dans le temps.

Malheureusement, le secteur du développement ne fournit pas beaucoup d'exemples de flexibilité intégrée dans les opérations, la gouvernance, l'architecture et le financement. Au contraire, ce domaine reste sous-exploré et représente un sujet pour lequel les recettes réussies font défaut.

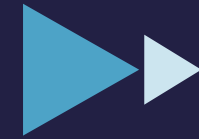
Cela ne veut pas dire qu'aucune initiative n'a réussi à inclure la flexibilité et l'adaptabilité, mais plutôt que les connaissances et les recherches sur la manière de le faire sont limitées. Même les initiatives réussies n'ont pas explicitement théorisé la manière d'assurer la mise en œuvre d'une approche évolutive. Les sections ci-dessous mettent en évidence quelques résultats et exemples de l'analyse du paysage, mais des lacunes subsistent en ce qui concerne les mesures concrètes à prendre par les partenaires de partage des données dans ce domaine.



## Recettes pour intégrer la flexibilité dans la gouvernance, les opérations et l'architecture

La flexibilité de la gouvernance et des opérations est difficile à atteindre en raison de la tension inhérente à la formalisation des modes de fonctionnement pour garantir la confiance des parties prenantes et à la création de marges de manœuvre pour la mise en œuvre des changements. La façon dont les organisations ont traditionnellement relevé ce défi consiste à adopter des processus de gestion du changement. Ceux-ci font référence à une série de tâches permettant une transition sans heurts d'un état actuel des choses à un nouvel état sans affecter les opérations de partage des données ni diminuer la confiance des parties prenantes. Les processus de gestion du changement peuvent être négociés au début d'un partenariat dans le cadre de conversations sur la gouvernance. Bien qu'aucun exemple de modèles de gestion du changement dans les partenariats de partage de données au sein du secteur du développement n'ait fait surface dans cette étude, les discussions avec les experts suggèrent qu'il s'agit d'une option pertinente pour les initiatives de partage de données qui sont construites pour durer.

D'un point de vue architectural, les experts s'accordent à dire que certains des plus grands défis sont l'évolutivité de l'infrastructure existante et la manière de financer les mises à jour des technologies de l'information (TI) à l'avenir. À cet égard, il existe un consensus sur l'importance de développer des plans informatiques évolutifs dès le début d'une initiative, en anticipant l'évolution des besoins informatiques en fonction, par exemple, de l'augmentation du nombre d'utilisateurs de données ou de l'émergence de nouvelles technologies et normes. À titre d'exemple, l'adoption des interfaces de programmation d'applications a nécessité la mise à jour d'une grande partie de l'infrastructure de données ouvertes existante pour s'adapter à une autre façon de consommer les données par les utilisateurs. Au cours des dernières années, les interfaces de programmation d'applications ont facilité l'accès et l'utilisation des données par les développeurs et les programmeurs, réduisant ainsi la nécessité de télécharger des fichiers.



Aucune personne ou initiative de partage de données ne peut prédire l'avenir. **Mais le fait de prévoir les mises à jour nécessaires de l'infrastructure** permet de clarifier les besoins de financement futurs et de s'assurer que l'initiative restera pertinente pour les parties prenantes et les utilisateurs.

## Reprogrammer les initiatives de partage de données pour faire face à de nouveaux défis

---

L'expérience de la [campagne Microsoft Open Data](#) montre combien il est important d'intégrer la flexibilité et le changement dans les collaborations de partage de données.

L'un des partenariats qu'elle a facilités visait à surveiller la qualité de l'air au Royaume-Uni et impliquait l'Alan Turing Institute – le centre national du Royaume-Uni pour la science des données et l'intelligence artificielle. Après l'apparition de la pandémie de COVID-19, ce projet a été rapidement réorienté pour aider les autorités londoniennes à mesurer l'impact des mesures de confinement sur la ville. Pour ce faire, l'Alan Turing Institute a déployé l'infrastructure de l'étude sur la qualité de l'air pour mesurer l'affairement de la ville et la réaction du public aux interventions gouvernementales.

Bien que ce ne soit pas l'intention initiale de la collaboration en matière de données, les flux de données avaient déjà été mis en place et l'infrastructure de données a pu [générer des informations](#) pertinentes pour la crise du COVID. Les niveaux élevés de confiance entre les partenaires de données permettent une telle flexibilité.

## Des recettes pour des modèles financiers durables

La capacité d'une initiative de partage de données à collecter des fonds et/ou à générer des revenus a un impact sur la durabilité de ses opérations et de ses investissements dans la croissance continue. Assurer la durabilité financière et explorer un ensemble diversifié de modèles de financement était une préoccupation commune à la plupart des initiatives de partage de données incluses dans l'analyse du paysage.

Alors que la majorité des initiatives restent financées par des donateurs, toutes ont reconnu l'importance de diversifier les sources de financement et/ou d'intégrer des modèles de revenus dans leurs opérations. La plupart des initiatives ont exprimé leur intérêt à passer d'une approche basée sur les donateurs à une approche génératrice de revenus (ou partiellement génératrice de revenus), mais il y a peu d'exemples de transitions réussies. De même, il n'y a pas beaucoup d'exemples d'initiatives qui ont réussi à diversifier l'origine de leur financement ou à assurer la pérennité sur la base de cette approche.

Parmi les **modèles de financement** actuellement à l'étude figurent les frais d'abonnement, les frais pour les services de données offerts et les frais de stockage/traitement des données. Toutefois, ces modèles n'en sont qu'à leurs débuts, et leur évolutivité et leur durabilité ne sont pas encore claires.

## Adopter des modèles financiers appropriés

La [California Data Collaborative](#) est un exemple d'initiative autofinancée basée sur un modèle générateur de revenus. La plupart des coûts opérationnels du CaDC sont couverts par la cotisation annuelle versée par les membres. Si le CaDC est donc plus coûteux qu'une association professionnelle, il est moins cher que le coût que chaque organisme devrait supporter pour embaucher un gestionnaire de données ou un développeur de logiciels. Le CaDC reçoit également des subventions supplémentaires pour financer ses projets de recherche.

Le [Humanitarian Data Exchange](#), quant à lui, s'appuie sur le financement de base des partenaires clés de l'initiative. Initialement, en 2014, HDX a reçu un financement de projet du programme d'innovation et de preuves humanitaires du ministère britannique du développement international, du gouvernement suédois et du Fonds d'innovation humanitaire. Aujourd'hui, la majeure partie du financement de HDX est assurée par le Bureau de la coordination des affaires humanitaires des Nations unies, ce qui permet la continuité des opérations de la plateforme. HDX recherche également des financements supplémentaires pour

des projets spécifiques (mises à jour informatiques) auprès d'autres donateurs, mais la disponibilité de ces fonds n'affecte pas le fonctionnement de base de l'initiative. La demande de financement auprès de divers donateurs a permis à HDX d'affiner sa vision et d'obtenir un retour constructif. La diversification des sources de financement pour les activités principales et supplémentaires permet d'atténuer les risques financiers et de puiser dans des ressources supplémentaires en cas de besoin.

Enfin, il est toujours possible de modifier l'approche de financement pendant que l'initiative est en cours. Le [Data for Children Collaborative](#), par exemple, est actuellement financé par le Scottish Funding Council et le Data Driven Innovation Programme et recevra ce soutien financier pendant les trois premières années de son existence. Cependant, le collaboratif a pour objectif de devenir un cabinet de conseil qui construit des collaborations de données en tant que service. Cela impliquera de passer d'un financement basé sur des donateurs à des approches génératrices de revenus dans un avenir proche.

# Conclusions

Le nombre florissant d'initiatives de partage de données multipartites au sein du secteur du développement témoigne de l'importance de la collaboration en matière de données pour la réalisation des Objectifs de développement durable. Si l'augmentation du partage des données est une bonne nouvelle, les nouvelles plateformes de données en elles-mêmes ne suffiront pas à garantir la réalisation de l'agenda 2030. Les initiatives de partage de données comme celles décrites dans ce livre de recettes ne soutiendront pas les objectifs de la communauté du développement si elles ne sont pas efficaces. Un partage efficace des données au sein de ce secteur nécessite davantage de partage des connaissances afin d'identifier les meilleures pratiques pouvant être reproduites dans le monde entier.

C'est précisément l'objectif des recettes proposées dans ce livre de recettes. Celui-ci jette les bases de discussions sur les ingrédients clés d'un partage efficace des données et fournit un certain nombre de recettes à titre d'inspiration. Ce n'est qu'un début. Il reste encore beaucoup à faire pour combler les lacunes dans les connaissances existantes, en raison notamment des limites de ce travail et de la relative nouveauté de nombreuses initiatives de partage des données, en particulier dans le secteur du développement.

## Contribuer à ce livre de recettes

Êtes-vous impliqué dans la mise en œuvre d'initiatives de partage de données dans ce secteur ? Il y a des leçons à tirer des initiatives réussies et moins réussies. Vos contributions sont les bienvenues et nous vous invitons à partager avec notre communauté les ingrédients et les recettes d'un partage efficace des données.

Si vous avez une recette ou une suggestion à partager, n'hésitez pas à nous faire part de vos commentaires par le biais de ce [formulaire Google](#), et nous vous contacterons pour savoir comment partager vos connaissances avec la communauté.

# Annexe A

Ce livre de recettes s'appuie sur l'analyse du paysage commandée par le Global Partnership et réalisée par Athena Infonomics entre mars et septembre 2022. Le rapport final de l'analyse est disponible [ici](#) et son résumé exécutif [ici](#). Les lecteurs intéressés peuvent également trouver les résumés des études de cas [ici](#).

Le Global Partnership est reconnaissant à l'équipe d'Athena Infonomics et d'Atlas AI (Shruti Viswanathan, Deepa Karthykeyan, Vivek Sakhrani et Saiyed Kamil) pour leur travail pertinent et perspicace.





# Annexe B

Le tableau ci-dessous contient la liste des membres composant le Groupe de référence pour le partage efficace des données, par ordre alphabétique.

**Alewaert, Caroline**

Spécialiste de la confidentialité des données  
[UN Global Pulse](#)

**Arai, Ayumi**

Professeur assistant de projet et  
co-responsable du Spatial Data Commons  
[Centre pour la science de l'information  
spatiale, Université de Tokyo](#)

**Barrett, Alyssa**

Stratège produit senior,  
Land & Carbon Lab  
[World Resources Institute](#)

**Calderon Machicado, Claudia**

Responsable des partenariats, Partenariat  
pour les données de développement  
[Banque mondiale](#)

**Cecconi, Gianfranco**

Responsable des écosystèmes  
de données collaboratives  
[Capgemini Invent](#)

**Clause, Victoria**

Responsable de l'agriculture  
intelligente face au climat  
[Mercy Corps AgriFin](#)

**Cooper, Sonia**

Avocat général adjoint, Équipe  
Innovation ouverte  
[Microsoft](#)

**Díaz, Alexander**

Directeur principal, IA pour le bien  
social et la réponse aux crises  
[Google.org](#)

**Hodes, Cyrus**

Cofondateur et président de l'AI Initiative  
[The Future Society](#)

**Hoffman, William**

Ancien chef de projet, Initiative  
"Data for Common Purpose"  
[Forum économique mondial](#)

**Jansen, Ronald**

Sous-directeur et chef du Service de  
l'innovation et de la capacité des données  
[Division de la statistique des Nations Unies](#)

**Kiragga, Agnes**

Chercheur en sciences des données  
[Centre de recherche sur la population  
et la santé en Afrique](#)

**Krambeck, Holly**

Directeur de programme et fondateur,  
Development Data Partnership  
[Banque mondiale](#)

**Maloney, Christopher**

Chargé de programme, Programme  
d'équité entre les sexes et de gouvernance  
[Fondation William et Flora Hewlett](#)

**Orrell, Tom**

Directeur  
[DataReady](#)

**Powell, Joshua**

Directeur général  
[Development Gateway :](#)  
[Une entreprise de l'IREX](#)

**Scrollini, Fabrizio**

Directeur exécutif  
[Open Data Latin American Initiative](#)

**Schüür, Friederike**

Ancien boursier, éthique de l'IA  
et gouvernance numérique  
[UN Global Pulse, maintenant UNICEF](#)

**Sellitto, Michael**

Affiliate Fellow  
[Institut de Stanford pour l'intelligence  
artificielle centrée sur l'homme](#)

**Teran, Javier**

Statisticien  
[Bureau de la coordination des affaires  
humanitaires des Nations Unies](#)

**Ure, John**

Directeur de l'Institut Fair Tech  
[AccessPartnership](#)

**Villarreal, Pablo**

Chef de la plateforme numérique nationale  
[Secrétariat exécutif du système national  
anticorruption, gouvernement mexicain](#)

**Vollenweider, Xavier**

Directeur des partenariats pour  
les données mobiles  
[Flowminder](#)



Ce document a été rédigé par Martina Barbero avec le soutien de Janet McLaren et révisé par Jenna Slotin, toutes membres de l'équipe du Global Partnership for Sustainable Development Data (Global Partnership).

Ce projet a été rendu possible grâce à une subvention de Google.org. Les auteurs remercient sincèrement l'équipe de recherche d'Athena Infonomics et d'Atlas AI (Shruti Viswanathan, Deepa Karthykeyan, Vivek Sakhrani et Saiyed Kamil), qui a réalisé l'analyse du paysage dont s'inspire ce document (voir annexe A), ainsi que les experts du groupe de référence pour le partage efficace des données : au delà des plateformes, qui ont consacré leur temps et leurs commentaires à l'élaboration des conclusions contenues dans ce document. La liste des membres de ce groupe figure à l'annexe B.

Avec le soutien de  
**Google.org**