BROADBAND COMMISSION
FOR SUSTAINABLE DEVELOPMENT

Manuel de gouvernance des données

Institutions, principes et processus pour une politique responsable

Les idées et les opinions exprimées dans cette publication sont celles des auteurs ; elles ne reflètent pas nécessairement les points de vue de l'UNESCO, de l'UIT, du PNUD, de la Commission de l'Union africaine (CUA) et de la Commission sur le large bande pour le développement durable. Ces organisations ne sont pas responsables du contenu et n'approuvent pas les opinions spécifiques qui y sont exprimées.

Le présent rapport a été préparé, avec le soutien d'experts tiers, par les membres du groupe de travail sur la gouvernance des données pour la Commission sur le large bande pour le développement durable coprésidé par l'UNESCO, le PNUD, l'UIT et l'Union africaine. La traduction a été réalisée par Juliette Cribier.

Les désignations employées dans cette publication et la présentation des données qui y figurent n'impliquent de la part de l'UNESCO, de l'UIT, du PNUD, de la CUA et de la Commission sur le large bande aucune prise de position quant au statut juridique des pays, territoires, villes ou zones, ou de leurs autorités, ni quant au tracé de leurs frontières ou limites.

Le présent rapport du Groupe de travail n'engage pas la Commission sur le large bande pour le développement durable ni aucun de ses membres ou organisations partenaires.

CI/DIT/2025/DT/01

BROADBAND COMMISSION
FOR SUSTAINABLE DEVELOPMENT

ITU

unesco

Manuel de gouvernance des données

Institutions, principes et processus pour une politique responsable

Juillet 2025

Remerciements

Le présent rapport a été rédigé en collaboration, en s'appuyant sur les idées et les contributions des commissaires et de leurs organisations, dont la liste figure ci-dessous. Par conséquent, les opinions qui y sont exprimées ne sont pas attribuées à une organisation ou à un individu en particulier. Le présent rapport a été rédigé par Stefaan Verhulst, rédacteur technique principal et expert externe, avec la contribution de Leona Verdadero de l'UNESCO. Le projet a été présidé par l'UNESCO, l'UIT, le PNUD et la Commission de l'Union africaine.

La révision éditoriale a été assurée par Guilherme Canela de Souza Godoi, Cédric Wachholz, Guy Berger, Jaco du Toit, David Uribe, Tarja Turtia, David Castillo, Emma Fischer, Lucia Bosio et Kamel El Hilali de l'UNESCO ; Philippa Biggs et Nancy Sundberg de l'UIT ; Alena Klatte et Alper Gucumengil du PNUD ; Souhila Amazouz de la Commission de l'Union africaine ; et Begoña Otero, Adam Zable et Roshni Singh du Governance Lab (GovLab). La production du rapport a été supervisée par l'UNESCO, sous la direction de Leona Verdadero, de Lucia Bosio, de Sonia Savci et d'Emma Fischer.

Nous tenons également à remercier le Secrétariat de la Commission sur le large bande de l'UIT, c'est-à-dire Nur Sulyna Abdullah, Anna Polomska et Julia Gorlovetskaya, pour son soutien inestimable tout au long de l'élaboration de ce rapport.

Table des matières

Remerciements 4

Liste des tableaux, schémas et encadrés 6

Membres du Groupe de travail 7

Acronymes et abréviations 8

Introduction au Manuel 9

Introduction à la gouvernance des données 12

Liste de contrôle et Auto-évaluation de la gouvernance des données 18

Éléments de base pour l'élaboration d'un cadre de gouvernance des données 26

I. **POURQUOI :** Définir la vision et la finalité des données et leur gouvernance 28

II. **COMMENT :** Définir les principes de gouvernance des données 34

III. **QUI :** Les personnes et les processus dans la gouvernance des données 46

IV. **QUOI :** Les politiques, les pratiques et les technologies qui régissent chaque étape du cycle de vie des données, en veillant à ce que ces dernières soient traitées dans une finalité précise et dans le respect des principes directeurs. 51

1. Planification 54

2. Collecte 57

3. Traitement 60

4. Partage 63

5. Analyse 68

6. Utilisation 71

Glossaire 74

Bibliographie sélective sur la gouvernance des données 79

Annexes : Listes de contrôle 85

Liste des tableaux, schémas et encadrés

Tableaux

Tableau 1. Cartographie des Manuels de gouvernance des données 27

Tableau 2. Liste de contrôle : Finalités potentielles des données et de leur gouvernance 31

Tableau 3. Liste de contrôle : Principes relatifs aux données et à leur gouvernance 39

Tableau 4. Cartographie des parties prenantes 49

Tableau 5. Liste de contrôle : Planification 54

Tableau 6. Liste de contrôle : Collecte 57

Tableau 7. Liste de contrôle : Traitement 60

Tableau 8. Liste de contrôle : Partage 63

Tableau 9. Liste de contrôle : Analyse 68

Tableau 10. Liste de contrôle : Utilisation 71

Schémas

Schéma 1. Cadre de gouvernance des données 14

Schéma 2 : Cycle de vie des données 15

Encadrés

Encadré 1. Comprendre le cycle de vie des données 15

Encadré 2. Principe émergent : Qu'est-ce que l'autodétermination numérique (ADN) ? 38

Encadré 3. FOCUS : Principes de gouvernance des données du Conseil des chefs de secrétariat (CCS) de l'ONU pour la coordination 41

Encadré 4. 10 Mécanismes de gouvernance des données 52

Encadré 5. Différents types de données 56

Encadré 6. Collaborations sur les données 65

Encadré 7. FOCUS : Emplacement des données dans l'écosystème mobile 66

Encadré 8. Infrastructure publique numérique (IPN) 67

Encadré 9. Intelligence artificielle (IA) et Gouvernance des données 69

Encadré 10. Considérations transversales sur le cycle de vie des données 73

Membres du Groupe de travail

Co-présidents du Groupe de travail

Mme Audrey Azoulay, Organisation des Nations Unies pour l'éducation, la science et la culture (représentée par M. Cedric Wachholz et Mme Leona Verdadero)

M. Achim Steiner, Programme des Nations Unies pour le développement (représenté par Mme Alena Klatte et M. Alper Gucumengil)

Mme Doreen Bogdan-Martin, Union internationale des télécommunications (représentée par Mme Phillippa Biggs et Mme Nancy Sundberg)

S. Exc. Mme Lerato Mataboge, Union africaine (représentée par Mme Souhila Amazouz)

Commissaires et points focaux

La baronne Beeban Kidron, Fondation 5 Rights (représentée par Mme Marie-Ève Nadeau et Mme Nicola White)

S. Exc. Deemah Al Yahya, Organisation de coopération numérique (représentée par M. Hassan Nasser et M. Ahmad Bhinder)

Dr Hessa Al Jaber, Es'Hailsat

M. Piotr Dmochowski-Lipski, EUTELSAT IGO

Dr Qu Dongyu, Organisation des Nations Unies pour l'alimentation et l'agriculture (représenté par M. Máximo Torero, M. Henry Burgsteden et M. Cristopher Marchini)

S. Exc. Jessica Rosenworcel, Commission fédérale des communications (représentée par Mme Roxanne McElvane Webber)

M. Mats Granryd, GSM Association (représenté par M. Noriswadi Ismail, M. Luca Elmosi et M. Roddy McGlynn)

Dr Rumman Chowdhury, Humane Intelligence

M. Chuen Hong Lew, Infocomm Media Development Authority, Singapour (représenté par M. Bertrand Chew et Mme Evelyn Goh)

Mme Pamela Coke-Hamilton, Centre du commerce international (représentée par M. James Howe et M. Gilles Chappell)

M. Lacina Koné, Smart Africa (représenté par Mme Aretha Mare, Mme Thelma Efua Quaye et Mme Denyse Ntaganda)

S. Exc. l'ingénieur Majed Al Mesmar, Autorité de réglementation des télécommunications et du gouvernement numérique, Émirats arabes unis (représenté par M. Humaid Ali Al Basti et M. Abdul Rahman Al Marzouqi)

Mme Sima Sami Bahous, ONU-Femmes (représentée par M. Papa Seck, Mme Jessamyn Encarnacion, Mme Helene Molinier et Mme Ramya Emandi)

M. Filippo Grandi, Haut-Commissaire des Nations Unies pour les réfugiés (représenté par M. Conor Flavin)

M. Ziyang Xu, Zhongxing Telecommunication Equipment (représenté par M. Zhi Cheng Qu et M. Dao Tian)

Experts externes

Dr Stefaan Verhulst, The GovLab

Dr Jennifer Prendki, Quantum of Data

Dr Lucie-Aimée Kaffee, Hugging Face

M. Bart Rosseau, Digital Flanders

Acronymes et abréviations

ADN	Autodétermination numérique
ANS	Accords de niveau de service
APD	Accords de partage de données
ASEAN	Association des nations de l'Asie du Sud-Est
Principes CARE	Intérêt collectif, Droit de regard, Responsabilité et Éthique
CCS de l'ONU	Conseil des chefs de secrétariat des Nations Unies
CCT	Clauses contractuelles types
DAMA-DMBOK	Ensemble de connaissances en matière de gestion des données
DPD	Délégué(e) à la protection des données
EICD	Études d'impact sur la confidentialité des données
EIDH	Études d'impact sur les droits de l'homme
EIVP	Études d'impact sur la vie privée
Principes FAIR	Facilement Trouvable, Accessible, Interopérable et Réutilisable
FTD	Flux transfrontaliers de données
HCR	Haut-Commissariat des Nations Unies pour les réfugiés
IA	Intelligence artificielle
IPA	Interface de programmation d'applications
IPN	Infrastructure publique numérique
MA	Mémorandums d'accord
MTP	Mesures techniques de protection
OCDE	Organisation de coopération et de développement économiques
ODD	Objectif(s) de développement durable
PNUD	Programme des Nations Unies pour le développement
RD	Responsable des données
TI	Technologie de l'information
TIC	Technologie de l'information et de la communication
UA	Union africaine
UIT	Union internationale des télécommunications
UNESCO	Organisation des Nations Unies pour l'éducation, la science et la culture

Introduction au Manuel

Dans un monde de plus en plus dominé par les données, la capacité à gouverner ces dernières de manière responsable, efficace et au bénéfice de tous constitue désormais un défi majeur pour les gouvernements, les institutions et les sociétés. Qu'il s'agisse de l'action climatique, des systèmes de santé, des transports, de l'éducation ou du développement de l'IA, les données jouent un rôle central dans la création de valeur publique. Pourtant, la plupart des pays et des institutions restent confrontés à des lacunes dans la manière dont ils gouvernent les données, en articulant les droits, les responsabilités et la réutilisation à travers des systèmes fragmentés.

Le présent Manuel de gouvernance des données a été élaboré par la Commission sur le large bande pour le développement durable, par l'intermédiaire de son Groupe de travail sur la gouvernance des données, présidé par l'UNESCO et coprésidé par l'UIT, le PNUD et l'Union africaine. Il est le fruit d'une collaboration entre experts, décideurs politiques, praticiens, représentants de la société civile et entreprises qui se sont engagés à promouvoir des approches de gouvernance des données équitables et fondées sur les droits de l'homme.

Le présent Manuel a également bénéficié du retour d'information et des idées recueillies à l'occasion de consultations régionales, garantissant ainsi une réponse aux divers défis et opportunités auxquels les pays sont confrontés dans les différentes régions. Ces consultations ont contribué à façonner le Manuel, le rendant plus adaptable aux contextes locaux, tout en conservant sa pertinence au niveau mondial.

Il vise à constituer une ressource modulaire pour aider les institutions publiques, la société civile, les acteurs du secteur et d'autres parties prenantes à concevoir et à mettre en œuvre des systèmes de gouvernance des données à la fois adaptés à la finalité visée et adaptables aux réalités locales. Plutôt que de prescrire un modèle unique, il propose des questions directrices, des cadres flexibles et des outils pratiques pour aider les utilisateurs à parcourir le cycle de vie complet des données, de la collecte et du stockage au partage, en passant par l'analyse et l'utilisation.

Dans ce contexte, la gouvernance n'est pas seulement une question de conformité ou de contrôle. Il s'agit de définir la finalité, les principes, les processus, les personnes et les pratiques nécessaires pour garantir une utilisation des données fiable, utile, équitable et conforme aux droits de l'homme. Le présent Manuel met l'accent sur l'intendance plutôt que sur la propriété, reconnaissant que les données sont souvent coproduites, à usages multiples et façonnées par des dynamiques sociales, culturelles et politiques plus vastes.

Contenu du Manuel

Après l'« Introduction à la gouvernance des données », le présent Manuel s'articule autour de quatre composantes fondamentales de la gouvernance des données, appelées « FPPP » :

- **POURQUOI** (Finalité) : cherche à aider le lecteur à identifier et à préciser les objectifs visés par les données et leur gouvernance, en plus de définir la valeur publique qu'elles cherchent à créer.
- **COMMENT** (Principes) : vise à aider le lecteur à dresser la liste des normes qui guideraient les processus et les pratiques tout au long du cycle de vie des données.
- **QUI** (Personnes et Processus) : vise à aider le lecteur à structurer et à décrire les rôles, les responsabilités et les processus institutionnels nécessaires pour concevoir, mettre en œuvre et superviser efficacement la gouvernance (notamment la fonction des Intendant(e)s des données).
- **QUOI** (Pratiques et Mécanismes) : vise à aider le lecteur à définir des pratiques et des mécanismes exploitables pour mettre en œuvre des décisions tout au long du cycle de vie des données, en accord avec les principes, afin de répondre à la finalité.

Le Manuel comprend également :

- Un cadre d'auto-évaluation pour aider les organisations à évaluer leurs capacités actuelles ;
- Un glossaire des termes clés pour favoriser une compréhension commune ; et
- Une liste précise d'autres manuels et cadres pour un approfondissement éventuel.

À qui s'adresse ce Manuel ?

Le présent Manuel s'adresse à plusieurs types d'utilisateurs, chacun ayant des priorités, des responsabilités et des capacités qui lui sont propres.

Le principal groupe cible de ce Manuel se compose des personnes suivantes :

- **Responsables et Stratèges politiques.** *Si vous élaborez des stratégies nationales ou sectorielles en matière de données.* Utilisez ce Manuel pour définir une vision et faire correspondre les données de votre pays ou de votre organisation aux principes mondiaux, aux objectifs de développement durable et aux besoins de l'infrastructure publique numérique.
- **Responsables des données, Intendant(e)s des données, Autorités chargées de la protection des données et Responsables de la gouvernance.** *Si vous supervisez la manière dont les données sont collectées, partagées et utilisées au sein de votre organisation.* Utilisez ce Manuel pour évaluer les degrés de maturité, définir les principes, normaliser les pratiques du cycle de vie et lancer une collaboration intersectorielle.

En outre, d'autres experts peuvent trouver de la valeur à ce Manuel :

- **Professionnels du droit et de la conformité.** *Si vous veillez à ce que les pratiques en matière de données soient conformes aux lois et aux normes nationales et internationales.* Utilisez ce Manuel pour évaluer les garanties actuelles, intégrer les protections des droits de l'homme et faire correspondre les politiques aux instruments et aux cadres de gouvernance des données.
- **Architectes technologiques et Concepteurs de systèmes.** *Si vous créez des plateformes, des IPA ou des infrastructures qui gèrent ou déplacent des données.* Utilisez ce Manuel pour intégrer la gouvernance par la conception, comme les contrôles d'accès, les architectures fédérées et les normes de métadonnées.
- **Agences multilatérales et de donateurs.** *Si vous conseillez ou financez des efforts de transformation des données au niveau national.* Utilisez ce Manuel pour soutenir les initiatives de renforcement des capacités et encourager les modèles de gouvernance équitables et fondés sur les droits.
- **Société civile et Défenseurs des communautés.** *Si vous souhaitez façonner une gouvernance des données plus équitable.* Utilisez ce Manuel pour identifier les points d'intervention et vous familiariser avec les principes et les pratiques tels que l'autodétermination numérique et l'intendance des données.

Comment utiliser ce Manuel ?

Le présent Manuel est modulaire et peut être utilisé en totalité ou en partie :

- Pour élaborer ou réviser un cadre de gouvernance des données, commencez par les quatre questions directrices : POURQUOI (Finalité), COMMENT (Principes), QUI (Personnes) et QUOI (Pratiques).
- Pour évaluer l'état actuel de votre organisation, utilisez l'outil d'auto-évaluation pour comprendre votre état de préparation et identifier les domaines à améliorer.
- Pour impliquer les parties prenantes, passez en revue les outils proposés pour co-concevoir les politiques et renforcer la légitimité.

Limites du Manuel

Bien que le présent Manuel de gouvernance des données offre un cadre complet et des outils pratiques pour soutenir une gouvernance des données responsable et efficace, il est important de reconnaître ses limites :

1. Besoin d'adaptation contextuelle

Le présent Manuel est conçu pour s'adapter à divers secteurs, institutions et zones géographiques. Toutefois, ses orientations générales ne peuvent manifestement pas tenir compte de toutes les nuances juridiques, politiques, culturelles ou technologiques de chaque contexte. Les utilisateurs sont encouragés à adapter les recommandations à leur environnement spécifique, tout en respectant la législation internationale en matière de droits de l'homme, et à impliquer les parties prenantes locales dans ce processus.

2. Ce n’est pas un instrument juridique

Le présent Manuel ne se substitue pas à un conseil juridique ou à une mise en conformité avec la réglementation. Bien qu’il fasse référence à des exemples de principes internationaux et de modèles de gouvernance, il ne fournit pas d’interprétation officielle de lois ou de traités contraignants. Les utilisateurs doivent consulter des professionnels du droit et des organismes de réglementation pour s’assurer de la conformité avec les cadres juridiques applicables (p. ex. les lois régionales et nationales sur la protection des données).

3. Un paysage en évolution

Le domaine de la gouvernance des données évolue rapidement, façonné par les technologies émergentes (telles que l’IA et les données synthétiques), les nouveaux instruments réglementaires et l’évolution des attentes sociétales. Bien que ce Manuel intègre les pratiques les plus récentes au moment de sa publication, une révision régulière sera nécessaire pour rester à jour avec les avancées à venir.

Perspectives d’avenir

En résumé, le présent Manuel n’a pas vocation à être prescriptif. Il s’agit plutôt d’une ressource flexible favorisant l’adaptation au contexte, permettant à chaque utilisateur d’adapter les structures de gouvernance des données à ses réalités institutionnelles et culturelles, en tenant compte des normes internationales émergentes et des impératifs éthiques.

Pour remédier à ces limitations, les utilisateurs sont encouragés à :

- Considérer le présent Manuel comme un point de départ pour le dialogue et la cocréation ;
- Documenter et partager les enseignements tirés de la mise en œuvre afin de renforcer le savoir collectif ; et
- Participer aux futures mises à jour et contribuer à la communauté de pratique grandissante autour de la gouvernance des données.

Introduction à la gouvernance des données

Le XXIe siècle se caractérise par un processus accéléré de « datafication », c'est-à-dire la transformation systématique des pratiques quotidiennes, des opérations institutionnelles et des systèmes sociétaux en données quantifiables. Les flux de données ont connu une croissance exponentielle en raison de la prolifération des interactions numériques, de l'activité des réseaux sociaux et du déploiement à grande échelle de capteurs et d'appareils personnels. L'émergence et le déploiement de nouvelles technologies, comme l'intelligence artificielle (IA), amplifient cette tendance, à la fois en intensifiant la demande en données et en en produisant de nouvelles formes, notamment des données synthétiques. Le développement de l'infrastructure publique numérique (IPN) met également en évidence la nécessité de disposer de systèmes de données sécurisés et interopérables permettant de garantir un accès équitable aux services numériques et de soutenir l'innovation à grande échelle fondée sur les données.

L'enrichissement rapide de la société en données présente des défis et des opportunités colossaux pour les politiques publiques. Les gouvernements doivent adapter leurs cadres pour s'assurer que les données sont gérées de manière sûre et équitable, tout en encourageant l'innovation et en protégeant les droits des citoyens. Cela nécessite des politiques qui abordent des questions telles que la confidentialité des données, l'interopérabilité et l'utilisation éthique de l'IA. Les flux de données devenant de plus en plus essentiels au fonctionnement des économies et des sociétés numériques, des cadres politiques efficaces sont indispensables pour garantir que les données profitent à tous et respectent les droits de l'homme fondamentaux.

Les questions relatives à la gouvernance des données prennent donc de plus en plus d'importance. La gouvernance des données est d'autant plus importante que les données, comme de nombreuses technologies numériques, constituent une arme à double tranchant : elles peuvent apporter des avantages sociaux considérables (p. ex. en améliorant les interventions ciblées en matière de soins de santé et d'éducation), mais peuvent aussi causer des dommages importants (p. ex. en permettant la surveillance, en favorisant les biais et en entraînant des violations de la confidentialité).

L'impact de ces résultats dépend en partie de l'efficacité et de la réactivité des cadres de gouvernance des données, c'est-à-dire des principes, processus et pratiques qui entourent la manière dont les données sont collectées, stockées et déployées. Cette introduction donne un aperçu du concept et de la pratique de la gouvernance des données.

1. Pourquoi la gouvernance des données est-elle importante ?

L'élaboration d'un cadre de gouvernance des données est essentielle pour plusieurs raisons. En voici certaines (liste non exhaustive) :

- **Réalisation des objectifs de développement durable (ODD) :** Des données de bonne qualité sont essentielles pour mesurer les avancées, identifier les lacunes et éclairer les politiques qui soutiennent le Programme de développement durable de l'ONU à l'horizon 2030. Un grand nombre d'éléments confirment aujourd'hui le rôle que les données peuvent jouer dans la réalisation des ODD. Sans une solide gouvernance des données, ce potentiel est sérieusement entravé.
- **Prise de décisions plus éclairée :** Les gouvernements et les organisations s'appuient de plus en plus sur des données opportunes, exactes et bien gérées pour faire de l'information une priorité des décisions éclairées dans divers secteurs, notamment la santé, l'éducation, l'agriculture, le développement économique et l'action climatique. Une gouvernance efficace peut contribuer à l'intégrité, à l'accessibilité et à la facilité d'utilisation des données, ce qui en fait une base fiable pour l'élaboration des politiques.

- **Plus grande collaboration en matière de données :** En orientant la manière dont les données dispersées sont rassemblées, la gouvernance des données favorise une collaboration plus systématique, équitable et responsable, garantissant l'intérêt public tout en préservant les droits fondamentaux.

- **Atténuation des risques :** Si la *datafication* offre de nombreux avantages, elle présente également une série de risques sociétaux, allant de la violation de la confidentialité et des atteintes à la sécurité à l'exclusion économique et à la polarisation politique. De solides cadres de gouvernance des données peuvent aider à se prémunir contre ces risques et contribuer à maximiser le potentiel positif des données, tout en limitant leurs inconvénients.

- **Développement de l'IA éthique et responsable :** En l'absence de données de qualité et représentatives et d'une solide gouvernance, les applications d'IA risquent de perpétuer les biais, de violer la confidentialité et la propriété intellectuelle, en plus d'amplifier les inégalités. La gouvernance des données peut donc être considérée comme l'une des pierres angulaires d'une IA éthique et responsable.

- **Conception de l'infrastructure publique numérique (IPN) :** Les données jouent un double rôle dans l'infrastructure publique numérique (IPN) : elles en sont à la fois un outil essentiel et, de plus en plus, un élément central. Selon la Banque mondiale, l'IPN fait référence aux « systèmes fondamentaux permettant la fourniture de fonctions et de services essentiels à l'échelle de la société », englobant l'identité numérique, les paiements et les plateformes d'échange de données. Pour que l'IPN fonctionne efficacement et tienne sa promesse de services inclusifs, évolutifs et résilients, les données doivent être accessibles, fiables et bien gérées. Les pays ayant réussi à développer l'utilisation des données l'ont fait en réutilisant à plusieurs reprises les composants de l'IPN, tels que les systèmes d'identité et les registres, dans des secteurs tels que la santé, l'éducation, l'inclusion financière, l'agriculture et la traçabilité des terres et du carbone. Cette réutilisation améliore l'efficacité et favorise également le développement de nouveaux services basés sur les données. Lorsque l'IPN comprend des ensembles de données ou des systèmes d'échange de données, elle devient une ressource pour les prestataires de services publics, ainsi que pour les acteurs du secteur privé, la société civile et les innovateurs, ce qui multiplie sa valeur dans l'ensemble de l'écosystème. Pour libérer ce potentiel, une solide gouvernance des données est essentielle pour garantir que ces systèmes sont interopérables, sécurisés et conformes aux normes juridiques, éthiques et relatives aux droits de l'homme.

- **Amélioration des politiques de gouvernement ouvert :** Des cadres efficaces de gouvernance des données jouent également un rôle crucial dans les politiques de gouvernance ouverte. En établissant des règles concrètes pour la gestion des documents, en incluant des ensembles de données clés et en garantissant l'accès à l'information de manière active et passive, la gouvernance des données devient un élément essentiel de toute stratégie de gouvernance ouverte. Elle garantit la transparence, la responsabilisation et la distribution équitable de l'information publique, ce qui favorise la confiance du public et la participation aux processus démocratiques.

2. Qu'est-ce que la gouvernance des données ?

Il n'existe pas de définition universellement reconnue de la gouvernance des données. Cette notion constitue une vaste matrice de lois, de politiques, de normes, de technologies, d'individus, d'institutions et d'autres mécanismes et parties prenantes qui comprend au moins deux aspects essentiels :

- Comment les décisions relatives aux données sont prises ;

- Comment les données sont gouvernées tout au long de leur cycle de vie.

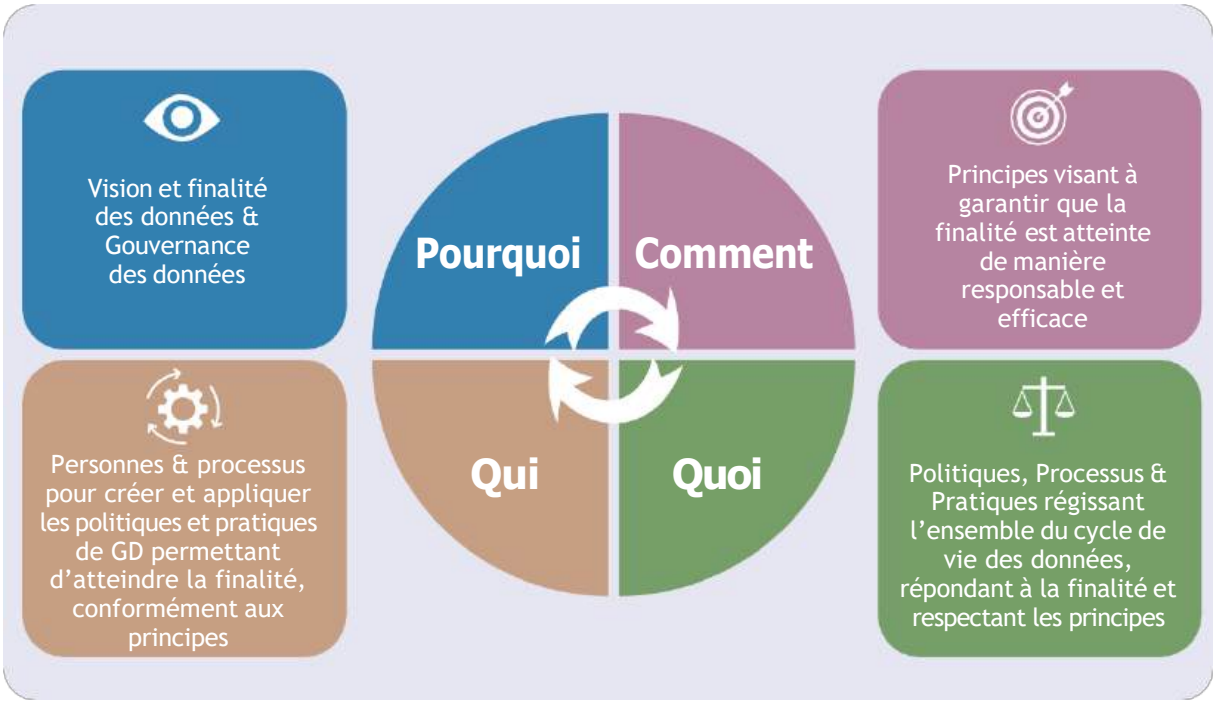
Sur la base de cette compréhension, le présent Manuel propose la définition de travail suivante pour la gouvernance des données :

***Les processus, les personnes, les politiques, les pratiques et la technologie** visant à régir le **cycle de vie des données afin** d'accroître la confiance, la valeur et l'équité, tout en minimisant les risques et les préjudices, conformément à un ensemble de **principes fondamentaux**.*

Comme indiqué précédemment, lors de la définition d'un cadre de gouvernance des données, quatre éléments et activités clés doivent être spécifiés :

- ***Pourquoi*** : Définir la vision et la(es) finalité(s) des données et de leur gouvernance.
- ***Comment :*** Préciser les **principes** qui guideront et détermineront la manière dont les décisions sont prises et les pratiques mises en œuvre pour atteindre la finalité de manière responsable et efficace.
- ***Qui*** : Établir des processus et impliquer les personnes nécessaires à la création et à l'application de politiques et de pratiques permettant d'atteindre la(es) finalité(s) dans le respect des principes.
- ***Quoi :*** Spécifier et mettre en œuvre les **politiques, les pratiques et les technologies** régissant les différentes étapes du cycle de vie des données, de manière à répondre à la finalité et à respecter les principes.

Schéma 1. Cadre de gouvernance des données



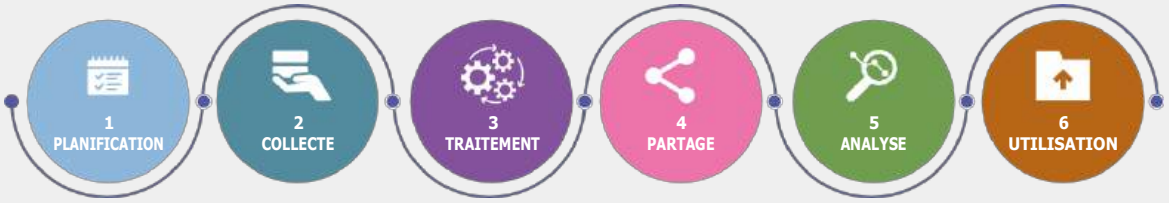
Source : Auteurs.

Encadré 1. Comprendre le cycle de vie des données

On entend par « cycle de vie des données » les différentes étapes par lesquelles passent les données, depuis leur planification initiale jusqu'à leur utilisation finale dans la prise de décisions.

Bien que les différents cadres puissent mettre l'accent sur différentes étapes ou utiliser une terminologie différente, les phases les plus communément reconnues sont les suivantes :

Schéma 2 : Cycle de vie des données



- **Planification** - Identifier les besoins en données, les utilisations prévues et les exigences en matière de gouvernance.
- **Collecte** - Collecte de données par le biais d'enquêtes, de capteurs, de transactions ou d'autres moyens.
- **Traitement** - Nettoyage, validation, organisation, stockage et conservation des données en vue de leur utilisation, y compris leur suppression si nécessaire, et garantie d'un traitement approprié tout au long de leur cycle de vie.
- **Partage** - Rendre les données accessibles à d'autres en vue de leur réutilisation, que ce soit par le biais de plateformes, d'IPA ou de collaborations de données.
- **Analyse** - Interpréter les données pour en tirer des enseignements.
- **Utilisation** - Appliquer ces connaissances pour éclairer les décisions, les politiques ou les services.

À chaque étape, des décisions sont prises en matière de gouvernance des données, notamment en ce qui concerne l'accès, le maintien de la qualité des données et la protection de la confidentialité. Ces décisions sont cumulatives et peuvent influencer de manière significative ce qui est possible aux stades ultérieurs. Une mauvaise gouvernance dès le départ (p. ex. une finalité peu claire ou une collecte non structurée) peut avoir des conséquences néfastes ou limiter la valeur ou l'utilité des données en aval.

3. Droits de l'homme et Gouvernance des données

Le présent Manuel préconise une approche de gouvernance des données fondée sur les droits de l'homme. Cela signifie que les pratiques en matière de données — tout au long de leur cycle de vie, de la collecte à la (ré)utilisation — doivent respecter, protéger et mettre en œuvre les droits et libertés des individus et des communautés. Cela implique également de traiter la gouvernance des données non pas comme une simple fonction technique ou de conformité, mais comme une pratique de gouvernance des données centrée sur les droits.

La section qui suit donne quelques exemples de la manière dont la gouvernance des données peut être alignée sur les cadres internationaux des droits de l'homme afin de prévenir les préjudices et de promouvoir la dignité, l'action et la justice.

Confidentialité & Protection des données :

Conflit : La collecte et le traitement des données peuvent contribuer à améliorer les résultats dans des domaines aussi divers que les performances scolaires ou les diagnostics et traitements médicaux. Toutefois, le traitement des données peut potentiellement porter atteinte au droit à la vie privée de la personne, en particulier lorsqu'il s'agit d'informations personnelles sensibles ou lorsque les garanties de sécurité ne sont pas régulièrement revues et mises à jour. *Réponse fondée sur les droits de l'homme :*

- Adopter des pratiques de minimisation des données, de protection dès la conception, de chiffrement et de stockage sécurisé.
- Garantir un consentement éclairé et significatif ainsi que la transparence dans la manière dont les données sont utilisées.
- Permettre aux individus d'accéder à leurs données, de les rectifier ou de les supprimer.
- Limiter la conservation des données au strict nécessaire, de manière proportionnée et justifiable au regard des normes relatives aux droits fondamentaux.

Discrimination & Biais :

Conflit : Les données biaisées et les systèmes algorithmiques peuvent renforcer la discrimination systémique dans l'analyse des données et la prise de décisions, entraînant des résultats discriminatoires (soit en favorisant une population, soit en produisant des résultats défavorables pour une autre, ou les deux à la fois) et refuser l'égalité de traitement à toutes les communautés. *Réponse fondée sur les droits de l'homme :*

- Vérifier les données et les modèles pour identifier, expliquer et atténuer les préjugés et les résultats discriminatoires.
- Engager diverses voix dans les processus de gouvernance des données.
- Veiller à ce que les ensembles de données reflètent et représentent diverses populations et soient adaptés au contexte.
- Contrôler en permanence les systèmes afin d'identifier et de corriger les impacts disparates.

Surveillance & Collecte massive de données :

Conflit : La surveillance de certains lieux (p. ex. distributeurs automatiques de billets, aéroports, établissements publics) ou d'événements (p. ex. matchs de football, élections ou manifestations) peut contribuer à prévenir la violence ou à améliorer la sécurité. Toutefois, une surveillance incontrôlée et sans limites appropriées peut porter atteinte à la vie privée, à la liberté d'expression, à l'autodétermination et au droit de réunion pacifique. *Réponse fondée sur les droits de l'homme :*

- Promulguer et appliquer des garanties juridiques solides, un contrôle et une procédure régulière.
- Limiter la surveillance et la collecte de données à des fins licites, spécifiques, nécessaires et proportionnées.
- Accroître la transparence autour des programmes et des technologies de surveillance, en les rendant accessibles et susceptibles d'être contestés.
- Mettre en place des mécanismes de responsabilisation indépendants pour examiner les abus.

Liberté d'expression & Accès à l'information :

Conflit : La gouvernance des données, notamment les contrôles d'accès aux données, peut conduire à la censure ou restreindre l'accès à des informations essentielles. *Réponse fondée sur les droits de l'homme :*

- Établir des politiques d'accès aux données équitables, transparentes et responsables.
- Promouvoir le libre accès aux informations et aux données d'intérêt public.
- Assurer un accès équitable aux données.
- Renforcer la culture numérique et l'engagement civique.
- Suivre les 9 principes d'un régime d'accès à l'information.

4. Responsabilisation & Recours

Conflit : Un cadre de gouvernance des données solide, équitable et efficace peut contribuer à protéger les individus et leurs données. Une gouvernance des données opaque ou non responsable peut laisser les individus sans recours en cas de violation. *Réponse fondée sur les droits de l'homme :*

- Créer des mécanismes de réclamation clairs, rapides et accessibles pour les violations des droits liés aux données.
- Garantir l'accès aux recours juridiques, notamment les voies administratives, judiciaires ou de médiation, avec des réparations exécutoires.
- Promouvoir la transparence dans les décisions de gouvernance des données et les études de risques.
- Inclure et garantir la participation du public dans les cadres de gouvernance des données.

Liste de contrôle et Auto-évaluation de la gouvernance des données

L'auto-évaluation de la gouvernance des données est conçue pour aider les gouvernements, les institutions publiques et les organisations à évaluer l'état actuel de leurs systèmes de gouvernance des données et à identifier les possibilités d'amélioration. Elle propose une méthode structurée pour passer de la réflexion à l'action, en s'appuyant sur le cadre « FPPP » : Finalités, Principes, Personnes et Pratiques.

ÉTAPES D'UTILISATION DE L'AUTO-ÉVALUATION

Vous trouverez ci-dessous des instructions étape par étape et des considérations clés pour utiliser cet outil de manière efficace.

Étape 1 : Constituer une équipe interfonctionnelle

Constituer un groupe interfonctionnel de parties prenantes internes et externes

Si possible, désigner un facilitateur ou un(e) Intendant(e) des données pour orienter le processus et consolider les contributions.

Étape 2 : Examiner chaque domaine et dimension

L'évaluation s'articule en six domaines principaux :

1. Vision & Finalité
2. Principes
3. Processus & Rôles
4. Pratiques & Politiques
5. Tendances émergentes & Gouvernance de l'IA
6. Capacité & Éducation

Pour chaque domaine, vous trouverez une série de questions Oui/Non destinées à évaluer si les éléments clés ont été instaurés.

Remarque : Il ne s'agit pas d'un exercice de conformité, mais d'un outil d'apprentissage et d'établissement de priorités.

Étape 3 : Déterminer le degré de priorité

En fonction de vos objectifs stratégiques et de vos contraintes, indiquez la priorité à accorder à l'amélioration de chaque dimension :

1. Élevée : Besoin d'action immédiate ; fondamentale pour d'autres domaines
2. Moyenne : Important, mais non urgent.
3. Faible : Pertinence moindre ou fonctionnement déjà satisfaisant.

Étape 4 : Analyser les résultats et planifier les prochaines étapes

Identifier les lacunes critiques.

Utiliser les résultats pour étayer votre stratégie de gouvernance des données, vos plans de renforcement des capacités ou vos révisions de politiques.

Partager les résultats en interne ou avec les partenaires pour favoriser l'alignement et coordonner les efforts.

À utiliser régulièrement (p. ex. une fois par an) pour suivre les avancées au fil du temps et assurer une amélioration continue.

Le Manuel de gouvernance des données

Les lecteurs peuvent se référer aux autres sections du Manuel de gouvernance des données pour obtenir un soutien ciblé :

- Utiliser la section POURQUOI (*Finalité*) pour affiner la stratégie ;
- Revoir la section COMMENT (*Principes*) pour l'aligner sur l'évolution des normes ;
- Consulter la section QUI (*Personnes et Processus*) pour la conception institutionnelle et les rôles d'intendance ;
- Appliquer la section QUOI (*Pratiques*) pour considérer les mécanismes à travers le cycle de vie des données.

Liste de contrôle et outil d'auto-évaluation

POURQUOI - Finalité :

Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Vision & Finalité	Existe-t-il une stratégie ou une vision nationale en matière de données qui lient explicitement la (ré)utilisation des données aux priorités (p. ex. le climat, la santé, la transformation numérique) ?	☐	☐		
	Les agences ou les secteurs disposent-ils de leurs propres stratégies ou programmes officiels en matière de données ?	☐	☐		
	Les objectifs ou les initiatives fondés sur les données sont-ils mentionnés dans les documents politiques officiels (p. ex. les stratégies des agences, les notes d'information, les plans de mise en œuvre) dans certains domaines au moins ?	☐	☐		
	Des cas d'usage à valeur publique ou des domaines prioritaires pour l'utilisation des données ont-ils été identifiés (p. ex. dans des secteurs tels que la santé, l'éducation ou la réponse aux crises) ?	☐	☐		
	Ces cas d'usage ou ces domaines prioritaires sont-ils élaborés en collaboration avec les principales parties prenantes, notamment les citoyens, la société civile ou les prestataires de services de première ligne ?	☐	☐		
	Existe-t-il un processus ou un mécanisme officiel permettant de revoir et d'actualiser régulièrement la stratégie nationale en matière de données en fonction de l'évolution des besoins ou des contributions des parties prenantes ?	☐	☐		

COMMENT - Principes :

Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Principes (Documen-tation	Les principes relatifs aux données (p. ex. la transparence, la responsabilisation, l'équité, la participation) sont-ils officiellement documentés ou approuvés publiquement au niveau politique ou organisationnel ?	☐	☐		
	Des principes techniques ou opérationnels distincts de gouvernance des données (p. ex. la qualité des données, l'intendance, le contrôle d'accès) sont-ils officiellement définis ou adoptés au niveau de l'agence/du service ?	☐	☐		
Principes (Mise en œuvre & Aligne-ment)	Ces principes sont-ils compris par le personnel concerné et intégrés dans les pratiques opérationnelles à tous les stades du cycle de vie des données (collecte, traitement, partage, réutilisation, stockage) ?	☐	☐		
	Les principes adoptés correspondent-ils aux principaux cadres internationaux tels que les principes **FAIR** (Facilement Trouvable, Accessible, Interopérable et Réutilisable), les principes **CARE** (Intérêt collectif, Droit de regard, Responsabilité et Éthique) et **FIPPS** (Norme fédérale de traitement des informations), garantissant la cohérence avec les bonnes pratiques mondiales en matière de gouvernance des données ?	☐	☐		
	Ces principes sont-ils applicables par le biais de contrats, de MA ou d'accords de partage de données avec des partenaires extérieurs ?	☐	☐		

QUI - Personnes & Rôles :

Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Personnes & Rôles	Existe-t-il un(e) Responsable des données (RD), un(e) Délégué(e) à la protection des données (DPD), un(e) Intendant(e) des données ou un(e) Responsable senior des données équivalent ayant une mission officielle et disposant des ressources adéquates ?	☐	☐		
	Existe-t-il une autorité de régulation ou une autorité indépendante de protection des données chargée de superviser les pratiques en matière de confidentialité et de protection des données et d'en assurer le respect par rapport aux législations nationales et internationales de protection des données ?	☐	☐		
	Les responsabilités des principaux acteurs de la gouvernance des données (p. ex. RD, DPD, comité d'éthique de l'IA, Intendant(e) des données) sont-elles clairement définies et documentées ?	☐	☐		
	Des rôles ou des structures officiels ou non ont-ils été instaurés pour l'intendance des données, la surveillance de la vie privée et l'utilisation éthique des données (p. ex. des Intendant(e)s des données, des Responsables de la protection de la vie privée, des comités d'éthique) ? Veuillez utiliser la section Commentaires pour expliquer.	☐	☐		
	Existe-t-il des mécanismes de coordination permettant d'aligner et de définir des politiques et des pratiques communes ?	☐	☐		
	Existe-t-il une coordination (officielle ou non) sur la gouvernance des données avec des acteurs extérieurs au gouvernement (p. ex. le monde universitaire, la société civile, le secteur privé) ? Veuillez utiliser la section Commentaires pour expliquer.	☐	☐		
	Existe-t-il une coordination (officielle ou non) sur la gouvernance des données avec des acteurs extérieurs au gouvernement (p. ex. le monde universitaire, la société civile, le secteur privé) ? Veuillez utiliser la section Commentaires pour expliquer.	☐	☐		

QUOI - Pratiques (Pratiques & Gouvernance)

Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Politiques & Gouver-nance	Des politiques officielles de gouvernance des données couvrant la collecte, le traitement, le partage et la réutilisation des données ont-elles été adoptées et sont-elles accessibles au public ?	☐	☐		
	Ces politiques ont-elles été élaborées selon une approche participative ?	☐	☐		
	Tous les services appliquent-ils ces politiques de manière cohérente ?	☐	☐		
	Ces politiques correspondent-elles aux normes éthiques et juridiques (p. ex aux lois nationales et régionales) ?	☐	☐		
	Les politiques sont-elles intégrées à la législation sur la liberté d'information, garantissant que l'accessibilité et la divulgation des données sont conformes aux exigences en matière d'accès du public à l'information ?	☐	☐		
	Les études de risques et d'impact sont-elles obligatoires avant l'utilisation ou le partage des données, en particulier pour les données sensibles ou les ensembles de données à haut risque (santé, finances, données démographiques) ?	☐	☐		
	Les besoins varient-ils entre les instituts de statistique, les unités informatiques ou les équipes d'intelligence artificielle ?	☐	☐		

QUOI - Pratiques (Cycle de vie des données)

Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Collecte & Stockage des données	Les processus de collecte des données correspondent-ils aux principes de limitation et de minimisation ?	☐	☐		
	L'expiration (fin de vie) ou la conservation des données est-elle spécifiée ?	☐	☐		
	Le consentement (notamment leconsentement dynamique) est-il obtenu si nécessaire ?	☐	☐		
	Les méthodes de collecte de données tiennent-elles compte de la représentation des groupes marginalisés ou mal desservis ?	☐	☐		
	Ces pratiques sont-elles communes à toutes les agences ou spécifiques à certains programmes ou unités ?	☐	☐		
Qualité des données, Gestion & Méta-données	Des contrôles de la qualité des données sont-ils mis en place (p. ex. exactitude, exhaustivité, actualité) ?	☐	☐		
	Des catalogues de données sont-ils utilisés et des normes de métadonnées (p. ex. DCAT, FAIR) sont-elles appliquées ?	☐	☐		
	Des pratiques de suivi des données ont-elles été mises en place pour garantir la traçabilité et accroître la responsabilisation ?	☐	☐		
	Les pratiques en matière de métadonnées et de qualité sont-elles harmonisées entre les départements ou dirigées par des unités spécifiques (p. ex. l'office national des statistiques) ?	☐	☐		
Partage des données	Existe-t-il des politiques et des outils techniques permettant un partage et une réutilisation sécurisés et éthiques des données (p. ex. des accords de partage des données, des IPA, des modèles d'accès fédérés) ?	☐	☐		
	Les données peuvent-elles être consultées par le biais de catalogues, de portails ou de registres ?	☐	☐		
	Les outils de partage de données et les portails sont-ils partagés entre les agences ou développés en silos ?	☐	☐		

Sécurité des données & Conformité	Existe-t-il des garanties juridiques et techniques pour la sécurité et la confidentialité des données ?	☐	☐		
	L'accès aux données et leur partage correspondent-ils aux cadres juridiques nationaux ou régionaux ?	☐	☐		
	Les éléments fondamentaux de l'infrastructure publique numérique (IPN), tels que les registres et les identifiants, sont-ils gérés de manière sécurisée ?	☐	☐		
Utilisation des données	L'analyse des données (notamment la visualisation) est-elle utilisée dans l'élaboration des politiques ?	☐	☐		
	Des mécanismes sont-ils mis en place pour une réutilisation responsable des données du secteur privé à des fins d'intérêt public ?	☐	☐		
	Les données sont-elles utilisées pour répondre aux crises (pandémie, catastrophes naturelles, etc.) ? Dans l'affirmative, quels sont les départements ou les domaines qui sont à la pointe des cas d'usage fondés sur les données ?	☐	☐		
Inclusivité, Équité, Autodétermination	Les groupes marginalisés sont-ils représentés dans les processus de gouvernance des données ?	☐	☐		
	Existe-t-il des cadres appropriés pour soutenir le contrôle et la gouvernance des données par les peuples autochtones ou les communautés, le cas échéant ?	☐	☐		
	L'accès équitable et l'utilisation des données font-ils l'objet d'une considération officielle dans la gouvernance ou la pratique ?	☐	☐		

QUOI - Pratiques (IA & Tendances émergentes) :

Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Tendances & Technologies émergentes	Existe-t-il des cadres de gouvernance de l'IA garantissant que les ensembles de données sont prêts pour l'IA (p. ex. impartiaux, représentatifs, de haute qualité) ?	☐	☐		
	Existe-t-il des lignes directrices éthiques pour le développement de l'IA à partir de données publiques ?	☐	☐		
	Les technologies de protection de la vie privée sont-elles utilisées pour minimiser les risques ?	☐	☐		

QUOI - Capacité & Éducation :

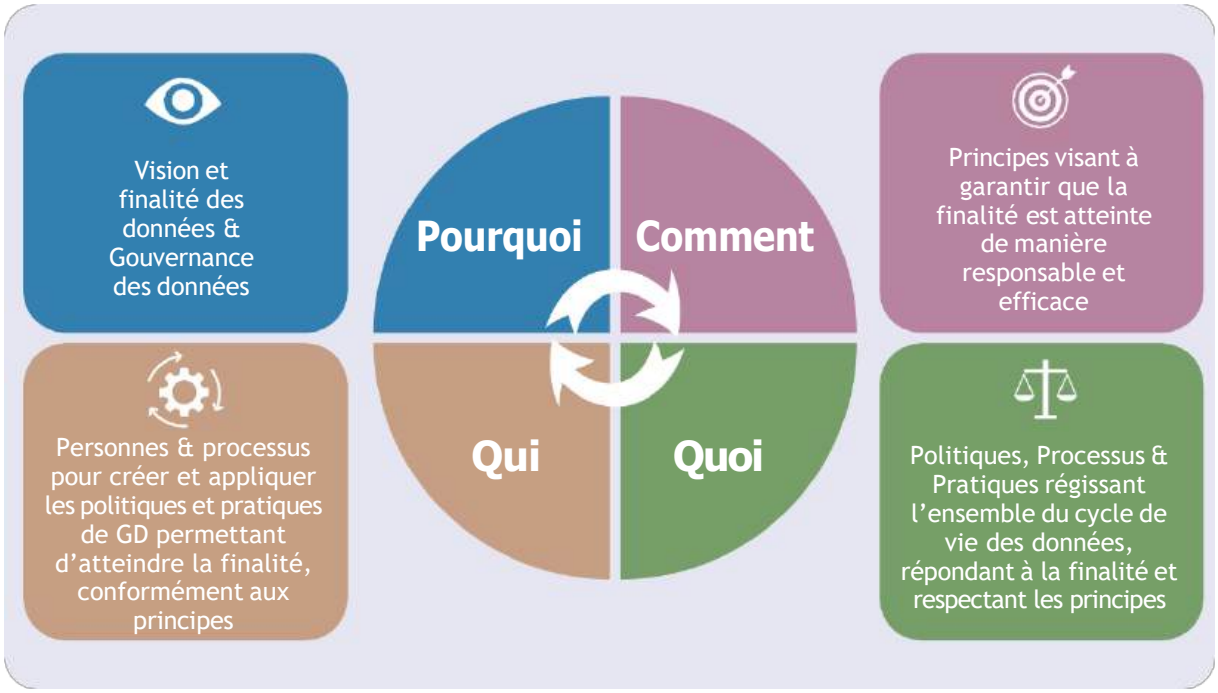
Dimension	Question	Oui	Non	Degré de priorité	Commentaires
Capacité & Éducation	Les acteurs à différents niveaux font-ils preuve d'une compréhension de base des concepts de gouvernance des données ?	☐	☐		
	Existe-t-il des programmes de formation réguliers sur la gouvernance des données, la protection de la vie privée, l'éthique de l'IA et la conformité ?	☐	☐		
	Existe-t-il des communautés de pratique officielles ou non travaillant sur la gouvernance des données (p. ex. des cercles d'apprentissage, des groupes de travail) ?	☐	☐		
	Existe-t-il des mécanismes de collaboration avec les universités, la société civile ou les organisations internationales pour renforcer le savoir institutionnel ?	☐	☐		
	Existe-t-il un budget dédié ou un plan d'investissement à long terme pour le renforcement des capacités en matière de gouvernance et d'intendance des données ?	☐	☐		
	Existe-t-il des communautés de pratique et des plateformes de partage des connaissances à l'intention des fonctionnaires ?	☐	☐		
	Les fonctionnaires sont-ils équipés pour s'engager dans la collaboration et les partenariats en matière de données (notamment intersectoriels) ?	☐	☐		

Éléments de base pour l'élaboration d'un cadre de gouvernance des données

Bien que les différents secteurs et juridictions puissent nécessiter des approches de gouvernance des données sur mesure, la section suivante vise à fournir des éléments de base pouvant être exploités et adaptés pour guider l'élaboration de nouvelles stratégies de gouvernance des données ou l'affinement de stratégies existantes. Comme expliqué précédemment, la gouvernance des données peut être décomposée selon les quatre questions suivantes (modèle « FPPP » de la gouvernance des données) et les actions qui doivent être spécifiées :

- ***Pourquoi :*** Définir la vision et la **finalité** des données et de leur gouvernance.
- ***Comment :*** Préciser les **principes** qui guideront et détermineront la manière dont les décisions sont prises et les pratiques mises en œuvre pour atteindre la finalité de manière responsable et efficace.
- ***Qui :*** Établir des **processus** et engager les **personnes** nécessaires à la création et à l'application de politiques et de pratiques permettant d'atteindre la finalité dans le respect des principes.
- ***Quoi :*** Spécifier et mettre en œuvre les **politiques, pratiques et technologies** régissant les différentes étapes du cycle de vie des données, de manière à répondre à la finalité et à respecter les principes.

Schéma 1.



Autres Manuels de gouvernance des données

Plusieurs autres Manuels de gouvernance des données existent déjà ; une sélection est présentée ci-dessous, avec une indication des publics auxquels ils s'adressent. Le présent Manuel n'a pas vocation à les reproduire, mais à les compléter. La valeur du présent Manuel réside dans son approche globale, en commençant par la finalité, les principes et les processus de gouvernance des données. Contrairement à de nombreux cadres existants qui se concentrent directement sur les pratiques, le présent Manuel met l'accent sur les éléments fondamentaux, en veillant à ce que la gouvernance des données soit fondée sur des considérations stratégiques et éthiques. Cette structure permet une gouvernance plus efficace, adaptable et pertinente dans des contextes divers.

Tableau 1. Cartographie des Manuels de gouvernance des données

Manuel	Public
Data Governance Toolkit (*Manuel de gouvernance des données*) (Agences du gouvernement de la Nouvelle-Galles du Sud)	Secteur public
Data Innovation Toolkit (*Manuel pour l'innovation en matière de données*) (Laboratoire d'innovation numérique de la Commission européenne)	Secteur public
Gouvernance des données de l'OCDE	Secteur public
Data to Policy Navigator (*Navigateur « Des données sur les politiques »*) (PNUD)	Secteur public
Cadre de politique des données (Union africaine)	Secteur public
Data Management Framework (*Cadre de gestion des données*) (ASEAN)	Secteur public
Navigating Data Governance (*Parcourir la gouvernance des données*) (UIT)	Autorités de régulation
Le PlayBook des données (*FICR et Solferimo Academy*)	Secteur humanitaire
Data Responsibility Journey (*Le parcours de la responsabilité des données*) (The GovLab)	Secteur public et privé
Data Governance and Management Toolkit (*Manuel de gouvernance et de gestion des données*) (Membres du SGIG DSC)	Gouvernements autochtones autonomes
Data Governance Workbook (*Manuel de gouvernance des données*) (Laboratoire de la société civile numérique)	Secteur à but non lucratif

I. POURQUOI

Définir la vision et la finalité des données et leur gouvernance

Yurchanka Siarhei/Shutterstock.com

Objectif principal

Établir et évaluer une vision et une finalité claires et réalisables pour l'utilisation et la gouvernance des données au sein d'une organisation ou d'un gouvernement.

L'importance de la finalité

Une vision et une/des finalité(s) claire(s) constituent le fondement d'une gouvernance des données efficace. Elles expliquent pourquoi une organisation ou un pays collecte, traite et utilise des données, reflétant ainsi ses valeurs fondamentales et ses priorités stratégiques. Les finalités ne sont pas monolithiques ; elles peuvent couvrir un large éventail d'objectifs, allant de l'optimisation de la valeur économique et sociale des données et de la promotion de l'innovation à l'établissement de l'équité, au soutien d'objectifs politiques spécifiques (tels que les résultats en matière de santé ou la protection de l'environnement), ou à la garantie d'un développement responsable de l'IA. Définir les finalités soulève des questions cruciales quant à savoir qui définit ces valeurs et veille à ce qu'elles soient conformes aux objectifs sociétaux plus larges et à la législation internationale en matière de droits de l'homme.

Évaluer la vision et la finalité

Pour déterminer si une vision et une finalité significatives pour les données et leur gouvernance sont instaurées et effectivement mises en œuvre, des éléments spécifiques doivent être évalués. Les questions suivantes portent sur l'existence, la nature et l'application de cette orientation stratégique.

1. Alignement et intention stratégiques : Relier la vision à la stratégie

- **Question : Existe-t-il une stratégie ou une vision nationale en matière de données qui associe explicitement la (ré)utilisation des données aux priorités politiques (p. ex. le climat, la santé, la transformation numérique) ?**
  - **Motif :** Une stratégie nationale est le principal instrument permettant de traduire une vision de haut niveau en un plan cohérent. Elle témoigne d'un engagement et doit préciser les finalités spécifiques des données qui sont prioritaires au sein de l'organisation au niveau national. La stratégie est-elle principalement axée sur l'optimisation de l'utilité et de la valeur des données pour la croissance économique ? L'accent est-il mis sur la promotion de l'innovation et du développement à long terme (p. ex. la promotion des ODD) ? Vise-t-elle explicitement à établir l'équité et l'autodétermination dans l'écosystème des données ? Ou bien se concentre-t-elle sur le soutien d'objectifs politiques spécifiques tels que l'ouverture des données pour la transparence, le partage transfrontalier des données ou la mobilisation des données pour la réponse aux crises ? Une stratégie claire permettrait de relier la vision globale à ces finalités sélectionnées et d'aligner la gouvernance des données sur les défis politiques urgents.
- **Question : Les agences ou secteurs individuels disposent-ils de leurs propres stratégies ou plans officiels en matière de données ?**
  - **Motif :** Il s'agit d'évaluer si la vision nationale et ses finalités prioritaires sont diffusées en cascade et traduites en plans réalisables dans des contextes opérationnels spécifiques. Les stratégies au niveau de l'agence rendent opérationnels les objectifs plus larges, en définissant la manière dont cette entité spécifique y contribuera. Par exemple, la stratégie d'une agence de santé peut détailler la manière dont elle utilisera les données pour faire progresser des ODD spécifiques liés à la santé (Soutenir des objectifs politiques spécifiques) tout en mettant en œuvre des cadres de gouvernance pour protéger les données des groupes vulnérables (Établir l'équité).

- **Question : Les objectifs ou les initiatives fondés sur les données sont-ils mentionnés dans les documents politiques officiels (p. ex. les stratégies des agences, les notes d'information, les plans de mise en œuvre) dans certains domaines au moins ?**
  - **Motif :** Cette question permet de vérifier si la stratégie (et les finalités qu'elle incarne) est intégrée dans le travail de base de l'organisation, ce qui garantit que les données sont traitées comme un atout essentiel pour atteindre les résultats, plutôt que de rester au stade de document isolé.

2. Définir la valeur et donner la priorité à l'utilisation : Traduire la finalité en action

- **Question : Des cas d'usage à valeur publique ou des domaines prioritaires pour l'utilisation des données ont-ils été identifiés (p. ex. dans des secteurs tels que la santé, l'éducation ou la réponse aux crises) ?**
  - **Motif :** Une vision et une stratégie claires doivent se traduire par des applications concrètes apportant une valeur concrète. Identifier les cas d'usage prioritaires démontre comment les finalités choisies sont poursuivies. Ces cas d'usage doivent refléter directement les priorités stratégiques. Par exemple :
    - Maximiser l'utilité/la valeur : Les cas d'usage se sont concentrés sur la connaissance de l'économie, l'optimisation des services ou la création de nouvelles sources de revenus.
    - Favoriser l'innovation/les ODD : Initiatives visant à promouvoir les écosystèmes de données ouvertes, la recherche fondée sur les données pour les objectifs de développement durable (santé, environnement, etc.) ou le développement des compétences.
    - Établir l'équité : Cas d'usage conçus pour remédier aux disparités, améliorer l'accès des groupes marginalisés ou protéger les droits des personnes concernées.
    - Soutenir les objectifs politiques spécifiques : Les cas d'usage ciblent l'amélioration de la transparence gouvernementale (données ouvertes), l'amélioration de la gestion des catastrophes (réponse aux crises), la promotion de la recherche internationale (partage transfrontalier), la garantie d'un déploiement équitable de l'IA (exploitation de l'IA) ou la mise en place de systèmes interopérables (alignement sur l'IPN). Cette évaluation permet de vérifier que l'intention stratégique conduit à une utilisation ciblée et efficace des données.

3. Inclusivité et engagement des parties prenantes :

- **Question : Ces cas d'usage ou ces domaines prioritaires sont-ils élaborés en collaboration avec les principales parties prenantes, notamment les citoyens, la société civile ou les prestataires de services de première ligne ?**
  - **Motif :** La gouvernance des données n'existe pas en vase clos. Le codéveloppement garantit que la finalité et les priorités définies reflètent les besoins et les valeurs réels des personnes concernées par l'utilisation des données. Elle favorise l'équité et l'autodétermination en incluant diverses perspectives, en particulier celles des groupes marginalisés ou vulnérables, ce qui permet d'établir la confiance et de garantir que les initiatives en matière de données sont réellement bénéfiques et minimisent les dommages. Cela va dans le sens d'une utilisation juste des données.

4. Adaptabilité et amélioration continue :

- **Question : Existe-t-il un processus ou un mécanisme officiel permettant de revoir et d'actualiser régulièrement la stratégie nationale en matière de données, avec la contribution d'un large éventail de parties prenantes, afin de s'assurer qu'elle correspond à l'évolution des besoins et des priorités du pays ?**
  - **Motif :** Le paysage des données est dynamique. Un processus de révision officiel permet de s'assurer que la vision et la finalité restent pertinentes et efficaces face à l'évolution des technologies, des attentes de la société et des objectifs politiques. Il permet de rectifier le tir et démontre un engagement en faveur de l'amélioration continue et de la responsabilisation.

Tableau 2. Liste de contrôle : Finalités potentielles des données et de leur gouvernance

Veuillez cocher toutes les finalités qui s'appliquent à votre initiative de gouvernance des données :

Maximiser l'utilité et la valeur des données / Minimiser les inconvénients et les coûts	
Exploiter les données pour obtenir de nouvelles informations et améliorer la prise de décisions.	☐
Créer une valeur concrète pour l'organisation et/ou la société grâce à l'utilisation des données.	☐
Réduire au minimum les inconvénients, les coûts et les conséquences négatives involontaires associés à la collecte et à l'utilisation des données.	☐
Favoriser l'innovation et le développement durable	
Stimuler l'innovation et l'esprit d'entreprise grâce à des approches fondées sur des données.	☐
Promouvoir et renforcer les opportunités économiques grâce aux données.	☐
Faire progresser les objectifs de développement durable (ODD) en utilisant des données pour le suivi et la mise en œuvre (p. ex. dans les domaines de la santé, de l'éducation et des infrastructures).	☐
Favoriser une culture axée sur les données et améliorer les compétences en matière de données au sein de l'organisation/société.	☐
Établir l'équité et l'autodétermination numérique	
Promouvoir un accès équitable aux données et veiller à ce que les bénéfices soient partagés équitablement.	☐
Veiller à ce que les groupes en situation de marginalisation et de vulnérabilité, notamment les enfants, soient activement inclus et protégés dans les écosystèmes de données.	☐
Respecter les principes d'autodétermination des données pour les individus et les communautés.	☐
Mettre en œuvre des cadres de gouvernance spécifiques pour protéger les données des populations vulnérables.	☐
Soutenir des objectifs politiques ou opérationnels spécifiques	
Améliorer la transparence, la responsabilisation et l'engagement des citoyens (p. ex. grâce à des initiatives de données ouvertes).	☐
Soutenir les priorités nationales en matière d'ODD, telles que la prise en compte de l'impact du changement climatique ou la promotion de l'éducation pour tous.	☐
Faciliter le partage transfrontalier de données et la collaboration internationale de manière sûre et efficace, tout en garantissant le respect de la législation.	☐
Mobiliser efficacement les données pour la préparation aux crises, la gestion des risques, la réponse et les efforts de rétablissement.	☐
Soutenir le développement et le déploiement responsables de l'intelligence artificielle (IA), en tenant compte des préjugés, de l'équité et des préoccupations éthiques.	☐
S'aligner sur les principes de l'infrastructure publique numérique (IPN) ou les mettre en œuvre afin de permettre un échange de données sécurisé, privé, interopérable et transparent.	☐

Méthodes

La gouvernance des données consiste à s'assurer que ces dernières sont gérées efficacement pour servir les objectifs et les valeurs d'un pays ou d'une organisation. La définition de la finalité de la gouvernance des données est la première étape cruciale, car elle dicte la portée, les priorités et le succès final de toute initiative de gouvernance. Les méthodes énumérées ci-dessous jouent chacune un rôle distinct, mais souvent interconnecté, dans ce processus de définition, et sont particulièrement renforcées lorsqu'elles sont envisagées sous l'angle des parties prenantes multiples.

Comprendre les besoins et les perspectives :

- Cartographie des parties prenantes : Il s'agit d'un élément fondamental. Avant de définir la finalité, il est nécessaire de savoir qui est concerné par la gouvernance des données et qui a un intérêt direct dans son résultat. La cartographie identifie les personnes clés, les équipes et les parties externes (comme les autorités de régulation ou les partenaires) ainsi que leurs liens actuels avec les données. Il est essentiel de comprendre leurs difficultés, leurs besoins et leurs attentes en matière de qualité des données, d'accès, de sécurité et de respect de la vie privée pour définir une finalité de gouvernance pertinente et utile. Exemple : Cartographie des parties prenantes et Engagement des parties prenantes

Définir une vision et des objectifs :

- Ateliers sur la vision et la mission : Ces sessions de collaboration rassemblent les parties prenantes pour réfléchir et articuler un état futur commun pour les données au sein de l'organisation. En définissant une vision convaincante (l'avenir idéal) et une mission claire (le rôle de l'organisation dans la réalisation de cet avenir grâce aux données) par le biais d'un processus multipartite, ces ateliers peuvent directement contribuer à définir la raison d'être de la gouvernance des données. La finalité doit s'aligner sur cette vision et cette mission plus larges en matière de données et en favoriser la réalisation. Exemple : Ateliers sur la vision et la mission et Ateliers et séances de consultation

- Cadre de la théorie du changement : Ce cadre permet de tracer logiquement les voies de causalité entre les activités de gouvernance des données et les résultats et l'impact souhaités à long terme. En formulant les hypothèses et les étapes nécessaires pour atteindre les objectifs de la gouvernance, il clarifie la manière dont la gouvernance des données est censée conduire le changement et atteindre sa finalité. Ce processus permet d'affiner la finalité en s'assurant qu'elle est réaliste et qu'elle a un impact clair. Exemple : Cadre de la théorie du changement

- Canevas de finalité (*Purpose Canvas*) : Il s'agit d'un outil spécialement conçu à des fins de clarification. Un canevas de finalité suscite généralement une réflexion sur des éléments tels que le problème central à résoudre, les bénéficiaires cibles, la proposition de valeur de l'initiative et les principales activités. L'application de ce principe à la gouvernance des données exige une articulation structurée des raisons pour lesquelles la gouvernance est nécessaire, des bénéficiaires et des fonctions clés qu'elle remplit. Exemple : Canevas de finalité

Assurer l'alignement et l'efficacité :

- Analyse comparative et examen des bonnes pratiques : L'examen de la manière dont d'autres organisations performantes ont abordé la gouvernance des données peut fournir des indications précieuses sur les défis communs, les stratégies efficaces et les déclarations de finalités potentielles. Bien qu'il ne s'agisse pas d'une méthode directe pour définir la finalité unique d'une organisation, il aide à comprendre ce qui est possible, en identifiant les domaines d'intérêt potentiels et en s'assurant que la finalité définie est complète et qu'elle tient compte des bonnes pratiques reconnues dans le domaine. Exemple : Analyse comparative et examen des bonnes pratiques

- Examen de la réglementation et de la conformité : La gouvernance des données est souvent dictée par des exigences réglementaires. Un examen approfondi des lois, notamment des obligations internationales et des règlements pertinents, est essentiel pour s'assurer que la finalité de la gouvernance des données inclut explicitement le respect de ces obligations. La conformité n'est pas seulement une tâche ; elle est souvent une raison fondamentale de la mise en œuvre de la gouvernance, et la finalité doit le refléter. Exemple : Examen réglementaire et de conformité et Examen des politiques et Analyse comparative

- Planification de scénarios : Il s'agit d'explorer différents états ou défis futurs potentiels liés aux données (p. ex. une violation de données, l'adoption d'une nouvelle technologie, une modification de la réglementation). En réfléchissant à la manière dont la gouvernance des données devrait fonctionner dans ces scénarios, les entreprises peuvent s'assurer que leur finalité définie est solide, adaptable et tournée vers l'avenir. Cela permet de définir une finalité qui n'est pas seulement pertinente aujourd'hui, mais qui le restera face à l'incertitude future. Exemple : Planification de scénarios

- Tableau de bord prospectif : Bien qu'il s'agisse avant tout d'un outil de gestion des performances utilisé après la définition de la finalité, le processus d'élaboration d'un tableau de bord prospectif pour la gouvernance des données peut rétrospectivement informer et affiner la compréhension de sa finalité. En identifiant les perspectives clés (p. ex. les finances, les clients, les processus internes, l'apprentissage et la croissance) et les objectifs auxquels la gouvernance des données contribue, il renforce la raison pour laquelle la gouvernance est importante et comment son succès sera mesuré par rapport aux objectifs de l'organisation. Cela renforce le lien entre la finalité de gouvernance et les objectifs stratégiques plus vastes. Exemple : Tableau de bord prospectif

Ressources pour une lecture plus approfondie :

- ASEAN Cadre de gestion des données
- GNUD Note d'orientation concernant les mégadonnées
- PNUD Data Futures Exchange (DFx)
- UNESCO Lignes directrices sur les données ouvertes
- Union africaine Résolution sur la promotion et l'exploitation de l'accès aux données en tant qu'outil de promotion des droits de l'homme et du développement durable à l'ère du numérique.

Études de cas nationales

- Évaluation des besoins en intelligence artificielle en Afrique
- Dotation Carnegie pour la paix internationale Peace's Data Governance : Asian Alternatives : How India and Korea Are Creating New Models and Policies (*Gouvernance des données, alternatives asiatiques : Comment l'Inde et la Corée créent de nouveaux modèles et de nouvelles politiques*).
- UNESCO Rapport d'évaluation de l'état de préparation à l'intelligence artificielle au Kenya
- UNESCO Chili : rapport d'évaluation de la préparation à l'intelligence artificielle
- UNESCO (2025). Afrique du Sud : Rapport d'évaluation de l'état de préparation à l'intelligence artificielle.
- UNESCO (2025). Mozambique : Rapport d'évaluation de l'état de préparation à l'intelligence artificielle.
- UNESCO (2025). Indonésie : Rapport d'évaluation de l'état de préparation à l'intelligence artificielle.

II. COMMENT

Définir les principes de gouvernance des données

TippaPatt/Shutterstock.com

Objectif principal

Définir et maintenir un ensemble clair de principes directeurs qui guident l'ensemble des processus, décisions et pratiques dans le cadre de la gouvernance des données, en garantissant la cohérence, la responsabilisation et l'alignement sur les valeurs éthiques, juridiques et sociétales.

L'importance des principes

Les principes constituent la base d'un cadre de gouvernance des données, façonnant l'ensemble des activités, des décisions et des processus afin de garantir la cohérence, la responsabilisation et la conformité aux normes convenues. Lorsqu'ils correspondent aux cadres reconnus au niveau mondial (tels que les droits de l'homme, la provenance des données et les normes d'interopérabilité ou les lignes directrices éthiques en matière d'IA), les principes peuvent contribuer à garantir que les pratiques en matière de données restent transparentes, équitables et adaptables. L'intégration de ces principes dans les structures de gouvernance de l'organisation facilite également la collaboration en matière de données, favorise la confiance entre les parties prenantes et peut encourager la collaboration transfrontalière en matière de données tout en préservant la vie privée et la sécurité.

Taxonomie des principes

Différentes catégories de principes guident les divers éléments constitutifs de la gouvernance des données, soulignant leur nature interconnectée et leur chevauchement fréquent. Les paragraphes suivants donnent un aperçu des principes les plus courants utilisés pour définir la manière dont les données sont régies dans les processus, la prise de décisions et le traitement des données.

1. Principes relatifs aux processus

Ces principes influencent la manière dont les processus de gouvernance doivent être menés, en cherchant à garantir que les décisions sont prises de manière juste, transparente et équitable. On peut notamment citer :

- **La transparence :** Veiller à ce que les processus de gouvernance, les méthodologies et les justifications soient ouverts, compréhensibles et accessibles aux parties prenantes concernées. Ce principe est essentiel pour faire respecter le droit de l'homme à rechercher, recevoir et communiquer des informations, pour permettre un examen public et encourager la responsabilisation.
- **La responsabilisation :** Établir une responsabilité claire pour les résultats de la gouvernance et veiller à ce qu'il existe des mécanismes permettant de tenir les individus et les organisations responsables de leurs actions liées aux données, en particulier lorsque ces actions ont un impact sur les droits de l'homme.
- **L'approche centrée sur les personnes :** Donner la priorité aux besoins, aux droits, au bien-être et aux intérêts des individus et des communautés tout au long du processus de gouvernance. Ce principe place explicitement la dignité humaine et les libertés fondamentales au cœur de la conception et de la mise en œuvre de la gouvernance des données.
- **L'équité :** Chercher à garantir l'égalité de traitement, à éviter les pratiques arbitraires et à atténuer de manière proactive les préjugés dans le cadre des procédures de gouvernance. Ce principe reflète directement le droit de l'homme à l'égalité et à la non-discrimination, en veillant à ce que la gouvernance elle-même ne perpétue pas ou ne crée pas de disparités injustes.
- **La participation :** Impliquer activement les parties prenantes concernées, notamment les individus et les groupes affectés en situation de marginalisation, dans la conception, la mise en œuvre et l'évolution des processus de gouvernance, dans la mesure du possible et de manière appropriée. Ceci est conforme aux droits liés à la liberté d'association, de réunion pacifique et à la participation aux affaires publiques, en reconnaissant que les personnes

concernées par la gouvernance des données doivent pouvoir s'exprimer.

- **La licéité :** Veiller à ce que tous les processus et activités de gouvernance soient pleinement conformes aux cadres légaux applicables, notamment aux lois et aux règlements internationaux, est un principe fondamental de la législation en matière de droits de l'homme. Ce principe exige que toute ingérence dans les droits soit encadrée par la loi (principe de légalité), que la loi elle-même soit compatible avec les normes en matière de droits de l'homme (principe de légitimité) et que les dispositions de la loi soient nécessaires et proportionnées afin d'atteindre l'objectif visé. Ces trois éléments, connus sous le nom de test en trois parties, garantissent que les actions juridiques respectent les droits fondamentaux des individus tout en maintenant une approche légale et équilibrée de la gouvernance.

- **L'inclusion :** Veiller à ce que les processus de gouvernance soient conçus pour représenter et inclure les perspectives et les besoins de tous les groupes concernés, en accordant une attention particulière aux communautés « invisibles », aux peuples autochtones (en lien avec les principes CARE, voir ci-dessous) et aux parties prenantes concernées par les impacts humains de la *datafication*. Ce principe est essentiel pour appliquer le droit de l'homme à la non-discrimination et garantir que la gouvernance des données sert et protège tout le monde de manière équitable.

2. Principes relatifs aux décisions

Ces principes guident les décisions prises dans le cadre de la gouvernance, en déterminant la manière dont les politiques sont définies et mises en œuvre. On peut notamment citer :

- **La transparence :** Communiquer clairement aux parties concernées la logique, les critères et l'impact des décisions de gouvernance. La transparence dans la prise de décisions est essentielle pour permettre aux individus de comprendre comment les décisions affectant leurs données (et potentiellement leurs droits) sont prises, pour soutenir le droit à l'information et faciliter les contestations le cas échéant.

- **La proportionnalité :** Veiller à ce que les décisions de gouvernance, les politiques et toutes les restrictions relatives à l'utilisation des données soient adaptées à l'ampleur de l'activité liée aux données, aux risques éventuels et aux avantages escomptés, tout en minimisant les effets négatifs potentiels sur les droits de l'homme. Ce principe est une pierre angulaire de la législation sur les droits de l'homme lorsqu'il s'agit de mettre en équilibre des droits concurrents ou des objectifs légitimes ; toute limitation d'un droit doit être nécessaire et proportionnée à l'objectif poursuivi.

- **La finalité définie :** Veiller à ce que toutes les décisions de gouvernance soient prises dans un but clair, légitime et spécifique, et que l'utilisation des données soit conforme à cette finalité. Ce principe, fondamental pour la protection des données et le droit à la vie privée, empêche les utilisations arbitraires ou abusives des données susceptibles de porter atteinte aux droits et aux libertés.

- **La responsabilisation :** Veiller à ce que les décideurs soient clairement identifiés et tenus responsables des résultats et des effets des politiques et des choix qu'ils font, en particulier en ce qui concerne les implications potentielles en matière de droits de l'homme. Cela renforce le droit à un recours effectif et encourage une prise de décisions responsable.

- **L'approche centrée sur les personnes :** Placer les besoins, les droits (notamment le respect de la vie privée, l'autonomie et la non-discrimination) et le bien-être des personnes au premier plan lors de la prise de décisions concernant la collecte, l'utilisation et le partage des données. Ce principe garantit que la prise de décisions privilégie l'impact sur l'homme par rapport à la commodité purement technique ou organisationnelle.

- **L'équité :** Veiller à ce que les décisions de gouvernance aboutissent à un traitement juste et équitable des individus et des groupes, en s'efforçant de manière proactive d'identifier et d'éliminer les sources potentielles de partialité dans les processus de prise de décisions ou les résultats. Cela soutient directement le droit de l'homme à l'égalité et à la non-discrimination.

- **La prévention des préjudices et la non-discrimination :** Concevoir explicitement des décisions et des politiques pour se prémunir des effets négatifs potentiels sur les individus ou les groupes et prévenir activement les utilisations discriminatoires des données, les résultats algorithmiques biaisés ou d'autres préjudices. Il s'agit d'un principe fondamental des droits de l'homme - l'obligation de ne pas causer de préjudice et de garantir l'égalité de traitement et de chances pour tous.
- **La participation :** Encourager et intégrer des contributions et des perspectives diverses dans les processus de prise de décisions, en particulier lorsque les décisions ont un impact sur différents groupes de parties prenantes, notamment les populations vulnérables ou marginalisées. Cela renforce la légitimité des décisions et contribue à garantir qu'elles respectent les droits et les réalités des personnes concernées.

3. Principes relatifs à la gestion des données

Ces principes guident la manière dont les données sont traitées, stockées et partagées, garantissant ainsi un traitement responsable et sûr. On peut notamment citer :

- **La confidentialité et la sécurité :** Mettre en œuvre des mesures techniques et organisationnelles solides pour protéger les données sensibles et confidentielles contre l'accès, la divulgation, l'altération ou la destruction non autorisés. Une sécurité adéquate n'est pas seulement une exigence technique ; elle est essentielle pour protéger les libertés fondamentales, en particulier le droit à la vie privée et potentiellement d'autres droits comme le droit à la sécurité de la personne, car les violations de données peuvent conduire à des violations importantes des droits de l'homme.
- **La proportionnalité :** Veiller à ce que les méthodes et l'étendue de la collecte, du traitement, du stockage et du partage des données soient appropriées et limitées au strict nécessaire et au moins intrusif pour la finalité définie et légitime. En ce qui concerne les données à caractère personnel, ce principe est essentiel pour le respect du droit à la vie privée, en veillant à ce que les activités de gestion des données n'empiètent pas indûment sur la sphère personnelle des individus au-delà de ce qui est justifiable et nécessaire.
- **L'accessibilité et la portabilité des données et de l'information :** S'efforcer de rendre les données accessibles et utilisables le cas échéant (p. ex. les principes des données ouvertes), tout en équilibrant les besoins de partage et de découverte avec les exigences en matière de protection de la vie privée, de sécurité, de propriété intellectuelle et de contrôle des utilisateurs (p. ex. le droit des individus d'accéder à leurs propres données et éventuellement, de les déplacer). Ce principe soutient différents droits, notamment celui de rechercher et de recevoir des informations (accès à l'information), celui au respect de la vie privée (liberté de limiter toute intrusion et contrôle de ses données) et potentiellement, dans certains pays, la liberté d'expression et d'association dans l'environnement numérique.
- **La protection de la vie privée :** Veiller à ce que les données à caractère personnel soient collectées, utilisées et gérées d'une manière qui respecte les droits de la personne en matière de protection de la vie privée et qui soit conforme aux lois et réglementations applicables en la matière. Le droit à la vie privée est un droit de l'homme fondamental dans de nombreux pays, consacré par divers instruments nationaux et internationaux, et ce principe fait de sa protection un élément central de toutes les activités de traitement des données.
- **La licéité :** Veiller à ce que toutes les activités de gestion des données soient conformes aux cadres juridiques applicables, notamment les réglementations en matière de protection des données, les lois sectorielles et les obligations contractuelles. La licéité du traitement des données signifie que les activités ont une base légale et surtout, que ces lois sont elles-mêmes conformes aux normes internationales en matière de droits de l'homme.
- **Consentement éclairé :** Lorsque le consentement est la base juridique du traitement, s'assurer que les personnes reçoivent des informations claires et compréhensibles sur la manière dont leurs

données seront utilisées et qu'elles donnent ou retirent leur accord librement et sans contrainte. Ce principe est profondément ancré dans les concepts des droits de l'homme que sont l'autonomie, la dignité et le droit de contrôler ses informations personnelles.

- **La qualité des données et des métadonnées :** S'engager à maintenir des normes élevées en matière d'exactitude, d'exhaustivité, de cohérence, d'actualité et de pertinence des données, notamment des données déduites. Veiller à ce que les métadonnées soient structurées, complètes et lisibles par machine facilite la découverte, la compréhension et la réutilisation responsable des données. La mauvaise qualité des données peut conduire à des résultats injustes ou discriminatoires et potentiellement violer les droits de l'homme (p. ex. des données incorrectes ayant un impact sur l'accès aux services sociaux, ou entraînant des décisions biaisées par des systèmes automatisés).
- **L'interopérabilité et la normalisation :** Respecter les formats, structures, protocoles et normes de données reconnus afin de promouvoir l'échange et l'intégration des données au sein et entre les systèmes, les organisations et les juridictions. Bien qu'essentiellement technique, ce principe peut soutenir les droits de l'homme en facilitant l'accès à l'information publique, en permettant la portabilité des données pour soutenir les droits individuels et en promouvant la transparence grâce à des flux de données normalisés, à condition qu'il soit mis en œuvre avec des garanties de protection de la vie privée et de sécurité.

Encadré 2. Principe émergent : Qu'est-ce que l'autodétermination numérique (ADN) ?

L'autodétermination numérique (ADN) est un nouveau principe de gouvernance des données qui transfère le contrôle des institutions centralisées vers les individus et les communautés, leur permettant de participer activement à l'utilisation de leurs données tout au long de leur cycle de vie. Contrairement aux modèles traditionnels, qui reposent sur un consentement unique ou une gouvernance centrée sur l'État, l'ADN introduit un contrôle continu et adaptatif et se concentre sur la prise de décisions participative basée sur une licence sociale. Les principes fondamentaux sont les suivants :

- **Contrôle dynamique et permanent :** Les individus et les communautés peuvent adapter l'utilisation de leurs données en fonction de l'évolution de la situation, en s'appuyant sur une licence sociale.
- **Pouvoir d'action effectif tout au long du cycle de vie des données :** Les individus participent non seulement à la collecte des données, mais aussi à leur traitement, à leur partage et à leur suppression, ce qui favorise une structure de gouvernance plus participative.
- **Transparence et confiance :** L'ADN exige une transparence permanente, garantissant que les personnes concernées sont continuellement informées de la manière dont leurs données sont traitées, stockées et partagées.
- **Aspect éthique et collectif :** L'ADN transcende les modèles de propriété et de marchandisation en intégrant des dimensions éthiques, sociales et communautaires dans la gouvernance des données. Elle met l'accent sur la gouvernance collective, permettant aux communautés de cogérer les données dans le respect des droits individuels et collectifs.
- **Gouvernance décentralisée et démocratique :** L'ADN décentralise l'autorité, encourageant la gouvernance partagée entre les individus, les communautés et les organisations plutôt que la concentration du pouvoir dans les institutions.

Source : International Network on Digital Self-Determination (Réseau international pour l'autodétermination numérique).

Tableau 3. Liste de contrôle : Principes relatifs aux données et à leur gouvernance

Cette liste de contrôle permet de déterminer et d'évaluer la présence de principes dans trois catégories :

1. **Principes relatifs aux processus** – comment les décisions de gouvernance sont prises
2. **Principes relatifs aux décisions** – ce qui éclaire ces décisions
3. **Principes relatifs à la gestion des données** – comment les données sont gérées dans la pratique

Principes relatifs aux processus	
Ces principes peuvent contribuer à garantir que les activités de gouvernance sont menées de manière juste, équitable et transparente.	
Transparence - Les processus de gouvernance sont ouverts et compréhensibles.	☐
Responsabilisation - Les acteurs sont responsables des résultats des processus de gouvernance.	☐
Approche centrée sur les personnes - Les besoins et les droits des personnes sont prioritaires.	☐
Équité - Traitement égal et impartial dans les procédures.	☐
Participation - Les parties prenantes sont impliquées de manière significative dans la gouvernance.	☐
Licéité - Toutes les actions sont conformes aux lois et aux règlements en vigueur.	☐
Inclusion - Les groupes marginalisés et les personnes affectées par la *datafication* et la numérisation croissante de la société sont représentés.	☐
Principes relatifs aux décisions	
Ces principes peuvent aider à guider la définition et la mise en œuvre des décisions en matière de gouvernance des données.	
Transparence - La logique des décisions est clairement communiquée.	☐
Proportionnalité - Les décisions sont adaptées à leur contexte et à leur impact.	☐
Finalité définie - Les décisions sont guidées par des objectifs clairs et spécifiques.	☐
Responsabilisation - Les décideurs doivent répondre de leurs choix.	☐
Approche centrée sur les personnes - Les besoins et les droits des personnes sont au premier plan.	☐
Équité - Les décisions sont justes et équitables. Protection contre les préjudices et non-discrimination - Les risques sont atténués et les préjugés sont évités.	☐
Participation - Des perspectives diverses sont intégrées dans la prise de décisions.	☐

Principes relatifs à la gestion des données	
Ces principes peuvent contribuer à garantir que les données sont gérées de manière responsable, sûre et conforme aux droits et aux attentes législatives.	☐
Confidentialité et Sécurité - Les données sensibles sont protégées contre tout accès non autorisé.	☐
Proportionnalité - Les pratiques en matière de données correspondent à la finalité et au besoin.	☐
Accessibilité et portabilité - Les données sont disponibles et portatives dans des conditions appropriées.	☐
Protection de la vie privée - Les données à caractère personnel sont protégées conformément aux lois sur la protection de la vie privée.	☐
Licéité - La gestion des données est conforme à toutes les normes juridiques applicables.	☐
Consentement éclairé - Les personnes concernées sont conscientes de l'utilisation des données et y consentent.	☐
Qualité des données et des métadonnées - Les données sont précises, fiables et bien documentées en vue de leur réutilisation.	☐
Interopérabilité et Normalisation - Les données adhèrent à des normes communes pour faciliter l'échange et l'intégration.	☐

Principes existants et spécifiques

Voici quelques exemples de cadres existants pouvant contribuer à l'élaboration de cadres de gouvernance des données :

- Principes CARE : Centrée sur les droits des peuples autochtones, la gouvernance des données favorise le bénéfice collectif, l'autorité de contrôle, la responsabilité et l'éthique.
- Principes de protection de la vie privée (Organisation de coopération numérique) : L'objectif est de fournir aux États membres de l'OCN une base commune de confidentialité des données et de constituer la base du mécanisme d'interopérabilité de l'OCN pour les flux transfrontaliers de données.
- Principes FIPPs (relatifs aux pratiques équitables en matière d'information) : Des principes fondamentaux tels que la transparence, la spécification de la finalité, la responsabilisation, la qualité des données et l'accessibilité contribuent à garantir un traitement équitable et responsable des données.
- Principes FAIR pour les données scientifiques : Il s'agit de rendre les données Facilement trouvables, Accessibles, Interopérables et Réutilisables, principalement pour les données de recherche.
- Principes FARR : Principes dans le domaine de l'apprentissage automatique, de la préparation à l'IA & de la Reproductibilité.
- CPI, Directives opérationnelles sur la responsabilité en matière de données dans l'action humanitaire.
- Principes de bonnes pratiques de l'OCDE pour l'éthique des données : Ils soutiennent la mise en œuvre de l'éthique des données dans les projets de gouvernance numérique, en mettant l'accent sur la confiance au cœur de la conception et de la mise en œuvre, mais aussi le maintien de l'intégrité publique par le biais d'actions spécifiques menées par les gouvernements et les organismes publics.
- Recommandation de l'OCDE sur l'amélioration de l'accès aux données et de leur partage : Ensemble de principes et d'orientations politiques convenus au niveau international sur la manière dont les gouvernements peuvent maximiser les avantages intersectoriels de nombreux types de données, tout en protégeant les droits des parties prenantes.
- Approche des données du HCR fondée sur les droits de l'homme : Ensemble de principes, de recommandations et de bonnes pratiques pour guider la collecte et la ventilation des données sous l'angle des droits de l'homme.
- Conseil des chefs de secrétariat (CCS) de l'ONU pour la coordination : Principes relatifs à l'utilisation éthique de l'IA par le système des Nations Unies.

Encadré 3. FOCUS : Principes de gouvernance des données du Conseil des chefs de secrétariat (CCS) de l'ONU pour la coordination

1. **Valeur**

- Maximiser la valeur des données : Mettre l'accent sur la promotion d'une culture valorisant les données en tant qu'atout essentiel pour favoriser le développement de tous. Il s'agit notamment de promouvoir la qualité des données, leur utilisation responsable et d'améliorer l'interopérabilité grâce à des définitions et à des classifications normalisées.
- Permettre l'utilisation et la réutilisation des données : Encourager l'accès, le partage et la réutilisation appropriés des données au-delà des frontières et des secteurs, en veillant à ce qu'elles contribuent au bien public. Intégrer les principes de mutualité et de solidarité dans la gouvernance des données afin de garantir que la valeur des données profite à la fois aux individus et à la société dans son ensemble, au lieu d'être exploitée pour le seul profit privé.
- Interopérabilité : Adopter des formats, des métadonnées et des définitions de données normalisés afin de garantir un échange de données et une collaboration transparente entre les systèmes et au-delà des frontières.
- Bâtir une société éduquée aux données : Promouvoir l'éducation aux données et l'accès à la technologie pour permettre aux individus et aux organisations d'utiliser, d'analyser et de comprendre efficacement les données. Mettre en place des initiatives et des infrastructures éducatives permettant aux personnes de travailler avec des données et de les comprendre de manière responsable.

2. **Confiance**

- Approche des données fondée sur les droits de l'homme : Ancrer toutes les pratiques de gouvernance des données dans les cadres internationaux des droits de l'homme, en garantissant le respect de la vie privée, la protection et la sécurité des données à caractère personnel, en particulier pour les groupes vulnérables. Prioriser la protection de l'accès aux données à caractère personnel avant d'en permettre l'utilisation et la réutilisation, en veillant à ce que la gouvernance des données respecte les droits fondamentaux avant de faciliter l'innovation.
- Responsabilisation et Transparence : Établir des rôles, des responsabilités et des mécanismes de contrôle clairs afin de garantir la responsabilisation en matière de gouvernance des données. Cela inclut des processus décisionnels transparents et la mise en place de mécanismes de réparation pour les personnes et les communautés affectées par l'utilisation abusive des données. De même, l'application des limites de la procédure régulière pour réglementer l'accès, le traitement et l'utilisation légaux des données.
- Qualité et Sécurité des données : Mettre en œuvre des mesures solides pour garantir la qualité des données et protéger la sécurité des données et de l'infrastructure qui les soutient tout au long de leur cycle de vie. Cela inclut l'utilisation d'études de la qualité des données en fonction du contexte et la protection contre la corruption ou la violation des données. Cela implique également la prévention de la réidentification non autorisée et la protection des individus et des groupes sociaux contre la discrimination découlant du traitement et de l'analyse des données.

3. **Équité**

- Promouvoir l'équité dans la gouvernance des données : Veiller à ce que les avantages des données soient équitablement répartis, en se concentrant sur la réduction de la pauvreté en matière de données et la prévention de la discrimination. Ce principe met l'accent sur les processus décisionnels participatifs, en veillant à ce que les individus et les communautés marginalisés aient le contrôle de leurs données et soient activement impliqués dans l'élaboration des politiques en matière de données et des structures de gouvernance. Il encourage également la représentation dans les organes de gouvernance et les initiatives communautaires en matière de données afin de garantir que la gouvernance des données reflète les besoins des diverses populations.
- Autodétermination numérique : Défendre les droits des individus à contrôler leurs données à caractère personnel, en leur permettant de prendre des décisions éclairées quant à leur utilisation et en garantissant leur représentation dans l'écosystème des données. Elle promeut des mécanismes de consentement clairs et accessibles et des cadres d'acceptation qui permettent aux utilisateurs de s'engager de manière significative sur la manière dont leurs données sont utilisées.
- Équité et Non-discrimination : Promouvoir un traitement équitable dans la collecte, l'analyse et l'utilisation des données, en s'efforçant activement d'atténuer les préjugés et de prévenir la discrimination dans les pratiques relatives aux données. Garantir la transparence et l'explicabilité des algorithmes, en exigeant que les modèles de prise de décisions soient interprétables, responsables et vérifiables, afin d'éviter les systèmes opaques ayant un impact disproportionné sur certaines communautés.
- Intendance des données et Réutilisation éthique pour le bien public : Mettre en place de solides pratiques d'intendance des données qui concilient leur protection avec une réutilisation responsable et éthique afin de maximiser les avantages pour la société. Veiller à ce que les données soient gérées de manière responsable, éthique et sécurisée afin de soutenir les intérêts publics tout en respectant les droits des personnes. Encourager les initiatives de données ouvertes et les modèles de gouvernance collaborative qui offrent un accès équitable aux données pour la recherche, l'élaboration des politiques et l'innovation sociale, tout en respectant la vie privée et les préoccupations en matière de sécurité.

Questions d'évaluation & leurs motifs :

Principes (Documentation)

1. **Question : Les principes relatifs aux données (p. ex. la transparence, la responsabilisation, l'équité, la participation) sont-ils officiellement documentés ou approuvés publiquement au niveau politique ou organisationnel ?**
   - **Motif :** Cette question vise à déterminer si l'organisation a établi des principes de haut niveau, souvent importants d'un point de vue éthique ou politique, pour orienter son approche globale des données. La documentation officielle et l'approbation des dirigeants sont des signes d'engagement et fournissent une base pour la confiance et la responsabilisation publique. Une réponse positive constitue la première étape de la définition de valeurs fondamentales pour l'utilisation des données.
2. **Question : Des principes techniques ou opérationnels distincts de gouvernance des données (p. ex. la qualité des données, l'intendance, le contrôle d'accès) sont-ils officiellement définis ou adoptés au niveau de l'agence/du service ?**
   - **Motif :** Si les principes de haut niveau définissent la vision, les principes opérationnels traduisent cette vision en orientations pratiques pour la gestion quotidienne des données. Cette question vérifie l'existence de règles ou de normes spécifiques relatives à la manière dont les données sont traitées par les praticiens. Une définition officielle garantit la cohérence et la clarté pour les créateurs, les utilisateurs et les Intendant(e)s de données. Une réponse positive indique que ces lignes directrices pratiques nécessaires ont été établies.

Principes (Mise en œuvre & Alignement)

3. **Question : Ces principes sont-ils intégrés dans les pratiques opérationnelles à tous les stades du cycle de vie des données (p. ex. la collecte, le traitement, le partage, la réutilisation) ?**

   - **Motif :** Les principes ne sont efficaces que s'ils sont mis en pratique. Cette question évalue le lien important entre les principes documentés et la mise en œuvre effective. L'intégration signifie que les principes influencent les flux de travail, les outils, la formation et la prise de décisions tout au long du cycle de vie et du traitement des données. Une réponse positive indique que les principes sont intégrés aux pratiques réelles de traitement des données.

4. **Question : Les principes adoptés correspondent-ils aux cadres internationaux ou nationaux et aux bonnes pratiques (p. ex. les principes FAIR, CARE, ceux de l'Union africaine et de l'OCDE) ?**

   - **Motif :** L'alignement sur des normes et des réglementations externes reconnues est essentiel pour l'interopérabilité, la conformité, la responsabilité éthique (en particulier pour des types de données spécifiques comme les données autochtones) et la démonstration de bonnes pratiques au niveau international ou national. Cette question permet d'évaluer si les principes internes s'accordent avec des écosystèmes de données juridiques, éthiques et techniques plus vastes. Une réponse positive témoigne de la prise en compte des bonnes pratiques et des exigences externes.

5. **Question : Ces principes sont-ils applicables par le biais de contrats, de MA ou d'accords de partage de données avec des partenaires extérieurs ?**

   - **Motif :** La gouvernance des données dépasse les frontières de l'organisation lorsque des données sont partagées ou reçues. Cette question vise à déterminer si l'organisation veille à ce que ses principes en matière de données (en particulier ceux relatifs à la sécurité, à la protection de la vie privée, aux limites d'utilisation et à la qualité) soient contractuellement contraignants pour les parties externes. L'applicabilité est essentielle pour gérer les risques, maintenir le contrôle sur les données partagées et s'assurer que les partenaires respectent les normes nécessaires. Une réponse positive indique que les principes sont légalement intégrés dans les collaborations externes en matière de données.

Méthodes

La définition des principes qui sous-tendent un cadre de gouvernance des données n'est pas seulement un exercice technique - elle établit des valeurs fondamentales. Les principes déterminent la manière dont les décisions sont prises, les données sont traitées et la confiance est établie. Ils déterminent ce qui est considéré comme juste, légitime et efficace dans les processus de gouvernance des données. Si la définition de finalités claires donne la direction à suivre, la formulation de principes garantit que le parcours vers ces finalités reflète les valeurs et les priorités des personnes concernées.

Pour définir ces principes de manière réfléchie et inclusive, il est possible de recourir à une série de méthodes, chacune d'entre elles offrant un éclairage unique sur les valeurs qui devraient guider la gouvernance.

Fonder les principes sur les besoins et les perspectives du monde réel

- La première étape consiste à comprendre qui le cadre de gouvernance affectera et dont les voix devraient être reflétées. La **cartographie et l'implication des parties prenantes** sont essentielles à cet égard. En identifiant les individus et les groupes concernés par la manière dont les données sont utilisées, des agences gouvernementales et des entreprises privées aux organisations communautaires et aux populations marginalisées, leurs préoccupations et leurs aspirations peuvent commencer à être identifiées. Les activités d'engagement, telles que les entretiens, les ateliers ou les forums de citoyens, permettent de clarifier les principes les plus importants : La transparence est-elle une exigence essentielle ? L'inclusivité est-elle négligée ? Existe-t-il des appels à une plus grande responsabilisation ou à des protections plus fortes contre les préjudices ? Exemple : HUD Exchange's Community Engagement Toolkit (*Manuel d'engagement communautaire de HUD Exchange*).

- Les initiatives tournées vers le public, telles que les assemblées de données ou les jurys de citoyens, peuvent davantage garantir que les principes en matière de données ne sont pas simplement imposés du haut vers le bas, mais qu'ils sont élaborés en collaboration avec les personnes les plus concernées. Ces processus participatifs démocratisent la définition des principes et les ancrent dans l'expérience vécue. Exemple : L'Assemblée des données

Faire émerger les valeurs par la délibération et la conception

- Un autre outil consiste à utiliser des **Canevas d'éthique des données**, des outils structurés incitant les équipes à réfléchir aux dimensions éthiques de l'utilisation des données. Basés sur la conception centrée sur l'homme, ces canevas guident les discussions sur les risques, les déséquilibres de pouvoir et l'impact des parties prenantes, en aidant à cristalliser en termes concrets des principes tels que la transparence, la non-discrimination et l'action de l'utilisateur. Ils sont particulièrement utiles lors des premières étapes de conception de la gouvernance, lorsque les équipes ont besoin d'expliciter leurs valeurs fondamentales. Exemple : Canevas d'éthique des données

- **Les ateliers d'ingénierie des normes** constituent une autre approche intéressante. Ces sessions réunissent diverses parties prenantes — experts techniques, conseillers juridiques, représentants de la société civile — pour débattre et concilier les tensions et les compromis (p. ex. entre ouverture et confidentialité, ou entre innovation et précaution). L'objectif est de co-construire des principes qui soient non seulement significatifs, mais aussi réalistes et adaptés aux contraintes institutionnelles et sociétales. Exemple : Ingénierie des normes

Ancrer les principes dans les droits et les responsabilités

- Pour s'assurer que les principes de gouvernance des données reflètent les normes internationalement reconnues, certaines organisations réalisent des **Études d'impact sur les droits de l'homme** (EIDH). Cette méthode permet d'examiner comment une initiative en matière de données pourrait affecter les libertés fondamentales, telles que le droit à la vie privée, l'égalité ou l'accès à l'information, et de distiller les principes fondamentaux à intégrer pour prévenir les préjudices et faire respecter la justice. Une bonne pratique pour évaluer les cadres juridiques dans le cadre d'une EIDH consiste à appliquer le **test en trois parties**, qui garantit que toute ingérence dans les droits est prévue par la loi, compatible avec les normes en matière de droits de l'homme, et nécessaire et proportionnelle. Ce test permet de garantir que les dispositions légales sont conformes aux principes des droits de l'homme et qu'elles maintiennent une approche équilibrée. Exemple : Étude d'impact sur les droits de l'homme

Impliquer le public et planifier l'avenir

- Les principes de gouvernance des données ne doivent pas seulement servir les besoins institutionnels, ils doivent aussi **trouver un écho dans la société**. À cette fin, des outils tels que les **évaluations de licence sociale et la charte civique** offrent des moyens significatifs d'évaluer l'acceptabilité par le public et de créer des cadres normatifs. Par le biais de jurys de citoyens, de groupes de discussion ou d'exercices de corédaction, les communautés peuvent exprimer les valeurs qui, selon elles, devraient régir l'utilisation des données, faisant souvent émerger des priorités telles que la dignité, le consentement et l'équité de manière nouvelle et localisée. Exemple : Boîte à outils sur la licence sociale, Charte civique et Boîte à outils sur la transformation numérique pour des villes et des communautés centrées sur les personnes

Des principes à la pratique

- Même les principes les mieux définis peuvent s'avérer insuffisants s'ils ne sont pas mis en œuvre. Comme indiqué précédemment, des outils tels que les **tableaux de bord prospectifs et les cartographies entre les principes et les pratiques** aident à intégrer les principes dans les routines institutionnelles et les structures de responsabilisation. Ils veillent à ce que des valeurs telles que l'inclusion ou la proportionnalité se traduisent par des résultats mesurables et des procédures concrètes, qu'il s'agisse d'accords de partage de données, de pistes d'audit ou de normes en matière de métadonnées. Voir la Matrice d'audit interne.

Ressources pour une lecture plus approfondie :

- UIT - Banque mondiale : Plateforme de réglementation numérique, Platform's Navigating Data Governance : Guiding Tool for Regulators (*Parcourir la gouvernance des données : outil d'orientation à l'intention des autorités de régulation*). Cet article fournit des conseils pratiques aux autorités de régulation des TIC, aux autres agences de régulation (notamment aux autorités chargées de la protection des données), ainsi qu'aux parties prenantes qui s'occupent de la

gouvernance des données, en surveillant et en guidant les pratiques de gouvernance des données des organisations, en mettant l'accent sur la classification des données, l'interopérabilité des données, la disponibilité des données, la qualité et l'intégrité, l'accès et le partage des données, la sécurité des données et la protection des données et de la vie privée.

- Conseil des chefs de secrétariat des Nations Unies Fondements normatifs proposés pour la gouvernance internationale des données : Objectifs et Principes
- Division de statistique de l'ONU Principes fondamentaux de la statistique officielle, Lignes directrices pour la mise en œuvre
- One Trust Guide des données
- L'état des lieux des politiques en matière de données ouvertes
- Transparency and Accountability as Trust Builders in the African Data Governance Ecosystem (*La transparence et la responsabilisation en tant que facteurs de confiance dans l'écosystème africain de gouvernance des données*)
- Institut danois des droits de l'homme : Guide sur l'évaluation de l'impact des activités numériques sur les droits de l'homme
- Union africaine Convention sur la cybersécurité et la protection des données à caractère personnel

III. QUI

Les personnes et les processus dans la gouvernance des données

NicoElNino/Shutterstock.com

Objectif principal

Identifier les principales parties prenantes et les rôles institutionnels responsables de la gouvernance des données, et veiller à ce que des mécanismes efficaces de coordination et de responsabilisation soient instaurés.

L'importance des personnes et des processus

La gouvernance des données nécessite une délimitation claire des rôles et des responsabilités, garantissant que les personnes et les processus appropriés sont instaurés pour prendre des décisions, superviser la conformité et gérer les données tout au long de leur cycle de vie. Des cadres de gouvernance efficaces devraient établir des responsabilités claires et peuvent également créer des mécanismes de coordination, de conformité et d'examen éthique afin de garantir la transparence et la responsabilisation dans tous les processus.

Taxonomie des processus, des rôles et des responsabilités

Pour garantir une gouvernance efficace et transparente, les rôles et les processus clés suivants figurent souvent dans un cadre solide de gouvernance des données (liste non exhaustive) :

1. Développement, Coordination et Supervision :

- Ces fonctions sont chargées de définir l'orientation stratégique, d'assurer l'alignement sur les objectifs de l'organisation et de superviser la mise en œuvre des politiques de gouvernance des données.
- Exemples :
  - Conseil d'administration ou Comité de coordination : assure la gouvernance globale et l'orientation de la stratégie.
  - Responsable des données (RD) : dirige l'élaboration et l'exécution de la stratégie de gouvernance des données et veille à l'alignement sur les objectifs généraux de l'organisation. Le/La RD peut présider un conseil ou un comité national de gouvernance des données et les ministères concernés et d'autres institutions gouvernementales comptent souvent des RD ou des coordinateurs des données.
  - Autorités indépendantes de régulation des données : créent le cadre nécessaire à la gouvernance des données qui s'applique aux agences gouvernementales et/ou aux organisations du secteur privé.
  - Équipes de travail interministérielles : au sein du gouvernement, il peut y avoir des ministères ou des agences dont les questions relatives aux données constituent régulièrement une partie essentielle de leur travail. Par exemple, les ministères du Commerce traitent les données relatives aux échanges commerciaux ; les ministères de la Justice celles relatives à la justice, à la criminalité et à l'emprisonnement ; les ministères des Investissements, du Travail & de l'Industrie celles économiques. Les accusations de vol de données ayant des implications en matière de sécurité peuvent être traitées par les agences nationales de sécurité. La coordination entre les différents départements et les agences intergouvernementaux est essentielle pour fournir un cadre cohérent de gouvernance des données.

2. Conformité et Arbitrage :

- • Ils veillent à ce que l'organisation respecte les lois sur la protection des données et résolvent les problèmes liés à l'utilisation abusive ou à la violation des données.
- Exemples :
  - Délégué(e) à la protection des données (DPD) : veille au respect des réglementations en matière de

confidentialité des données et gère les exigences réglementaires en matière de rapports (des fonctions similaires peuvent également être exercées par des juristes, p. ex. concernant les accords de non-divulgation).

- Les responsables de l'éthique peuvent également avoir des responsabilités en ce qui concerne la divulgation de données contraires à l'éthique, la protection des lanceurs d'alerte, etc.

3. Facilitation et Gestion :

- Ces personnes sont chargées de la gestion quotidienne des données et veillent à ce qu'elles soient correctement traitées, stockées, partagées et réutilisées d'une manière systématique, éthique et responsable.
- Exemples :
  - Intendant(e)s des données : gèrent et maintiennent la qualité, la sécurité, l'accès responsable et la réutilisation des données dans différents domaines au sein de l'organisation et avec des tiers.
  - Gardien(ne)s des données : responsables de la saisie, du stockage et de l'élimination des données pour les outils techniques d'approvisionnement en données. Ils/Elles travaillent avec les Intendant(e)s des données pour garantir la qualité de ces dernières.
  - Archivistes : gèrent la conservation à long terme des données et des documents, en veillant à ce que les dossiers historiques et les ensembles de données critiques soient stockés en toute sécurité, accessibles et correctement entretenus en vue d'une utilisation future.
  - Évaluateur(ice)s de maturité de la gouvernance des données : évaluent les capacités de gestion des données de l'organisation par rapport aux modèles de maturité de la gouvernance des données.

4. Examen et Orientation :

- Ces fonctions assurent une surveillance et une orientation éthiques afin de garantir que les processus de gouvernance des données respectent les normes morales et éthiques et sont conformes aux cadres juridiques.
- Exemples :
  - Comité d'examen éthique : évalue et oriente les décisions impliquant l'utilisation de données sensibles, en veillant au respect des principes et des cadres éthiques.
  - Équipe chargée de la conformité : vérifie la conformité avec les cadres juridiques et organisationnels existants.

Établir la preuve de la décision de gouvernance des données

La documentation du pouvoir de décision est essentielle pour la transparence et la responsabilisation. En identifiant les personnes impliquées dans chaque étape du cycle de vie des données, les organisations peuvent s'assurer que les décisions sont traçables et conformes aux normes de gouvernance.

Les rôles clés à définir sont les suivants :

- Responsable : Qui est responsable de la prise de décisions ? Qui est subordonné ? Qui est responsable de la mise en œuvre de la décision ?
- Garant : Qui est garant en dernier ressort du résultat de la décision ?
- Consulté : Qui doit être consulté avant qu'une décision ne soit prise ?
- Informé : Qui doit être informé de la décision une fois qu'elle a été prise ?
- La définition des rôles impliqués par la cartographie des parties prenantes garantit la clarté et la responsabilisation des décisions de gouvernance tout au long du cycle de vie des données, de la collecte et du traitement des données à leur partage et leur utilisation.

Tableau 4. Cartographie des parties prenantes

Modèle RACI	Planification	Collecte	Traitement	Partage	Analyse	Utilisation
Responsable						
Garant						
Consulté						
Informé						

Évaluation et questions

Direction et Mandat stratégique

1. **Question : Existe-t-il un(e) Responsable des données (RD), un(e) Délégué(e) à la protection des données (DPD), un(e) Intendant(e) des données ou un(e) Responsable senior des données équivalent ayant une mission officielle et disposant des ressources adéquates ?**

- **Motif :** L'existence d'un(e) Responsable senior des données est un signe d'engagement stratégique et garantit la responsabilisation de la gouvernance des données dans l'ensemble de l'organisation. Cette fonction confère l'autorité et les ressources nécessaires pour mettre en œuvre des initiatives en matière de données, assurer la coordination entre les départements et représenter les priorités en matière de données au niveau de la direction.

Clarté des rôles et Responsabilisation

2. **Question : Les responsabilités des principaux acteurs de la gouvernance des données (p. ex. RD, DPD, comité d'éthique de l'IA, Intendant(e) des données) sont-elles clairement définies et documentées ?**

- **Motif :** La clarté des rôles réduit la confusion, évite les lacunes ou les chevauchements de responsabilités et favorise une mise en œuvre plus efficace des politiques et des garanties. Des responsabilités clairement définies facilitent également l'intégration, la formation et la collaboration entre les équipes.

3. **Question : Des rôles ou des structures officiels ou non ont-ils été instaurés pour l'intendance des données, la surveillance de la vie privée et l'utilisation éthique des données (p. ex. des Intendant(e)s des données, des responsables de la protection de la vie privée, des comités d'éthique) ?**

- **Motif :** La présence de ces rôles, qu'ils soient officiellement institutionnalisés ou officieusement reconnus, indique la capacité d'une organisation à gérer ses responsabilités en matière d'éthique, de protection de la vie privée et d'intendance. Ces structures contribuent à garantir que l'utilisation des données n'est pas seulement légale, mais aussi responsable, éthique et conforme aux valeurs de la société.

4. **Question : Existe-t-il une autorité de régulation ou une autorité indépendante de protection des données pour superviser les pratiques en matière de confidentialité et de protection des données, en veillant au respect des lois nationales et internationales de protection des données et en défendant les droits de l'homme ?**

   - **Motif :** Une autorité de régulation indépendante veille à ce que les pratiques en matière de protection de la vie privée et de confidentialité des données ne soient pas soumises à des pressions politiques, économiques ou autres, et maintient l'intégrité de la gouvernance des données. Ces autorités de régulation jouent un rôle essentiel en obligeant les organisations à respecter les lois en vigueur, en procédant à des examens périodiques et en veillant à ce que les structures de gouvernance soient conformes aux normes internationales en matière de droits de l'homme.

Coordination interinstitutionnelle

5. **Question : Existe-t-il une collaboration (structurée ou non) sur la gouvernance des données entre les agences gouvernementales (p. ex. des comités de données interorganisations, des stratégies communes, des plateformes de données partagées) ?**

   - **Motif :** Les données couvrent souvent plusieurs départements. La collaboration interdéveloppementale ou interorganisations permet une approche cohérente et unifiée de la gouvernance, améliore l'interopérabilité des données et réduit les redondances. Elle est également essentielle pour relever des défis communs tels que la réponse aux crises ou la transformation numérique.

Implication multipartite

6. **Question : Pour les gouvernements, existe-t-il une coordination (officielle ou non) en matière de gouvernance des données avec des acteurs extérieurs au gouvernement (p. ex. le monde universitaire, la société civile, le secteur privé) ?**

   - **Motif :** L'implication des parties prenantes externes renforce la légitimité, la qualité et l'impact des cadres de gouvernance des données. Elle apporte une diversité d'expertise, contribue à combler les lacunes et garantit que la gouvernance reflète les besoins et les valeurs du public, ce qui est particulièrement important pour instaurer la confiance et permettre l'utilisation des données pour le bien public.

Méthodes et Outils

Pour une gouvernance des données efficace et responsable, il est essentiel de définir clairement les rôles et les responsabilités de chacun. Si l'on ne sait pas clairement qui est responsable de quoi, les efforts risquent d'être fragmentés, ce qui peut entraîner un manque d'efficacité, des dédoublements de fonctions ou des lacunes en matière de surveillance. Les exemples suivants présentent quelques méthodes et outils existants pouvant contribuer à identifier les acteurs clés, à répartir les responsabilités et à favoriser la coordination au sein de l'écosystème des données :

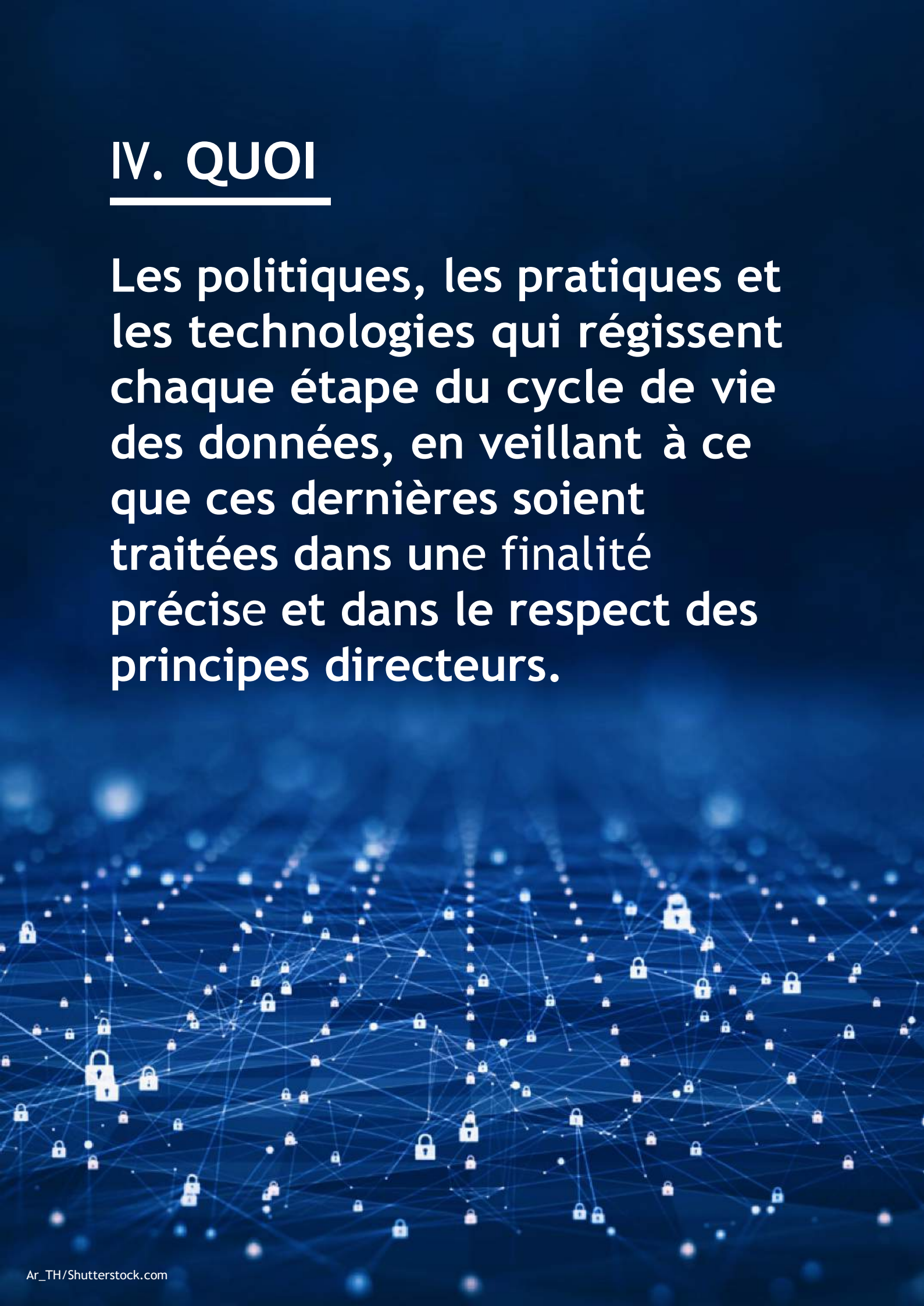
- Modèle RACI de provenance de la décision : Cet outil permet de définir et de documenter les rôles décisionnels en identifiant qui est Responsable, Garant, Consulté et Informé à chaque étape du processus de gouvernance des données.
- Cartographie de votre écosystème de données : Cet outil aide les décideurs politiques à identifier et à comprendre les acteurs clés, les sources de données et les interactions au sein d'un écosystème de données. Cartographier les flux de données, l'infrastructure et les échanges de valeur entre les institutions publiques, les organisations privées et les communautés révèle des possibilités d'amélioration de la gestion, de l'accessibilité et de la gouvernance des données.
- Cartographie des acteurs existants de la gouvernance des données : Exercice de cartographie complet pour identifier tous les acteurs actuels impliqués dans la gouvernance des données, leurs rôles et leurs responsabilités au sein de l'organisation. Cela permet de savoir clairement qui est impliqué dans les différentes étapes de gestion et de supervision des données.

Ressources pour une lecture plus approfondie :

- UIT Navigating Data Governance : A Guiding Tool for Regulators (*Parcourir la gouvernance des données : outil d'orientation à l'intention des autorités de régulation*).
- UIT Rapport technique D4.1 - Cadre pour la sécurité, la protection de la vie privée, le risque et la gouvernance dans le traitement et la gestion des données

IV. QUOI

Les politiques, les pratiques et les technologies qui régissent chaque étape du cycle de vie des données, en veillant à ce que ces dernières soient traitées dans une finalité précise et dans le respect des principes directeurs.



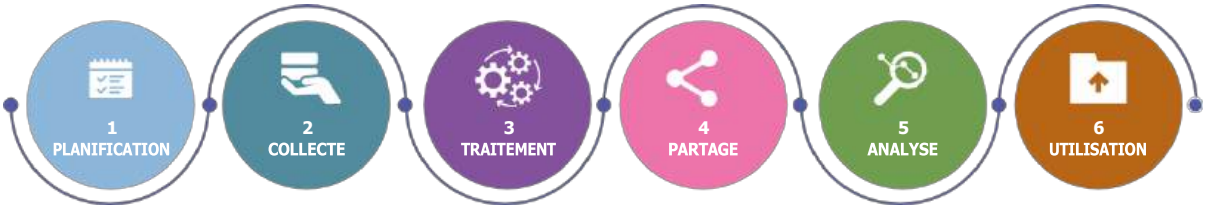
Ar_TH/Shutterstock.com

Objectif principal :

Veiller à ce que les données soient gouvernées efficacement tout au long de leur cycle de vie, de la planification et de la collecte à leur utilisation et réutilisation, afin qu'elles remplissent leur(s) finalité(s) prévue(s) tout en correspondant aux principes établis, aux normes éthiques et à la conformité réglementaire.

Analyser le cycle de vie des données ?

Comme indiqué précédemment, le cycle de vie des données fait référence aux différentes étapes par lesquelles passent les données, depuis leur planification initiale jusqu'à leur utilisation finale dans la prise de décisions. Bien que les différents cadres puissent mettre l'accent sur différentes étapes ou utiliser une terminologie différente, les phases les plus communément reconnues sont les suivantes :



- **Planification** - Identifier les besoins en données, les utilisations prévues et les exigences en matière de gouvernance.
- **Collecte** - Rassembler des données par le biais d'enquêtes, de capteurs, de transactions, de cartographie, d'imagerie ou d'autres moyens.
- **Traitement** - Nettoyage, validation, organisation, stockage et conservation des données en vue de leur utilisation, y compris leur suppression si nécessaire, et garantie d'un traitement approprié tout au long de leur cycle de vie.
- **Partage** - Rendre les données accessibles à d'autres, que ce soit par le biais de plateformes, d'IPA ou de collaborations de données.
- **Analyse** - Interpréter les données pour en tirer des enseignements.
- **Utilisation** - Appliquer ces connaissances pour éclairer les décisions, les politiques ou les services.

À chaque étape, plusieurs préoccupations et considérations doivent être prises en compte pour garantir que la gouvernance des données correspond aux principes et atteint les finalités établies. La section suivante met en lumière ces considérations, ainsi que les développements et outils récents.

Encadré 4. 10 Mécanismes de gouvernance des données

Pour mettre en œuvre les principes et les décisions de gouvernance des données tout au long de leur cycle de vie, plusieurs mécanismes peuvent être envisagés. Cela n'inclut pas la législation ou les réglementations telles que la protection des données ou la loi sur la vie privée, car elles servent de cadres juridiques généraux plutôt que d'outils opérationnels ou procéduraux pour la gouvernance au jour le jour. On peut citer :

1. **Mécanismes contractuels :**
   - Accords juridiquement contraignants qui fixent les conditions et attribuent les responsabilités en matière d'accès, de partage et d'utilisation des données, ainsi que les limites de l'interaction des tiers avec l'accès aux données et leur utilisation, les droits des tiers, le contrôle des mécanismes techniques d'accès aux données tels que les IPA, etc. Ces accords peuvent être adaptés à un cas particulier ou définis en termes généraux.
   - Exemples : **Accords de partage de données (APD), Mémorandums d'accord (MA), Accords de niveau de service (ANS), Accords de licence pour l'utilisateur final, conditions de service de l'IPA, Politiques d'utilisation de l'IPA.**

2. Politiques & Lignes directrices

- Lignes directrices institutionnelles et gouvernementales décrivant la manière dont les principes de la gouvernance des données doivent être mis en œuvre.
- Exemples : **Politiques en matière de données ouvertes, Lignes directrices en matière d'éthique de l'IA.**

3. Technologie & Gouvernance par la conception

- Solutions techniques intégrées dans les systèmes pour appliquer les principes de gouvernance.
- Exemples : **Confidentialité différentielle, apprentissage fédéré, chiffrement, contrôles d'accès, mesures techniques de protection (MTP), architectures de données et conception de l'intégration.**

4. Normes et Vocabulaire

- Protocoles et définitions communs pour documenter et garantir la qualité, la sécurité, l'interopérabilité, la cohérence et la facilité d'utilisation des données.
- Exemples : **F ISO 27001 (Sécurité de l'information), DCAT (Data Catalog Vocabulary).**

5. Codes de conduite

- Cadres volontaires ou obligatoires guidant l'utilisation responsable des données.
- Exemples : **Code de conduite de l'UE en matière de désinformation.**

6. Acquisition & Gestion des fournisseurs

- Intégrer les exigences en matière de gouvernance des données dans les processus d'acquisition.
- Exemples : **Exigences du secteur public en matière de partage de données dans les contrats de fournisseurs.**

7. Licences

- Mécanismes définissant les autorisations de réutilisation et de distribution des données.
- Exemples : **Licences Creative Commons, Licences Open Data**

8. Intendance des données & Dispositions institutionnelles

- Établir les rôles et les responsabilités en matière de gestion des données en fonction des objectifs de gouvernance.
- Exemples : **Intendant(e)s des données, Fiducies de données, Auditeurs indépendants**

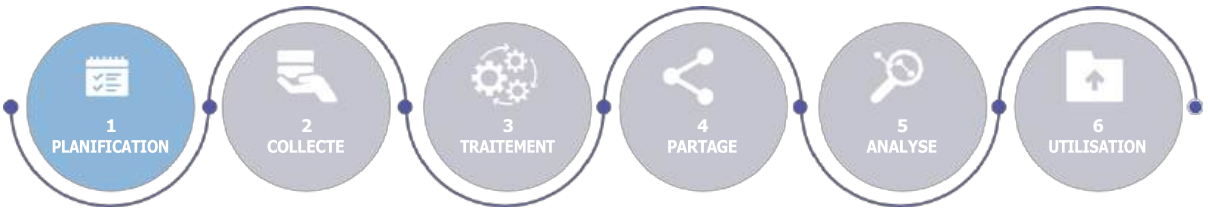
9. Mécanismes d'audit & de conformité

- Méthodes de contrôle et d'application des politiques de gouvernance.
- Exemples : **Études d'impact, Audits de conformité, Rapports sur la transparence algorithmique.**

10. Formation et Initiatives de changement culturel

- Mécanismes permettant d'intégrer la formation aux données et le changement culturel dans l'organisation, tels que des cours de formation obligatoires (protection de la vie privée, sécurité des données), sessions en personne ou virtuelles, formation à la gouvernance des données.

1. Planification



- Lors de la phase de planification, des décisions sont prises pour définir la portée, la finalité, la faisabilité et l'initiative de gouvernance des données. Ces premières étapes déterminent fondamentalement le succès et l'intégrité de l'utilisation des données par la suite. Une planification efficace nécessite de comprendre la valeur publique prévue des données, en cartographiant les acteurs et les systèmes clés, en identifiant des lacunes en matière de gouvernance et la conception de processus conformes aux réalités juridiques, éthiques et contextuelles.
- Cette étape implique également la réalisation d'études précoces des risques et des coûts-bénéfices, la définition des rôles et des responsabilités, et l'établissement de modèles de gouvernance qui seront mis en œuvre tout au long du cycle de vie des données. Une attention particulière doit être accordée à l'alignement sur les principes de limitation de la finalité, de minimisation des données, d'interopérabilité et de fiabilité, en particulier dans les environnements impliquant une collaboration intersectorielle ou transfrontalière.

Tableau 5. Liste de contrôle : Planification

Tâche	
La finalité et la valeur des données sont clairement définies et documentées	☐
Les parties prenantes et les communautés concernées ont été identifiées et cartographiées	☐
Le contexte juridique et politique, notamment les efforts antérieurs, ont été examinés	☐
Le modèle de gouvernance (rôles, responsabilités, prise de décisions) est conçu	☐
La portée, les objectifs et les limites du projet de données sont définis	☐
Les ressources financières, techniques et humaines sont évaluées et garanties	☐
Des accords de partage de données (MA, APD) et des modèles juridiques sont préparés	☐
Les besoins en matière d'interopérabilité et les vocabulaires partagés ont été examinés	☐
Des études de risques (p. ex. protection de la vie privée, sécurité, risques liés à l'IA) ont été effectuées	☐
Une stratégie d'engagement des parties prenantes et d'instauration d'un climat de confiance a été mise en place	☐
Un plan de communication et de transparence a été élaboré	☐
Des boucles de rétroaction et des mesures d'évaluation ont été définies	☐
Les besoins en matière de conservation et de stockage des données ont été évalués et planifiés	☐

10 questions d'évaluation et leurs motifs

1. La(es) finalité(s) de collecte et d'utilisation des données a/ont-elle(s) été clairement définie(s) ?

Motif : Faire en sorte que la finalité corresponde aux objectifs et aux principes sociétaux, tels que la limitation de la finalité, pour éviter le détournement de fonction ou l'utilisation abusive des données.

2. Avez-vous identifié et planifié les problèmes éthiques ou une opposition sociale ?

Motif : L'anticipation des questions sociétales et éthiques permet d'obtenir une licence sociale et de conserver la confiance du public.

3. Toutes les parties prenantes concernées, notamment les groupes marginalisés, sont-elles identifiées et impliquées ?

Motif : L'inclusion améliore la qualité des données, renforce la légitimité et garantit que les cadres de gouvernance reflètent des valeurs et des besoins divers.

4. Les hypothèses, les indicateurs de réussite et les critères de décision ont-ils été documentés ?

Motif : La clarté des hypothèses et des mesures d'évaluation favorise la responsabilisation et permet d'adapter le projet en fonction des nouvelles informations.

5. Existe-t-il des accords juridiques documentés couvrant l'accès aux données, leur partage et la protection de la vie privée ?

Motif : La clarté juridique renforce la confiance entre les partenaires et garantit le respect des réglementations en matière de protection des données.

6. Avez-vous évalué les coûts et les avantages de cette initiative en matière de données ?

Motif : Une bonne planification financière et des risques permet de s'assurer que les investissements dans les données apportent la valeur publique attendue tout en atténuant les inconvénients potentiels.

7. Des exigences techniques ont-elles été définies pour l'interopérabilité et l'évolutivité ?

Motif : La planification de l'intégration et de l'évolutivité des données permet d'éviter les goulets d'étranglement et favorise l'utilisation à long terme des données dans tous les secteurs.

8. Les parties prenantes ont-elles convenu d'un vocabulaire commun ou d'une taxonomie ?

Motif : Un langage partagé garantit une compréhension commune et permet la découverte et la réutilisation des données.

9. Existe-t-il des procédures de révision et d'adaptation régulières de la stratégie de gouvernance des données ?

Motif : La gouvernance doit être itérative et s'adapter aux risques et opportunités émergents, ainsi qu'au retour d'information des parties prenantes.

10. Les efforts passés ou les enseignements tirés de l'analyse comparative d'initiatives similaires en matière de données ont-ils été examinés ?

Motif : L'examen des initiatives antérieures permet d'éviter les doubles emplois, d'intégrer les bonnes pratiques et de s'appuyer sur l'infrastructure existante ou les relations de confiance.

Méthodes et Outils

- Cadre d'impact des données ouvertes : Tableau périodique des éléments des données ouvertes détaillant les conditions favorables et les facteurs défavorables qui déterminent souvent l'impact des initiatives en matière de données ouvertes.
- Cadres de gestion des données : Utilisez des cadres comme le DAMA-DMBOK (Ensemble de connaissances en matière de gestion des données) pour guider des pratiques de gestion normalisées et transparentes dès le départ.
- Outils de flux de travail : Concevoir des diagrammes de processus et des flux de décision pour attribuer les responsabilités, structurer les processus de décision et clarifier les données prévues, comme le modèle « FPPP ».
- Modèles légaux : Élaborer des avant-projets d'Accords de partage de données (APD), de Mémorandums d'accord (MA) et des études d'impact sur la vie privée (EIVP) afin d'officialiser les partenariats et de garantir la conformité avec les lois sur la protection des données.
- Études de risques en matière de protection de la vie privée et de sécurité : Effectuer des études initiales à l'aide d'outils tels que les Études d'impact sur la vie privée (EIVP) afin d'identifier et d'atténuer les risques en matière de protection de la vie privée.

- Glossaires et vocabulaires partagés : Élaborer dès le départ un vocabulaire commun à l'aide de glossaires, de taxonomies et de systèmes de gestion des connaissances afin de garantir que tous les partenaires s'accordent sur les termes et les concepts clés.
- Classification des données publiques de l'UNESCO (à venir).

Ressources

PNUD Modèle de cadre de gouvernance pour le système d'identité juridique numérique

UIT Navigating Data Governance : A Guiding Tool for Regulators (*Parcourir la gouvernance des données : outil d'orientation à l'intention des autorités de régulation*)

UIT Rapport technique D4.1 - Cadre pour la sécurité, la protection de la vie privée, le risque et la gouvernance dans le traitement et la gestion des données

Infocomm Media Development Authority, Singapour (IMDA) Guide d'évaluation des données en vue de leur partage

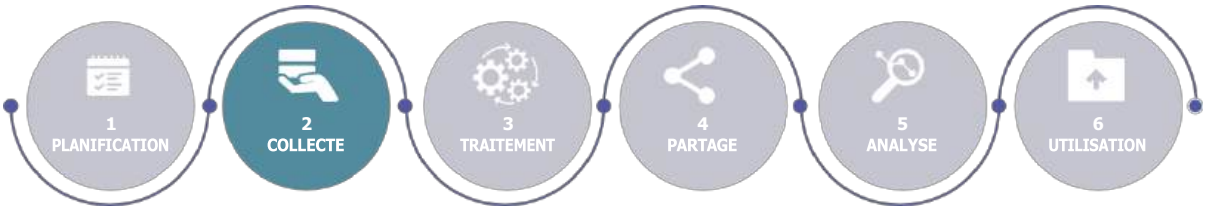
Encadré 5. Différents types de données

Les différents types de données requièrent des considérations de gouvernance distinctes afin de garantir un traitement adéquat et le respect des normes juridiques et éthiques.

- **Les données à caractère personnel (DCP)** telles que les noms, les adresses et les numéros de sécurité sociale, exigent une gouvernance plus stricte en raison des risques pour la vie privée et la sécurité. La protection des DCP passe par la mise en place de contrôles d'accès rigoureux, le chiffrement et l'adhésion à des instruments de protection des données tels que la Convention 108 du Conseil de l'Europe, ratifiée par de nombreux pays en dehors de l'espace européen. La différenciation des stratégies de gouvernance en fonction de la sensibilité des données permet de protéger les droits individuels tout en maximisant l'utilité des informations moins sensibles pour l'innovation et le bien public. L'anonymisation des données (p. ex. par le masquage, l'agrégation des données ou la pseudonymisation) peut contribuer à résoudre certaines difficultés ; toutefois, l'anonymisation doit être gérée avec prudence, car il a été souligné que l'identité peut être déduite (p. ex. par des attaques de lien, des références croisées ou des liens entre les données, en particulier les données corrélées). La pseudonymisation est souvent assimilée à l'anonymisation des données, mais elle n'est pas toujours efficace - Why Are Legacy Data Anonymization Techniques Failing?) (*Pourquoi les techniques classiques d'anonymisation des données échouent-elles ?*)
- À l'inverse, les **données à caractère non personnel**, telles que les ensembles de données agrégées ou anonymisées, ou les ensembles de données relatives à des objets ou à des lieux (p. ex. les données sur les bâtiments, les stocks de médicaments, etc.), bien qu'elles posent moins de problèmes en matière de protection de la vie privée, nécessitent toujours une gouvernance rigoureuse pour maintenir la qualité des données, garantir une utilisation éthique et favoriser la transparence.
- Les **données non structurées**, qui n'ont pas de format prédéfini ou de cadre organisationnel, nécessitent des outils et des stratégies spécialisés pour libérer leur potentiel. Du point de vue de la planification, cela implique différents défis tels que la découverte et la classification des données, leur intégration, les considérations relatives à l'évolutivité et au stockage, le respect de la vie privée et de la sécurité, ainsi que l'affectation des ressources pour gérer tous les défis potentiels.

Il est toutefois important de signaler que le niveau de gouvernance doit être basé sur la sensibilité des données et pas seulement sur leur type. Par exemple, les données à caractère non personnel peuvent également être très sensibles en raison de leur valeur commerciale, de leurs implications éthiques ou des exigences réglementaires.

2. Collecte



- L'étape de collecte détermine quelles données seront recueillies, auprès de qui, par quels moyens et dans quelles conditions juridiques et éthiques. Il s'agit de la première étape pour garantir une utilisation responsable et légale des données. À ce stade, la bonne gouvernance des données se concentre sur la légalité (respect des lois et des règlements nationaux et régionaux), la proportionnalité (collecte de ce qui est strictement nécessaire à des fins spécifiques), le consentement (individuel et/ou collectif), la transparence (informer les personnes concernées), l'équité (représentation inclusive dans les sources de données) et l'éthique (veiller à ce que l'utilisation des données soit conforme aux principes moraux tels que l'équité, la responsabilisation et le respect des droits).
- Traditionnellement, en termes de statistiques, il existait un compromis entre la quantité de données collectées et la précision des résultats statistiques (p. ex. pour déduire les caractéristiques d'une population à partir d'un échantillon de taille limitée), ainsi que les coûts de la collecte de données (p. ex. des méthodes d'échantillonnage moins coûteuses par rapport au recensement de la population). Une priorité considérable a été accordée, dans le domaine des statistiques (ainsi que dans les offices de statistique et les cabinets de conseil), à la réduction des coûts d'échantillonnage, tout en cherchant à préserver la représentativité et l'exactitude. Toutefois, les techniques récentes, les données déduites et l'analyse automatique peuvent contribuer à résoudre certains de ces problèmes.
- Les données doivent être collectées de manière à refléter la finalité et respecter les obligations en matière de protection des données. Dans le même temps, des considérations pratiques telles que l'interopérabilité, la documentation des métadonnées, la mise en œuvre de normes claires et de méthodologies standardisées pour la collecte des données et l'identification des biais sont essentielles pour garantir l'exploitabilité et la fiabilité des données par la suite.

Tableau 6. Liste de contrôle : Collecte

Tâche	
La collecte des données est conforme aux finalités déclarées et aux principes de minimisation	☐
Les mécanismes de consentement sont établis et documentés (notamment le consentement dynamique)	☐
Les populations marginalisées ou mal desservies sont représentées	☐
Les principes de protection des données dès la conception sont appliqués aux systèmes de collecte	☐
Les formats de données sont normalisés et interopérables	☐
Les pratiques en matière de métadonnées et de documentation sont établies	☐
Les flux transfrontaliers de données et les exigences en matière d'emplacement ont été évalués	☐

Des outils ont été instaurés pour détecter et atténuer les biais lors de la collecte des données	☐
Des méthodes de chiffrement ou d'anonymisation sont appliquées à l'étape de collecte	☐
Les obligations légales (p. ex. la base légale, les questions juridictionnelles) sont examinées et respectées	☐
Des mécanismes de retour d'information ont été instaurés pour adapter les processus de collecte en cas de problème	☐

10 questions d'évaluation et leurs motifs

1. Les méthodes et les pratiques de collecte de données respectent-elles le principe de proportionnalité (uniquement ce qui est nécessaire) ?

Motif : La minimisation de la collecte de données peut réduire les risques pour la vie privée, diminuer les coûts de traitement et correspondre aux exigences légales.

2. Un consentement valable et éclairé a-t-il été obtenu et enregistré le cas échéant ?

Motif : Le consentement renforce la légitimité et répond aux obligations légales, ce qui est particulièrement important pour les données sensibles ou à caractère personnel.

3. Les droits et les perspectives des groupes en situation de marginalisation sont-ils pris en compte dans la conception de la collecte des données ?

La **représentation** garantit que les données étayent l'élaboration de politiques pour tous et évite de renforcer les préjugés structurels.

4. Existe-t-il un cadre de protection des données dès la conception appliqué aux outils ou plateformes de collecte de données ?

Motif : L'intégration de la protection de la vie privée dès le départ renforce les protections et évite des ajustements rétroactifs coûteux.

5. Les métadonnées et les informations contextuelles sont-elles collectées en même temps que les données brutes ?

Motif : Les métadonnées sont essentielles pour évaluer la pertinence et la qualité et permettre une réutilisation future responsable.

6. Des formats et des protocoles normalisés sont-ils utilisés pour favoriser l'interopérabilité des données ?

Motif : L'interopérabilité renforce le potentiel de collaboration, de réutilisation intersectorielle et d'intégration des systèmes.

7. Des mesures de sauvegarde (chiffrement, anonymisation) sont-elles appliquées à l'étape de collecte ?

Motif : La protection des données sensibles à un stade précoce réduit l'exposition aux violations et favorise un traitement sécurisé du cycle de vie.

8. Des outils ou des méthodes ont-ils été instaurés pour détecter les biais potentiels lors de la collecte des données ?

Motif : L'identification précoce des biais permet de corriger le tir et d'améliorer l'impartialité de l'analyse ultérieure.

9. Les règles locales et internationales de protection des données sont-elles prises en compte dans la collecte transfrontalière ?

Motif : Comprendre les exigences juridictionnelles permet de garantir la conformité et d'éviter les risques juridiques lorsque les données franchissent les frontières.

10. Le processus de collecte des données est-il régulièrement revu et mis à jour pour garantir le respect des obligations légales ?

Motif : L'examen continu permet de corriger les oublis et d'adapter les pratiques aux nouveaux risques ou aux nouvelles opportunités.

Méthodes et Outils

Pratiques de collecte

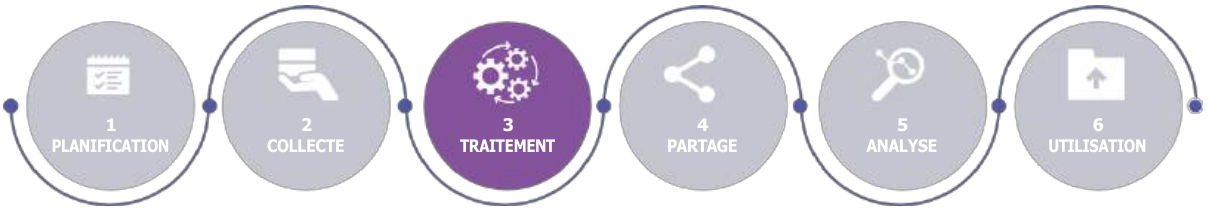
- Outils de minimisation des données : Mettre en place des systèmes garantissant que seules les données essentielles sont collectées, grâce à des techniques telles que la confidentialité différentielle pour minimiser la collecte de données inutiles.
- Plateformes de gestion des consentements : Utiliser des outils permettant d'obtenir, de gérer et d'enregistrer un consentement valable de la part des personnes concernées, en veillant à ce que les processus de consentement soient transparents et bien documentés.

- Études d'impact sur la vie privée (EIVP) : Réaliser régulièrement des EIVP pour évaluer les risques en matière de protection de la vie privée et veiller à ce que les processus de collecte des données soient conformes aux normes éthiques et juridiques. Par exemple, utiliser les outils de prise de décisions RD4C pour identifier et traiter les risques liés à la collecte de données. Ces cadres peuvent guider une conception respectueuse de la vie privée, même en dehors du contexte des données relatives aux enfants.
- Collecte de données d'intelligence collective : Pour les organisations du secteur public en particulier, la mise en œuvre de méthodes de collecte participatives telles que la production participative, les sciences participatives et les contributions basées sur des capteurs peut améliorer l'exhaustivité et la légitimité des données, en particulier pour les ensembles de données relatives à l'environnement, à la santé et à la communauté. Toutefois, l'utilisation de méthodes d'enquête et d'échantillonnage en ligne peut également introduire certains biais (p. ex., il peut être difficile de faire des inférences sur les populations âgées, malades ou hors ligne à partir d'enquêtes en ligne).
- Cadres de protection des données dès la conception : Adopter des cadres qui intègrent les considérations relatives au respect de la vie privée dans la conception des systèmes de collecte de données dès le départ, en garantissant une protection solide de la vie privée tout au long du cycle de vie.
- Outils de conformité réglementaire : Utiliser des systèmes de conformité pour garantir le respect des lois locales, nationales et internationales en matière de protection des données.
- Évaluer les ressources de données
- Outils de conservation de données : Ils peuvent contribuer à définir, à respecter et à soutenir l'adoption de normes appropriées.
- Outils d'inventaire des données : Utiliser des outils et des plateformes pour auditer et cataloguer les ensembles de données existants, en veillant à ce que toutes les données disponibles soient correctement suivies et organisées.
- Analyse du contexte des données : Mettre en place des outils permettant d'évaluer la pertinence, l'exactitude et l'applicabilité des données par rapport au cas d'usage spécifique ou au problème traité.
- Outils de gestion de la qualité des données : Utiliser des systèmes pour évaluer et maintenir la qualité des données, en garantissant leur exhaustivité, leur cohérence et leur exactitude tout au long de leur cycle de vie.
- Outils de détection des biais : Appliquer des algorithmes de détection des biais et consulter des experts externes pour identifier et traiter les biais potentiels dans les données collectées.
- Sauvegarde des données et Documentation : Mettre en place des plateformes qui documentent le processus de collecte des données, notamment les métadonnées, les limites et les biais, afin de promouvoir la transparence et la responsabilisation.

Ressources

- Royal Society From privacy to partnership: the role of Privacy Enhancing Technologies in data governance and collaborative analysis (*De la protection de la vie privée au partenariat : rôle des technologies de renforcement de la confidentialité dans la gouvernance des données et l'analyse collaborative*)
- UIT Navigating Data Governance : A Guiding Tool for Regulators (*Parcourir la gouvernance des données : outil d'orientation à l'intention des autorités de régulation*)
- UNESCO Groupes invisibles aux données et minimisation des données dans le déploiement de solutions d'IA : note d'orientation

3. Traitement



L'étape du traitement consiste à préparer les données collectées en vue de leur stockage, de leur analyse ou de leur échange. Cela comprend des tâches telles que le nettoyage, la catégorisation, la transformation, la sécurisation et la validation des données. À ce stade, les décisions de gouvernance peuvent contribuer à déterminer dans quelle mesure les données conservent leur intégrité et leur utilité, dans quelle mesure elles sont gérées de manière sécurisée et sont conformes aux normes juridiques et éthiques.

Le traitement des données prend souvent plus de temps et s'avère plus ardu que prévu, malgré la plus grande disponibilité de la puissance de calcul et des ressources. Il implique des considérations de haut niveau sur qui et pourquoi, ainsi que des considérations plus matérielles et détaillées (p. ex. la capacité de traitement hors site/sur site, la largeur de bande, les capacités de transmission, l'alimentation, la sauvegarde, la gestion des dossiers, le stockage).

Le traitement des données doit trouver un équilibre entre le coût, la facilité d'utilisation et la protection, en veillant à ce que les données soient adaptées à la finalité visée tout en minimisant les risques tels que la réidentification, l'accès non autorisé ou l'utilisation abusive. Il comprend également des décisions relatives à l'architecture du système (p. ex. le stockage centralisé ou décentralisé), à l'interopérabilité des données et à l'évolutivité à long terme.

Des politiques de conservation des archives et des données publiques, ainsi que des lignes directrices structurées pour l'archivage et la conservation des données, devraient être établies pour garantir l'accessibilité à long terme et la conformité avec les cadres juridiques et réglementaires.

Tableau 7. Liste de contrôle : Traitement

Tâche	
Les processus de nettoyage et de transformation des données sont documentés	☐
La qualité des données (leur exactitude, cohérence, exhaustivité) est validée	☐
Les données sont catégorisées et classées (p. ex. sensibles ou non)	☐
La provenance des données et l'historique des versions sont à jour	☐
Un chiffrement et des contrôles d'accès sont mis en œuvre	☐
Des protocoles d'accès internes ont été définis (p. ex. l'accès par paliers, les journaux d'audit)	☐
Des technologies de renforcement de la confidentialité sont envisagées ou appliquées	☐
Des systèmes de sauvegarde, d'archivage et de suppression efficaces et sécurisés ont été instaurés	☐
Des politiques de conservation des données (y compris des archives publiques) ont été établies	☐
Des systèmes d'archivage et des mécanismes de stockage à long terme ont été instaurés	☐

La compatibilité avec les normes d'interopérabilité est assurée	☐
Les activités de traitement font l'objet d'un contrôle de conformité juridique et éthique	☐
Des mécanismes pour éviter les liens involontaires entre les données ont été instaurés	☐

10 questions d'évaluation et leurs motifs

1. Les données ont-elles été nettoyées, validées et préparées à l'analyse ?

Motif : Des données nettoyées et validées garantissent la fiabilité des informations en aval et évitent la propagation des erreurs.

2. Les données sont-elles classées en fonction de leur sensibilité et de leur utilisation (p. ex. DCP, publiques, restreintes) ?

Motif : La catégorisation permet de prendre des mesures de protection appropriées et de procéder à une réutilisation responsable.

3. La provenance des données (origines, transformations et versions) est-elle suivie ?

Motif : La provenance favorise la transparence, la reproductibilité et la responsabilisation, en particulier pour les systèmes d'intelligence artificielle.

4. Des protocoles de sécurité solides sont-ils instaurés (p. ex. un chiffrement, une anonymisation, un contrôle d'accès) ?

Motif : La sécurité garantit que les données sensibles ou précieuses sont protégées contre les violations ou les modifications non autorisées.

5. Les contrôles d'accès internes sont-ils basés sur les rôles et vérifiables ?

Motif : L'accès hiérarchisé et les journaux d'audit renforcent la sécurité et permettent de responsabiliser quant à l'utilisation des données.

6. Des technologies de renforcement de la confidentialité sont-elles envisagées ou mises en œuvre au cours du traitement ?

Motif : Des technologies de l'information et de la communication, telles que le chiffrement homomorphe ou l'apprentissage fédéré, réduisent les risques pour la vie privée tout en permettant l'utilisation des données.

7. L'environnement de traitement des données est-il protégé contre les accès non autorisés ou les pannes de système ?

Motif : Les environnements de traitement sécurisés réduisent l'exposition aux violations et garantissent la continuité des activités.

8. Les protocoles de sauvegarde, de conservation, d'archivage et de suppression sécurisée sont-ils clairement définis et respectés ?

Motif : Des protocoles clairs de conservation, d'archivage et de suppression sécurisée sont essentiels pour la conformité légale, l'intégrité des données et la réduction des risques. Une bonne conservation garantit un accès à long terme aux données critiques, tandis qu'une suppression sécurisée minimise l'exposition et répond aux exigences réglementaires.

9. Les formats et méthodes de traitement sont-ils compatibles avec les besoins d'interopérabilité prévus ?

Motif : L'utilisation de formats de données et de normes communs garantit que les données traitées peuvent être réutilisées d'une plateforme à une autre.

10. Des mesures ont-elles été prises pour éviter les agrégations nuisibles ou les déductions involontaires ?

Motif : Même des données non sensibles peuvent présenter des risques lorsqu'elles sont associées - le fait d'anticiper ces risques permet d'éviter les abus.

Outils et Pratiques

- **Élaboration** de lignes directrices en matière de stockage et d'exploitation
  - Systèmes de sauvegarde et d'archivage : Mettre en œuvre des services de sauvegarde basés sur le cloud ou sur site garantissant des sauvegardes automatiques des données et fournissant des solutions d'archivage fiables.
  - Outils de suppression de données : Utiliser des outils et des méthodes de suppression de données sécurisées pour garantir la suppression complète des données sensibles et empêcher toute récupération non autorisée.

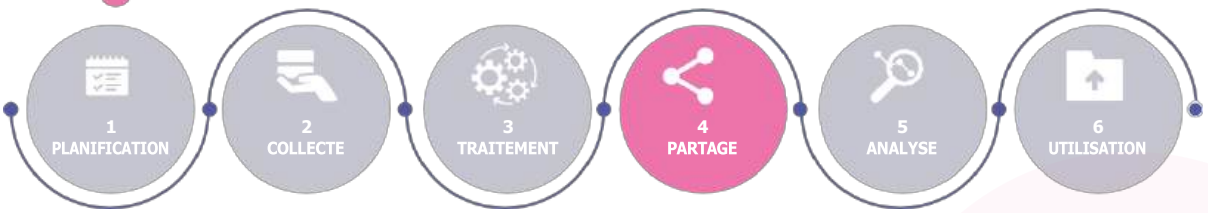
- Gestion de parc : Déployer des systèmes pour gérer, suivre et retirer en toute sécurité les appareils qui stockent des informations sensibles.

- Bâtir une sécurité robuste
  - Outils de sécurité informatique : Mettre en œuvre des mesures de sécurité avancées telles que des pare-feu, des systèmes de détection d'intrusion et des technologies de chiffrement pour protéger les données contre les accès non autorisés et les violations.
  - Expertise en matière de sécurité : Consulter régulièrement des experts internes et externes en matière de sécurité des données afin de garantir une protection solide et le respect des bonnes pratiques.
  - Formation à la sécurité : Organiser des programmes de formation pour le personnel afin de le sensibiliser et de le préparer à faire face à d'éventuels incidents liés à la sécurité des données.
  - Stratégie de communication de crise : Élaborer et tenir à jour des plans d'intervention en cas d'incident afin de réagir et de communiquer efficacement en cas de violation de données ou de crise de sécurité.
- Établir des protocoles d'accès interne & de sécurité
  - Gestion de l'accès par paliers : Mettre en place des systèmes de gestion des identités et des accès offrant différents niveaux d'accès en fonction du rôle et des responsabilités au sein de l'organisation.
  - Authentification multifacteurs : Utiliser des processus d'authentification à plusieurs facteurs pour renforcer la sécurité de l'accès aux données sensibles.
  - Pistes d'audit : Mettre en place des systèmes d'enregistrement et d'audit pour documenter et suivre qui accède aux données, quand, et quelles modifications ou actions sont effectuées au fil du temps.
  - Chiffrement des données : Garantir le chiffrement des données sensibles lors de leur traitement et de leur transmission afin de préserver la vie privée et la sécurité.
- Catégorisation, Classification & Taxonomisation des données (Architecture des données)
  - Outils de taxonomie des données : Mettre en œuvre des systèmes de catégorisation et de classification des données, en saisissant la provenance et en garantissant une organisation adéquate pour l'analyse.
  - Outils de cartographie des données : Utiliser des outils d'intégration de données et de cartographie pour transformer les ensembles de données dans des formats adaptés à l'analyse tout en maintenant la précision.
  - Outils de documentation : Établir des pratiques de documentation et de contrôle des versions pour suivre les hypothèses, les processus de nettoyage et les autres activités de traitement des données.
  - Outils pour l'intégrité des données : Déployer des mesures de protection pour assurer l'intégrité et la compatibilité des données lors de l'agrégation ou de la corrélation de différents ensembles de données, afin d'éviter les erreurs de traitement.

Ressources

- Guide de l'ASEAN sur l'anonymisation des données
- Guide des Nations Unies sur les Technologies de renforcement de la confidentialité pour les statistiques officielles
- Recommandation de l'UNESCO concernant la préservation et l'accessibilité du patrimoine documentaire, y compris le patrimoine numérique
- Prochaine classification des données publiques de l'UNESCO (à confirmer)

4. Partage



La phase de partage régit comment, avec qui et dans quelles conditions les données sont échangées. Il s'agit d'une étape cruciale pour la confiance, l'interopérabilité et la conformité légale, en particulier dans le cadre de partenariats public-privé ou de collaborations transfrontalières. L'accès aux données en vue de leur réutilisation peut débloquer une valeur collective et publique, mais soulève également des risques de réputation, juridiques et éthiques s'il est mal géré.

À ce stade, la bonne gouvernance exige des accords juridiques clairs, des garanties techniques, une communication transparente et une compréhension commune des finalités par les partenaires. Elle doit également tenir compte de la diversité des modèles de partage des données, des portails ouverts aux échanges fédérés et aux collaborations en matière de données, et anticiper les conséquences imprévues telles que la surexposition ou la mauvaise utilisation des ensembles de données combinés.

Le partage et l'échange de données ne font pas référence à la vente de données à des fins financières ou commerciales, comme la vente de données à caractère personnel à des spécialistes du marketing, à des publicitaires ou à des vendeurs tiers (p. ex. la vente de numéros de téléphone à des publicitaires ou à des expéditeurs de spams). Toutefois, il est important de faire la distinction entre ces ventes commerciales et les situations où l'accès aux données peut être fourni contre rémunération, par exemple à des fins universitaires, de recherche, ou dans les cas où les données sont utilisées à des fins non commerciales qui contribuent au bien public ou à l'innovation. Ces scénarios peuvent toujours impliquer des accords de partage de données, mais sont généralement soumis à des considérations et à des réglementations différentes, notamment en ce qui concerne la transparence, le consentement et la finalité de l'utilisation des données.

Tableau 8. Liste de contrôle : Partage

Tâche	
La finalité et la portée du partage des données sont clairement définies et convenues	☐
Les accords juridiques (p. ex. les APD, les MA, les licences) ont été formalisés	☐
Les parties prenantes internes sont prises en compte et informées des attentes en matière de gouvernance	☐
Une stratégie de communication externe a été développée	☐
Les risques de réputation et les plans d'atténuation ont été identifiés	☐
Des protocoles d'interopérabilité (p. ex. un vocabulaire et des formats communs) ont été établis	☐
Des garanties techniques (p. ex. des IPA sécurisées, un accès fédéré) ont été instaurées	☐
Les risques pour la vie privée et la sécurité liés à la combinaison des données ont été anticipés	☐
La gouvernance des intermédiaires tiers ou des plateformes a été clarifiée	☐
La licence sociale et des considérations éthiques ont été examinées	☐
Des mécanismes de coordination et d'examen permanents ont été mis en place	☐

10 questions d'évaluation et leurs motifs

1. La finalité du partage des données est-elle clairement définie et documentée ?

Motif : La clarté de la finalité permet d'éviter les abus, de garantir la pertinence et de soutenir la conformité légale, par exemple la limitation des finalités.

2. Des accords juridiques ont-ils été conclus pour régir les droits, les responsabilités et l'accès ?

Motif : Les contrats protègent les parties concernées, clarifient les responsabilités et garantissent le respect de la loi et de l'éthique.

3. Les parties prenantes internes (p. ex. les équipes juridiques, informatiques et éthiques) ont-elles été consultées et prises en compte ?

Motif : L'alignement interfonctionnel garantit que les risques opérationnels, juridiques et de réputation sont correctement gérés.

4. Les communications externes expliquent-elles clairement pourquoi les données sont partagées et avec qui ?

Motif : La transparence renforce la confiance du public et protège contre la perception d'abus ou d'intentions cachées.

5. Les risques de réputation sont-ils évalués, notamment ceux liés au partenaire de partage ou au cas d'usage ?

Motif : La gestion des risques protège la crédibilité institutionnelle et anticipe la réaction du public.

6. Les formats de données, les vocabulaires et les protocoles d'accès sont-ils normalisés pour assurer l'interopérabilité ?

Motif : L'alignement technique permet d'éviter la fragmentation des données et d'améliorer leur réutilisation dans les différents systèmes ou secteurs.

7. Les risques de violation de la vie privée, de sécurité et de réidentification liés à la combinaison des ensembles de données ont-ils été évalués ?

Motif : Les ensembles de données combinés peuvent révéler par inadvertance des informations sensibles - ces risques devant être gérés de manière préventive.

8. Existe-t-il des structures de gouvernance claires pour les intermédiaires ou les plateformes partagées (p. ex. les fiducies de données) ?

Motif : Les plateformes tierces doivent être régies de manière à ce qu'elles ne contournent pas les règles et n'affaiblissent pas la responsabilisation.

9. La licence sociale a-t-elle été prise en compte au-delà du consentement individuel ?

Motif : Les attentes de la communauté ou de la collectivité (en particulier dans les domaines sensibles) doivent être respectées pour éviter les réactions négatives.

10. Existe-t-il un plan d'évaluation continue et de correction de trajectoire dans le cadre du partenariat de partage ?

Motif : Une supervision permanente permet de s'assurer que le partenariat est pris en compte, qu'il est efficace et qu'il s'adapte au changement.

Encadré 6. Collaborations sur les données

- Collaborations en matière de données : offrent des moyens structurés pour l'échange de données entre les secteurs public et privé tout en répondant aux préoccupations en matière de confiance, de respect de la vie privée et de concurrence. Les différents modèles permettent différents niveaux de contrôle et de gouvernance :
- Mise en commun des données : Les organisations combinent les données dans un pool partagé pour l'analyse collaborative ou l'innovation, mais une gouvernance solide est essentielle pour assurer la protection de la vie privée et de la propriété intellectuelle.
- Intermédiaires de données : Des tiers neutres gèrent l'échange de données, appliquent les règles de gouvernance et garantissent la conformité, réduisant ainsi les risques pour les entités partageant des données.
- Partage d'analyses : Les organisations partagent des informations ou des analyses dérivées de leurs données sans exposer les données brutes, ce qui permet de collaborer tout en gardant le contrôle.
- Défis et Prix : Encouragent la résolution de problèmes complexes à l'aide de données partagées par le biais de concours offrant des récompenses ou une reconnaissance.
- Partenariats de recherche : Les secteurs public et privé soutiennent le partage de données pour la recherche universitaire ou politique, souvent régi par des accords garantissant un traitement éthique et la protection des droits de propriété.
- Interfaces publiques : Fournissent un accès libre à des ensembles de données ou à des outils spécifiques, permettant à un large éventail de parties prenantes — chercheurs, développeurs et organisations de la société civile — d'analyser les données et de générer des idées.

Encadré 7. FOCUS : Emplacement des données dans l'écosystème mobile

- **Énoncé du problème :**
  Le paysage des flux transfrontaliers de données (FTD) est de plus en plus complexe. L'état actuel de la localisation des données dans le contexte des FTD (flux transfrontaliers de données) est fragmenté en raison de la divergence des cadres politiques et réglementaires. Certaines juridictions favorisent l'emplacement des données sur la base de la souveraineté nationale, de la sécurité et d'une politique publique légitime. Certaines juridictions sont favorables à l'emplacement des données sur la base d'exigences législatives sectorielles, tandis que d'autres ne sont pas favorables à l'emplacement des données sur la base d'une confiance partagée, d'une responsabilisation, d'une responsabilité et de cadres harmonisés.

- **Motivation :**
  Les opérateurs de réseaux mobiles (ORM) traitent de volumineux ensembles de données relatives aux opérations commerciales et aux consommateurs dans l'ensemble de l'écosystème mobile et numérique. Des initiatives sont en cours sur la libre circulation des données en toute confiance (DFFT) par le biais de cas d'usage et de solutions technologiques, d'une plus grande sécurité juridique, de mécanismes de transfert complémentaires et d'une coopération réglementaire internationale en matière de confidentialité des données.

  Source GSMA :
  Mobile Policy Handbook
  Cross-Border Data Flows: The impact of data localisation on IoT *(Manuel de politique mobile - Flux transfrontaliers de données : l'impact de l'emplacement des données sur l'IoT)*

  Source pertinente :
  Moving Forward on Data Free Flow with Trust (DFFT) *(Aller de l'avant en matière de libre circulation des données en toute confiance (DFFT))*

- **Impact de l'emplacement des données sur les services d'argent mobile en Afrique et en Asie :**
  Dans le contexte de l'argent mobile, les opérateurs d'argent mobile sont confrontés à plusieurs défis : 1) Des procédures de conformité complexes qui entravent les activités des opérateurs d'argent mobile ; 2) Une efficacité réduite des opérations d'argent mobile en raison de la lenteur des procédures d'approbation ; 3) Des contraintes sur l'expansion des opérateurs d'argent mobile dans d'autres régions ; 4) Un accès limité à la technologie disponible pour les opérateurs d'argent mobile ; 5) Une augmentation des coûts d'exploitation ; 6) Une limite du partage d'informations pour la prévention de la fraude qui augmente les risques en matière de cybersécurité.

  Source GSMA :
  Cross-border data flows: impact of data localisation on mobile money services in Africa and Asia *(Flux transfrontaliers de données : l'impact de l'emplacement des données sur les services d'argent mobile en Afrique et en Asie)*

À paraître :
The GSMA Ministerial Programme 2025: Data Privacy Track *(Le programme ministériel 2025 de la GSMA : Protection de la vie privée)*

Outils et Pratiques

- Gestion du risque de réputation : Mettre en œuvre des outils et des pratiques permettant de surveiller les risques pour la réputation liés au partage des données, en garantissant la transparence et en répondant à toute préoccupation susceptible de nuire à la crédibilité de la collaboration.
- Systèmes de contrôle de la protection de la vie privée : Engager des experts internes en matière de protection de la vie privée et mettre en place des systèmes pour veiller à ce que le traitement des données soit conforme aux réglementations et aux normes applicables en matière de protection de la vie privée.
- Normes et solutions d'interopérabilité : Veiller à ce que les données partagées entre les organisations respectent des formats et des protocoles normalisés, ce qui permet de faciliter

une intégration et une compatibilité fluides entre différents systèmes.

- Atténuation des risques : Appliquer des méthodes d'évaluation et d'atténuation des risques, tels que les atteintes à la vie privée ou les problèmes de sécurité, en particulier lors de la combinaison d'ensembles de données susceptibles d'avoir des conséquences inattendues.

Ressources pour une lecture plus approfondie

- Association des nations de l'Asie du Sud-Est (ASEAN), "Model Contractual Clauses (ASEAN MCCs) and Ibero-American Data Protection Network's Model Contractual Clauses (RIPD MCCs)" *(Clauses contractuelles types (CCT de l'ASEAN) et clauses contractuelles types du Réseau ibéro-américain de protection des données (CCT du RIPD), Janvier 2025.*
- Organisation de coopération numérique (OCN), "Enabling Cross-Border Data Flows Amongst the Digital Cooperation Organization Member States" *(Permettre les flux transfrontaliers de données entre les États membres de l'Organisation de coopération numérique)*. Octobre 2023.
- Union européenne, Contrat type pour les transferts de données
- UIT Navigating Data Governance: A Guiding Tool for Regulators (*Parcourir la gouvernance des données : outil d'orientation à l'intention des autorités de régulation*)
- Infocomm Media Development Authority, Singapour (IMDA), Trusted Data Sharing Framework (*Cadre de partage des données de confiance*)
- HCR. « Passerelle d'interopérabilité PRIMES (PING) »

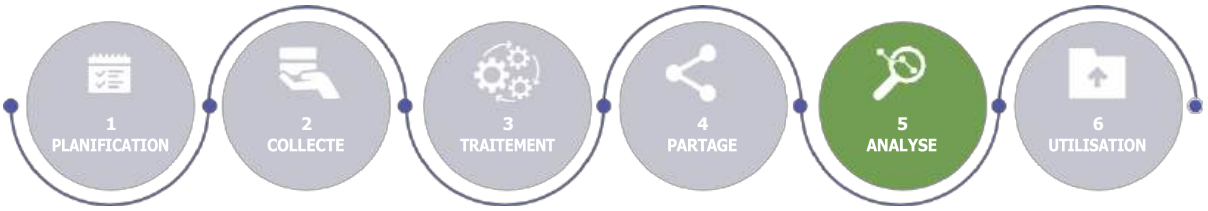
Encadré 8. Infrastructure publique numérique (IPN)

L'**infrastructure publique numérique (IPN)** est une approche émergente de la transformation numérique pouvant améliorer la prestation de services à l'échelle de la société. L'IPN est définie comme un ensemble de systèmes numériques partagés, sécurisés, interopérables et basés sur des technologies ouvertes, garantissant un accès équitable aux services publics et privés à l'échelle de la société. L'IPN utilise des normes communes et des composants technologiques réutilisables. Les **catégories de base de l'IPN** constituent la structure de base permettant l'interopérabilité et la réutilisation entre les systèmes et les cas d'usage.

À cet égard, le Manuel peut contribuer à améliorer l'IPN de la manière suivante :

- **Confidentialité des données, sécurité et consentement :** Le Manuel promeut la confidentialité et l'utilisation éthique des données en recommandant l'utilisation d'outils de minimisation des données et de gestion du consentement. De même, dans le contexte de l'IPN, cela signifie que les institutions gouvernementales ne peuvent généralement accéder qu'aux données nécessaires, ce qui réduit les risques et favorise la confiance. Les cadres de protection des données dès la conception garantissent que les systèmes sont configurés de manière à assurer la sécurité et la protection de la vie privée dès le départ, ce qui constitue un élément fondamental de l'IPN. Les plateformes de gestion du consentement peuvent contribuer à garantir la transparence de la collecte et du stockage des données, ce qui est crucial dans le cadre de l'IPN, où la confiance du public est essentielle.
- **Qualité et intégrité des données :** Le Manuel fournit des outils pour l'inventaire des données, la gestion de la qualité et la détection des biais, garantissant que les données échangées au sein des systèmes de l'IPN sont pertinentes, exactes et impartiales. Le maintien de l'intégrité des données est essentiel pour l'IPN, car les entités du gouvernement et du secteur public s'appuient sur les mêmes sources de données pour l'élaboration des politiques et la prestation des services.
- **Interopérabilité :** Le Manuel préconise le partage des données entre les organisations grâce à des formats et des protocoles normalisés, ce qui permet une intégration et une compatibilité harmonieuses entre les différents systèmes. Les systèmes de l'IPN sont techniquement interopérables, ce qui permet d'accéder efficacement aux données, de les utiliser et de les réutiliser dans tous les secteurs.
- **Environnements de stockage et de traitement des données :** Les environnements de stockage et de traitement sécurisés, la gestion des accès et les pratiques de catégorisation des données créent une infrastructure solide pour l'IPN. Cela permet de s'assurer que les données ne sont pas modifiées lors du traitement et qu'elles restent sécurisées.

5. Analyse



La phase d'analyse consiste à interpréter les données afin de générer des informations, d'éclairer les décisions et d'orienter les mesures politiques ou programmatiques. C'est à ce stade que les finalités et le potentiel des données peuvent être atteints, mais aussi que des risques d'utilisation abusive, de partialité ou d'opacité dans la prise de décisions peuvent apparaître. De plus en plus, l'analyse des données inclut des modèles d'IA et d'apprentissage automatique, qui nécessitent des garanties supplémentaires.

Compte tenu de la sensibilité de certains types et méthodes d'analyse statistique, il est essentiel de prendre en considération des facteurs tels que les valeurs aberrantes, la robustesse, le biais algorithmique, l'analyse de sensibilité et le biais de la variable omise (certains types d'analyse sont sensibles aux facteurs et aux données qui ont été omis des données, ainsi qu'à ceux qui ont été inclus dans les échantillons de données). Le choix de la ou des méthodes de traitement et d'analyse a des conséquences importantes sur les résultats, leur robustesse et leur sensibilité. À ce stade, les principales considérations de gouvernance consistent à aligner l'analyse sur la(es) finalité(s) initiales, à garantir la transparence méthodologique, à vérifier l'absence de biais et à maintenir un contrôle humain. Les implications éthiques doivent faire l'objet d'un suivi permanent, en particulier lorsque les informations peuvent affecter les droits individuels, les politiques publiques ou les résultats sociaux.

Tableau 9. Liste de contrôle : Analyse

Tâche	
Le plan d'analyse correspond aux objectifs initiaux du projet	☐
Les méthodes analytiques, les algorithmes et les hypothèses sont documentés	☐
Des tests de biais ont été effectués sur les données et les modèles	☐
Des mécanismes d'interprétabilité et d'explicabilité ont été appliqués	☐
Les résultats ont été validés par rapport aux observations du monde réel ou à la vérité terrain	☐
Des mécanismes de contrôle humain ont été instaurés pour les décisions importantes	☐
Un examen éthique de l'analyse a été effectué (en particulier pour les systèmes d'intelligence artificielle)	☐
Les données synthétiques ou les méthodes de préservation de la vie privée sont utilisées en cas de besoin	☐
L'engagement des parties prenantes ou des approches participatives ont été envisagés	☐
Les modèles ont été réentraînés ou recalibrés en cas de changements significatifs du contexte	☐

24. 10 questions d'évaluation et leurs motifs

1. L'analyse correspond-elle à la finalité initiale et à la définition du problème ?

Motif : Le respect de la finalité empêche les dérives et garantit que les résultats sont pertinents, exploitables et conformes à la législation.

2. Les méthodes d'analyse, les algorithmes et les hypothèses sont-ils documentés de manière transparente ?

Motif : La transparence permet la reproductibilité, l'audit et l'examen éclairé par les parties prenantes.

3. Les données et le modèle ont-ils été évalués en termes de biais et de représentativité ?

Motif : Les biais dans l'analyse peuvent conduire à des résultats discriminatoires ou inexacts. Les tests permettent de détecter et d'atténuer ce risque.

4. Les résultats sont-ils interprétables et explicables, en particulier dans le cas des systèmes d'intelligence artificielle ?

Motif : Le caractère explicable renforce la confiance et permet un contrôle efficace des décisions automatisées.

5. Les résultats analytiques sont-ils validés par rapport à des résultats réels ou à d'autres critères de référence ?

Motif : La validation garantit l'exactitude et la fiabilité des données et permet d'éviter les ajustements excessifs ou les fausses corrélations.

6. Le contrôle humain est-il intégré dans le processus de l'analyse à la décision ?

Motif : La conservation du jugement humain est essentielle pour la prise de décisions éthique et la correction des erreurs.

7. Les risques éthiques (p. ex. l'exclusion, la surveillance, le préjudice) ont-ils été examinés ?

Motif : L'examen éthique réduit les préjudices et garantit que l'utilisation des données est conforme aux valeurs publiques et aux droits de l'homme.

8. Les technologies de préservation de la vie privée (p. ex. les données synthétiques, les PETs) sont-elles utilisées lors de l'analyse de données sensibles ?

Motif : Ces techniques permettent une analyse pertinente tout en protégeant les droits des personnes.

9. Les parties prenantes concernées sont-elles informées ou impliquées dans l'interprétation et la mise en contexte des résultats ?

Motif : L'analyse participative renforce la pertinence et réduit le risque d'interprétation erronée et préjudiciable.

10. Les modèles sont-ils régulièrement mis à jour ou réajustés pour tenir compte des nouvelles données ou de l'évolution des conditions ?

Motif : Le contexte évolue et la mise à jour des modèles garantit une pertinence et une précision constantes au fil du temps.

Encadré 9. Intelligence artificielle (IA) et Gouvernance des données

À mesure que l'intelligence artificielle (IA), notamment l'apprentissage automatique et l'IA générative, devient partie intégrante de l'analyse moderne des données, de nouvelles considérations apparaissent aux stades du partage et de l'analyse des données. Les outils d'IA, qui s'appuient sur de vastes ensembles de données pour former des modèles, présentent à la fois des opportunités et des risques.

- D'une part, l'IA peut permettre une analyse plus efficace, nuancée et évolutive, mettant au jour des schémas et des connaissances auparavant inaccessibles.
- Cependant, l'utilisation de l'IA introduit des préoccupations concernant la qualité des données, les biais et l'interprétabilité.
- Les modèles d'apprentissage automatique peuvent amplifier involontairement les biais présents dans les données d'entraînement, tandis que l'IA générative peut créer des résultats difficiles à vérifier ou à interpréter.
- Il est essentiel de s'assurer que ces systèmes d'IA sont transparents et éthiques et que leurs décisions peuvent être comprises par les humains, en particulier lorsque les chercheurs ne savent pas exactement pourquoi ou comment leurs modèles sont arrivés aux résultats obtenus.
- En outre, les connaissances issues de l'IA doivent être rigoureusement testées par rapport aux résultats du monde réel afin d'éviter que des variables cachées ou des prédictions erronées n'influencent la prise de décisions.
- Enfin, la supervision humaine reste cruciale ; si l'IA améliore les capacités d'analyse, le maintien du contrôle humain sur les décisions finales garantit la responsabilisation et l'utilisation éthique des technologies de l'IA dans la gouvernance des données.

Outils et Pratiques

- Modélisation et visualisation des données : Concevoir des structures de données optimisées pour une analyse ciblée. Veiller à ce que les outils de visualisation des données traduisent des données complexes en informations claires et exploitables.

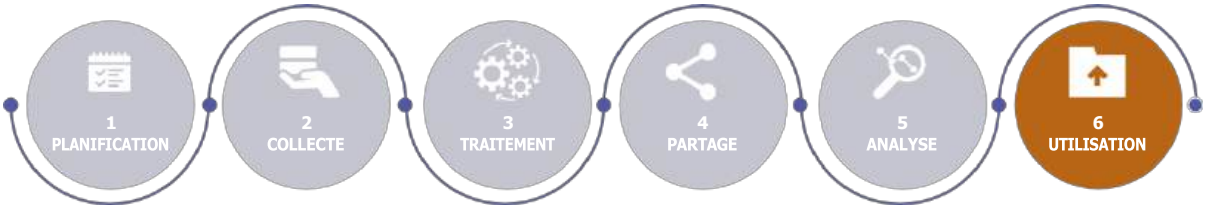
Ressources

- Navigateur du PNUD « Des données sur les politiques », AI for Public Policy (*L'IA au service des politiques publiques*)
- ASEAN Advisory Guidelines on Use Of Personal Data in AI Recommendation and Decision Systems (*Lignes directrices consultatives sur l'utilisation des données à caractère personnel dans les systèmes de recommandation et de décision basés sur l'IA*)
- ASEAN AI governance and ethics – generative AI (*Gouvernance et éthique de l'IA - IA générative*)

Cas d'usage

- Étude de cas du PNUD sur le projet de déchets marins au Ghana : L'IA peut collecter des données plus rapidement et avec plus de précision que les humains. Les approches fondées sur l'IA peuvent contribuer à améliorer la disponibilité et la qualité des données pour répondre aux besoins de suivi des ODD.
- Étude de cas du PNUD sur l'analyse des données relatives au genre : Breaking Barriers: Reinforcing Gender Data Analysis and Use with the Gender Data Lab Initiative (*Briser les barrières : renforcer l'analyse et l'utilisation des données sur le genre grâce à l'initiative « Gender Data Lab » (laboratoire de données sur le genre*))

6. Utilisation



L'étape de l'utilisation permet de traduire en actions les informations obtenues à partir des données. C'est là que les données et l'analyse qui en résulte peuvent conduire à des décisions, influencer les politiques ou informer les services publics. C'est également à ce stade que les dommages peuvent être encourus ou amplifiés si les résultats sont déformés, mal compris, mal appliqués ou utilisés au-delà de la finalité prévue (p. ex. via les réseaux sociaux).

À ce stade, la bonne gouvernance garantit que l'utilisation des données est conforme aux intentions initiales, qu'elle respecte les droits individuels et collectifs et qu'elle comprend des processus d'examen, de responsabilisation et d'apprentissage. Cela inclut les vérifications finales du consentement, de l'alignement des finalités et des garanties par rapport à toute interprétation erronée ou application discriminatoire. Il s'agit par exemple de l'examen par les pairs dans les revues universitaires, ainsi que du principe scientifique important de la vérification et de la vérifiabilité. Lorsque les résultats sont partagés ou mis en œuvre, l'importance d'une communication claire et d'une amélioration continue devient primordiale. En effet, la manière dont les données sont publiées et mises en évidence dans différentes sphères peut avoir un impact majeur. Prenons l'exemple de la publication des résultats et des taux de vaccination contre le COVID dans les revues médicales par rapport à la discussion sur la perception des taux de vaccination par le public sur les réseaux sociaux.

Tableau 10. Liste de contrôle : Utilisation

Tâche	
L'utilisation finale des données est conforme aux objectifs initiaux du projet et au consentement	☐
La conformité juridique, éthique et de protection de la vie privée a été vérifiée avant l'utilisation	☐
Les risques de conséquences involontaires (p. ex. la réidentification, les biais) ont été évalués	☐
Les résultats ont été examinés par des experts pour en vérifier la qualité et l'impartialité	☐
Une stratégie de communication et de messages claire a été établie	☐
Les actions basées sur les données ont été examinées pour évaluer leur efficacité et leur impact	☐
Des mécanismes de retour d'information ont été instaurés pour recueillir les réactions des utilisateurs/communautés	☐
Des politiques de conservation des données ou de suppression sécurisée ont été établies	☐
Les modèles ont été recyclés ou révisés sur la base de nouvelles données ou d'analyses d'impact	☐
De la documentation sur les résultats, les apprentissages et les occasions manquées a été créée	☐

10 questions d'évaluation et leurs motifs

1. Les données sont-elles utilisées conformément aux objectifs initiaux et aux accords de consentement ?

Motif : Le maintien de l'alignement permet d'éviter les dérives de la mission et de garantir une utilisation légale et éthique des données.

2. Le cas d'usage final a-t-il fait l'objet d'un examen juridique et éthique ?

Motif : Cela permet de confirmer la conformité et d'éviter les préjudices involontaires, en particulier lorsque l'utilisation des données a une incidence sur les droits ou les services.

3. Existe-t-il des processus permettant de détecter et d'atténuer les conséquences involontaires potentielles ?

Motif : L'anticipation des effets en aval réduit le risque de réactions négatives ou de préjudices de la part du public (p. ex. à la suite de décisions politiques erronées).

4. Des experts indépendants ou pluridisciplinaires ont-ils examiné les conclusions ou les résultats ?

Motif : L'examen par les pairs améliore la qualité et la crédibilité et permet de déceler les zones d'ombre.

5. Les données ou les informations sont-elles communiquées clairement, avec les mises en garde qui s'imposent ?

Motif : Une communication claire permet d'éviter les interprétations erronées et favorise un dialogue public et des décisions politiques éclairées.

6. Des mécanismes ont-ils été instaurés pour mesurer l'impact réel des décisions prises sur la base des données ?

Motif : L'évaluation des résultats garantit la responsabilisation et contribue à améliorer les efforts futurs fondés sur les données.

7. Existe-t-il un retour d'information de la part des parties prenantes ou des utilisateurs finaux sur la manière dont les données ont été utilisées ?

Motif : Le retour d'information favorise la réactivité, l'inclusion et l'amélioration des pratiques.

8. Les politiques de conservation et de destruction des données sont-elles respectées après utilisation ?

Motif : Une clôture correcte du cycle de vie favorise la protection de la vie privée, la conformité légale et la minimisation des données.

9. Les modèles analytiques sont-ils recyclés ou recalibrés au fil du temps ?

Motif : La mise à jour des modèles permet de garantir leur pertinence et d'éviter de s'appuyer sur des connaissances dépassées.

10. Les enseignements tirés du projet sont-ils documentés pour la mémoire institutionnelle ?

Motif : L'enregistrement des réussites et des échecs favorise l'apprentissage, l'élargissement et l'adaptation des politiques.

Outils et Pratiques

- Systèmes de révision complète : Mettre en place des mécanismes d'examen de la qualité des données, des résultats algorithmiques et de l'intégrité de l'analyse, afin d'éviter les erreurs ou les biais dans les rapports.
- Protocoles de protection des données : Mettre en œuvre des stratégies solides pour se prémunir contre les risques de réidentification et protéger les données sensibles, en particulier lors de leur diffusion publique.
- Outils de conservation et de destruction des données : Utiliser des systèmes de gestion des données pour définir des processus de conservation à long terme ou de suppression sécurisée, en veillant au respect des réglementations en matière de protection de la vie privée, telles que le droit à l'oubli.
- Maintenance du modèle et Recyclage : Mettre régulièrement à jour et entraîner à nouveau les modèles d'apprentissage automatique avec de nouvelles données pour que les analyses restent pertinentes et précises au fil du temps.
- Pratiques de réflexion et de documentation : Mettre en place des processus pour évaluer et documenter les résultats des projets, en se concentrant sur les succès et les occasions manquées en matière d'utilisation des données, afin d'étayer les initiatives futures.

Encadré 10. Considérations transversales sur le cycle de vie des données

Lors de la gestion des données tout au long de leur cycle de vie, il est essentiel de prendre en compte les questions transversales qui favorisent un flux continu de la collecte à l'utilisation.

- Une infrastructure de données bien structurée est fondamentale, car elle garantit que chaque étape, qu'il s'agisse de la collecte, du traitement, du partage ou de l'utilisation, bénéficie du soutien technique et organisationnel nécessaire. Il s'agit notamment d'investir dans le stockage sécurisé des données, dans des plateformes et des protocoles interopérables et dans des cadres de gouvernance qui permettent l'accès aux données et la collaboration.
- En outre, le renforcement des capacités est essentiel à tous les stades. Les parties prenantes, notamment les fonctionnaires, les Intendant(e)s de données et les analystes, ont besoin d'une formation continue pour développer leur expertise technique, leur connaissance des données et leur éducation aux considérations éthiques.
- Il est tout aussi important de favoriser une culture des données qui encourage la transparence, la collaboration et l'utilisation éthique des données. L'instauration d'une culture de la confiance et de la responsabilité, par le biais de pratiques de gestion du changement, garantit que tous les acteurs du cycle de vie, des équipes de collecte des données aux décideurs, comprennent leur rôle dans la protection de l'intégrité des données et l'utilisation des données pour obtenir des résultats significatifs.

Glossaire

Anonymisation des données

Processus de modification des données à caractère personnel de manière à empêcher l'identification des individus tout en préservant l'utilité des données pour l'analyse et la prise de décisions.

Atténuation des biais dans les données

Techniques et stratégies permettant d'identifier, de réduire ou d'éliminer les biais présents dans les ensembles de données et les algorithmes, afin de garantir que les idées et les décisions fondées sur les données sont justes et équitables.

Autodétermination numérique (ADN)

Principe selon lequel les individus et les communautés ont l'autonomie pour gérer leur présence numérique, notamment le contrôle des données à caractère personnel, des identités numériques et de la manière dont ils utilisent les technologies numériques. Ce concept met l'accent sur la capacité des utilisateurs à prendre des décisions éclairées sur la manière dont leurs données sont collectées, partagées et utilisées, en veillant à ce que les interactions numériques soient conformes à leurs valeurs et à leurs intérêts. Il englobe des dimensions à la fois individuelles et collectives, plaidant en faveur de l'action et des droits tout au long du cycle de vie des données numériques.[1]

Bacs à sable de données

Environnements contrôlés permettant de tester et d'expérimenter en toute sécurité des innovations basées sur des données, en conciliant la surveillance réglementaire nécessaire et la possibilité d'explorer de nouvelles applications.

Cadres de responsabilisation des données

Mécanismes et politiques garantissant que les organisations et les institutions assument la responsabilité de la manière dont les données sont des structures d'audit, d'établissement de rapports et de contrôle.

Cadres éthiques de l'IA

Lignes directrices et structures de gouvernance garantissant que les systèmes d'IA sont développés et déployés de manière responsable, en tenant compte de préoccupations telles que la partialité, l'équité, la transparence et la responsabilisation.

Classification de la sensibilité des données

Catégorisation des données en fonction de leur niveau de confidentialité, des exigences réglementaires et de l'impact potentiel sur les individus ou les organisations en cas d'utilisation abusive, permettant d'orienter les mesures de gouvernance appropriées.

Commun(s) de données

Modèle de gouvernance partagée permettant à de multiples parties prenantes de contribuer, d'accéder et d'utiliser un ensemble collectif de ressources de données tout en veillant à ce que les considérations éthiques, juridiques et sociales soient respectées.

Confidentialité différentielle

Technique de protection de la vie privée permettant l'analyse statistique d'ensembles de données tout en garantissant que les données individuelles ne sont pas divulguées, ce qui réduit les risques de réidentification.

Conservation des données

Rôle des organisations ou des personnes responsables de la maintenance et de la sécurisation des données pour le compte des propriétaires de données ou de l'intérêt public au sens large, en veillant à la conformité avec les cadres de gouvernance.

1 Verhulst, S, 'Operationalizing Digital Self-Determination' (*Opérationnalisation de l'autodétermination numérique*) (2023) 5 Data & Policy e14 https://doi.org/10.1017/dap.2023.11

Cycle de vie des données

Ensemble des processus d'une application qui transforme les données brutes en connaissances exploitables.[2] Ces processus comprennent la collecte, le traitement, le partage, l'analyse et l'utilisation des données régis par des politiques et des pratiques spécifiques à chaque phase.[3]

Données synthétiques

Données artificielles générées à partir de données originales et d'un modèle entraîné à reproduire les caractéristiques et la structure des données originales.[4]

Droit à la portabilité des données

Droit légal permettant aux individus d'obtenir et de transférer leurs données à caractère personnel entre les fournisseurs de services dans un format structuré, couramment utilisé et lisible par machine.

Écosystème de données

Intégration et interaction entre les différentes parties prenantes concernées, notamment les détenteurs de données, les producteurs de données, les intermédiaires de données et les personnes concernées, qui sont impliquées ou affectées par les dispositions relatives à l'accès aux données et à leur partage, en fonction de leurs différents rôles, responsabilités et droits, technologies et modèles d'entreprise.[5]

Éthique des données

Nouvelle branche de l'éthique qui étudie et évalue les problèmes moraux liés aux données (notamment la création, l'enregistrement, la conservation, le traitement, la diffusion, le partage et l'utilisation), aux algorithmes (notamment l'intelligence artificielle, les agents artificiels, l'apprentissage automatique et les robots) et aux pratiques correspondantes (notamment l'innovation responsable, la programmation, le piratage et les codes professionnels), afin de formuler et d'encourager des solutions moralement bonnes (p. ex. des conduites ou des valeurs justes).[6]

Féminisme des données

Perspective mettant en évidence la manière dont le pouvoir et les privilèges influencent la collecte, l'utilisation et la gouvernance des données, en plaidant pour des approches plus réactives et équitables de la prise de décisions fondée sur les données.

Fiducies de données

Structures juridiques et de gouvernance permettant une intendance collective des données, où les administrateurs gèrent les actifs de données au nom des bénéficiaires tout en veillant au respect des cadres éthiques et juridiques.

Gestion des consentements

Système ou cadre permettant aux individus de donner, de réviser et de retirer leur consentement pour la collecte, le traitement et le partage de leurs données à caractère personnel, garantissant ainsi le respect des réglementations en matière de protection de la vie privée telles que le RGPD.

---

2 NIST, 'Data Life Cycle' (*Cycle de vie des données*) (Glossaire NIST) https://csrc.nist.gov/glossary/term/data_life_cycle, consulté le 10 février 2025.

3 Young A, Zahuranec, A J, Verhulst, S et Gazaryan, K, The Third Wave of Open Data Toolkit: Operational Guidance on Capturing the Institutional and Societal Value of Data Re-Use (*Troisième version de la boîte à outils pour les données ouvertes : guide opérationnel pour saisir la valeur institutionnelle et sociétale de la réutilisation des données*) (The GovLab 2021) https://files.thegovlab.org/The-Third-Wave-of-Open-Data-Toolkit.pdf, consulté le 10 février 2025.

4 CEPD, Rapport Techsonar 2022-2023 (Novembre 2022) https://www.edps.europa.eu/data-protection/our-work/publications/reports/2022-11-10-techsonar-report-2022-2023_en, consulté le 10 février 2025.

5 OCDE, Recommandation du Conseil sur l'amélioration de l'accès aux données et de leur partage (Instruments juridiques de l'OCDE, 2019) https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463, consulté le 10 février 2025.

6 Floridi L et Taddeo M, 'What Is Data Ethics?' (*Qu'est-ce que l'éthique des données ?*) (2016) 374 Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 20160360, https://doi.org/10.1098/rsta.2016.0360.

Gouvernance de l'IA

Cadres, politiques et structures juridiques qui déterminent le développement, le déploiement, la surveillance et les effets sociétaux des systèmes d'IA, en veillant à ce qu'ils soient éthiques, transparents, responsables et respectueux des droits de l'homme et des valeurs sociétales. Pour fournir efficacement des systèmes d'IA responsables, ces cadres doivent être ancrés dans les principes de gouvernance des données.[7]

Gouvernance de l'identité numérique

Politiques et normes régissant la façon dont les identités numériques sont créées, authentifiées et gérées afin de garantir la sécurité, la protection de la vie privée et le contrôle des données à caractère personnel par l'utilisateur.

Gouvernance des données fédérées

Modèle décentralisé de gouvernance des données dans lequel différentes entités gardent le contrôle de leurs données respectives tout en suivant des normes et des protocoles communs pour une interopérabilité et une utilisation éthique.

Gouvernance des données transfrontalières

Politiques, cadres juridiques et normes techniques qui réglementent le mouvement, le stockage et le traitement des données dans les juridictions nationales ou régionales.

Graphes de connaissances dans la gouvernance des données

Représentations structurées des liens entre les éléments de données améliorant l'interopérabilité, l'intégration des données en fonction du contexte et le partage responsable des données.

IA juste (Équité, Responsabilisation et Transparence dans l'IA)

Cadre d'évaluation des systèmes d'IA afin de s'assurer qu'ils sont équitables, impartiaux et explicables, en alignant les pratiques de gouvernance des données sur les principes éthiques.

Infrastructure publique numérique (IPN)

Ensemble de systèmes numériques partagés qui devraient être sécurisés et interopérables et pouvant être construits sur la base de normes et de spécifications ouvertes afin de fournir un accès équitable aux services publics et/ou privés à l'échelle de la société et qui sont régis par des cadres juridiques applicables et des règles habilitantes afin de favoriser le développement, l'inclusion, l'innovation, la confiance et la concurrence et de respecter les droits de l'homme et les libertés fondamentales.[8]

Intendant(e) des données

Responsables ou équipes d'organisations habilités à créer de la valeur publique en réutilisant les données (et l'expertise en matière de données) de leur organisation, en identifiant les possibilités de collaboration intersectorielle productive et en répondant activement aux demandes externes d'accès fonctionnel aux données, au savoir ou à l'expertise. Ils sont actifs à la fois dans le secteur public et dans le secteur privé, promouvant la confiance à l'intérieur et à l'extérieur de leur organisation.[9]

Intermédiaires de données

Tiers de confiance facilitant l'échange éthique et sécurisé de données entre les fournisseurs et les utilisateurs de données, garantissant la conformité avec les cadres de gouvernance tout en maximisant l'utilité des données.

---

7 Verhulst S et Schüür F 'Interwoven Realms: Data Governance as the Bedrock for AI Governance' (*La gouvernance des données, fondement de la gouvernance de l'IA*) (Data G Policy Blog, 20 novembre 2023) https://medium.com/data-policy/interwoven-realms-data-governance-as-the-bedrock-for-ai-governance-ffd56a6a4543, consulté le 10 février 2025.

8 PNUD, Accelerating the SDGs through Digital Public Infrastructure: A Compendium of the Potential of Digital Public Infrastructure (*Accélérer la réalisation des ODD grâce aux infrastructures publiques numériques : Recueil des possibilités offertes par les infrastructures publiques numériques*) (21 août 2023) https://www.undp.org/publications/accelerating-sdgs-through-digital-public-infrastructure-compendium-potential-digital-public-infrastructure, consulté le 10 février 2025.

9 TheGovLab, 'Wanted: Data Stewards. (Re-)Defining the Roles and Responsibilities of Data Stewards for an Age of Data Collaboration' (*Recherchés : intendant(e)s de données. (Re)définir les rôles et responsabilités des intendant(e)s de données à l'ère de la collaboration en matière de données*) (2020) https://thegovlab.org/static/files/publications/wanted-data-stewards.pdf, consulté le 10 février 2025.

Interopérabilité

Capacité de deux ou plusieurs systèmes ou applications à échanger des informations et à utiliser mutuellement les informations échangées.[10]

Maillage de données

Approche de l'architecture des données consistant à répartir la propriété des informations entre des équipes interfonctionnelles qui fournissent ensuite des produits de données aux utilisateurs finaux.[11]

Métadonnées

Données sur des données.[12]

Minimisation des données

Principe selon lequel la collecte d'informations à caractère personnel doit se limiter à ce qui est directement pertinent et nécessaire pour atteindre une finalité spécifique.[13]

Monétisation des données

Processus de création de valeur économique à partir de données grâce à l'octroi de licences, aux partenariats, à la production d'idées ou au développement de produits, qui nécessite souvent des mécanismes de gouvernance pour garantir une utilisation éthique et responsable.

Provenance des données

Historique des données, détaillant leurs origines, leurs transformations et la manière dont elles ont été utilisées au fil du temps pour garantir la responsabilisation, la reproductibilité et la confiance dans la prise de décisions fondée sur les données.

Proportionnalité dans la gouvernance des données

Principe selon lequel les mesures de gouvernance des données doivent être proportionnées au niveau de risque, en équilibrant l'accès aux données et l'innovation avec la protection de la vie privée, la sécurité et les considérations éthiques.

Protection des données dès la conception

Approche visant à protéger la vie privée des individus et la protection des données par des choix de conception intentionnels. Contrairement aux méthodes traditionnelles de protection de la vie privée qui considèrent la protection de la vie privée comme une réflexion après coup, la prise en compte du respect de la vie privée dès la conception fait de la protection de la vie privée un élément central dès les premières étapes de la conception.[14]

Réciprocité des données

Principe garantissant un partage juste et équitable des données entre les entités, en mettant l'accent sur les avantages mutuels, la transparence et la gestion responsable dans le cadre d'accords de collaboration en matière de gouvernance des données.

Responsabilisation

Obligation pour les décideurs de veiller à ce que les stratégies, les politiques et les pratiques liées aux données soient participatives, transparentes et éthiques. Elle exige des responsables de la gouvernance des données qu'ils répondent de leurs actes, en favorisant la confiance, l'équité et le respect des normes éthiques et juridiques.

---

10 Organisation internationale de normalisation (ISO), ISO/IEC 19941:2017 - Technologies de l'information - Informatique en nuage - Interopérabilité et portabilité (ISO, 2017) https://www.iso.org/obp/ui/#iso:std:iso-iec:19941:ed-1:v1:en, consulté le 10 février 2025.

11 Thoughtworks, 'Data Mesh' (Maillage des données) https://www.thoughtworks.com/en-de/insights/decoder/d/data-mesh, consulté le 10 février 2025.

12 Voir (n. 2).

13 Contrôleur européen de la protection des données (CEPD), https://www.edps.europa.eu/data-protection/data-protection/glossary/d_en#data_minimization, consulté le 10 février 2025.

14 IEEE, Digital Privacy, 'What is Privacy by Design and why it is important?' (*Confidentialité numérique, Qu'est-ce que la confidentialité dès la conception et pourquoi est-elle importante ?*) https://digitalprivacy.ieee.org/publications/topics/what-is-privacy-by-design-and-why-it-s-important, consulté le 10 février 2025.

Souveraineté des données

Terme fréquemment utilisé de manière interchangeable avec la souveraineté numérique, la cybersouveraineté et la souveraineté technologique, n'ayant pas de définition claire et largement acceptée en raison de ses implications géopolitiques. D'une manière générale, il s'agit de la capacité d'une entité, qu'elle soit publique ou privée, à exercer un contrôle sur son avenir numérique à travers trois couches distinctes. Ces couches comprennent la couche physique (englobant l'infrastructure et la technologie), la couche de code (comprenant les normes, les réglementations et la conception) et la couche de données (concernant l'accès, le flux et l'utilisation des données).

Souveraineté des données autochtones

La souveraineté des données autochtones affirme le droit des communautés autochtones à régir la collecte, l'utilisation et l'intendance de leurs données conformément à leurs lois, coutumes et systèmes de connaissances. Elle remet en question les modèles centrés sur l'État et le marché en plaidant pour l'autodétermination, les pratiques éthiques en matière de données et les structures de gouvernance culturellement appropriées.

Spectre d'ouverture des données

Cadre catégorisant l'accessibilité des données le long d'un spectre allant de fermé (accès restreint) à totalement ouvert (disponible publiquement), avec des mécanismes de gouvernance définissant les niveaux appropriés d'ouverture basés sur la vie privée, la sécurité et les considérations éthiques.

Technologies de renforcement de la confidentialité (PETs)

Outils et méthodologies permettant l'analyse de données tout en minimisant l'exposition aux données à caractère personnel (DCP), parmi lesquels figurent le chiffrement homomorphe, le calcul multipartite sécurisé et la création de données synthétiques.

Transparence

Dans la gouvernance des données, selon le contexte, la transparence fait référence à l'un des principes des processus et aux principes qui guident la prise de décisions. En tant que principe des processus de gouvernance des données, la transparence garantit que les processus de gouvernance sont ouverts et compréhensibles. En tant que principe de prise de décisions, la transparence signifie communiquer clairement la raison d'être des décisions.

Transparence algorithmique

Principe selon lequel les processus de prise de décisions alimentés par des algorithmes et l'IA doivent pouvoir être expliqués et compris par les personnes concernées. Cela inclut la capacité de vérifier, d'interpréter et de contester les décisions prises par les systèmes automatisés.

Bibliographie sélective sur la gouvernance des données

1. **Cadres de gouvernance des données :** documents sur les modèles, les principes et les cadres internationaux de gouvernance des données.

Institut ADA Lovelace. "Participatory data stewardship: A framework for involving people in the use of data" (*Intendance participative des données : cadre pour impliquer les citoyens dans l'utilisation des données*). Septembre 2021.

Barbero, Martina et Janet McLaren. « Partage efficace et éthique à grande échelle ». Partenariat mondial pour les données du développement durable. 27 mai 2024.

Dotation Carnegie pour la paix internationale. "Data Governance, Asian Alternatives: How India and Korea Are Creating New Models and Policies" (*Gouvernance des données, alternatives asiatiques : Comment l'Inde et la Corée créent de nouveaux modèles et de nouvelles politiques*). Édité par Evan A. Feigenbaum et Michael R. Nelson. Août 2022.

Centre pour l'innovation dans la gouvernance internationale. "The Role of Governance in Unleashing the Value of Data" (*Le rôle de la gouvernance dans la libération de la valeur des données*). Novembre 2024.

Digital Cooperation Organization (*Organisation de coopération numérique, OCN*). "DCO Privacy Principles" (*Principes de protection de la vie privée*). Mai 2025.

"Data Governance and Policy in Africa" (*Gouvernance et politique des données en Afrique*). Édité par Bitange Ndemo, Njuguna Ndung'u, Scholastica Odhiambo, Abebe Shimeles. Cham : Springer, 2023.

Initiative Datasphere. "Datasphere Governance Atlas 2022" (*Atlas 2022 de gouvernance de Datasphere*). Avril 2022.

Commission européenne. « Une stratégie européenne pour les données ». 19 février 2020.

Union européenne. « La législation sur les services numériques. » 27 octobre 2022.

Infocomm Media Development Authority. "Trusted Data Sharing Framework" (Cadre de partage de données fiables). 1er juillet 2024.

UIT. "Navigating Data Governance: A Guiding Tool for Regulators" (*Parcourir la gouvernance des données : outil d'orientation à l'intention des autorités de régulation*). 31 octobre 2024.

Lights on Data. "The complete guide to data governance roles and responsibilities" (*Guide complet des rôles et responsabilités en matière de gouvernance des données*).

MacFeely, Steve, Angela Me, Friederike Schueuer, Joseph M Costanzo, David Passarelli, Malarvizhi Veerappan et Stefaan Verhulst. 2025. "Towards a Set of Universal Data Principles" (*Vers un ensemble de principes universels de données*). *Journal statistique de l'AISO : Journal de l'Association internationale pour les statistiques officielles* 41 (1) : 150-55.

Marcucci, Sara, Natalia González Alarcón, Stefaan G. Verhulst et Elena Wüllhorst. 2023. "Informing the Global Data Future: Benchmarking Data Governance Frameworks" (*Informer l'avenir des données mondiales : Analyse comparative des cadres de gouvernance des données)*. Data & Policy 5 (Janvier).

Gouvernement de la Nouvelle-Zélande. "Co-designing Māori data governance" (Co-conception de la gouvernance des données Māori). 2021.

OCDE. « Recommandation du Conseil sur l'amélioration de l'accès aux données et de leur partage » Octobre 2021.

OCDE. Boîte à outils « Going Digital »

Open Data Institute (ODI). "Mapping the Wide World of Data Sharing" (*Cartographie du vaste monde du partage des données*). 4 juillet 2019.

Partenariat mondial sur l'intelligence artificielle (PMIA). "Data Governance Working Group: A Framework Paper for GPAI's Work on Data Governance 2.0" (*Groupe de travail sur la gouvernance des données : Document-cadre pour les travaux du PMIA sur la gouvernance des données 2.0*). Novembre 2022.

The GovLab. "Facilitating Data Flows through Data Collaboratives" (*Faciliter les flux de données grâce aux collaborations de données*). 19 octobre 2023.

PNUD - Stratégie numérique 2022-2025.

UNESCO « Le partage des données pour promouvoir l'information comme bien public », 2023.

UNICEF. « Cadre de qualité des données de l'UNICEF ». Avril 2022.

Conseil des chefs de secrétariat (CCS) de l'ONU. "Proposed Normative Foundations for International Data Governance: Goals and Principles" (*Fondements normatifs proposés pour la gouvernance internationale des données : Objectifs et principes*). Novembre 2024.

Nations Unies. « Pacte numérique mondial ». 22 septembre 2024.

Verhulst, Stefaan G. "Operationalizing Digital Self-Determination" (*Opérationnalisation de l'autodétermination numérique*). Data G Policy 5 (2023) : e11. 24 avril 2023.

Forum économique mondial. "Advancing Digital Agency: The Power of Data Intermediaries" (*Faire progresser les agences numériques : le pouvoir des intermédiaires de données*). Février 2022.

2. **Cadres sectoriels : modèles de gouvernance** des données adaptés à des secteurs spécifiques, comme la santé, l'humanitaire et les données statistiques.

Data Responsibility Guidelines (*Lignes directrices sur la responsabilité des données*). OCHA, Centre de données humanitaires, 2025.

Dodgson, Kate, et al. A Framework for The Ethical Use of Advanced Data Science Methods in the Humanitarian Sector (*Cadre pour l'utilisation éthique des méthodes avancées de science des données dans le secteur humanitaire*). Groupe de science des données et d'éthique (DSEG), Organisation internationale pour les migrations (OIM), 1er avr. 2020.

GSMA. GSMA Guidelines on the Protection of Privacy in the Use of Mobile Phone Data for Responding to the Ebola Outbreak (*Lignes directrices de la GSMA sur la protection de la vie privée dans l'utilisation des données de téléphonie mobile pour répondre à l'épidémie d'Ebola*). GSMA, Octobre 2014.

Hastie, Rachel, et Amy O'Donnell. Kit de formation sur la gestion responsable de données. Édité par Sally Bolton, Oxfam, 2017.

Organisation internationale pour les migrations. "DTM G Partners Toolkit: Enhancing Responsible Data Sharing" (*Boîte à outils de DTM & Partners : améliorer le partage responsable des données*). 16 novembre 2018.

Marelli, Massimo, et al, éditeurs. Handbook on Data Protection in Humanitarian Action (*Manuel sur la protection des données dans l'action humanitaire*). Troisième édition, Cambridge University Press, 2024.

OCDE. « Recommandation du Conseil sur la gouvernance des données de santé ». Adoptée le 13 janvier 2017.

Organisation panaméricaine de la santé (OPS). "National Data Governance Framework: Information Systems for Health" (*Cadre national de gouvernance des données : Systèmes d'information pour la santé*). 25 octobre 2024.

Association Sphere. The Sphere Handbook: Humanitarian Charter and Minimum Standards in Humanitarian Response (*Le manuel de Sphere : Charte humanitaire et normes minimales dans la réponse humanitaire*). Quatrième édition, Association Sphere, 2018.

Transformer la santé. "Health Data Governance Principles" (*Principes de gouvernance des données de santé*). Mai 2021.

Verhulst, Stefaan. 2024. "The Need for Climate Data Stewardship: 10 Tensions and Reflections Regarding Climate Data Governance" (*Le besoin d'une intendance responsable des données climatiques : 10 tensions et réflexions concernant la gouvernance des données climatiques*).

Division de statistique de l'ONU (UNSTATS). « Principes fondamentaux de l'ONU de la statistique officielle, Lignes directrices de mise en œuvre ». Janvier 2015.

CPI. « Directives opérationnelles sur la responsabilité en matière de données dans l'action humanitaire. » Avril 2023.

OMS. « Principes de l'OMS en matière de données ». Juin 2020.

3. **Protection des données, Vie privée et Sécurité :** sources couvrant la confidentialité des données, les lois de protection et la sécurité des données.

Union africaine « Convention sur la cybersécurité et la protection des données à caractère personnel » Juin 2014.

De France, James, et Luciia Laplante. Politique de la FICR en matière de protection des données à caractère personnel. Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge, 25 mars 2019.

Union européenne. « Règlement général sur la protection des données (RGPD). » 27 avril 2016.

Union européenne. « Outils de protection des données et de vie privée ».

Chapitre 5 du Bureau du commissaire à l'information : "Privacy-Enhancing Technologies (PETs) in Anonymisation and Pseudonymisation" (*Technologies de renforcement de la confidentialité dans l'anonymisation et la pseudonymisation*), Septembre 2022.

Manuel de protection des données de l'OIM. Organisation internationale pour les migrations, 2010.

UIT - Union internationale des télécommunications. « Rapport technique D4.1 - Cadre pour la sécurité, la protection de la vie privée, le risque et la gouvernance dans le traitement et la gestion des données ». Décembre 2019.

OCDE. « Technologies émergentes renforçant la protection de la vie privée : Approches réglementaires et politiques actuelles ». Documents de l'OCDE sur l'économie numérique, mars 2023.

OCDE. « Recommandation du Conseil concernant les Lignes directrices régissant la protection de la vie privée et les flux transfrontaliers de données de caractère personnel ». Adoptée le 23 septembre 1980 ; révisée le 11 juillet 2013.

Bureau du Commissaire australien à l'information. "Undertaking a Privacy Impact Assessment" (*Entreprendre une évaluation de l'impact sur la vie privée*). Mai 2020.

Responsible Program Data Policy (*Politique de données d'un programme responsable*). Oxfam International, 17 février 2025.

The Engine Room. "Hand-book of the Modern Development Specialist: Being, a Complete, Illustrated Guide to Responsible Data Usage, Manners, and General Deportment" (*Manuel du spécialiste du développement moderne : guide complet et illustré sur l'utilisation responsable des données, les bonnes pratiques et le comportement général*).

Équipe de travail sur les Technologies de renforcement de la confidentialité du Comité d'experts en mégadonnées et sciences des données en statistique officielle. "UN Guide on Privacy-Enhancing Technologies for Official Statistics" (*Guide de l'ONU sur les Technologies de renforcement de la confidentialité pour les statistiques officielles*). 2023.

Verhulst, Stefaan. 2025. "Data Stewardship Decoded: Mapping Its Diverse Manifestations and Emerging Relevance at a Time of AI" (*L'intendance des données décodée : Cartographie de ses diverses manifestations et sa pertinence émergente à l'heure de l'IA*). Document de travail SSRN.

UNICEF. Politique de l'UNICEF de protection des données personnelles. 15 juillet 2020.

Groupe des Nations Unies pour le développement durable (GNUD). « Confidentialité, éthique et protection des données : Note d'orientation du GNUD concernant les mégadonnées à l'appui de la réalisation du Programme 2030. » Novembre 2017.

Initiative Global Pulse de l'ONU. « Évaluation des risques, des préjudices et des avantages. » Juin 2018.

UNESCO « Liberté d'information : étude juridique comparative ». 2008.

UNESCO « Droits de l'homme et chiffrement ». 2016.

UNESCO « Lignes directrices à l'intention des acteurs judiciaires sur la protection de la vie privée et des données ». 2022.

UNHCR. « Politique relative à la protection des données des personnes relevant de la compétence du HCR ». Mai 2015.

HCR, Politique générale sur la protection et la confidentialité des données à caractère personnel, 2022.

Haut-Commissariat des Nations Unies pour les réfugiés (HCR). Politique relative à la protection des données des personnes relevant de la compétence du HCR, Mai 2015.

4. **Pouvoir des données, Justice, Éthique et Droits numériques :** travaux explorant l'éthique, la justice et les implications sociales des données, notamment la dynamique du pouvoir et les droits individuels.

Institut Ada Lovelace. "Rethinking Data and Rebalancing Digital Power" (*Repenser les données et rééquilibrer le pouvoir numérique*). Novembre 2022.

Access Now et le Gouvernement de Catalogne. « Déclaration de Genève sur la surveillance ciblée et les droits de l'homme ». 29 septembre 2022.

Centre pour les droits de l'homme et la justice mondiale (CHRGJ). "Paving a Digital Road to Hell? A Primer on the Role of the World Bank and Global Networks in Promoting Digital ID" (*Une route numérique vers l'enfer ? Introduction au rôle de la Banque mondiale et des réseaux mondiaux dans la promotion de l'identification numérique*). Juin 2022.

D'Ignazio, Catherine, et Lauren F. Klein. Data Feminism (*Féminisme des données*). The MIT Press, Mars 2020 ; réimpression en Octobre 2023.

Commission d'éthique des données (Allemagne). "Opinion of the Data Ethics Commission" (*Avis de la Commission de déontologie des données*). Octobre 2019.

Partenariat mondial sur l'intelligence artificielle (PMIA). "Data Justice : A Primer on Data and Economic Justice" (*Justice des données : abécédaire des données et de la justice économique*). Novembre 2022.

Heeks, Richard, et Jaco Renken. "Data Justice for Development: What Would It Mean?" (*La justice des données pour le développement : Qu'est-ce que cela signifierait ?)*. Information Development, vol. 34, no. 1, 2018, p. 90–102.

Organisation internationale pour les migrations (OIM). « Manuel de protection des données de l'OIM ». 2010.

Open Data Institute (ODI). "The Data Ethics Canvas" (*Canevas de l'éthique des données*). Dernière mise à jour en juin 2021.

Taylor, Linnet. "What is Data Justice? The Case for Connecting Digital Rights and Freedoms Globally" (*Qu'est-ce que la justice des données ? Arguments en faveur d'une connexion mondiale des droits numériques et des libertés*). Big Data G Society, vol. 4, no. 2, Juillet-Décembre 2017.

Verhulst, Stefaan G. "Reimagining data responsibility: 10 new approaches toward a culture of trust in re-using data to address critical public needs" (*Réimaginer la responsabilité en matière de données : 10 nouvelles approches pour une culture de la confiance dans la réutilisation des données pour répondre aux besoins publics essentiels*). Data G Policy 3 (2021).

Zahuranec, Andrew J., Hannah Chafetz et Stefaan Verhulst. 2023. "Data Sharing Agreements: Moving from Idea to Practice" (*Accords de partage des données : Passer de l'idée à la pratique*). Brooklyn, New York.

5. **Droits numériques de l'enfant :** documents et lignes directrices traitant spécifiquement des droits de l'enfant dans les espaces numériques.

Conseil de l'Europe. « Lignes directrices relatives au respect, à la protection et à la réalisation des droits de l'enfant dans l'environnement numérique. » Adoptées le 4 juillet 2018.

Assemblée mondiale pour la protection de la vie privée. « Résolution sur les droits numériques des enfants ». 25 octobre 2021.

« Principes ». Responsible Data for Children.

UNICEF et The GovLab. "Responsible Data for Children Synthesis report" (*Rapport de synthèse des données responsables concernant les enfants*).

Bureau de recherche de l'UNICEF - Innocenti. "The Case for Better Governance of Children's Data: A Manifesto" (*Arguments en faveur d'une meilleure gouvernance des données relatives aux enfants : manifeste*). 2020. Consulté le 25 octobre 2024.

Nations Unies. « Les droits de l'enfant dans l'environnement numérique ». 19 décembre 2023.

Nations Unies. « Note d'orientation du Secrétaire général sur la transversalisation des droits de l'enfant ». Juillet 2023.

Nations Unies. « Promotion et protection des droits de l'enfant ». Décembre 2023.

Fondation 5 Rights. « Une enfance ébranlée : Le coût de la conception persuasive ». 11 avril 2023.

Fondation 5 Rights. "Risk by Design" (Le risque dès la conception).

Comité des droits de l'enfant de l'ONU. Voir également l'Observation générale no 25 (2021) sur les droits de l'enfant en relation avec l'environnement numérique. 2021.

6. **Données pour le développement :** approches fondées sur les données pour soutenir le développement durable.

AI in Global Development Playbook (*Guide sur l'IA dans le développement mondial*). USAID, DOS, 2024.

Union africaine. "The Digital Transformation Strategy for Africa (2020–2030)" (*Stratégie de transformation numérique pour l'Afrique (2020-2030)*). 2020.

OCDE. « Mesurer la valeur des données et des flux de données ».

Paul, Amy, et al. Reflecting the Past, Shaping the Future: Making AI Work for International Development (*Réfléchir au passé, façonner l'avenir : Mettre l'IA au service du développement international*). Centre pour le développement numérique de l'USAID, 2018.

Paul, Amy, et al. Managing Machine Learning Projects in International Development (*Gestion de projets d'apprentissage automatique dans le domaine du développement international*). USAID, DAI, Vital Wave, 2021.

Verhulst, Stefaan, Laura Sandor, Elena Murray et Peter Addo. 2024. « Réutilisation éthique des données dans les pays en développement : une licence sociale à travers l'engagement public ».

PNUD. "Accelerating The SDGs Through Digital Public Infrastructure: A Compendium of the Potential of Digital Public Infrastructure" (*Accélérer la réalisation des ODD grâce aux infrastructures publiques numériques : recueil sur le potentiel des infrastructures publiques numériques*). Publié le 21 août 2023.

PNUD. « Principes de données pour le PNUD ». 2020.

Verhulst, Stefaan "Social License for Data: Going Beyond Consent to Protect Vulnerable Populations" (*Licence sociale pour les données : Au-delà du consentement pour protéger les populations vulnérables*).

Banque mondiale. « Rapport sur le développement dans le monde 2021 : Des données au service d'une vie meilleure ». Publié en mars 2021.

Verhulst, Stefaan G. "Reusing data responsibly to achieve development goals" (*Réutiliser les données de manière responsable pour atteindre les objectifs de développement*). Dans le *Rapport de l'OCDE de coopération pour le développement 2021 : pour une transformation numérique juste*, 289-297. Paris : Éditions de l'OCDE, 2021.

Banque mondiale. « Guide du praticien ID4D. »

7. Flux transfrontaliers de données et Économie numérique

Brookings. "Data portability and interoperability: A primer on two policy tools for regulation of digitized industries" (*Portabilité et interopérabilité des données : introduction à deux outils politiques pour la réglementation des industries numérisées*). Mai 2023.

Catapult. "City Data Sharing Toolkit" (*Boîte à outils pour le partage des données des villes*). Septembre 2019.

Initiative Datasphere. "Sandboxes for Data: Creating Spaces for Agile Solutions Across Borders" (*Bacs à sable de données : Créer des espaces pour des solutions agiles au-delà des frontières*). Mai 2022.

Organisation de coopération numérique. "Enabling Cross-Border Data Flows amongst the Digital Cooperation Organization Member States" (*Faciliter les flux transfrontaliers de données entre les États membres de l'Organisation de coopération numérique*). Octobre 2023.

Digital Public Goods Charter. « Charte des biens publics numériques ». 1er juin 2022.

Gebru, Timnit et al. "Datasheets for Datasets" (*Fiches de données pour les ensembles de données*). 1er décembre 2021.

The GovLab "Facilitating Data Flows through Data Collaboratives" (*Faciliter les flux de données grâce à la collaboration en matière de données*).

Global Data Alliance. "Cross-Border Data Policy Principles" (*Principes relatifs à la politique en matière de données transfrontalières*). 2 mars 2021.

Partenariat mondial pour les données du développement durable. « Livre de recettes pour un partage de données ». Mai 2024.

NESTA, gouvernement britannique "Data Sharing Toolkit" (*Boîte à outils pour le partage des données*).

Open Data Institute. "Mapping the Wide World of Data Sharing" (*Cartographier le vaste monde du partage de données*).

Verhulst, Stefaan. 2023. "Policy Paper 10 (Document d'orientation n° 10) – Data Collaboratives : Enabling a Healthy Data Economy through Partnerships (*Collaborations en matière de données : Favoriser une économie des données saine grâce à des partenariats*). Dans The Digital Revolution and the New Social Contract (*La révolution numérique et le nouveau contrat social*). IE University.

CNUCED. « Rapports sur l'économie numérique ».

8. Gouvernance et Réglementation de l'IA : réglementations, principes et lignes directrices éthiques pour le développement et l'application de l'IA.

Adams, Rachel, et al. Indice mondial de l'IA responsable. Global Center on AI Governance (*Centre mondial sur la gouvernance de l'IA*), 2024.

Explorateur de la loi sur l'IA. https://artificialintelligenceact.eu/fr/ai-act-explorer/

Artificial Intelligence Plan: Charting the Course for Responsible AI in USAID Programming (*Plan pour l'intelligence artificielle : DéfInir les orientations pour une IA responsable dans les programmes de l'USAID*). USAID, 2022.

Commission européenne. (2024). *Règlement (UE) 2024/1689 du Parlement européen et du Conseil du 13 juin 2024 établissant des règles harmonisées concernant l'intelligence artificielle*. https://eur-lex.europa.eu/eli/reg/2024/1689/ oj/eng

The GovLab. "Resources on Data Sharing Agreements" (*Ressources sur les accords de partage des données*).

Partenariat mondial sur l'intelligence artificielle (PMIA). "Global Index on Responsible AI" (*Indice mondial sur l'IA responsable*).

Hugging Face. "Building Better AI: The Importance of Data Quality" (*Construire une meilleure IA : l'importance de la qualité des données*).

OCDE. « Principes de l'OCDE sur l'IA ». Adoptés le 22 mai 2019.

OCDE. "Regulatory Sandboxes in Artificial Intelligence" (*Bacs à sable réglementaires dans le domaine de l'intelligence artificielle*). Mai 2022.

"The IEEE Global Initiative 2.0 on Ethics of Autonomous and Intelligent Systems" (*Initiative mondiale 2.0 de l'IEEE sur l'éthique des systèmes autonomes et intelligents*). Association de normalisation de l'IEEE.

Conseil des chefs de secrétariat (CCS) de l'ONU pour la coordination : Principes relatifs à l'utilisation éthique de l'IA par le système des Nations Unies.

UNESCO « Recommandation de l'UNESCO sur l'éthique de l'IA ». 23 novembre 2021.

9. Interopérabilité

Commission européenne. "ISA² - Interoperability solutions for public administrations, businesses and citizens" (*ISA² - Solutions d'interopérabilité pour les administrations publiques, les entreprises et les citoyens*). 25 avril 2018.

Union européenne. "Access to Base Registries: Good Practices on Building Successful Interconnections of Base Registries" (*Accès aux registres de base : Bonnes pratiques pour établir des interconnexions efficaces entre les registres de base*). 2016.

Union européenne : Good Practices on Building Successful Interconnections of Base Registries (*Bonnes pratiques pour établir des interconnexions efficaces entre les registres de base*).

HCR. « Passerelle d'interopérabilité PRIMES (PING) »

10. Gouvernance des données pour les ressources de l'entreprise, axée sur l'application des principes et stratégies de gouvernance des données dans le contexte de l'entreprise.

Aubert, Benoit, et al. Data Governance: Governing Data for Sustainable Business (*Gouvernance des données : Gérer les données pour une activité durable*). Édité par Alison Holt, BCS, The Chartered Institute for IT, 2020.

Bhansali, Neera. Data Governance: Creating Value from Information Assets (*Gouvernance des données : Créer de la valeur à partir des actifs informationnels*). CRC Press, Taylor & Francis Group, 2014.

Deloitte. "Data Valuation: Understanding the Value of Your Data Assets" (*Évaluation des données : Comprendre la valeur de vos actifs de données*). 2020.

Eryurek, Evren, et al. Data Governance: The Definitive Guide: People, Processes, and Tools to Operationalize Data Trustworthiness (*Gouvernance des données : Le guide définitif : Personnes, processus et outils pour opérationnaliser la fiabilité des données*). 1re édition, O'Reilly Media, 2021.

Infocomm Media Development Authority, Singapour (IMDA) et Commission de protection des données personnelles (PDPC). "Guide to Data Valuation for Data Sharing" (*Guide d'évaluation des données en vue de leur partage*). 2019.

International, DAMA. The DAMA Guide to the Data Management Body of Knowledge (*Le guide DAMA sur le corpus de connaissances en matière de gestion des données*) - Édition imprimée. Première édition, Technics Publications, LLC, 2010.

Ladley, John. Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program (*Gouvernance des données : Comment concevoir, déployer et maintenir un programme efficace de gouver-nance des données*). Newnes, 2012.

Madsen, Laura B. Disrupting Data Governance: A Call to Action (*Perturber la gouvernance des données : appel à l'action*). Technics Publications, 2019.

Maydanchik, Arkady. Data Quality Assessment (*Évaluation de la qualité des données*). Technics Publications, LLC, 2012.

OCDE. "Measuring the Value of Data and Data Flows" (*Mesurer la valeur des données et des flux de données*). Documents de l'OCDE sur l'économie numérique. Paris.

Éditions de l'OCDE, 14 décembre 2022.

Reichental, Jonathan. Data Governance For Dummies (*La gouvernance des données pour les nuls*). 1re édition, Pour les Nuls, 2022.

Annexes : Listes de contrôle

POURQUOI : Déterminer les finalités des données et de leur gouvernance

Veuillez cocher toutes les finalités qui s'appliquent à votre initiative de gouvernance des données :

Maximiser l'utilité et la valeur des données / Minimiser les inconvénients et les coûts	
Exploiter les données pour obtenir de nouvelles informations et améliorer la prise de décisions.	☐
Créer une valeur concrète pour l'organisation et/ou la société grâce à l'utilisation des données.	☐
Réduire au minimum les inconvénients, les coûts et les conséquences négatives involontaires associés à la collecte et à l'utilisation des données.	☐
Favoriser l'innovation et le développement durable	
Stimuler l'innovation et l'esprit d'entreprise grâce à des approches fondées sur des données.	☐
Promouvoir et renforcer les opportunités économiques grâce aux données.	☐
Faire progresser les objectifs de développement durable (ODD) en utilisant des données pour le suivi et la mise en œuvre (p. ex. dans les domaines de la santé, de l'éducation et des infrastructures).	☐
Favoriser une culture axée sur les données et améliorer les compétences en matière de données au sein de l'organisation/société.	☐
Établir l'équité et l'autodétermination numérique	
Promouvoir un accès équitable aux données et veiller à ce que les bénéfices soient partagés équitablement.	☐
Veiller à ce que les groupes en situation de marginalisation et de vulnérabilité, notamment les enfants, soient activement inclus et protégés dans les écosystèmes de données.	☐
Respecter les principes d'autodétermination des données pour les individus et les communautés.	☐
Mettre en œuvre des cadres de gouvernance spécifiques pour protéger les données des populations vulnérables.	☐
Soutenir des objectifs politiques ou opérationnels spécifiques	
Améliorer la transparence, la responsabilisation et l'engagement des citoyens (p. ex. grâce à des initiatives de données ouvertes).	☐
Soutenir les priorités nationales en matière d'ODD, telles que la prise en compte de l'impact du changement climatique ou la promotion de l'éducation pour tous.	☐
Faciliter le partage transfrontalier de données et la collaboration internationale de manière sûre et efficace, tout en garantissant le respect de la législation.	☐
Mobiliser efficacement les données pour la préparation aux crises, la gestion des risques, la réponse et les efforts de rétablissement.	☐
Soutenir le développement et le déploiement responsables de l'intelligence artificielle (IA), en tenant compte des préjugés, de l'équité et des préoccupations éthiques.	☐
S'aligner sur les principes de l'infrastructure publique numérique (IPN) ou les mettre en œuvre afin de permettre un échange de données sécurisé, privé, interopérable et transparent.	☐

COMMENT : Déterminer les principes de gouvernance des données

Cette liste de contrôle permet de déterminer et d'évaluer la présence de principes dans trois catégories :

1. **Principes relatifs aux processus** - comment les décisions de gouvernance sont prises
2. **Principes relatifs aux décisions** - ce qui éclaire ces décisions
3. **Principes relatifs à la gestion des données** - comment les données sont gérées dans la pratique

Principes relatifs aux processus	
Ces principes garantissent que les activités de gouvernance sont menées de manière équitable, éthique et transparente.	
Transparence - Les processus de gouvernance sont ouverts et compréhensibles.	☐
Responsabilisation - Les acteurs sont responsables des résultats des processus de gouvernance.	☐
Approche centrée sur les personnes - Les besoins et les droits des personnes sont prioritaires.	☐
Équité - Traitement égal et impartial dans les procédures.	☐
Participation - Les parties prenantes sont impliquées de manière significative dans la gouvernance.	☐
Licéité - Toutes les actions sont conformes aux lois et règlements en vigueur.	☐
Inclusivité - Les groupes marginalisés et les personnes affectées par la *datafication* sont représentés.	☐
Principes relatifs aux décisions	
Ces principes guident la définition et la mise en œuvre des décisions en matière de gouvernance des données.	☐
Transparence - La logique des décisions est clairement communiquée.	☐
Proportionnalité - Les décisions sont adaptées à leur contexte et à leur impact.	☐
Finalité définie - Les décisions sont guidées par des objectifs clairs et spécifiques.	☐
Responsabilisation - Les décideurs doivent répondre de leurs choix.	☐
Approche centrée sur les personnes - Les besoins et les droits des personnes sont au premier plan.	☐
Équité - Les décisions sont justes et équitables.	☐
Protection contre les préjudices et non-discrimination - Les risques sont atténués et les préjugés sont évités.	☐
Participation - Des perspectives diverses sont intégrées dans la prise de décisions.	☐
Principes relatifs à la gestion des données	
Ces principes garantissent que les données sont gérées de manière responsable, sûre et conforme aux droits et aux attentes.	
Confidentialité et Sécurité - Les données sensibles sont protégées contre tout accès non autorisé.	☐

Proportionnalité - Les pratiques en matière de données correspondent à la finalité et au besoin.	☐
Accessibilité et Portabilité - Les données sont disponibles et portatives dans des conditions appropriées.	☐
Protection de la vie privée - Les données à caractère personnel sont sauvegardées conformément aux lois sur la protection de la vie privée.	☐
Licéité - Le traitement des données est conforme à toutes les normes juridiques applicables.	☐
Consentement éclairé - Les personnes concernées sont conscientes de l'utilisation des données et y consentent.	☐
Qualité des données et des métadonnées - Les données sont précises, fiables et bien documentées en vue de leur réutilisation.	☐
Interopérabilité et normalisation - Les données adhèrent à des normes partagées pour faciliter l'échange et l'intégration.	☐

QUI : Détermination des rôles et des responsabilités

Modèle RACI	Planification	Collecte	Traitement	Partage	Analyse	Utilisation
Responsable						
Garant						
Consulté						
Informé						

QUOI : Planification

Tâche	
La finalité et la valeur des données sont clairement définies et documentées	☐
Les parties prenantes et les communautés concernées ont été identifiées et cartographiées	☐
Le contexte juridique et politique, notamment les efforts antérieurs, ont été examinés	☐
Le modèle de gouvernance (rôles, responsabilités, prise de décisions) est conçu	☐
La portée, les objectifs et les limites du projet de données sont définis	☐
Les ressources financières, techniques et humaines sont évaluées et garanties	☐
Des accords de partage de données (MA, APD) et des modèles juridiques sont préparés	☐
Les besoins en matière d'interopérabilité et les vocabulaires partagés ont été examinés	☐
Des études de risques (p. ex. protection de la vie privée, sécurité, risques liés à l'IA) ont été effectuées	☐
Une stratégie d'engagement des parties prenantes et d'instauration d'un climat de confiance a été mise en place	☐
Un plan de communication et de transparence a été élaboré	☐
Des boucles de rétroaction et des mesures d'évaluation ont été définies	☐

QUOI : Collecte

Tâche	
La collecte des données est conforme aux finalités déclarées et aux principes de minimisation	☐
Les mécanismes de consentement sont établis et documentés (notamment le consentement dynamique)	☐
Les populations marginalisées ou mal desservies sont représentées	☐
Les principes de protection des données dès la conception sont appliqués aux systèmes de collecte	☐
Les formats de données sont normalisés et interopérables	☐
Les pratiques en matière de métadonnées et de documentation sont établies	☐
Les flux transfrontaliers de données et les exigences en matière d'emplacement ont été évalués	☐
Des outils ont été instaurés pour détecter et atténuer les biais lors de la collecte des données	☐
Des méthodes de chiffrement ou d'anonymisation sont appliquées à l'étape de collecte	☐
Les obligations légales (p. ex. la base légale, les questions juridictionnelles) sont examinées et respectées	☐
Des mécanismes de retour d'information ont été instaurés pour adapter les processus de collecte en cas de problème	☐

QUOI : Traitement

Tâche	
Les processus de nettoyage et de transformation des données sont documentés	☐
La qualité des données (leur exactitude, cohérence, exhaustivité) est validée	☐
Les données sont catégorisées et classées (p. ex. sensibles ou non)	☐
La provenance des données et l'historique des versions sont à jour	☐
Un chiffrement et des contrôles d'accès sont mis en œuvre	☐
Des protocoles d'accès internes ont été définis (p. ex. l'accès par paliers, les journaux d'audit)	☐
Des technologies de renforcement de la confidentialité sont envisagées ou appliquées	☐
Des systèmes de sauvegarde, d'archivage et de suppression efficaces et sécurisés ont été instaurés	☐
La compatibilité avec les normes d'interopérabilité est assurée	☐
Les activités de traitement font l'objet d'un contrôle de conformité juridique et éthique	☐
Des mécanismes pour éviter les liens involontaires entre les données ont été instaurés	☐

QUOI : Partage

Tâche	
La finalité et la portée du partage des données sont clairement définies et convenues	☐
Les accords juridiques (p. ex. les APD, les MA, les licences) ont été formalisés	☐
Les parties prenantes internes sont prises en compte et informées des attentes en matière de gouvernance	☐
Une stratégie de communication externe a été développée	☐
Les risques de réputation et les plans d'atténuation ont été identifiés	☐
Des protocoles d'interopérabilité (p. ex. un vocabulaire et des formats communs) ont été établis	☐
Des garanties techniques (p. ex. des IPA sécurisées, un accès fédéré) ont été instaurées	☐
Les risques pour la vie privée et la sécurité liés à la combinaison des données ont été anticipés	☐
La gouvernance des intermédiaires tiers ou des plateformes a été clarifiée	☐
La licence sociale et des considérations éthiques ont été examinées	☐
Des mécanismes de coordination et d'examen permanents ont été mis en place	☐

QUOI : Analyse

Tâche	
Le plan d'analyse correspond aux objectifs initiaux du projet	☐
Les méthodes analytiques, les algorithmes et les hypothèses sont documentés	☐
Des tests de biais ont été effectués sur les données et les modèles	☐
Des mécanismes d'interprétabilité et d'explicabilité ont été appliqués	☐
Les résultats ont été validés par rapport aux observations du monde réel ou à la vérité terrain	☐
Des mécanismes de contrôle humain ont été instaurés pour les décisions importantes	☐
Un examen éthique de l'analyse a été effectué (en particulier pour les systèmes d'intelligence artificielle)	☐
Les données synthétiques ou les méthodes de préservation de la vie privée sont utilisées en cas de besoin	☐
L'engagement des parties prenantes ou des approches participatives ont été envisagés	☐
Les modèles ont été réentraînés ou recalibrés en cas de changements significatifs du contexte	☐

QUOI : Utilisation

Tâche	
L'utilisation finale des données est conforme aux objectifs initiaux du projet et au consentement	☐
La conformité juridique, éthique et de protection de la vie privée a été vérifiée avant l'utilisation	☐
Les risques de conséquences involontaires (p. ex. la réidentification, les biais) ont été évalués	☐
Les résultats ont été examinés par des experts pour en vérifier la qualité et l'impartialité	☐
Une stratégie de communication et de messages claire a été établie	☐
Les actions basées sur les données ont été examinées pour évaluer leur efficacité et leur impact	☐
Des mécanismes de retour d'information ont été instaurés pour recueillir les réactions des utilisateurs/communautés	☐
Des politiques de conservation des données ou de suppression sécurisée ont été établies	☐
Les modèles ont été recyclés ou révisés sur la base de nouvelles données ou d'analyses d'impact	☐
De la documentation sur les résultats, les apprentissages et les occasions manquées a été créée	☐

BROADBAND COMMISSION
FOR SUSTAINABLE DEVELOPMENT

Photo de couverture : metamorworks/Shutterstock.com