

# RAPPORT GLOBAL

Première édition



**Global Data Barometer**



**D4D.net**  
Data for Development

# Table des matières

|   |            |
|---|------------|
| <b>Remerciements</b> .....  | <b>3</b>   |
| <b>Introduction</b> .....   | <b>7</b>   |
| ■ À propos du baromètre.....  | <b>9</b>   |
| ■ Données pour le bien public.....  | <b>12</b>  |
| <b>Méthodologie</b> .....   | <b>16</b>  |
| <b>Piliers</b> .....  | <b>24</b>  |
| ■ Gouvernance.....  | <b>25</b>  |
| ■ Capacités.....  | <b>37</b>  |
| ■ Disponibilité.....  | <b>50</b>  |
| ■ Utilisation et impact.....  | <b>67</b>  |
| <b>Analyse régionale</b> .....  | <b>80</b>  |
| ■ Afrique.....  | <b>81</b>  |
| ■ Europe de l'Est et Asie centrale.....   | <b>86</b>  |
| ■ L'Union européenne, l'Amérique du Nord, Israël,<br>l'Australie, et la Nouvelle-Zélande..... | <b>91</b>  |
| ■ Amérique latine et Caraïbes.....  | <b>96</b>  |
| ■ Moyen-Orient et Afrique du Nord.....  | <b>102</b> |
| ■ Asie du Sud et de l'Est.....  | <b>105</b> |
| <b>Conclusions et travaux futurs</b> .....  | <b>111</b> |
| <b>Annexe: méthodologie</b> .....   | <b>115</b> |

## Remerciements

**Référence suggérée** : Baromètre mondial des données (2022). Rapport de première édition – Baromètre mondial des données. ILDA. DOI : <https://doi.org/10.5281/zenodo.6488349>

Ce rapport et l'ensemble du projet du Baromètre mondial des données ont été possibles grâce à la participation d'un vaste réseau de collaborateurs. Avec les remerciements suivants (par ordre alphabétique) :

**Auteurs du rapport** : Tim Davies & Silvana Fumega avec les contributions de Jonathan Gray

**Direction du projet** : Silvana Fumega

**Recherche et coordination des données** : Nicolas Grossman

**Recherche et renforcement de capacités** : David Zamora

**Conception et calcul d'indice**: Nicolas Grossman et David Zamora

**Processus de données et soutien analytique**: Ana Florido

**Communication** : Fiona Chawana avec le soutien-clé de Catrina Cortes et Violeta Belvet

**Début du développement du projet** : Katie Clancy, Veronica Cretu, Tim Davies, Amy Johnson, Oscar Montiel, Fernando Perini, Stephen Walker

**Développement d'outils d'enquête** : David Raznick (Open Data Services Co-operative)

**Outils visuels du site** : Socio Publico

**Hôte administratif** : ILDA

**Appui financier et en nature** : CRDI, Global Affairs Canada, et William and Flora Hewlett Foundation (via Open Data for Development), the Data for Development Global Research Hub (D4D.net) et FNUAP, partenaires thématiques et régionaux (voir ci-dessous)



Comité consultatif sur la recherche:

- Rachel Adams - Human Sciences Research Council, Afrique du Sud
- Ana Brandusescu - Université McGill, Canada
- Michael Canares - Step Up Consulting, Philippines
- Craig Hammer - Banque mondiale
- Arturo Muenté Kunigami - Banque Interaméricaine de Développement
- Leonida Mutuku - LDRI, Kenya
- Suzana Russell - University of the West Indies, Jamaïque
- Fabrizio Scrollini - ILDA, Uruguay
- Stefaan Verhulst - The GovLab, États-Unis d'Amérique
- Aimee Whitcroft - Experte indépendante, Nouvelle-Zélande

**Partenaires thématiques :** The Global Initiative for Fiscal Transparency, Land Portal, The Open Contracting Partnership, The Open Government Partnership, Open Ownership et Transparency International.



Partenaires régionaux:

- Access Info Europe
- Université de Birzeit
- ILDA
- Caribbean Open Institute
- Open Data China
- Open Data Kosovo
- InfoCulture
- Internet Society (chapitre de Hong Kong)
- Institut pour le développement de la liberté de l'information
- Institut de recherche pour le développement local
- Data for Development (D4D) Asia Hub
- La Communauté du Pacifique

Chercheurs locaux :

Pour Access Info Europe : Michal Škop, Niels Erik Kaaber Rasmussen, Open Knowledge Estonia, TIEKE Information Society Development Centre, Caroline Murgue, Elsa Perreau, Denis Parfenov, Openpolis, Liene Gatere, Rugile Trumpyte, Open State Foundation, Gustavo Magalhaes, Asedie, Elenor Weijmar, Javier Ruiz, Stephan Anguelov, Gong, Erasmia Tsipou & Sophocles P. Geroules, Christoph Schwaiger, Transparency International Romania, Lucia Cizmaziova, Rebecca Williams, Suk Kyoung Kim, Guy Zomer, Elsa Foucraut (chapitre TI).

Pour le Caribbean Open Institute : Michael Yee Shui, Fitzgerald Yaw, Victor Gonzalez, Laurajan Obermuller, Michelle McLeod, Louise Mathurin-Serieux, Shurland George

Pour le Centre de formation continue de l'Université de Birzeit : Hatem Ben Yacoub, Fatma Mosaad, Issa Al Mahasni, Raed Al Sahrif, Ahmad Qadi

Pour l'Open Development Initiative (East-West Management Institute) : Nikesh Balami, Vino Lucero, Ramathi Bandaranayake, Tong Khanh Linh

Pour l'IDFI Géorgie : Arpine Mazhinyan, Igbal Safarov, Mikhail Doroshevich, David Eristavi, Sholpan Aitenova, Gulnura Toralieva, Elberel Tumenjargal, Livia Turcanu, Sandro Rochikashvili, Rustam Muhamedov, Nadiia Babynska, Hurriyat Khudoykulova

Pour ILDA : Virginia Brussa, Milenka Villegas Taguasi, Larissa Galdino de Magalhães Santos, Carlos David Carrasco Muro, Juan Pablo Marín Díaz, Jorge Umaña Cubillo, Eduardo Bejar, Iris Bertila Palma Recinos, Julio Roberto Herrera Toledo, Daniel Emilio Rodriguez Rivera, Aura Eréndira Martínez Oriol, David Riveros García, Aída Martínez Mórtola, Ana Isabel Fiafilio Rodriguez, Eliana Álvarez, Pablo Secchi (chapitre TI)

Pour InfoCulture Russie : Olga Parkhimovich

Pour le chapitre de l'ISOC Hong Kong : Benjamin Zhou

Pour le Local Development Research Institute (LDRI) : Tsandzana Dércio, Kossi Amessinou, Oarabile Mudongo, Abzeta Kouloum Ouedraogo, Poncelet O. Ileleji, Donatien Abel Gbessi Gbala, Mobile Web Ghana, Local Development Research Institute, Lamii Kpargoi, Logos Open Culture, Hatem Ben Yacoub, Tsandzana Dércio, Nashilongo Gervasius, Yusuf Suleiman, Richard Ndicunguye, Open Data pour Elles (Open Data for Her Network), Abdulai Kallon, Mbongeni Hlabano, Emmanuel Vitus, Asma Cherifi

Pour Open Data China : Feng Gao

Pour Open Data Kosovo : Dafina Olluri

Pour la Communauté du Pacifique : Ida Carapelle

Données supplémentaires : Mailén García, Larissa Magalhães, Aura Martinez Oriol et Esbeidy Torres Hondal.

## Avertissement

Les données utilisées pour produire ce rapport ont été produites avec le soutien des organisations partenaires. Toutefois, les données et l'analyse ne représentent pas nécessairement les vues, les positions ou les opinions de ces différentes organisations, et toute erreur ou omission relève de la seule responsabilité du projet de Baromètre mondial de données. Plus particulièrement, les opinions exprimées ici ne représentent pas nécessairement celles du CRDI ou de son conseil des gouverneurs. Les données utilisées dans ce rapport correspondent à la période mai 2019- mai 2021.

# Introduction

Les données sont une source de pouvoir. Elles peuvent être exploitées à des fins privées, et utilisées pour limiter la liberté, ou peuvent être déployées pour le bien public : en tant que ressource pour relever les défis sociaux, permettre la collaboration, stimuler l'innovation et améliorer la responsabilité.

Au cours de la dernière décennie, les données se sont hissées au sommet des programmes politiques nationaux et mondiaux : alors que les nations cherchent à développer leurs économies, celles-ci utilisent les données pour relever les défis sociaux et répondre aux préoccupations des citoyens quant à l'utilisation et l'abus de ces données. Les progrès vers une gouvernance efficace des données et vers la réalisation de la valeur publique des données demeurent très inégaux d'un pays, d'une région et d'un secteur à l'autre. Par exemple, alors que les lois sur la protection des données sont maintenant répandues à travers le monde entier, beaucoup manquent de mécanismes de recours clés permettant aux individus et aux communautés d'exercer efficacement leurs droits en matière de données, et peu abordent de manière globale les problèmes émergents liés aux données de localisation ou à la prise de décision algorithmique. Dans des domaines critiques tels que l'action climatique, d'importantes lacunes dans les données peuvent entraver l'action locale visant à protéger les écosystèmes et à répondre à la vulnérabilité climatique. Et lorsque nous considérons au-delà de la simple disponibilité des ensembles de données, pour examiner si les données fournies répondent aux besoins des utilisateurs, nous trouvons des cas de données collectées et partagées, mais manquant de caractéristiques-clés ou d'assurances de qualité leur permettant d'alimenter correctement l'action civique, l'amélioration des services publics et le développement économique.

Cette première édition du Baromètre mondial de données fournit des données probantes et des informations sur le développement des écosystèmes de données nationales et mondiales : elle offre des comparaisons, des analyses et des exemples critiques qui peuvent aider à mener une action descendante et ascendante à réaliser les opportunités de la « révolution des données »<sup>[1,2]</sup>, tout en naviguant au sein de ses risques. En ce sens, le Baromètre vise non seulement à produire des évaluations des pays sur la base de l'état de leurs données, mais aussi à rassembler et à soutenir l'apprentissage collectif autour de ce qui fonctionne, et sur la manière d'avenir efficacement avec et autour des données. Dans ce rapport, nous résumons un certain nombre de constatations clés et soulignons les données recueillies par le Baromètre peuvent être utilisées dans d'autres travaux exploratoires.

## Points forts du baromètre

### **Il est possible de façonner les données pour le bien public. Mais le chemin est encore long.**

Presque tous les points de repère établis par nos indicateurs d'enquête ont été atteints *quelque part* dans le monde -montrant qu'il est à la fois possible de bien gérer les données et de les rendre disponibles et utilisés pour le bien public. Aucun pays n'a toutefois obtenu un score supérieur à 70 sur 100, et le score moyen du pays par rapport au baromètre de référence était 34.38 sur 100. Ceci montre que chaque pays a du travail à faire afin de mettre à jour les politiques, renforcer les capacités, partager les données et promouvoir l'utilisation des données afin de s'assurer que les données fonctionnent comme une ressource pour le développement durable, et souligner l'importance de continuer à mettre l'accent sur l'élaboration des politiques et des pratiques en matière de données pour assurer le bien public.

**Les programmes de données ouvertes sont actifs : mais ne se propagent pas.** En appliquant strictement des critères de données ouvertes, le pourcentage d'ensembles de données entièrement lisibles par machine, sous licence ouverte, librement disponibles et en vrac (10.63%) n'a guère

augmenté depuis la dernière mesure mondiale effectuée par le Baromètre des données ouvertes en 2016. Bien que de nouvelles initiatives nationales en matière de données ouvertes aient été lancées depuis 2016, d'autres initiatives ont disparu. Cependant, là où les initiatives ont été soutenues, elles sont souvent mieux réalisées en ressources et plus intégrées que de par le passé, et les principes de données ouvertes sont également intégrés dans un certain nombre d'initiatives sectorielles.

**Les lacunes en matière de capacité continuent de freiner la production de valeur à partir des données.** Alors que les fractures numériques en termes d'accès peuvent diminuer, amenant plus de personnes dans une société où les données sont fiables, les lacunes en termes de capacité du gouvernement, du secteur privé et de la société civile à créer et à utiliser des données pour le bien public restent importantes. L'offre limitée de formation et de renforcement des capacités au sein du gouvernement risque de compromettre la production et la gouvernance de données de haute qualité, et il est urgent de passer d'une formation ad hoc et à petite échelle, à un renforcement de la maîtrise des données à l'échelle des sociétés.

**Des cadres bien rédigés fournissent de meilleures données.** Nous avons examiné 7 politiques sectorielles différentes, et examiné à quel point les règles de gouvernance exigent que les données soient collectées et partagées sous des formes structurées. L'examen de la mise en œuvre entre les règles et la disponibilité des données souligne que lorsque les règles sont explicites sur la collecte et le partage des données, les données sont plus susceptibles d'être disponibles, ouvertes et de contenir les champs de données qui comptent.

**Les partenariats alimentent l'utilisation des données.** La collaboration entre la société civile traditionnelle et les technologies civiles, ou entre les journalistes et les fournisseurs d'applications du secteur privé stimule de nouvelles utilisations des données pour mettre en évidence la corruption, promouvoir l'intégrité publique, surveiller les questions environnementales et façonner les débats politiques.

En examinant les pratiques en matière de données dans des secteurs spécifiques, cette première édition du Baromètre mondial des données contribue aux travaux explorant, entre autres, les sujets suivants :

**Action pour le climat.** La réponse mondiale à la COVID-19 a démontré que de nouvelles infrastructures de données peuvent être construites rapidement, mais il existe des lacunes importantes et urgentes dans la disponibilité des ensembles de données sur les émissions, la biodiversité et la vulnérabilité climatique. Les données qui devraient être disponibles pour soutenir l'action locale visant à combattre les changements climatiques et à s'y adapter ne sont souvent disponibles que sous des formes agrégées et périmées. Nos données probantes ont le potentiel de soutenir l'action participative sur l'amélioration des écosystèmes de données climatiques, en aidant les communautés à identifier et comparer les bonnes pratiques et les lacunes.

**Intégrité politique.** Si les pays qui fournissent déjà des informations sur l'intégrité politique en ligne devaient passer des processus sur papier à la collecte de données structurées, ils pourraient débloquent de nouvelles approches en matière de responsabilité. Bien que le manque d'interopérabilité entre les ensembles de données sur l'intégrité politique demeure un problème clé dans de nombreux pays, nos données peuvent être utilisées pour explorer les exemples de points positifs et de meilleures pratiques où les systèmes de divulgation numérique sont à l'origine du changement. Nous fournissons également de nouvelles données de référence sur la prévalence des règles de divulgation des activités de lobbying et mettons en évidence le manque de données structurées sur le lobbying, ce qui permet de suivre les progrès vers une plus grande divulgation à l'avenir.

**Finances publiques et passation de marchés.** Un examen attentif des données disponibles révèle cependant également que, si des données sont de plus en plus disponibles du côté des « intrants » de l'investissement public (par exemple, allocations budgétaires, appels d'offres et adjudications de contrats, etc.), d'importants progrès restent à faire dans le suivi du côté des « extrants », en fournissant des données conjointes sur la mise en œuvre des contrats ou sur les incidences des dépenses, en

particulier sur les questions d'équité et de développement durable. Les niveaux relativement élevés de publication de données structurées et ouvertes détectés par notre enquête pour les données sur le budget et les dépenses du gouvernement, et pour les données sur les marchés publics, suggèrent une influence positive des campagnes mondiales et des initiatives de renforcement des capacités dans la promotion de la publication et de l'utilisation des données.

L'analyse régionale et les recommandations concernant chacun des quatre piliers du Baromètre mondial de données mettent en évidence les domaines d'action, adaptés aux différents contextes nationaux. Il s'agit notamment d'appels à :

**Renforcer le leadership et la stratégie** pour développer et intégrer les compétences, les institutions et les libertés nécessaires pour que les données soient régies et utilisées pour le bien public.

**Développer des cadres solides de partage de données**, y compris au niveau infranational, de sorte que les abus potentiels de données soient limités et que la réutilisation positive des données, qu'elles proviennent du secteur public, du secteur privé ou du secteur à but non lucratif, soit possible.

**Mettre davantage l'accent sur l'équité et l'inclusion**, en reconnaissant que la gouvernance, les capacités, la disponibilité et l'utilisation des données doivent prendre explicitement en compte les besoins des populations marginalisées.

**Accroître la transparence de l'utilisation des données gouvernementales** et sensibiliser le public au moment où les gouvernements recueillent, partagent ou utilisent des données. Ceci peut favoriser une pratique des données plus responsable et favoriser une plus grande collaboration entre les secteurs dans l'utilisation efficace des données.

Dans le chapitre sur l'analyse régionale, nous identifions les forces et les faiblesses potentielles de chaque pays, ainsi que les endroits où les pays peuvent offrir des exemples de points forts potentiels avec un potentiel d'apprentissage par les pairs. Ce chapitre fournit également des indications précieuses sur l'amélioration ou le recul des politiques et initiatives en matière de données ouvertes dans chaque pays depuis les dernières mesures comparables en 2016, 2017 ou 2020.

## À propos du baromètre

Le Baromètre \* mondial \* de données (ci-après, le Baromètre) s'appuie sur l'étude du Open Data Barometer (ODB), menée par la Web Foundation et ses partenaires entre 2013 et 2020 (l'édition mondiale finale a été achevée en 2016, et l'édition régionale finale, basée en Amérique latine et dans les Caraïbes, complétée par ILDA en 2020). Le Baromètre s'appuie sur les données primaires d'une enquête mondiale d'experts réalisée à la mi-2021, et en examinant les données probantes pour la période allant du 1er mai 2019 au 1er mai 2021. Ceci est combiné avec des données secondaires provenant de sources fiables pour générer une gamme de mesures. Le baromètre fournit:

- **Un score et non un classement** - les indicateurs primaires et les scores sont basés sur une échelle de 0 à 100, où 100 est conçu pour mesurer les « meilleures pratiques », définies par rapport aux normes ou cadres convenus au niveau international. Peu de pays obtiennent un score de 100 sur 100 sur quelque indicateur. Bien que les comparaisons entre les pays puissent être utilisées à la recherche des forces et des faiblesses relatives, la plus grande valeur de ce modèle est de montrer les domaines individuels à améliorer dans chaque pays.
- **Indicateurs et données probantes** - Les indicateurs primaires sont composés d'un certain nombre de sous-questions structurées et sont étayés par des données probantes qualitatives pouvant être utilisées pour mieux comprendre le contexte de chaque pays.

- **Évaluation réactive** - Les indicateurs ont été conçus pour prendre en compte les différents systèmes politiques et structures étatiques (par exemple, les systèmes fédéraux ou non), tout en réservant les scores les plus élevés aux cas où les règles de gouvernance, les capacités en matière de données ou les données disponibles couvrent l'ensemble de la population d'un pays.
- **Un réseau mondial** - la recherche a été menée à travers un réseau de pôles régionaux, principalement par des chercheurs nationaux. Les conclusions ont ensuite fait l'objet d'un recoupement avec un réseau mondial de partenaires experts thématiques. Ce modèle contribue au renforcement des capacités mondiales, en créant une communauté de chercheurs et de praticiens qui explorent les données pour le bien public.
- **Perspectives concrètes** - chaque indicateur du Baromètre a été conçu pour mesurer les caractéristiques de la gouvernance, de la capacité, de la disponibilité et de l'utilisation des données que les gouvernements ont le pouvoir d'aborder, et la société civile nationale et internationale d'appuyer et d'influencer.
- **Données ouvertes** - toutes les notes et données probantes recueillies par le Baromètre sont publiées avec ce rapport en tant que données ouvertes, ce qui permet de poursuivre à la fois recherches et analyses. Nous nous félicitons des travaux supplémentaires visant à remixer les résultats du baromètre.

Le Baromètre mondial des données va au-delà des conventions juridiques et techniques des données ouvertes pour explorer des paysages plus larges de disponibilité, de gouvernance, d'utilisation et d'impact des données, dans le but de placer les problèmes et les communautés au centre. Plutôt que de se concentrer uniquement sur la manière dont les données sont mises à disposition par les gouvernements, il explore aussi la vie sociale et les paramètres des données et la manière dont celles-ci ont été mises en œuvre en relation avec les problèmes – y compris via des exemples de « points forts », ainsi que les groupes de citoyens, de journalistes et de militants qui produisent et utilisent des données. En ce sens, le Baromètre vise non seulement à produire des évaluations des pays sur la base de l'état de leurs données, mais aussi à rassembler et à soutenir l'apprentissage collectif autour de ce qui fonctionne, et sur la manière d'avenir efficacement avec et autour des données.

Au final, nous espérons que cette première édition du Baromètre mondial des données contribuera non seulement à améliorer la gouvernance, la disponibilité et l'utilisation des données, mais aussi à développer la capacité de la société civile à façonner démocratiquement la « révolution des données ». L'élargissement de la participation autour de la promotion des données publiques et de l'élaboration des politiques dans le monde augmente les chances de changer ce qui compte, et de composer des infrastructures de données pouvant faire une différence.

## À propos de ce rapport

Ce rapport offre un aperçu des résultats du Baromètre. Le Baromètre complet comprend 39 indicateurs primaires et plus de 500 sous-questions, couvrant 109 pays (fournissant plus de 60 000 points de données au total). Dans ce rapport, nous n'en sélectionnons que quelques-uns à explorer, offrant un aperçu non exhaustif de certains des sujets qui pourraient être explorés plus avant à l'aide des données du Baromètre.

- **La section 1** fournit un bref aperçu des concepts clés utilisés dans le baromètre et une brève description de la méthodologie
- **La section 2** examine les quatre piliers clés du Baromètre : la gouvernance, la capacité, la disponibilité et l'utilisation, et fournit les titres de chacun.
- **La section 3** fournit une analyse régionale, s'appuyant sur les connaissances des pôles

régionaux du Baromètre pour comprendre le contexte unique de chaque région et les forces et faiblesses relatives des pays.

- **La section 4** fournit un bref résumé des enseignements tirés de la première édition et met en évidence les orientations pour les travaux futurs

La méthodologie complète et les détails sur la façon d'accéder et de travailler davantage avec les données du Baromètre sont contenus aux Annexes.

## À propos du Réseau de données pour le développement (D4D.net)

D4D.net est une alliance mondiale de recherche qui s'efforce de renforcer la collaboration entre un large réseau d'acteurs en développant et en mobilisant les connaissances nécessaires pour faire progresser l'utilisation des données pour répondre aux critiques des défis du développement dans les pays du Sud. D4D.net a été lancé en 2021 en tant qu'initiative du Centre de recherches pour le développement international (CRDI) et s'appuie sur les activités du réseau OD4D et de partenaires mondiaux intéressés par le renforcement de la collaboration sur les données ouvertes, l'IA responsable, les mégadonnées, les droits à la vie privée, la propriété intellectuelle, cybersécurité, surveillance en ligne, CRVS, etc.

## Données pour le bien public

### Que signifie l'évaluation des données pour le bien public ?

Un certain nombre de penseurs ont discuté de l'idée que les données numériques devraient être traitées *comme* un bien public<sup>[3]</sup>. En d'autres termes, parce qu'une personne copiant ou utilisant un ensemble de données numériques ne réduit pas la disponibilité des données pour que quelqu'un d'autre puisse les copier ou les utiliser, la plus grande valeur peut être créée en fournissant un accès libre et ouvert aux données. Pour de nombreux ensembles de données, la valeur est également augmentée par l'échelle et la normalisation : les effets de réseau<sup>[4]</sup> signifient que, bien qu'avoir, par exemple, des données d'un pays soit utile, il peut être plus de deux fois plus utile de faire des connexions entre des ensembles de données de deux pays ou plus. C'est là l'une des raisons pour lesquelles les multinationales disposant de vastes ressources de données sont devenues si puissantes. Et c'est l'une des raisons pour lesquelles le développement des données ouvertes en tant que bien public implique souvent des efforts pour créer de nouvelles infrastructures de données publiques interopérables qui peuvent connecter les ensembles de données de différents pays et régions. L'idée que les données sont un bien public était au cœur du Baromètre des données ouvertes (notre étude précédente).

Il existe cependant des exceptions à la notion selon laquelle toutes les données devraient être « ouvertes par défaut », par exemple lorsque les données sont soumises à des intérêts légitimes en matière de protection de la vie privée ou lorsqu'un risque important de préjudice résultant de l'utilisation des données existe. Le spectre de données<sup>[5]</sup> est un outil largement utilisé mettant en évidence que, même si certains ensembles de données peuvent être considérés en **tant** que bien public et fournis ouvertement, d'autres ensembles de données peuvent être mieux gérés en tant que biens partagés du club, communs ou propriété privée, et d'autres ensembles de données (par exemple, les données personnelles sensibles) doivent être gérés en tant que ressources fermées, soigneusement protégées et rendues disponibles uniquement dans le cadre d'accords strictement contrôlés.

Dans le Baromètre mondial des données, lorsque nous parlons de données **pour** le bien public, nous cherchons finalement à poser deux questions connexes :

- Les données sous toutes leurs formes (fermées, partagées et ouvertes) sont-elles gérées de manière à promouvoir le bien public?
- Les données pertinentes sont-elles collectées, partagées et ouvertes afin de soutenir des biens publics plus étendus (santé, éducation, développement durable, justice, etc.)?

Nous reconnaissons surtout que **le** bien public est un concept contesté. Il y a plusieurs publics, plusieurs visions différentes sur la façon dont la société devrait être organisée, et il y a beaucoup de points de vue sur les objectifs auxquels nous devrions travailler individuellement et collectivement. Dans le Baromètre, les Objectifs de développement durable (ODD), convenus dans le cadre d'un vaste processus international, fournissent un point de référence commun l'identification d'un ensemble de biens publics particuliers que les données pourraient aider à fournir, et que nous pouvons fournir une évaluation globale par rapport à : une bonne santé pour tous, à l'action pour le climat et à des institutions justes et fortes.

Parallèlement aux ODD, nous nous appuyons sur d'autres accords et normes internationales largement adoptés pour guider les indicateurs que nous avons utilisés. Nous rendons également disponibles toutes nos données sous-jacentes pour permettre une analyse plus approfondie et/ou alternative. Le Baromètre contient, parmi d'autres points de données, des détails sur : les caractéristiques particulières des règles de gouvernance des données ; les types de parties prenantes ciblées par les interventions

de renforcement des capacités ; les propriétés particulières des ensembles de données publiées ; et des exemples d'utilisation et d'impacts des données. Ceci reflète la nécessité d'aller au-delà de la simple question de la disponibilité en ligne de certains ensembles de données, mais plutôt de se demander : si la fourniture de données repose sur des bases solides ; si les données ont les caractéristiques requises pour être utilisées pour résoudre des problèmes publics particuliers ; et si les communautés ont des capacités largement réparties à travailler avec les données.

La motivation pour explorer ces questions peut se retrouver dans quatre hypothèses intégrées au baromètre sur ce que signifie travailler à l'obtention de données pour le bien public :

- La collecte et le partage des données devraient être régis par des règles publiques légitimes respectant les droits fondamentaux;
- Les pays et les communautés nécessitent des capacités largement réparties pour le partage, la publication et l'utilisation des données afin que celles-ci soient régies et utilisées au niveau national pour le bien public;
- Des ensembles de données particulièrement prioritaires doivent être disponibles, accessibles, ouverts et, dans de nombreux cas, interopérables pour une réutilisation pour soutenir une utilisation généralisée des données pour des biens publics particuliers;
- Il existe une boucle de rétroaction positive entre l'utilisation et la fourniture des données pouvant contribuer à la prospérité des écosystèmes nationaux et sectoriels de données pour le bien public.

Il convient plus particulièrement de noter que dans le Baromètre, nous demandons à plusieurs reprises si les cadres régissant la disponibilité et l'utilisation des données sont clairement définis et s'ils ont force de loi ou non. En d'autres termes, est-il théoriquement possible les acteurs puissants puissent rendre des comptes aux citoyens et, dans les États où le public prend part, dans une certaine mesure, à l'élaboration des lois, les règles elles-mêmes sont-elles soumises à une certaine forme de contrôle ou de surveillance de la part des citoyens ? Bien entendu, les États varient considérablement en ce qui concerne le fonctionnement de la primauté du droit, la mesure dans laquelle les lois reflètent les intérêts publics ou privés et la mesure dans laquelle tous les peuples relevant de leur juridiction se voient accorder des droits de citoyenneté. Les individus peuvent également avoir de plus en plus accès à de nouveaux arrangements de gouvernance des données dépendant moins directement des juridictions des États<sup>[6,7]</sup>. Dans cette première édition du Baromètre, nous ne prenons pas encore pleinement tous ces éléments en compte.

Vous remarquerez aussi tout au long de cette première édition que nous nous concentrons principalement sur la disponibilité et l'utilisation des **données publiques** pour le **bien public**. Autrement dit, bien que nous incluons des indicateurs qui traitent de la gouvernance des données privées (par exemple, la présence de lois et d'institutions de protection des données), et que nous ayons demandé aux chercheurs de nous aviser, lorsque le gouvernement ne fournit pas de données, si elles sont disponibles auprès d'autres sources, y compris le secteur privé, nous ne mesurons pas **l'efficacité** des mécanismes de gouvernance pour protéger les citoyens d'un pays donné envers les préjudices causés par les utilisations ou les abus des données par le secteur privé, et nous ne saisissons pas non plus de données comparatives entre les pays sur la mesure dans laquelle les données du secteur privé sont utilisées pour fournir le bien public dans des contextes particuliers. Cette omission ne signifie pas que ces enjeux ne sont pas importants : ils étaient plutôt plus difficiles à aborder dans l'ensemble de la méthodologie actuelle de la recherche au niveau national.

Fondamentalement, notre approche du bien public reconnaît que la construction du bien public est un processus continu, inachevé et contesté. Un rôle plus important pour les données dans la société ne conduira pas inévitablement à de meilleures sociétés : les politiques et les pratiques en matière de données doivent être intentionnelles si elles veulent mener à des changements, et c'est, en fin de compte, l'objectif et le progrès de ce travail intentionnel que le Baromètre mesure.

### Au-delà des données pour le bien ?

Il existe un certain nombre de projets « données pour le bien » (ou de plus en plus, en mettant l'accent sur les méthodes d'analyse des données utilisées, « l'intelligence artificielle pour le bien ») à travers le monde. Ceux-ci cherchent à utiliser des ensembles de données publiques ou privées pour atteindre certains objectifs sociaux (souvent mal définis). Les discussions sur les « données pour le bien » se limitent généralement à l'examen de la question à savoir si certaines utilisations d'un ensemble de données à des fins sociales ont été fructueuses par rapport à leurs objectifs déclarés.

Un ensemble de données qui est toutefois utilisé « pour le bien » dans un cas, peut également être utilisé « pour le mal » dans d'autres, où peut avoir été collecté, géré ou plus largement utilisé de manière à agir contre le bien public. Pour cette raison, une donnée dans le but du bien public cherche à examiner plus largement les concepts de gouvernance, de capacité, de disponibilité et d'utilisation des données.

# Méthodologie

## Création du baromètre

Le Baromètre mondial des données (GDB) est issu de travaux antérieurs sur le Baromètre des données ouvertes<sup>[8]</sup> (ODB), « *une mesure mondiale de la manière dont les gouvernements publient et utilisent des données ouvertes pour la responsabilité, l'innovation et impact social* ». Ceci à son tour, a été motivé par l'Open Data Index (l'index des données ouvertes)<sup>[9]</sup>, produit le *crowd sourcing* (approvisionnement par la foule) cherchant à fournir « le cliché le plus complet disponible de l'état de la publication de données gouvernementales ouvertes ». L'index des données ouvertes a débuté par une liste de dix domaines de données « que la plupart des gouvernements pourraient raisonnablement s'attendre à collecter », et a proposé sept questions sur la disponibilité des données dans ces domaines. Le Baromètre des données ouvertes a ajouté d'autres questions quant à la préparation, la mise en œuvre et l'impact afin d'évaluer l'état des données ou vertes dans le monde. Quatre éditions complètes du Baromètre des données ouvertes ont été produites par la World Wide Web Foundation entre 2013 et 2016, et une plus petite édition « Leaders » de 30 pays a été publiée en 2018. Une édition séparée du Baromètre des données ouvertes pour l'Amérique latine et les Caraïbes<sup>[10]</sup> a été produite en 2020/21 par le Latin American Open Data Initiative (ILDA).

À la mi-2019, la discussion a débuté avec le réseau Open Data for Development (OD4D) et l'équipe de recherche du Open Government Partnership (OGP) quant au « redémarrage » de l'ODB : plus particulièrement afin de fournir des données actualisées quant à la disponibilité de certains ensembles de données clés. Via un premier processus de mesure, éclairé par les conclusions du livre *The State of Open Data : Histories and Horizons*<sup>[11]</sup> et des entrevues avec les anciens utilisateurs d'ODB, nous avons identifié le besoin d'un cadre plus large réfléchissant les débats actuels quant aux données : en examinant non seulement les données ouvertes, mais aussi le partage de données et la gouvernance des données privées. Au début de 2020, nous avons réuni les membres du réseau OD4D et invité des experts à façonner la conception d'un nouvel instrument d'étude et d'enquête. Grâce au regroupement d'indicateurs prospectifs, l'atelier a élaboré la structure à quatre piliers de la BDG : gouvernance, capacité, disponibilité et utilisation ainsi qu'impact. Cela a entraîné un cadrage autour des « données pour le bien public », et le choix de passer d'indicateurs notés sur une échelle guidée de 0 à 10, à l'utilisation de scores d'indicateurs basés sur des sous-questions détaillées et discrètes qui génèrent un score de 0 à 100.. Les participants à l'atelier de conception ont demandé que chaque point de données de l'étude soit basé sur les cadres normatifs existants, les accords internationaux, les normes ou les données probantes, et que la conception de l'étude soit mieux à même de faire apparaître les « faits saillants » et encourager l'apprentissage par les pairs, plutôt que de mettre l'accent sur le classement par pays.

| Score global  | Gouvernance  | Capacités  | Disponibilité   |
|---|--|--|---|
|            |                     |                                  |  |
| Couvrant la gestion ; la protection et le partage des données ainsi que des données ouvertes, | Couvrant les fondations (par exemple la connectivité ; l'éducation), les capacités et les institutions | Aborder l'existence des données, les caractéristiques (la qualité), l'ouverture et la couverture (étendue) à travers | Analyse préliminaire basée sur des exemples de cas d'utilisation.                     |

| Score global   | Gouvernance  | Capacités                                  | Disponibilité |
|--|--|--|---------------|
| ainsi que de l'examen de la manière dont les données sont traitées au sein des réglementations sectorielles. | gouvernementales en matière de données, et les capacités au sein du secteur privé et de la société civile. | un certain nombre de domaines thématiques. |               |

Grâce au soutien financier du Centre de recherches pour le développement international (CRDI) du Canada par l'emplacement du nouveau réseau Données pour le développement (D4D), une équipe restreinte du Baromètre a commencé à travailler à la mi-2020 afin de mieux conceptualiser la nouvelle étude, en explorant un large éventail de thèmes de données. En collaboration avec des partenaires thématiques et régionaux, des projets d'indicateurs et un manuel de chercheur ont été publiés au début 2021. En réponse à la perturbation du projet suite à la pandémie de la COVID-19, et à des tests initiaux qui ont révélé un compromis potentiel significatif entre la quantité de questions et la qualité des réponses, la première édition du Baromètre a été réduite dans sa portée afin de mettre l'accent sur un nombre limité d'indicateurs thématiques.

Les indicateurs de la première édition ont été classés par ordre de priorité en fonction de la présence de partenariats susceptibles d'examiner, d'analyser et de réutiliser des données, et pour combler les lacunes critiques au sein des domaines d'importance mondiale. Nous avons choisi de mettre l'accent plus particulièrement sur l'intersection des données avec les problèmes de responsabilité de longue date, de pouvoir et d'argent, et d'y inclure un accent sur les problèmes mondiaux urgents, y compris l'urgence climatique et la pandémie de la COVID-19. L'objectif du Baromètre est d'élargir la couverture thématique d'une année à l'autre, en gérant les changements au fil du temps dans l'échantillon des ensembles de données thématiques, des capacités, des règles de gouvernance et des cas d'utilisation étudiés : équilibrer la comparabilité longitudinale et la réactivité aux questions de données émergentes. Un comité de recherche consultatif a examiné la conception finale du cadre d'indicateurs du baromètre, fournissant de la rétroaction critique qui a servi à renforcer l'étude.

De mai 2021 à fin 2021, le travail sur le terrain a eu lieu dans 109 pays, géré par un réseau de centres régionaux. Un chercheur expert de chaque pays a mené une enquête approfondie à laquelle des réponses ont été soumises pour examen aux niveaux régional et mondial. Les données préliminaires ont été partagées avec les partenaires thématiques pour fins de validation supplémentaire, avec des réponses revérifiées, les valeurs aberrantes examinées et des vérifications de validation finales effectuées par l'équipe du Baromètre. Parallèlement, des sources de données secondaires ont été consultées, adoptées et normalisées. (Veuillez noter que bien que les données utilisées pour produire ce rapport aient été produites avec le soutien des organisations partenaires, les données et l'analyse ne représentent pas nécessairement les points de vue, les positions ou les opinions de ces organisations individuelles, et toute erreur ou omission relève de la seule responsabilité du projet du Baromètre mondial de données.)

Une fois les données initiales disponibles, les partenaires régionaux et thématiques ont utilisé les tableaux de bord de l'équipe de baromètre global, ainsi que leur propre analyse de données afin d'identifier les principaux thèmes et messages. Celles-ci ont éclairé la rédaction du présent rapport mondial. Le chapitre régional présente les rapports de chaque pôle de recherche.

Tout le long de ce processus, nous avons effectué une réflexion critique sur la méthodologie du Baromètre, en identifiant les défis particuliers de recherche à relever dans les itérations futures du projet. La publication de ce rapport, ainsi que la présentation qui l'accompagne et les ensembles de données ouvertes disponibles sur le site Web du Baromètre, marque la fin de la première phase de construction du Baromètre. Le travail reste à faire demeure considérable. La section finale de ce rapport décrit certains des futurs domaines à développer, ainsi que de l'apprentissage à mener aux prochaines étapes de ce travail.

## Profil régional

La diversité de la région ressort clairement des données du Baromètre. Pour refléter une perspective inclusive et centrée sur les citoyens et sur la facilitation d'utilisation et la gouvernance ultime des données, tout en tenant compte des systèmes fédéraux, les indicateurs ont été conçus pour attribuer les scores les plus élevés lorsque les règles, les interventions ou les données ont une couverture nationale sans omission significative. Les indicateurs sont cependant également sensibles à l'identification des bonnes pratiques au niveau infranational, même lorsque ce n'est pas réfléchi à l'échelle nationale.

En bref, bien qu'il devrait, par exemple, être possible d'identifier et d'explorer des exemples de bonnes pratiques d'un pays disposant de données de haute qualité sur la propriété foncière pour une grande ville ou un État infranational, les scores les plus élevés devraient être réservés aux pays qui fournissent des données sur la propriété foncière qui répondent aux besoins de tous les citoyens, qu'ils soient urbains ou ruraux, ou quelque soit le type d'occupation pouvant les intéresser.

Le Baromètre a également été conçu pour mettre l'accent sur les questions sur lesquelles les gouvernements ont un pouvoir d'influence. Les gouvernements disposent de toute une gamme d'outils, allant de l'élaboration des politiques et de la législation à la mise en place d'initiatives ou d'infrastructures communes pour l'accès et l'utilisation des données.

## Méthodologie d'enquête

| Années couvertes par l'enquête | Enquêtes de pays réalisées | Centres d'appui régionaux  | Chercheurs            | Partenaires thématiques | Indicateurs secondaires |
|--------------------------------|----------------------------|--|-----------------------|-------------------------|-------------------------|
| 2                              | 109                        | 12   | 113                   | 6                       | 14                      |
| Indicateurs primaires          | Variables primaires        | Points de données collectés (y compris les données justificatives) | URL uniques examinées | Mot de justification    |                         |
| 39                             | 607                        | 107 389  | 17799                 | 581040                  |                         |

Notre enquête d'experts est au cœur de la collecte de données du Baromètre. Dans plusieurs pays, les données issues de l'enquête gouvernementale (une version abrégée de l'enquête d'experts) ont également été prises en compte au cours du processus d'examen. Chaque pôle régional a recruté et formé des chercheurs nationaux, qui ont procédé à la collecte initiale de données, à la suite d'un manuel de recherche détaillé (disponible en ligne sur <https://handbook.globaldatabarometer.org/2021/>). Les réponses au sondage ont ensuite été examinées par les centres régionaux et par d'autres chercheurs nationaux, et les commentaires ont été renvoyés aux chercheurs afin qu'ils procèdent à une collecte et à une vérification supplémentaires des données probantes. Une nouvelle série de révisions et de mises à jour menées par des chercheurs a eu lieu avec la contribution de l'équipe mondiale et des partenaires thématiques, comparant les réponses à des questions spécifiques dans tous les pays. Au cours de la phase finale de validation, de nouvelles corrections et mises à jour ont été effectuées par l'équipe mondiale en s'appuyant sur les données recueillies auprès des chercheurs nationaux et des partenaires thématiques.

La plupart des indicateurs d'enquête d'experts sont basés sur un schéma commun, composé de trois sous-sections :

- **Existence** - évaluer la présence de données probantes de l'existence d'un cadre de gouvernance, d'une capacité, d'un type ou d'une utilisation des données au sein du pays, et la nature de cette existence.
- **Scores des composantes.**
  - **Caractéristiques liées à la qualité** d'une règle, d'un ensemble de données ou d'une capacité. Le choix des caractéristiques, là où c'est possible, était fondé à la fois sur des normes internationales largement acceptées et sur des cas

d'utilisation clairs.

- **Caractéristiques (ouvertes) liées aux données** d'une règle ou d'un ensemble de données, en utilisant un ensemble commun de sous-questions dans le but d'évaluer des questions telles que les références/la présence de données structurées et de licences.
- **Étendue** - évaluer si la gouvernance, la capacité ou les données évaluées sont applicables, pertinentes ou utiles à travers le pays et pour tous les citoyens, ou si elles présentent des limites. Dans le cas des indicateurs d'utilisation de données, ces visent l'étendue des données probantes que l'utilisation entraîne des impacts.

Les chercheurs ont fourni une justification écrite et des sources pour chaque indicateur, de nombreuses sous-questions ont également exigé des informations complémentaires à l'appui, telles que :

- Il existe peu de données probantes accessibles au public que les gouvernements investissent dans les compétences en matière de compréhension des données dans la fonction publique en dehors des unités statistiques ou des bureaux nationaux de statistique ;
- Le contenu des sites Web des organismes, des ministères et des départements est très fragmenté, ce qui fait qu'il est difficile pour le public et les autres parties prenantes de trouver des données pertinentes ;
- Dans certains pays, les plateformes gouvernementales ne disposent pas de moyens faciles de diffuser l'information à un public plus étendu.
- La plus récente date de mise à jour d'un ensemble de données

Il semble enfin y avoir une difficulté avec les gouvernements locaux lorsqu'il s'agit de rendre les données disponibles et ouvertes au public, car l'enquête montre qu'il y a eu un mouvement minimal sur le front des données ouvertes et du gouvernement ouvert au niveau régional ou municipal au-delà de quelques exemples exceptionnels.

La majorité des questions de la sous-section de l'**élément** pouvaient être notées sur une échelle de « Non », « Partiellement » ou « Oui ». L'option « partiellement » a été introduite afin de répondre à une variation entre les pays pouvant signifier, par exemple, qu'un ensemble de données particulier n'existe pas tel que décrit par la question, mais disponible sous une forme comparable, ou sous une forme avec une limitation notable, ou pour saisir les cas où une sous-question demande à savoir si une liste de composants est présente dans un ensemble de données, et que seuls quelques-uns ont été trouvés. Dans certaines indications sélectionnées, des conseils précis quant au moment où la réponse partielle doit être utilisée, ont été fournis aux chercheurs. La plupart des réponses « partielles » ont déclenché une question complémentaire demandant des justifications plus détaillées. Les réponses partielles à une sous-question reçue 50 % de la note qu'une réponse « oui » recevrait.

Ces notes d'élément en sous-question totalisent 100 points, ce qui signifie que, par exemple, un ensemble de données avec toutes les caractéristiques de qualité et d'ouverture débutera l'évaluation avec une note de 100. Ce score est ensuite réduit proportionnellement si les sous-questions relatives à l'existence ou à l'étude indiquent qu'il existe des facteurs affaiblissant la disponibilité, l'applicabilité ou la portée des ensembles de données, des cadres ou des interventions pertinentes. Il peut, par exemple, y avoir un cadre de gouvernance solide pour la protection des données, mais qui n'est applicable qu'à un secteur particulier tel que la santé, ou qui ne s'applique que dans l'un des États d'un système fédéral.

Pour ce faire, les questions relatives à l'**existence** et à l'**étendue** des sous-sections fonctionnent comme des multiplicateurs (à l'exception de l'indicateur de gouvernance « Dans quelle mesure les lois, règlements, politiques et orientations pertinents exigent-ils que la collecte et la publication de données soient accessibles aux personnes handicapées ? Par exemple, si l'existence et la nature d'un cadre d'un ensemble de données répondent à la barre la plus élevée fixée par le Baromètre (par exemple, qui a force de loi, ou est publié par le gouvernement), la valeur du multiplicateur est alors de 1. Mais si un ensemble de données est disponible, mais pas suite à des actions gouvernementales, le multiplicateur sera alors de 0,9. Ainsi, le score maximal disponible pour les ensembles de données non fournis par le

gouvernement est de 90 points au total ( $100 * 0,9$ ). La pareille se produit avec la sous-section Étendue : une capacité, un cadre ou une disponibilité de données avec une large couverture conservera tous les points gagnés à la sous-section Éléments, tandis qu'un exemple isolé sera affecté par le multiplicateur réduisant son score.

Un tableau complet des multiplicateurs est fourni en annexe.

### Données secondaires

En plus de 39 indicateurs primaires, les scores des piliers et des modules s'appuient également sur 14 indicateurs secondaires. Celles-ci sont prises de sources externes soigneusement examinées et ont été transformées sur une échelle de 0 à 100, les valeurs manquantes étant imputées s'il y a un lieu. Une liste complète des indicateurs secondaires est comprise dans l'appendice de la méthodologie.

### Score et calcul

Les scores individuels des indicateurs primaires et secondaires appartenant chacun à un pilier, et à un module, sont utilisés pour calculer :

- **Une note globale** représentant la performance d'un pays au travers des quatre piliers.
- **Les scores des piliers** pour la gouvernance, la capacité, la disponibilité et l'utilisation et l'impact des piliers.
- **Scores des modules** pour chacun des modules thématiques (Informations sur l'entreprise ; Terrain ; Intégrité politique ; Finances publiques ; Achats ; Action climatique ; Santé et Covid-19)

On attribue d'abord à chaque indicateur une pondération fondée sur son type, les indicateurs primaires étant plus pondérés que les indicateurs secondaires, et les indicateurs de gouvernance, de capacité et de disponibilité étant plus pondérés que les indicateurs d'utilisation et d'impact. Les coefficients de pondération d'indicateurs secondaires sont également différenciés en fonction du type de source secondaire, les variables secondaires d'indice étant pondérées en fonction d'une source secondaire supérieure à une seule métrique et en fonction d'une variable secondaire supérieure à une variable secondaire dichotomique. Une ventilation complète de la pondération est fournie à l'annexe sur la méthodologie. Les pondérations sont mises à l'échelle pour placer chaque score de pilier ou de module sur une échelle de 0 à 100, et chaque indicateur est ensuite multiplié par sa pondération et par les résultats additionnés.

Conformément à la conception du baromètre en tant qu'instrument de notation, plutôt que de classement, nous ne procédons pas à une mise à l'échelle min-max des résultats, ce qui signifie qu'un pays ne marquerait 100 sur 100 sur un pilier donné que si tous les indicateurs d'entrée obtenaient également 100 sur 100. L'écart entre un score de pays et 100 sur un score de pilier ou de module représente donc l'écart entre les performances actuelles et un idéal normatif que les indicateurs Baromètre représentant.

Les normes du Baromètre ne sont cependant pas conçues pour être inaccessibles. Si nous prenons le score maximum donné à chaque indicateur, et construisons un pays imaginaire qui combine les meilleures performances trouvées à travers chacun des pays dans le Baromètre, il marquerait 95,92, prouvant ainsi que pratiquement tous les repères fixés par le Baromètre sont, en théorie, réalisables aujourd'hui.

## Apprentissage et limites

En cherchant à fournir une vue d'ensemble des données pour le bien public, à travers plus de 100 pays différents, le Baromètre poursuit un objectif ambitieux. Bien que ce rapport et l'ensemble des données publiées en parallèle nous rapprochent de la réalisation de cet objectif, il est important de prendre en note certaines limites de la méthode Baromètre et des données recueillies dans cette première édition en particulier :

- Nous n'avons pas été en mesure d'inclure autant d'indicateurs de gouvernance des données ou de capacité que nous l'avions initialement prévu. Plus particulièrement, nous avons une couverture plus limitée des utilisations des données par l'intelligence artificielle, et seulement des aperçus initiaux sur les cadres et les capacités de partage de données. Le BMD est cependant un outil de mesure novateur offrant des perspectives uniques sur l'utilisation des données à travers le monde. À ce titre, cet outil doit être affiné et évalué davantage afin d'augmenter l'exactitude, et éventuellement inclure de nouveaux indicateurs et thèmes.
- Peu de sources secondaires que nous avons identifiées offraient une désagrégation par sexe, et notre méthode d'enquête d'experts n'offre que des possibilités limitées à recueillir des données probantes solides sur la mesure dans laquelle les révolutions de données sont équilibrées par sexe ou ont des impacts significativement sexospécifiques. Certains indicateurs comprennent bien sûr des sous-questions quant au genre et à l'inclusion, il ne nous a pas toujours été possible de les inclure parmi les scores des indicateurs. Des efforts supplémentaires seront ajoutés aux éditions futures pour fournir de nombreux autres points de données sur l'inclusion, en général, et plus particulièrement sur l'égalité des sexes.
- Certaines questions de l'enquête ont moins bien réussi à produire des données comparables solides. De multiples cycles d'examen ont eu lieu, mais à la fois de faux positifs et de faux négatifs sont possibles au sein des données du baromètre lorsque l'on examine si certaines règles de gouvernance, certains ensembles de données ou certaines utilisations existent et que l'on évalue les éléments ou les caractéristiques que ceux-ci comportent. Nous offrons par conséquent aux utilisateurs la possibilité de contacter le Baromètre et de partager leurs opinions sur des réponses particulières, si nécessaire.
- Nous n'avons pu inclure qu'un petit nombre d'indicateurs d'utilisation et d'impact au sein de cette édition. Ces indicateurs ont également posé des problèmes particuliers en matière de collecte de données, comme la disponibilité de données sur l'utilisation et l'impact des données varient considérablement d'un pays à l'autre, n'étant pas toujours proportionnellement aux niveaux réels d'utilisation des données. Conséquences, le poids du pilier d'utilisation et d'impact est limité dans les calculs des scores finaux pour refléter ceci - et la comparaison entre pays des scores d'utilisation et d'impact doit être traitée prudemment. Les données qualitatives de ces indicateurs offrent cependant des indications valables quant à l'utilisation des données à travers le monde.

Tout au long de ce rapport, des limites méthodologiques spécifiques sont pertinentes si c'est applicable. Comme beaucoup des ensembles de données étudiés par le Baromètre, nos données devraient être abordées de manière critique et en tant qu'outil parmi d'autres pour la construction de la connaissance et de l'action.

## Utilisation des données

Toutes les données et données probantes collectées par le Baromètre mondial de données sont disponibles comme données ouvertes pour une analyse plus approfondie. Vous pouvez

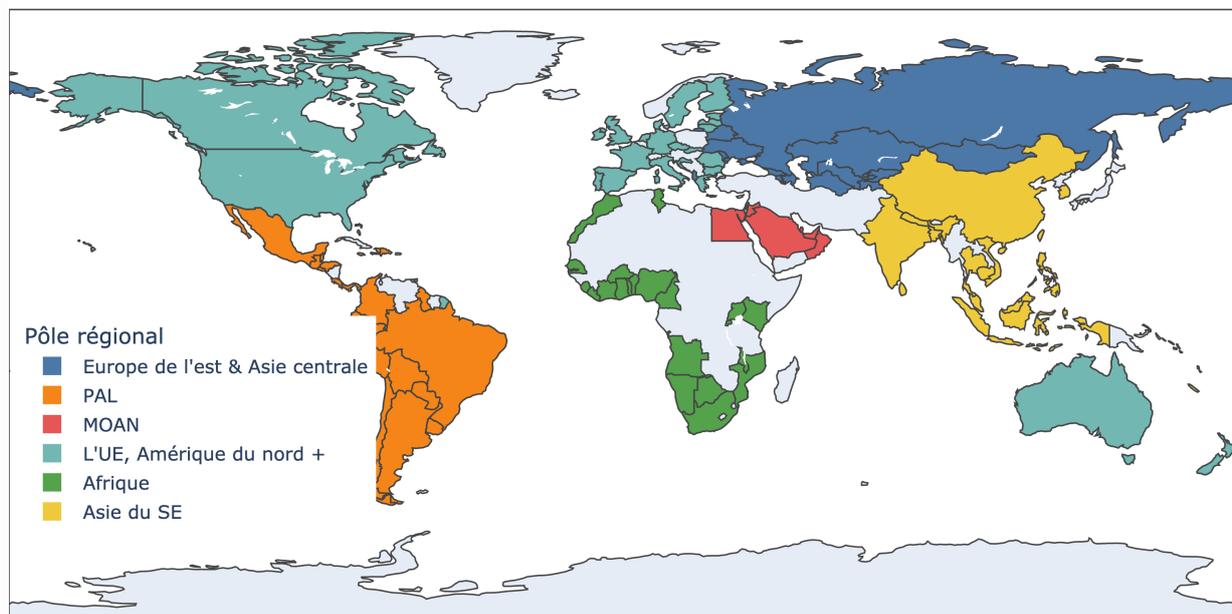
- **Lire le manuel de recherche** afin de voir comment chaque indicateur primaire a été défini et réviser les conseils donnés aux chercheurs.
- **Accédez aux scores globaux** par pilier, module, indicateur ou pays.
- **Politique et initiative en matière de données ouvertes.**

Découvrez comment accéder aux ensembles de données du Baromètre sur <https://www.globaldatabarometer.org>

## Pôles régionaux

Dans ce rapport, nous utilisons six pôles régionaux pour présenter les données. Les régions utilisées étaient basées à la fois sur la proximité géographique et sur la façon dont nos pôles de recherche régionaux traitaient les pays. Chaque pôle de recherche a été invité à sélectionner les pays sur lesquels il se concentrerait, en fonction des priorités régionales et de considérations pratiques. La liste complète des pays de chaque région peut être trouvée à la section 3, ainsi que des détails sur les pôles qui travaillent dans cette région. Les régions que nous utilisons dans le rapport sont les suivantes :

- Afrique
- Europe de l'Est & Asie centrale
- L'Union européenne, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande
- Amérique latine et les Caraïbes
- Moyen-Orient et Afrique du Nord
- Asie du Sud et de l'Est



**Regroupements régionaux:** carte indiquant l'affectation des pays aux groupes régionaux.

# Piliers

- Gouvernance
- Capacités
- Disponibilité
- Utilisation et impact

Le Baromètre mondial des données est basé sur quatre piliers, chacun représentant un aspect différent des données pour le bien public :

- **La gouvernance** consiste à s'assurer que la collecte et le partage des données sont régis par des règles publiques légitimes qui respectent les droits fondamentaux. Ceci couvre les règles qui limitent l'accès à certains types de données et leur utilisation **et** les règles qui favorisent la disponibilité et la réutilisation d'autres types de données.
- **Les capacités** se rapportent aux ressources (connectivité, compétences, institutions, formation, etc.) et aux possibilités de les utiliser (libertés politiques, environnement favorable, etc.), qui soutiennent la collecte, la gestion, le partage et l'utilisation des données de manière à contribuer au développement durable.
- Des enquêtes de **disponibilité** sur la présence, l'ouverture et les principales caractéristiques d'ensembles de données sélectionnés afin de comprendre dans quelle mesure chaque pays rend les ensembles de données clés accessibles sous forme structurée en ligne, adaptée aux besoins du public.
- **L'utilisation et l'impact** recherche des données probantes de cas particuliers d'utilisation des données et explore quels groupes de parties prenantes utilisent les données dans chaque pays.

Les sections suivantes présentent les données clés de chacun de ces piliers, en s'appuyant à la fois sur les indicateurs quantitatifs du baromètre et sur des données qualitatives.

## Gouvernance

Les gouvernements peuvent façonner la collecte, la disponibilité et l'utilisation des données de plusieurs manières clés, notamment en établissant des règles, en adoptant des normes, en suivant les bonnes pratiques et en établissant des institutions et des réseaux de surveillance ou en collaborant avec eux.

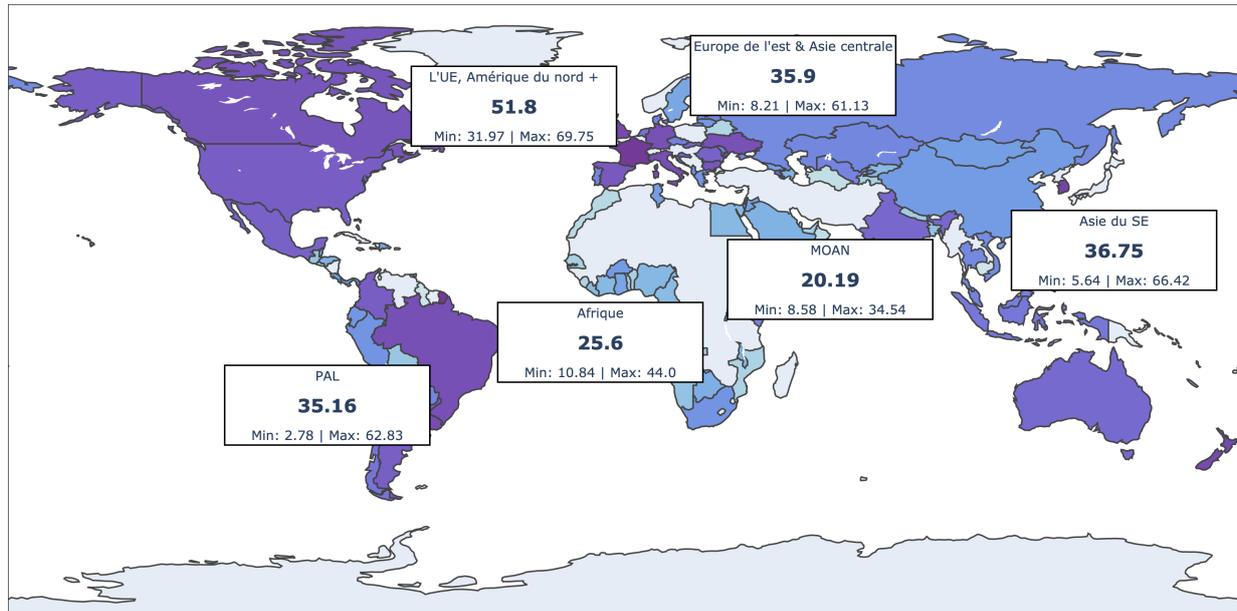
La gouvernance des données pour le bien public combine une bonne gestion des données avec la mise en place de politiques et de cadres solides qui protègent les droits des individus et des communautés en matière de données et qui rendent les données non sensibles largement disponibles pour la réutilisation. La gouvernance des données pour le bien public devrait aborder explicitement les questions d'inclusion : s'assurer que la production et l'utilisation des données réduisent, plutôt qu'élargissent, les inégalités sociales.

Dans cette première édition du Baromètre, nous avons spécifiquement examiné dans quelle mesure la gouvernance des données se base sur des règles contraignantes. En d'autres termes, existe-t-il des lois, des réglementations, des politiques ou des directives étant exécutoires et qui, en supposant un état de droit fonctionnel, fournissent les bases d'une approche cohérente et stable de la manière dont les problèmes liés aux données seront gérés.

### Résumé

- Les pays ont de plus en plus de lois protégeant les données personnelles. 98 Sur les 109 pays étudiés par le Baromètre ayant une certaine forme de cadre, bien qu'au sein des 13 pays ceux-ci n'aient pas force de loi, et que dans 12 les pays les protections soient limitées à des secteurs particuliers, ne couvrant pas pleinement l'utilisation des données des secteurs public et privé. La majorité des cadres intègrent des principes clairs de choix ou de consentement, ainsi que des droits d'accès et de correction des données. Il y a cependant, moins de notifications de violation de couverture, et il y a des données probantes d'un besoin de moderniser de nombreux cadres afin qu'ils répondent mieux aux risques particuliers liés aux données de localisation et concernant l'utilisation algorithmique des données personnelles.
- La gestion des données et de leurs politiques d'ouverture sont de plus en plus correctement établies, bien que beaucoup moins de pays disposent de cadres de données solides régissant le partage des données, ce qui risque de créer des lacunes en ce qui a trait à la fois à la facilitation et à la réglementation de l'échange de données sensibles entre les pouvoirs publics et d'autres secteurs, limitant les possibilités d'assurer une utilisation fiable des données pour le bien public. 30 les pays disposent désormais des politiques juridiquement contraignantes en matière de données ouvertes, bien que plusieurs de ceux-ci n'aient pas encore abordé de manière adéquate la normalisation et l'interopérabilité des données ouvertes publiées.
- Un peu plus de la moitié des lois et cadres sectoriels que nous avons identifiés visaient la collecte et la publication de données dans le cadre de règles contraignantes. Les règles en matière de finances publiques, de droit à l'information (RTI) et de déclaration de patrimoine étaient les plus susceptibles de spécifier que les données collectées devraient être publiées en tant que données ouvertes structurées. Bien que de nombreux cadres traitent de la qualité des données par la mise en place d'institutions de surveillance, moins de la moitié des cadres établissent des processus de vérification et seulement 36.69% soutiennent explicitement la collecte de données structurées. Il existe un lien évident entre les lois

spécifiant que les données doivent être publiées et les données mises à disposition, bien que la «différence de mise en œuvre» entre les exigences de publication et celles de publication varie selon le secteur.



**Scores régionaux du pilier de gouvernance:** Notre regroupement régional UE, Amérique du Nord + obtient les scores les plus élevés sur le pilier de gouvernance. Les pays du Moyen-Orient et d'Afrique du Nord ont le score moyen le plus bas.

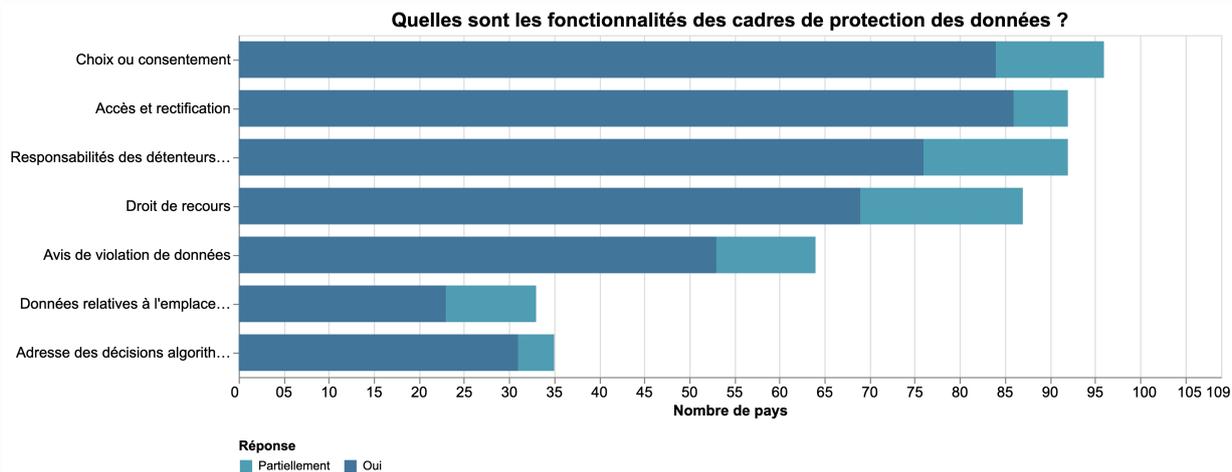
### Une note sur les « cadres »

Dans le Baromètre, nous abrégeons souvent « lois, règlements, politiques et directives pertinents » comme « cadre ». En effet, la base de la gouvernance, de la collecte ou de la publication des données se répartit souvent entre plusieurs lois, règlements, politiques et documents d'orientation. Par exemple, une loi peut habiliter un organisme à recueillir des données, un autre règlement ou mémorandum peut préciser que les données doivent être fournies sous forme structurée, et une autre loi peut prescrire que lorsque des données sont fournies, ce doit être fait sous licence ouverte. Le terme « cadre » est utilisé pour représenter l'ensemble des lois, des règlements, des politiques et des directives pertinents, il ne signifie pas qu'un gouvernement lui-même les présente que ceux-ci sont nécessairement compris en tant que cadre unifié.

## Dans quelle mesure les cadres de protection des données sont-ils bien développés ?

La croissance constante de la couverture mondiale des lois sur la protection des données a été largement étudiée, [greenleafSheherezade101Data2013;<sup>[12]</sup>], et les données du baromètre reflètent la tendance, avec 98 (89.9%) des pays interrogés ayant au moins une certaine forme de cadre de protection des données. Un certain nombre d'entre eux sont cependant limités à des secteurs particuliers, tels que les services financiers, ou ne s'appliquent qu'à la collecte de données par le gouvernement central, n'offrant aucune protection aux citoyens contre les abus de données par le secteur privé. Dans d'autres cas, le secteur public est exempt d'exigences en matière de protection des données. En Amérique latine et dans les Caraïbes, au Moyen-Orient et en Afrique du Nord, en Asie du Sud et de l'Est et en Afrique, certains pays n'ont pas encore mis en place de cadre de protection des données ou dans lesquels les cadres doivent être considérablement renforcés.

Des cadres solides de protection des données devraient être pris en considération en tant que condition préalable à la réalisation de travaux visant à promouvoir la réutilisation des données pouvant être dérivées en dernier ressort de données individuelles ou liées à ces données, bien qu'elles ne constituent qu'une partie d'un régime robuste de gouvernance des données, qui doit également envisager la manière de régler d'autres formes de préjudice lié aux données. L'évolution récente des normes mondiales pour les cadres de protection des données a mis un accent particulier sur l'amélioration de la notification des violations, la reconnaissance des sensibilités particulières des données de localisation et l'utilisation des données dans les applications et algorithmes d'intelligence artificielle<sup>[13,14]</sup>.



**Quelles fonctionnalités les cadres de protection des données contiennent-ils ?** La majorité des cadres de protection des données portent sur le choix et le consentement, l'accès et la correction, la responsabilité des détenteurs de données et les droits de recours. Moins d'avis de violation de données sont couverts, et une minorité abordent explicitement les données liées à l'emplacement, et l'utilisation des données dans la prise de décision algorithmique.

Les données du baromètre indiquent que, si certains aspects de la réglementation en matière de protection des données sont largement répandus, les pays dotés de cadres (45.9%) semblent ne pas avoir de dispositions solides en matière de notification des violations de données et 29 (29.6%) ont un droit de recours limité en ce qui a trait au préjudice résultant d'abus de données. Seuls 23 (23.5%) des cadres disponibles traitent les données de localisation de manière robuste, avec un peu plus de (31/31.6%) de prise de décision algorithmique.

Notamment, dans 21.4 % des pays dotés de cadres, les chercheurs ont signalé au moins un certain degré de limitations imposées au fonctionnement des cadres de protection des données dans le contexte de pandémie de la COVID -19, allant des régulateurs annonçant qu'ils adopteraient une approche plus légère en matière de signalement et d'application pour les organisations impliquées dans la réponse à la pandémie, jusqu'à la suspension des exigences de consentement pour le partage de données liées à la santé, ou l'autorisation d'activer les enregistrements de téléphone mobile pour la surveillance des mouvements de population. Alors que dans certains cas, ceci impliquait l'utilisation de dispositions déjà en place pour les situations d'urgence, les militants ont exprimé la crainte que certains processus de protection des données aient été affaiblis pendant la réponse à la pandémie, soulignant la nécessité de suivre à la fois la mesure dans laquelle les exemptions ou les changements restent en vigueur après la pandémie, et d'explorer si les régulateurs reviennent pour examiner les pratiques qui auraient pu se développer lors d'une période de réglementation plus douce.

Il convient également de prendre note que l'on craint de plus en plus que les cadres de protection des données fondés sur les droits individuels ne couvrent qu'une partie des préjudices potentiels de la collecte et de l'utilisation des données devant être gérés. Les données qualitatives recueillies par le Baromètre fournissent un aperçu du fonctionnement des systèmes nationaux de protection des données, mais les travaux futurs devront porter sur un plus vaste éventail de modalités permettant de s'assurer que les données ne sont pas utilisées pour menacer les droits des individus et des communautés.

## Dans quelle mesure les cadres de gestion des données, de partage et d'accès sont-ils bien développés ?

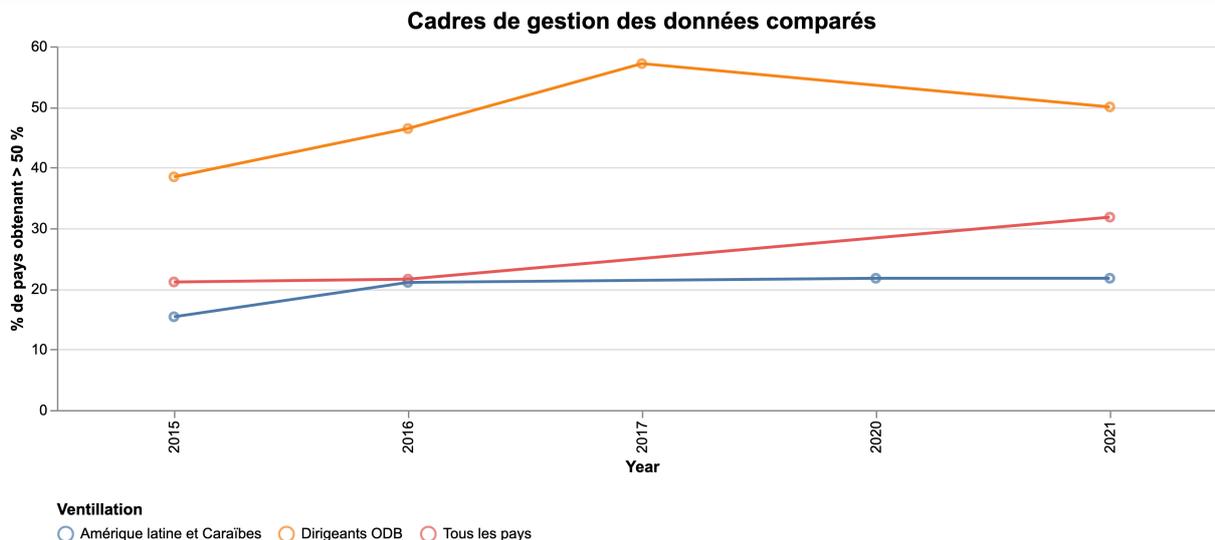
| Pays dotés de cadres de gestion des données | Pays dotés de cadres de partage de données | Pays ayant des politiques de données ouvertes |
|---|--|---|
| 71.6%                                       | 62.4%                                      | 67.9%   |
| Score moyen: 30.28 sur 10                   | Note moyenne : 24.05 sur 10                | Score moyen : 37.69 sur 10                    |

Tableau : Indicateurs retenus pour le pilier de gouvernance

La valeur des données pour le bien public, peu importe qu'il s'agisse de données ouvertes ou non, est accrue lorsque les données sont plus facilement découvrables, lorsqu'elles sont accompagnées d'une documentation claire, lorsque la qualité des données a été assurée, lorsque des normes techniques appropriées sont utilisées et lorsque l'on cherche à obtenir l'avis des utilisateurs pour améliorer la gestion des données. Les gouvernements peuvent promouvoir des approches cohérentes et de haute qualité de la gestion des données par divers moyens, y compris des stratégies nationales en matière de données, des directives de gestion des données et des normes de gestion des données. Des cadres clairs pour la gestion et le partage des données peuvent également protéger contre les méfaits de l'usage abusif des données, en veillant à ce que les données soient traitées via des processus transparents et à ce qu'il existe des lignes de responsabilité autour de l'utilisation des données.

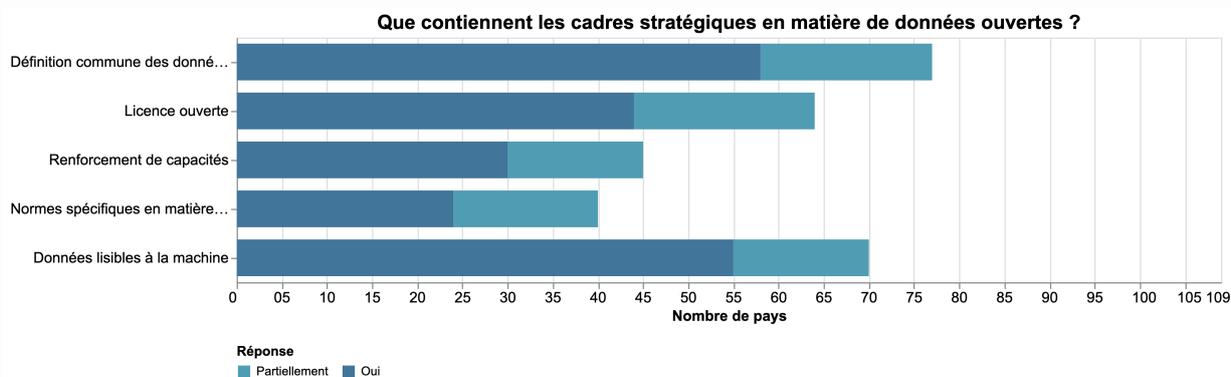
Le Baromètre a demandé aux chercheurs de vérifier la présence de cadres de gestion des données et d'en explorer les caractéristiques. Nous avons constaté que 36 (33%) des pays disposaient de cadres de gestion des données contraignants et que 42 (38.5%) d'autres pays disposaient de cadres moins contraignants (par exemple, des directives ou des stratégies). Parmi ceux-ci, seulement 8 (10.3%) présentaient des données probantes de mécanismes documentés robustes pour solliciter ou intégrer les commentaires des utilisateurs externes afin d'améliorer la qualité des données, bien que 49 (62.8%) a référencé des normes minimales pour les métadonnées.

En tout et partout, 29 pays ont obtenu une note supérieure à 50 sur 100 pour la qualité et la couverture de leurs cadres de gestion des données. Comme notre indicateur de gestion des données a été conçu pour être largement comparable à un indicateur précédent du Baromètre des données ouvertes, collecté depuis 2015, nous pouvons suivre l'évolution de la qualité des cadres de gestion des données au fil du temps. Comme indiqué par le graphique ci-dessous, en ne considérant uniquement que les pays inclus à la fois dans l'ODB et la GDB qui ont obtenu un score supérieur à 50% sur cet indicateur lors d'une année donnée, il y a des données probantes de progrès réguliers, avec une augmentation de 21% des pays atteignant ce seuil en 2016, à 31% en 2021. Le graphique compare également les scores pour seulement les 28 pays inclus dans l'édition des leaders de l'ODB 2017 (montrant une légère baisse du score), et pour jusqu'à 23 pays de la région d'Amérique latine et de Caraïbes (aucune croissance récente) pour lesquels d'autres données ont été collectées en 2020. L'ajout de ces comparaisons suggère, plus particulièrement, qu'une tendance mondiale modeste vers une gestion plus forte des données provient de pays autres que les signataires de la Charte sur les données ouvertes couverts par l'édition des dirigeants de l'ODB.



**Comparaison des scores de gestion des données du Baromètre de données ouvertes et du Baromètre de données global.** De modestes progrès ont été réalisés à l'échelle mondiale sur la voie de cadres de gestion des données plus solides entre 2015 et 2021, bien que cela soit moins évident dans les pays couverts par « l'édition des leaders » du baromètre des données ouvertes. Il semble que les progrès vers l'amélioration de la gestion des données en Amérique latine soient au point mort. Remarque : les différences de méthodologie entre la BDO et la BDG peuvent expliquer une partie de la variation indiquée ci-dessus.

En ce qui a trait les politiques en matière de données ouvertes, nous avons trouvé des 74 pays disposant d'une certaine forme 30 de politique, dont les politiques étaient juridiquement contraignantes. 91.9% des politiques offrent une définition commune des données ouvertes, avec 83.8% nécessitant des données informatiquement lisibles, et 78.4% promouvant les licences ouvertes sans restriction au-delà de l'attribution et de l'actionnariat. Cependant, seulement 47.3% traitent des normes communes pour les données, ce qui suggère une opportunité pour que les futurs travaux de se concentrent davantage sur la normalisation et l'interopérabilité des données ouvertes publiées.

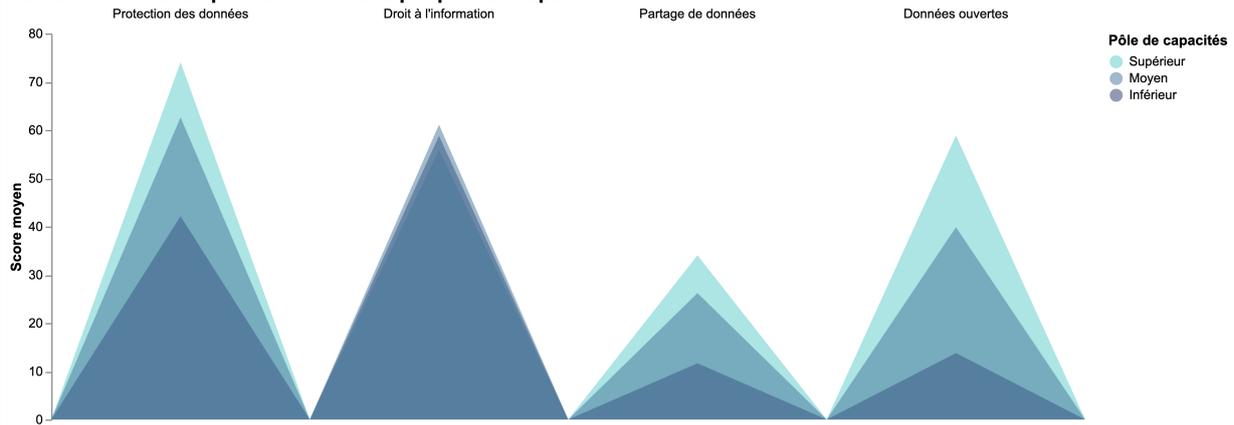


**Que contiennent les cadres stratégiques en matière de données ouvertes ?** Les politiques de données ouvertes sont plus susceptibles de couvrir les formats de données qu'elles ne le sont pour répondre aux exigences de licence ou de renforcement des capacités des fonctionnaires.

## Protection et accès sur l'ensemble du spectre de données

Réalisation de la valeur des données pour le bien public implique de trouver la bonne approche pour chaque ensemble de données qui peut maximiser l'accès et la réutilisation, tout en respectant les droits individuels et collectifs sur les données. L'Open Data Institute a développé le Data Spectrum<sup>[5]</sup> en tant qu'outil permettant de déterminer si les données doivent être fermées (conservées en toute sécurité), partagées (fournies à des tiers de confiance, souvent sur la base d'accords formels) ou ouvertes (accessible à tous et réutilisable sans restriction).

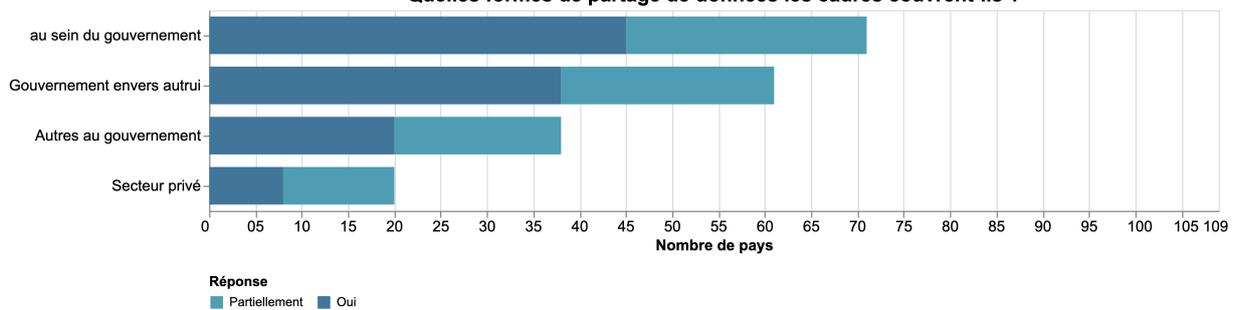
**Cadres à travers le spectre de données par pôle de capacité**



**Le spectre de données exploré :** Les pays à capacités élevées et moyennes progressent dans la mise en place de cadres de données ouvertes et de partage de données solides et dotés de ressources suffisantes, tandis que les pays à capacités faibles ont moins progressé dans ce domaine. Il existe une différence limitée en raison des capacités numériques entre la qualité des cadres du droit à l'information des pays, tandis que ceux dotés de capacités élevées sont plus susceptibles de disposer de cadres solides de protection des données. Visualisation basée sur le spectre des données ouvertes développé par l'Open Data Institute.

En divisant les pays du Baromètre en trois pôles de capacité (voir le chapitre suivant), nous pouvons explorer la maturité relative des cadres de gouvernance pour chaque partie du spectre de données. Cependant, à travers tous les pôles, la gouvernance du partage des données est moins bien développée (du moins lorsqu'il s'agit d'être placé sur une base juridique), et les pays à faible capacité accusent un retard important en termes de protection et de partage des données et de cadres de données ouvertes. En mettant l'accent sur le partage de données pour de bonnes initiatives, y compris dans les pays à faible capacité, cela peut être une source de préoccupation particulière, et met en évidence les domaines nécessitant un développement important. Comme le montre le graphique ci-dessus, il y a relativement peu de différences entre les pays en fonction de leurs capacités lorsqu'il s'agit de la présence de cadres solides en matière de droit à l'information, et les pays à forte et moyenne capacité disposent de cadres de données ouvertes raisonnablement solides.

**Quelles formes de partage de données les cadres couvrent-ils ?**



**Que couvrent les cadres de partage de données ?** Les cadres de partage des données sont plus de trois fois plus susceptibles de réglementer le partage des données au sein du gouvernement que les ententes de partage des données avec le secteur privé.

En examinant davantage les types de partage de données couverts par les cadres existants, nous trouvons la majorité (92.6% des 68 cadres identifiés) régissent le partage des données au sein du gouvernement, avec 79.4% couvrant la façon dont le gouvernement devrait partager les données avec d'autres secteurs, et 51.5% examinant comment les autres secteurs devraient partager les données avec le gouvernement. Seulement 16.2% traitent explicitement des utilisations des données par l'intelligence artificielle, et uniquement 26.5% regardent le partage des données au sein du secteur privé. Une gouvernance accrue du partage des données du secteur privé a été à l'ordre du jour de l'Union européenne grâce au concept « d'espaces de données » introduit dans la stratégie européenne en matière de données<sup>[15]</sup>, mais cela ne semble pas encore avoir été intégré dans les cadres nationaux. Dans un certain nombre de cas, lorsque les chercheurs n'étaient pas en mesure de trouver des cadres de partage des données, ils ont examiné les dispositions relatives au partage des données dans le cadre de la législation sur la protection des données. Au cours des prochaines éditions du Baromètre, nous chercherons à renforcer la définition des cadres de partage des données afin que ces

cas ne soient pas comptés, car ils ne mettent généralement pas l'accent sur les mécanismes qui régissent spécifiquement le partage des données. En tous les cas, nous prévoyons que cela a conduit à un comptage marginal des cadres de partage de données dans cette édition.

En somme, les données qualitatives recueillies par l'enquête Baromètre ont révélé plusieurs approches différentes à travers le monde à la gouvernance du partage des données, des approches axées sur l'établissement de règles, la fourniture de plateformes, la promotion de l'interopérabilité, la création de nouveaux pouvoirs gouvernementaux et la fourniture de conseils au gouvernement et à l'industrie. La création de nouveaux pouvoirs gouvernementaux et la fourniture de conseils au gouvernement et à l'industrie. Une grande partie de la pratique demeure au niveau des politiques, plutôt que des règles juridiques contraignantes.

## Dans quelle mesure la gouvernance tient-elle compte de l'accessibilité et de l'inclusion ?

Deux des principaux obstacles au caractère inclusif de la politique de données observées au début de la montée en puissance du mouvement des données ouvertes étaient la tendance à ne fournir des données et des portails que dans une seule langue, même dans les pays où de nombreuses langues sont parlées, et à concevoir des plateformes de données sans accorder beaucoup d'attention aux caractéristiques d'accessibilité, telles que la compatibilité avec les technologies d'assistance telles que les lecteurs d'écran, ou en appliquant des principes de conception universelle lors de la création de sites Web et d'outils liés aux données.

Le Baromètre comprenait deux indicateurs conçus pour déterminer dans quelle mesure les pays disposaient de cadres de gouvernance pour promouvoir l'accessibilité inclusive des données. L'un d'entre eux, qui porte sur la langue, a été exclu de la notation, car nous n'avons pas été en mesure d'obtenir des résultats suffisamment fiables à partir de notre question de sondage. L'autre, sur l'accessibilité, révèle que si 66 des pays disposent d'une forme de loi, de réglementation, de politique ou d'orientation qui exige que la collecte et la publication des données soient accessibles aux personnes handicapées, cela découle dans la plupart des cas des règles générales d'accessibilité, seuls 17 des pays disposant d'un cadre d'accessibilité dédié ou de dispositions spécifiques relatives aux données.

Au niveau régional, nous observons des variations significatives dans les cadres d'accessibilité, avec les scores les plus faibles dans la région du Moyen-Orient et de l'Afrique du Nord, et le score moyen le plus bas suivant en Afrique.

| Afrique | L'Europe de l'Est et l'Asie Centrale | L'Union européenne, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande | Amérique latine et Caraïbes | Moyen-Orient et Afrique du Nord | L'Amérique du Sud |
|---------|--------------------------------------|--|-----------------------------|---------------------------------|-------------------|
| 19.11   | 22.33                                | 58.97  | 30.4                        | 5.06                            | 30.78             |

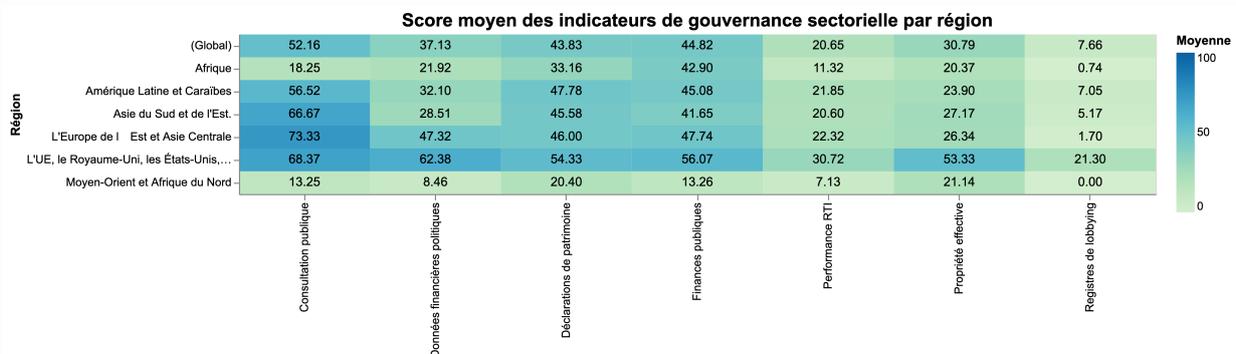
Tableau : Notes moyennes sur l'indicateur d'accessibilité dans le pilier gouvernance par région

Dans cette première édition du Baromètre, nous n'avons pas été en mesure d'explorer la mesure dans laquelle les cadres de gouvernance individuels pour l'inclusion des données abordent l'inclusion, bien qu'il s'agisse d'un domaine pour les travaux futurs, en s'appuyant sur un certain nombre de questions plus larges axées sur l'inclusion dans le Baromètre, y compris une question contextuelle posée pour identifier des modèles d'exclusion et de marginalisation particulièrement pertinents dans chaque pays.

## Les lois, politiques ou orientations sectorielles sont-elles « sensibles aux données » ?

Près de la moitié de la note attribuée au pilier gouvernance du baromètre provient d'indicateurs de gouvernance sectorielle. Au sein de ces indicateurs, nous examinons dans quelle mesure certaines lois et certains cadres thématiques sont « sensibles aux données ». Autrement dit, les lois ou les politiques tiennent-elles compte du fait que les processus gouvernementaux, tels que l'établissement des budgets nationaux, la tenue de consultations publiques ou la mise en place de mécanismes visant à améliorer l'intégrité politique, génèrent tous des données et s'appuient sur celles-ci. Les premiers efforts dans le but d'ouvrir les données gouvernementales reposaient souvent sur des décisions administratives pour la publication des ensembles de données. Les décisions pouvant être révoquées en tout temps. Lors de la collection et la publication des données reposent sur des bases solides, la société civile et les entreprises sont plus susceptibles de pouvoir s'appuyer sur elles et de mettre en place des processus se servant des données. Et lorsque la collecte et la publication de données sont explicitement abordées dans les lois et règlements, elles peuvent gagner en légitimité et faire l'objet d'un examen public plus approfondi.

Chaque indicateur de gouvernance sectorielle est jumelé à un indicateur de disponibilité des données pour soutenir la recherche sur la relation entre la collecte de données et les règles de publication, et les données partagées ou ouvertes. Au sein de chaque indicateur, nous examinons les caractéristiques particulières importantes pour la qualité des données pouvant y être produites, y compris les caractéristiques générales de tous les types de données (telles que la fourniture de données structurées, les mises à jour en temps opportun et la mise en place de processus d'assurance ou de vérification de qualité) et les caractéristiques spécifiques au sujet (telles que les identifiants interopérables ou la désagrégation via d'importantes variables).

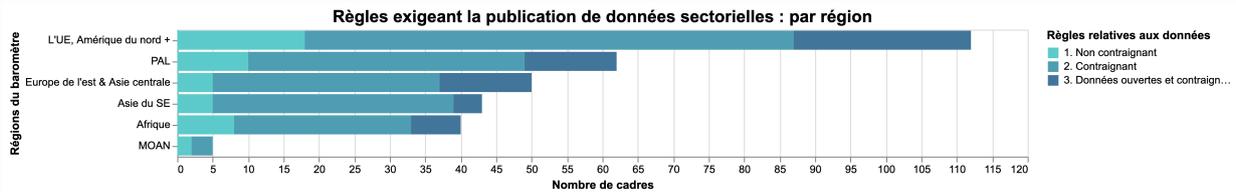


**Résultats globaux des indicateurs de gouvernance sectorielle par région.** Les notes moyennes les plus élevées, indiquant que les lois ou les cadres sont plus largement en place et présentent davantage de caractéristiques pertinentes, ont été trouvées pour les « déclarations de patrimoine ». Les scores les plus faibles ont été trouvés pour les registres de lobbying.

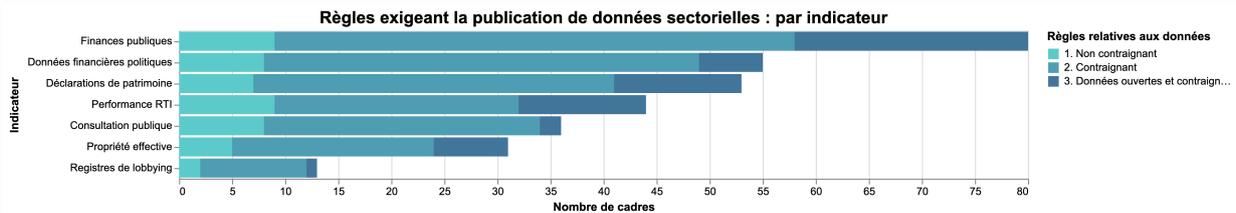
Le graphique ci-dessus indique les scores moyens pour chacun de ces indicateurs, ventilés par région. 186 des 507 (36.69%) cadres de gouvernance opérationnelle identifiés par notre enquête soutiennent la collecte de données structurées, et 246 (48.52%) nécessitent une certaine forme de processus de vérification pour assurer la qualité des données. Reconnaisant l'importance des institutions dans l'adoption des règles de gouvernance, 358 occurrences de cadres étudiés (70.61%) habilite une agence ou un fonctionnaire à superviser des éléments de collecte et de publication des données.

Lorsqu'il s'agit spécifiquement de la publication de données ouvertes, nous ne trouvons aucune mention de la publication de données dans certains 195 (38.46%) des lois, politiques, réglementations ou orientations opérationnelles, avec 48 (9.47%) décrivant certaines exigences de publication de données dans des politiques ou orientations non contraignantes, 202 (39.84%) y compris dans des politiques, réglementations ou lois contraignantes, et 62 (12.23%) fournissant les formes les plus fortes d'exigence de publication de données ouvertes. Parmi ceux-ci, 25 nous trouvons dans l'Union

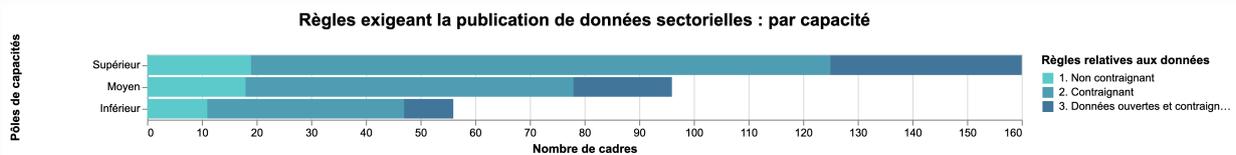
européenne, au Royaume-Uni, en Amérique du Nord, en Israël, en Australie et en Nouvelle-Zélande, ainsi que dans les pays à plus forte capacité, les lois sur les finances publiques comptant 35 parmi les exigences les plus strictes en matière 22 de données ouvertes.



**Règles sectorielles exigeant la publication de données par région:** toutes les régions, à l'exception de la région MENA, sont tenues de publier des données ouvertes dans le cadre des cadres de gouvernance sectoriels



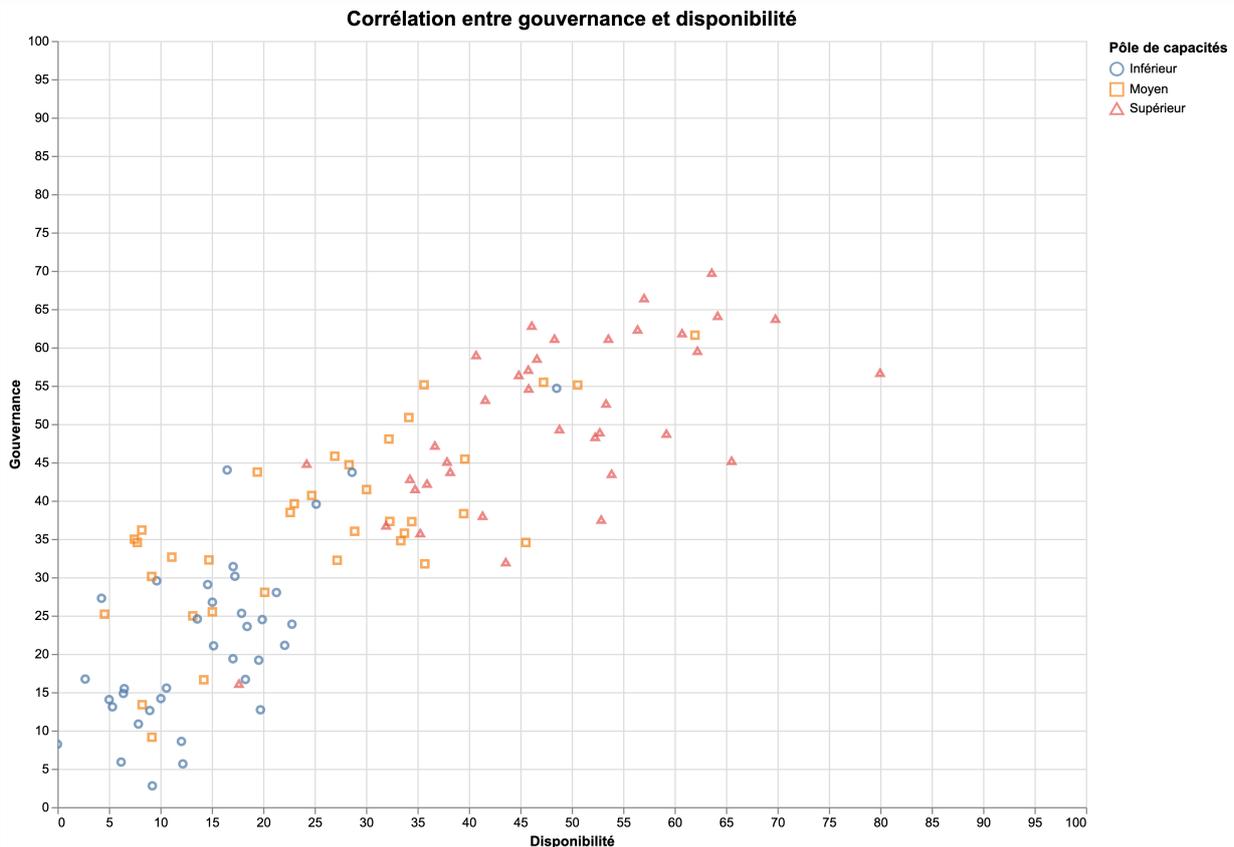
**La règle sectorielle exigeant la publication de données par thème:** les cadres des finances publiques sont les plus susceptibles d'avoir des exigences contraignantes pour la publication à la fois des données et des données ouvertes



**Les règles sectorielles exigeant la publication de données par groupe de capacités:** la majorité des cadres ayant des exigences contraignantes pour la publication de données ouvertes existent dans les pays à capacité supérieure

## Gouvernance et mise en œuvre

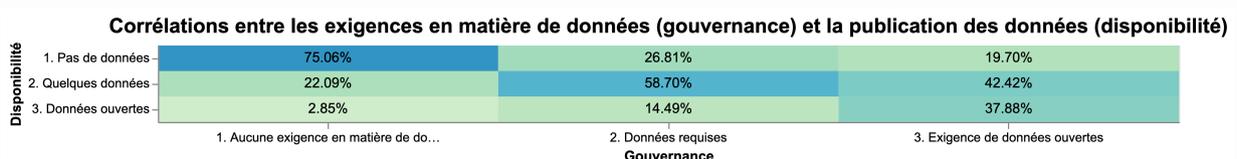
Nous pouvons utiliser les variables du Baromètre pour explorer le lien entre la qualité de la gouvernance des données et la qualité de la disponibilité des données. À un niveau élevé, le graphique ci-dessous indique un diagramme de dispersion des scores du pilier de gouvernance par rapport au pilier disponibilité, indiquant une corrélation large. Nous pouvons cependant voir que cette relation peut, dans une certaine mesure, s'expliquer aussi par la capacité des pays.



**Corrélation entre les piliers de gouvernance et de responsabilité:** il existe une relation positive entre la gouvernance et les scores de disponibilité

Nous pouvons aussi examiner le niveau de gouvernance des données appariées et les indicateurs de disponibilité des données, dans le but développer la matrice trois par trois ci dessous, explorant la mesure dans laquelle il existe des cas où, pour chaque paire de pays et d'ensembles de données (109 x 7 = 763 cas au total), il existe des cadres ou des règles qui exigent que les données soient collectées et publiées, et si ces données sont alors disponibles ou non.

Comme indiqué par la matrice ci-dessous, dans 75.06% des cas où il n'existe pas de cadre imposant la collecte ou la publication de données, il n'existe pas non plus de données disponibles. Lorsque la publication de données est requise par les cadres de gouvernance, dans 58.7% des cas, les données sont disponibles sous une forme ou une autre, et dans 14.49% des cas, les données répondent à la définition ouverte. Ceci laisse toujours un écart de mise en œuvre dans 26.81% des cas, où les cadres de gouvernance nécessitent la publication de données, mais aucune donnée n'a été révélée par notre enquête. L'écart de mise en œuvre est plus important quant aux données ouvertes, où seuls 37.88% des exigences contraignantes de publication de données ouvertes semblent entraîner la disponibilité de données ouvertes, bien que certaines données, bien qu'en deçà de la définition ouverte, soient disponibles dans un autre 42.42% de ces cas.



**La comparaison des exigences de gouvernance des données avec la disponibilité des données:** L'utilisation de toutes les paires d'indicateurs de gouvernance et de disponibilité montre que les exigences de données plus fortes entraînent une plus grande disponibilité des données. Les pourcentages sont donnés par colonne.

L'écart de mise en œuvre varie également d'une catégorie de données à l'autre. Pour les données relatives aux finances publiques, il existe une quantité importante de données ouvertes disponibles même lorsque des règles de formatage n'ont pas été trouvées qui l'exigent, alors que pour les

déclarations d'intérêts et d'actifs des politiciens, les exigences strictes en matière de données ouvertes dans les cadres de gouvernance ne se reflètent pas encore dans la disponibilité de données structurées et ouvertes.

## Comment les pays peuvent-ils améliorer la gouvernance des données à l'avenir ?

Pour chaque pays, le site Web du Baromètre mondial de données contient un profil indiquant la répartition des scores individuels par rapport aux indicateurs de gouvernance. Il peut être utilisé pour identifier des domaines particuliers d'amélioration. Sur la base des données quantitatives et qualitatives recueillies pour ce pilier de gouvernance, nous mettons en évidence les domaines d'action communs comme suit :

### Revoir et actualiser les cadres fondamentaux

Il y a 16 pays couverts par le Baromètre qui ont obtenu un score inférieur à 5 à notre indicateur de protection des données. Vu l'importance fondamentale des cadres de protection des données pour gérer les risques d'utilisation abusive des données tout en soutenant l'utilisation des données dans l'intérêt public, cela suggère un besoin prioritaire de créer, de mettre en œuvre ou de renforcer les cadres de protection des données dans ces pays. **Tous** les pays doivent cependant revoir régulièrement leurs cadres de protection des données et de gouvernance, en veillant à ce qu'il existe des processus pour examiner l'efficacité des mécanismes conçus pour présenter les abus de données et pour répondre à l'évolution du paysage des risques et des opportunités liés aux données. Il convient plus particulièrement de veiller à ce que le renforcement des règles de protection des données ne porte pas atteinte aux intérêts publics légitimes en matière de données de responsabilité.

Les données ouvertes et les cadres de gestion des données, dans plusieurs pays, doivent également être renforcés, notamment en ce qui concerne la normalisation et l'interopérabilité des données.

Pour que les données puissent être utilisées en vue d'un modèle inclusif du bien public, les pays doivent également se concentrer sur des cadres visant à garantir que les données et les plateformes liées aux données soient accessibles aux personnes handicapées.

### Établir et itérer des cadres clairs pour le partage de données

Au cours de la prochaine décennie, nous assisterons probablement à une augmentation des accords volontaires et obligatoires de partage de données entre les entreprises, au sein des secteurs industriels, entre les entreprises et le gouvernement, et dans les accords de collaboration en matière de données visant à relever les défis humanitaires et de développement. En l'absence de cadres clairs facilitant et régissant ces arrangements, il y a des risques que des utilisations positives des données soient manquées et que les abus de données continuent sans contrôle.

Les pays doivent identifier des modèles appropriés pour régir le partage des données impliquant le gouvernement et le partage des données dans l'ensemble de l'économie. Ceci nécessitera probablement une combinaison de vastes consultations nationales pour élaborer des approches légitimes, ainsi qu'un engagement international qui peut aider à aligner les approches en matière de partage de données impliquant des flux de données internationaux.

## Prendre en compte la collecte et la publication des données dans la législation et dans les politiques sectorielles

Le Baromètre a exploré un certain nombre de secteurs spécifiques dans lesquels les pays pourraient établir ou renforcer des règles exigeant la collecte et la publication de données structurées, notamment:

- Les registres de lobbying ;
- Les données sur l'exercice du droit à l'information ; et
- Les Registres des bénéficiaires effectifs.

Ce ne sont là cependant que quelques exemples des types de cadres juridiques sectoriels que les pays peuvent envisager d'établir. Le point sous-jacent exploré par le Baromètre est que, chaque fois que la législation ou la politique sont susceptibles d'impliquer la création de données, une attention explicite devrait être accordée aux règles régissant la manière dont les données doivent être fournies, y compris la manière dont les données seront vérifiées, protégées contre la vie privée et les données non privées mises à disposition dans le cadre de licences ouvertes.

Les prochaines éditions du Baromètre couvriront d'autres règles de gouvernance sectorielle. Les gouvernements devraient chercher à déterminer comment une bonne gestion des données et des pratiques en matière de données ouvertes peuvent être intégrées dans toute réforme sectorielle en cours, et les organisations de la société civile devraient renforcer leur capacité d'examiner de près l'impact de chaque réforme sur la production et le partage des données.

## Capacités

Pour que les données profitent au bien public, les pays nécessitent un large éventail de ressources, allant des fondements largement accessibles de l'accès à l'Internet et des infrastructures de données aux compétences numériques de base dans l'ensemble d'une population, en passant par les institutions appuyant la réalisation des droits relatifs aux données et la bonne gestion des données, et la disponibilité d'outils d'analyse et de compétences de pointe dans les administrations publiques, le secteur privé et la société civile. Dans le contexte du Baromètre, la capacité implique non seulement la présence de ressources numériques et l'accès à la connaissance des données, mais aussi les libertés et les possibilités de compétences et de ressources à mettre en œuvre au service d'objectifs sociaux plus larges.

Les paramètres de ce pilier du Baromètre reflètent à la fois les conditions de base dans un pays pour l'utilisation des données pour le bien public, et les points d'intervention particuliers qui peuvent être mis à la disposition des gouvernements, ou que les partenaires externes peuvent soutenir, qui peuvent contribuer à un environnement dans lequel les données sont gouvernées, mises à disposition et utilisées, pour relever les défis sociaux et atteindre les objectifs sociaux.

## Résumé

- Les fractures numériques en termes de connectivité Internet se réduisent. Si l'on ne s'efforce pas cependant aussi de réduire les lacunes en matière de maîtrise des données, de mettre en place des institutions efficaces capables de réglementer les données et de favoriser un large engagement à l'égard des données, il est à craindre qu'un plus grand nombre de personnes voient leurs données collectées et utilisées à des fins privées, sans avoir également la capacité d'utiliser les données pour promouvoir leurs propres intérêts et le bien public au sens large.
- Un grand nombre des ensembles de données les plus pertinents pour la vie quotidienne des communautés relèvent souvent de la responsabilité des gouvernements locaux. Cependant, seulement 22% des pays semblent disposer de données probantes d'une capacité soutenue et institutionnalisée à gérer efficacement les données au niveau infranational, avec peu de pays montrant des données probantes de politiques locales généralisées sur les données ouvertes (16 / 14.68%) et le partage de données (16 / 14.68%). Les 47 pays ont cependant des points lumineux parmi les États ou les municipalités démontrant des capacités de données infranationales raisonnables, mettant en évidence les opportunités d'apprentissage entre pairs entre les États, les villes et les municipalités, ainsi que la mise en évidence des retours potentiels sur les investissements dans les capacités locales.
- Le niveau de formation dispensée par les pouvoirs publics dans le but de développer les connaissances et les compétences des fonctionnaires en matière de données reste faible. Seuls 23 des pays disposaient de données probantes d'une formation étendue et régulière des fonctionnaires sur les questions de données dans le cadre d'une stratégie planifiée et durable, bien que la majorité des pays disposent au moins d'une certaine formation. Il existe d'importantes possibilités d'élargir l'éventail des sujets couverts par la formation liée aux données et à la portée de la formation dispensée dans l'ensemble des organismes gouvernementaux.
- Les capacités varient considérablement d'un pays et d'une région à l'autre. Le pilier de la capacité du baromètre a la plus grande plage entre les scores les plus élevés et les plus

les Caraïbes, par exemple, les opportunités créées par un capital humain et des libertés politiques plus élevés semblent entravées par des niveaux inférieurs de compétences numériques au sein de la population dans son ensemble, tandis qu'au Moyen-Orient et en Afrique du Nord, les capacités des entreprises et des gouvernements à utiliser les données sont fortes, mais les libertés de la société civile demeurent faibles.

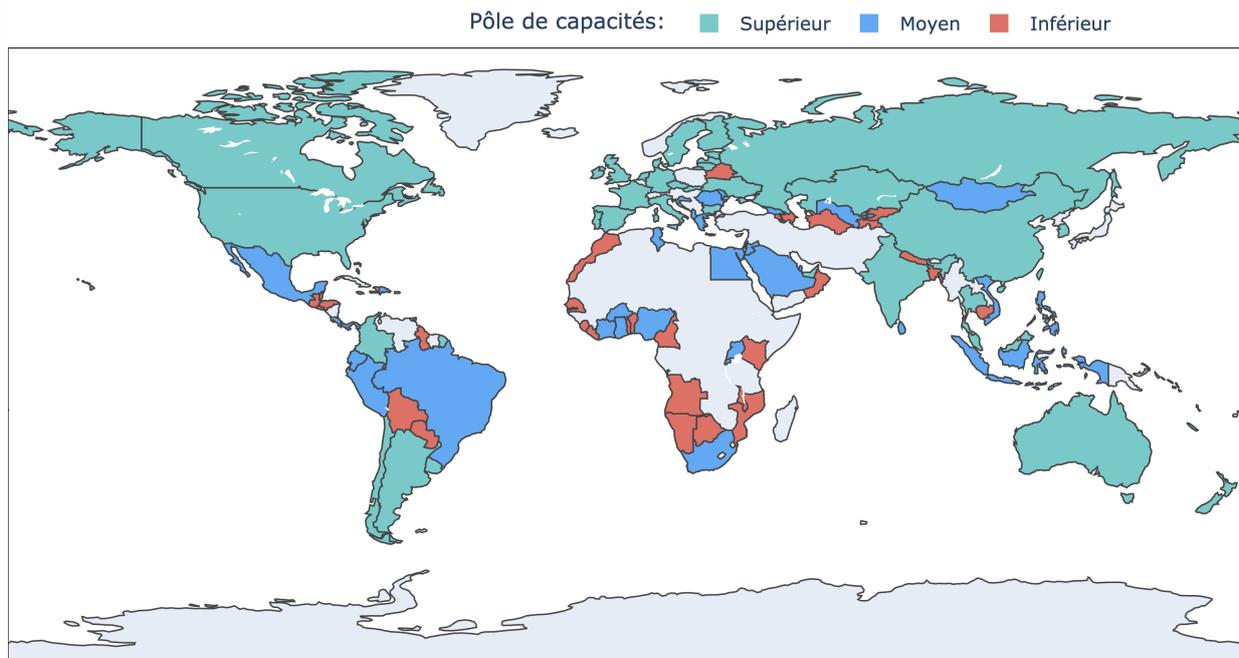
## Note méthodologique

La composante de capacité du baromètre est construite depuis une combinaison de dix sources de données secondaires et de quatre indicateurs primaires. Les indicateurs primaires visent à combler des lacunes particulières parmi les connaissances sur la présence d'initiatives de renforcement des capacités soutenues par le gouvernement pour les fonctionnaires, autour des données ouvertes et de la réutilisation des données, ainsi que des capacités de gestion des données au niveau infranational. Vu que notre principale méthode d'enquête repose sur des données probantes publiques de renforcement des capacités, de formation ou de soutien pour la réutilisation des données, il est possible qu'elle n'ait pas identifié de cas non documentés de prestation de formation ou de renforcement des capacités au sein des institutions gouvernementales.

En élaborant cette composante de capacité, nous avons fait face à un certain nombre d'autres lacunes importantes au sein des données qui ont limité notre capacité à générer une vision comparative robuste des capacités d'utilisation des données de la société civile, des médias, du milieu universitaire et du gouvernement. Le [cadre des indicateurs de performance statistique de la Banque mondiale<sup>16\]</sup>](#) décrit certains des principaux problèmes de données.. Nous avons également identifié des limites dans les mesures existantes des compétences numériques générales, notamment la [couverture limitée des pays et la comparabilité des rapports par rapport aux composantes à forte intensité de données de l'ODD 4.4.1.](#) Dans un certain nombre de cas tel que l'évaluation des capacités de données des médias ou de la société civile, nous n'avons pas pu identifier un indicateur primaire pouvant générer des données suffisamment solides dans le cadre des contraintes de notre méthode d'enquête. En raison de ces limites, le volet capacités de cette première édition du Baromètre est fortement axé sur l'évaluation des capacités gouvernementales, et des travaux seront requis au sein des prochaines éditions du Baromètre de manière à élaborer un ensemble plus global d'indicateurs.

## Comment la capacité varie-t-elle de par le monde ? Pays à capacités plus élevées, moyennes et plus faibles

En prenant le score global de la composante de capacité du Baromètre (une moyenne pondérée des indicateurs de capacité), nous divisons les pays en trois groupes de taille égale, désignés ici comme en tant que pays de capacité « inférieure », « moyenne » et « supérieure ». Nous utilisons ces groupes tout au long du rapport afin d'identifier les actions potentielles pour les pays sur la base de leurs points de départ relatifs et de leurs niveaux de capacité. Les données recueillies pour ce volet peuvent également être utilisées pour explorer les relations entre la gouvernance, la capacité, la disponibilité et l'utilisation des données, en explorant les questions relatives aux types d'intervention favorisant le mieux les données pour le bien public.



**Affectation par pays aux pôles de capacités** Les pôles sont générées à 33 % et 66 % de coupures de pourcentage dans le score de la composante de capacité pondérée, ce qui donne trois groupes de taille égale.

| Pays du groupe des capacités supérieures | Pays du groupe de capacités moyennes | Pays du groupe de capacité inférieur |
|--|--------------------------------------|--------------------------------------|
| Allemagne,                               | Afrique du Sud,                      | Angola,                              |
| Argentine,                               | Albanie,                             | Arménie,                             |
| Australie,                               | Arabie Saoudite,                     | Azerbaïdjan,                         |
| Bulgarie,                                | Brésil,                              | Bahamas,                             |
| Canada,                                  | Burkina Faso,                        | Bahrain,                             |
| Chili,                                   | Costa Rica,                          | Bangladesh,                          |
| Chine,                                   | Croatie,                             | Biélorussie,                         |
| Colombie,                                | Côte d'Ivoire,                       | Bolivie,                             |
| Danemark,                                | Philippines,                         | Botswana,                            |
| EAU,                                     | Ghana,                               | Bélize,                              |
| Espagne,                                 | Grèce,                               | Bénin,                               |
| Estonie,                                 | Géorgie,                             | Cambodge,                            |
| Finlande,                                | Indonésie,                           | Cameroun,                            |
| Hong Kong,                               | Jordanie,                            | Gambie,                              |
| Inde,                                    | Kosovo,                              | Guatémala,                           |
| Irlande,                                 | Malte,                               | Guyane,                              |
| Israël,                                  | Mexique,                             | Haïti,                               |
| Italie,                                  | Moldavie,                            | Honduras,                            |
| Kazakstan,                               | Mongolie,                            | Jamaïque,                            |
| Lettonie,                                | Nigéria,                             | Kenya,                               |
| Lituanie,                                | Nouvelle-Calédonie,                  | Lesotho,                             |
| Malaisie,                                | Ouganda,                             | Libéria,                             |
| Nouvelle-Zélande,                        | Ouzbékistan,                         | Malawi,                              |
| Pays-Bas,                                | Panama,                              | Maroc,                               |
| Portugal,                                | Pérou,                               | Mozambique,                          |
| Ranska,                                  | Qatar,                               | Namibie,                             |
| Royaume-Uni,                             | Roumanie,                            | Népal,                               |
| Russie,                                  | Rwanda,                              | Oman,                                |
| République Tchèque,                      | République dominicaine,              | Paraguay,                            |
| République de Corée,                     | Sainte-Lucie,                        | Salvador,                            |
| Slovaquie,                               | Sri Lanka,                           | Sierra Leone,                        |
| Suède,                                   | Tunisie,                             | Sénégal,                             |
| Taiwan,                                  | Vietnam,                             | Tadjikistan,                         |
| Thaïlande,                               | Égypte,                              | Togo,                                |
| Ukraine,                                 | Équateur,                            | Trinité-et-Tobago,                   |
| Uruguay,                                 | État de la Palestine                 | Turkménistan                         |
| États-Unis                               |                                      |                                      |

Les pays ayant les capacités les plus élevées affichent des niveaux élevés de connectivité à Internet et de capital humain abordables (mesurés par l'indice du capital humain de l'Enquête des Nations Unies sur le gouvernement électronique), ainsi que des données probantes claires du soutien du gouvernement à la pratique du numérique et des données, et des entreprises qui utilisent des outils numériques dans leurs opérations. Alors que la plupart des pays du groupe de plus grande capacité obtiennent de bons résultats en ce qui concerne les mesures de la liberté politique, le groupe comprend également un certain nombre de valeurs aberrantes (China, Kazakhstan, Russian Federation and United Arab Emirates) avec des niveaux inférieurs de liberté politique en vertu des classements de la Freedom House. Même au sein du groupe de capacités élevées, des variations importantes existent dans la mesure dans laquelle le gouvernement offre de la formation pour développer la formation et les compétences des fonctionnaires, et dans la mesure dans laquelle les gouvernements soutiennent la

réutilisation des données. Même les pays dotés de capacités élevées doivent pour la plupart développer des infrastructures d'interopérabilité robustes pour les données publiques, comme en témoignent les faibles scores obtenus sur l'indicateur d'interopérabilité de l'intégrité politique.

Dans les pays à capacité modérée, le gouvernement est moins susceptible d'avoir adopté des normes et des pratiques solides dans la production de statistiques officielles, et les pays obtiennent des résultats inférieurs à l'indice des services en ligne du gouvernement des Nations Unies. Les 25 pays de ce groupe ont cependant encore une certaine forme d'initiative quant aux données ouvertes, bien que beaucoup moins susceptibles d'avoir été évalués dans notre enquête comme ayant une capacité efficace de gestion des données au niveau des villes, des régions ou des collectivités locales. Les gouvernements de ce groupe sont également moins susceptibles de fournir un soutien actuel pour la réutilisation des données que ceux du groupe ayant les capacités les plus élevées.

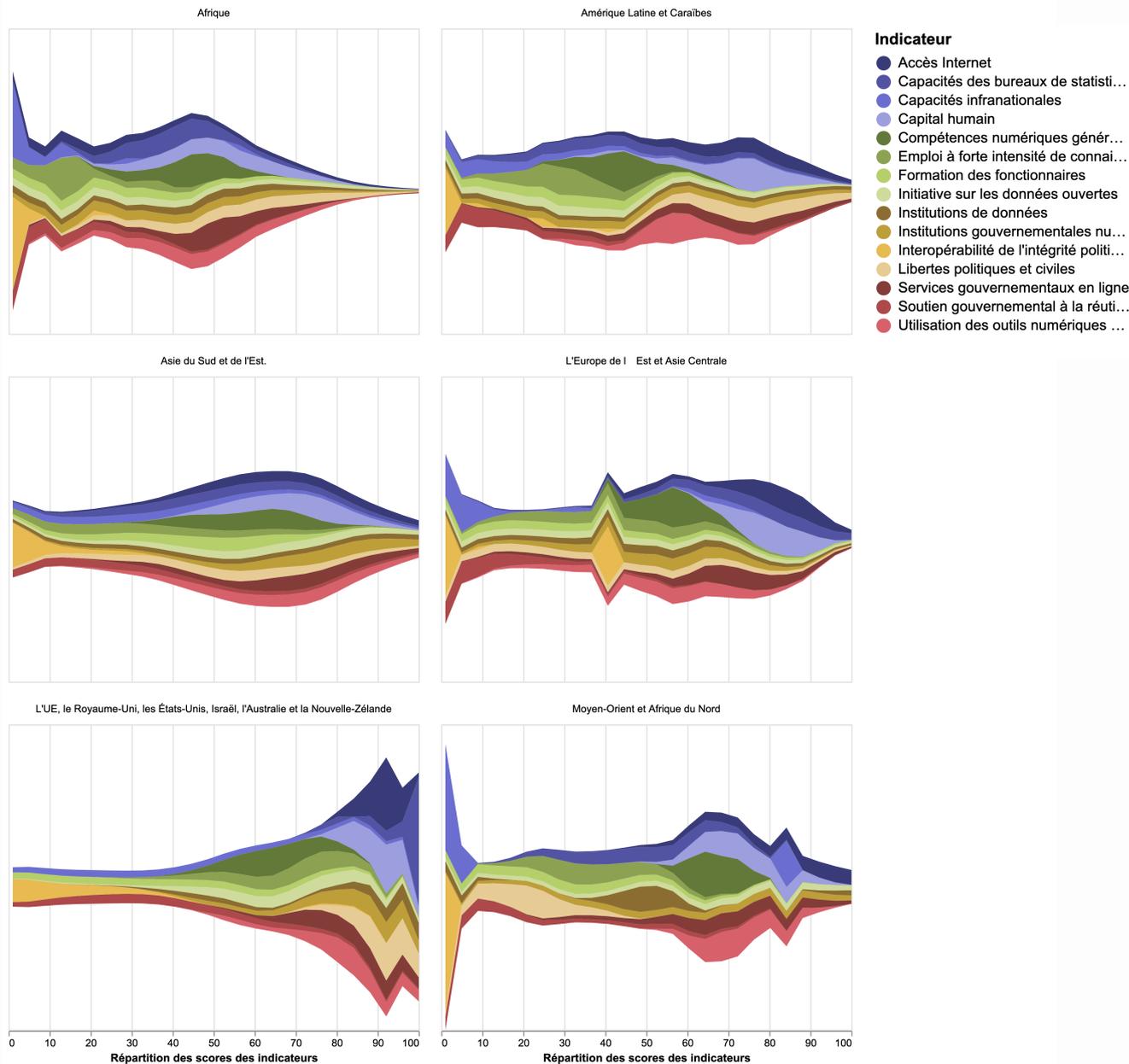
Le groupe de capacités inférieur comprend de nombreux pays où les institutions de données manquent ou sont faibles, où les scores de capital humain sont inférieurs, où moins d'entreprises se servent présentement d'outils numériques et où moins de personnes sont employées dans des industries à forte intensité de connaissances. L'accès à l'Internet est aussi souvent plus coûteux, et moins étendu, au sein de ce groupe. Ces pays peuvent avoir des difficultés particulières à fournir et à utiliser des données pour le bien public. Au cours de la période à l'étude, 11 des pays de ce groupe avaient cependant une initiative sur les données ouvertes, et un certain nombre d'entre eux avaient des données probantes de renforcement des capacités, plus particulièrement en ce qui a trait aux données statistiques.

## Quels facteurs façonnent la capacité dans chaque région ?

Le graphique de densité ci-dessous indique la distribution des scores non pondérés pour chaque indicateur de capacité de baromètre primaire et secondaire par région. Une bande plus large à la gauche de chaque graphique montrerait que les scores de cet indicateur pour cette région penchent vers l'extrémité inférieure, tandis que des bandes plus larges à droite indiquent que les scores penchent plus haut. Si la bande de couleur pour un indicateur particulier est relativement uniforme sur l'ensemble du graphique, elle indique une dispersion des valeurs pour les pays de la région.

Répartition des scores des indicateurs de capacité par région

Barometer Region



**Diagramme de densité montrant la répartition des indicateurs de capacité par région:** les indicateurs primaires et secondaires sont présentés, fournissant une vue des forces et des faiblesses relatives de chaque région et de la façon dont chaque variable est répartie.

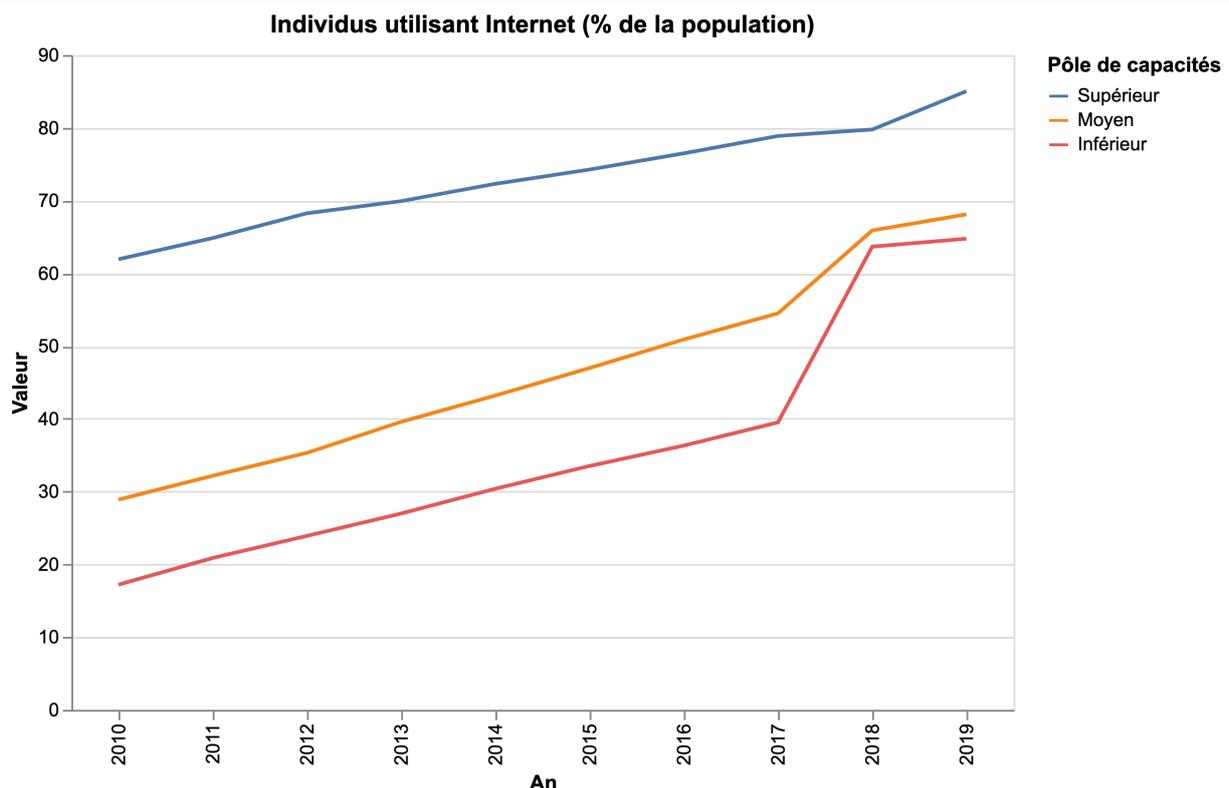
Les données qualitatives recueillies par le sondage du Baromètre suggèrent que la pandémie de COVID-19 a été un test puissant des capacités de données du gouvernement, de la société civile et du secteur privé. Dans un certain nombre de cas, les chercheurs rapportent des données probantes que les gouvernements infranationaux ont plus particulièrement démontré leur capacité à répondre rapidement aux demandes de données de la pandémie, assumant rapidement la responsabilité de la création et du maintien de nouveaux systèmes de données. Dans d'autres cas, la pandémie a toutefois mis en évidence des faiblesses dans la capacité de l'État ou de la société civile à utiliser les données pour répondre à une crise ou à contrôler la prise de décision d'urgence.

Les capacités infranationales ne sont pas en bonne place dans le graphique ci-dessus, toutes les régions voyant à la fois de faibles niveaux de capacités infranationales, et une large distribution des scores en ce qui concerne l'étendue des pratiques infranationales robustes autour des données ouvertes, de la gestion des données et du renforcement des capacités de la fonction publique, et aucune de nos variables secondaires choisies ne nous permettant d'explorer si les capacités identifiées sont concentrées dans les centres urbains, ou si elles sont plus uniformément réparties dans un pays.

Cependant, des données qualitatives mettent en évidence l'importance d'un équilibre entre les capacités nationales et infranationales pour soutenir l'inclusivité et l'innovation. Les chercheurs rapportent, par exemple, de la situation en Espagne, où une forte capacité de gestion des données dans les 17 régions autonomes n'est pas toujours assortie d'infrastructures ou de processus intégrant les données au niveau national : ce qui conduit à des ressources de données solides pour l'engagement au niveau régional, mais rend difficile pour les réutilisateurs d'obtenir une image nationale globale sur les questions clés. Pour les populations vivant près des frontières nationales, impliquées dans le commerce ou travaillant sur des questions transfrontalières, l'interopérabilité internationale des données et les pratiques partagées entre les pays deviennent aussi importantes.

## Comment les capacités de données sont-elles développées ?

Le Baromètre aborde les capacités via quatre dimensions : les fondations (couvrant l'accès à Internet de base, les compétences numériques et le capital humain) ; les capacités gouvernementales (pour gouverner, produire et partager des données) ; la capacité du secteur privé (à utiliser les données afin de générer de la valeur économique) ; et la capacité de la société civile et des médias (y compris les libertés politiques qui soutiennent une large utilisation des données pour le bien public, et la surveillance de la gouvernance des données par la société civile).



**Pourcentage de la population ayant accès à Internet (Source : UIT) par groupe de capacités au fil du temps:** convergence vers des niveaux élevés d'accès à Internet.

Au cours de la dernière décennie, l'accès à Internet s'est rapidement développé dans le monde entier, les paramètres de connectivité augmentant beaucoup plus rapidement que ceux des compétences numériques ou du capital humain. Même dans les pays de capacité moyenne et faible, la probabilité qu'un individu ait des aspects de sa vie numérisés, et aura au moins un point d'accès où ils pourraient s'engager avec les services numériques, a fortement augmenté. L'augmentation de la connectivité offre à la fois des opportunités et des menaces. Pour les gérer, il est essentiel d'établir un équilibre entre les

capacités gouvernementales, les libertés politiques pour soutenir les contrôles et les équilibres sur le gouvernement, et la capacité et l'engagement des autres parties, y compris les universités, le secteur privé, la société civile et les médias.

## Comment les capacités de données sont-elles développées ?

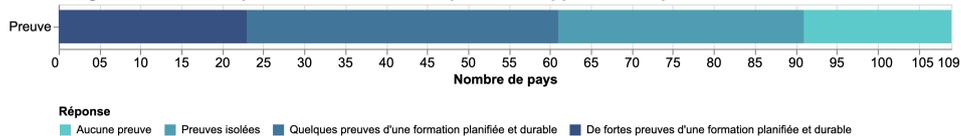
| -         | Services gouvernementaux en ligne | Stratégie de gouvernement numérique | Institutions de données | Normes en matière de statistiques |
|-----------|-----------------------------------|-------------------------------------|-------------------------|-----------------------------------|
| Supérieur | 84                                | 83.78                               | 81.76                   | 86.08                             |
| Moyen     | 65.59                             | 54.95                               | 53.47                   | 55.28                             |
| Inférieur | 49.64                             | 36.57                               | 22.92                   | 41.81                             |

Tableau : Notes moyennes par pôle pour certains indicateurs de capacité secondaires

Les gouvernements nationaux ont un rôle essentiel à jouer dans la gouvernance et dans la gestion des données, notamment en ce qui a trait à l'élaboration des stratégies en matière de données, la mise en place et le financement des principales institutions de gouvernance, tel que les autorités chargées de la protection des données, la fourniture de services numériques pour la collecte et l'accès aux données, et l'établissement et l'adoption de normes claires en matière de données. Les indicateurs de gouvernance secondaires du Baromètre montrent que les pays à faible capacité sont confrontés à des défis particuliers en ce qui concerne la présence d'institutions chargées de la gouvernance, de la gestion et de la protection des données, et de disposer à la fois d'infrastructures clés (plateformes gouvernementales en nuage) et de stratégies (y compris des stratégies technologiques et d'interopérabilité).

Le Baromètre explore la mesure dans laquelle il existe des données probantes que les gouvernements investissent dans le renforcement des capacités, en fournissant des programmes de formation aux fonctionnaires. Bien que plus de 90 pays aient des données probantes de l'existence d'une formation, moins de 25 % des pays paraissent avoir des programmes de renforcement des capacités en cours et durables à l'intention des fonctionnaires sur la maîtrise des données et les compétences en matière de données.

Dans quelle mesure le gouvernement dispense-t-il une formation pour développer les compétences des fonctionnaires en matière de données ?



**Fourniture par le gouvernement de formation à la maîtrise des données et aux compétences:** Bien que plus de 90 pays aient des données probantes de l'existence d'une formation, moins de 25 % des pays semblent avoir mis en place des programmes de renforcement des capacités généralisés, permanents et durables, à l'intention des fonctionnaires en matière d'éducation aux données et de compétences en matière de données.

Notamment, la disponibilité de la formation est particulièrement limitée à l'échelle infranationale. Même lorsque les pays disposent de programmes de formation, les chiffres communiqués sur le nombre de fonctionnaires formés indiquent que seul un très faible pourcentage de fonctionnaires ont accès à une formation ciblée quant aux compétences liées aux données. En Bulgarie, par exemple, les chercheurs notent que, selon les données de l'Institut d'administration publique, « au cours de la période 2019-2021, toutes les informations liées aux [données ont rassemblé 68 groupes et un total de 1102] stagiaires ». C'est là moins de deux personnes pour chacun des 570+ organismes exécutifs ou d'administration locale au sein du pays. Un portrait similaire est observé au Royaume-Uni, où le Data Science Campus britannique rapporte avoir formé 681 analystes aux outils de science des données en avril 2021. Des chercheurs gambiens ont noté que, bien que la formation relative aux données fournies par le gouvernement soit rare, le personnel technique des ministères nationaux a souvent accès aux compétences par l'entremise de programmes d'études, généralement à l'étranger. Dans l'ensemble, il y a un sentiment au sein des pays et des régions que la demande réelle ou potentielle de renforcement des capacités excède largement l'offre.

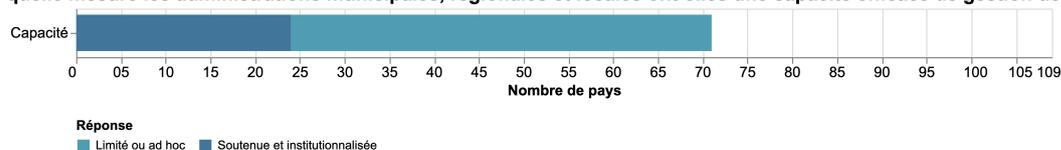
L'intensification de l'initiative de renforcement des capacités liées aux données demande un leadership et une stratégie. La République de Corée est l'un des rares pays à avoir mis systématiquement l'accent sur la formation en vertu de l'article 25 (Éducation et formation aux données publiques) de la loi sur la promotion de la fourniture et de l'utilisation de données publiques, qui fait obligation au ministre de l'Intérieur et de la Sécurité de formuler des politiques d'éducation et de formation à l'intention des fonctionnaires concernant les données publiques. En revanche, le chercheur suédois a noté que les stratégies nationales en matière de données ne contiennent pas de mention de la maîtrise des données et de la formation aux compétences en matière de données, et une cartographie des cours d'introduction dans 193 autorités gouvernementales n'a trouvé aucune preuve d'éducation liée à la connaissance et aux compétences numériques.

Pour les pays à faible capacité, les donateurs ont souvent été un catalyseur essentiel du renforcement des capacités. Des chercheurs du Kosovo signalent, par exemple, que des apports importants en matière de formation ont été fournis lors du lancement de l'initiative de données ouvertes du pays, faisant souvent appel à l'appui d'ONG pour dispenser la formation. Cette formation n'a toutefois pas été maintenue ni systématisée au fil du temps.

## Comment les capacités infranationales se développent-elles?

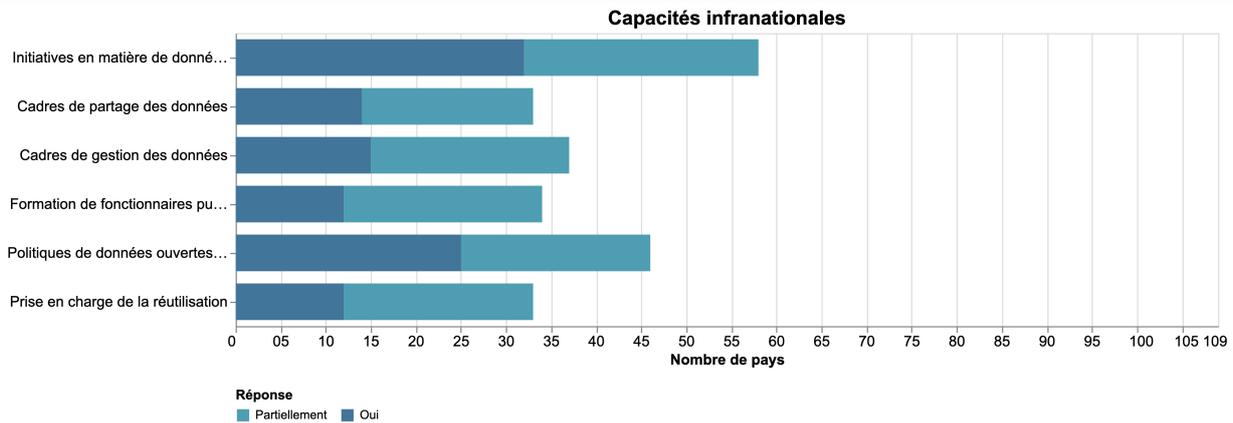
Bien que 71.6% des pays couverts par le Baromètre disposent d'une certaine forme de politique ou de législation nationale en matière de gestion des données, nous n'avons trouvé de données probantes d'une capacité de gestion des données soutenue et institutionnalisée que dans 22 moins de % des pays.

**Dans quelle mesure les administrations municipales, régionales et locales ont-elles une capacité efficace de gestion des données ?**



**Capacités infranationales:** Moins de 25 % des pays peuvent démontrer une capacité de gestion des données durable et institutionnalisée au niveau infranational

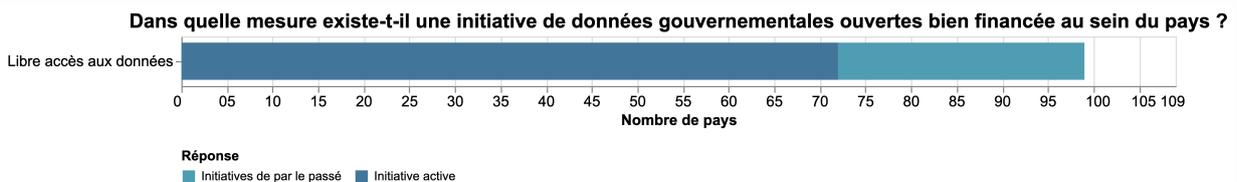
Les capacités infranationales identifiées par notre enquête centrée sur la présence de politiques et d'initiatives de données ouvertes, avec moins de données probantes que les gouvernements infranationaux fournissent, de façon systématique, une formation aux fonctionnaires, ou établissent des cadres clairs de partage de données locales ou de gestion des données. Il s'agit là d'une lacune particulière, vu plus particulièrement l'intérêt croissant porté au rôle que peuvent jouer les données du secteur privé dans la prise de décisions au niveau des villes et des États. Tel que travaux récents de<sup>[17]</sup> le soulignent, les gouvernements locaux font face à des obstacles importants et à des déséquilibres de pouvoir lorsqu'ils négocient l'accès aux données du secteur privé, et sans effort pour développer la capacité et la coordination infranationales pour s'engager dans le partage de données d'entreprise à gouvernement, il y a des risques que les gouvernements locaux ne soient pas en mesure de garantir le meilleur bien public des accords de partage de données.



**Domaines prioritaires des capacités infranationales:** Les capacités infranationales identifiées par notre enquête étaient centrées sur la présence de politiques et d’initiatives de données ouvertes, avec moins de données probantes que les gouvernements infranationaux ont des cadres clairs de partage ou de gestion des données.

## Comment se développent les initiatives de données ouvertes ?

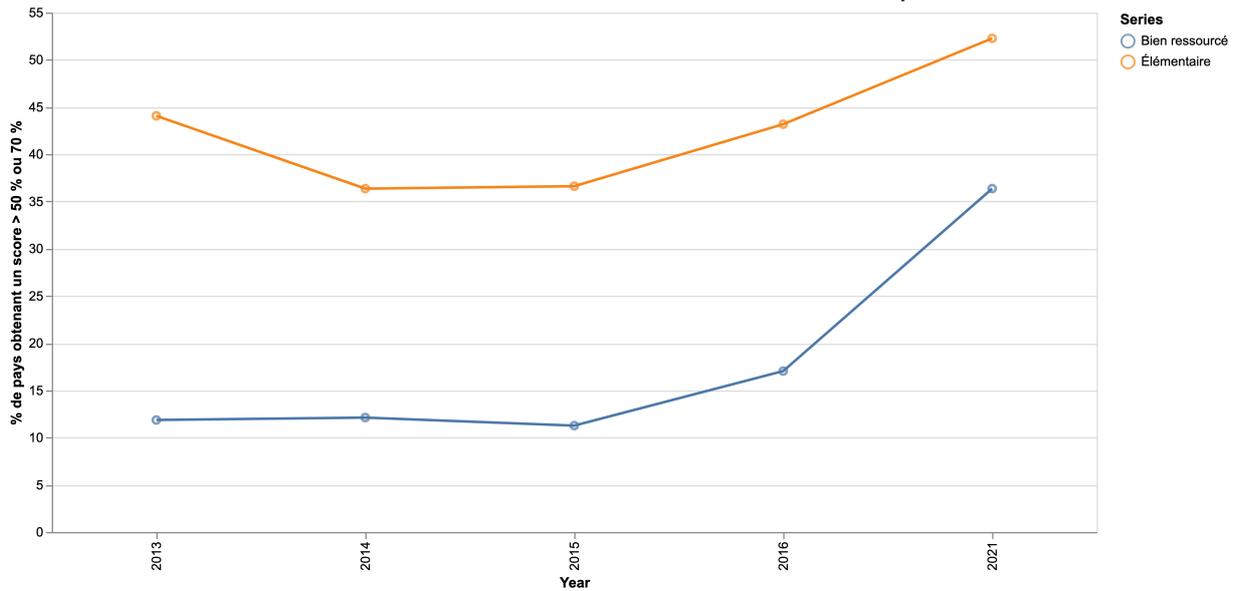
Près de 58 avec une certaine forme d’initiative locale sur les données ouvertes, près de 100 pays ont, à un moment donné, lancé une initiative nationale sur les données ouvertes. Notre enquête a cependant révélé que seules ces 72 initiatives restent actives, suggérant qu’au cours des dernières années, un certain nombre d’initiatives de données ouvertes ont été annulées ou sont devenues inactives.



Près de **100 pays ont lancé des initiatives** en matière de données ouvertes, cependant seulement 72 ne demeurent encore actifs.

Le Baromètre des données ouvertes suit la présence d’initiatives en matière de données ouvertes depuis 2013. En prenant les pays couverts à la fois par le Baromètre mondial des données et l’ODB, nous pouvons faire une comparaison approximative de la présence d’initiatives actives de données ouvertes (les pays obtenant plus de 5 sur 10 sur l’indicateur comparable de l’Initiative de données ouvertes des deux études), et le pourcentage d’initiatives qui semblent être bien dotées en ressources (les pays obtenant plus de 7 sur 10). Comme démontré par le graphique ci-dessous, le nombre d’initiatives actives ne semble avoir connu qu’une croissance modeste depuis 2016, et un certain nombre de pays ayant pris des initiatives de premier plan en 2016, tels que le Mexique et le Royaume-Uni, ont vu leur score d’indicateur considérablement réduit, les activités de données ouvertes ayant été déshéritées ou négligées.

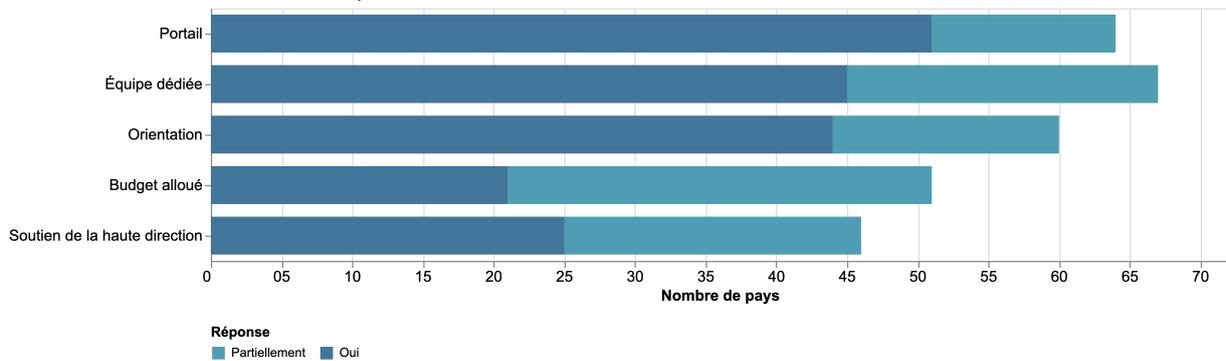
Présence et solidité des initiatives de données ouvertes dans le temps



**Force des initiatives en matière de données ouvertes:** Plus de 50 % des pays couverts à la fois par l’ODB et la GDB ont maintenant des initiatives actives en matière de données ouvertes, ce qui représente une croissance modeste par rapport à 2016. Parallèlement, le nombre d’initiatives bien dotées en ressources a considérablement augmenté depuis 2016.

Dans l’ensemble, toutefois, le nombre de pays disposant de ressources suffisantes s’est accru, ce qui invite à penser que là où les initiatives en matière de données ouvertes sont demeurées actives, elles sont devenues de plus en plus institutionnalisées, avec des équipes spécialisées et une infrastructure technique implantée. Un examen des sous-éléments de la question de l’initiative de données ouvertes du Baromètre mondial des données révèle que le manque de budget alloué et le soutien limité de la haute direction sont des faiblesses pressantes pour de nombreuses initiatives, représentant potentiellement les défis d’attirer l’attention sur les activités de données ouvertes lorsque l’attention politique s’est déplacée vers d’autres programmes connexes tels que l’intelligence artificielle et la gouvernance des données.

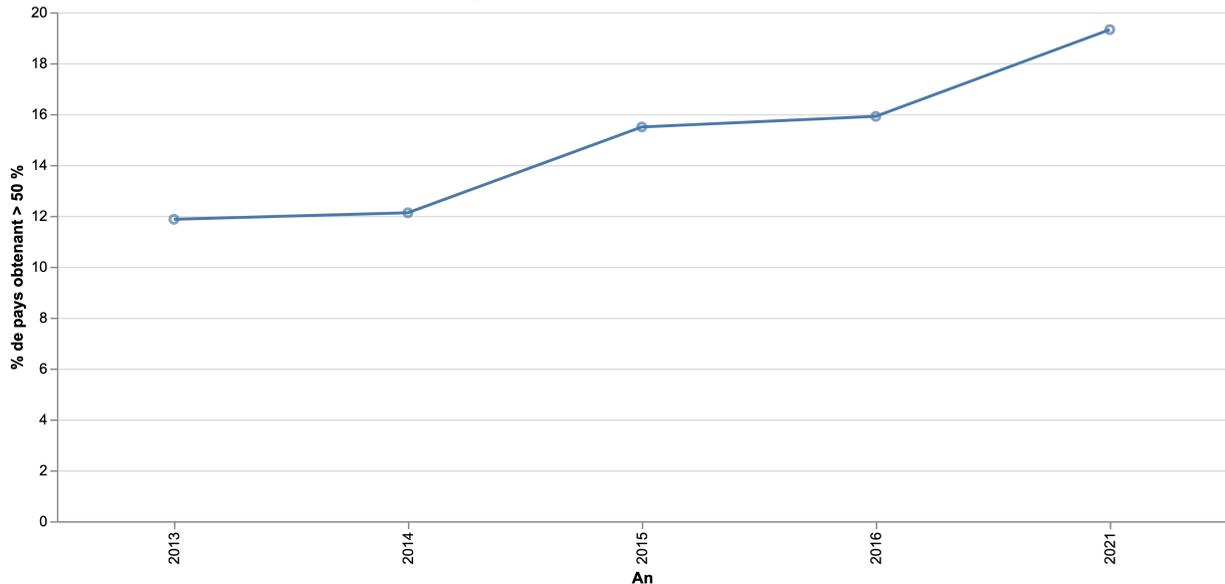
Quelles sont les fonctionnalités des initiatives actives de données ouvertes ?



**Les caractéristiques des initiatives actives de données ouvertes** ont généralement des équipes dédiées et une infrastructure technique sous forme de portails, mais manquent souvent de budget alloué et manquent de plus en plus de soutien clair de la part des hauts dirigeants du gouvernement

En comparant les indicateurs du Baromètre mondial des données et du Baromètre ouvert des données sur le soutien gouvernemental à la réutilisation des données, nous pouvons également voir une tendance positive, avec une croissance constante du nombre de pays obtenant des scores supérieurs à 5 sur cet indicateur au fil du temps, bien qu’une minorité de pays atteignant ce seuil.

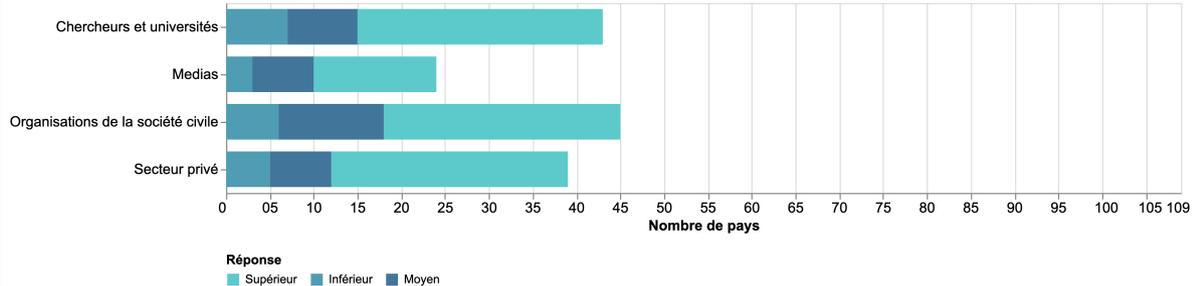
Prise en charge de la réutilisation des données comparée



**Soutien à la réutilisation des données:** Le nombre de pays fournissant un soutien à la réutilisation des données a légèrement augmenté au fil du temps

En somme, moins de 50 % des pouvoirs publics soutenant la réutilisation des données sur une base régulière, le soutien le plus important étant apporté à la société civile et aux universitaires, et le soutien le moins important est accordé aux médias.

Quels sont les secteurs visés par le soutien gouvernemental à la réutilisation des données ? (Par pôle de capacités)



**Les secteurs bénéficiant d'un soutien (par pôle de capacités):** Moins de 50 % du gouvernement fournit un soutien régulier à la réutilisation des données, avec le plus grand soutien à la société civile et aux universitaires, et le moins de soutien disponible pour les médias

## De quelles capacités plus larges les pays disposent-ils ou ont-ils besoin ?

La capacité ultime d'un pays à se servir des données pour le bien public n'est pas déterminée uniquement, ni même majoritairement, par la capacité technique directe, ou les interventions, des pouvoirs publics. La présence de compétences techniques, et la liberté de déployer celle-ci, dans l'ensemble de la population sont des éléments clés de l'utilisation des données envers le bien public.

Il existe peu de données comparables à l'échelle mondiale sur la prévalence des types de compétences numériques qui sont requises pour gouverner et utiliser les données pour le bien public (et, de fait, il existe des lacunes dans la recherche pour identifier l'éventail complet des compétences qui devraient être suivies). Nous nous sommes tournés vers un certain nombre d'indicateurs secondaires dans cette composante du Baromètre, bien que demeurant conscients que ceux-ci ne nous offrent pas de désagrégation par dimensions clés du genre ou du statut marginal. Pour obtenir des résultats satisfaisants pour le public depuis les données, il faut non seulement que les pays disposent de citoyens qualifiés libres d'examiner ou d'utiliser efficacement les données, mais que ces compétences soient largement réparties entre l'âge, le sexe, l'origine ethnique, la géographie et le groupe social.

| -         | Utilisation des outils numériques par les entreprises | Emploi à forte intensité de connaissances | Capital Humain | Libertés politiques et civiles | Compétences numériques au sein de la population |
|-----------|---|---|----------------|--------------------------------|---|
| Supérieur | 79.92   | 60.89                                     | 86.93          | 78.16                          | 64.36   |
| Moyen     | 58.96   | 34.36                                     | 71.21          | 58.74                          | 51.71   |
| Inférieur | 49.38   | 21.86                                     | 61.02          | 49.14                          | 45.94   |

Tableau : Notes moyennes par pôle pour certains indicateurs de capacité secondaires

## Comment les pays peuvent-ils améliorer leurs capacités en matière de données ?

Le site Web du Baromètre mondial de données contient des profils de pays qui indiquent les scores pour chaque indicateur et une ventilation de la manière dont ceux-ci ont été générés. Ceux-ci peuvent indiquer des domaines particuliers d'amélioration sur lesquels chaque pays à se concentrer, et l'examen des profils d'autres pays peut mettre l'accent sur des exemples à apprendre. En s'appuyant sur les données quantitatives et qualitatives recueillies dans le cadre du pilier des Capacités, nous avons identifié trois domaines d'action intersectoriels :

### Le soutien et l'institutionnalisation du renforcement des capacités

Les pays à faible capacité doivent passer d'interventions ponctuelles ou pilotes de renforcement des capacités à une prestation plus soutenue de formation et d'appui centré autour de la collecte, de la gouvernance, de la fourniture et de l'utilisation des données. Ceci implique des programmes dans les secteurs public et privé, ainsi que l'appui à l'institutionnalisation d'un soutien indépendant au renforcement des capacités des médias et de la société civile. Le renforcement des capacités doit porter à la fois sur les compétences techniques et essentielles en sciences sociales et humaines afin de travailler avec les données. Pour que les pays à faible capacité améliorent la gouvernance et la gestion des données, il faut également mettre l'accent sur le renforcement des institutions nationales qui supervisent la protection des données, les données ouvertes et la gestion de celles-ci, plus particulièrement dans le développement des capacités autour de l'utilisation de normes de données communes.

Les stratégies des donateurs doivent être axées sur l'intégration du renforcement des capacités en matière de données, plutôt que sur la fourniture de programmes à court terme exécutés à l'externe. Ceci peut entraîner des retours plus lents, cependant plus durables.

Une attention particulière devrait être accordée à la participation des groupes marginalisés en tant que partenaires dans la mise en œuvre des interventions de renforcement des capacités, de façon à promouvoir une meilleure compréhension de la manière dont les données affectent les différentes populations.

### Développer la stratégie et le leadership

Les pays à capacité moyenne doivent mettre l'accent sur le leadership et la stratégie qui fourniront des capacités accrues. Il devrait y avoir une direction politique de haut niveau non seulement pour développer le rôle des données dans l'économie, mais aussi pour s'assurer que le pays a les compétences, les institutions et les libertés nécessaires pour que les données soient gouvernées et utilisées pour le bien public. Les stratégies de formation et de renforcement des capacités devraient

fixer des objectifs clairs en vue d'accroître les compétences en matière de données de la fonction publique et de rendre la formation plus accessible à d'autres secteurs par l'intermédiaire des institutions établies.

Un certain nombre de pays à capacité moyenne ou plus élevée ont des capacités techniques relativement élevées dans les administrations publiques et les entreprises, mais leur environnement politique est plus limité, où il est plus difficile pour la société civile indépendante d'élaborer et d'appuyer des programmes sur la bonne gouvernance et l'utilisation des données. Dans ces pays, des efforts sont nécessaires afin d'accroître l'espace d'action indépendante et critique avec des données et pour développer des modèles de résolution de problèmes fondés sur des données qui permettent aux gouvernements, aux universités, au secteur privé et aux acteurs indépendants de travailler ensemble de manière fiable.

## Capacité de mise à l'échelle et d'intégration

Pour les pays dotés de capacités plus élevées, le défi principal à relever est de mettre à l'échelle et d'intégrer les activités de renforcement des capacités, plus particulièrement au niveau infranational. Des objectifs ambitieux devraient être fixés pour la prestation de la formation et pour l'inclusion des compétences techniques et critiques en matière de données dans les cadres du développement professionnel. Un leadership national renouvelé est nécessaire pour les activités relatives aux données ouvertes et pour faire en sorte que les pays puissent obtenir des résultats équilibrés en matière de biens publics grâce aux initiatives de partage de données avec le secteur privé.

Dans les pays à plus forte capacité, les innovations autour des données pour le bien public proviennent souvent du niveau infranational, avec des villes, des États ou des municipalités qui sont des « valeurs aberrantes » par rapport à la situation nationale. Lorsque c'est le cas, les pays devraient soutenir les efforts visant à partager les enseignements tirés de l'expérience de ces dirigeants locaux et aborder la question des cadres nationaux de normalisation et d'interopérabilité afin d'éviter la création ou l'approfondissement d'importantes divisions de données au sein du pays.

## Disponibilité

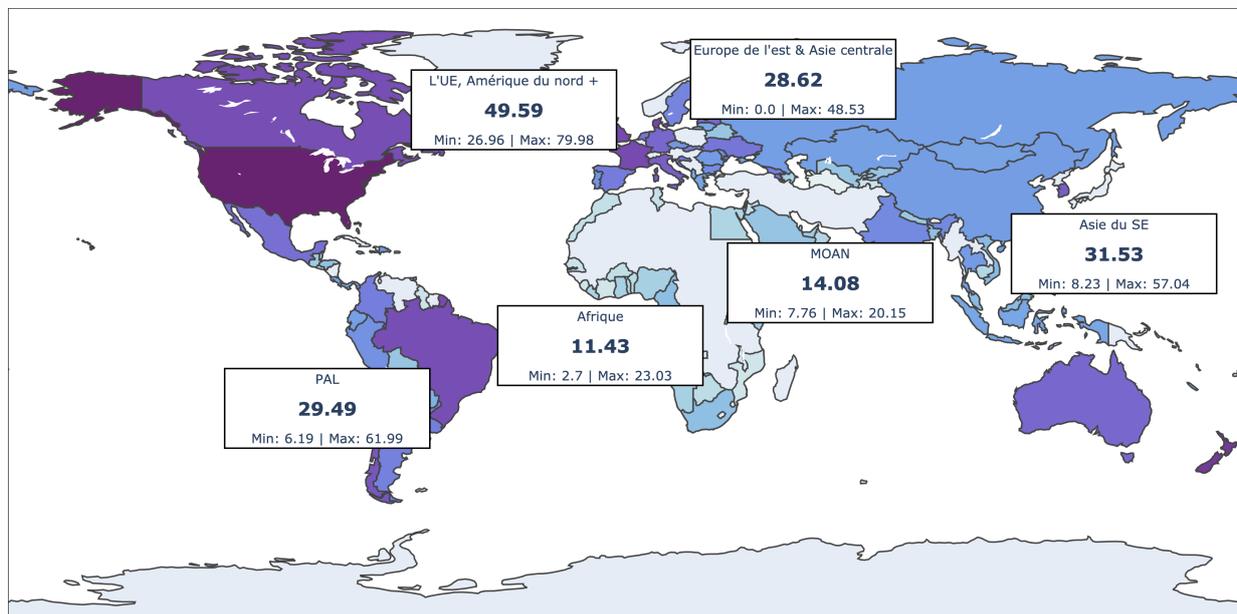
Les données ne sont pas toutes créées égales. La question de savoir si un pays est en mesure de tirer parti des données pour le bien public dépendra de la diffusion à grande échelle des principaux ensembles de données utiles pour résoudre les problèmes sociaux, politiques et environnementaux. Dans cette première édition du Baromètre, nous évaluons la disponibilité de 19 ensembles de données, sélectionnés en fonction de leur potentiel dans le but d'aborder des questions clés telles que le changement climatique, la santé publique, l'intégrité politique et les droits fonciers.

### Résumé

- Basé sur une définition stricte, la proportion d'ensembles de données publiés en tant que données ouvertes (gratuitement, sous des formes en vrac et lisibles à la machine, et sous licence ouverte) est demeurée relativement stable au cours de la dernière décennie, le 10.63 pourcentage des ensembles de données interrogés répondant à la définition ouverte. Ce chiffre n'est que légèrement supérieur au sommet de 10 % enregistré au baromètre des données ouvertes de 2015. En adoptant cependant une approche plus flexible de l'évaluation, nous constatons qu'un total de 17% d'ensembles de données qui soit répondent, ou se rapprochent, d'être fournis en tant que données ouvertes.
- La pandémie de COVID-19 a démontré l'importance du partage des données à l'échelle mondiale et à la capacité de la communauté mondiale à établir des infrastructures de données partagées, à la fois par des mesures officielles et via les efforts de collaboration de chercheurs indépendants, de citoyens scientifiques et de militants de données. Par l'utilisation de métriques secondaires, nous constatons que les données sont disponibles sur les tests COVID et la positivité pour 84.4% des pays et les données statistiques sur la capacité du système de santé sont disponibles pour 86.2%. Cependant, les mesures de la capacité du système de santé en temps réel sont disponibles dans seulement 46.8% des pays, et les points de données clés pouvant être utilisés pour l'évaluation des questions d'équité dans l'accès au système de santé et au déploiement de la vaccination COVID-19 sont souvent absents, même lorsque les données sont disponibles.
- D'importantes lacunes existent également en ce qui a trait à la disponibilité d'ensembles de données clés à l'appui des réponses nationales à la crise climatique. Bien que les statistiques agrégées soient largement communiquées aux systèmes internationaux, des données détaillées et ventilées sur les émissions, la biodiversité et la vulnérabilité climatique sont rarement disponibles comme données ouvertes à l'utilisation au sein du pays.
- Lors d'initiatives mondiales concertées, la disponibilité et la qualité des données paraissent relativement élevées. D'importants efforts ont par exemple été déployés à l'échelle mondiale au cours de la dernière décennie dans le cadre de forums, incluant le Partenariat pour un gouvernement ouvert pour promouvoir la transparence des dépenses et des achats et de renforcer les capacités en matière de données dans ce domaine. Notre enquête révèle que 106 les pays disposent d'au moins quelques informations sur le budget et les dépenses en ligne, et 100 disposent d'au moins un exemple de données sur les marchés publics fournies (bien que souvent provenant d'agences, d'États ou de municipalités particuliers, plutôt que d'ensembles de données nationaux intégrés).
- Dans les domaines d'intérêt émergent, tels que la propriété effective, l'intégrité politique et le régime foncier, le Baromètre offre une mesure de base de la disponibilité et de l'ouverture des données, trouvant moins de 50 % des pays avec des données disponibles par rapport à

ces indicateurs en moyenne, et moins de 10 % avec des ensembles de données ouverts. Les prochaines éditions du Baromètre seront en mesure de traquer les progrès accomplis par rapport à ces niveaux de référence.

- Les chiffres généraux sur la disponibilité et l'ouverture des données cachent d'importantes variations régionales. À titre d'exemple, plus de la moitié des ensembles de données ouvertes identifiés par le Baromètre provenaient de l'Union européenne, du Royaume-Uni, de l'Amérique du Nord, d'Israël, de l'Australie et de la Nouvelle-Zélande ; tandis qu'en Afrique, moins de 40 % des ensembles de données recherchés étaient disponibles en ligne sous quelque forme que ce soit.
- Il y a relativement peu de données probantes que, lorsque des ensembles de données clés ne sont pas disponibles auprès du gouvernement, que d'autres fournisseurs de données offrent des substituts adéquats. Dans les 84 829 cas où les gouvernements ne fournissaient pas d'ensembles de données, les chercheurs pouvaient localiser les données issues d'une autre source.



**Résultats régionaux du pilier Disponibilité:** Les scores les plus élevés sont observés dans l'UE, l'Amérique du Nord et le groupement régional, avec les scores les plus faibles en Afrique.

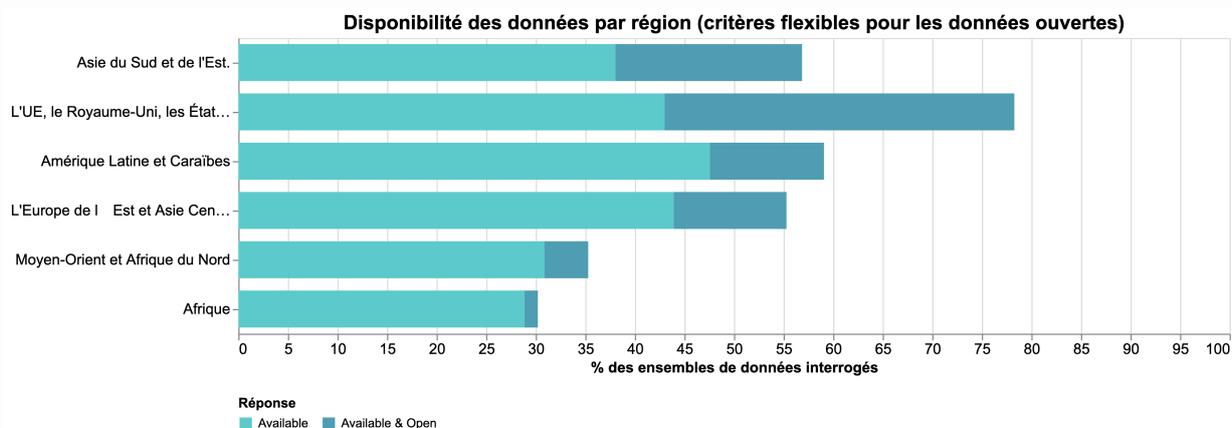
## À quel point les données couvertes par le Baromètre sont-elles ouvertes ?

Selon la définition ouverte<sup>[18]</sup>, un ensemble de données « est ouvert si quelqu'un est libre d'y accéder, de l'utiliser, de le modifier et de le partager — sous réserve, tout au plus, de mesures qui préservent la provenance et l'ouverture ». En pratique, ceci se traduit par le fait que les données soient disponibles sous forme numérique, lisible à la machine et non exclusive, soit gratuitement, soit à un coût ne dépassant pas le coût de la reproduction, et dans des conditions explicites permettant la réutilisation (plutôt que, par exemple, des conditions restrictives du droit d'auteur). En pratique, ceci se traduit par le fait que les données soient disponibles sous forme numérique, lisible à la machine et non exclusive, soit gratuitement, soit à un coût ne dépassant pas le coût de la reproduction, et dans des conditions explicites permettant la réutilisation (plutôt que, par exemple, des conditions restrictives du droit d'auteur). L'idée centrale de la définition ouverte c'est que, lorsque les obstacles techniques ou juridiques (explicites ou dus à l'incertitude) à la réutilisation sont supprimés, les données peuvent être utilisées de manière beaucoup plus vaste.

Les données barométriques nous offrent deux façons de calculer le nombre total d'ensembles de données ouvertes disponibles à travers notre échantillon. Une approche stricte, ne comptant que les ensembles de données qui sont disponibles auprès du gouvernement, et que les chercheurs ont évaluées comme ayant une bonne lisibilité à la machine, un accès en vrac, licenciée et étant disponibles gratuitement, révèle que les 197 ensembles de données (10.63%) répondent à la définition ouverte. Ceci constitue un progrès modeste par rapport au chiffre de 7 % des ensembles de données entièrement ouverts enregistrés par le Baromètre des données ouvertes en 2016, bien qu'il corresponde au nombre trouvé en 2014 et 2015 (les différences dans l'échantillon des deux pays et les ensembles de données entre les études signifient que ces comparaisons ne donnent qu'une indication approximative).

Si nous adoptons un modèle d'évaluation plus flexible et incluons des données qui présentent des faiblesses mineures lorsqu'elles sont évaluées par rapport à la définition ouverte (par exemple, une licence n'est pas explicitement indiquée à côté de l'ensemble de données, mais peut être identifiée avec des recherches supplémentaires ; ou les données peuvent être converties sous une forme lisible par machine, mais ne sont pas fournies à la source dans les structures ou formats idéaux. En d'autres termes, nous incluons des ensembles de données où les chercheurs du Baromètre répondent « partiellement » en réponse aux questions sur la lisibilité des machines, les données en vrac, la licence et le coût), nous trouvons un total de 315 ensembles de données (17%) qui pourraient être considérés, en fait, comme des données ouvertes.

Cette image principale cache des variations significatives entre régions et secteurs. Le graphique ci-dessous indique la proportion des ensembles de données que nos chercheurs ont trouvés à la fois en ligne et disponibles en tant que données ouvertes (soit en utilisant des critères flexibles, pour inclure les cas où des conditions telles que l'accès en bloc ou l'octroi de licences sont partiellement remplies), par région.



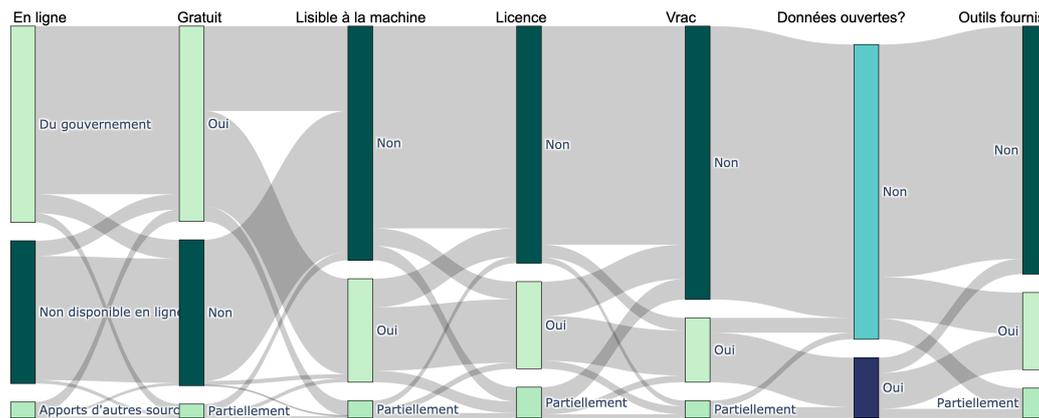
**Disponibilité des données par région:** la variation entre le nombre d'ensembles de données ouvertes d'une région à l'autre est beaucoup plus significative que la variation entre la disponibilité des données sous quelque forme que ce soit, ce qui met en évidence une double division des données – d'abord en termes d'accès à toutes les données, puis en termes d'accès pour la réutilisation des données.

Dans l'ensemble, plus de la moitié des ensembles de données 315 ouverts identifiés par le Baromètre se trouvent dans la région « Union européenne, Royaume-Uni, Amérique du Nord, Israël, Australie et Nouvelle-Zélande » du Baromètre, alors que seuls 11 des 374 ensembles de données vérifiées dans la région d'Afrique ont été considérés comme données ouvertes.

En 2016, le Baromètre des données ouvertes a conclu que le principal facteur freinant la diffusion des données ouvertes était l'absence de licences ouvertes<sup>[19]</sup>. Comme le montre le diagramme ci-dessous, en 2021, le plus grand facteur limitatif paraît être l'absence de fourniture de données en vrac.

Les évaluations de la disponibilité du Baromètre mondial de données comprennent aussi une variable demandant s'il existe des « outils officiels accessibles et ouverts disponibles afin d'aider les utilisateurs à explorer un ensemble de données ». Bien que la fourniture de données en vrac importe au soutien de

la réutilisation innovante des données, dans de nombreux cas, les utilisateurs veulent aussi pouvoir accéder directement aux faits ou aux informations depuis un ensemble de données sans nécessairement avoir à télécharger un ensemble de données complet ou complexe. Nous avons constaté que dans 46.4% des cas où les données sont disponibles en ligne, et dans la majorité des cas où les données ont été publiées en tant que données ouvertes, une certaine forme d'outil en ligne était disponible afin de les explorer. L'ensemble de données du Baromètre contient des liens vers tous les outils identifiés, offrant une opportunité future d'explorer les types d'interfaces qui pourraient être créés dans le but d'augmenter la gamme d'utilisateurs qui peuvent bénéficier de la fourniture de données structurées et ouvertes.



**Les obstacles à la disponibilité des données ouvertes:** Le diagramme de sankey montre l'évaluation globale de tous les ensembles de données par rapport aux critères de définition ouverts (en ligne, gratuit, lisible par machine, licence ouverte, données en vrac) ainsi que la présence d'outils accessibles pour explorer les données disponibles. Les réponses partielles sont présentées séparément, mais sont comptées pour répondre aux critères de définition ouverte (c'est-à-dire en utilisant une approche flexible de la définition des données ouvertes).

## D'autres organisations fournissent-elles des données alors que le gouvernement ne le fait pas ?

Pour chaque ensemble de données interrogé, si les chercheurs ne pouvaient pas trouver une source de données fournie par le gouvernement, ou en raison d'une action du gouvernement (p. ex. un mandat du gouvernement selon lequel un tiers devrait publier l'ensemble de données), nous leur avons demandé de déterminer si des données pertinentes étaient disponibles auprès d'une autre source. Dans 84 des 829 cas où les gouvernements ne fournissaient pas d'ensembles de données, les chercheurs pouvaient localiser les données issues d'une autre source. Ces autres sources comprenaient des plateformes universitaires, des sites Web de la société civile et des plateformes de données fournies par des donateurs ou des organisations multilatérales telles que la Banque mondiale ou l'Organisation mondiale de la santé. Dans un faible nombre de cas, plus particulièrement en ce qui a trait à l'information sur les entreprises, les fournisseurs de données du secteur privé ont été identifiés, mais ils n'offrent généralement que des données payantes plutôt qu'un accès libre.

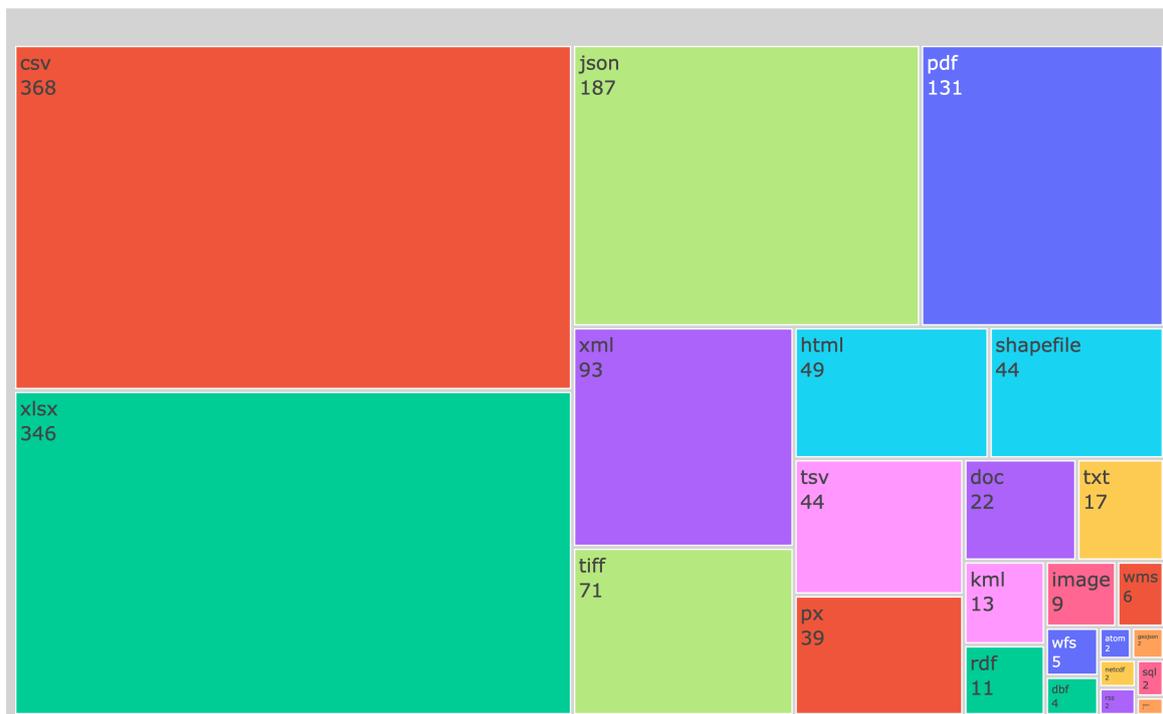
Dans l'ensemble, nous voyons peu de données probantes que d'autres sources de données offrent présentement des substituts efficaces lorsqu'il y a des lacunes dans les données fournies par les gouvernements, bien que pour les pays à faible capacité, les organisations internationales et les plateformes mondiales de la société civile ou de recherche puissent fournir des environnements d'hébergement et de publication de données utiles. Pour les ensembles de données fournis via ces

plateformes alternatives, en moyenne 45 % des champs ou fonctionnalités souhaitables que nous avons recherchés ont été fournis, par rapport à 58% des fonctionnalités lorsque les données ont été fournies par ou en raison d'une action gouvernementale.

## Les données disponibles sont-elles adaptées à l'objectif ?

Alors que le test véritable pour déterminer si les données sont adaptées à l'objectif est de savoir si celles-ci sont utilisées avec succès dans le but de résoudre un éventail de défis sociaux, le Baromètre offre un certain nombre d'indicateurs qui saisissent les aspects de la qualité des données. Nous constatons que 73.78% des ensembles de données en ligne ont été jugés raisonnablement « opportuns et mis à jour », et 61.34% des ensembles de données en ligne ont fourni un certain degré de données historiques qui permettraient aux utilisateurs de suivre les changements au fil du temps (par exemple, des dossiers sur l'utilisation passée des terres ou des informations sur le budget de l'année précédente et de l'année en cours). Ceci laisse encore près de 40 % des ensembles de données où seules les données courantes sont publiées, ce qui peut générer des défis pour un certain nombre de cas de responsabilité ou d'utilisation analytique.

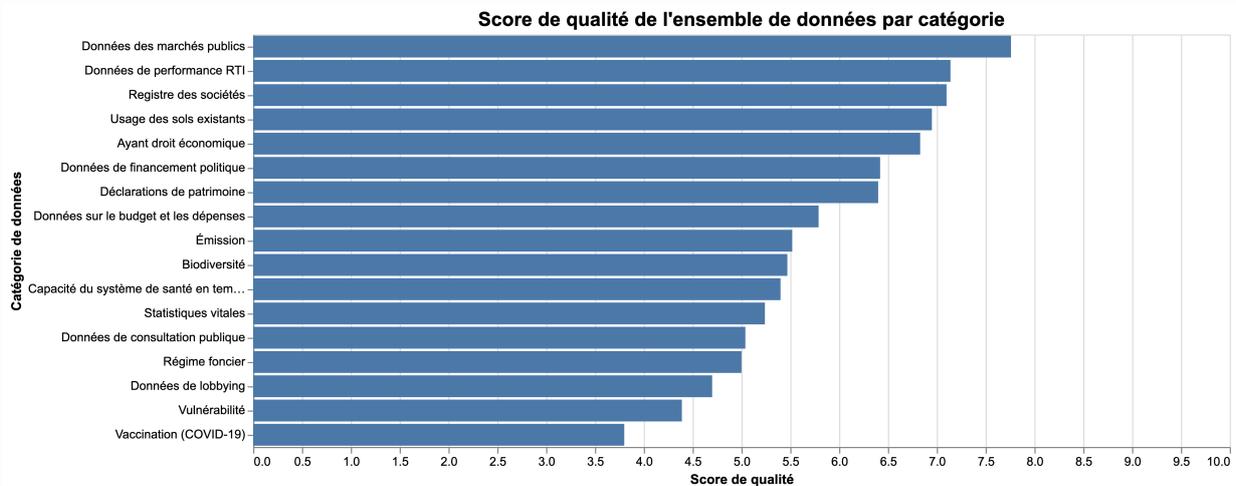
Le (s) format(s) de fichier dans lequel un ensemble de données est fourni peut avoir un impact sur sa convivialité. Les chercheurs ont fourni des rapports en texte libre des principaux formats de fichiers utilisés par les ensembles de données lisibles à la machine, que nous avons recodés pour y découvrir une préférence étroite parmi les données publiées pour le format CSV non-exclusif, et des données probantes croissantes de l'utilisation d'autres formats structurés tels que json et XML.



**Ce treemap des extensions de fichiers détectées dans les rapports en texte libre sur les formats de fichiers utilisés par les ensembles de données.** : CSV est le format de fichier le plus couramment rencontré, suivi des fichiers Excel (xlsx), json, pdf et xml

Chaque indicateur de disponibilité primaire au sein du baromètre comprend un certain nombre de sous-questions conçues pour vérifier la disponibilité de certains champs ou caractéristiques d'ensembles de données importantes aux cas d'utilisation d'intérêt public des données et qu'il est raisonnable de s'attendre à ce que les éditeurs de données fournissent. Bien que ces listes de champs

et de caractéristiques soient adaptées à chaque catégorie de données, et que certaines puissent être plus difficiles à utiliser que d'autres, en examinant le score moyen pondéré sur ces sous-questions pour chacune, nous pouvons identifier les ensembles de données les plus susceptibles d'être adaptés à l'objectif.



**Score de qualité de l'ensemble de données par catégorie** Les données d'approvisionnement affichent les scores les plus élevés pour la présence de champs et de caractéristiques-clés au sein des données pubs liés, tandis que les ensembles de données sur la santé et le climat sont moins susceptibles de contenir toute la gamme de caractéristiques vérifiées par l'enquête du Baromètre.

Les données sur l'approvisionnement viennent à la tête de ce tableau, bien que, comme indiqué ci-dessous, bon nombre des ensembles de données évalués pour cet indicateur soient des exemples infranationaux ou d'organismes uniques qui peuvent agir comme des valeurs aberrantes positives et qui ne sont pas toujours représentatives de la qualité des données disponibles dans l'ensemble du pays. Les ensembles de données sur la santé et le climat sont plus particulièrement parmi les plus susceptibles d'avoir des lacunes de qualité significatives lorsqu'ils sont mesurés à l'aide des indicateurs du Baromètre.

Un certain nombre de champs et de caractéristiques d'ensembles de données que nous avons vérifiés peuvent, si présentes, être utilisés pour se concentrer spécifiquement sur les modèles d'inclusion et d'exclusion. Ceci comprend les champs de données relatifs au sexe, au handicap ou à la situation au sein d'un groupe marginalisé ou de populations autochtones. En comparant le nombre de fois où ces champs ont été identifiés au nombre moyen de fois que des champs non liés à l'inclusion ont été identifiés, nous constatons que les champs liés à l'inclusion sont à près de 50 % moins susceptibles d'être fournis que d'autres types de champs. En d'autres termes, il y a d'importantes lacunes à combler dans la mesure dans laquelle les données disponibles appuient les mesures d'inclusion et de lutte envers les formes de discrimination ou de marginalisation.

## Quels ensembles de données sont disponibles ?

Dans cette première édition du Baromètre, nous avons révisé la disponibilité de 19 ensembles de données, organisés en un nombre de modules thématiques. Cette section présente un certain nombre d'observations sélectionnées pour chaque module, ainsi que des entrées de tableau récapitulatif pour chacun des ensembles de données.

## Note méthodologique

Les tableaux récapitulatifs présentent quatre valeurs pour chaque ensemble de données couvert par notre enquête principale.

- **Disponible en ligne**, représentant le nombre de pays dans lesquels certaines données répondant à la définition de l'indicateur ont été trouvées en ligne. Ceci peut comprendre des données disponibles uniquement au niveau infranational, ou des données présentant des limitations importantes (telles que présentées sous des formes non lisibles à la machine, ou sur des sites Web nécessitant un enregistrement ou un paiement)
- **Les données ouvertes** représentent le nombre de pays ayant au moins certaines données répondant à la définition ouverte. Nous adoptons notre définition flexible des données ouvertes, y compris les pays où les chercheurs ont répondu « partiellement » à une ou plusieurs des questions d'évaluation des données ouvertes.
- **Le score de qualité** représente une moyenne pondérée du nombre de champs ou de caractéristiques souhaitables pour l'ensemble de données qui ont été trouvés disponibles dans les données en ligne. Notez que les critères de qualité sont différents pour chaque catégorie de données.
- **Le score d'ouverture** représente une moyenne pondérée des questions d'ouverture standard posées pour chaque ensemble de données. Il s'agit notamment de questions sur la présence de données actuelles et historiques, ainsi que sur la disponibilité d'outils accessibles pour l'exploration des données. Les critères utilisés sont communs à chaque catégorie de données.

## Santé et la COVID-19

La pandémie de la COVID-19 a souligné l'importance de données de santé fiables, accessibles et dignes de confiance pour pouvoir permettre une action coordonnée. Alors que la gouvernance des données de santé est traitée par des indicateurs sur la présence de cadres de protection des données, dans le pilier de disponibilité du baromètre, nous avons recherché des données probante que les données sont disponibles pour soutenir la planification de la santé à long terme et pour permettre une réponse rapide spécifiquement dans le contexte de la pandémie de la COVID-19.

| Ensemble de données                        | Disponible en ligne | Données ouvertes | Score moyen de qualité | Score moyen d'ouverture |
|--|---------------------|------------------|------------------------|-------------------------|
| Vaccination (COVID-19)                     | 95                  | 31               | 3.8                    | 6.37                    |
| Capacité du système de santé *             | 94                  | s/o              | s/o                    | s/o                     |
| Données de test (COVID-19)*                | 92                  | s/o              | s/o                    | s/o                     |
| Statistiques vitales                       | 85                  | 29               | 5.24                   | 6.43                    |
| Capacité du système de santé en temps réel | 51                  | 17               | 5.4                    | 5.83                    |

\* Indicateurs basés sur des données secondaires qui ne nous permettent pas de calculer des scores de qualité ou d'ouverture.

Bien que les données statistiques rétrospectives sur les systèmes de santé et les données agrégées sur le dépistage de la COVID-19 soient largement disponibles, les données en temps réel sur la capacité du système de santé et les points de données pouvant soutenir l'action sur les inégalités de santé sont beaucoup moins susceptibles d'être fournis en ligne ou en formats ouverts.

Au moment de notre collecte de données, le projet Our World in Data avait réussi à rassembler des données sur les tests de la COVID provenant de plus de 100 pays, dont 92 sont couverts par le Baromètre. Bien qu'ils aient été en mesure de rendre les données recueillies depuis des sites Web nationaux du ministère de la Santé ou de sources régionales tels que les Centres africains de contrôle et de prévention des maladies, largement comparables, la documentation des différentes sources de données<sup>[20]</sup> met en évidence les défis importants liés à la localisation des données et à la recherche de documentation pouvant expliquer si les chiffres des tests se rapportent à des individus ou à des échantillons, et quels types de tests (par exemple, PCR ou antigène) sont couverts. Rassembler des données provenant de tant de pays en un si court laps de temps est une réalisation remarquable : mais elle demeure aussi plus complexe qu'elle ne devrait l'être, avec une marge de manœuvre importante pour renforcer les pratiques de publication des données à la source, y compris par une publication plus lisible par machine, et une plus grande fourniture et adoption de normes pour la publication des données<sup>[21]</sup>.

Les mêmes activités, telles que le suivi de la disponibilité des lits d'hôpitaux de soins réguliers et intensifs, peuvent, en fonction des infrastructures de données en place dans un pays, être utilisées afin de générer des données statistiques agrégées et fournir des informations en temps réel sur la capacité des services de santé. Dans certains pays, ces données en temps réel sont essentielles afin de permettre aux individus de savoir où ils peuvent accéder aux soins, et à des moments critiques, elles peuvent également être d'une grande valeur pour les journalistes ou les organisations de la société civile d'examiner attentivement la fourniture, la performance et l'équité des soins de santé. Toutefois, dans un certain nombre de pays, les données en temps réel ne sont pas collectées ou ne sont partagées qu'avec des organismes publics ou des professionnels de la santé. Si nous nous basons sur

les données de l'inventaire des données ouvertes (ODIN), nous constatons que 86.2% des pays Baromètres fournissent des données statistiques sur la capacité du système de santé, mais que seulement 46.8% ont des données en temps réel, ou presque en temps réel, disponibles, et de celles-ci, seuls les pays 33 fournissent des données au niveau de l'établissement (par exemple, hôpital ou clinique), et uniquement les 35 disposent d'informations sur la disponibilité des lits. Dans un certain nombre de pays, des données en temps réel ont été mises à disposition pour la première fois dans le contexte de la COVID-19, soulevant la question de savoir si les États vont cesser cette publication à l'avenir, ou mettre la fourniture d'informations sur la capacité de soins de santé sur une base plus stable.

Les sous-questions de qualité de notre indicateur d'enquête sur les taux de vaccination ont porté une attention particulière à l'équité de la distribution des vaccins, en demandant si les données disponibles étaient ventilées par âge, sexe, géographie, statut d'invalidité et situation quant aux soins de santé. Bien que 46.3% des ensembles de données de vaccination disponibles (67.7% de ceux publiés en tant que données ouvertes) contiennent une désagrégation par âge, juste 33.7% (51.6% des ensembles de données ouverts) ont été désagrégés par sexe ou genre, et seulement 13.7% (29% des ensembles de données ouverts) a fourni une ventilation par statut de handicap.

Comme le module santé du Baromètre est entièrement basé sur des indicateurs de disponibilité, nous pouvons utiliser les scores du module pour construire une image de la performance régionale comparative sur la disponibilité des données de santé. Le tableau ci-dessous montre les cinq premières questions dans chaque région Baromètre, ainsi que leurs scores de module.

| Afrique  | L'Europe de l'Est et l'Asie Centrale  | L'Union européenne, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande               | Amérique latine et Caraïbes  | Moyen-Orient et Afrique du Nord  | L'Amérique du Sud  |
|--|---|--|--|--|--|
| Ouganda (44.97), Burkina Faso (36.03), Kenya (34.03), Afrique du Sud (32.22), Rwanda (30.08) | Ukraine (55.15), Arménie (49.14), Moldavie (44.77), Mongolie (41.75), Lesotho (41.27) | États-Unis (85.4), Allemagne (83.02), Slovaquie (82.4), Nouvelle-Zélande (79.19), Italie (78.56) | Pérou (77.08), Brésil (73.22), Chili (73.21), Argentine (63.63), Uruguay (58.11) | EAU (41.58), Qatar (41.09), Arabie Saoudite (36.59), Oman (32.62), Bahrain (31.47) | République de Corée (75.62), Hong Kong (67.17), Taiwan (65.41), Bangladesh (55.69), Inde (55.06) |

### Action pour le climat

Comme clairement indiquée dans le programme 2030 des Nations unies, la lutte contre la crise climatique est un bien public convenu à l'échelle mondiale. Les changements climatiques et les mesures que les gouvernements et les citoyens peuvent prendre pour les atténuer et s'y adapter sont des questions d'une importance vitale à travers le monde. Les données peuvent s'avérer un outil puissant pour hiérarchiser et évaluer l'action climatique, ainsi qu'un moyen d'augmenter le nombre de groupes qui peuvent suivre et proposer une action climatique, prendre des mesures significatives dans leurs propres communautés et organisations, et tenir les gouvernements responsables de leurs actions ou de leur inaction.

Le Open Data Charter' s Open Up Guide to climate action<sup>[22]</sup> identifie une gamme d'ensembles de données pertinents pour le climat que les États pourraient publier, allant des statistiques sur les émissions aux indicateurs de biodiversité, en passant par les statistiques sur la production d'électricité et les données sur l'exposition aux risques climatiques. Le guide prend bien note que :

«les pays peuvent déjà être tenus de communiquer certaines de ces informations à un organisme régional ou international... il importe d'examiner comment les informations déjà disponibles ou communiquées pourraient être rendues plus accessibles et utiles à des tiers par leur publication sous forme de données ouvertes.»

Les indicateurs du Baromètre offrent un aperçu de la mesure dans laquelle les données sont disponibles au niveau national, en mettant en évidence certains points positifs et certaines lacunes importantes.

| Ensemble de données | Disponible en ligne | Données ouvertes | Score moyen de qualité | Score moyen d'ouverture |
|---------------------|---------------------|------------------|------------------------|-------------------------|
| Émission            | 73                  | 28               | 5.52                   | 5.81                    |
| Biodiversité        | 68                  | 16               | 5.47                   | 4.75                    |
| Vulnérabilité       | 58                  | 14               | 4.39                   | 5.49                    |

Bien que 73 les pays fournissent des données sur les émissions, et le 28 font en tant que données ouvertes, la façon dont ces données sont fournies, leur actualité et leur degré de détail varient considérablement. Dans de nombreux cas, les pays communiquent des données d'émission au titre de la Convention-cadre des Nations Unies sur les changements climatiques (CCNUCC), mais n'offrent pas l'accès à ces données aux citoyens nationaux. Ceci permet de croire que, lors de l'élaboration de normes internationales en matière d'établissement de rapports, il faudrait aussi envisager des normes ou des recommandations visant à rendre les données également disponibles au niveau national, en tirant le meilleur parti possible des activités de collecte de données.

Les lacunes en matière de données sont encore plus importantes lorsqu'il s'agit de données sur la vulnérabilité climatique. En dépit de l'importance des données localisées sur la vulnérabilité climatique pour soutenir l'adaptation, seulement 53.2% des pays interrogés ont offert des données pertinentes (14 sous forme d'ensembles de données ouverts), et les ensembles de données que nous avons pu localiser étaient particulièrement faibles en ce qui a trait à la fourniture d'informations sur la manière dont les vulnérabilités peuvent affecter différemment les populations marginalisées.

Le 62.4 pourcentage de pays fournissant des informations sur la biodiversité en ligne masque également des variations notables en termes de formes de vie représentées dans les données disponibles. Par exemple, 60.6% des pays fournissent des données sur les oiseaux en danger, et 58.7% couvrent les mammifères, mais seulement 37.6% fournissent des données sur les champignons et les lichens. Ce manque comparatif d'attention aux champignons est particulièrement remarquable compte tenu du rôle vital que ceux-ci jouent dans les écosystèmes du monde entier, tout en servant de puits de carbone mondial important.

Les premiers travaux menés par les partenaires du Baromètre à l'aide des données préliminaires du module d'action pour le climat ont porté attention sur l'importance d'utiliser les informations qualitatives de l'enquête pour comprendre les diverses pratiques de surveillance de la biodiversité, reconnaissant qu'il existe de nombreuses histoires et pratiques de collecte d'informations sur la biodiversité et les espèces menacées, des taxonomies scientifiques aux inventaires coloniaux, en passant par les pratiques de connaissance locales. Dans un point qui peut être adapté à de nombreux ensembles de données couverts par notre enquête, le chercheur Dr Jonathan Gray soutient que : *“Rendre les données sur la biodiversité plus largement disponibles peut non seulement servir à permettre leur*

*réutilisation pour communiquer l'état de la biodiversité à des publics plus larges, mais aussi pour faciliter la participation du public, l'apprentissage et la mise en évidence de différents types d'expertise locale et d'action pour comprendre et prendre soin de la vie que les données dépeignent.*" (Source : Atelier basé sur les résultats préliminaires du module d'Action pour le climat du Baromètre mondial de données au Festival Mozilla le 7 mars intitulé « GDB as a collective learning device to explore climate action ».)

## Intégrité politique

Le bien public peut être le mieux servi lorsqu'une sphère publique ouverte, responsable et équitable existe dans laquelle l'argent ne fausse pas la prise de décision équitable ou l'accès à la fonction politique. Les données peuvent être un outil puissant à identifier les intérêts influenceurs sur la façon dont les décisions en matière de gouvernance sont prises et mises en œuvre. Au sein des systèmes politiques démocratiques, ceci implique la transparence des finances des partis politiques, l'information sur les intérêts des décideurs politiques, sur les interventions des lobbyistes et sur les processus de consultation publique dans l'élaboration des règlements, ainsi qu'un solide système d'accès à l'information qui aide les membres du public à évaluer et à demander des comptes aux personnes au pouvoir.

L'accent que nous mettons sur les données relatives à l'intégrité politique est conforme à l'ODD 16 : Paix, justice et institutions fortes, en particulier ses cibles concernant l'état de droit (16.3) ; les institutions transparentes et responsables (16.6) ; la prise de décision réactive, inclusive, participative et représentative (16.7) ; et l'accès du public à l'information (16.10).

| Ensemble de données              | Disponible en ligne | Données ouvertes | Score moyen de qualité | Score moyen d'ouverture |
|----------------------------------|---------------------|------------------|------------------------|-------------------------|
| Données de consultation publique | 62                  | 8                | 5.04                   | 4.15                    |
| Données de financement politique | 57                  | 10               | 6.42                   | 5.33                    |
| Données de performance RTI       | 43                  | 8                | 7.14                   | 5.62                    |
| Déclarations de patrimoine       | 50                  | 4                | 6.4                    | 4.73                    |
| Données de lobbying              | 19                  | 4                | 4.7                    | 5.19                    |

Nous trouvons, à travers ce module, une image très mitigée de la disponibilité de l'information et des données.

En ce qui concerne la consultation publique, nous constatons que bien que, dans 56.9% des pays interrogés, il existe au moins certaines informations en ligne qui peuvent être utilisées pour soutenir l'engagement ou l'examen des processus de consultation publique autour de l'élaboration de règles

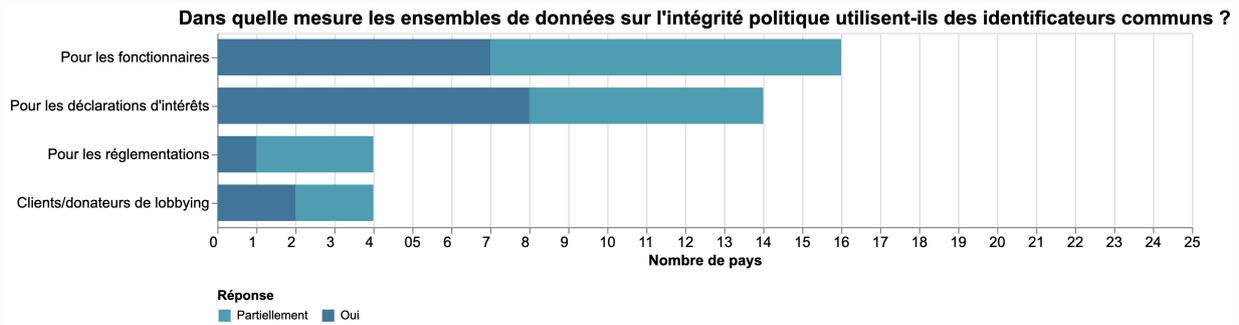
exécutives (parfois appelées « législation secondaire »), très peu de pays (17.4%) fournissent des données lisibles à la machine ou en vrac (11%) d'une manière qui pourrait faciliter l'utilisation d'outils innovants pour soutenir l'engagement des citoyens, ou pour surveiller l'influence des intérêts spéciaux par le biais de processus de consultation. Au-delà des pays principaux, la qualité des données fournies est nettement en baisse, de nombreux pays comptant parmi ceux qui fournissent *certaines* données en ligne, soit en ne couvrant qu'un petit nombre de consultations possibles, soit en ne disposant pas d'informations sur les commentaires formulés, soit en ne disposant pas d'autres caractéristiques essentielles : en se concentrant principalement sur la publicité de la possibilité de participer. Un certain nombre des pays les plus performants dans cet indicateur disposent de plateformes électroniques nationales pour faciliter l'élaboration de règles, telles que Regulation.gov des États-Unis et le Eelnõude Infosüsteem d'Estonie, bien que d'autres adoptent une approche plus distribuée, avec la responsabilité de chaque agence d'accueillir les détails des consultations et des réponses aux consultations.

Bien que la grande majorité des pays que nous avons étudiés disposent d'un cadre juridique relatif au droit à l'information (RTI), l'indicateur de disponibilité RTI fournit de nouvelles informations sur le nombre de pays publiant des informations sur les performances pour surveiller la mise en œuvre de leurs règles RTI. Il indique que 39.4% des pays disposent d'informations sur les performances en ligne, et que 88.4% d'entre eux fournissent une couverture raisonnablement large des régions et des agences. Lorsque des informations sont disponibles en ligne, 97.7% des pays fournissent des données sur le nombre de demandes faites, bien que seulement 58.1% fournissent des données sur quand et pourquoi les informations ont été refusées, et 58.1% fournissent des données sur la rapidité des réponses aux demandes : suggérant qu'une grande partie des informations fournies seront d'une utilité limitée pour surveiller le respect des délais légaux. Les données sur les recours contre les refus de RTI sont disponibles dans 58.1% des pays disposant d'informations en ligne, et dans certains cas, **seules** les données sur les recours sont disponibles, car les statistiques ne sont collectées que pour les procédures de recours, plutôt que pour les demandes initiales.

Il est important de se souvenir qu'une note élevée sur l'indicateur de disponibilité des RTI ne prouve pas qu'un régime RTI de pays *fonctionne* efficacement dans la pratique. La présence d'information sur le rendement démontre cependant un engagement à l'égard de la mise en œuvre de la RTI et fournit une ressource pour l'examen du fonctionnement des règles de la RTI. Les données probantes à l'appui et les liens recueillis pour cet indicateur (disponibles sur le site Web du Baromètre) pourraient soutenir la poursuite des travaux sur l'élaboration de mesures comparatives à l'échelle nationale pour la performance de la RTI, ainsi que fournir des indications sur les sources qui pourraient être utilisées pour des recherches qualitatives plus approfondies.

Les ensembles de données sur le financement politique et sur les déclarations d'intérêts et d'actifs sont particulièrement valables lorsqu'ils sont fournis sous forme de données en vrac et de données lisibles à la machine qui permettent à l'analyse de trouver des modèles ou de rechercher des liens qui pourraient indiquer l'influence politique d'entreprises ou de particuliers. Bien que 52.3% des pays disposent d'une certaine forme de données de financement politique en ligne et 45.9% d'informations sur les déclarations de patrimoine, moins de 10% des pays fournissent des ensembles de données ouverts dans les deux cas, avec des systèmes de divulgation souvent sur papier ou insuffisamment numériques pour soutenir la fourniture de données fiables lisibles à la machine. Ceci, jumelé à la faible disponibilité d'identifiants communs pour les donateurs (Just 29.8% des pays disposant d'informations en ligne) ou des personnes divulguant des actifs (40%) signifie que la réutilisation de ces données doit s'appuyer sur des intermédiaires qui peuvent numériser les ressources et utiliser diverses techniques de correspondance floue pour soutenir l'exploration des données et les enquêtes.

L'absence d'une infrastructure d'interopérabilité pour les données d'intégrité politique est illustrée par un indicateur de notre pilier capacité, qui voit que très peu de pays ont mis en place une utilisation généralisée d'identificateurs communs qui pourraient lier les ensembles de données. Bien qu'il existe des préoccupations légitimes en matière de protection de la vie privée quant à l'activation de certaines formes de couplage de données, les pays ayant une marge de manœuvre importante pour développer des approches plus numériques et conjointes pour les processus de divulgation liés à l'intégrité.



**Présence d'identificateurs communs soutenant l'interopérabilité des données d'intégrité politique:** Très peu de pays ont mis en place des infrastructures d'interopérabilité pour soutenir les données d'intégrité politique conjointes : créant des obstacles évitables à l'utilisation des données d'intégrité politique pour promouvoir une surveillance et une responsabilisation accrues.

Les occasions manquées et les leçons tirées de la faible interopérabilité des données quant au financement politique et à la divulgation des actifs devraient être instructives pour la conception des futurs cadres de lobbying et la publication des données. Le Baromètre révèle, pour la première fois, la faible disponibilité de données solides sur la manière dont les intérêts privés cherchent à influencer les politiques publiques. Avec seulement 17.4 % de pays fournissant des données en ligne et 4 n'offrant que des données ouvertes, cette première édition du Baromètre offre une base de référence pour suivre les progrès si le lobbying en faveur de la transparence gagne, comme nous l'espérons, plus de visibilité dans des forums tels que le Partenariat pour un gouvernement ouvert dans les années à venir.

En tout, bien que la comparaison des indicateurs de gouvernance associés à chaque indicateur de disponibilité de l'intégrité politique révèle une corrélation positive claire entre la qualité des règles exigeant la publication de données et la disponibilité des données (voir le chapitre sur la gouvernance), elle révèle également des lacunes importantes dans la mise en œuvre. Par exemple, alors que 103 pays ont des règles exigeant des déclarations d'intérêts et d'actifs, et 53 incluent des exigences concernant la collecte et la publication de données structurées, seulement 50 ont les informations disponibles en ligne, avec seulement 4 qui fournissent des ensembles de données ouverts. Les données figurant sur le site Web du Baromètre peuvent aider les pays à identifier des exemples de bonnes pratiques parmi leurs pairs afin de soutenir les actions visant à combler cette lacune à l'avenir.

## Informations sur l'entreprise

Les entreprises privées peuvent être des moteurs du développement, de l'innovation et de la fourniture de services essentiels et de biens de consommation. En même temps, les activités d'entreprise peuvent causer des dommages environnementaux et sociaux, les structures d'entreprise peuvent être utilisées à mauvais escient pour le blanchiment d'argent, la corruption et pour dissimuler des actes répréhensibles, et les structures d'entreprise opaques peuvent nuire au commerce international. Les informations sur l'enregistrement, la propriété et les activités des entreprises peuvent fournir des données probantes essentielles pour la compréhension et la réglementation par le public des activités des entreprises, et peuvent soutenir le fonctionnement d'un secteur privé productif qui soutient les résultats du développement durable.

Au cours des dernières années, des normes internationales ont été élaborées exigeant des niveaux minimums de collecte de données sur l'enregistrement et la propriété effective des entreprises, des fiducies et d'autres arrangements juridiques, la société civile plaidant pour que ces informations soient rendues accessibles au public, et des normes régionales et sectorielles de divulgation publique se développant dans le cadre des mesures visant à une plus grande transparence<sup>[23]</sup>. Le mouvement des données ouvertes a particulièrement mis l'accent sur les identifiants d'entreprise en tant que point de liaison entre différents ensembles de données publiques, y compris les ensembles de données sur les finances publiques, les marchés publics et l'intégrité politique<sup>[24]</sup>.

| Ensemble de données    | Disponible en ligne | Données ouvertes | Score moyen de qualité | Score moyen d'ouverture |
|------------------------|---------------------|------------------|------------------------|-------------------------|
| Ayant droit économique | 33                  | 10               | 6.83                   | 4.96                    |
| Registre des sociétés  | 78                  | 24               | 7.1                    | 5.55                    |

À l'heure actuelle, des progrès rapides dans la mise en place de registres ouverts des bénéficiaires effectifs peuvent être retardés par l'absence de pratiques établies en matière de données ouvertes dans les registres des sociétés.

Bien que 77 pays aient au moins un accès en ligne de base aux registres des sociétés, et 82.1% d'entre eux donnent l'accès à certaines de ces informations gratuitement, juste 46 fournissent des données d'entreprise lisibles à la machine, et seuls 24 pays fournissent des données accessibles en vrac, et sous des licences ouvertes pour répondre à la définition des données ouvertes. Dans de nombreux autres cas, les données structurées ne sont disponibles que dans le cadre d'accords commerciaux d'échange de données.

### Finances et marchés publics

Les données sur les dépenses et les marchés publics peuvent constituer une ressource puissante améliorant l'efficacité et l'efficacité du gouvernement, stimulant l'innovation dans le secteur privé, soutenant la lutte contre la corruption et favorisant une plus grande participation du public aux décisions quant aux dépenses publiques.

Les données sur le budget, les dépenses et les marchés publics figurent parmi les indicateurs de disponibilité les plus élevés du Baromètre, avec des exemples de données structurées et lisibles à la machine disponibles sur chaque continent. La disponibilité relativement élevée d'ensembles de données ouvertes bien structurées paraît refléter l'impact de campagnes mondiales de longue date et la fourniture d'une assistance technique aux pays pour mettre en œuvre le partage de données ouvertes, par des organisations, y compris nos partenaires thématiques le Partenariat pour les contrats ouverts (OCP) et l'Initiative mondiale pour la transparence fiscale (GIFT). Les systèmes de passation des marchés et des finances publiques ont également été largement numérisés, ce qui fait que la publication des données dépend, dans de nombreux cas, de la volonté politique plutôt que des capacités techniques. Notamment, tous les pays disposant de données ouvertes sur les marchés publics ne disposent pas de données ouvertes sur les finances publiques, et vice-versa.

| Ensemble de données                   | Disponible en ligne | Données ouvertes | Score moyen de qualité | Score moyen d'ouverture |
|---------------------------------------|---------------------|------------------|------------------------|-------------------------|
| Données sur le budget et les dépenses | 106                 | 33               | 5.79                   | 6.47                    |
| Données des marchés publics           | 100                 | 37               | 7.76                   | 6.57                    |

Les indicateurs du baromètre fournissent une ventilation des caractéristiques des ensembles de données disponibles, complétant des études telles que l'enquête ouverte sur le budget<sup>[25]</sup>, qui ont longtemps suivi la mesure dans laquelle les documents budgétaires clés sont mis à disposition, ainsi que les possibilités de participation du public aux processus budgétaires. Via ce détail, nous pouvons voir des domaines significatifs d'amélioration autour de la fourniture de la désagrégation budgétaire par des programmes transversaux ou des questions telles que les ODD, l'action climatique ou la budgétisation de genre (actuellement disponible pour seulement 21.2% des ensembles de données ouverts), et inclusion d'identifiants pour lier le budget et les dépenses aux grands projets (présents dans 45.5% des cas de données ouvertes).

Il existe des lacunes similaires dans les données disponibles sur les marchés publics, bien que les chercheurs aient trouvé de nombreux cas où au moins certaines informations des appels d'offres et des attributions sont publiées (81.7% et 79.8% des pays respectivement), des données sur la phase de planification (56%) et la mise en œuvre ou la livraison de biens ou de travaux contractuels (30.3%) étaient beaucoup moins susceptibles d'être disponibles. En général, la majorité des ensembles de données vérifiés contenaient de nombreux éléments essentiels à la compréhension des marchés publics, tels qu'une description des biens, des travaux ou des services achetés, le nom de l'entreprise à laquelle le marché a été attribué, et la valeur des marchés. Ceci cependant ne garantit pas que ces champs de données sont *toujours* présents. Dans 33.9% des cas, les chercheurs ont signalé soit des données probantes qu'une partie des données prescrites manquaient, soit des données probantes d'omissions généralisées parmi les données (l'approvisionnement figurait parmi les ensembles de données avec le plus grand nombre de pays avec des données manquantes signalées). Bon nombre des sources de données sur les achats recensées ne couvrent également qu'un nombre limité d'organismes ou, dans certains cas, représentent des exemples de grandes villes qui n'ont pas encore été reproduits dans l'ensemble du pays. Les réponses qualitatives ont également fait état d'une série de défis spécifiques liés à la qualité des données d'approvisionnement, tels que les avis d'attribution manquants ou les avis dont les variables clés n'étaient pas remplies. Ceci souligne que, lorsqu'on examine la disponibilité des données pour le bien public, ce ne sont pas seulement les champs énumérés aux en-têtes de colonne qui comptent - mais aussi la mesure dans laquelle ces champs sont remplis de données significatives. Un instrument mondial tel que le Baromètre peut servir d'indicateur de surface de la qualité probable des données, mais la qualité et la valeur complètes de tout ensemble de données ne peuvent être comprises que par l'utilisation des données.

## Terrain

Les données sur le régime foncier identifient les détenteurs de droits fonciers. Ces données peuvent être utilisées afin de comprendre le paysage de la propriété foncière dans un pays, pour identifier la concentration des terres, pour comprendre l'accès à la terre et la sécurité foncière, et à des fins de lutte contre la corruption<sup>[26]</sup>. Le Baromètre examine la disponibilité et l'ouverture des données foncières, ainsi que la mesure dans laquelle les données disponibles couvrent différents types d'occupation, y compris l'occupation urbaine et rurale, les terres appartenant à l'État, les terres communales, les terres appartenant à l'entreprise et les terres détenues par des particuliers.

Les données sur l'utilisation des terres décrivent les fins auxquelles celles-ci sont consacrées et ont une valeur significative pour la planification économique, la protection de l'environnement et les travaux sur l'atténuation du changement climatique et la résilience. Le Baromètre recherche des ensembles de données structurées qui détaillent les types d'activités se déroulant dans des emplacements particuliers, avec des références géospatiales y étant associées. Nous recherchons des données sur l'utilisation des terres comprenant des métadonnées décrivant la façon dont la classification des terres est effectuée, ainsi que des informations spécifiques sur les zones protégées et les zones forestières.

| Ensemble de données      | Disponible en ligne | Données ouvertes | Score moyen de qualité | Score moyen d'ouverture |
|--------------------------|---------------------|------------------|------------------------|-------------------------|
| Régime foncier           | 53                  | 11               | 5                      | 4.08                    |
| Usage des sols existants | 77                  | 31               | 6.95                   | 5.57                    |

Alors qu'un certain nombre de pays publient des données ouvertes structurées sur l'utilisation des terres, dans de nombreux cas, les cartes en ligne de l'utilisation des terres sont présentées sous forme de PDF ou d'images sans les données sous-jacentes. Il est possible de rendre beaucoup plus de données sur l'utilisation des sols disponibles en renforçant la capacité des principaux propriétaires de données à les publier dans des formats lisibles par machine.

Pour ce qui est des données sur le régime foncier, le tableau est mitigé à travers le monde. Dans certains cas, les données ne sont pas encore numérisées. Il existe des données dans d'autres pays, mais elles ne sont disponibles que suite à des demandes officielles ou du paiement de redevances. Si certains pays ont mis au point des modèles robustes dans le but d'équilibrer la fourniture ouverte de données foncières granulaires et la protection des données personnelles des propriétaires fonciers, il est urgent d'adopter des approches plus cohérentes pour gérer l'accessibilité et la protection des données. Le Baromètre fournit une nouvelle référence pour le suivi des progrès futurs dans la mise à disposition en ligne de types particuliers d'informations foncières : y compris les terres appartenant à l'État (actuellement 41.3% des pays ont des informations disponibles), des terres appartenant à des entreprises (36.7%), tenure individuelle (41.3%), et le régime foncier des populations autochtones ou marginalisées (15.6%).

## Comment les pays peuvent-ils améliorer la disponibilité des données ?

Pour chaque pays, le site Web du Baromètre mondial de données contient un profil qui montre la répartition des scores individuels par rapport aux indicateurs de gouvernance. Il peut être utilisé pour identifier des domaines particuliers d'amélioration en fonction des catégories de données couvertes dans cette édition du Baromètre. Les prochaines éditions du Baromètre porteront sur un certain nombre d'autres secteurs, ainsi que sur la mesure des progrès par rapport aux niveaux de référence établis dans la présente édition.

Sur la base des données quantitatives et qualitatives recueillies pour ce pilier de gouvernance, nous mettons en évidence les domaines d'action communs suivants :

### Développer l'infrastructure de données

Pour la publication des données, les pays à faible capacité sont souvent tributaires d'infrastructures obsolètes, ou fournies de l'extérieur. Ceci peut limiter leur capacité à gérer les données et à les publier en formats accessibles. Les gouvernements ont besoin d'aide afin de développer des systèmes de données nationaux intégrés<sup>[27]</sup>, et pour s'assurer qu'ils sont en mesure de soutenir les flux de données ouverts.

Pour les pays de capacité moyenne et supérieure, l'accent devrait être mis sur l'interopérabilité des données, par l'adoption de normes et d'identificateurs communs qui permettent d'établir des liens entre les ensembles de données, tout en tenant compte des risques potentiels de certains couplages de données.

## Actualiser les pratiques de publication des données

Les pays à capacité moyenne doivent se concentrer à combler les lacunes en matière de données et à améliorer la qualité des données. Ceci nécessite de passer de l'approche des données ouvertes comme exercice de case à cocher basé sur des portails de données à un accent sur les « données en tant que service », en reconnaissant que les ensembles de données publiés doivent être conçus pour les besoins des utilisateurs et doivent être activement maintenus. Dans certains cas, ceci impliquera des processus plus larges de reconception des services, en examinant comment les données provenant des systèmes d'utilisation des terres et des registres fonciers, des registres des sociétés, des systèmes d'information sur l'environnement ou d'autres éléments de l'infrastructure nationale de données peuvent être fournies sous forme ouverte et durable.

Les pays fédéraux, ou les pays où les villes ou les États sont en tête du programme de données ouvertes, peuvent devoir se concentrer plus particulièrement sur l'interopérabilité, pour évoluer vers une situation dans laquelle les données sont disponibles de manière fiable dans l'ensemble du pays, plutôt que d'être limitées à des zones urbaines particulières.

## Prise en charge de la fourniture de données et d'interfaces

Même pour les pays à plus forte capacité, les lacunes dans la fourniture de données en vrac, les accords de licences restrictifs et les modèles de tarification des données freinent les progrès vers une large disponibilité des données pour la réutilisation. En même temps, le manque d'interfaces accessibles pour explorer les données signifie qu'un grand nombre d'utilisateurs potentiels de données ne sont pas en mesure de trouver les points de données ou les faits particuliers qui pourraient leur être pertinents.

Les pays devraient se concentrer sur la compréhension des cas d'utilisation locaux et nationaux plus larges pour chaque grand ensemble de données fournis, et devraient chercher à offrir des moyens d'accès différents à différents publics : soit directement, soit en soutenant activement des intermédiaires durables.

## Utilisation et impact

Il est particulièrement difficile de mesurer la mesure dans laquelle les données sont utilisées pour produire des effets de valeur sociale. Au fil des ans, de nombreuses estimations ont été fournies sur la valeur économique ou sociale que les données peuvent apporter, présentant des chiffres importants sur la valeur potentielle. Il est toutefois difficile de démêler les avantages pour le bien public des données tirées de la croissance du PIB ou d'autres indicateurs agrégés, notamment parce que les statistiques économiques agrégées peuvent décrire les recettes provenant à la fois de l'utilisation et de l'utilisation abusive des données. Les approches qualitatives ont également leurs difficultés, bien que la compréhension des histoires sur la façon dont les données ont été utilisées puisse être précieuse pour inspirer la reproduction et l'adaptation des idées à de nouveaux contextes. Toutefois, lorsque les études de cas sont utilisées comme mesure de l'impact, ceci permet souvent de saisir à quel point les pays écrivent leurs utilisations des données ouvertes, plutôt que dans quelle mesure les données sont effectivement utilisées à grande échelle.

Dans la première édition du Baromètre, nous avons exploré une approche centrée sur les cas d'utilisation représentatifs afin d'acquérir une compréhension comparative approximative de l'utilisation et de l'impact des données. Nous avons présenté quatre exemples spécifiques d'utilisation des données que les chercheurs ont été invités à rechercher. Ces utilisations, chacune liée à un module thématique particulier du Baromètre, ont été conçues pour saisir une gamme de formes d'utilisation des données : des données permettant à la société civile d'examiner les gouvernements ; à la réutilisation par le secteur privé des données pour fournir des services qui apportent des avantages publics ; et de la recherche et du plaidoyer fondés sur les données pour promouvoir une plus grande équité ; à l'utilisation par les gouvernements des données pour fournir des services plus efficaces. En plus de ces indicateurs primaires, nous incluons un indicateur secondaire sur l'utilisation des données nationales par les organisations internationales.

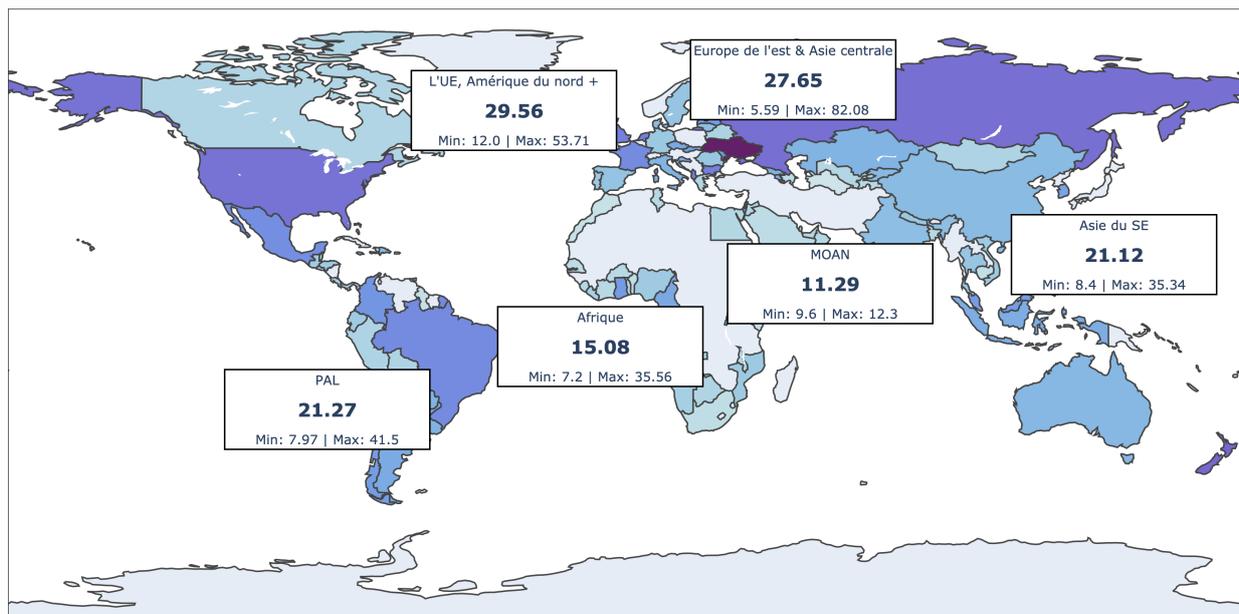
Bien que leur portée soit limitée et qu'elles ne couvrent qu'un très petit nombre d'utilisations possibles des données, les données recueillies pour ce pilier donnent un aperçu des moteurs de l'utilisation des données et des défis auxquels sont confrontés les utilisateurs actuels et potentiels des données.

### Résumé

- Parmi les exemples étudiés au Baromètre, il y a des niveaux relativement faibles d'utilisation des données, des 20 pays n'ayant signalé aucune utilisation notable de données par rapport aux exemples de cas d'utilisation fournis, et seuls des 22 pays ayant des utilisations pour rendre compte des quatre. Cependant, plus encourageant, dans 56.3% des 247 cas où l'utilisation des données a été identifiée, au moins certaines données probantes d'impact (par exemple, l'utilisation des données conduisant à un changement de politique ou à une valeur sociale ou politique) ont été identifiées.
- La majorité des utilisations des données identifiées ont été considérées comme des cas d'utilisation isolés (par exemple, une seule organisation de la société civile utilisant un ensemble de données pour une campagne de plaidoyer particulière), plutôt que de représenter des utilisations établies et répandues des données (par exemple, des données utilisées par de multiples organisations de la société civile (OSC) et médias dans un certain nombre de cas). Alors que les pays aux capacités plus élevées et plus faibles ont à peu près le même nombre de cas isolés d'utilisation des données, une utilisation généralisée semble corrélée à une capacité plus élevée. Les pays disposant plus particulièrement de capacités de données infranationales plus élevées sont plus susceptibles de voir une utilisation étendue des données.

- De nombreux utilisateurs de données œuvrent autour de la faible disponibilité des ensembles de données structurés en rassemblant des données provenant de sources multiples, en les convertissant et en les nettoyant. Dans certains cas, la société civile, les universitaires et parfois les utilisateurs du secteur privé deviennent des intermédiaires fournissant des données structurées à d'autres utilisateurs. Ceci n'est cependant pas souvent durable et peut engendrer des obstacles à l'intensification de l'utilisation des données.
- Les médias (36.4 % des cas vérifiés) et les organisations de la société civile (37.6 %) étaient les plus susceptibles d'être identifiés en tant qu'utilisateurs de données au sein de notre enquête, suivie par le gouvernement (30.3 %) et le milieu universitaire (25.7 %). Nous n'avons trouvé que 62 des cas d'utilisation par le secteur privé, principalement dans des pays où les secteurs technologiques prévalent.
- Nous avons trouvé peu de données probantes de l'utilisation des données depuis l'intelligence artificielle, avec moins de 20 cas cités dans l'ensemble de l'enquête. Dans un certain nombre de cas, l'IA a été utilisée pour nettoyer, plutôt qu'analyser, les données disponibles.

## Où les données sont-elles utilisées ?



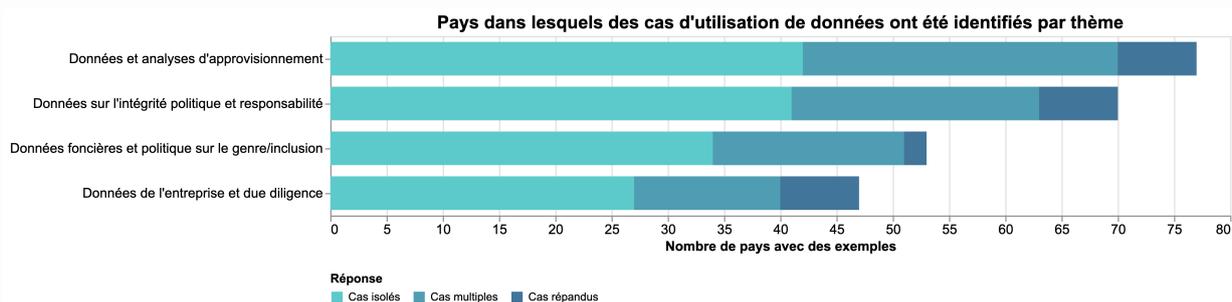
**L'utilisation et impact des scores régionaux du pilier :** les faibles scores du pilier Utilisation et impact indiquent que relativement peu de cas spécifiques d'utilisation des données recherchés par l'enquête du Baromètre ont été identifiés dans chaque région.

| Afrique   | L'Europe de l'Est et l'Asie Centrale  | L'Union européenne, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande                   | Amérique Latine et Caraïbes  | Moyen-Orient et Afrique du Nord  | Asie du Sud et de l'Est  |
|---|---|--|--|--|--|
| Ghana (35.56), Cameroun (31.49), Rwanda (24.1), Ouganda (19.32), Mozambique (17.82) | Ukraine (82.08), Russie (50.81), Albanie (45.46), Moldavie (36.77), Géorgie (28.21) | Nouvelle-Zélande (53.71), États-Unis (49.42), Pays-Bas (45.75), Royaume-Uni (45.12), Bulgarie (45.0) | Brésil (41.5), Mexique (40.33), Colombie (37.82), Chili (35.36), Argentine (30.29) | Jordanie (12.29), Oman (12.0), EAU (12.0), Qatar (12.0), Égypte (12.0) | Malaisie (35.34), République de Corée (32.57), Hong Kong (30.08), Indonésie (29.77), Chine (24.28) |

Le pilier de l'utilisation est principalement axé sur quatre indicateurs d'enquête, demandant dans quelle mesure il existe des données probantes de l'utilisation des données pour :

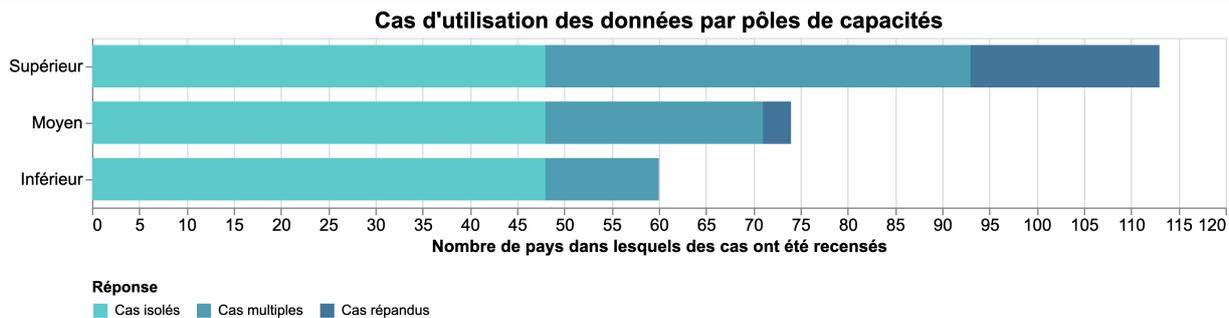
- **Améliorer les pratiques d'approvisionnement:** y compris par le biais de tableaux de bord analytiques, effectuer une analyse des signaux d'alerte, améliorer la diversité de l'approvisionnement, ou évaluer et améliorer les impacts environnementaux de l'approvisionnement.
- **Identifier, exposer ou mettre en évidence les défaillances gouvernementales:** en mettant l'accent sur l'utilisation des données d'intégrité politique par un large éventail de parties prenantes, y compris la société civile, les journalistes, les universités et le secteur privé.
- **Influencer les politiques dans l'intérêt d'un régime foncier et d'une utilisation équitable et inclusive :** notamment par le biais de journalistes, d'universitaires ou d'organisations de la société civile.
- **Soutenir la diligence raisonnable des entreprises :** notamment en fournissant des services aux organisations du secteur privé pour vérifier les références des partenaires commerciaux potentiels, pour que les gouvernements effectuent des vérifications avant de conclure des contrats, ou aux médias de rendre compte de la propriété de l'entreprise.

Pour chaque indicateur, nous avons demandé aux chercheurs d'identifier les cas d'utilisation et de vérifier s'il s'agissait d'exemples isolés ou représentatifs d'une utilisation généralisée. Comme le montre le graphique ci-dessous, les données d'approvisionnement ont été les plus utilisées, avec des utilisations de diligence raisonnable des données de l'entreprise identifiées dans le plus petit nombre de pays.



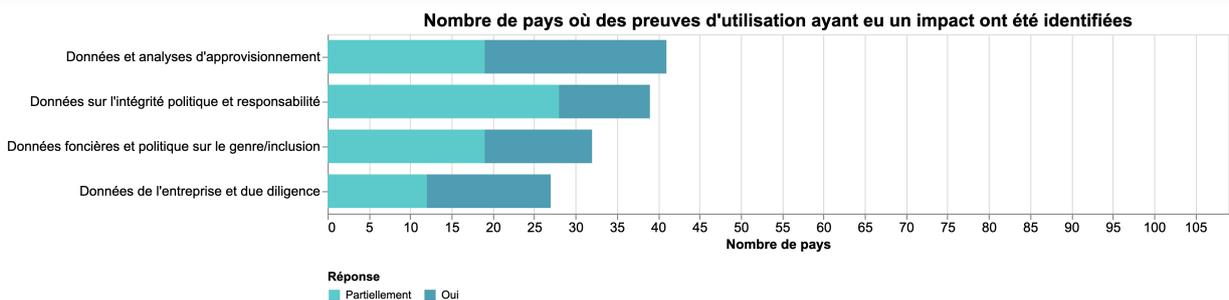
**Cas d'utilisation de données par sujet :** les cas d'utilisation de l'analyse d'achats ont été les plus fréquemment identifiés, suivis de l'utilisation des données à des fins de responsabilité politique et d'intégrité, puis de l'utilisation des données liées à la gouvernance foncière. Bien que moins de pays puissent démontrer que les entreprises privées nationales utilisent les données d'entreprise afin de soutenir les activités de diligence raisonnable, là où les données et les marchés à l'appui sont disponibles, il est plus probable que ceci se traduise par une utilisation généralisée des données.

En examinant les cas d'utilisation des données par les pôles de capacités (voir le chapitre Capacités), nous constatons des niveaux étonnamment élevés d'utilisation des données dans les pays à faible capacité, bien que la plupart des utilisations identifiées aient été des cas isolés. Un examen des justifications qualitatives de ces scores révèle que, si la société civile ou les journalistes dans les pays disposant de peu de données peuvent être en mesure de mener des projets pilotes ou des analyses ponctuelles, ou si les donateurs peuvent soutenir des initiatives de démonstration dans ces contextes, le maintien et la mise à l'échelle de l'utilisation des données nécessitent à la fois une fourniture plus durable de données de la part du gouvernement et une plus grande capacité de la population dans son ensemble à travailler avec les données mises à disposition. Notamment, plus de 50 % des cas d'utilisation des données du secteur privé recensés provenaient de pays à forte capacité.



**Cas d'utilisation des données par pôle de capacités:** Les pays à capacité supérieure, moyenne et faible ont tous un nombre similaire de cas isolés d'utilisation des données, tandis que l'utilisation généralisée des données est corrélée avec la capacité.

On a également demandé aux chercheurs de déterminer s'il existait des données probantes que les utilisations particulières des données couvertes par l'enquête avaient des effets positifs significatifs et d'évaluer la force des données probantes. Notamment, bien qu'il y ait peu de différence entre les indicateurs lorsque nous prenons en considération la présence de données probantes (certaines données probantes existent dans 56.3% des cas), lorsque nous recherchons des données probantes solides de l'impact, et des niveaux élevés de confiance que les données sont à l'origine de l'impact, les scores pour les analyses d'approvisionnement des données étaient deux fois plus élevés que pour l'intégrité politique. En bref : il semble que, bien que les données soient un outil puissant pour identifier, exposer ou révéler les défaillances du gouvernement, les obstacles à la traduction de cette information en impact sont substantiels, les problèmes persistants d'impunité et les mécanismes de responsabilisation faibles ne doivent pas être sous-estimés lorsqu'on pense à la façon dont les données peuvent mener au changement.



**Preuve d'impact:** de solides données probantes d'impact ont été identifiées dans deux fois plus de pays quant à l'utilisation des données analytiques d'approvisionnement, que pour l'impact de l'utilisation des données d'intégrité politique.

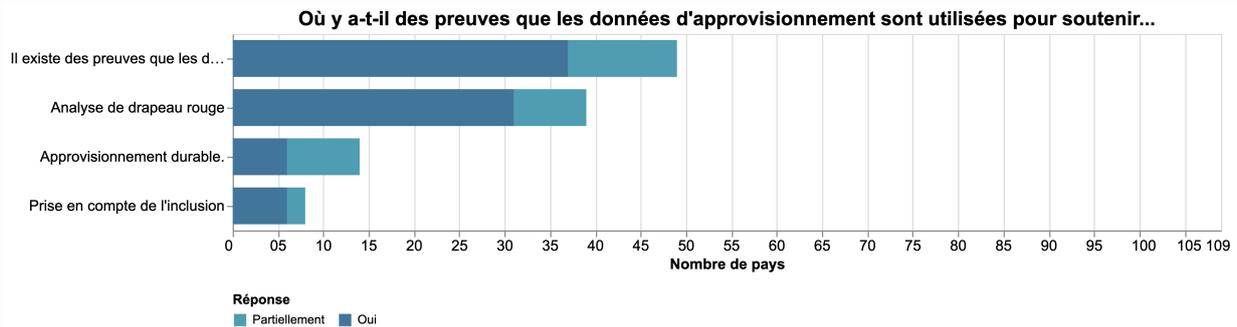
## Comment l'utilisation des données évolue-t-elle ?

Bien que cette première édition du Baromètre ne puisse fournir de données longitudinales sur l'utilisation, les riches justifications qualitatives et la collecte de liens et de brefs résumés fournis par les chercheurs peuvent fournir un aperçu critique de la façon dont les pratiques d'utilisation des données se développent dans le monde.

### Tableaux de bord et détection

En ce qui a trait à l'utilisation des données d'approvisionnement, nous constatons que l'accent est mis sur la création de tableaux de bord et d'outils d'analyse des signaux d'alerte pouvant utiliser les données dans le but d'indiquer les risques potentiels de corruption. L'enquête du Baromètre a identifié plus de 50 exemples de tableaux de bord d'approvisionnement accessibles au public produits par un éventail de parties prenantes, d'agences officielles d'approvisionnement public aux organisations de la société civile, aux programmes de recherche internationaux et aux journalistes d'enquête. Vous trouverez ci-dessous des captures d'écran sélectionnées de portails, indiquant certains des éléments de commentaire, et des approches uniques prises parmi chacune. En particulier, le nombre de

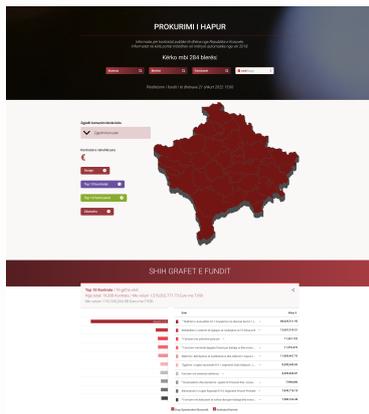
plateformes qui ont été adaptées ou lancées avec un accent particulier sur les achats d'urgence de COVID-19, démontrant la capacité des plateformes de données ouvertes à réagir rapidement aux événements actuels.



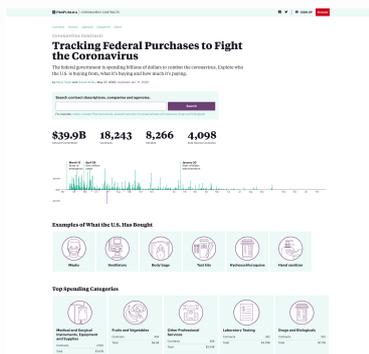
**Utilisation de données d'approvisionnement:** les données d'approvisionnement sont largement utilisées dans le but de fournir des tableaux de bord et pour l'analyse des signaux d'alerte, mais il y a moins de données probantes que les données soient utilisées au soutien, à l'inclusion ou à l'approvisionnement environnemental

Les utilisations des données sur les achats appuyant les achats durables ou pour promouvoir des pratiques d'achat plus inclusives (telles que l'analyse des incidences des décisions d'achat sur l'égalité des sexes) étaient moins évidentes. Bien qu'un certain nombre de pays aient élaboré des politiques de "marchés publics écologiques", peu de données probantes que les données étaient utilisées pour les suivre existent, seuls la République de Corée et la Lettonie ayant fourni des données probantes dans notre enquête des statistiques sur le respect de ces politiques. Ceci met en évidence un futur domaine de travail, une meilleure intégration des plateformes solides de transparence des données d'approvisionnement, avec des défis critiques d'approvisionnement durable.

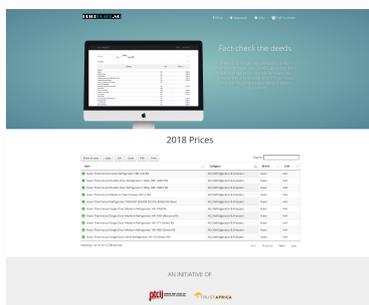
## Tableaux de bord d'approvisionnement comparés



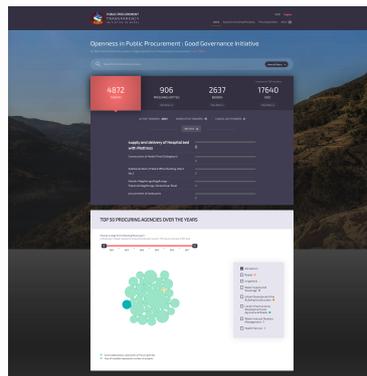
Le Kosovo Open and Transparent Procurement Portal - (PPHT) est géré par une organisation de la société civile, avec un financement de l'USAID et la coopération de l'agence des marchés publics. Des développements récents du portail ont ajouté la possibilité de rechercher des contrats par le nom du propriétaire de l'entreprise recevant un contrat.



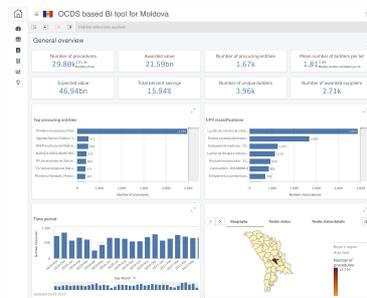
Aux États-Unis, la salle de presse du journalisme d'investigation ProPublica a utilisé les données d'approvisionnement pour créer un tableau de bord axé sur les contrats fédéraux liés à la COVID-19.



Au Nigéria, le Premium Times Centre for Investigative Journalism a utilisé les données du Bureau nigérian des marchés publics pour fournir un tableau de bord des prix des produits de base, répondant en particulier aux problèmes de corruption et d'inflation des coûts.



Au Népal, une plateforme initialement prototypée par la société civile a été adoptée par le Bureau officiel de suivi des marchés publics et est mise à jour à l'aide des données provenant des systèmes d'approvisionnement en ligne.



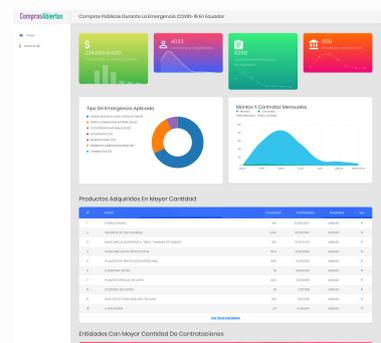
En Moldavie, il existe une gamme de tableaux de bord basés sur les données de la norme du Open Contracting Data Standard, y compris un tableau de bord basé sur le logiciel Qlik Business Intelligence - répliquant un modèle développé initialement par la plateforme ProZorro Business Intelligence en Ukraine



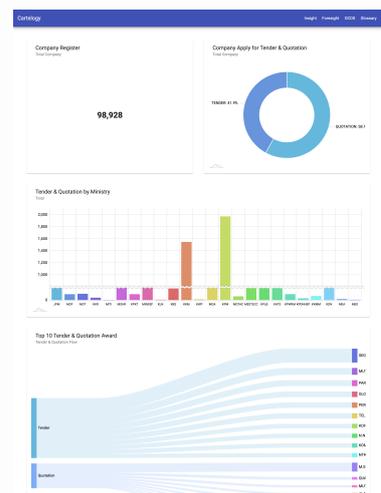
En République Kirghize, un tableau de bord financé par la Banque européenne pour la reconstruction et le développement s'appuie sur les données du portail national des marchés publics. Les chercheurs rapportent que ce tableau de bord est parfois utilisé en soutien des reportages des médias, mais qu'il est plus couramment utilisé par les entreprises à la recherche d'opportunités d'approvisionnement.



En Italie, l'Authority National AntiCorruzione (ANAC) héberge un tableau de bord analytique sur la base de données nationale des marchés publics

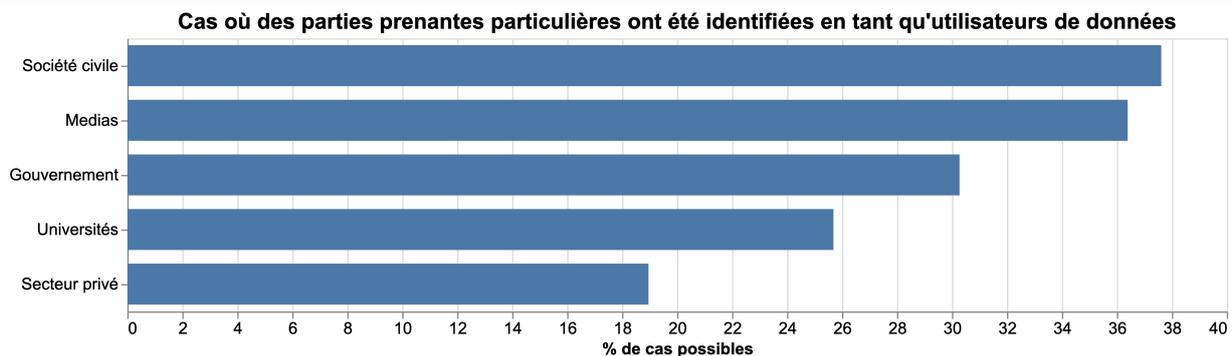


En Équateur, un tableau de bord conçu par la Fundapi, organisme à but non lucratif a pris des données au format Open Contracting Data Standard du Service national des marchés publics afin de visualiser l'utilisation des procédures d'urgence dans la réponse au COVID-19



En Malaisie, la plateforme « Cartology » a été développée pour relier les données d'approvisionnement, les profils d'entreprises et les données sur les personnes politiquement exposées pour créer un outil « d'alerte » conçu pour aider les fonctionnaires à détecter de manière proactive les activités d'approvisionnement potentiellement suspects.

## Partenariats prometteurs



**Parties prenantes utilisant des données:** Les organisations de la société civile ont été identifiées en tant qu'utilisateurs de données dans 37.6% des cas où ceci a été vérifié, tandis que les utilisateurs du secteur privé ont été identifiés dans seulement 19% des cas.

De par le passé, certains des principaux obstacles à l'utilisation des données de la société civile et des médias ont été liés à des déficits de compétences. Bien que d'importantes lacunes en matière de maîtrise des données dans les médias traditionnels et les organisations de la société civile subsistent, nous observons un certain nombre de tendances prometteuses, notamment la professionnalisation en cours du journalisme de données, la preuve de partenariats plus étroits entre les organisations traditionnelles de la société civile et de technologie civique, et l'utilisation par les médias de plateformes commerciales fournies pour le soutien à la fois la recherche et de la visualisation de données dans des histoires médiatiques complexes.

Aux États-Unis, les données sur l'intégrité politique sont largement utilisées à la fois par les plateformes spécialisées dans le journalisme de données et le journalisme d'investigation ainsi que dans les salles de presse traditionnelles et par les universitaires qui s'intéressent aux tendances plus larges du lobbying politique ou du financement des partis. Des utilisations journalistiques étendues des données sur l'intégrité politique ont également été identifiées au sein de pays tels que le Ghana, le Chili, l'Argentine et le Népal, bien que les journalistes se soient tournés vers des lois sur le droit à l'information plutôt que vers des sources de données ouvertes dans un certain nombre de cas.

La capacité des médias à rendre compte de leurs activités de manière à avoir des effets substantiels sur l'intégrité politique est déterminée par un environnement politique plus élargi. Là où les institutions officielles sont solides, les médias peuvent toutefois jouer un rôle important en portant l'attention sur des questions qui déclenchent alors une activité d'application de la loi. Rédigé avant l'invasion illégale de l'Ukraine en 2022, notre chercheur a rapporté que les rapports des médias basés sur le registre des déclarations de patrimoine ont été observés pour déclencher une action d'organismes officiels de lutte anti-corruption, qui pourraient vérifier les données utilisées, et parfois engager des poursuites pénales.

Dans les rapports sur l'utilisation des données de diligence raisonnable des entreprises, nous avons observé un certain nombre de cas où les médias s'appuyaient sur des plateformes commerciales pour mener des recherches sur la propriété de l'entreprise, en utilisant des visualisations générées par ces plateformes dans leurs reportages médiatiques. Un certain nombre de plateformes semblent en avoir reconnu la valeur en offrant aux organisations nationales de la société civile et aux journalistes un accès gratuit à leurs plateformes tout en appliquant un modèle commercial payant pour les utilisateurs commerciaux.

Un certain nombre de cas d'utilisation identifiés dans l'histoire du Baromètre semblent impliquer des collaborations entre des organisations de la société civile traditionnelle axées sur les causes et organisations de technologie civique plus techniquement orientées. Ceci représente une maturation du paysage s'éloignant des organisations de technologie civique menant leurs propres campagnes indépendantes, et suggère un modèle de pratique plus durable parmi lequel les organisations ayant un engagement à long terme pour une cause particulière combinent leurs connaissances de domaine aux

compétences aux connaissances des spécialistes des données. En France, un consortium d'ONG a, par exemple, collaboré avec la coopérative Dataactivist dans le but de développer une plateforme usant de l'apprentissage automatique pour analyser si les entreprises publient des plans légalement requis sur les activités de due diligence.

Il est également prouvé que des sources de données publiques sont utilisées par des acteurs hors-pays, plus particulièrement dans les cas où les libertés politiques nationales sont restreintes. Bien que plusieurs pays ne signalent aucune utilisation des données sur l'intégrité politique pour demander des comptes aux gouvernements, en raison soit d'un manque de données ou d'espace civique indépendant, dans un certain nombre de cas, les chercheurs signalent que des organisations non gouvernementales internationales ou des citoyens expatriés ou des journalistes d'investigation outre-mer utilisent des données pour mettre en évidence les échecs ou les sujets de préoccupation gouvernementaux.



## Travailler pour combler les lacunes dans les données

L'utilisation des données du baromètre à la comparaison de la disponibilité des données foncières et la preuve de l'utilisation des données foncières révèle une corrélation étonnamment faible. Comme nous avons plus de pays où l'utilisation des données est évidente que de pays fournissant des données entièrement ouvertes, il est là suggéré que les acteurs utilisent les données *en dépit de* leur faible disponibilité, plutôt *qu'en raison des* données disponibles.

L'examen des réponses détaillées à l'enquête révèle de nombreux cas où les médias et les acteurs de la société civile ont réuni des données probantes documentaires provenant de registres fonciers ou de documents officiels, et les ont combinées avec d'autres sources dans le but de produire une analyse fondée sur des données. Comme noté par le chercheur du pays, cette affaire démontre « *d'avantage la réalisation de l'OSC qui a mené l'affaire, plutôt que la disponibilité des données pour un usage public* », plus particulièrement en termes de la transformation d'une observation de déforestation illégale en une plainte formelle. Par exemple, au Paraguay, l'ONG environnementale IDEA s'est servie des demandes d'accès à l'information afin d'obtenir des données sur la propriété foncière, et les a ensuite combinées avec des images satellites et des informations sur les licences pour détecter le déboisement illégal, puis pour formuler une plainte officielle.

Nos données d'enquête suggèrent aussi que la société civile joue actuellement le plus grand rôle dans l'utilisation des données foncières pour conduire un changement de politique (40 pays), suivi par les universités et les médias (en 35 et finalement par le secteur privé (en 19). Dans de nombreux cas où l'utilisation des données foncières par le secteur privé a été signalée, l'accent a été plus souvent mis sur le développement de services pour les acheteurs et vendeurs immobiliers, plutôt que sur un soutien explicite à une plus grande équité et inclusion sur les marchés fonciers. À la fois en France et à Taiwan, les chercheurs ont signalé que les principaux utilisateurs de données foncières étaient des entreprises privées opérant ainsi, Taiwan fournissant l'un des seuls 5 exemples identifiés d'IA utilisée pour le travail avec des données foncières.

Les questions de concentration de la propriété foncière et de politique agricole figurent en bonne place parmi celles abordées dans le cadre de l'utilisation des données. Au Brésil, par exemple, la « carte des inégalités » se sert des dossiers de plus de 5,3 millions de propriétés rurales de l'Atlas de l'agriculture brésilienne pour contribuer aux discussions en cours et intensément politiques sur la réforme agraire. Ce projet, utilisant l'analyse de données géographiques, offre la possibilité de fournir des perspectives alternatives sur l'inégalité foncière par rapport à celles disponibles uniquement à partir de données statistiques officielles. Dans un autre exemple, couvrant à la fois la politique agricole et les questions d'accapement des terres, l'ONG mexicaine, le Civil Council for Sustainable Forestry (CCMSS) s'appuie sur les données du Registre agraire national pour développer des visualisations axées à prôner et montrer les écosystèmes, les activités économiques et le régime foncier. Ceci alimente à la fois le soutien consultatif aux producteurs par le biais du travail sur le terrain, et la promotion des politiques publiques éclairée par une combinaison d'expérience de base et d'analyse de données.

Les difficultés à la reconstitution d'une image de la propriété foncière, même au sein des économies fortement numérisées, sont évidentes dans les cas d'utilisation du Royaume-Uni et des Pays-Bas. Parmi ces deux pays, où les médias et la société civile ne disposent pas d'accès massif aux informations sur le régime foncier, les journalistes et les militants ont assemblé des dossiers pour présenter des données sur les plus grands propriétaires fonciers, remettant en question la réflexion classique sur la répartition de la propriété foncière, y compris en fournissant des données probantes sur l'étendue des terres appartenant à l'État. Dans une analyse, avec l'appui des données de l'Organisation des Nations Unies pour l'alimentation et l'agriculture (FAO), les journalistes rapportent que seulement 6 % des terres néerlandaises appartiennent à des femmes.

Au Rwanda, l'un des rares pays de notre enquête à collecter et publier des données sur le sexe des propriétaires fonciers, le système d'information sur l'administration des biens fonciers et le tableau de bord foncier rwandais qu'il pilote, a été utilisé pour l'analyse des politiques par les universitaires, la

société civile, les médias et le gouvernement. Nos chercheurs ont noté que sa création, depuis 2005, a été une réalisation importante, mais que les défis à relever pour affiner les résultats statistiques de haut niveau demeurent et que le tableau de bord présente en informations exploitables pour les décideurs. Cette question, dont les utilisateurs dépendent souvent des statistiques publiées par les organismes publics, mais ne peuvent accéder aux données sous-jacentes pour élaborer des paramètres d'analyse plus détaillés, est un thème courant dans un certain nombre de pays et limite sans doute la marge de manœuvre dont disposent les acteurs extérieurs au gouvernement pour élaborer des contributions solides aux débats d'orientation. Les chercheurs ont plus particulièrement noté que lorsque la société civile doit s'appuyer sur des sources de données limitées et ponctuelles, ceci peut conduire à une fragmentation du débat politique et du plaidoyer, lorsque, comme nos réponses à l'enquête sur l'utilisation des données foncières de la Sierra Leone l'ont indiqué, « *le plaidoyer va souvent dans tous les sens avec des objectifs différents simplement parce que chacune des organisations utilisent des données non standardisées* ».

## Comment les pays peuvent-ils améliorer la gouvernance des données à l'avenir ?

Pour chaque pays, le site Web du Baromètre mondial de données contient un profil indiquant la répartition des scores individuels vis-à-vis les indicateurs de gouvernance. Ceci peut être utilisé à identifier les domaines possibles d'amélioration, bien qu'il faille noter que les futures éditions du Baromètre sont susceptibles de comprendre un plus large éventail de cas d'utilisation et d'évaluation d'impact que ceux inclus dans cette édition, et peuvent voir un certain développement de l'utilisation et de la méthodologie de mesure d'impact.

Nos recommandations portent par conséquent moins sur les cas d'utilisation particuliers couverts par l'enquête Baromètre que sur des approches particulières visant à appuyer la réutilisation diversifiée des données axées sur le bien public. Ceci repose sur une reconnaissance du fait que la demande de données et l'utilisation des données sont des facteurs essentiels pour l'amélioration de la qualité des données et que la transformation des infrastructures nationales de données requiert la participation active d'un large éventail d'utilisateurs de données.

## Les gouvernements devraient faire un plus large et plus transparent usage de leurs propres données.

Avec l'amélioration des cadres de gestion et de partage des données, en renforçant des capacités internes d'utilisation des données et à l'établissement de partenariats avec divers secteurs, les gouvernements peuvent accroître la valeur publique générée par les données déjà recueillies et détenues.

Nous avons constaté qu'il était souvent difficile de déterminer comment le gouvernement se sert des données. Cependant, pour promouvoir l'utilisation de données qui servent le bien public et pour se protéger envers celles pouvant nuire aux minorités ou aux populations marginalisées, ou ayant autrement des répercussions négatives, il est toutefois important que les gouvernements rendent compte aux citoyens de la façon dont les données sont utilisées. Lors de la publication d'ensembles de données, de la commande de nouvelles analyses ou de nouveaux outils fondés sur des données ou l'utilisation de systèmes algorithmiques, les gouvernements devraient trouver des moyens de sensibiliser le public à la façon dont ceux-ci sont utilisés.

## Soutenir les partenariats intersectoriels pour l'utilisation des données

Certains des exemples les plus intéressants et les plus durables de données identifiées impliquent les partenariats multi-sectoriels : y compris la société civile traditionnelle travaillant avec des organisations à but non lucratif et des entreprises de technologie civique ou plateformes de données initialement développées par des organisations non gouvernementales adoptées et hébergées en permanence par des agences gouvernementales. Les gouvernements, bailleurs de fonds et autres parties prenantes devraient tous réfléchir aux moyens d'appuyer ces formes de collaboration et d'éliminer les obstacles s'y opposant.

## Mettre davantage l'accent sur l'équité et l'inclusion

Nous avons trouvé relativement peu de données probantes que les données étaient utilisées pour aborder les questions d'équité et d'inclusion. L'utilisation des données pour le bien public doit répondre aux besoins de tous. Il peut s'agir d'accorder une attention particulière à la promotion d'une utilisation des données alignée sur les objectifs de développement durable, ou la mise en place des cadres garantissant que l'utilisation des données tient compte des besoins particuliers des populations marginalisées.

# Analyse régionale

- Afrique
- Europe de l'Est et Asie centrale
- L'Union européenne, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande
- Amérique latine et Caraïbes
- Moyen-Orient et Afrique du Nord
- Asie du Sud et de l'Est

Dans cette section, nous explorons les résultats du Baromètre dans 6 régions mondiales, sélectionnées pour refléter la structure de notre réseau de pôles de recherche : l'Afrique ; l'Europe de l'Est et l'Asie centrale ; l'Union européenne, le Royaume-Uni, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande ; l'Amérique latine et les Caraïbes ; le Moyen-Orient et l'Afrique du Nord ; et l'Asie du Sud et de l'Est. Vous trouverez ci-dessous des données sommaires pour chaque région, ainsi que des informations sélectionnées provenant du ou des pôles de recherche régionaux responsables de la coordination de la recherche dans cette région.

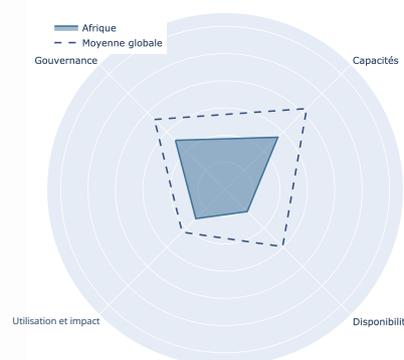
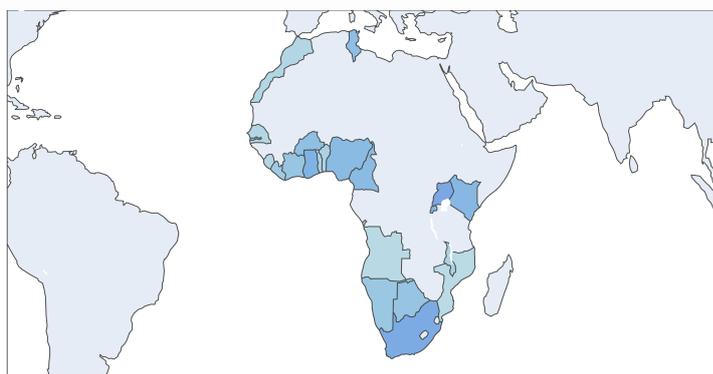
## Lecture des tableaux régionaux

- **Les forces et les faiblesses** sont sélectionnées en fonction de l'indicateur primaire pour lequel un pays s'écarte le plus de la moyenne régionale. En d'autres termes, parmi tous les indicateurs disponibles, la colonne de la force indique les indicateurs où le pays est le plus susceptible d'avoir un exemple de bonne pratique à partager avec d'autres pays de la région. La colonne des faiblesses indique les domaines dans lesquels un pays peut être confronté à des défis particuliers et/ou ceux où il y a le plus de possibilités de progrès rapides en apprenant de pairs régionaux. Ceux-ci sont déterminés uniquement par les scores d'indicateurs et ne reflètent pas une décision éditoriale.
- **Dirigeants régionaux (RL) et mondiaux (GL).** Nous utilisons « RL » par rapport à un indicateur pour indiquer lorsqu'un pays a l'un des scores les plus élevés de la région sur un indicateur particulier, et « GL » quand il a l'un des scores les plus élevés au monde. Ceci peut être utile pour localiser les pays qui peuvent démontrer les meilleures pratiques.
- **Scores des composantes.** Le tableau comprend la note pondérée pour chaque pays pour les composantes Gouvernance, Capacités, Disponibilité et Utilisation et l'Impact du Baromètre. Chacune de ces composantes se voit attribuer une note sur 10.
- **Politique et initiative en matière de données ouvertes.** Les tableaux comprennent des scores d'indicateurs spécifiques pour la présence de la politique d'Open Data et des initiatives d'Open Data (voir les chapitres Gouvernance et Capacités). Ces indicateurs sont globalement comparables aux indicateurs d'éditions précédentes du Baromètre de données ouvertes, qui permettent de savoir si les politiques et pratiques nationales en matière de données ouvertes se sont améliorées ou ont été négligées au cours des dernières années. Un changement de moins de 1 sur 10 au score de l'indicateur est indiqué comme . Une augmentation de plus d'un point sur 10 des scores ODB antérieurs est indiqué à l'aide de , suggérant une amélioration probable de la qualité de la politique ou de l'initiative. Une chute de plus de 1 sur 10 est représentée à l'aide de . Les mesures de l'ODB peuvent provenir de 2020 (Amérique latine), 2017 (dirigeants de l'ODB) ou 2016 (autres pays). Lorsqu'il n'y a pas de données ODB antérieures auxquelles comparer,  est indiqué.

Ces tableaux doivent être lus en gardant à l'esprit les limites de cette première édition du Baromètre. Il est possible que dans certains cas, les forces et les faiblesses affichées représentent des erreurs de mesure, ou des valeurs aberrantes présentes en raison d'un aspect particulier du contexte du pays qui ne peut être pleinement compris qu'en lisant les données probantes qualitatives à l'appui. De même, bien que certains indicateurs de la BDG et de la BDO aient été conçus à des fins de comparabilité, il existe des différences dans la méthode de mesure pouvant affecter des pays particuliers. Les tableaux ci-dessous sont destinés au début, plutôt qu'à la fin, ou à la discussion et à l'exploration. Étant donné les limites de l'espace, nous n'affichons qu'une seule force ou faiblesse pour chaque pays.

# Afrique

Le Baromètre couvre 22 pays d'Afrique. La région obtient des résultats inférieurs à la moyenne mondiale pour tous les piliers, avec un besoin important d'investissement dans les institutions de gouvernance des données, des infrastructures de données robustes et complètes, et dans la promotion de vastes capacités de gestion et d'utilisation des données pour le bien public. Il existe toutefois d'importantes possibilités pour l'apprentissage par les pairs de progresser sur des sujets tels que la gestion et le partage des données des finances publiques, et de transférer les apprentissages à d'autres secteurs tels que les données pour l'action climatique. La région dispose aussi d'un certain nombre de cadres politiques solides et améliorés en matière de données ouvertes sur lesquels s'appuyer, bien que le maintien des ressources et de l'appui aux initiatives en matière de données ouvertes demeure un défi majeur pour l'avenir.



| Score global | Gouvernance | Capacités | Disponibilité | Impact et utilisation |
|--------------|-------------|-----------|---------------|-----------------------|
| 19.61        | 25.6        | 27.19     | 11.43         | 15.08                 |

| Pays | Force comparative (régionale)                                  | Faiblesse comparée (régionale)          | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données       |
|------|--|---|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------------|
|      | (G) Données des finances publiques                             | (A) Données des marchés publics         | 10.6         | 16.7        | 15.9     | 2.7           | 15.2                  | 0.0 <input type="checkbox"/>  | 0.0 <input type="checkbox"/>  |
|      | (G) Protection des données                                     | (G) Gestion des données                 | 14.4         | 14.2        | 23.1     | 10.1          | 8.8                   | 0.0 <input type="checkbox"/>  | 40.0 <input type="checkbox"/> |
|      | (G) Données de consultation publique <sup>RL GL</sup>          | (A) Données des marchés publics         | 22.5         | 36.2        | 32.6     | 8.2           | 10.8                  | 54.0 <input type="checkbox"/> | 72.0 <input type="checkbox"/> |
|      | (G) Gestion des données  | (G) Déclarations d'actifs               | 20.2         | 29.5        | 28.0     | 9.6           | 13.0                  | 0.0 <input type="checkbox"/>  | 32.0 <input type="checkbox"/> |
|      | (C) Soutien gouvernemental à la réutilisation <sup>RL GL</sup> | (G) Déclarations d'actifs               | 19.8         | 25.2        | 41.3     | 4.6           | 10.8                  | 0.0 <input type="checkbox"/>  | 51.0 <input type="checkbox"/> |
|      | (A) Utilisation actuelle des terres <sup>RL</sup>              | (G) Gestion des données                 | 24.1         | 23.9        | 25.6     | 22.8          | 31.5                  | 0.0 <input type="checkbox"/>  | 36.0 <input type="checkbox"/> |
|      | (G) Propriété effective <sup>RL GL</sup>                       | (A) Données des marchés publics         | 27.7         | 32.3        | 43.2     | 14.7          | 35.6                  | 0.0 <input type="checkbox"/>  | 70.0 <input type="checkbox"/> |
|      | (G) Données de consultation publique <sup>RL GL</sup>          | (A) Données des marchés publics         | 20.5         | 29.0        | 20.7     | 14.6          | 17.2                  | 0.0 <input type="checkbox"/>  | 0.0 <input type="checkbox"/>  |
|      | (G) Données de consultation publique <sup>RL GL</sup>          | (C) Initiative sur les données ouvertes | 25.7         | 44.0        | 21.9     | 16.5          | 8.4                   | 0.0 <input type="checkbox"/>  | 0.0 <input type="checkbox"/>  |
|      | (G) Performance RTI <sup>RL</sup>                              | (G) Protection des données              | 17.2         | 24.6        | 14.7     | 13.6          | 14.9                  | 0.0 <input type="checkbox"/>  | 20.0 <input type="checkbox"/> |
|      | (A) Registre du commerce et des sociétés                       | (G) Données des finances publiques      | 12.4         | 10.8        | 23.0     | 7.9           | 7.2                   | 0.0 <input type="checkbox"/>  | 0.0 <input type="checkbox"/>  |
|      | (G) Cadres de partage des données                              | (G) Protection des données              | 10.3         | 14.0        | 13.5     | 5.0           | 17.8                  | 0.0 <input type="checkbox"/>  | 0.0 <input type="checkbox"/>  |
|      | (G) Cadres de partage des données                              | (G) Gestion des données                 | 14.6         | 27.3        | 17.3     | 4.3           | 11.6                  | 28.0 <input type="checkbox"/> | 16.8 <input type="checkbox"/> |
|      | (A) Vaccination (COVID-19) <sup>RL</sup>                       | (G) Données des finances publiques      | 18.9         | 21.1        | 23.2     | 15.2          | 15.4                  | 0.0 <input type="checkbox"/>  | 0.0 <input type="checkbox"/>  |

| Pays  | Force comparative (régionale) | Faiblesse comparée (régionale)                             | Score global                              | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |        |
|---|-------------------------------|--|---|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|--------|
|  | Nigéria                       | (C) Infra-national <sup>RL</sup>                           | (G) Données des finances publiques        | 24.3        | 25.5     | 40.0          | 15.1                  | 17.5                          | 6.3 ▼                   | 80.0 ▲ |
|  | Rwanda                        | (G) Politique d'ouverture des données                      | (G) Protection des données                | 24.8        | 32.6     | 39.1          | 11.1                  | 24.1                          | 63.0 ▲                  | 72.0 ▲ |
|  | Sénégal                       | (A) Données des marchés publics                            | (A) Données sur le budget et les dépenses | 12.1        | 14.8     | 18.9          | 6.4                   | 9.6                           | 0.0 ■                   | 0.0 ■  |
|  | Sierra Leone                  | (G) Finances politiques                                    | (G) Protection des données                | 12.1        | 13.1     | 22.8          | 5.3                   | 11.6                          | 0.0 ▼                   | 28.0 ▼ |
|  | Togo                          | (G) Déclarations d'actifs                                  | (G) Données des finances publiques        | 14.6        | 15.5     | 20.8          | 10.6                  | 11.6                          | 0.0 ■                   | 14.0 ▼ |
|  | Tunisie                       | (G) Politique d'ouverture des données <sup>RL</sup>        | (G) Données des finances publiques        | 23.1        | 35.0     | 37.5          | 7.5                   | 10.8                          | 80.0 ▲                  | 40.0 ■ |
|  | Ouganda                       | (G) Gestion des données                                    | (G) Propriété effective                   | 31.4        | 39.6     | 37.8          | 23.0                  | 19.3                          | 0.0 ■                   | 50.0 ▲ |
|  | Afrique du Sud                | (A) Données sur le budget et les dépenses <sup>RL GL</sup> | (C) Fonction publique                     | 30.4        | 38.5     | 37.3          | 22.6                  | 9.2                           | 0.0 ■                   | 28.0 ▼ |

[G = Pilier gouvernance ; C = Pilier capacités ; A = Pilier disponibilité ; U = Pilier utilisation et impact ; RL = Leader régional (parmi les scores les plus élevés au niveau régional) pour cet indicateur ; GL = Leader mondial (parmi les scores les plus élevés au niveau mondial) pour cet indicateur. ▲ = score GDB supérieur au score ODB sur l'indicateur connexe ; ▼ = score GDB inférieur au score ODB sur l'indicateur connexe ; ■ = changement de score inférieur à 1 pt ; □ = pas de données à comparer.]  
 {.regional-table-legend}

## Perspectives des carrefours : Institut de recherche sur le développement local (Afrique subsaharienne)

# LDRI

Local Development  
Research Institute

Par Keziah Munyao

### Profil régional

L'Afrique subsaharienne est un continent diversifié avec une population de plus d'un milliard d'habitants. Elle dispose de ressources humaines et naturelles susceptibles de générer une croissance équitable et d'atténuer la pauvreté. Alors que le continent continue de réaliser des progrès significatifs, malheureusement, la région possède toujours le taux de mortalité le plus élevé au monde pour les moins de 5 ans ; près d'un quart de la population souffre de malnutrition ; 42 % de la population ne dispose pas d'un approvisionnement en eau de base ; 72 % n'a pas d'assainissement de base ; et 58 % des 15 à 17 ans ne vont pas à l'école<sup>[28]</sup>. La région ne parvient pas à atteindre l'égalité de genre et se classe au dernier rang parmi les régions sur l'indice de genre des ODD 2019<sup>[29 en]</sup> raison des taux élevés de mortalité maternelle et de l'accès inégal à l'eau et à l'électricité. L'Afrique subsaharienne est également confrontée aux taux de croissance démographique les plus élevés et à une croissance rapide de la population urbaine, qui devrait atteindre 1,3 milliard d'ici 2050. En dépit des progrès significatifs réalisés dans le domaine de la santé mondiale, plus particulièrement dans la lutte contre le VIH/sida, de nombreuses tendances plus larges risquent de mettre à rude

épreuve les systèmes d'alimentation, de santé, d'éducation et d'infrastructures publiques déjà surchargés, ainsi que de submerger les marchés de l'emploi qui ne sont pas en mesure de répondre à la demande d'emploi de populations de jeunes en croissance.

Le continent est également une région caractérisée par des disparités politiques et socioéconomiques dramatiques. Dans quelques pays, les gouvernements autocratiques, les troubles civils, la corruption, la médiocrité des institutions et la lenteur des systèmes politiques continuent d'entraver les initiatives de réforme. Les femmes et les jeunes autonomisés continuent de s'impliquer davantage dans la politique au sein de l'Union africaine et par son intermédiaire. Les organisations internationales ont été en mesure d'apporter leur aide en élargissant la coopération régionale, en appuyant la gouvernance démocratique et du secteur de la sécurité dans l'ensemble des pays subsahariens, en promouvant un discours constructif entre les partis et en renforçant les capacités des partis politiques. Même si les pays d'Afrique subsaharienne ont commencé à s'ouvrir à la démocratie il y a trois décennies, la région est encore caractérisée par une grande variété de systèmes politiques. Les démocraties fragiles sont souvent confrontées à de multiples problèmes et failles, et elles partagent des frontières avec certains des pays les moins démocratiques du monde. Les élections multipartites, quasiment inexistantes en 1990, sont désormais la norme, mais elles n'aboutissent encore que rarement à des transferts de pouvoir. Les tendances récentes à la récession démocratique ne sont pas passées inaperçues dans la région, mais ont affecté différemment des nations distinctes. Certains changements démocratiques imprévus se sont produits parallèlement à un déclin démocratique plus large. En outre, alors que la pandémie de coronavirus frappe la région moins durement que d'autres régions du monde, son influence sur la démocratie et les droits de l'homme est considérable.

## Bilan des missions sur le terrain

La première édition du Baromètre a été une expérience d'apprentissage pour le pôle régional de l'Afrique subsaharienne, avec certains des thèmes importants observés au cours de l'étude sur le terrain, notamment :

- Il existe peu de preuves accessibles au public que les gouvernements investissent dans les compétences en matière de compréhension des données dans la fonction publique en dehors des unités statistiques ou des bureaux nationaux de statistique ;
- Le contenu des sites Web des organismes, des ministères et des départements est très fragmenté, ce qui fait qu'il est difficile pour le public et les autres parties prenantes de trouver des données pertinentes ;
- Dans certains pays, les plates-formes gouvernementales ne disposent pas de moyens faciles de diffuser l'information à un public plus étendu.

En plus de ces questions liées à la création et à la diffusion des données, nous avons constaté que la faible connaissance des données pour l'interprétation des données publiques était un problème courant dans la région. Il était, par exemple, difficile pour les individus de comprendre les ensembles de données sur les systèmes d'enregistrement des faits d'état civil et de statistiques de l'état civil. Ceci met en évidence les défis à la fois du côté de « l'offre » et du côté de la « demande » des données pour le bien public. Il soulève également des questions sur la mesure dans laquelle les données sont gérées de manière à soutenir l'analyse et la réutilisation, tant à l'intérieur qu'à l'extérieur du gouvernement, plutôt que de simplement permettre la publication de rapports statistiques. La plupart des données que nous avons trouvées dans le cadre de l'étude sur le terrain n'étaient pas adaptées à une analyse continue, et elles ne semblaient pas non plus liées à des processus au sein du gouvernement qui permettraient de les tenir à jour.

Les travaux sur le terrain ont également mis en évidence des lacunes en ce qui concerne la protection des données ou les normes de confidentialité dans un certain nombre de pays, même lorsque des efforts sont en cours pour promouvoir une utilisation et une ouverture plus larges des données. L'absence de cadres

juridiques solides parallèlement aux nouvelles avancées technologiques paraît être une préoccupation croissante, en particulier dans les pays où il n'existe pas de cadres pour superviser l'utilisation des technologies émergentes telles que l'intelligence artificielle (IA).

Il semble enfin y avoir une difficulté avec les gouvernements locaux lorsqu'il s'agit de rendre les données disponibles et ouvertes au public, car l'enquête montre qu'il y a eu un mouvement minimal sur le front des données ouvertes et du gouvernement ouvert au niveau régional ou municipal au-delà de quelques exemples exceptionnels.

## Orientations futures

L'étude du Baromètre s'est avérée une activité cruciale pour la région, car elle pourrait aider les pays à gagner en confiance, à prendre des décisions sur les stratégies en matière de données et à être plus proactifs dans l'élaboration de politiques et d'initiatives fondées sur les données qui peuvent avoir un impact plus important sur les gens à l'avenir. Le Baromètre peut aider les pays à mieux comprendre la dynamique et la complexité de la génération, de la publication et de l'analyse des données, et à rassembler les avis d'experts et les points de vue des fonctionnaires gouvernementaux pour comprendre les mesures pratiques à entreprendre. En mettant l'accent sur l'apprentissage par les pairs constructif dans la région et partout dans le monde, nous espérons que les outils complets et équilibrés fournis par le Baromètre aideront à désamorcer les considérations politiques qui retardent ou empêchent si souvent l'ouverture des données en Afrique.

Lorsque nous examinons les prochaines étapes et les priorités, les pays africains doivent obtenir et fournir davantage de ressources pour fournir une assistance continue, le renforcement des capacités et le renforcement aux bureaux nationaux de statistique, qui sont un élément moteur des projets nationaux de données ouvertes. Soutenir la réutilisation nécessite des données accessibles, compréhensibles et utilisables. Pour y parvenir, nous devons surmonter divers obstacles, tels que les barrières technologiques, politiques, financières et socioculturelles. Compte tenu de ce qui précède, il faudrait davantage mettre l'accent sur le renforcement des capacités, l'apprentissage par

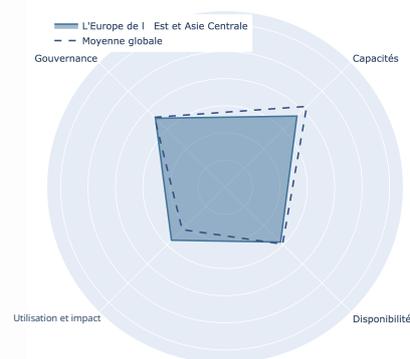
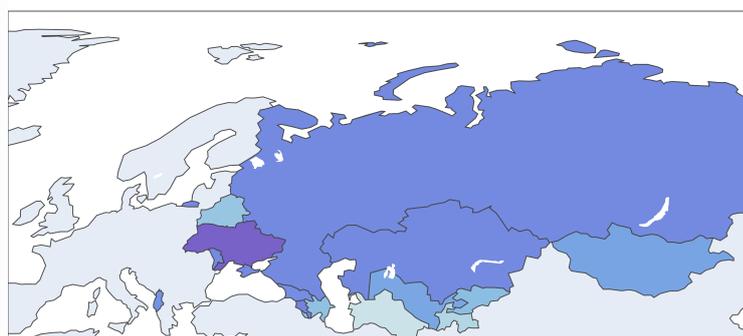
les pairs et la formation au sein de la communauté des universitaires, des journalistes et des autres praticiens qui peuvent stimuler la demande et l'utilisation des données.

Enfin, au fur et à mesure que le baromètre s'étendra, espérons-le, à d'autres domaines thématiques à l'avenir, et compte tenu de l'état actuel du système alimentaire mondial, qui est soumis à des tensions en raison des incertitudes climatiques et d'une population croissante, les

gouvernements d'Afrique subsaharienne devraient envisager de mettre un accent particulier sur le partage de données sur le secteur agricole pour la consommation publique, en soutenant les efforts de partage de données menés par les gouvernements et en s'assurant que les données agricoles de toutes les parties prenantes sont disponibles en tant que ressource pour le bien public.

## Europe de l'Est et Asie centrale

Le baromètre couvre 15 pays d'Europe orientale et d'Asie centrale. La région obtient des résultats légèrement inférieurs à la moyenne mondiale en ce qui a trait au pilier capacités, avec une grande variabilité dans la mesure dans laquelle les pays ont les compétences, les institutions et les libertés nécessaires pour utiliser les données pour le bien public. La région obtient des résultats légèrement supérieurs à la moyenne mondiale en ce qui concerne l'utilisation et l'impact, reflétant en particulier les utilisations des données dirigées par la société civile. Le renforcement de la gouvernance des données et l'adoption d'approches plus multipartites dans le but de promouvoir la gestion, la disponibilité et l'utilisation des données pour le bien public sont des domaines clés au développement de la région dans les années à venir.



| Score global | Gouvernance | Capacités | Disponibilité | Impact et utilisation |
|--------------|-------------|-----------|---------------|-----------------------|
| 32.79        | 35.9        | 37.05     | 28.62         | 27.65                 |

| Pays | Force comparative (régionale)                           | Faiblesse comparée (régionale)           | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|------|---|--|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|
|      | (A) Utilisation actuelle des terres <sup>RL GL</sup>    | (A) Émissions                            | 38.3         | 41.4        | 47.7     | 30.1          | 45.5                  | 40.0                          | 54.0                    |
|      | (G) Données des finances publiques <sup>RL GL</sup>     | (C) Initiative sur les données ouvertes  | 44.6         | 54.7        | 27.7     | 48.5          | 28.0                  | 0.0                           | 0.0                     |
|      | (A) Émissions   | (G) Déclarations d'actifs                | 21.8         | 26.7        | 29.5     | 15.1          | 9.6                   | 10.0                          | 0.0                     |
|      | (G) Données de consultation publique <sup>RL GL</sup>   | (G) Protection des données               | 19.5         | 12.7        | 28.4     | 19.7          | 13.7                  | 0.0                           | 0.0                     |
|      | (A) Émissions <sup>RL</sup>                             | (G) Données de consultation publique     | 40.3         | 34.5        | 40.1     | 45.5          | 28.2                  | 0.0                           | 36.0                    |
|      | (A) Données de performance RTI <sup>RL</sup>            | (G) Déclarations d'actifs                | 41.7         | 42.8        | 55.5     | 34.3          | 27.6                  | 50.0                          | 80.0                    |
|      | (G) Propriété effective                                 | (G) Données de consultation publique     | 23.5         | 24.5        | 28.3     | 19.9          | 25.0                  | 0.0                           | 35.0                    |
|      | (G) Déclarations d'actifs                               | (C) Fonction publique                    | 41.4         | 45.4        | 40.4     | 39.6          | 36.8                  | 36.0                          | 36.0                    |
|      | (A) Vulnérabilité <sup>RL</sup>                         | (A) Registre du commerce et des sociétés | 32.8         | 34.8        | 32.7     | 33.4          | 13.5                  | 54.0                          | 60.0                    |
|      | (C) Fonction publique <sup>RL</sup>                     | (A) Statistiques vitales                 | 41.7         | 41.5        | 52.3     | 34.8          | 50.8                  | 54.0                          | 36.0                    |
|      | (G) Politique d'ouverture des données                   | (G) Données de consultation publique     | 12.2         | 15.5        | 18.5     | 6.5           | 10.8                  | 37.8                          | 7.0                     |
|      | (C) Fonction publique                                   | (G) Données de consultation publique     | 6.4          | 8.2         | 15.5     | 0.0           | 5.6                   | 0.0                           | 0.0                     |
|      | (U) Diligence raisonnable d'entreprise <sup>RL GL</sup> | (A) Données de consultation publique     | 55.5         | 61.1        | 56.5     | 48.3          | 82.1                  | 80.0                          | 80.0                    |
|      | (C) Initiative sur les données ouvertes <sup>RL</sup>   | (G) Déclarations d'actifs                | 31.7         | 43.7        | 41.6     | 19.4          | 12.1                  | 68.8                          | 80.0                    |
|      | (G) Performance RTI <sup>RL</sup>                       | (A) Utilisation actuelle des terres      | 40.5         | 50.8        | 41.0     | 34.2          | 25.5                  | 54.0                          | 40.0                    |

## Perspectives des plateformes : IDFI (Partenariat du Centre et de l'est de l'Asie)



Par Teona Turashvili et David Eristavi

### Profil régional

Les pays d'Asie centrale et du partenariat oriental ont un dénominateur commun majeur, à savoir le passé soviétique, dont les conséquences sont encore assez évidentes en termes de gouvernance et d'administration. Ces pays partagent de nombreuses caractéristiques communes aux gouvernements post-soviétiques à des degrés divers, telles que des bureaucraties complexes, un leadership autoritaire, des infrastructures numériques sous-développées et l'alphabétisation, et des défis à la liberté de la presse. Malgré les similitudes, il existe cependant des écosystèmes de données très différents, avec des exemples assez développés comme l'Ukraine à une extrémité du spectre et des structures plus étouffées tel que le Turkménistan et le Bélarus à l'autre.

En ce qui a trait aux groupes marginalisés, la plupart des chercheurs n'ont pas identifié les minorités ethniques résidant dans leurs pays respectifs, comme la population ouzbèke au Tadjikistan, les Arméniens et Azéris de souche en Géorgie, divers groupes en Azerbaïdjan, etc. Pour les zones affectées par le conflit tel que la Géorgie et l'Ukraine, un thème commun était les populations isolées dans les zones rurales et les personnes déplacées à l'intérieur du pays. Les

groupes suivants ont plus généralement été mentionnés en tant que groupes potentiellement marginalisés dans le cas de plusieurs pays : les femmes, la communauté LGBTQ+, les Roms, les personnes handicapées, les personnes âgées et les populations rurales. Une tendance intéressante qui a été découverte a trait à la reconnaissance officielle des groupes marginalisés/vulnérables par l'État. Si certains pays font référence à ces groupes dans leur législation ou leurs documents d'orientation et s'emploient activement à relever les défis auxquels ils sont confrontés, d'autres ne semblent pas avoir pris de mesures officielles dans ce sens.

### Bilan des missions sur le terrain

La catégorie d'utilisation et d'impact du Baromètre semblait être la plus difficile pour la région, à l'exception de la bonne performance de l'Ukraine. Les chercheurs n'ont trouvé que des exemples sporadiques de cas d'utilisation de données, principalement dirigés par les médias et les organisations de la société civile. Ceci dénote un manque d'engagement à l'égard des données provenant du milieu universitaire et des acteurs du secteur privé dans de nombreux pays. Même lorsque les médias utilisaient des données, ceci reposait souvent sur la société civile en tant qu'intermédiaire. En l'absence d'ensembles de données et d'outils analytiques ouverts, la réalisation d'un traitement approfondi des données peut prendre trop de temps pour les médias à un rythme rapide et, pour cette raison, elles couvrent principalement les études et les rapports publiés par les OSC, en particulier ceux basés sur des données financières politiques et les déclarations de patrimoine des fonctionnaires.

En outre, dans plusieurs pays de la région, il n'existe pratiquement pas d'outils de contrôle public des fonctionnaires et des processus de gouvernance. Leurs systèmes politiques restent fermés à tout engagement sociopolitique significatif des citoyens et des acteurs de la société civile. Il y a un manque d'institutions véritablement indépendantes capables d'identifier, d'exposer et de mettre en évidence les échecs du gouvernement dans tous les pays autoritaires de la région. Les chercheurs nationaux de ces pays

ont rapporté que les rares militants qui s'élèvent contre le gouvernement dans leurs pays respectifs sont régulièrement victimes d'intimidation, de harcèlement et d'emprisonnement.

De toutes les catégories de données, les données sur les achats semblent être les plus disponibles dans la région. La plupart des pays interrogés avaient mis en place des portails d'approvisionnement novateurs et utilisaient activement la norme ouverte de données contractuelles (OCDS) à des degrés divers d'adhésion. En revanche, les données de lobbying paraissent être les plus difficiles dans la région, la plupart des pays ne disposant pas d'un cadre fonctionnel pour collecter ces données ou ne les rendant pas publiques. L'absence de données ventilées sur la vaccination est un autre problème qui se pose, en particulier dans la période de rétablissement post-pandémique. Bien que des statistiques globales soient disponibles dans la plupart des pays de la région, il y a un manque important d'ensembles de données ouvertes et lisibles par machine à cet égard.

Bien que cela puisse être vrai pour de nombreuses autres régions, il a été constaté que les données sur l'intégrité politique dans les pays du Partenariat oriental et d'Asie centrale manquaient d'interopérabilité. Certains pays progressent en termes d'amélioration de l'accès à la législation et aux pratiques en matière de données ouvertes, tandis que d'autres sont dans l'impasse et qu'aucune réforme fondamentale n'a été entreprise au cours des dernières années, tandis que d'autres sont même en recul.

## Orientations futures

Tout d'abord, les pays de la région ont besoin de cadres réglementaires solides sur l'accès aux données ouvertes, car la majorité d'entre eux ne disposent pas de normes communes en la matière. À cette fin, il est nécessaire d'avoir une forte volonté politique, qui garantira que les décideurs sont conscients de l'importance, des avantages et des facteurs clés pour des écosystèmes de données ouvertes solides dans le pays.

Un autre élément important est le renforcement des capacités de toutes les parties prenantes concernées par la gestion des données ouvertes, qui comprend la collecte, le traitement, la publication et l'utilisation des données. En particulier, les qualifications des fonctionnaires en matière de gestion des données et de données ouvertes doivent également être hiérarchisées, étant donné qu'ils sont responsables de la production ou de la collecte de la plupart des données publiques. Parallèlement, le renforcement des capacités avec d'autres parties prenantes, telles que la société civile, les médias et les entreprises, est nécessaire pour accroître l'impact généré par l'utilisation de données ouvertes.

Enfin, il est d'une importance cruciale que les différentes parties prenantes établissent des partenariats les unes avec les autres sur les questions de collecte, de publication et d'utilisation des données ouvertes. Une telle coopération multisectorielle et multipartite accroîtra considérablement l'impact des initiatives en matière de données pour le bien public. Différentes parties prenantes ayant des antécédents et une expérience différente peuvent mieux combiner leurs efforts, et potentiellement concevoir des services innovants et de nouveaux produits ayant une incidence positive sur le bien-être des citoyens. À ce stade, les avantages pour la société civile et les médias sont déjà évidents, mais le potentiel économique des données ne semble pas être pleinement généré ou encouragé à l'heure actuelle. Un recentrage sur les aspects économiques des données ouvertes pourrait encourager et jouer un rôle crucial en facilitant et en faisant progresser toutes les pratiques en matière de données ouvertes dans la région. En outre, mettre l'accent sur les aspects économiques des données ouvertes peut aider à convaincre les décideurs d'inscrire les réformes liées aux données ouvertes à l'ordre du jour politique, et peut engager le secteur privé dans un dialogue avec la société civile et le secteur public sur la nécessité d'améliorer les systèmes de gestion des données et d'améliorer l'accès aux données publiques pour le bien public.

## Perspective du Hub : Open Data Kosovo (Kosovo et Albanie)



Par Dafina Olluri

### Profil régional

Le Kosovo et l'Albanie sont deux nouvelles démocraties au passé historique difficile. L'Albanie a souffert d'un long état de transition qui a commencé au début des années 90 et a suivi un demi-siècle de dictature communiste, période pendant laquelle le gouvernement a eu un contrôle total sur l'économie et a révoqué les libertés civiles. La République du Kosovo, qui existe en tant que pays depuis environ treize ans, a également un passé de dictature communiste, ainsi qu'une économie qui se remet des destructions de la guerre. Les deux pays sont actuellement des sociétés démocratiques et ont des constitutions modernes en vigueur. Par l'entremise de ces deux pays, ils expriment leur détermination à bâtir des pays libres, démocratiques et épris de paix pour tous les citoyens ; leur engagement à l'égard du développement d'un État qui assure la liberté civile, l'égalité et les droits de chaque citoyen devant la loi ; et leur engagement à promouvoir le bien-être économique et la prospérité sociale.

### Bilan des missions sur le terrain

Le processus d'ouverture des données en Albanie et au Kosovo s'est intensifié au cours des quatre à cinq dernières années, principalement en raison des facteurs venant de la société civile, où les

organisations non gouvernementales telles que Open Data Kosovo ont joué un rôle crucial. Les portails gouvernementaux ouverts sont gérés par les agences gouvernementales respectives pour la société de l'information, mais ne sont pas suffisamment remplis avec des données provenant d'autres entités publiques, ne contenant finalement qu'un choix limité d'ensembles de données. Cependant, les données ouvertes sont publiées par diverses entités publiques et agences gouvernementales d'État sous différents formats, et celles-ci agissent comme une source précieuse d'informations et de prise de décision pour les citoyens et les fonctionnaires. Les données statistiques sur les marchés publics, les finances publiques, l'intégrité politique (données publiées par l'Agence de lutte contre la corruption) sont assez complètes et ouvertes, ce qui conduit à un pays plus transparent. Il reste encore, malgré cela, beaucoup à faire pour faire participer le secteur privé, car les exemples de données ouvertes ou partagées provenant du secteur privé sont encore très rares.

En plus des données ouvertes, nous prenons note que le cadre juridique quant à la protection des données et l'accès à l'information publique est assez avancé et conforme aux exigences de l'Union européenne. Ces données, disponibles et facilement accessibles en ligne gratuitement, contiennent plus de dix ans de déclarations, fournissant une base de contrôle pour suivre la richesse des politiciens du Kosovo. D'autre part, les données concernant le régime foncier ne sont pas disponibles en tant que données ouvertes et, par conséquent, rien ne prouve qu'elles sont utilisées pour influencer les politiques dans l'intérêt d'un régime foncier et d'une utilisation des terres équitables et inclusives. En plus, le pays ne dispose pas d'un système d'information sanitaire permettant d'héberger les données des citoyens de manière centralisée.

### Orientations futures

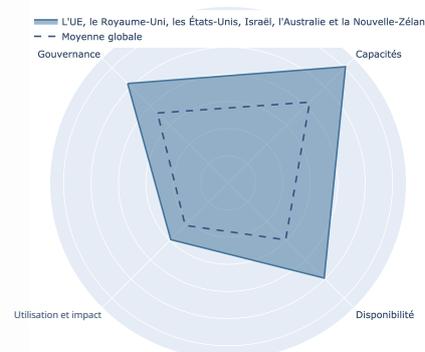
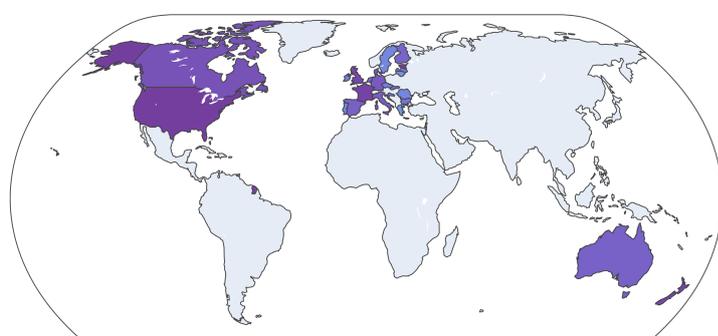
Le Kosovo et l'Albanie, malgré leurs efforts pour ouvrir toutes les données publiques, en sont encore aux premiers stades de ce processus et une exigence juridique plus stricte sera vitale pour contribuer à stimuler positivement le processus d'ouverture des données et soutenir l'utilisation des données pour le bien public. Il convient plus

particulièrement d'accorder une attention à la manière dont le respect de la gouvernance des données, des données ouvertes et des règles de partage des données sera assuré. En ce qui concerne l'utilisation plus large des données pour

le bien public, des efforts spécifiques devraient être faits pour sensibiliser le public aux données produites dans le secteur privé et à l'importance de leur utilisation et de leur ouverture aux parties prenantes des secteurs public et privé.

## L'Union européenne, l'Amérique du Nord, Israël, l'Australie et la Nouvelle-Zélande

Le baromètre couvre 20 pays de l'Union européenne, qui ont été analysés aux côtés du Royaume-Uni, des États-Unis, du Canada, de l'Australie, de la Nouvelle-Zélande et d'Israël dans un groupe défini par des niveaux généralement élevés de revenu national et des capacités liées aux données. Des scores de gouvernance comparables masquent certaines différences significatives entre les pays, certains d'entre eux se prononçant fortement sur les règles sectorielles en matière de divulgation des données, tandis que d'autres se prononcent davantage sur l'applicabilité universelle des cadres de protection des données. La poursuite des travaux sur l'interopérabilité des données, l'amélioration des capacités infranationales, la prise en compte des questions d'équité et d'inclusion dans la collecte et l'utilisation des données, et le renforcement des données ouvertes et du partage des données liées à l'action pour le climat devraient être des domaines prioritaires pour les pays de ce groupe.



| Score global | Gouvernance | Capacités | Disponibilité | Impact et utilisation |
|--------------|-------------|-----------|---------------|-----------------------|
| 52.08        | 51.8        | 60.55     | 49.59         | 29.56                 |

| Pays               | Force comparative (régionale)  | Faiblesse comparée (régionale)                 | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|--------------------|--|--|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|
| Australie          | (A) Biodiversité   | (G) Données de consultation publique           | 55.5         | 52.7        | 67.7     | 53.3          | 25.7                  | 81.0                          | 80.0                    |
| Bulgarie           | (U) Influence de la politique en faveur de l'égalité des sexes et de l'inclusion | (A) Registre du commerce et des sociétés       | 49.7         | 59.0        | 54.4     | 40.7          | 45.0                  | 70.0                          | 70.0                    |
| Canada             | (A) Données sur le lobbying <sup>RL GL</sup>                                     | (G) Données de consultation publique           | 60.8         | 59.5        | 68.0     | 62.2          | 12.8                  | 90.0                          | 80.0                    |
| République Tchèque | (G) Performance RTI  | (G) Données de consultation publique           | 45.0         | 45.1        | 59.7     | 37.9          | 31.4                  | 80.0                          | 72.0                    |
| Allemagne          | (C) Infra-national   | (A) Registre du commerce et des sociétés       | 58.1         | 61.1        | 68.8     | 53.6          | 18.2                  | 50.0                          | 81.0                    |
| Danemark           | (A) Propriété effective <sup>RL GL</sup>   | (G) Politique d'ouverture des données          | 58.2         | 45.2        | 67.3     | 65.5          | 23.6                  | 0.0                           | 51.0                    |
| Espagne            | (C) Infra-national <sup>RL GL</sup>  | (G) Données des finances publiques             | 55.8         | 58.5        | 74.4     | 46.6          | 20.6                  | 70.0                          | 80.0                    |
| Estonie            | (C) Interopérabilité de l'intégrité politique <sup>RL GL</sup>                   | (A) Capacité du système de santé en temps réel | 67.4         | 61.9        | 91.2     | 60.7          | 35.1                  | 50.0                          | 90.0                    |
| Finlande           | (A) Biodiversité <sup>RL GL</sup>  | (A) Capacité du système de santé en temps réel | 54.5         | 48.3        | 68.9     | 52.3          | 37.9                  | 30.0                          | 80.0                    |
| Ranska             | (C) Soutien gouvernemental à la réutilisation <sup>RL</sup>                      | (C) Fonction publique                          | 66.2         | 69.7        | 70.5     | 63.6          | 41.3                  | 80.0                          | 100.0                   |
| Royaume-Uni        | (A) Propriété effective  | (A) Capacité du système de santé en temps réel | 64.5         | 64.1        | 68.9     | 64.2          | 45.1                  | 50.0                          | 63.0                    |
| Grèce              | (G) Propriété effective <sup>RL</sup>  | (G) Finances politiques                        | 36.6         | 37.3        | 47.4     | 32.3          | 12.0                  | 60.0                          | 63.0                    |
| Croatie            | (A) Données de performance RTI   | (G) Politique d'ouverture des données          | 47.9         | 55.4        | 42.8     | 47.3          | 28.8                  | 0.0                           | 45.0                    |

| Pays  | Force comparative (régionale) | Faiblesse comparée (régionale)                        | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|---|-------------------------------|---|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|
|  | Irlande                       | (G) Gestion des données                               | 46.0         | 47.2        | 66.5     | 36.7          | 12.8                  | 72.0 ▲                        | 90.0 ■                  |
|  | Israël                        | (A) Registre du commerce et des sociétés              | 42.1         | 38.0        | 49.6     | 41.3          | 35.5                  | 53.6 ■                        | 76.5 ▲                  |
|  | Italie                        | (A) Propriété effective                               | 56.5         | 62.3        | 54.5     | 56.4          | 26.9                  | 50.0 ▼                        | 60.0 ■                  |
|  | Lituanie                      | (A) Capacité du système de santé en temps réel        | 37.3         | 36.8        | 49.7     | 31.9          | 23.0                  | 36.0 □                        | 56.0 □                  |
|  | Lettonie                      | (A) Propriété effective                               | 49.2         | 43.5        | 50.4     | 53.9          | 35.1                  | 81.0 ▲                        | 50.0 ▲                  |
|  | Malte                         | (G) Données des finances publiques                    | 36.5         | 45.8        | 44.8     | 27.0          | 18.0                  | 0.0 □                         | 16.0 □                  |
|  | Pays-Bas                      | (U) Diligence raisonnable d'entreprise <sup>RL</sup>  | 54.0         | 49.3        | 70.4     | 48.8          | 45.8                  | 42.5 ▼                        | 90.0 ▲                  |
|  | Nouvelle-Zélande              | (A) Régime foncier <sup>RL GL</sup>                   | 65.6         | 63.7        | 62.4     | 69.8          | 53.7                  | 81.0 ■                        | 70.0 ▼                  |
|  | Portugal                      | (C) Fonction publique                                 | 41.9         | 43.8        | 50.2     | 38.2          | 17.5                  | 70.0 ▲                        | 50.0 ■                  |
|  | Roumanie                      | (G) Données de consultation publique <sup>RL GL</sup> | 43.0         | 55.1        | 44.8     | 35.6          | 19.1                  | 30.0 □                        | 51.0 □                  |
|  | Slovaquie                     | (A) Capacité du système de santé en temps réel        | 50.9         | 54.6        | 57.9     | 45.8          | 34.0                  | 63.0 ■                        | 80.0 ■                  |
|  | Suède                         | (A) Biodiversité                                      | 42.8         | 32.0        | 58.6     | 43.6          | 20.2                  | 0.0 ▼                         | 80.0 ■                  |
|  | États-Unis                    | (A) Données sur le lobbying                           | 68.0         | 56.7        | 64.4     | 80.0          | 49.4                  | 42.0 ▼                        | 54.0 ▼                  |

## Perspectives du Hub : Access Info Europe (Europe et au-delà)



*Par Helen Darbishire et Marta Morcuende*

### Profil régional

Access Info Europe a servi de plaque tournante régionale pour un groupe de 27 pays, la plupart d'entre eux dans la région européenne. Parmi ces 27 pays, 20 sont des États membres de l'Union européenne (UE) et partagent des politiques communes dans de nombreux domaines, tels que la protection des données. Le Royaume-Uni, également situé dans la région européenne et inclus dans l'étude, a officiellement quitté l'Union européenne en 2020, mais il partage également de nombreuses normes et politiques avec les États membres de l'UE qui ont été approuvées avant son départ. À l'extérieur de la région européenne, Access Info a coordonné la recherche dans plusieurs pays de différentes régions du monde, tel que l'Amérique du Nord (Canada et États-Unis), l'Océanie (Australie et Nouvelle-Zélande), l'Asie occidentale (Israël) et l'Asie de l'Est (République de Corée). La Corée est comprise dans les tableaux et statistiques « Asie du Sud et de l'Est », mais tous ces autres pays sont traités dans le Baromètre comme faisant partie du groupe « UE, Amérique du Nord + ».

Ces pays sont généralement parmi les plus développés au monde, tant en termes de développement économique que de développement démocratique. Ce sont des pays qui ont mené le mouvement des données

ouvertes, et les pays de l'Union européenne ont ouvert la voie à des normes communes sur l'ouverture des informations du secteur public dans le cadre de la garantie de conditions de concurrence équitables pour la réutilisation des données, avec un nombre important et croissant d'entreprises utilisant des données publiques, ainsi que l'utilisation par la société civile et les journalistes d'investigation.

Il s'agit de pays jouissant d'une solide réputation en matière de bonne gouvernance et qui réussissent à lutter contre la corruption, bien qu'il existe des variations importantes dans la région, tel qu'il peut être perçu, par exemple, dans l'indice de perception de la corruption, qui affiche de meilleurs résultats dans le nord et l'ouest du continent européen que dans les pays d'Europe méridionale et orientale. Ceci n'a pas empêché de multiples scandales de corruption reliés au blanchiment d'argent, au financement offshore, aux avoirs volés dans d'autres régions du monde et aux marchés publics irréguliers pendant la pandémie de Covid.

Ces pays sont tous membres du Partenariat pour un gouvernement ouvert, ce qui signifie que tous les deux ou quatre ans, ils doivent soumettre un plan d'action co-créé avec la société civile décrivant des engagements concrets pour améliorer la transparence, la responsabilité et la participation du public au gouvernement.

### Bilan des missions sur le terrain

Les pays de ce groupe régional obtiennent, dans l'ensemble, des résultats supérieurs à la moyenne mondiale. Ceci comprend les quatre piliers clés du baromètre, ainsi que tous les modules thématiques. La région obtient les scores les plus élevés en ce qui a trait au pilier capacités, avec les scores les plus bas dans le pilier utilisation et impact. L'intégrité politique est le module thématique le moins bien coté, ce qui montre qu'un nombre comparativement moins élevé de pays respectent la barre supérieure de la gouvernance et de la publication des données dans l'ensemble des sujets couverts par ce module.

Cette région obtient de bons résultats quant à l'existence de cadres de protection des données, grâce en partie aux règles de protection des données obligatoires de l'Union depuis 1995 et au règlement général sur la protection des données renforcées de 2016, qui est entré en vigueur dans tous les pays de l'Union en mai 2018. Il existe dans la plupart des pays de la région des cadres juridiques exigeant la collecte et la publication de données sur le financement politique et les déclarations de patrimoine, bien qu'il soit important de noter que la plupart des données sur les déclarations de patrimoine ne sont pas disponibles sous une forme lisible à la machine.

Les autres ensembles de données qui étaient généralement disponibles dans cette région étaient ceux sur les vaccinations contre la COVID-19, les déclarations d'actifs, les marchés publics, les émissions, les finances politiques, les données budgétaires et de dépenses et les statistiques de l'état civil. L'enquête a aussi révélée qu'il existe peu de réglementations exigeant la collecte et la publication de données sur les performances et le lobbying du droit à l'information, et l'absence de réglementation se traduit directement par un manque de disponibilité de ces informations dans la pratique.

Un certain nombre d'autres observations peuvent être tirées des données relatives à ce groupe de pays :

- Il est important d'avoir des règlements exigeant la collecte et/ou la publication pour garantir que plus de données sont disponibles. Dans la plupart des cas, l'absence d'un cadre juridique exigeant la collecte/publication de données se traduit par un manque de disponibilité des données au sein de la pratique.
- Les données sur les bénéficiaires effectifs ne sont toujours pas disponibles dans de nombreux pays, et lorsqu'elles le sont, elles ne sont généralement pas disponibles gratuitement, sous des licences ouvertes ou dans des formats lisibles à la machine. Les informations sur l'entreprise sont disponibles dans plus de pays, mais elles ne sont pas entièrement gratuites dans beaucoup d'entre eux.
- Lorsque des informations sur les entreprises et des données sur les marchés publics sont disponibles, il est

clairement établi qu'elles sont bien utilisées par un éventail de parties prenantes.

- Bien que le changement climatique soit l'un des problèmes les plus urgents et que la région ait une grande capacité à produire et à utiliser des données, il existe encore étonnamment peu de données disponibles sur la vulnérabilité climatique telles que des informations sur les risques naturels futurs, les phénomènes météorologiques extrêmes ou la variabilité climatique.

## Orientations futures

Les pays de ce groupe obtiennent de meilleurs résultats que la moyenne mondiale, mais beaucoup reste encore à faire.

Un écart entre les pays d'Europe septentrionale/occidentale et ceux d'Europe méridionale/orientale peut encore être perçu. Les politiques en matière de données ouvertes appliquées dans tous les pays de l'UE, telles que la directive sur les données ouvertes, si elles sont correctement mises en œuvre, égaliseraient les chances et amélioreraient le paysage des données dans la région. Un certain nombre de cadres juridiques spécifiques font également encore largement défaut, y compris les réglementations en matière de lobbying : l'adoption de règles exigeant la collecte de données sur les activités de lobbying est une priorité afin de permettre le suivi de l'influence du lobbying sur la prise de décision. Lorsque des cadres sont déjà largement en place, les pays doivent alors s'attacher à améliorer la disponibilité, l'ouverture et la facilité d'utilisation des principaux ensembles de données, notamment les registres des sociétés et des bénéficiaires effectifs. Ces données se sont avérées très utiles pour les utilisateurs, et des preuves de leur utilisation ont été trouvées dans près de 20 pays, leur utilisation étant essentielle pour lutter efficacement contre la corruption et le blanchiment d'argent.

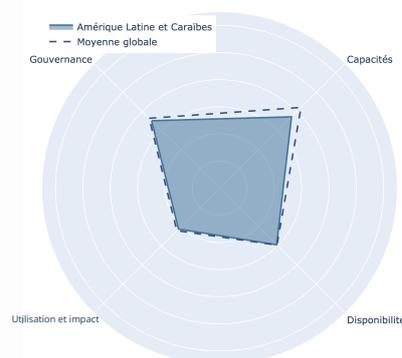
Les gouvernements doivent également prendre davantage de mesures pour la promotion de la réutilisation des données publiques afin d'accroître les avantages que les données apportent aux sociétés et aux économies. Pour ce faire, les données devraient toujours être mises à disposition gratuitement, dans le cadre de licences ouvertes qui permettent tout type de réutilisation,

dans des formats lisibles par machine et disposant d'ensembles de données lisibles par machine dans leur ensemble, diffusés en temps utile, mis à jour, et avec des données historiques disponibles pour permettre aux utilisateurs de suivre l'évolution au fil du temps, disponibles dans toutes les langues du pays, et disposant d'outils accessibles et ouverts pour aider les utilisateurs à explorer les données. Les gouvernements devraient également promouvoir une plus grande interopérabilité entre les ensembles de données

de manière à libérer véritablement tout le potentiel que les données ouvertes peuvent apporter à la société. Ceci dit, la publication de données ouvertes n'est pas suffisante. Des stratégies pertinentes en matière de données doivent être mises en œuvre en cours de route, en veillant à ce que les fonctionnaires soient correctement formés et à ce qu'il y ait une promotion active des avantages de la publication et de la réutilisation des données.

## Amérique latine et Caraïbes

Le Baromètre couvre 23 pays d'Amérique latine et des Caraïbes. Malgré des scores moyens inférieurs à la moyenne mondiale en matière de gouvernance et de capacités, la région correspond à la moyenne mondiale pour ce qui est de la disponibilité, de l'utilisation et de l'impact des données, ce qui reflète, en particulier, le rôle des initiatives de données ouvertes menées par les communautés dans l'appui à la publication et à l'utilisation des données. Le développement et l'intégration d'approches participatives et inclusives en matière de politique des données, ainsi que l'élaboration de cadres plus solides pour régir le partage des données et l'utilisation algorithmique des données, devraient figurer à l'ordre du jour de la région dans les années à venir, parallèlement aux travaux visant à renforcer les initiatives sectorielles en matière de données ouvertes grâce à l'apprentissage par les pairs.



| Score global | Gouvernance | Capacités | Disponibilité | Impact et utilisation |
|--------------|-------------|-----------|---------------|-----------------------|
| 32.71        | 35.16       | 37.18     | 29.49         | 21.27                 |

| Pays                   | Force comparative (régionale)                               | Faiblesse comparée (régionale)            | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|------------------------|---|---|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|
| Argentine              | (G) Finances politiques <sup>RL</sup>                       | (A) Émissions                             | 50.4         | 56.4        | 56.0     | 44.8          | 30.3                  | 72.0 ■                        | 72.0 ■                  |
| Bahamas                | (G) Données de consultation publique <sup>RL</sup>          | (C) Initiative sur les données ouvertes   | 23.9         | 31.4        | 28.5     | 17.1          | 12.4                  | 63.0 ▲                        | 0.0 ■                   |
| Belize                 | (G) Données de consultation publique <sup>RL</sup>          | (A) Données sur le budget et les dépenses | 24.4         | 30.1        | 31.3     | 17.3          | 14.6                  | 36.0 ▲                        | 35.0 ▲                  |
| Bolivie                | (A) Utilisation actuelle des terres                         | (G) Données de consultation publique      | 22.0         | 19.2        | 31.1     | 19.6          | 13.4                  | 63.0 ▲                        | 33.6 ■                  |
| Brésil                 | (A) Émissions <sup>RL</sup>                                 | (G) Finances politiques                   | 58.0         | 61.6        | 49.5     | 62.0          | 41.5                  | 72.0 ▲                        | 36.0 ▼                  |
| Chili                  | (A) Données sur le lobbying <sup>RL</sup>                   | (G) Performance RTI                       | 52.9         | 48.7        | 50.0     | 59.2          | 35.4                  | 50.4 ▲                        | 70.0 ■                  |
| Colombie               | (C) Soutien gouvernemental à la réutilisation <sup>RL</sup> | (G) Performance RTI                       | 53.8         | 57.1        | 66.3     | 45.8          | 37.8                  | 80.0 ■                        | 80.0 ■                  |
| Costa Rica             | (A) Émissions   | (G) Données de consultation publique      | 34.5         | 35.8        | 37.5     | 33.7          | 15.1                  | 37.8 ▼                        | 45.0 ■                  |
| République dominicaine | (G) Données de consultation publique <sup>RL</sup>          | (A) Utilisation actuelle des terres       | 35.2         | 31.7        | 40.3     | 35.7          | 24.0                  | 54.0 ▼                        | 80.0 ■                  |
| Équateur               | (A) Données de consultation publique                        | (G) Protection des données                | 34.6         | 37.3        | 34.2     | 34.4          | 17.7                  | 76.5 ■                        | 56.0 ▼                  |
| Guatemala              | (A) Données sur le budget et les dépenses                   | (G) Données de consultation publique      | 18.8         | 19.4        | 21.3     | 17.1          | 16.9                  | 44.1 ▼                        | 13.6 ▼                  |
| Guyane                 | (C) Fonction publique                                       | (G) Données de consultation publique      | 11.2         | 5.9         | 27.3     | 6.2           | 8.0                   | 0.0 ■                         | 24.0 ▲                  |
| Honduras               | (A) Données de performance RTI                              | (G) Données de consultation publique      | 24.9         | 28.0        | 28.8     | 21.3          | 17.2                  | 0.0 ▼                         | 28.0 ▼                  |
| Haïti                  | (A) Utilisation actuelle des terres                         | (G) Données de consultation publique      | 8.0          | 2.8         | 11.9     | 9.2           | 10.4                  | 0.0 ▼                         | 0.0 ▼                   |

| Pays  | Force comparative (régionale)   | Faiblesse comparée (régionale)           | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|---|---|--|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|
|  | Jamaïque<br>(G) Déclarations d'actifs RL                              | (G) Politique d'ouverture des données    | 31.0         | 39.5        | 30.7     | 25.2          | 29.6                  | 0.0 ▼                         | 25.5 ▼                  |
|  | Sainte-Lucie<br>(G) Politique d'ouverture des données RL              | (G) Données de consultation publique     | 21.3         | 30.1        | 32.5     | 9.2           | 16.0                  | 81.0 ▲                        | 34.0 ■                  |
|  | Mexique<br>(A) Vulnérabilité RL                                       | (A) Émissions                            | 50.6         | 55.1        | 46.9     | 50.6          | 40.3                  | 80.0 ▲                        | 28.8 ▼                  |
|  | Panama<br>(G) Données des finances publiques RL                       | (G) Données de consultation publique     | 34.6         | 36.0        | 46.4     | 28.9          | 12.7                  | 63.0 ■                        | 72.0 ■                  |
|  | Pérou<br>(A) Capacité du système de santé en temps réel RL            | (G) Données de consultation publique     | 37.6         | 38.3        | 37.6     | 39.5          | 13.4                  | 54.0 ▼                        | 28.0 ▼                  |
|  | Paraguay<br>(G) Données de consultation publique RL GL                | (A) Statistiques vitales                 | 33.5         | 43.7        | 30.4     | 28.6          | 27.5                  | 34.0 ▼                        | 16.8 ▼                  |
|  | Salvador<br>(U) Utilisations de l'IP à des fins de responsabilisation | (G) Données de consultation publique     | 13.4         | 12.6        | 22.3     | 9.0           | 13.6                  | 0.0 ▼                         | 0.0 ▼                   |
|  | Trinité-et-Tobago<br>(G) Données de consultation publique RL GL       | (G) Données des finances publiques       | 22.4         | 25.3        | 28.2     | 17.9          | 13.5                  | 0.0 ▼                         | 13.6 ▼                  |
|  | Uruguay<br>(C) Soutien gouvernemental à la réutilisation              | (A) Registre du commerce et des sociétés | 55.2         | 62.8        | 66.2     | 46.1          | 27.9                  | 80.0 ■                        | 72.0 ■                  |

## Perspectives du Hub : Iniciativa Latinoamericana de Datos Abiertos - ILDA (Amérique latine)



Par Fabrizio Scrollini et David Zamora

### Profil régional

L'Amérique latine est la région du monde la plus inégalitaire. Les problèmes auxquels la région est confrontée sont complexes, du fait de la faiblesse des institutions politiques, de la mauvaise gestion économique et de l'inefficacité et de l'insuffisance des politiques sociales et d'innovation.

Néanmoins, L'Amérique latine est également un endroit où l'innovation sociale et commerciale prospère et où un ensemble de nouvelles « licornes numériques » a émergé au cours des dix dernières années. La plupart des gouvernements latino-américains ont rejoint l'Open Government Partnership (OGP), une alliance d'organisations gouvernementales et de la société civile, où beaucoup d'entre eux ont joué un rôle de premier plan en ouvrant la voie à l'ouverture et aux transformations numériques. Le Baromètre des données ouvertes de la Fondation Web - parmi d'autres indicateurs - a contribué à démontrer ces réalisations. Mais l'ouverture en est, pour le moment, au point mort dans la région.

Le dernier baromètre régional de données ouvertes coordonné par l'ILDA indique des progrès limités de la part des gouvernements, ce qui reflète un tableau plus large et inquiétant se dessinant présentement en Amérique latine. De plus, la démocratie se trouve sur la défensive.

Bien que la région ait toujours été confrontée à des défis, une nouvelle vague de gouvernements et de mouvements sociaux est apparue, avec une tendance à ignorer les données en tant que source valable pour le débat et la prise de décision politique. Ceci alimente à son tour les campagnes de désinformation qui affectent le débat politique et les institutions politiques.

La gouvernance dans son ensemble est maintenant fortement débattue et de nouvelles formes de représentation et de participation politiques sont requises. La crise environnementale frappe également les Latino-Américains. Certains gouvernements possèdent un piètre bilan en ce qui a trait à la diffusion d'ensembles de données sur l'environnement et les ressources naturelles, mais il existe également des problèmes de capacité qui les empêchent de le faire. Sans données appropriées sur l'environnement, la lutte contre le changement climatique demeure un vœu pieux. L'émergence de l'IA peut accroître plusieurs risques sociaux et économiques. Les pays d'Amérique latine n'ont toujours pas d'approche commune pouvant établir un équilibre entre la protection de la vie privée, la productivité et l'innovation et la justice.

### Bilan des missions sur le terrain

Le score moyen pour l'Amérique latine dans le pilier gouvernance des données du Baromètre était similaire à la moyenne mondiale, montrant des progrès généraux limités dans la région par rapport au reste du monde. Un examen plus approfondi des indicateurs de gouvernance individuels révèle que certains affichent de meilleurs résultats que d'autres : avec des scores plus élevés pour les cadres de protection des données et les politiques de données ouvertes. D'autre part, les indicateurs relatifs aux cadres de partage des données figuraient parmi les plus faibles de la région. Le partage des données est l'un des domaines les plus complexes à traiter, mais est essentiel à la progression de l'utilisation des données pour le développement.

Le score moyen pour l'Amérique latine pour les capacités de données était légèrement inférieur à la moyenne mondiale. Plus précisément, le

soutien gouvernemental à la réutilisation des données montre une faiblesse importante, suivie par des préoccupations quant à la qualité et au financement des initiatives de données ouvertes. Au total, 14 des 15 pays participants ne disposent pas de mécanismes de financement pour le soutien de la réutilisation des données, et 10 ont déclaré que le gouvernement ne mène pas de séances d'information sur la réutilisation des données ni ne les organise. Et un total de 7 pays n'ont pas de hauts dirigeants politiques pour le soutien de l'initiative de données ouvertes, et 7 autres n'ont pas de budget alloué aux activités de données ouvertes. Malgré ces faiblesses régionales, plus de la moitié des pays de la région (selon le nombre) sont au-dessus de la moyenne mondiale. Les activités de renforcement des capacités demeurent une priorité pour l'Amérique latine et un domaine dans lequel la coopération internationale, les banques de développement, le secteur privé et les gouvernements peuvent collaborer.

Alors que globalement, le module sur l'intégrité politique avait les plus basses moyennes, en Amérique latine, ce module était le deuxième plus bas, quelques points supérieurs à la moyenne mondiale. Les cadres relatifs aux déclarations de patrimoine et au financement politique étaient les deux indicateurs aux scores les plus élevés dans la région, suivis par la disponibilité des données sur le financement politique. Contrairement à ces indicateurs, l'interopérabilité des données sur l'intégrité politique et la disponibilité des données sur le lobbying étaient les deux indicateurs ayant la moyenne la plus faible de la région. Aussi, la région présente un écart important entre les pays les plus performants et ceux dont les résultats sont médiocres. Dans ce deuxième groupe, 4 pays affichent des scores significativement faibles. En bref, l'Amérique latine a une grande marge d'amélioration au niveau de ce domaine.

La note régionale moyenne pour les données sur l'action climatique était supérieure à la moyenne mondiale au sein du Baromètre. Les trois indicateurs évalués (disponibilité des données sur les émissions, biodiversité et vulnérabilité) présentaient une moyenne supérieure dans la région que dans le monde, les données sur la vulnérabilité climatique étant l'indicateur pour lequel la région présente le plus grand écart positif par rapport à la moyenne mondiale. Ceci est la résultante d'efforts positifs de certains pays au sein de la publication de données climatiques. Malgré cela, les résultats généraux sur les données de l'action climatique ne peuvent être

considérés comme satisfaisants dans la région, les scores globaux n'ayant encore obtenu que la moitié des points possibles disponibles, ce qui indique qu'il reste encore beaucoup de tâches à accomplir, à la fois en ce qui a trait à l'ouverture des données et, pour quelques pays, la mise à disposition en ligne des données climatiques de base sous quelque forme que ce soit.

Sur le sujet spécifique de la vie privée, le Baromètre a identifié que 14 des 15 pays participants d'Amérique latine ont un cadre de protection des données, et 13 de ces cadres ont force de loi, montrant un progrès général sur le sujet. La plupart des cadres accordent aux personnes concernées des droits de choix ou de consentement et des droits d'accès et de rectification des données les concernant. Cependant, aucun d'entre eux ne couvre explicitement la protection des données liées à la localisation, et un seul aborde dans un certain contexte la prise de décision algorithmique, ce qui donne généralement l'impression que les cadres de protection des données dans la région nécessitent des mises à jour pour répondre aux problèmes de confidentialité modernes.

## Orientations futures

Notre récent événement régional - Abrelatam Condatos - était axé sur l'avenir du domaine des données ouvertes. Parmi la pandémie, des penseurs et acteurs de la région se sont réunis pour réfléchir aux possibilités que l'avenir pourrait offrir. Ceci fait partie des leçons à tirer de l'Amérique latine : les forums régionaux pour établir des dialogues et fixer des normes sont incroyablement efficaces. En synthétisant ces dialogues et en s'appuyant sur les recherches du Baromètre, nous pensons qu'il y a trois domaines clés où le domaine doit être amélioré.

Premièrement, l'ouverture doit être défendue. Le bilan en matière d'ouverture dans la région est au point mort et il ne semble pas y avoir suffisamment de soutien de la part des gouvernements pour améliorer l'offre de données ouvertes dans plusieurs domaines. Il est urgent de réunir une coalition de parties prenantes pouvant se concentrer sur la défense de l'ouverture dans les secteurs où la région en a le plus besoin : l'environnement, les finances publiques et l'intégrité. Pour le maintien de l'ouverture, il est important de soutenir un groupe émergent de la société civile, des médias et des universitaires utilisant ces données dans ces secteurs.

Deuxièmement, les données vont au-delà de l'ouverture. La façon dont les données sont actuellement traitées par les entreprises et les gouvernements est dans une certaine mesure incertaine. Les cadres de protection de la vie privée en Amérique latine évoluent, la région n'a cependant pas une vision cohérente et partagée de la façon dont l'ouverture, la protection de la vie privée et l'innovation sont liées. À l'ère de l'IA, il devient crucial de récolter les bénéfices de l'innovation. De nouveaux modèles de gouvernance doivent émerger sur la base de preuves et du respect des droits de l'homme.

Troisièmement, la question de la capacité et d'inclusion. De nombreuses décisions en matière de conception et de mise en œuvre sont toujours entre les mains d'un petit nombre de personnes qui ne sont pas en mesure ou ne sont pas disposées à créer des processus de conception

de politiques inclusifs et à mettre en œuvre les nouvelles conceptions. L'inclusion de critères permettant d'inclure de manière adéquate et sécurisée les différents genres, ainsi que la représentation des communautés exclues telles que les populations autochtones, est importante. Il est aussi nécessaire d'investir dans des infrastructures publiques solides et fiables pour soutenir le domaine des données afin de continuer à renforcer les capacités dans la région.

Dans l'ensemble, l'Amérique latine a du travail à faire pour que le domaine des données contribue au véritable développement régional. L'évolution sera inextricablement liée à l'évolution des institutions politiques dans la région et, conséquemment, la question cruciale à se poser est comme suit : dans quelle mesure le domaine des données servira-t-il la démocratie ?

## Perspectives du Hub : Caribbean Open Institute (Les Caraïbes)



*Par Suzana Russell, Lila Rao-Graham, Maurice McNaughton*

### Profil régional

Le Caribbean Open Institute (COI), basé à la Mona School of Business & Management, UWI Mona Campus, fait partie d'une communauté mondiale de hubs régionaux qui a soutenu le premier projet du baromètre mondial de données en s'engageant dans la collecte de données, l'analyse de données et la diffusion des résultats dans le cadre de la première édition de l'étude mondiale 2021. Huit pays des Caraïbes ont pris

part à cette étude initiale : les Bahamas, le Belize, la Guyane, la Jamaïque, Sainte-Lucie, Trinité-et-Tobago, Haïti et la République dominicaine.

Les Caraïbes sont l'une des régions du monde les plus diversifiées au plan culturel et politique. Avec une longue histoire d'« ouverture » culturelle et commerciale, située tel qu'elle l'est, à cheval sur les principales voies de navigation est-ouest, les histoires coloniales ont conféré à la région des Caraïbes un héritage éclectique d'influences dans les institutions politiques, sociales, culturelles et administratives. Tous les pays de cette région sont considérés comme de petits États insulaires en développement, ce qui pose des problèmes particuliers en ce qui a trait aux ressources publiques destinées aux activités liées aux données. Quant aux données ouvertes, la région a été caractérisée par une « forte demande, mais une offre lente ».

### Bilan des missions sur le terrain

En termes d'évaluation globale, les Caraïbes, en tant que région, se situent juste au-dessous de la moyenne mondiale pour les quatre piliers. La marge par rapport à la moyenne était plus élevée pour la disponibilité et plus faible pour l'utilisation et l'impact. C'est là conforme aux conclusions d'études antérieures (Baromètre des données ouvertes) suggérant que si les entreprises et

entrepreneurs de la région démontrent la capacité et la volonté de tirer parti des possibilités sociales et économiques offertes par les données pour le développement, ce n'est là pas suffisamment compensé par un engagement et une action proportionnés de la part des gouvernements en termes d'initiatives et de politiques appropriées quant aux données.

L'exploration du pilier de Gouvernance a entre autres révélé que les Caraïbes étaient les plus fortes dans les mécanismes de protection des données, mais les plus faibles dans les cadres de leur partage. Quant au pilier Capacités, bien que des preuves d'initiatives gouvernementales actives et ouvertes en matière de données dans plusieurs pays existent, les activités connexes visant à soutenir le renforcement des capacités des fonctionnaires en matière de données et à encourager une réutilisation plus large des données étaient moins évidentes. En ce qui concerne les indicateurs secondaires reflétant les capacités, les « libertés politiques et libertés civiles », le « capital humain » et « l'utilisation commerciale des outils numériques » étaient tous des facteurs favorables importants pour les Caraïbes, tandis que les plus faibles étaient le « gouvernement numérique » et les « institutions de données ». Cette constatation corrobore le discours selon lequel les gouvernements de la région doivent être plus proactifs dans la mise en place d'un environnement propice et dans le renforcement des capacités institutionnelles pour soutenir l'écosystème de données.

Allant plus loin dans la disponibilité, l'utilisation et l'impact des modules thématiques, la région a obtenu les meilleurs résultats en ce qui concerne les données relatives aux achats et à la santé. Le module thématique le plus bas concerne étonnamment la disponibilité et l'utilisation des données de l'Action pour le climat, compte tenu de la vulnérabilité particulière des Caraïbes aux changements climatiques.

## Orientations futures

Le plus récent rapport sur le développement mondial<sup>[27]</sup> souligne que les pays ne peuvent réaliser toute la valeur des données pour le développement qu'en établissant un système de données national intégré s'appuyant sur :

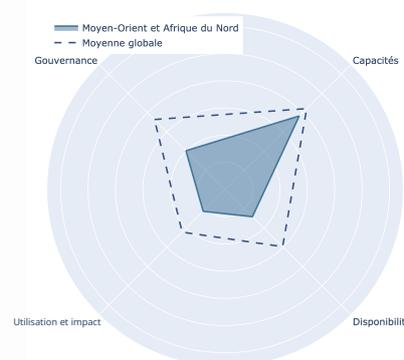
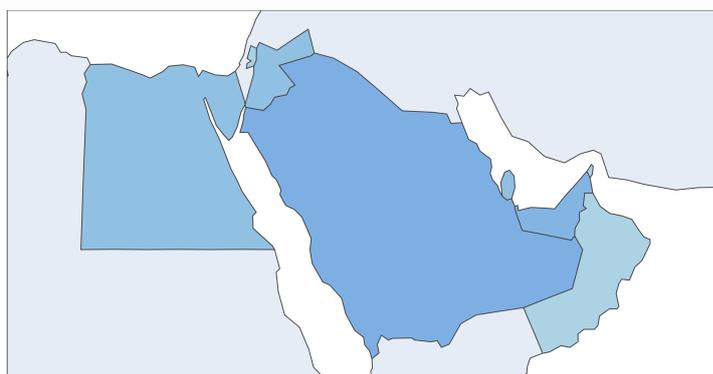
- une gouvernance intentionnelle, pangouvernementale et multipartite des données ainsi que des institutions qui protègent et protègent les données envers toute utilisation abusive ;
- des ressources humaines hautement qualifiées au sein du gouvernement, de la société civile, du monde universitaire et du secteur privé ;
- les données disponibles qui sont produites, ouvertes, contrôlées par la qualité, utilisées et réutilisées.

Le Baromètre mondial des données est un indice multidimensionnel important pour les Caraïbes, car celui-ci aide les régulateurs et les décideurs régionaux des secteurs publics et privés à mesurer et à évaluer les progrès parmi tous ces domaines. Des systèmes de données nationaux et régionaux sains sont considérés comme élément essentiel d'écosystèmes numériques équitables et d'une condition préalable explicite à la réalisation des objectifs de développement durable (ODD) de l'agenda 2030.

Les premiers résultats ci-dessus suggèrent un besoin particulier de financement adéquat et d'investissements ciblés dans les institutions de données et des compétences techniques des organismes du secteur public afin de coordonner les efforts de partage des données, ainsi que pour oeuvrer au renforcement des processus critiques et des canaux de prestation de services liés aux données.

## Moyen-Orient et Afrique du Nord

Le Baromètre couvre 8 pays de la région du Moyen-Orient et de l’Afrique du Nord. Bien que la région soit nettement inférieure à la moyenne mondiale en matière de gouvernance, de disponibilité, d’utilisation et d’impact des données, elle est relativement forte en termes de capacité du gouvernement et du secteur privé à gérer les données, les limites des scores de capacité étant façonnées par les restrictions à la liberté politique. La mise au point de modèles de partenariat permettant de tirer parti de la capacité du secteur privé de promouvoir l’utilisation des données pour le bien public, plus particulièrement en ce qui concerne les défis du développement durable, offre un moyen d’intensifier le soutien à la réutilisation des données. Grâce aux travaux sur les cadres de partage des données et aux activités de renforcement des capacités, il est aussi nécessaire de s’attacher à assurer l’égalité d’accès aux avantages des données et à la protection envers les préjudices potentiels.



| Score global | Gouvernance | Capacités | Disponibilité | Impact et utilisation |
|--------------|-------------|-----------|---------------|-----------------------|
| 21.6         | 20.19       | 38.26     | 14.08         | 11.29                 |

| Pays                 | Force comparative (régionale)                           | Faiblesse comparée (régionale)          | Score global | Gouvernance | Capacité | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|----------------------|---|---|--------------|-------------|----------|---------------|-----------------------|-------------------------------|-------------------------|
| EAU                  | (C) Infra-national <sup>RL</sup>                        | (G) Protection des données              | 26.7         | 16.1        | 58.2     | 17.6          | 12.0                  | 20.0                          | 72.0                    |
| Bahrain              | (G) Protection des données <sup>RL</sup>                | (C) Initiative sur les données ouvertes | 22.0         | 23.6        | 28.1     | 18.4          | 10.8                  | 63.0                          | 20.0                    |
| Égypte               | (A) Émissions <sup>RL</sup>                             | (G) Politique d’ouverture des données   | 21.8         | 25.0        | 34.7     | 13.2          | 12.0                  | 0.0                           | 16.8                    |
| Jordanie             | (G) Données de consultation publique <sup>RL GL</sup>   | (G) Protection des données              | 22.2         | 34.5        | 33.7     | 7.8           | 12.3                  | 63.0                          | 80.0                    |
| Oman                 | (A) Données sur le budget et les dépenses <sup>RL</sup> | (C) Initiative sur les données ouvertes | 14.1         | 8.6         | 25.1     | 12.1          | 12.0                  | 54.0                          | 0.0                     |
| État de la Palestine | (C) Initiative sur les données ouvertes <sup>RL</sup>   | (G) Protection des données              | 14.7         | 9.1         | 32.3     | 9.2           | 9.6                   | 37.8                          | 81.0                    |
| Qatar                | (G) Protection des données                              | (A) Données des marchés publics         | 22.2         | 16.6        | 44.9     | 14.2          | 12.0                  | 40.0                          | 80.0                    |
| Arabie Saoudite      | (G) Gestion des données <sup>RL</sup>                   | (G) Déclarations d’actifs               | 29.0         | 28.0        | 49.1     | 20.1          | 9.6                   | 63.0                          | 80.0                    |

## Perspectives du Hub : Université CEE Birzeit (Moyen-Orient et Afrique du Nord)



Par Raed Sharif avec Abed Khooli

### Profil régional

Au cours de la dernière décennie, les changements mondiaux et plusieurs cycles de soulèvement dans la région MOAN ont fondamentalement changé la relation entre gouvernement et citoyens. Ces changements ont aujourd'hui transformé la perception des citoyens envers leurs droits et leurs devoirs, plus particulièrement dans les domaines liés à la responsabilité, à la transparence et à l'ouverture. La numérisation rapide des économies et des sociétés en réponse à la pandémie de COVID-19 a plus récemment créé des pressions pour des stratégies, des outils et des applications de données plus rapides et plus efficaces.

En même temps, il est important de reconnaître certains des défis fondamentaux en matière d'infrastructures et de politiques auxquels la région fait face. Selon l'Union internationale des télécommunications (UIT), 55 % des résidents des pays arabes avaient accès à Internet en 2019, et plus des deux tiers (67 %) d'entre eux étaient des jeunes âgés de 15 à 24 ans. Ventilés par sexe, 47 % des femmes et 61 % des hommes avaient accès à Internet en 2019. Ceci indique une disparité entre les sexes de 0,77. Du point de vue géographique, 74 % des zones urbaines (principalement la 4G) et 34 % des zones rurales

avaient accès à Internet, les zones rurales dépendant principalement de connexions mobiles 3G plus lentes.

Une ventilation plus détaillée des données par sous-région et par pays révèle de nouvelles disparités au niveau de l'accès à l'Internet et dans son utilisation. Le pourcentage d'internautes dans la région du Golfe est par exemple supérieur à 90 %, entre 60 et 65 % dans les pays arabes d'Afrique du Nord (Maroc, Algérie et Tunisie) et inférieur à 40 % dans d'autres parties de la région MENA telles que le Soudan (30,9 %), le Yémen (26,7 %), la Syrie (34,3 %) et la Mauritanie (20,8 %). Des tendances similaires se dégagent également quant aux compétences individuelles en matière de TIC (p. ex., la capacité d'utiliser des appareils, des logiciels et des applications).

Au niveau politique et juridique, seuls six pays arabes ont des lois relatives au droit à l'information (La Jordanie (2007), le Maroc (2011), le Yémen (2012), le Soudan (2015), la Tunisie (2016) et le Liban (2017)), et trois d'entre eux sont membres de l'Open Government Partnership (OGP).

### Bilan des missions sur le terrain

Malgré les différences au niveau des pays, l'enquête a révélé des progrès modestes vers plus de partage de données, d'ouverture et de facilité d'utilisation dans la région MOAN. Il existe aussi une volonté politique croissante et un soutien officiel aux initiatives en matière de données et de données ouvertes dans des pays tels que les Émirats arabes unis, le Qatar, Oman et l'Arabie saoudite. De plus, des signes encourageants montrent que la société civile et d'autres acteurs non gouvernementaux jouent un rôle plus actif dans la promotion d'une gestion, d'un partage et d'une convivialité efficaces des données et aident les gouvernements à traduire leurs politiques et leurs visions en produits et services concrets.

Les lacunes persistantes en matière de capacités peuvent toutefois avoir un impact négatif sur l'impact potentiel des initiatives nationales. L'insuffisance ou le manque d'infrastructures et la faiblesse des connaissances numériques peuvent entraver la capacité des citoyens à utiliser

pleinement les données disponibles et à prendre part de manière significative aux consultations publiques et aux processus décisionnels. L'une des pratiques qui se sont par exemple révélées populaires et efficaces dans la région pour inciter les citoyens et les principales parties prenantes à s'impliquer dans le domaine du gouvernement ouvert et des données (et dans le processus de transformation numérique plus large du gouvernement) a été les idéathons et les hackathons de données, où les citoyens, innovateurs, militants, chercheurs et les scientifiques de données explorent et testent en collaboration le potentiel d'un large éventail d'ensembles de données pour la création de solutions aux défis sociaux, économiques et environnementaux auxquels leurs sociétés sont confrontées.

### Orientations futures

L'enquête barométrique a révélé des domaines spécifiques nécessitant une action, notamment:

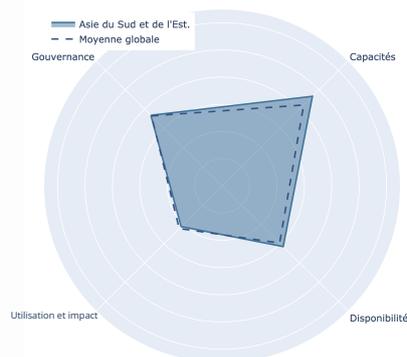
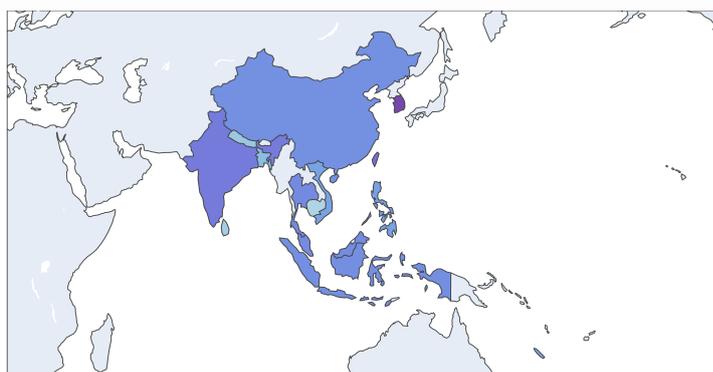
- 

Nous encourageons de plus les agences de données et les initiatives dans la région MENA à établir des engagements stratégiques avec les organisations de la société civile, le secteur privé et d'autres parties prenantes clés pour renforcer les capacités nécessaires à une utilisation efficace et stratégique des données rendues disponibles. Il est aussi nécessaire d'accroître la formation professionnelle et les programmes universitaires liés aux données pouvant accroître et améliorer le capital humain de la région autour de données pour le bien public.

Nous encourageons finalement les agences de données et les initiatives dans la région MOAN à explorer le potentiel de collaborations régionales et l'échange de connaissances, de compétences et d'enseignements tirés, pour renforcer les capacités régionales en matière d'IA responsable et de l'analyse de données. Ceci est plus particulièrement important dans les domaines liés aux outils et techniques d'apprentissage automatique en langue arabe, à l'analyse avancée des données et à la visualisation des données, en mettant davantage l'emphase sur l'infrastructure et les cadres réglementaires régionaux.

## Asie du Sud et de l'Est

Le Baromètre couvre 15 pays d'Asie du Sud et de l'Est. La diversité des pays au sein de la région se traduit également par des écarts importants dans les résultats régionaux à travers les quatre piliers. Malgré la présence généralisée de politiques et d'initiatives en matière de données ouvertes, la disponibilité et la publication des données ouvertes des ensembles de données clés pourraient être renforcées, plus particulièrement en ce qui a trait à l'intégrité politique, l'information des entreprises et l'action pour le climat.



| Score global |   | Gouvernance                                    | Capacités    | Disponibilité | Impact et utilisation |               |                       |                               |                         |
|--------------|---|--|--------------|---------------|-----------------------|---------------|-----------------------|-------------------------------|-------------------------|
| 32.79        |   | 35.9   | 37.05        | 28.62         | 27.65                 |               |                       |                               |                         |
| Pays         | Force comparative (régionale)                 | Faiblesse comparée (régionale)                 | Score global | Gouvernance   | Capacité              | Disponibilité | Utilisation et impact | Politique de Données ouvertes | Libre accès aux données |
|              | (A) Données financières politiques            | (G) Données de consultation publique           | 23.8         | 21.1          | 32.1                  | 22.1          | 10.8                  | 63.0 ▲                        | 49.0 ■                  |
|              | (G) Performance RTI                           | (A) Vaccination (COVID-19)                     | 39.8         | 35.8          | 55.5                  | 35.3          | 24.3                  | 18.9 ■                        | 68.0 ▲                  |
|              | (A) Émissions RL                              | (G) Gestion des données                        | 48.6         | 37.5          | 58.2                  | 52.9          | 30.1                  | 54.0 □                        | 80.0 □                  |
|              | (G) Performance RTI RL                        | (G) Protection des données                     | 40.2         | 48.0          | 46.2                  | 32.2          | 29.8                  | 0.0 ▼                         | 76.5 ■                  |
|              | (G) Gestion des données                       | (G) Cadres de partage des données              | 46.7         | 53.2          | 51.4                  | 41.6          | 22.9                  | 90.0 ▲                        | 80.0 ■                  |
|              | (C) Fonction publique                         | (G) Données de consultation publique           | 13.2         | 5.6           | 25.0                  | 12.2          | 8.4                   | 0.0 □                         | 0.0 □                   |
|              | (A) Propriété effective RL                    | (A) Vulnérabilité                              | 64.5         | 66.4          | 80.4                  | 57.0          | 32.6                  | 90.0 ■                        | 100.0 ■                 |
|              | (G) Performance RTI                           | (G) Données de consultation publique           | 16.3         | 13.4          | 35.0                  | 8.2           | 12.0                  | 0.0 □                         | 40.0 □                  |
|              | (C) Soutien gouvernemental à la réutilisation | (G) Gestion des données                        | 41.6         | 44.8          | 68.8                  | 24.2          | 35.3                  | 54.0 ▲                        | 63.0 ■                  |
|              | (G) Protection des données RL                 | (G) Données de consultation publique           | 30.7         | 32.2          | 37.9                  | 27.2          | 13.0                  | 80.0 □                        | 54.0 □                  |
|              | (U) Analyse des données d'approvisionnement   | (G) Données de consultation publique           | 18.9         | 16.7          | 23.0                  | 18.3          | 18.6                  | 0.0 ■                         | 0.0 ▼                   |
|              | (A) Propriété effective                       | (A) Registre du commerce et des sociétés       | 34.0         | 44.7          | 33.4                  | 28.3          | 18.2                  | 50.4 ▼                        | 40.8 ▼                  |
|              | (G) Finances politiques                       | (G) Données des finances publiques             | 41.7         | 42.2          | 55.0                  | 35.9          | 19.7                  | 70.0 ▲                        | 80.0 ▲                  |
|              | (G) Registre du lobbying RL                   | (C) Fonction publique                          | 51.1         | 48.9          | 55.4                  | 52.7          | 23.4                  | 72.0 □                        | 80.0 □                  |
|              | (G) Données des finances publiques            | (A) Capacité du système de santé en temps réel | 33.3         | 40.7          | 41.8                  | 24.7          | 17.7                  | 36.0 ▲                        | 40.8 ▲                  |

## Perspectives du Hub : D4DAsia (Asie du Sud-Est)



**ASIA HUB**

Data for Development

*Par Khairil Yusof et Pyrou Chung*

### Profil régional

L'Asie est une région peuplée et diversifiée, avec des langues, cultures, formes de gouvernement, niveaux de développement économique et des contextes politiques différents, même au sein des sous-régions. En tant que tel, lors de l'examen des tendances générales et des scores globaux, il faudra approfondir chaque pays et chaque pilier/module en détail pour mieux comprendre les défis et les opportunités en matière de données pour le développement.

Pour l'Asie du Sud-Est, comme pour plusieurs pays, en plus des questions de santé, la Covid-19 a aussi entraîné un déclin rapide de l'espace civique, plus particulièrement à Hong Kong, et en moindre mesure en Malaisie, en Thaïlande et aux Philippines, l'instabilité politique et le recours aux mesures d'urgence ayant des impacts sur la disponibilité des données dans un certain nombre de cas.

L'Asie de l'Est et du Sud-Est est également une région géopolitiquement importante pour le commerce et les investissements internationaux. De nombreuses perturbations dans la chaîne d'approvisionnement mondiale dues aux confinements de la Covid-19 peuvent être attribuées aux exportations manufacturières clés et à l'expédition de marchandises au sein de la région. La capacité des gouvernements à fournir

des données de qualité et des services numériques liés au commerce et à l'investissement sera essentielle à la reprise économique dans la région.

À la fin de 2021, l'Asie du Sud-Est a également été confrontée à des inondations désastreuses et à des tempêtes tropicales et, avec la crise climatique, les conditions météorologiques extrêmes continueront probablement à affecter la région plus fréquemment dans les années à venir. Les forêts tropicales détiennent plus de 68 % des stocks mondiaux de carbone, et l'Asie du Sud-Est abrite 15 % des forêts tropicales mondiales.

### Bilan des missions sur le terrain

L'Asie du Sud-Est est une région exceptionnelle parmi les autres sous-régions Asiatiques, ayant une capacité supérieure à la disponibilité des données. Les résultats globaux du Baromètre dans la région sont généralement en corrélation avec le niveau de développement économique (par habitant/revenu) d'un pays. L'atout notable dans la région est Taïwan, où les données sont très disponibles, mais où les résultats en matière de gouvernance sont médiocres.

À l'exception de la Corée du Sud, le pays asiatique ayant obtenu les meilleurs résultats dans l'étude, les capacités des pays en matière de données ne semblent pas bien se traduire en effets. Ceci soulève des questions sur les types d'interventions pouvant être nécessaires au soutien d'une utilisation plus large des données.

Quant aux données ouvertes, les pays de la région semblent avoir de bonnes bases législatives, juridiques et politiques pour les données ouvertes, à l'exception notable du Cambodge, où la disponibilité et les capacités des données sont fournies par les organisations locales et internationales de la société civile, plutôt que par le gouvernement.

La diversité de la région ressort clairement des données du Baromètre. Pour chaque pays de la région, quelques modules existent ou indicateurs dont les résultats sont médiocres par rapport aux résultats globaux. Les points forts des données nationales ne semblent cependant pas toujours suivre les problèmes les plus urgents auxquels ils sont confrontés : ce qui suggère qu'il peut y avoir

une déconnexion entre les infrastructures de données dont disposent les pays et celles dont ils ont besoin pour relever les défis futurs. Par exemple, un certain nombre de pays récemment gravement touchés par les changements climatiques ont des scores relativement faibles dans le module d'action climatique par rapport à leurs capacités globales.

## Orientations futures

Plutôt que de mettre l'accent sur les scores globaux, les pays peuvent utiliser le site Web du Baromètre mondial de données pour approfondir les indicateurs et les modules où leurs scores sont relativement faibles, en examinant comment les mesures individuelles pourraient être ciblées pour l'amélioration. Il est tout aussi important d'examiner les points forts comparatifs d'autres pays. Chaque pays a quelque chose d'unique qu'il fait bien, et dont d'autres pourraient en apprendre. Les pays d'Asie du Sud tels que, par exemple, le Sri Lanka, l'Inde et le Népal ont de très solides lois et capacités en matière de droit à l'information dont d'autres pays peuvent tirer des leçons. Un grand potentiel existe à développer davantage les possibilités de partager les bonnes pratiques dans toute la région.

Dans l'ensemble, l'Asie du Sud-Est paraît posséder une bonne capacité théorique à assurer le bien public depuis des données, mais par rapport à d'autres régions, il existe un écart important entre la collecte de données et la disponibilité pour la réutilisation. Les données recueillies par l'enquête du Baromètre fournissent une mine d'informations sur les endroits où ces données verrouillées peuvent exister, et peuvent aider à élaborer des stratégies pour déverrouiller ces données par le biais d'initiatives gouvernementales et de la société civile. Pour une région qui a plus particulièrement été touchée par des phénomènes météorologiques violents dus au changement climatique, en termes de données pour l'action climatique, il reste encore beaucoup à faire. De nombreux pays ont de grandes populations vivant dans des plaines inondables et sont affectés chaque année par les inondations et

tempêtes tropicales ainsi que par la déforestation, les scores enregistrés dans le module Action pour le climat des baromètres pour la région étaient cependant relativement faibles.

Quant aux utilisations futures du Baromètre, nous prenons également note que, dans un monde connecté, des données pertinentes pour chaque pays peuvent également être trouvées au niveau mondial ou au-delà des frontières. Nous devrions par exemple prêter attention aux systèmes mondiaux de publication de données tels que le COMTRADE pour les statistiques commerciales de l'ONU et la mission GEDI (Global Ecosystem Dynamics Investigation) des États-Unis et de la NASA, qui peuvent fournir des données sur la couverture forestière pour les pays qui n'ont pas la capacité de générer ou de publier de telles données. De même, les organisations non-gouvernementales internationales et les banques de développement ont mis et devraient mettre à disposition les données qu'elles recueillent dans les pays où elles travaillent, en fournissant d'autres sources de données disponibles pour les pays qui n'ont pas actuellement la capacité de fournir des ensembles de données clés.

À l'aide du baromètre actuel, nous pouvons déjà examiner des données transfrontalières potentielles pouvant aider à résoudre des problèmes nationaux, telles que les données effectives sur les bénéficiaires et la divulgation d'actifs disponibles dans les pays ayant des secteurs financiers importants tel que le Royaume-Uni et les États-Unis, qui peuvent être utilisées pour l'obtention d'informations sur les actifs étrangers des politiciens, ou les flux de financement politique en provenance d'autres pays. Pour soutenir la prise de décision fondée sur les données dans toute l'Asie de l'Est et du Sud-Est, nous devons continuer à développer la gouvernance et la disponibilité des données régionales, mais aussi à développer la sensibilisation aux ressources de données régionales sur le changement climatique, les exportations de vaccins, le commerce et de nombreux autres sujets disponibles ailleurs dans le monde.

## Perspectives du Hub : Internet Society Hong Kong (Hong Kong)



Par Benjamin Zhou

### Profil régional

Hong Kong, officiellement la Région administrative spéciale de Hong Kong de la République populaire de Chine (RASHK), est régie par le principe d'« un pays, deux systèmes », qui permet à la ville de maintenir un système juridique différent de celui de la Chine continentale. Dans la pratique, l'administration municipale jouit d'un niveau élevé d'autonomie pour les affaires locales, y compris la formulation de politiques en matière de données. Hong Kong est aussi une plaque tournante financière internationale qui reconnaît le chinois et l'anglais en tant que langues officielles, et ces deux langues sont largement utilisées.

Quant aux politiques et aux mesures relatives aux données, le gouvernement de Hong Kong a adopté une approche sectorielle plutôt qu'holistique. Il n'y a pas de loi ou de document de politique établissant une stratégie de collecte, de partage et de protection des données, ni de preuve indiquant qu'ils ont l'intention de ce faire. Sauf la loi sur la confidentialité des données - Personal Data (Privacy) Ordinance - promulguée en 1996, les initiatives gouvernementales liées aux données sont principalement des mesures administratives sans législation. Par exemple, bien qu'un portail de données ouvertes ait été lancé en 2011 dans le cadre d'un projet pilote, ce n'est qu'en décembre 2017 que le gouvernement a officiellement abordé la question des « données

ouvertes » dans un document d'orientation de haut niveau et dans le discours politique annuel du chef exécutif. Le portail data.gov.hk est géré par l'Office of Government Chief Information Officer (OGCIO) en tant qu'une plate-forme permettant à tous les départements, organismes publics et certaines entreprises privées de (volontairement) publier des données. Entre-temps, d'autres départements ont mis en place leurs propres plateformes de données, comme un projet en cours sur l'infrastructure commune de données spatiales élaboré par le département des Terres.

### Bilan des missions sur le terrain

Les résultats de l'évaluation indiquent que, parmi les quatre piliers du Baromètre, Hong Kong obtient de meilleurs résultats en matière de disponibilité des données que les trois autres piliers grâce au classement mondial, tandis que la gouvernance est la plus faible à cet égard.

Les ensembles de données examinés par les modules thématiques du Baromètre sont pour la plupart disponibles, à l'exception de trois : la propriété effective, le lobbying et le financement politique. L'indicateur de propriété effective dans le module d'information sur les sociétés est nul, car l'ordonnance sur les sociétés de Hong Kong (modifiée en 2018), obligeant les sociétés à tenir à jour les informations de leurs contrôleurs importants, ne prévoit pas de registre centralisé de propriété effective. Hong Kong ne reçoit pas non plus de crédit sur les indicateurs de lobbying et de financement politique dans le module sur l'intégrité politique, vu qu'il n'existe pas de législation réglementant les partis politiques et leur financement bien que la question ait été débattue publiquement depuis de nombreuses années. En fait, ces données ne sont pas disponibles auprès du gouvernement ou d'un tiers autorisé, mais auprès de certains partis politiques à la suite d'une divulgation volontaire. Comparativement à d'autres juridictions dans les évaluations du Baromètre, la disponibilité des données de Hong Kong généralement bonne.

L'enjeu majeur de Hong Kong réside dans la gouvernance : elle a des scores inférieurs aux indicateurs de gouvernance plutôt qu'aux autres piliers, y compris la capacité, la ville n'ayant pas encore mis en place de structure de gouvernance

ni fourni de stratégie de données. Parmi tous les indicateurs de ce module, seul l'indicateur de protection des données obtient une note supérieure à 80 %, vu la législation en vigueur sur la protection de la vie privée. Une politique sur les données ouvertes existe, mais il suffit de demander à chaque ministère de dresser un plan annuel de données ouvertes par lui-même et de s'assurer que les ensembles de données publiés sur le portail de données ouvertes satisfont à quelques critères (c.-à-d. lisibles à la machine, opportuns, métadonnées et licence ouverte). Les indicateurs de gestion des données et des cadres de partage obtiennent les résultats les plus bas, car ils ne sont pas clairement documentés, voir inexistant. Certaines initiatives sectorielles de partage de données existent, par exemple une plateforme développée par l'administration hospitalière afin de partager les données de santé des hôpitaux publics à des fins de recherche, mais aucun cadre de partage de données pour la société en général n'a été trouvé lors de la recherche. Le cadre de gestion des données, qu'il s'agisse de données ouvertes ou non, peut prendre la forme d'une stratégie, d'orientations ou de normes nationales ou locales en matière de données, conformément aux directives du Baromètre. Aucun de ces outils n'est cependant accessible au public à Hong Kong.

## Orientations futures

Bien que les preuves démontrent une bonne base de disponibilité des données, le simple fait de pousser les ministères à publier davantage d'ensembles de données n'est pas suffisant pour relever les défis de l'ère des mégadonnées et pour en libérer les valeurs. L'absence de structure et de stratégies de gouvernance des données au sein de l'administration risque de nuire aux réalisations actuelles et à la capacité du gouvernement à assurer la prestation de la fonction publique, faisant face à une quantité croissante de données de partout. Le gouvernement de Hong Kong devrait investir davantage dans la formulation d'une feuille de route à long terme sur la coordination de la collecte, de la gestion, du partage, de la publication et de l'utilisation des données dans tous les secteurs de la société en général. La vision et le leadership sont essentiels pour que la structure de gouvernance s'assure que les stratégies et les politiques soient constamment revues et adaptées pour relever les nouveaux défis.

# **Conclusions et travaux futurs**

Ce rapport passe en revue les données de la première édition du Baromètre mondial de données, organisées autour de quatre piliers : la gouvernance, les capacités, la disponibilité, l'utilisation et les impacts des données. Via une enquête mondiale et des données secondaires, le Baromètre a exploré différentes dimensions de la gouvernance des données incluant la prévention de l'utilisation abusive et la gestion des données pour s'assurer qu'elles sont aptes à l'utilisation et la promotion de la disponibilité de celles-ci pour la réutilisation. Il s'est penché sur la mesure dans laquelle les ressources et les compétences liées aux données et la liberté et la possibilité de les déployer, existent et sont distribuées dans chaque pays. Il a examiné la mesure dans laquelle les données nécessaires à répondre aux problèmes urgents sont partagées ou ouvertes dans chaque pays incluant les données liées à la crise climatique ; à la pandémie de COVID-19 ; à l'intégrité politique ; à la propriété de l'entreprise ; à l'utilisation et à la gestion des terres ; et à la surveillance des marchés publics et des finances. Et il a rassemblé des exemples d'utilisation des données, ainsi que des preuves d'impacts, dans quatre cas d'utilisation pertinents pour les parties prenantes des secteurs public, privé, de la société civile et des médias. Les données probantes et les indicateurs recueillis dans le cadre de notre enquête comblent les lacunes critiques quant aux connaissances, et le modèle en réseau des pôles de recherche régionaux du Baromètre et la collecte de données par les chercheurs nationaux ont contribué au renforcement des capacités en vue d'une action plus globale sur les données pour le bien public.

À l'aire d'un indice pondéré, le Baromètre résume des centaines de points de données pour chacun des 109 pays couverts en évaluations détaillées. Au sein de ce rapport, nous avons utilisé ces scores quantitatifs et les mesures individuelles qui les génèrent en tant que point de départ pour l'analyse comparative et l'exploration. Nous n'avons pu couvrir qu'un petit sous-ensemble des éléments de preuve recueillis par le Baromètre. Ce rapport se trouve côte à côte avec l'ensemble de données du Baromètre, aux profils de pays et thématiques et aux histoires de données <https://www.globaldatabarometer.org> approfondies disponibles sur ainsi que de l'analyse et des résultats produits par les pôles régionaux. Nous invitons et encourageons les lecteurs à s'engager avec ces ressources complémentaires pour d'approfondir les conclusions du Baromètre et à interroger comment celles-ci peuvent éclairer vos propres priorités et pratiques.

## Des données pour le bien public : de l'objectif à la réalité ?

Si nous prenons le score le plus élevé obtenu pour chaque indicateur du Baromètre, et que nous imaginons que ces scores ont tous été atteints par un seul pays, ceci donnerait un score de 95.92 sur 100 sur l'indice pondéré global du Baromètre. Comme le Baromètre fixe un niveau de référence élevé, fondé, dans la plupart des cas, sur des normes et critères convenus à l'échelle mondiale plutôt que sur la pratique de fait d'un pays ou d'une région en particulier, c'est là une conclusion frappante. Pour chacun des thèmes étudiés, il est *possible* de régir les données, de les rendre disponibles et de les utiliser de manière à répondre aux aspirations généralement acceptées et à maximiser les possibilités d'obtenir des résultats favorables au bien public. Le score global moyen dans tous les pays n'était cependant que 34.38 sur 100, indiquant là un écart important entre le point de référence idéal et les réalités actuelles.

Bref, il est *possible* de créer les conditions nécessaires pour que les données contribuent au bien public, mais elles ne sont nullement garanties. Les choix politiques et actions concrètes au cours des prochaines années doivent explicitement viser à renforcer les capacités, les institutions et les infrastructures autour des données pour garantir aux individus et aux communautés une protection

efficace contre les préjudices liés aux données, tandis que les ensembles de données critiques sont collectés, gérés, partagés et publiés de manière fiable avec les fonctionnalités nécessaires à une réutilisation généralisée, collaborative et à fort impact.

Il y a, dans certains cas, des gains rapides à avoir. En examinant la pratique des pairs proches, ou en se concentrant sur les champs et les caractéristiques manquants des données ouvertes publiées, les gouvernements pourraient être en mesure de faire des progrès rapides par rapport à certains indicateurs du Baromètre. Dans d'autres cas, la réalisation de progrès exige des interventions plus vastes et à plus long terme : tel qu'identifier les possibilités de réformer les cadres juridiques pour mieux refléter les risques et les possibilités actuels liés aux données et placer la divulgation de données structurées sur une base légale, ou investir dans des programmes de renforcement des capacités durables et inclusifs pour la production et l'utilisation des données. Les initiatives mondiales soutenant l'apprentissage par les pairs et fournissant une assistance pratique aux gouvernements pour la traduction d'engagements politiques en données pratiques semblent avoir été particulièrement utiles pour l'accélération des progrès dans un certain nombre de secteurs tels que les finances et marchés publics, et il est de plus en plus justifié pour les donateurs d'investir dans la construction d'infrastructures de données. Les données de référence fournies par cette première édition du BGD seront précieuses pour le suivi à savoir dans quelle mesure d'autres programmes émergents, tels que la transparence de la propriété des entreprises, l'utilisation des données pour la gouvernance foncière ou une plus grande divulgation des lobbies politiques, sont en mesure d'intégrer et de diffuser les bonnes pratiques autour de la publication de données structurées, ouvertes et interopérables au cours de la prochaine décennie, et de créer divers écosystèmes de données qui nous aident à atteindre les objectifs de développement durable.

Il ne faut pas sous-estimer le défi que représente le fait de veiller à ce que les données servent le bien public. Les scores relativement faibles observés dans les piliers et indicateurs du Baromètre permettent une vérification de la réalité. Il y a relativement peu de pays proches d'avoir des ensembles de données publiques interopérables couvrant un large éventail de sujets. De nombreux pays aux revenus allant de faible revenu à intermédiaire ont d'importantes divisions de données à surmonter, et, pour les pays à revenu élevé, les systèmes existants peuvent faire échouer les tentatives de création d'infrastructures de données modernes regroupées. En faisant l'examen des faibles scores obtenus par rapport aux questions liées à l'inclusion dans le Baromètre, on constate que les divisions des données au sein des pays peuvent être aussi importantes que celles entre les pays. Dans de nombreux cas, les ensembles de données publiques ou les compétences pouvant s'avérer nécessaires pour piloter de nouveaux systèmes algorithmiques, soutenir l'analyse en vue de concevoir des politiques plus inclusives ou permettre un examen minutieux de la prise de décision publique au moyen de données ne sont tout simplement pas présents sous forme exhaustive, et encore moins largement disponibles. Un mouvement en faveur d'une prise de décision fondée sur les données peut avoir une dynamique très différente dans un pays qui dispose déjà de données publiques raisonnablement inclusives et d'une communauté active d'utilisateurs de données vis-à-vis un pays où les données nécessaires ne sont disponibles qu'auprès d'acteurs privés, et où les données ont été initialement collectées dans le seul but privé. De même, l'impact des nouveaux efforts de collecte de données peut varier considérablement entre les pays qui disposent de solides cadres et institutions de protection des données, et ceux n'en ayant pas. Le Baromètre offre un moyen d'interroger les affirmations sur la façon dont les données pourraient transformer le développement dans différents contextes de manière critique, et fournit un point de départ pour la conception d'interventions plus personnalisées et adaptées au contexte.

La comparaison des données de cette première édition du Baromètre des données *Global* aux indicateurs comparables des cinq éditions précédentes du Baromètre des données ouvertes (2012 - 2020) révèle un programme de données ouvertes qui est toujours d'actualité, mais qui est plus ou moins au point mort. Avec un certain nombre d'anciens leaders des données ouvertes perdant leur focus sur l'ouverture, et un roulement parmi les pays évalués comme ayant des politiques ou des initiatives fortes, nous voyons les données ouvertes avoir du mal à retenir l'attention. Bien que certains pays intègrent les principes de l'ouverture des données dans la législation sectorielle, un risque important existe que l'idée que les données publiques devraient être « ouvertes par défaut » continue

de perdre de son élan et que des gains en matière d'ouverture des données publiques pourraient être perdus. Si on les aborde avec soin, les travaux visant à renforcer les cadres pour le partage des données pourraient être l'occasion de réaffirmer également l'importance de l'ouverture, en encadrant le partage des données comme alternative de repli lorsque « l'ouverture par défaut » ne peut pas être adoptée.

Le défi de s'assurer que les données sont disponibles pour une réutilisation significative est bien illustré par les lacunes des données autour de l'action climatique et de la COVID-19. Des données désagrégées, disponibles pour une utilisation pratique de la résolution de problèmes par les parties prenantes nationales, et en particulier infra-nationales, peuvent et doivent être disponibles. Pourtant, dans des centaines de cas, les données rapportées globalement aux parties prenantes mondiales ne sont pas accessibles aux communautés, aux journalistes ou aux entrepreneurs dans le pays, ou travaillant au niveau populaire. Alors que notre enquête sur l'utilisation et l'impact des données a révélé de nombreux exemples de solutions de contournement et d'aménagements que les utilisateurs peuvent apporter aux données manquantes ou non structurées, la myriade d'obstacles à l'accès et à l'utilisation efficace des données continue de limiter la mesure dans laquelle l'innovation et la résolution de problèmes fondées sur les données peuvent s'étendre. Dans ce rapport, nous avons identifié un éventail de moyens possibles pour combler certaines de ces lacunes, allant des gouvernements qui se concentrent davantage sur l'utilisation de leurs propres ensembles de données (ouvertes) et qui développent des infrastructures de données plus robustes, en passant par un partenariat et une collaboration accrue autour de la fourniture et de l'utilisation des données. Tout comme le détail du bien public est ultimement défini dans chaque pays et dans chaque communauté, les infrastructures de données doivent également être définies et façonnées par l'engagement et le dialogue.

En somme, cette première édition du Baromètre mondial de données est, au final, une invitation au dialogue. Nous espérons que vous pourrez :

- Utiliser les conclusions et les preuves présentées dans ce rapport pour susciter la discussion et le débat sur la façon dont les données sont régies, mises à disposition et utilisées dans votre région, pays ou secteur ;
- Explorez les visualisations ou les histoires de données sur le site Web du Baromètre mondial de données afin d'accéder à de nouvelles idées, questions ou aperçus, et pour trouver des pairs potentiels avec lesquels parler et apprendre ;
- Téléchargez les données détaillées de l'enquête et les utiliser dans le cadre de nouvelles recherches, analyses ou explorations - qu'il s'agisse d'une analyse statistique des scores des indicateurs, ou en prenant les URL fournies comme point de départ pour une étude qualitative plus représentative à l'échelle mondiale ;
- Partagez vos commentaires avec nous pour nous aider à améliorer les données et la méthodologie de l'étude. Nous savons qu'il y a beaucoup à faire pour améliorer à la fois la méthode et les données de l'étude. Recherchez des détails sur le site Web de la façon de contribuer afin de façonner les itérations futures du baromètre.
- Restez au courant des développements en vous inscrivant pour recevoir des mises à jour sur le site Web du Baromètre.

### Réflexions méthodologiques et orientations futures

Cette première édition a, bien sûr, des limites importantes. Tel que discuté à la section méthodologie, les contraintes causées par le COVID et l'ampleur du défi de la recherche dans une première édition (où toute la recherche implique une *nouvelle* collecte de preuves, plutôt que la *mise à jour* des preuves passées), signifient que nous avons couvert moins de domaines thématiques qu'initialement prévu, et que nous n'avons pas été en mesure d'élaborer des indicateurs négatifs examinant directement les preuves de préjudices provenant d'une mauvaise utilisation des données ou de cadres de gouvernance médiocres, ou qui mesurent de manière plus robuste les questions de gouvernance *dans la pratique* plutôt que *dans le droit*. Il nous a été particulièrement difficile de trouver des données secondaires et d'élaborer des questions d'enquête primaires solides pour aborder explicitement les questions de genre et d'équité, et nous sommes déterminés à poursuivre nos efforts pour combler ces lacunes. Nous avons également déterminé que nous ne serions pas en mesure de rendre justice aux questions d'intelligence artificielle au sein du Baromètre, et nous avons réduit l'attention que nous avions prévue sur ce sujet. Nous attendons avec impatience le prochain Index mondial sur l'IA responsable qui, nous l'espérons, sera en mesure d'éclairer les prochaines éditions du BGD. Notre intention est que, lors des prochaines éditions de cette étude, nous puissions inclure de nouveaux indicateurs (primaires et secondaires), et affiner certaines des mesures existantes pour développer davantage une approche pour mesurer non seulement les *conditions* de gouvernance et d'utilisation des données pour le bien public, mais aussi pour fournir de meilleures preuves sur la question à savoir si tous les résidents des pays à travers le monde vivent dans des contextes où l'effet net des données dans la société est de conduire un développement durable et équitable qui soit aligné sur les visions définies à l'échelle mondiale et communautaire du bien public.

Nous avons identifié des domaines où un certain nombre de nos indicateurs, et la formation des chercheurs pour les opérationnaliser, pourraient être renforcés : et ceci a parfois eu des impacts sur la qualité des données générées. Avec 107 389 points de données dans l'enquête, nous ne pouvons pas garantir que chaque valeur soit exempte d'erreur. En fait, nous sommes certains qu'il y a des réponses demeurant discutables : que ce soit en raison d'une interprétation différente des directives entre les chercheurs et les évaluateurs, de faux négatifs lorsque des sources sont passées inaperçues ou de faux positifs lorsqu'une source a été interprétée de façon trop généreuse. Bien que le processus d'examen ait résolu bon nombre de ces questions, il ne les a pas toutes saisies. Plutôt que de cacher ces limites des données en gardant notre matériel source fermé, nous préférons le chemin du partage de nos données de recherche complètes, erreurs et tout, afin que les réutilisateurs puissent effectuer leurs propres évaluations quant à l'exactitude des indicateurs particuliers, et calibrer la façon dont ils utilisent les données pour leurs propres cas d'utilisation particuliers. En conclusion, lorsque l'on envisage l'utilisation de données pour le bien public, toutes les données doivent être abordées de manière critique, y compris la nôtre.

# Annexe : Méthodologie

Cette annexe contient des détails sur les pondérations appliquées pour le calcul des scores du Baromètre mondial de données. Plus de détails sur la méthode d'enquête utilisée pour la collecte des données peuvent être trouvés dans le manuel de recherche qui n'est pas reproduit ici.

## Structure

Le baromètre est structuré autour de quatre piliers et se compose de 39 indicateurs primaires tirés d'une enquête d'experts et de 14 indicateurs secondaires.

| Gouvernance  | Capacités   | Disponibilité   | Utilisation et impact  |
|--|---|---|--|
|    |   |    |  |
| Couvrant la gestion ; la protection et le partage des données ainsi que des données ouvertes, ainsi que de l'examen de la manière dont les données sont traitées au sein des réglementations sectorielles. | Couvrant les fondations (par exemple la connectivité ; l'éducation), les capacités et les institutions gouvernementales en matière de données, et les capacités au sein du secteur privé et de la société civile. | Aborder l'existence des données, les caractéristiques (la qualité), l'ouverture et la couverture (étendue) à travers un certain nombre de domaines thématiques. | Analyse préliminaire basée sur des exemples de cas d'utilisation.                    |
| Pondération : 30%  | Pondération : 24%   | Pondération : 42%   | Pondération : 4%   |

Compte tenu de la qualité des données, de la sensibilité de chaque indicateur et du nombre d'indicateurs au sein de chaque pilier ou module, et à la suite de discussions avec le conseil consultatif, sept principes de pondération ont été utilisés : La contribution de chaque indicateur aux scores des modules et des piliers est aussi pondérée.

Chaque indicateur primaire est calculé sur la base des réponses pondérées données aux sous-questions fermées organisées en un ensemble de sous-sections. Chaque indicateur primaire a une justification écrite et des liens de source. Les sous-questions d'indicateurs peuvent avoir des questions ouvertes à l'appui demandant des preuves.

À quelques exceptions près, les indicateurs sont structurés dans les sous-sections suivantes:

- **Existence** - évaluer la présence de preuves de l'existence d'un cadre de gouvernance, d'une capacité, d'un type ou d'une utilisation des données au sein du pays, et la nature de celle-ci.
- **Éléments** - généralement divisés en deux parties pour fins d'évaluation :
  - **Caractéristiques liées à la qualité** d'une règle, d'un ensemble de données ou d'une capacité. Le choix des caractéristiques, la ou c'est possible, était fondé à la fois sur des normes internationales largement acceptées et sur des cas

- d'utilisation clairs.
- **Caractéristiques (ouvertes) liées aux données** d'une règle ou d'un ensemble de données, en utilisant un ensemble commun de sous-questions dans le but d'évaluer des questions telles que les références/la présence de données structurées et de licences.
- **Étendue** - évaluer si la gouvernance, la capacité ou les données évaluées sont applicables, pertinentes ou utiles à travers le pays et pour tous les citoyens, ou si celles-ci présentent des limites.

Les données du baromètre sont donc organisées en la hiérarchie suivante, avec différentes informations capturées ou calculées à chaque niveau.

- **Score global** - Score pondéré
  - **Piliers** ou module (1) Score pondéré
    - **Indicateur** - Score pondéré, justifications et preuves
      - **Sous-section** - Score pondéré ou multiplicateur
        - **Sous-question** - Réponses brutes, scores, scores pondérés, données à l'appui

Le manuel de recherche complet, avec les détails de chaque indicateur principal, sous-questions et points de données justificatifs demandés, peut être consulté à l'adresse <https://handbook.globaldatabarometer.org/2021/>.

## Multiplicateurs et pondérations au niveau des questions

Le design du baromètre vise à fournir à la fois une gamme de points de données discrets pouvant soutenir la recherche exploratoire ou de test d'hypothèses sur des ensembles de données **spécifiques**, les utilisations de données, les capacités ou les règles de gouvernance et de fournir des indicateurs et des mesures qui soutiennent une comparaison de plus haut niveau entre pays et contextes. Nous présentons, pour cette raison, les données brutes collectées depuis notre enquête d'experts et, pour calculer des indicateurs comparables, nous appliquons un certain nombre de règles de pondération et de multiplicateur aux points de données. Nous cherchons à trouver, à cet égard, un équilibre entre le respect du caractère unique de chaque question ou thème étudié et l'adoption d'une approche commune.

- **Des pondérations fixes** sont attribués à chaque sous-question **d'éléments**.
- **Les multiplicateurs** sont utilisés pour modifier le score global de l'indicateur en fonction des sous-questions relatives à l'**existence** et à l'**étendue**.

### Score et pondération d'éléments

La section d'éléments de chaque indicateur est initialement notée sur une plage de 0 à 100 avant l'application des multiplicateurs.

La section d'élément est divisée en groupes (e1, e2 et e3). Au sein du groupe, les sous-questions sont généralement pondérées de manière égale, avec des ajustements mineurs pour des raisons théoriques. La pondération de chaque groupe est fixé en fonction de la justification de la question et cherche à équilibrer l'importance relative des sous-questions pour une évaluation globale de l'indicateur en s'appuyant sur les cadres établis et les précédents, ainsi qu'à soutenir la comparabilité entre les

indicateurs. Dans les indicateurs de disponibilité, si par exemple, plus de 60 % du score de l'élément est constitué d'une liste de contrôle commune des propriétés de l'ensemble de données interrogé, tandis qu'environ 40 % du score provient d'éléments spécifiques de l'ensemble de données.

Pour certains indicateurs, une quatrième série de questions élémentaires est présente (étiquetées « eb ») qui comptent pour le score de cet indicateur. Ceux-ci peuvent généralement enlever jusqu'à 20 points du score des éléments.

La pondération attribuée à chaque groupe d'éléments et sous-questions peuvent être trouvés dans la colonne de pondération de l'ensemble de données du baromètre.

Chaque sous-question individuelle peut être répondue par « Oui » (1), « Partiellement » (0,5) ou « Non » (0). Le manuel de recherche fournit des conseils généraux sur la manière d'utiliser la réponse partielle, et des conseils spécifiques sont fournis pour certaines questions (détaillées dans le manuel sous ces sous-questions).

## Exemple traité

L'exemple suivant indique les éléments de l'indicateur « Dans quelle mesure les lois, réglementations, politiques et orientations pertinentes fournissent-elles un cadre complet pour la protection des données à caractère personnel ? »

| Nom de la variable                               | Sous-question   | Pondération | Remarque  |
|--|---|-------------|---|
| G.GOVERNANCE.DPL.e                               | Éléments  | <b>100</b>  | Le score le plus élevé pour tous les éléments s'élèverait à 100   |
| G.GOVERNANCE.DPL.e.e1                            | (Groupe d'éléments) Droits et responsabilités   | <b>80</b>   | Pour cet indicateur, le groupe d'éléments e1 vaut 80 des 100 points disponibles. Il y a cinq éléments dans ce groupe, chacun vaut donc 16 points. |
| G.GOVERNANCE.DPL.e.e1.CONSENT                    | Le cadre confère aux personnes concernées des droits de choix ou de consentement.   | 16          | -   |
| G.GOVERNANCE.DPL.e.e1.ACCESS_CORRECTION          | Ce cadre confère aux sujets des données le droit d'accès et de correction aux données les concernant.                         | 16          | -   |
| G.GOVERNANCE.DPL.e.e1.DATA_HOLDER_RESPONSIBILITY | Le cadre définit clairement les responsabilités des détenteurs de données.  | 16          | -   |
| G.GOVERNANCE.DPL.e.e1.RECTIFICATION              | Le cadre prévoit des droits de recours  | 16          | -   |
| G.GOVERNANCE.DPL.e.e1.BREACH                     | Le cadre exige des responsables du traitement des données qu'ils informent une autorité compétente des violations de données. | 16          | -   |
| G.GOVERNANCE.DPL.e.e2                            | (Groupe d'éléments) Considérations particulières  | <b>20</b>   | Pour cet indicateur, le groupe d'éléments e2 vaut 20 des 100 points disponibles. Il y a deux éléments dans ce groupe, chacun vaut donc 10 points. |
| G.GOVERNANCE.DPL.e.e2.LOCATION                   | Les cadres couvrent explicitement la protection des données liées à la localisation.  | 10          | -   |

| Nom de la variable                    | Sous-question  | Pondération | Remarque  |
|---------------------------------------|--|-------------|---|
| G.GOVERNANCE.DPL.e.e2.IA              | Le cadre traite de la prise de décision algorithmique.   | 10          | -   |
| G.GOVERNANCE.DPL.e.eb                 | (Groupe d'éléments) Notation négative  | -20         | Pour cet indicateur, le groupe d'éléments eb peut enlever jusqu'à 20 points du score obtenu |
| G.GOVERNANCE.DPL.e.eb.COVIDEXCEPTIONS | Des exceptions au cadre habituel de protection des données ont été faites dans le cadre de la réponse du pays à la COVID-19. | -20         | -   |

Si nous imaginons un pays où le chercheur répond « Oui » au CONSENTEMENT (1 (score) x (pondération) 16), ACCESS\_CORRECTION (1 x 16) et DATA HOLDER RESPONSIBILITY (1 x 16), « Partiellement » à la RÉPARATION (0,5 x 16), à la VIOLATION (0,5 x 16) et à l'EMPLACEMENT (0,5 x 10) et COVIDEXCEPTIONS (0,5 x - 20), et « Non » à l'IA (0 x 10), alors nous calculerions le score de l'élément comme 59.

## Multiplicateurs d'existence

Un multiplicateur est appliqué au score des éléments en fonction des questions d'existence. Différentes approches multiplicatrices sont adoptées dans les différents piliers du baromètre.

### Gouvernance

Les questions fondamentales de gouvernance demandent dans quelle mesure un cadre de gouvernance particulier existe et quelle forme celui-ci prend. Là où les chercheurs n'ont trouvé aucun cadre, il leur a été demandé : « En l'absence d'un cadre juridique solide, existe-t-il d'autres normes ou coutumes qui jouent ce rôle dans le pays ? » et, dans certains cas, ont fourni des données structurées à ce sujet. Pour supprimer les valeurs de notation des réponses où aucun cadre n'est présent, dans de tels cas, le score total de l'indicateur est alors multiplié par 0.

Lorsqu'un cadre existe mais n'a pas force de loi, le score global de l'indicateur est légèrement réduit. Ceci signifie que s'il y a deux pays ayant des caractéristiques égales dans leurs cadres, mais que l'un a force de loi, et que l'autre ne l'a pas, le premier pays obtiendra des résultats plus élevés.

| Réponse   | Multiplicateur |
|---|----------------|
| Aucun cadre n'existe.   | 0              |
| Il existe un cadre juridique, mais celui-ci n'a pas force de loi. | 0,9            |
| Un cadre existe et a force de loi.                                | 1              |

Les questions thématiques sur l'existence de la gouvernance se divisent en deux parties : à savoir s'il existe des cadres ou sont à l'état d'ébauche, et de connaître la force des règles relatives aux données.

Les multiplicateurs suivants sont appliqués :

| Réponse   | Multiplicateur |
|---|----------------|
| Les cadres n'existent pas.  | 0              |
| Les cadres sont en cours d'élaboration ou ne sont pas encore appliqués. | 0.6            |
| Des cadres existent et sont opérationnels.                              | 1              |

| Réponse  | Multiplicateur |
|--|----------------|
| Il n'y a aucune mention de données ou de leur publication dans les lois, politiques ou directives.                                   | 0.6            |
| Les exigences relatives à la publication des données sont énoncées dans une politique ou des directives non-contraignantes.          | 0.85           |
| Les exigences relatives à la publication des données sont énoncées dans des politiques, des règlements ou des lois contraignantes.   | 0,95           |
| Les exigences relatives à la publication des données sont énoncées dans des politiques, des règlements ou des lois contraignantes ou | 1              |

Les questions de gouvernance « Couverture d'accessibilité & données » n'utilisent pas de multiplicateur d'existence, et au lieu de ça les deux questions d'existence sont traitées en tant qu'éléments, allouant jusqu'à 40 points au score de l'indicateur de base.

## Questions sur les capacités

Les questions sur l'existence des capacités utilisent une échelle de trois ou quatre points, avec des multiplicateurs attribués en conséquence.

| Réponse | Multiplicateur sur une échelle de 3 points | Multiplicateur sur une échelle de 4 points |
|---------|--|--|
| 0       | 0  | 0  |
| 1       | 0,8  | 0.6  |
| 2       | 1  | 0,8  |
| 3       | -  | 1  |

## Disponibilité

Les indicateurs de disponibilité demandent si les données sont disponibles en ligne, et si cela résulte d'une action gouvernementale ou non. Les notes les plus élevées sont réservées aux cas où les données sont fournies par le gouvernement, mais les chercheurs sont en mesure d'utiliser des questions élémentaires pour décrire les caractéristiques d'autres méthodes d'accès aux données, offrant des aperçus qualitatifs clés sur la façon dont les données sont fournies lorsque les gouvernements n'offrent pas un accès fiable à celles-ci. Les multiplicateurs suivants sont appliqués:

| Réponse  | Multiplicateur |
|--|----------------|
| Les données ne sont pas disponibles en ligne   | 0              |
| Les données sont disponibles, mais ne résultent pas d'une action gouvernementale.      | 0,5            |
| Les données sont disponibles du gouvernement, ou en raison de mesures gouvernementales | 1              |

## Utilisation

Les indicateurs d'utilisation demandent s'il y a des cas d'utilisation identifiés et s'il s'agit de cas isolés ou répandus. Les scores les plus élevés des indicateurs sont réservés à une utilisation généralisée des données.

| Réponse          | Multiplicateur |
|------------------|----------------|
| Aucune preuve    | 0              |
| Cas isolés       | 0,5            |
| Un nombre de cas | 0,9            |
| Cas répandus     | 1              |

## Multiplicateurs d'étendue

Les indicateurs ont des questions de portée différente, selon que l'on puisse considérer que la couverture nationale, celle de tous les ministères et organismes et ainsi de suite sont complètes. Certains indicateurs comportent une seule question sur la portée à trois éléments. D'autres ont deux questions d'étendue, conduisant à 6 valeurs possibles ou plus dans la note d'étendue. Le calcul du multiplicateur est basé sur un seuil tel que :

- Le niveau d'étendue le plus limité tel qu'évalué par les sous-questions reçoit une pondération de 0,7
- Les réponses atteignant le niveau intermédiaire sur l'étendue reçoivent une pondération de 0,85
- Lorsqu'il y a plus de trois options d'étendue, les réponses entre le niveau intermédiaire et le niveau supérieur reçoivent 0,9
- Les valeurs les plus élevées de l'étendue évaluées par les sub-questions reçoivent une pondération de 1

## Application des multiplicateurs : exemple concret

À l'aide de l'exemple concret présenté précédemment, l'indicateur de gouvernance « Dans quelle mesure les lois, règlements, politiques et lignes directrices pertinents fournissent-ils un cadre complet pour la protection des données à caractère personnel ? » a :

- Une question d'existence demandant si le cadre a la pleine force de loi
- Deux questions d'étendue, se demandant dans quelle mesure le cadre s'applique (par exemple, à différents secteurs) et à la couverture géographique (c'est-à-dire est-ce applicable à l'ensemble du pays ?).

Imaginons que le meilleur cadre de protection des données au pays manque de pleine force de loi (multiplicateur d'existence de 0,9), et qu'il s'applique à l'ensemble du pays mais pas à tous les secteurs (donnant une réponse entre le niveau moyen et le niveau le plus élevé disponible, donc un multiplicateur d'étendue de 0,85).

Pour obtenir le score final de l'indicateur, nous multiplions le score des éléments comme suit :

| Score de l'élément x Multiplicateur d'existence x Multiplicateur d'étendue

Dans notre exemple, ceci donne :

|  $59 * 0,9 * 0,85 = 45,135$

Le score final de l'indicateur est donc de 45.135.

## Trouver des pondérations de sous-question

Toutes les pondérations de questions utilisées se trouvent dans l'ensemble de données du Baromètre publié (Disponible à l'adresse <https://www.globaldatabarometer.org>) contenant un certain nombre de champs clés, qui, combinés, indiquent les pondérations ayant été appliquées.

Pour trouver la pondération de chaque sous-question :

- Filtrer sur la colonne de niveau à niveau =4 (niveau hiérarchique = 4 = sous-questions)
- Filtrer sur la colonne data\_type vers data\_type='response' (pour voir les réponses plutôt que les justifications, les données de soutien, les exemples ou tout autre contenu)

La pondération peut alors être lu depuis les champs suivants :

| champ                     | description  |
|---------------------------|--|
| response                  | La réponse sélectionnée ou saisie par le chercheur. Pour data_type='response', il s'agira généralement d'une option fixe fournie par l'enquête.  |
| response_value            | La valeur numérique attribuée au résultat. Pour les questions avec « Non » (0), « Partiellement » (1), « Oui » (2), les réponses seront comprises entre 0 et 2. Pour certaines questions d'existence et d'étendue, ceci peut aller de 0 à 3 ou de 0 à 4. |
| normalized_response_value | Chaque score est réévalué sur une échelle de 0 à 1   |
| weight                    | Un multiplicateur appliqué au score calculé pour obtenir la contribution pondérée de cette question à l'indicateur. À quelques exceptions près, les pondérations des sous-questions sont gardées constantes dans la même sous-section de l'indicateur.   |
| score                     | normalized_response_value*weight   |

## Pondérations au niveau de l'indicateur

La contribution de chaque indicateur aux scores des modules et des piliers est aussi pondérée. Compte tenu de la qualité des données, de la sensibilité de chaque indicateur et du nombre d'indicateurs au sein de chaque pilier ou module, et à la suite de discussions avec le conseil consultatif, sept principes de pondération ont été utilisés :

- Les données primaires de l'un des deux modules de base de la BDG (gouvernance et capacité) reçoivent la priorité de la première pondération.
- Les données primaires du module de base Gouvernance pèsent légèrement plus que celles provenant de Capacités.
- Les données primaires de l'un des modules thématiques de la BDG (de l'un des quatre piliers) reçoivent la priorité de pondération suivante.
- Les données primaires provenant de la disponibilité pèsent légèrement plus que les autres données des modules thématiques de la BDG.
- Les données secondaires présentées sous forme d'indice (traitement de plusieurs variables) reçoivent la cinquième priorité de pondération.

- Les données secondaires présentées en tant que métrique (traitement de quelques variables qui appartiennent généralement à un indice) reçoivent la sixième priorité de pondération.
- Les données secondaires présentées en tant que variables dichotomiques uniques reçoivent la septième priorité de pondération.

Sur la base de ces règles et en étudiant diverses autres options et alternatives, les piliers individuels ont été pondérés comme suit : gouvernance 0,32 ; capacité 0,20 ; disponibilité 0,40 ; et utilisation 0,08.

Les tableaux ci-dessous présentent tous les indicateurs de chaque pilier, ainsi que leur pondération.

## Gouvernance

Le pilier gouvernance du Baromètre est composé de 14 indicateurs (13 primaires, 1 secondaire). Un indicateur (couverture linguistique et données) n'est pas compris dans la notation du pilier, parce que les vérifications post-enquête sur la qualité des données suggèrent que les réponses n'étaient pas suffisamment solides pour être fiables, bien que les données qualitatives pour cet indicateur soient toujours incluses dans l'ensemble de données du Baromètre pour soutenir les travaux futurs.

| Nom de l'indicateur                      | Module                    | Type     | Question / Source   | Pondération au sein du pilier | Pondération au sein du module |
|--|---------------------------|----------|---|-------------------------------|-------------------------------|
| Protection des données                   | Gouvernance               | Primaire | Jusqu'à quel point les lois, réglementations, politiques et orientations pertinentes fournissent-elles un cadre complet pour la protection des données personnelles ?   | 0,1                           | 0,2000                        |
| Politique de Données ouvertes            | Gouvernance               | Primaire | Jusqu'à quel point les lois, réglementations, politiques et orientations pertinentes fournissent-elles un cadre complet pour la génération et la publication des données ouvertes ?   | 0,1                           | 0,2000                        |
| Cadres de partage des données            | Gouvernance               | Primaire | Jusqu'à quel point les lois, réglementations, politiques et orientations pertinentes fournissent-elles un cadre complet pour le partage des données ?   | 0,1                           | 0,2000                        |
| Gestion des données                      | Gouvernance               | Primaire | Jusqu'à quel point les lois, réglementations, politiques et orientations pertinentes fournissent-elles un cadre complet pour la gestion et la publication cohérentes des données ?  | 0,1                           | 0,2000                        |
| Couverture linguistique et données       | Gouvernance               | Primaire | Jusqu'à quel point les lois, règlements, politiques et directives pertinents exigent-ils que les processus de collecte et de publication des données soient disponibles dans les langues officielles ou nationales du pays ? Si le pays n'a pas de langue officielle ou nationale, ces processus sont-ils disponibles dans les langues utilisées dans le pays ? | 0                             | 0                             |
| Couverture de l'accessibilité et données | Gouvernance               | Primaire | Jusqu'à quel point les lois, règlements, politiques et les lignes directrices pertinentes exigent-ils que la collecte et la publication des données soient accessibles aux personnes handicapées ?  | 0,1                           | 0,2000                        |
| Ayant droit économique                   | Informations d'entreprise | Primaire | Jusqu'à quel point les lois, règlements, politiques et lignes directrices pertinentes fournissent-elles une base pour la collecte et la publication de données sur les bénéficiaires effectifs des entreprises ?  | 0,065                         | 0,275                         |
| Finances politiques                      | Intégrité politique       | Primaire | Jusqu'à quel point les lois, règlements, politiques et les directives pertinentes fournissent-ils une   | 0,065                         | 0,075                         |

| Nom de l'indicateur              | Module              | Type       | Question / Source   | Pondération au sein du pilier | Pondération au sein du module |
|----------------------------------|---------------------|------------|---|-------------------------------|-------------------------------|
| Déclarations de patrimoine       | Intégrité politique | Primaire   | base pour la collecte et la publication de données sur le financement des campagnes et des partis ?<br>Jusqu'à quel point les lois, règlements, politiques et directives pertinents fournissent-ils une base pour la collecte et la publication de données sur les intérêts et les biens des agents publics ? | 0,065                         | 0,075                         |
| Registre de lobbying             | Intégrité politique | Primaire   | Jusqu'à quel point les lois, règlements, politiques et orientations pertinents fournissent-ils une base pour la collecte et la publication de données sur les activités de lobbying ?   | 0,065                         | 0,075                         |
| Données de consultation publique | Intégrité politique | Primaire   | Jusqu'à quel point les lois, règlements, politiques et les lignes directrices pertinentes fournissent-ils une base pour la collecte et la publication des données générées via la consultation publique sur l'élaboration des règlements et sur la consultation publique à ce sujet ?                         | 0,065                         | 0,075                         |
| Performance RTI                  | Intégrité politique | Primaire   | Jusqu'à quel point les lois, règlements, politiques et orientations pertinents fournissent-ils une base pour la collecte et la publication de données sur l'exécution des processus relatifs au droit à l'information (RTI) /à la liberté d'information (FOI) ?   | 0,065                         | 0,075                         |
| Finances publiques               | Finances publiques  | Primaire   | Jusqu'à quel point les lois, règlements, politiques et orientations pertinents fournissent-ils une base pour la collecte et la publication de données sur les finances publiques ? (Par exemple, les budgets gouvernementaux, les dépenses gouvernementales, la dette et les emprunts.)                       | 0,065                         | 0,45                          |
| Cadre RTI                        | Intégrité politique | Secondaire | <u>Cote RTI</u> : Cote RTI  | 0,045                         | 0,052                         |

## Capacité

Le pilier Capacités du Baromètre est composé de 5 indicateurs primaires et de 10 indicateurs secondaires.

| Nom de l'indicateur                       | Module              | Type     | Question / Source  | Pondération au sein du pilier | Pondération au sein du module |
|---|---------------------|----------|--|-------------------------------|-------------------------------|
| Fonction publique                         | Capacités           | Primaire | Jusqu'à quel point le gouvernement dispense-t-il une formation pour développer les compétences des fonctionnaires en matière de données ?  | 0,12                          | 0,9500                        |
| Initiative sur les données ouvertes       | Capacités           | Primaire | Jusqu'à quel point existe-t-il une initiative de données gouvernementales ouvertes bien financée au sein du pays ?   | 0,12                          | 0,9500                        |
| Soutien gouvernemental à la réutilisation | Capacités           | Primaire | Jusqu'à quel point des preuves existe-t-elles de l'appui du gouvernement à la réutilisation des données ?  | 0,12                          | 0,9500                        |
| Infra-national                            | Capacités           | Primaire | Jusqu'à quel point les administrations municipales, régionales et locales ont-elles une capacité efficace de gestion des données ?   | 0,12                          | 0,9500                        |
| Interopérabilité de l'intégrité politique | Intégrité politique | Primaire | Jusqu'à quel point les données sur l'intégrité politique sont-elles interopérables entre différents ensembles de données sur celle-ci, ainsi qu'avec d'autres ensembles de données associés aux flux d'informations pertinents ? | 0,08                          | 0,075                         |

| Nom de l'indicateur   | Module    | Type       | Question / Source   | Pondération au sein du pilier | Pondération au sein du module |
|---|-----------|------------|---|-------------------------------|-------------------------------|
| Gouvernement numérique  | Capacités | Secondaire | <u>Banque mondiale - Ensemble de données DGSS</u> : Existe-t-il une stratégie DG/GovTech ?; Existe-t-il une institution GovTech dédiée ; Existe-t-il une stratégie nationale sur les technologies perturbatrices ?; Existe-t-il un cloud gouvernemental (plateforme partagée) ?; Existe-t-il une plateforme bus gouvernementale de service/d'interopérabilité ? | 0,0575                        | -                             |
| Services gouvernementaux en ligne                                 | Capacités | Secondaire | <u>Enquête de l'ONU sur le gouvernement électronique</u> : Indice des services en ligne eGov de l'ONU (2020)  | 0,0575                        | 0,07250                       |
| Capital humain  | Capacités | Secondaire | <u>Enquête des Nations Unies sur l'e-gouvernement</u> : Indice du capital humain eGov des Nations Unies (2020)  | 0,0575                        | 0,07250                       |
| Libertés politiques et civiles                                    | Capacités | Secondaire | <u>FreedomHouse</u> : Score des Droits Politiques & Score des Libertés Civiles  | 0,0575                        | 0,07250                       |
| Utilisation des outils numériques par les entreprises             | Capacités | Secondaire | <u>Forum EM</u> : Utilisation des outils numériques par les entreprises   | 0,035                         | 0,05500                       |
| Institutions de données   | Capacités | Secondaire | <u>Banque mondiale - Ensemble de données DGSS</u> : Existe-t-il une entité gouvernementale chargée de la gouvernance ou de la gestion des données ? Existe-t-il une autorité de protection des données ?  | 0,035                         | 0,05500                       |
| Accès Internet  | Capacités | Secondaire | <u>Union internationale des télécommunications (UIT)</u> : Panier fixe à large bande en % du RNB p.c. ; utilisateurs d'internet, total (%)  | 0,035                         | 0,05500                       |
| Emploi à forte intensité de connaissances                         | Capacités | Secondaire | <u>Indice mondial de l'innovation/OIT</u> : Emploi dans les professions à forte intensité de connaissances (% de la main-d' œuvre)  | 0,035                         | 0,05500                       |
| Utilisation de normes et méthodes dans les bureaux de statistique | Capacités | Secondaire | <u>Indicateurs de performance statistique</u> : Dimension 5.2 : Normes et méthodes  | 0,035                         | 0,05500                       |
| Compétences numériques  | Capacités | Secondaire | <u>Forum EM</u> : Compétences numériques au sein de la population active (score)  | 0,035                         | 0,05500                       |

## Disponibilité

Le pilier de la disponibilité est composé de 17 indicateurs primaires et de 2 indicateurs secondaires. Un autre indicateur secondaire est compris dans l'ensemble de données, mais on lui a attribué une pondération nulle lors de l'examen de l'indice parce que celui-ci s'est avéré ne permettant pas de suivre adéquatement les différences entre les pays.

| Nom de l'indicateur         | Module                    | Type     | Question / Source   | Pondération au sein du pilier | Pondération au sein du module |
|-----------------------------|---------------------------|----------|---|-------------------------------|-------------------------------|
| Ayant droit économique      | Informations d'entreprise | Primaire | Jusqu'à quel point les informations de l'entreprise sont-elles disponibles sous forme de données ouvertes structurées ? | 0,0570                        | 0,325                         |
| Registre du commerce et des | Informations d'entreprise | Primaire | Jusqu'à quel point les informations de l'entreprise sont-elles disponibles sous forme de données ouvertes structurées ? | 0,0570                        | 0,325                         |

| Nom de l'indicateur                        | Module                | Type       | Question / Source  | Pondération au sein du pilier | Pondération au sein du module |
|--|-----------------------|------------|--|-------------------------------|-------------------------------|
| sociétés                                   |                       |            |  |                               |                               |
| Régime foncier                             | Terrains              | Primaire   | Jusqu'à quel point des informations détaillées sur le régime foncier sont-elles disponibles sous forme de données ouvertes ?   | 0,0570                        | 0,45                          |
| Utilisation actuelle des terrains;         | Terrains              | Primaire   | Jusqu'à quel point les informations existantes sur l'utilisation des terrains sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,45                          |
| Données financières politiques             | Intégrité politique   | Primaire   | Jusqu'à quel point les informations sur le financement politique sont-elles disponibles sous forme de données ouvertes ?   | 0,0570                        | 0,095                         |
| Déclarations de patrimoine                 | Intégrité politique   | Primaire   | Jusqu'à quel point les informations relatives aux déclarations d'intérêts et d'actifs sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,095                         |
| Données de lobbying                        | Intégrité politique   | Primaire   | Jusqu'à quel point les informations des registres de lobbying sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,095                         |
| Données de consultation publique           | Intégrité politique   | Primaire   | Jusqu'à quel point les informations issues de la consultation publique sont-elles disponibles sous forme de données ouvertes ?   | 0,0570                        | 0,095                         |
| Données de performance RTI                 | Intégrité politique   | Primaire   | Jusqu'à quel point les informations détaillées sur le rendement RTI sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,095                         |
| Données sur le budget et les dépenses      | Finances publiques    | Primaire   | Jusqu'à quel point les informations sur le budget et les dépenses du gouvernement (l'exécution du budget) sont-elles disponibles sous forme de données ouvertes structurées ?            | 0,0570                        | 0,55                          |
| Données des marchés publics                | Achats                | Primaire   | Jusqu'à quel point des données structurées détaillées sur les procédures de passation des marchés publics sont-elles disponibles comme données ouvertes ?                                | 0,0570                        | 0,825                         |
| Émission                                   | Action pour le climat | Primaire   | Jusqu'à quel point les informations sur les émissions sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,3334                        |
| Biodiversité                               | Action pour le climat | Primaire   | Jusqu'à quel point les informations sur les espèces et les écosystèmes menacés sont-elles disponibles sous forme de données ouvertes ?   | 0,0570                        | 0,3333                        |
| Vulnérabilité                              | Action pour le climat | Primaire   | Jusqu'à quel point les informations sur la vulnérabilité climatique sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,3333                        |
| Statistiques vitales                       | Santé & Covid-19      | Primaire   | Jusqu'à quel point les informations relatives à l'enregistrement des faits d'état civil et aux statistiques de l'état civil (CRVS) sont-elles disponibles en tant que données ouvertes ? | 0,0570                        | 0,26                          |
| Capacité du système de santé en temps réel | Santé & Covid-19      | Primaire   | Jusqu'à quel point les informations sur la capacité en temps réel du système de santé sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,26                          |
| Vaccination (COVID-19)                     | Santé & Covid-19      | Primaire   | Jusqu'à quel point les informations de vaccination contre la COVID-19 sont-elles disponibles sous forme de données ouvertes ?  | 0,0570                        | 0,26                          |
| Capacité du système de santé               | Santé & Covid-19      | Secondaire | <u>ODW Open Data Index</u> : Établissements de santé   | 0,02                          | 0,14                          |
| Données de test (COVID-19)                 | Santé & Covid-19      | Secondaire | <u>Organisation mondiale de la santé (OMS)</u> : Cas - total cumulatif   | 0,011                         | 0,08                          |

## Utilisation et impact

Le pilier d'utilisation et d'impact est composé de quatre indicateurs primaires et d'un indicateur secondaire.

| Nom de l'indicateur  | Module                    | Type       | Question / Source  | Pondération au sein du pilier | Pondération au sein du module |
|--|---------------------------|------------|--|-------------------------------|-------------------------------|
| Diligence raisonnable d'entreprise   | Informations d'entreprise | Primaire   | Jusqu'à quel point existe-t-il des produits ou des services qui utilisent des données d'entreprise ouvertes au soutien de la diligence raisonnable ?   | 0,22                          | 0,075                         |
| Influence de la politique en faveur de l'égalité des sexes et de l'inclusion | Terrains                  | Primaire   | Jusqu'à quel point est-il prouvé que les données foncières sont utilisées pour influencer les politiques dans l'intérêt d'un régime foncier et d'une utilisation équitable et inclusive des terrains ? | 0,22                          | 0,1                           |
| Utilisations de l'IP à des fins de responsabilisation                        | Intégrité politique       | Primaire   | Jusqu'à quel point les données sur l'intégrité politique sont-elles utilisées pour identifier, exposer ou mettre en évidence les défaillances gouvernementales ?                                       | 0,22                          | 0,023                         |
| Analyse des données d'approvisionnement                                      | Achats                    | Primaire   | Jusqu'à quel point les données sur les marchés publics sont-elles analysées pour l'amélioration des pratiques d'approvisionnement ?  | 0,22                          | 0,175                         |
| Utilisation des données par les organisations internationales                | Utilisation et impact     | Secondaire | <u>Indicateurs statistiques de résultats:</u><br>Dimension 1.5 : Utilisation des données par les organisations internationales   | 0,12                          | -                             |

## Références

1. UN IEAG. A world that counts—mobilising the data revolution for sustainable development. UN Secretary-General's Independent Expert Advisory Group on the Data Revolution for Sustainable Development; 2014. Available from: <https://www.undatarevolution.org/report/>
2. Kitchin R. The data revolution: Big data, open data, data infrastructures and their consequences. Sage; 2014.
3. Multiple Authors. Public good (economics). In: Wikipedia. Wikipedia; 2022. Available from: [https://en.wikipedia.org/w/index.php?title=Public\\_good\\_\(economics\)&oldid=1074988450](https://en.wikipedia.org/w/index.php?title=Public_good_(economics)&oldid=1074988450)
4. Multiple Authors. Network effect. In: Wikipedia. 2022. Available from: [https://en.wikipedia.org/w/index.php?title=Network\\_effect&oldid=1081034493](https://en.wikipedia.org/w/index.php?title=Network_effect&oldid=1081034493)
5. Open Data Institute. The Data Spectrum. 2016. Available from: <https://theodi.org/about-the-odi/the-data-spectrum/>
6. Baack S, Maxwell M. Alternative Data Governance Approaches: Global Landscape Scan and Analysis. Mozilla Foundation; 2020. Available from: <https://assets.mofoprod.net/network/documents/DataGovernanceApproaches.pdf>
7. De La Chapelle B, Porciuncula L. We Need to Talk About Data: Framing the Debate Around the Free Flow of Data and Data Sovereignty. Internet & Jurisdiction Policy Network; 2021. Available from: <https://www.internetjurisdiction.net/news/aboutdata-report>
8. World Wide Web Foundation. Open Data Barometer. 2013–2017. Available from: <https://opendatabarometer.org/>
9. Open Knowledge Foundation. Open Data Index. 2011–2017. Available from: <https://index.okfn.org/>

10. ILDA. Regional Open Data Barometer Latin America and the Caribbean 2020. 2020. Available from: <https://barometerlac.org/wp-content/themes/odbpress/reporte-ILDA-EN.pdf>
11. Davies T, Walker SB, Rubinstein M, Perini F. The state of open data: Histories and horizons. African Minds; 2019. Available from: <https://stateofopendata.od4d.net/>
12. Greenleaf G. Global Data Privacy Laws 2017: 120 National Data Privacy Laws, Including Indonesia and Turkey. Rochester, NY: Social Science Research Network; 2017. Available from: <https://papers.ssrn.com/abstract=2993035>
13. Union E, Europarat, editors. Handbook on European data protection law. 2018 edition. Luxembourg: Publications Office of the European Union; 2018.
14. Council of Europe. Artificial Intelligence and Data Protection. Council of Europe; 2019. Available from: <https://rm.coe.int/2018-lignes-directrices-sur-l-intelligence-artificielle-et-la-protecti/168098e1b7>
15. EC. European Data Strategy. 2022. Available from: <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>
16. Dang H-AH, Pullinger J, Serajuddin U, Stacy B. Statistical Performance Indicators and Index: A New Tool to Measure Country Statistical Capacity. The World Bank; 2021. Available from: <http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-9570>
17. Micheli M. Public bodies' access to private sector data: The perspectives of twelve European local administrations. First Monday 2022. Available from: <https://firstmonday.org/ojs/index.php/fm/article/view/11720>
18. OKF. The Open Definition - Open Definition - Defining Open in Open Data, Open Content and Open Knowledge. 2016. Available from: <https://opendefinition.org/>
19. Brandusescu A, Iglesias C, Robinson K, Alonso JM, Fagan C, Jellema A, et al. Open Data Barometer : Global report : Fourth edition. 2017. Available from: <https://idl-bnc-idrc.dspacedirect.org/handle/10625/57676>
20. Ritchie H, Mathieu E, Rodés-Guirao L, Appel C, Giattino C, Ortiz-Ospina E, et al. Coronavirus Pandemic (COVID-19). 2022. Available from: <https://ourworldindata.org/coronavirus-testing#source-information-country-by-country>
21. Mathieu E. Commit to transparent COVID data until the WHO declares the pandemic is over. Nature 2022. 602(7898):549–9. Available from: <https://www.nature.com/articles/d41586-022-00424-9>
22. Open Data Charter. Open Up Guide to Climate Action. Open Data Charter; 2020. Available from: <https://open-data-charter.gitbook.io/>
23. Davila J, Barron M, Law T. Towards a Global Norm of Beneficial Ownership Transparency - A scoping study on a strategic approach to achieving a global norm. 2019. Available from: <https://adamsmithinternational.com/app/uploads/2019/07/Towards-a-Global-Norm-of-Beneficial-Ownership-Transparency-Phase-2-Paper-March-2019.pdf>
24. Garcia Aceves R. Open Up Guide: Using Open Data to Combat Corruption (Version 2.0). Open Data Charter; 2018. Available from: <https://open-data-charter.gitbook.io/open-up-guide-using-open-data-to-combat-corruption/>
25. International Budget Partnership. Open Budget Survey. 2020. Available from: <https://www.internationalbudget.org/open-budget-survey>
26. Bayer C-T, Booth K. Open data for improved land governance. Land Portal and Open Data Charter; 2021. Available from: <https://landportal.org/library/resources/open-guide-land-governance-version-20-following-public-comments>
27. World Bank. World development report 2021: Data for better lives. The World Bank; 2021. Available from: <https://www.worldbank.org/en/publication/wdr2021>
28. Interaction. Regional Trends: Sub-Saharan Africa. 2020. Available from: <https://www.interaction.org/globalsituationreport2020/global-situation-report/sub-saharan-africa/>
29. Equal Measures 2030. 2019 SDG Gender Index Report. Equal Measures; 2019. Available from: <https://www.equalmeasures2030.org/2019-sdg-gender-index-report/>