Stratégie Fédérale de la Donnée

Rapport d'analyse

© AdobeStock.com

economie

SPF Economie, P.M.E., Classes moyennes et Energie

.be

Avertissement : la présente étude a été réalisée par un consultant externe. Les analyses, conclusions et recommandations qui y figurent relèvent de la seule responsabilité de ce dernier. Elles ne sauraient être interprétées comme engageant le SPF Economie, ni comme constituant une prise de position ou un engagement politique.

economie

SPF Economie, P.M.E., Classes moyennes et Energie

Rue du Progrès 50 — 1210 Bruxelles

Numéro d'entreprise : 0314.595.348

- 0800 120 33 (numéro gratuit)
- SPFEco
- @spfeconomie
- linkedin.com/company/fod-economie (page bilingue)
- instagram.com/spfeco
- youtube.com/user/SPFEconomie
- economie.fgov.be

Éditrice responsable :
Séverine Waterbley
Présidente du Comité de direction
Rue du Progrès 50 — 1210 Bruxelles

Version internet

182-25

Table des matières

Introduction 5
Partie 1. Rapport sur l'économie belge des données 6
Introduction 6
1. La valeur des données au sein des organisations 7
Principaux résultats 7
1.1. La chaîne de valeur des données 8
1.2. L'utilisation faite des données au sein des entreprises belges 15
2. L'économie de la donnée en Belgique 25
Principaux résultats 25
2.1. Introduction 27
2.2. Les entreprises de données 27
2.3. Le marché des données 31
2.4. L'impact du marché des données sur l'économie belge 35
2.5. Les retombées futures potentielles de l'IA générative 37
3. Les compétences 39
Principaux résultats 39
3.1. Les compétences en data 40
3.2. Les compétences nécessaires à l'ère de l'IA 41
3.3. Le marché du travail dans le secteur des données 44
3.4. Le gender gap 48
4. Culture de la donnée 50
Principaux résultats 50
4.1. L'ouverture des statistiques publiques 52
4.2. Partage de données par les entreprises 54
4.3. Préoccupations citoyennes face au partage de leurs données 54
5. Infrastructure 59
Principaux résultats 59
6. Réglementations 60
Principaux résultats 60
Partie 2. Analyse comparative des stratégies nationales en France, aux Pays-Bas, en Allemagne et au Royaume-Uni 65
1. Stratégies nationales de 4 pays européens 65
1.1. Allemagne 65
1.2. Royaume-Uni 66
1.3. Pays-Bas 67
1.4. France 68
2. Analyse comparative des 4 stratégies 69

2.1. Compétences ... 71
2.2. Valeur des données ... 75
2.3. Collaborations et partages en matière de données ... 82
2.4. Infrastructures de données ... 85
2.5. Réglementations ... 88
3. Entretiens parties prenantes ... 94
3.1. Pays-Bas ... 94
Partie 3. Analyse SWOT et consultation des parties prenantes belges ... 97
Analyse SWOT ... 97
Consultation des parties prenantes ... 100
Partie 4. Une proposition de stratégie fédérale pour la Belgique ... 102
1. Volet stratégique ... 102
Objectif transversal : délivrer des effets concrets grâce à une organisation efficace portée par une gouvernance centrale forte orientée résultats ... 103
Objectif stratégique #1 : augmenter le nombre de talents au service de l'économie de la donnée 103
Objectif stratégique #2 : faire de la Belgique un des territoires européens les plus dynamiques en termes de collaboration et d'innovation entre structures publiques et privées autour de la donnée ... 104
Objectif stratégique #3 : offrir aux parties prenantes belges un cadre réglementaire clair et stable autour de la donnée ... 106
Objectif stratégique #4 : encourager la mise en place d'infrastructures de partage de la donnée sécurisée et performante ... 107
2. Volet opérationnel ... 108
Préambule ... 108
2.1. National Data Facilitator (NDF) ... 108
2.2. Chantiers liés à la collaboration et la standardisation de la donnée ... 110
2.3. Chantiers de réglementation ... 114
2.4. Chantiers Infrastructure ... 115
2.5. Chantiers Formation ... 116
Partie 5. Conclusions et recommandations ... 119

Introduction

La présente étude a pour but d'établir une stratégie à l'échelle de la Belgique permettant de capitaliser sur l'économie des données afin de générer un impact sur l'économie belge dans son ensemble.

L'étude débute par une phase d'analyse visant à déterminer, au travers d'une recherche documentaire, une compréhension de l'économie belge des données (Partie 1) et d'établir une analyse comparative des stratégies nationales des pays voisins de la Belgique (Partie 2).

Ensuite, au travers de consultations de parties prenantes, une analyse SWOT de l'économie belge des données (Partie 3) a été établie ainsi qu'une proposition de stratégie fédérale de la donnée comportant un volet stratégique et un volet opérationnel (Partie 4).

Enfin, le rapport se termine sur les conclusions (Partie 5) reprenant les impacts envisageables via l'implémentation de cette stratégie pour la Belgique et les recommandations concernant les actions à entreprendre afin de permettre à cette stratégie de porter ses fruits.

La combinaison de ces phases d'analyse documentaire et de consultations variées de parties prenantes a permis de construire une proposition concrète avec des chantiers à réaliser bien délimités s'insérant dans un contexte économique mesurable.

La ligne de force de la stratégie fédérale de la donnée résultant de ce travail est une volonté d'exécution pragmatique via un National Data Facilitator. Celui-ci bénéficierait d'un mandat clair provenant des instances à même de soutenir la démarche afin de déployer au maximum les effets de cette stratégie.

La démarche serait également basée sur l'implication maximale des parties prenantes publiques et privées ainsi que de la société civile afin de disséminer la culture de la donnée au travers de toute la société belge et libérer son potentiel de croissance économique.

Partie 1. Rapport sur l'économie belge des données

Introduction

Les données constituent une ressource cruciale sur laquelle reposent les entreprises, les administrations publiques et l'ensemble de notre société. Elles sont au cœur des processus opérationnels de diverses organisations et entités gouvernementales, facilitant la fourniture de services diversifiés, allant des consultations financières aux soins médicaux, en passant par la diffusion d'informations essentielles via les médias. Au-delà de leur utilité dans le cadre professionnel, les données jouent un rôle prépondérant dans la prise de décisions stratégiques, optimisant par exemple la distribution de produits et services pour répondre au mieux aux besoins des consommateurs.

Sur le plan individuel, l'impact des données est tout aussi significatif. Elles permettent aux personnes d'accéder à une multitude de services, de maintenir des liens sociaux avec amis et famille, de gérer efficacement leurs obligations personnelles et professionnelles, et de prendre des décisions éclairées dans presque tous les aspects de leur vie quotidienne. De l'exercice physique à la recherche de partenaires amoureux, les données influencent nos choix et nos actions de manière de plus en plus marquée.

Au-delà de cela, les données ouvrent de nouvelles frontières dans les sciences et les sciences humaines. Elles élargissent notre compréhension de l'univers, approfondissent notre connaissance des dynamiques du changement climatique et révèlent l'influence d'un enseignant sur les résultats scolaires d'un élève. La promesse de valeur que recèlent les données est immense, témoignant de leur potentiel transformateur.

En 2017, The Economist a publié un article intitulé « La ressource la plus précieuse du monde n'est plus le pétrole, mais les données »[1]. Cette affirmation marquait un tournant dans notre compréhension et notre valorisation de ces ressources dans l'économie. Toutefois, si la métaphore des données en tant que « nouveau pétrole » ou « monnaie » s'efforce de saisir leur importance croissante, ces comparaisons ne rendent pas pleinement compte de la faculté des données à être partagées et réutilisées, engendrant de la valeur sans être consommées de manière exclusive.

Cette singularité des données, capables d'être agrégées, réutilisées et partagées sans s'épuiser, remet en question notre compréhension traditionnelle de la propriété et de la valeur, structurée autour de ressources physiques finies.

En effet, une caractéristique distinctive des données est leur non-épuisabilité. D'un point de vue économique, les données sont un bien non rival, ce qui signifie que la consommation de données par un individu n'empêche pas d'autres individus de les utiliser également. Cependant, les données sont excluables, ce qui signifie qu'un agent économique peut empêcher sa consommation par un autre simplement en le gardant caché (Buchanan, 1965). Toutefois, cela ne s'applique qu'aux données privées. Une fois qu'une donnée devient publiquement disponible, elle devient un bien public. Les deux caractéristiques, bien non rival, et bien excluable, font de la donnée un bien de club.

Une autre caractéristique consiste en la capacité des données à être combinées de différentes façons afin de générer de nouvelles sources de valeur. Une analogie intéressante peut être de comparer les données aux éléments du tableau périodique[2]. Comme les éléments d'un tableau périodique, les données peuvent avoir une valeur indépendamment mais interagissent entre elles pour créer de nouvelles combinaisons. De plus, la valeur des données dépend souvent du contexte et de l'application, tout comme les éléments du tableau périodique ont des valeurs différentes selon les domaines d'activité. En combinant différents éléments de données, on peut créer des compositions plus utiles et de valeur pour un plus grand nombre de personnes et d'organisations.

[1] « The world's most valuable resource is no longer oil, but data ». Source : https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data.

[2] Suggestion faite dans le cadre de l'étude « Delivering Value Through Data » qui a rassemblé 900 experts à travers le monde pour débattre de la valeur des données. Cette notion a acquis un large consensus.

1. La valeur des données au sein des organisations

Principaux résultats

Acquisition et prétraitement des données :

- Une grande majorité (95 %) des entreprises belges s'engagent activement dans le stockage de données.
- Ces entreprises ont recours surtout au stockage de données provenant de sources internes : les processus commerciaux, les informations liées aux clients et les systèmes d'information internes.
- Alors que l'utilisation de sources internes est conforme à la tendance observée au niveau de l'UE, les entreprises belges manifestent une propension nettement plus marquée à exploiter des données externes, en particulier les données publiques et celles provenant de plateformes numériques, par rapport à la moyenne européenne.
- Les secteurs de l'informatique, de la finance, de l'immobilier, et autres domaines à haute intensité de données, privilégient davantage l'utilisation de sources variées (internes et externes).

Utilisation du Cloud pour le stockage de données

- Une grande majorité (86 %) des entreprises belges qui ont recours au stockage de données utilisent le Cloud.
- Bien que le taux d'utilisation du Cloud en Belgique soit proche de la moyenne de l'UE, les entreprises belges utilisent le Cloud de manière plus intensive, avec 70 % d'entre elles y stockant une quantité modérée de données, contre 60 % pour la moyenne de l'UE-27.

Analyse des données

- La majorité des entreprises belges (83 %) qui stockent des données procèdent également à leur analyse.
- La Belgique pratique l'analyse de données de manière plus intensive que la moyenne européenne (UE-27) et que ses voisins, à l'exception des Pays-Bas.
- Le pays compte également une proportion élevée, tout comme les Pays-Bas, d'entreprises pour lesquelles l'analyse de données représente un élément central de leur activité.
- Les entreprises de taille moyenne et grande analysent leurs données plus fréquemment que les petites entreprises.

Profils d'utilisation des données

- 33 % des entreprises qui stockent des données ont un profil **d'utilisateur régulier et optimisateur**.
- 20 % des entreprises qui stockent des données ont un profil d'utilisateur avancé et monétiseur de données.
- Ces deux profils se distinguent par une fréquence élevée d'analyse de données et une grande importance accordée aux données internes. Toutefois, les utilisateurs avancés sont plus enclins à stocker des données externes et à adopter les technologies du Cloud, de l'IoT ainsi que des systèmes d'information et de gestion.

- La Belgique compte proportionnellement plus d'utilisateurs avancés et réguliers que l'Union européenne et que ses pays voisins, à l'exception des Pays-Bas.

Raisons majeures pour le stockage de données

- La majorité des entreprises évoque comme raison principale le stockage des données pour la prise de décision et la planification stratégique (69 %), le contrôle des activités (67 %), et l'amélioration de l'efficacité commerciale (55 %).
- L'amélioration de l'offre aux clients par le développement de nouveaux produits ou services (45 %) et l'amélioration de l'expérience client (51 %) sont également considérées comme deux raisons importantes de stocker des données.
- Comparativement à l'UE, les entreprises belges semblent accorder plus d'importance à la planification stratégique et au contrôle de l'empreinte environnementale de l'entreprise comme raison majeure de stocker des données (39 %).
- Seules 14 % considèrent la monétisation des données (vente de données ou des informations basées sur les données) comme une raison majeure pour stocker les données.

Adoption de l'intelligence artificielle (IA)

- En 2023, 13,8 % des entreprises ont adopté au moins une technologie d'intelligence artificielle, contre 10,3 % en 2021.
- La Belgique se distingue par une adoption élevée de l'IA, se classant en quatrième position dans l'UE.
- L'augmentation du taux d'utilisation de l'IA dans les entreprises belges est généralisée, mais l'augmentation est plus importante pour certaines fonctions liées au traitement automatisé du langage et l'optimisation des processus de travail, ce qui suggère une adoption de l'IA générative.
- Les grandes entreprises utilisent l'IA de manière significativement plus élevée par rapport aux PME.
- Le manque d'expertise et les coûts représentent les principaux obstacles à l'utilisation de l'IA. Toutefois, on constate en 2023 une augmentation des préoccupations liées à la protection des données, aux implications juridiques et aux considérations éthiques.

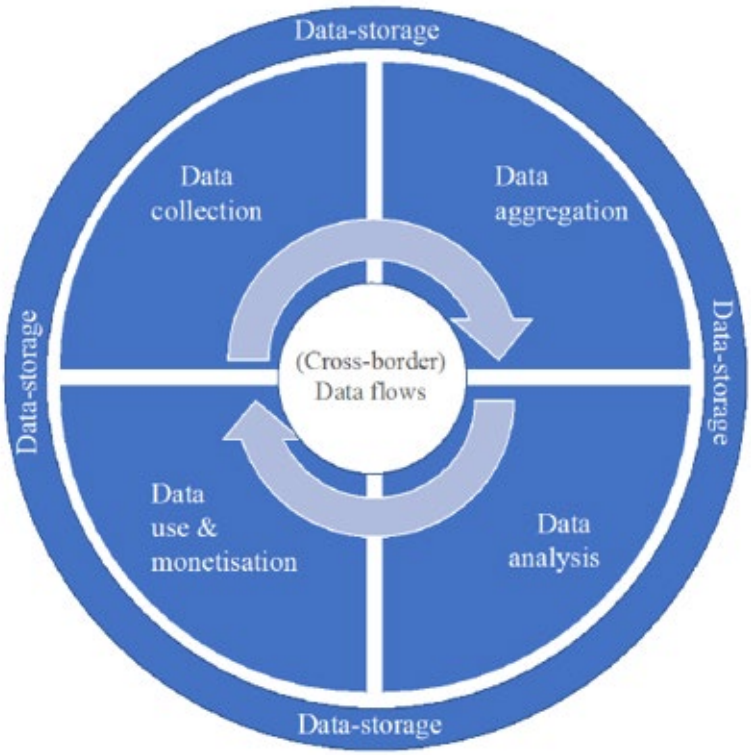
1.1. La chaîne de valeur des données

Les évolutions technologiques récentes dans la collecte, le stockage, et le traitement des données ont ouvert de nouvelles avenues pour l'innovation et la croissance des organisations. Pour explorer pleinement ces opportunités, il est essentiel d'adopter un cadre pour comprendre la chaîne de valeur des données dans les organisations.

Une étude de l'OCDE (Nguyen, D. and Paczos M. (2020))[3] propose un modèle de la chaîne de valeur des données qui se décline en quatre parties : la collecte, l'agrégation, l'analyse et l'utilisation et la monétisation des données. Sous-jacentes à ces quatre étapes se trouvent le stockage des données et les flux de données qui nécessitent également une attention particulière, ainsi que les avancées récentes dans les technologies numériques qui facilitent et amplifient les capacités à chaque niveau de cette chaîne.

[3] Nguyen, D. and Paczos M. (2020), « Measuring the economic value of data and cross-border data flows: A business perspective », OECD Digital Economy Papers, No. 297, OECD Publishing, Paris, https://dx.doi.org/10.1787/6345995e-en.

Figure 1. Le cycle de valeur de la donnée



Source : Nguyen, D. and M. Paczos (2020).

Collecte

La première étape pour utiliser et tirer de la valeur des données est l'étape de la collecte des données. Le développement d'internet, conjointement aux avancées dans les technologies numériques telles que les plateformes en ligne et l'Internet des objets (IoT), a ouvert la porte à une accessibilité sans précédent à une immense quantité de données. Ces innovations technologiques ont facilité l'émergence des « flux de données », désignant la méthode par laquelle les données sont recueillies, stockées, et potentiellement échangées de façon intentionnelle.

En diminuant considérablement les coûts de transaction, les données peuvent désormais traverser de nombreuses frontières en un temps record sans que l'expéditeur ou le destinataire en ait connaissance. L'augmentation des flux transfrontaliers de données conduirait à une augmentation de l'efficacité économique et de la productivité, entraînant elle-même une augmentation du bien-être et des niveaux de vie (OCDE, 2012)[4].

Au cœur de la collecte de données se trouvent les questions relatives aux sources de données, aux types de données et aux techniques spécifiques pour acquérir ces données.

Encadré 1. Sources et types de données

Les sources de données et les types de données vont souvent de pair. En effet, de nombreuses taxonomies[5] utilisent la source de données comme critère pour classer ces informations, distinguant par exemple les données privées des données publiques, les données internes des données externes ou même les données créées par l'utilisateur des données créées par la machine.

[4] OECD Internet Economy Outlook 2012 | en | OECD. (n.d.). https://www.oecd.org/sti/ieconomy/oecd-internet-economy-outlook-2012-9789264086463-en.htm.

[5] (par exemple, OCDE 2013 et 2015, George et al. 2014).

Les types spécifiques de données sont également mis en avant dans la littérature, et une classification importante distingue les données structurées des données non structurées. Baars et Kemper (2008)[6] ont défini les données structurées comme des données attribuées à des champs dédiés et pouvant ainsi être directement traitées avec des équipements informatiques, tandis que de nombreuses sources d'information sont non structurées ou au mieux semi-structurées, par exemple les e-mails clients, les pages web avec des informations concurrentielles, les rapports de force de vente, les référentiels de documents de recherche, et ainsi de suite. Cela signifie que ces données ne suivent pas un modèle ou un format spécifique, ce qui rend difficiles leur stockage et leur traitement dans des bases de données relationnelles traditionnelles et complique la réutilisation des données par d'autres parties prenantes.

Une autre classification est proposée dans le contexte du RGPD. De nombreuses entreprises ont utilisé la norme ISO/IEC 19944 sur « Flux de données, catégories de données et utilisation des données » qui distingue quatre types de données, à savoir les données de contenu client (incluant les informations biométriques et de contact), les données dérivées (identifiant les utilisateurs finaux ou les organisations), les données fournies par les prestataires de services Cloud, et les données de compte (données de paiement ou administratives).

Aujourd'hui, les entreprises peuvent collecter une large variété de données à partir de nombreuses sources. Cependant, les données sont plus à même de générer de la valeur si elles possèdent certaines caractéristiques (Nguyen, D. and Paczos, M. 2020) :

- Liables – Les données peuvent être croisées avec d'autres ensembles de données ;
- Accessibles – facilement récupérables et/ou intégrées dans les processus métier ;
- Désagrégées – au niveau souhaité ;
- Périodiques – mises à jour avec une fréquence suffisante pour répondre aux exigences de l'utilisation qui en est faite, par exemple annuellement, quotidiennement ou en temps réel ;
- Dignes de confiance - jugées crédibles par ceux qui l'utilisent ; les données sont impartiales et neutres, et ne dépendent pas du jugement, de l'interprétation ou de l'évaluation des individus ;
- Représentatives – les enregistrements ne contiennent pas de champs manquants, les données sont suffisamment représentatives pour répondre aux exigences ;
- Rares – propriétaires ou secrètes, difficiles à obtenir.

Agrégation

La valeur des données se révèle pleinement lorsque diverses sources et ensembles de données sont intégrés, ce qui souligne l'importance des caractéristiques des données précitées. En effet, partir de données déjà agrégées et/ou peu liables rend difficile la recombinaison de données, limitant ainsi les nouvelles opportunités d'analyse et de création de valeur.

L'agrégation et la liaison de donnée, du point de vue technique, amènent des défis significatifs liés au rassemblement de données hétérogènes, notamment le nettoyage des données, l'identification des clés uniques, la détermination du degré d'agrégation approprié, l'application de méthodes d'anonymisation et, de façon plus générale, la gestion efficace (ou curation) des données.

Le processus de nettoyage des données, bien qu'essentiel à leur exploitation, est souvent perçu comme laborieux. La majorité des données, dans leur état initial, ne sont pas prêtes à être analysées et nécessitent une préparation minutieuse pour en éliminer les imperfections et assurer leur cohérence. Des recherches telles que celle menée par Chen et al. (2014)[7] ont mis en lumière le coût élevé du nettoyage des données, estimant une dépense annuelle de plus de 3 trillions de dollars pour l'économie américaine.

---

[6] Baars, H., & Kemper, H.-G. (2008). Management Support with Structured and Unstructured Data—An Integrated Business Intelligence Framework. Information Systems Management, 25(2), 132–148. https://doi.org/10.1080/10580530801941058.

[7] Chen, L., Jia, Y., Sellis, T., & Liu, G. (eds.). (2014). Web Technologies and Applications: 16th Asia-Pacific Web Conference, APWeb 2014, Changsha, China,September 5-7, 2014. Proceedings (Vol. 8709). Springer International Publishing. https://doi.org/10.1007/978-3-319-11116-2.

De plus, les données doivent souvent être anonymisées pour garantir leur échange conformément aux dispositions du RGPD. Cela soulève d'importantes questions techniques, notamment car l'anonymisation ou la pseudonymisation sont soupçonnées d'être sujettes à une rétro-ingénierie. À titre d'exemple, Rocher et al. (2019)[8] ont développé une méthode capable de réidentifier avec une grande précision des personnes à partir d'ensembles de données ne comportant que 15 attributs démographiques.

Enfin, il convient également de mentionner le concept de curation des données à ce stade. Dans leur livre, Cavanillas et al. (2016)[9] définissent la curation des données comme la gestion active des données tout au long de leur cycle de vie pour garantir qu'elles répondent aux exigences de qualité nécessaires à leur utilisation efficace. Le sujet de la curation des données est donc un sujet plus vaste et plus transversal, recouvrant certains des aspects décrits dans la présente revue de la littérature. Toujours selon Cavanillas et al. (2016), la curation des données est effectuée par des conservateurs de données dont le rôle est de garantir que les données soient fiables, découvrables, accessibles, réutilisables et adaptées à leur usage.

Encadré 2. Principes FAIR

En 2016, les « Principes directeurs FAIR pour la gestion et la gouvernance des données scientifiques » ont été publiés dans la revue *Scientific Data*. Ces principes visent à améliorer la Découvrabilité (Findability), l'Accessibilité (Accessibility), l'Interopérabilité (Interoperability) et la Réutilisation(Reusability) des actifs numériques, en insistant particulièrement sur l'automatisation des processus permettant aux systèmes informatiques de trouver, accéder, interagir et réutiliser les données avec peu ou aucune intervention humaine. Cette approche répond à la nécessité croissante de soutien informatique face à l'augmentation du volume, de la complexité et de la vitesse de génération des données.

Les **principes FAIR** s'articulent autour de quatre axes principaux :

Découvrabilité : les données et les métadonnées doivent être aisément identifiables à la fois par les humains et par les machines. Cela implique d'assigner aux données des identifiants uniques et persistants, de les décrire avec des métadonnées détaillées, et de les référencer dans des bases de données consultables.

Accessibilité : après localisation des données nécessaires, il doit être clair comment celles-ci peuvent être consultées, y compris les modalités d'authentification et d'autorisation requises. Les métadonnées doivent demeurer accessibles même en cas d'indisponibilité des données elles-mêmes.

Interopérabilité : les données doivent pouvoir être intégrées avec d'autres ensembles de données et être compatibles avec des applications ou des processus d'analyse, de stockage et de traitement. Ceci requiert l'utilisation de langages formels, accessibles, partagés et largement applicables pour la représentation des connaissances, ainsi que l'adoption de vocabulaires respectant les principes FAIR.

Réutilisabilité : l'objectif ultime des principes FAIR est de maximiser la réutilisation des données. Pour cela, les données et les métadonnées doivent être précisément décrites pour permettre leur réplication et/ou leur combinaison dans différents contextes. Cela inclut la mise à disposition des données sous une licence d'utilisation transparente et accessible, l'association des données à une provenance détaillée et leur conformité aux standards de la communauté concernée.

Les principes FAIR s'adressent à trois catégories d'entités : les données (ou tout objet numérique), les métadonnées (les informations décrivant ces objets numériques) et l'infrastructure (par exemple, les systèmes où les données et les métadonnées sont enregistrées ou indexées pour être facilement trouvables).

[8] Rocher, L., Hendrickx, J.M. & de Montjoye, YA. Estimating the success of re-identifications in incomplete datasets using generative models. *Nat Commun* **10**, 3069 (2019). https://doi.org/10.1038/s41467-019-10933-3.

[9] Cavanillas, J., Curry, E., Wahlster, W. (eds) New Horizons for a Data-Driven Economy. Springer, Cham. https://doi.org/10.1007/978-3-319-21569-3_16.

Pour ceux qui cherchent une guidance pratique sur l'application de ces principes, le Cadre de FAIRification en trois points offre une méthodologie détaillée. Pour plus d'informations et pour télécharger un résumé des Principes FAIR, on peut visiter la page GitHub dédiée aux FAIR-nanopubs.

Source : https://www.go-fair.org/fair-principles/.

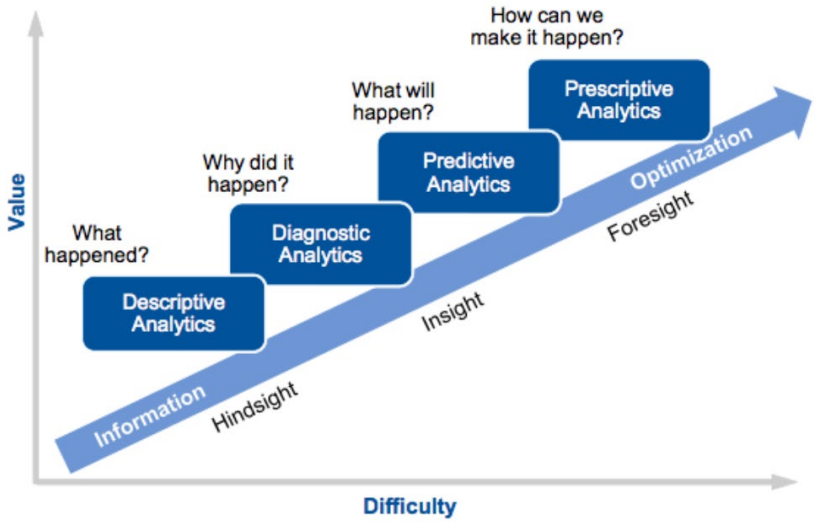
Analyse

L'analyse de données est la troisième étape dans la chaîne de valeur des données. Elle consiste à examiner minutieusement les données pour identifier les causes, les facteurs clés et les résultats possibles.

Gartner[10] a mis au point un modèle d'ascendance analytique basé sur quatre types d'analyses de données : descriptive, diagnostique, prédictive et prescriptive. Chaque type répond à une question spécifique, allant de la simple description d'un phénomène à la recommandation d'actions futures. Ces quatre types d'analyses de données sont décrits comme suit :

- Analyse descriptive : l'objectif est essentiellement de décrire un phénomène grâce à différentes mesures. La question clé à laquelle nous cherchons à répondre est : qu'est-il arrivé ?
- Analyse diagnostique : le but ici est de trouver les causes profondes d'un phénomène ou d'un problème. La question considérée est donc : pourquoi cela s'est-il produit ?
- Analyse prédictive : comme son nom l'indique, nous passons du passé au futur et tentons de prédire les résultats futurs potentiels. La question est : que va-t-il se passer ?
- Analyse prescriptive : le niveau le plus élevé, allant au-delà de la simple description, de l'explication et de la prédiction, l'analyse prescriptive cherche à conseiller sur les options à considérer et les décisions à prendre. La question finale est donc : que devrait-on faire ?

Figure 2. Le modèle d'ascendance analytique de Gartner



Source : https://computd.nl/4-levels-of-data-maturity/.

D'après une enquête réalisée par Gartner en 2018, soit six ans après la création de leur modèle, il s'avère que plus de 87 % des organisations sont encore à un niveau bas de maturité dans ce domaine. Cette constatation amène certains à remettre en question la pertinence du modèle de Gartner, jugé trop linéaire et désormais dépassé face aux défis actuels. Les critiques pointent du doigt la difficulté à générer de la valeur à partir du big data et de l'analyse prédictive mais aussi la sous-estimation de la vitesse à laquelle les entreprises doivent adapter leurs stratégies de données face à un environnement technologique en constante évolution et le besoin d'agilité dans la gestion de projet[11].

[10] https://www.gartner.com/en/information-technology/glossary/analytics.

[11] https://www.linkedin.com/pulse/why-only-13-achieved-gartner-analytics-maturity-model-denis-sproten/.

Il est important de souligner également le rôle crucial du capital humain et de l'expertise à côté des actifs technologiques dans les environnements axés sur les données. Les différents types d'analyses peuvent nécessiter différents niveaux de compétences et/ou de technologies. Par exemple, alors que l'analyse descriptive peut simplement être effectuée grâce à des tableurs de base, les analyses prédictives ou prescriptives nécessitent des techniques plus avancées telles que les algorithmes d'optimisation ou de simulation et des personnes capables de les construire et de les comprendre pour fournir des informations utiles.

Utilisation des données et monétisation

Afin d'élaborer une taxonomie permettant de catégoriser les données et de conceptualiser et mesurer leur valeur économique, Nguyen, D. et Paczos, M. (2020) propose de séparer les concepts d'entreprises habilitées par les données (*data-enabled businesses*) et d'entreprises améliorées par les données (*data-enhanced businesses*).

Cette distinction se concentre sur la fonction principale des données au sein de chaque modèle commercial : les entreprises habilitées par les données sont des sociétés qui ont élaboré des stratégies de génération de revenus entièrement dépendantes des données et qui n'existeraient pas sans l'accès à de grandes quantités de données et à des analyses de données avancées. En revanche, les entreprises améliorées par les données exploitent les données pour mieux coordonner les opérations commerciales préexistantes, faciliter la prise de décision et introduire de nouveaux biens et services ; les données ne modifient ni ne déterminent cependant leurs modèles commerciaux de base.

Ainsi, les auteurs identifient quatre catégories de modèles commerciaux liés aux données :

Catégorie 1. Vente ou concession de données brutes ou agrégées ;

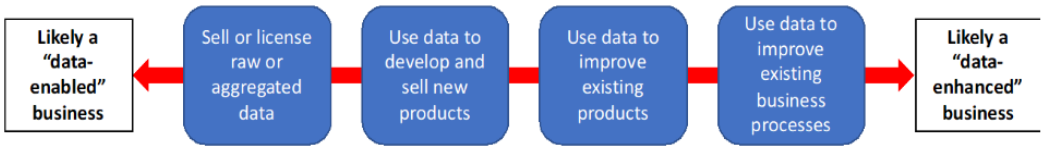
Catégorie 2. Développement et vente de nouveaux produits liés aux données ;

Catégorie 3. Utilisation des données pour améliorer les produits existants ;

Catégorie 4. Utilisation des données pour améliorer les processus de production ou l'efficacité commerciale.

Les entreprises habilitées par les données généreront des revenus liés aux catégories 1 et 2, tandis que les entreprises améliorées par les données concernent les catégories 3 et 4. Il convient toutefois de noter qu'une entreprise n'est pas l'un ou l'autre, mais que les entreprises peuvent adopter plusieurs types de modèles commerciaux autour des données.

Figure 3. « Entreprises habilitées par les données » vs « entreprises améliorées par les données »



Source : Nguyen, D. et Paczos, M.(2020).

La gouvernance des données

Nombre d'organisations aujourd'hui se heurtent à des défis significatifs tels que l'absence de politiques intégrales autour de la gestion de la qualité des données, le manque de structures de gouvernance transversales, et une définition floue des rôles liés aux données, ainsi que des processus d'affaires insuffisamment formalisés. Ces obstacles entravent non seulement la capacité d'une organisation à exploiter pleinement ses données mais augmentent également les risques liés à leur gestion.

Ainsi, reconnaître la data comme un actif unique et stratégique implique d'aller au-delà de la simple collecte et stockage. Il est impératif de lui accorder une attention particulière, à l'instar des autres actifs de l'organisation, en élaborant une stratégie propre qui encadre sa collecte, son stockage, son

agrégation et son utilisation. Cette stratégie doit donner lieu à la mise en place d'une politique de gestion des données.

Encadré 3. La gouvernance des données

La gestion des données est définie par l'Association de gestion des données (*Data Management Association*) comme le développement, l'exécution et la supervision de plans, de politiques, de programmes et de pratiques qui contrôlent, protègent, fournissent et améliorent la valeur des données et des actifs informationnels.

La gouvernance des données est définie par la même association comme l'exercice de l'autorité, du contrôle et de la prise de décision partagée (planification, surveillance et mise en œuvre) sur la gestion des actifs de données.

En ce qui concerne les sujets et les réglementations qui peuvent influencer l'évolution de la gestion des données et de la gouvernance dans les entreprises, la confidentialité des données et la sécurité des données sont deux sujets importants. Ceux-ci sont des sous-composants d'un concept plus général lié aux données appelé protection des données.

Elle implique de devoir transformer la culture organisationnelle, ce qui est un défi important. Pour que les données constituent un avantage concurrentiel, les gestionnaires doivent développer de nouvelles compétences et de nouveaux styles de gestion.

L'adoption d'une gouvernance des données transversale, impliquant une coordination étroite entre les différents départements, est cruciale. Elle permet d'établir des standards et des règles claires, de définir précisément les rôles et responsabilités autour des données, et de formaliser les processus d'affaires qui les englobent. Cela conduit à une meilleure qualité des données, à une prise de décision éclairée et à une réduction des risques liés à la gestion des données. La gestion des données devient alors un processus cross-fonctionnel, impliquant une collaboration étroite entre les départements informatiques, opérationnels et stratégiques.

Enfin, la réussite d'une stratégie de gouvernance des données repose sur l'engagement de la direction. Sans cet engagement, il est difficile d'aligner les initiatives de gouvernance des données avec les objectifs stratégiques de l'organisation. La direction doit reconnaître les données comme un atout stratégique et soutenir activement leur gouvernance, en allouant les ressources nécessaires et en promouvant une culture organisationnelle qui valorise et protège les données. Cet engagement en haut lieu est le garant de l'efficacité de la gouvernance des données, facilitant une transformation organisationnelle vers une entité réellement pilotée par les données.

Encadré 4. La Pyramide DMBOK

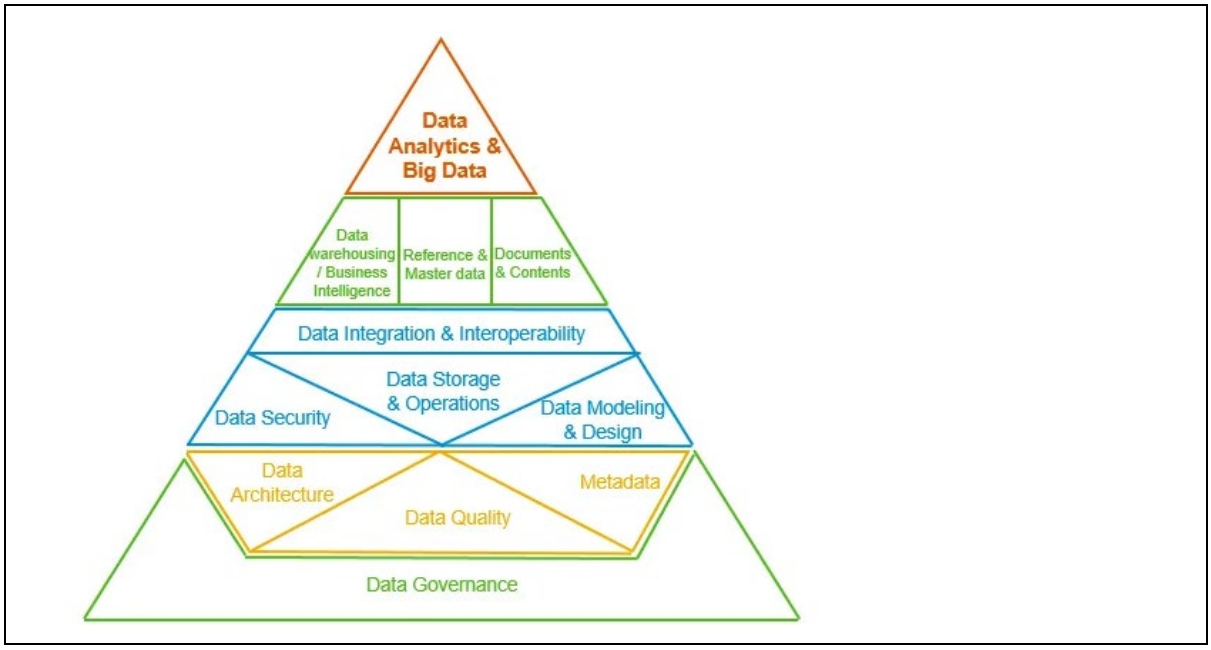
La Pyramide DMBOK (Data Management Body of Knowledge) illustre la relation hiérarchique entre les domaines fonctionnels de la gestion des données. Elle montre comment les différents composants de la gestion des données s'appuient les uns sur les autres pour former une approche globale et cohérente. La pyramide se compose de trois niveaux : stratégique, tactique et opérationnel.

Au niveau **stratégique**, on pose les fondations avec la gouvernance des données, qui définit la direction et les politiques de gestion des données, et l'architecture des données, qui conçoit la structure globale des actifs de données en alignement avec les objectifs de l'organisation. Ce niveau sert de base pour établir un cadre responsable et décisionnel autour des données.

Le **niveau tactique** se base sur le stratégique pour opérationnaliser les stratégies et plans de gestion des données. Il inclut le développement des données, qui couvre la création et la modification des actifs de données, les opérations des données, gérant la maintenance quotidienne, la sécurité des données, garantissant la protection des données, et l'intégration et interopérabilité des données, assurant l'échange fluide des données entre systèmes.

Enfin, le niveau **opérationnel** met en œuvre les activités de gestion des données au quotidien. Il englobe les entrepôts de données (*data warehouses*), la *Business Intelligence*, la gestion des documents et de contenu pour les données non structurées, la gestion des données de référence et maîtres (en vue d'une standardisation des données), la qualité des données et les métadonnées.

Figure 4. La pyramide de DMBOK



1.2. L'utilisation faite des données au sein des entreprises belges

Dans ce chapitre, nous observons comment les entreprises en Belgique se positionnent et se distinguent de leurs homologues européens vis-à-vis de leur utilisation des données. Pour ce faire, nous mobilisons principalement les données fournies par l'étude « Survey of Businesses on the Data Economy »[12] qui se base sur un échantillon de 421 entreprises belges[13] et 10.000 entreprises au niveau de l'Union européenne.

Suivant le cadre proposé par l'étude, nous déclinerons nos observations en fonction des trois étapes de la chaîne de valeur exposées dans l'étude, à savoir :

- l'acquisition et le prétraitement des données
- l'analyse de données
- l'utilisation et la validation de données

Acquisition et prétraitement des données

En ce qui concerne l'acquisition et le prétraitement des données, une grande majorité des entreprises belges (95 %) déclare en 2022 s'engager activement dans le stockage de données issues d'au moins une source.

Les entreprises en Belgique semblent avoir surtout recours au stockage de données provenant de sources internes, une tendance qui se reflète également à l'échelle européenne. Les processus commerciaux internes représentent la source la plus courante, avec 70 % des entreprises belges utilisant ces données, contre 67 % dans l'UE. Les informations client et les systèmes d'information internes suivent de près, avec respectivement 63 % et 55 %, ce qui montre une tendance similaire à celle observée dans les pays voisins européens, à l'exception notable des Pays-Bas où l'usage des sources internes est plus prononcé.

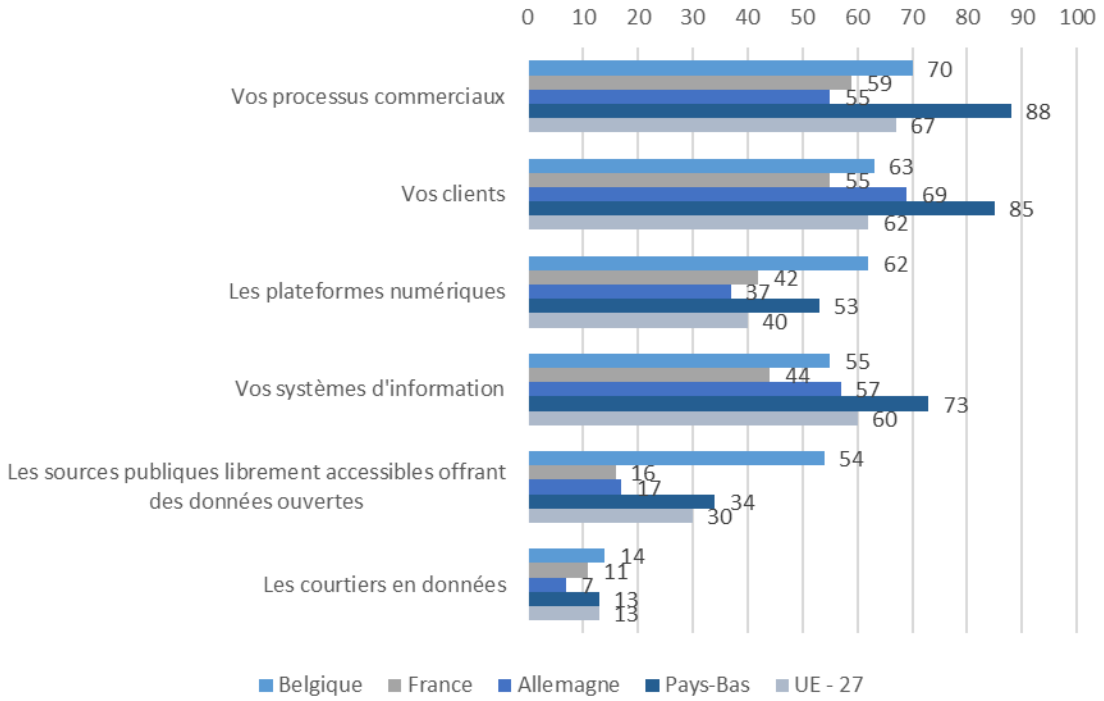
[12] European Commission DG Communications Networks, Content & Technology, Survey of Businesses on the Data Economy: A study prepared by Ipsos Belgium and iCite, 28 July 2022.

[13] Les entreprises sélectionnées appartiennent aux codes NACE suivants : NACE A, B, C, D, E, F, G, H, I, J, K, L, M, N, P, Q, R et S.

Concernant les données externes, la Belgique se distingue par une approche particulièrement ouverte, et une utilisation plus large que la moyenne européenne et les pays voisins. Un peu plus de la moitié des entreprises belges, 54 %, tirent parti des données publiques, un taux bien supérieur aux 30 % observés dans l'UE. Les plateformes numériques sont également une source importante, avec 62 % des entreprises belges les utilisant pour la collecte de données, contre 40 % à l'échelle de l'UE. En revanche, l'acquisition de données via des courtiers reste une pratique moins courante, concernant seulement 14 % des entreprises, soit un taux légèrement supérieur à la moyenne européenne de 13 %.

Figure 5. Sources de données utilisées

Parmi les entreprises qui stockent les données.



Source : Élaboration d'Idea Consult sur base des données du « Survey of Businesses on the Data Economy ».

L'utilisation du Cloud pour le stockage de données par les entreprises révèle des tendances intéressantes tant à l'échelle globale qu'au niveau sectoriel. En effet, une vaste majorité des entreprises qui ont recours au stockage de données (86 %), déclare avoir recours au Cloud pour conserver une petite quantité de données sur le Cloud, un taux très proche de la moyenne européenne fixée à 85 %. Toutefois, à y regarder de plus près, les entreprises en Belgique semblent utiliser plus intensivement du Cloud puisque pas moins de 70 % déclarent stocker une quantité modérée de leurs données sur le Cloud, contre 60 % en moyenne pour l'UE-27.

Lorsqu'on explore les habitudes de stockage des données selon les secteurs d'activité en Europe, les disparités sont marquantes. Bien que l'emploi du Cloud soit quasi-universel parmi les entreprises européennes, quel que soit leur domaine, l'origine des données stockées varie grandement en fonction du secteur spécifique. Les secteurs tels que l'informatique, la finance, l'immobilier, et d'autres domaines techniques et scientifiques, caractérisés par une intensité élevée de données, présentent une tendance à stocker des informations issues de diverses sources, internes et externes, à un niveau supérieur à la moyenne.

L'industrie manufacturière en Europe présente une fréquence plus élevée de stockage de données provenant de sources internes (c'est-à-dire des processus commerciaux) par rapport à tous les autres secteurs (72 % contre une moyenne de 67 %). Le secteur de l'éducation, de la santé humaine et des services sociaux est plus enclin à stocker des données provenant de sources externes, principalement issues de plateformes numériques, que tout autre secteur (48 % contre une moyenne de 40 %). En

revanche, le secteur de l'agriculture, de la sylviculture et/ou de la pêche (55 %) ainsi que le secteur de la construction, de la gestion des déchets, de l'approvisionnement en eau et d'électricité (58 %) sont moins susceptibles de stocker des données provenant de leurs clients par rapport à la moyenne (62 %) et par rapport aux entreprises de tous les autres secteurs.

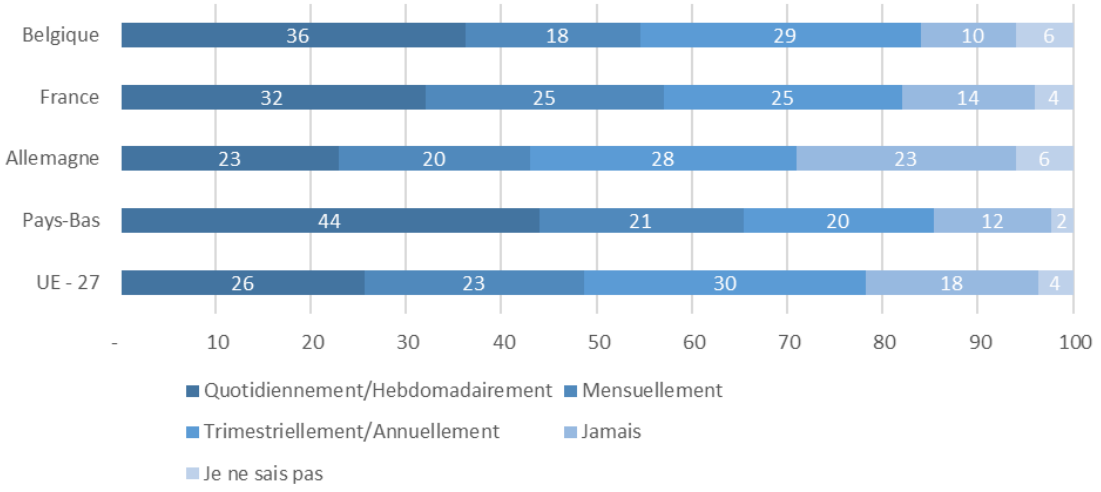
La place de l'analyse des données

En ce qui concerne l'analyse des données, il est à noter que la grande majorité des entreprises qui stocke des données en Belgique procède également à leur analyse (83 % contre 79 % dans l'UE). En ce qui concerne la fréquence à laquelle ces entreprises analysent leurs données (Figure 6), on peut les répartir en différentes catégories : près d'un dixième d'entre elles ne le font jamais, près d'un tiers (29 %) le font annuellement ou trimestriellement, environ un cinquième (18 %) le font mensuellement, et un peu plus d'un tiers (36 %) le font de manière hebdomadaire ou quotidienne.

De manière générale, comme la figure 6 le montre, la Belgique pratique l'analyse de données de manière plus intensive que la moyenne européenne (UE-27) et que ses voisins (à l'exception des Pays-Bas), tandis que la part d'entreprises ne réalisant pas d'analyse de données est moins nombreuse.

Figure 6. Fréquence à laquelle les entreprises analysent leurs données

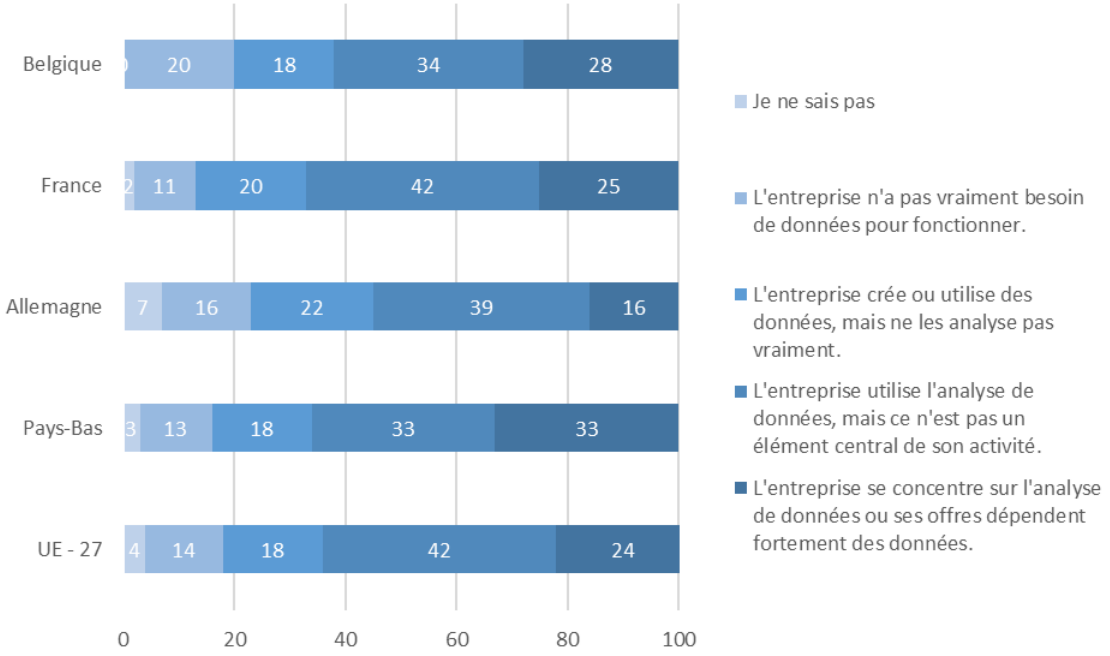
Parmi les entreprises qui stockent les données.



Source : Élaboration d'Idea Consult sur base des données du « Survey of Businesses on the Data Economy ».

Lorsque l'on examine l'importance accordée à l'analyse de données dans le modèle économique de l'organisation, la proportion d'entreprises déclarant ne pas avoir besoin de données pour fonctionner est plus élevée en Belgique que dans les pays voisins et au sein de l'ensemble de l'UE. En revanche, le pays compte également une proportion élevée, tout comme les Pays-Bas, d'entreprises pour lesquelles l'analyse de données représente un élément central de leur activité (28 % contre 33 % aux Pays-Bas et 24 % dans l'UE).

Figure 7. Importance de l'analyse de données dans le modèle d'affaire des entreprises



Source : Élaboration d'Idea Consult sur base des données du « Survey of Businesses on the Data Economy ».

Cependant, ces résultats varient en fonction de la taille de l'entreprise, avec un quart des micro-entreprises déclarant ne jamais analyser les données (25 % contre une moyenne de 18 % dans l'UE). De plus, les entreprises de taille moyenne (34 %) et grandes (40 %) qui stockent des données ont tendance à analyser les données fréquemment (c'est-à-dire quotidiennement ou hebdomadairement).

Le type d'utilisateur de données

En ce qui concerne les profils d'utilisation des données, ils permettent de classer les entreprises en quatre groupes : les non-utilisateurs, les utilisateurs passifs, les utilisateurs réguliers et les optimiseurs internes, les utilisateurs avancés et les monétiseurs de données[14].

Les **non-utilisateurs** sont les entreprises qui déclarent ne pas stocker de données de quelque type que ce soit, ou celles qui, bien qu'elles stockent des données, déclarent ne jamais les analyser ou ne pas savoir si elles le font. Les **utilisateurs passifs** sont ceux qui stockent des données et les analysent, mais rarement (une fois par trimestre ou une fois par an).

Les **utilisateurs avancés** se distinguent non seulement par la fréquence à laquelle ils analysent ces données, mais aussi par l'importance vitale de celles-ci pour leur activité. Ces entreprises ont également plus de chances de stocker des données provenant de toutes les sources, à la fois internes et externes à l'entreprise. En ce qui concerne leur utilisation de la technologie, ces utilisateurs avancés montrent des taux d'adoption plus élevés des technologies de stockage et de traitement des données par rapport aux utilisateurs de données moins avancés.

En revanche, les **utilisateurs réguliers** et les optimiseurs internes sont caractérisés par l'utilisation intensive de données spécifiques à leur propre entreprise. Bien qu'ils stockent des données provenant de sources internes dans une mesure similaire aux utilisateurs avancés, la principale différence réside dans le fait que les utilisateurs réguliers et les optimiseurs internes sont moins enclins à stocker des données externes. Les technologies utilisées par ces entreprises reflètent également cette perspective d'optimisation interne, avec des taux d'adoption inférieurs aux utilisateurs avancés pour les technologies du Cloud, des systèmes d'information de gestion et de l'Internet des objets, mais des

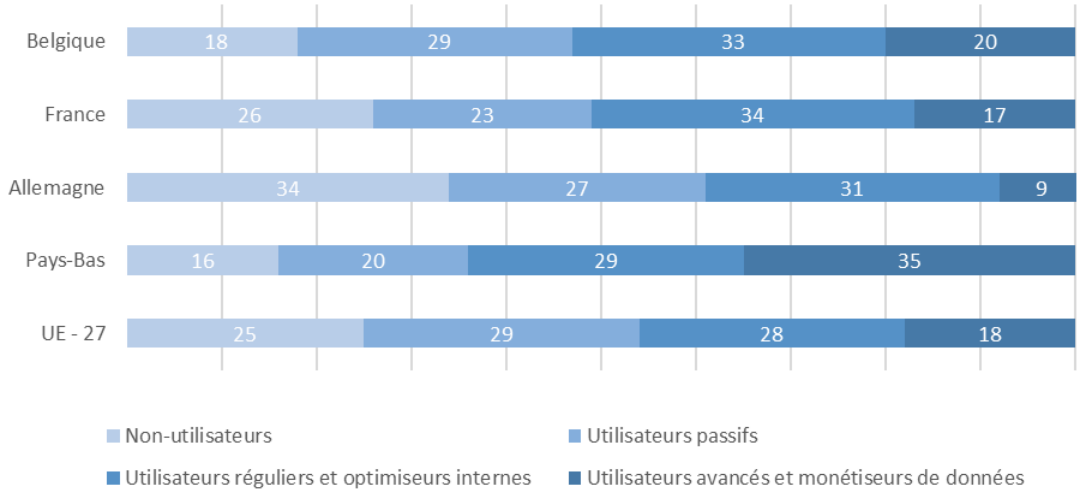
[14] Ces catégories ont été élaborées en fonction du stockage des données, de la fréquence d'analyse des données, des motivations pour les stocker et de l'importance des données pour les entreprises.

proportions plus proches des utilisateurs avancés par rapport à l'écart dans l'adoption de ces technologies entre les non-utilisateurs et les utilisateurs passifs.

En Belgique, on observe une répartition assez équilibrée entre ces groupes (Figure 8). Les non-utilisateurs représentent 18 % des entreprises, tandis que les utilisateurs passifs en constituent 29 %. Les utilisateurs réguliers et les optimiseurs internes représentent 33 %, et les utilisateurs avancés et les monétiseurs de données sont à 20 %.

En comparaison avec d'autres États membres de l'Union européenne, la Belgique se démarque par une proportion notablement élevée d'utilisateurs réguliers (33 %) et avancés (20 %), signe d'un recours significatif aux données pour l'optimisation interne ainsi qu'une adoption conséquente de technologies avancées de gestion de données. Parmi les pays comparés, seuls les Pays-Bas affichent de meilleurs résultats dans ces domaines.

Figure 8. Entreprises selon leur profil d'utilisation des données



Source : Élaboration d'Idea Consult sur base des données du « Survey of Businesses on the Data Economy ».

Lorsqu'on regarde les profils par secteur, la moyenne des entreprises de l'Union européenne (données pour la Belgique non disponibles) révèle que celles opérant dans les secteurs de la construction, de la gestion des déchets, de l'approvisionnement en eau et en électricité (NACE B, D, E et F) utilisent moins de données comparées à d'autres secteurs. Près des deux tiers des entreprises appartenant à ce regroupement sectoriel sont soit des non-utilisateurs (31 %), soit des utilisateurs passifs (32 %). En revanche, les secteurs de l'informatique, de la finance, de l'immobilier et d'autres secteurs techniques/scientifiques (NACE J, K, L, M, N et S) sont plus susceptibles d'être des utilisateurs avancés et des monétiseurs de données (22 %) par rapport aux autres.

Raisons pour stocker les données

Les entreprises de stockage de données ont été interrogées sur leurs motivations pour stocker des données, qui se répartissent en plusieurs catégories principales :

OPÉRATIONS :

- Améliorer les processus et l'efficacité commerciale
- Contrôle ou surveillance de vos activités
- Alimenter la prise de décision et la planification stratégique
- Contrôle ou surveillance de l'empreinte environnementale de votre entreprise

AMÉLIORATIONS DES OFFRES :

- Modifier les produits/services existants ou développer de nouveaux produits/services
- Améliorer l'expérience client

MONÉTISATION DES DONNÉES :

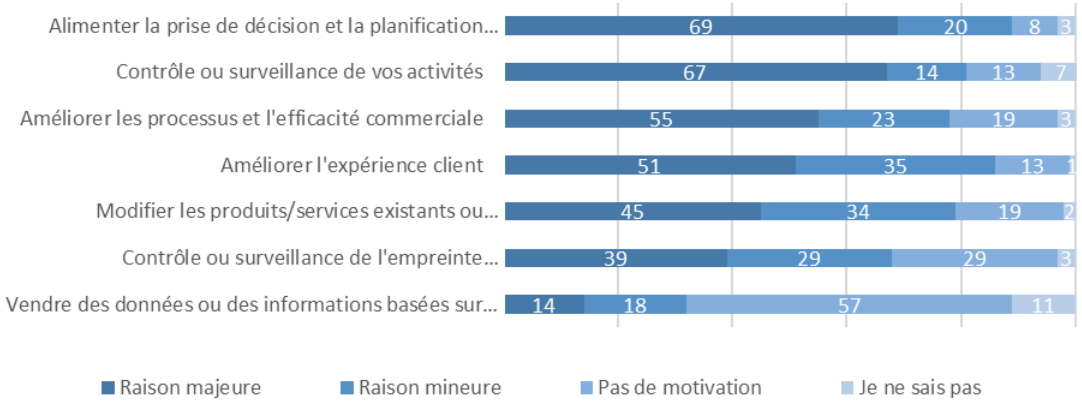
- Vendre des données ou des informations basées sur les données

L'enquête sur les motivations des entreprises belges pour le stockage de données révèle une prédominance des objectifs opérationnels comme facteurs déterminants. En effet, les résultats montrent que les principales motivations des entreprises pour stocker des données sont liées à l'amélioration des opérations internes des entreprises (en termes d'efficacité commerciale ou de prise de décision). En Belgique, la prise de décision et la planification stratégique (69 %), le contrôle ou la surveillance des activités des entreprises (67 %), et l'amélioration des processus et de l'efficacité commerciale (55 %) sont les trois principales raisons qui poussent les entreprises à stocker des données. En revanche, le stockage de données pour contrôler ou surveiller l'empreinte environnementale des entreprises semble être une motivation moins importante, dans la mesure où seules 39 % déclarent qu'il s'agit d'une raison majeure pour stocker les données.

L'amélioration de l'offre aux clients par le développement de nouveaux produits ou services (45 %) et l'amélioration de l'expérience client (51 %) sont également considérées comme deux raisons importantes de stocker des données.

En ce qui concerne la monétisation des données, pour la majorité des entreprises (68 %), la vente de données ou d'informations basées sur les données n'est pas du tout une motivation pour stocker des données. Ce schéma se maintient quelle que soit la taille de l'entreprise, car la monétisation n'est pas couramment choisie comme une cause mineure ou majeure de stockage de données pour les entreprises de toutes tailles.

Figure 9. Raisons évoquées pour stocker les données



Source : Survey of Businesses on the Data Economy.

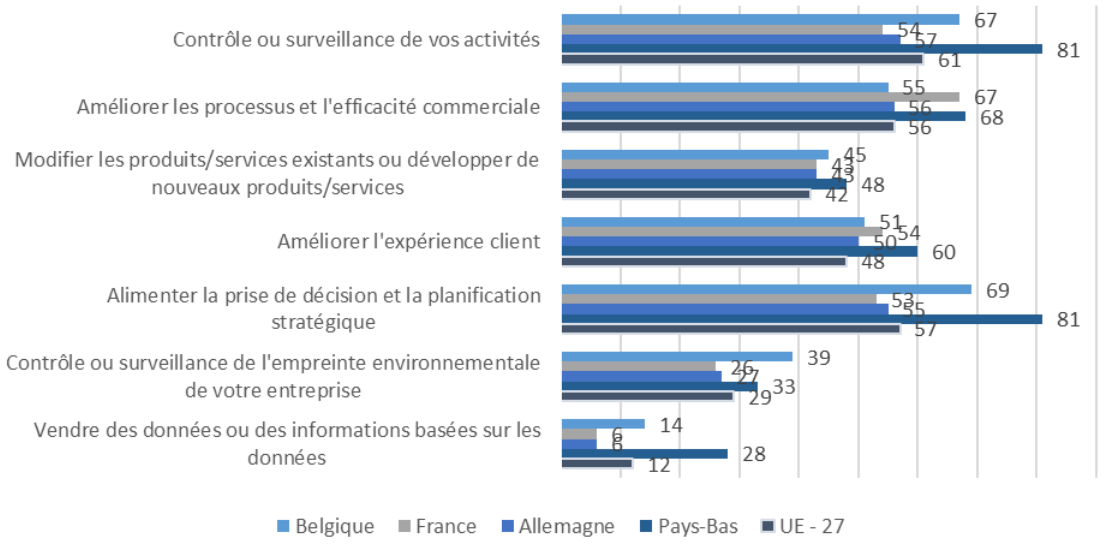
Cependant, en ce qui concerne les autres motivations de stockage de données, l'analyse en fonction de la taille révèle des différences statistiquement significatives entre les motivations majeures de stockage de données pour les grandes et moyennes entreprises par rapport aux micro et petites entreprises. Les grandes et moyennes entreprises sont plus susceptibles de stocker des données pour améliorer les opérations internes, y compris l'efficacité commerciale. Bien que moins courant, le stockage de données pour contrôler l'empreinte environnementale d'une entreprise semble également être une motivation plus importante pour les grandes et les moyennes entreprises par rapport aux micro et petites entreprises.

Par rapport aux pays voisins (Figure 10), la Belgique dénote par une plus forte prévalence de la motivation liée à la prise de décision et à la planification stratégique (69 % contre 57 % au niveau UE-27) ainsi qu'au contrôle ou à la surveillance des activités (67 % contre 61 % au niveau UE-27). De plus, la Belgique se distingue par une préférence plus marquée pour le contrôle de l'empreinte environnementale de l'entreprise par rapport à d'autres pays (39 % contre 29 % pour l'UE), bien que cette motivation reste relativement faible dans l'ensemble.

Parmi les pays voisins, les Pays-Bas se démarquent avec des pourcentages nettement plus élevés dans plusieurs catégories, notamment le contrôle ou la surveillance des activités (81 %), l'alimentation de

la prise de décision et de la planification stratégique (81 %) ainsi que la monétisation des données (28 %).

Figure 10. Raisons majeures évoquées pour stocker les données, par pays



Source : Survey of Businesses on the Data Economy.

Utilisation de l'intelligence artificielle

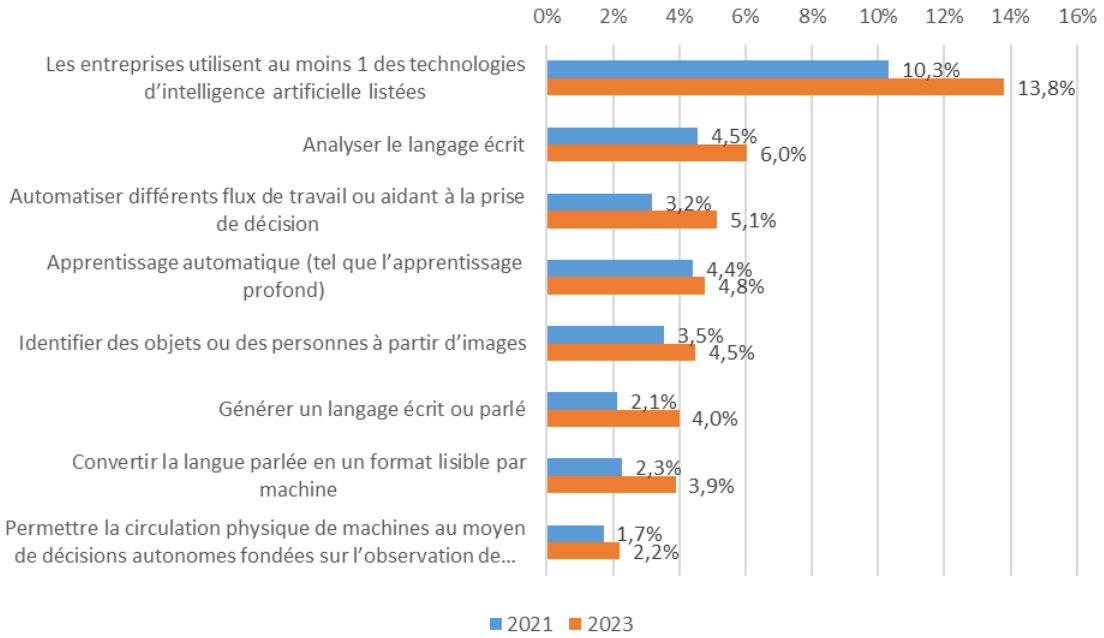
Entre 2021 et 2023, l'adoption de l'intelligence artificielle (IA) dans les entreprises a nettement progressé en Belgique, passant de 10,3 % à 13,8 % pour les entreprises de 10 employés ou plus. Cette hausse est particulièrement marquée dans le traitement automatisé du langage et l'optimisation des processus de travail.

Bien que l'apprentissage automatique (*machine learning)*, y compris l'apprentissage, ait aussi enregistré une augmentation, celle-ci reste inférieure aux autres catégories. Le succès relatif observé dans les autres catégories suggère que les technologies nouvelles et immédiatement applicables comme celles de l'IA générative (ex : ChatGPT) attirent l'attention des entreprises au-delà des approches d'apprentissage machine traditionnelles.

En effet, l'IA générative se distingue par sa facilité d'utilisation sans exigences de données importantes, puisqu'elle est déjà formée sur de vastes ensembles de données, permettant son application immédiate à diverses tâches. De plus, ces modèles, qui peuvent être exploités en langage naturel sans compétences en codage spécifiques, sont souvent disponibles gratuitement en ligne, sans nécessiter d'infrastructures ou de configurations de *machine learning* locales. De ce fait, l'analyse et la génération de langage, des domaines auparavant réservés aux spécialistes, sont rendues largement accessibles et pratiques pour un usage professionnel.

Figure 11. Utilisation de l'intelligence artificielle

% d'entreprises de 10 employés ou plus.



Source : Enquête Utilisation des TIC et de l'e-commerce dans les entreprises, Statbel.

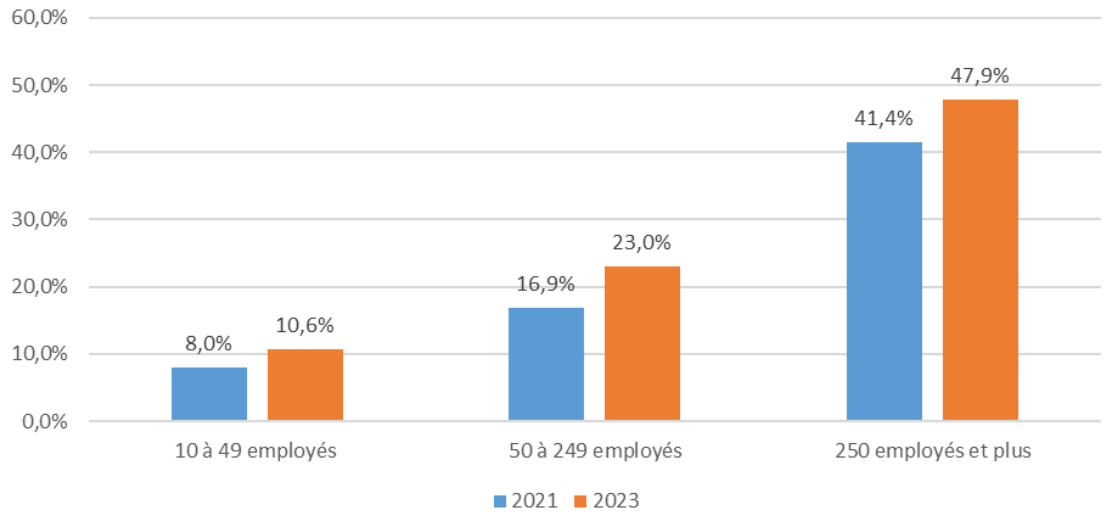
Comme dans d'autres pays européens, l'adoption de l'IA par les petites et moyennes entreprises (PME) en Belgique est plus lente comparativement aux grandes entreprises (Figure 12). Chez ces dernières (celles comptant plus de 250 employés), l'utilisation de l'IA est significativement plus élevée (47 %).

Cette tendance se vérifie également dans le domaine du big data. En 2020, 21,8 % des PME belges ont exploité les big data. La taille de l'entreprise est fortement corrélée à l'utilisation des big data : 56,8 % des grandes entreprises s'y adonnaient, contre seulement 19,4 % des petites entreprises[15].

[15] Source : Belgian Digital Economy Overview.

Figure 12. Utilisation de l'intelligence artificielle

% d'entreprises, en fonction de la taille.



Source : Enquête Utilisation des TIC et de l'e-commerce dans les entreprises, Statbel.

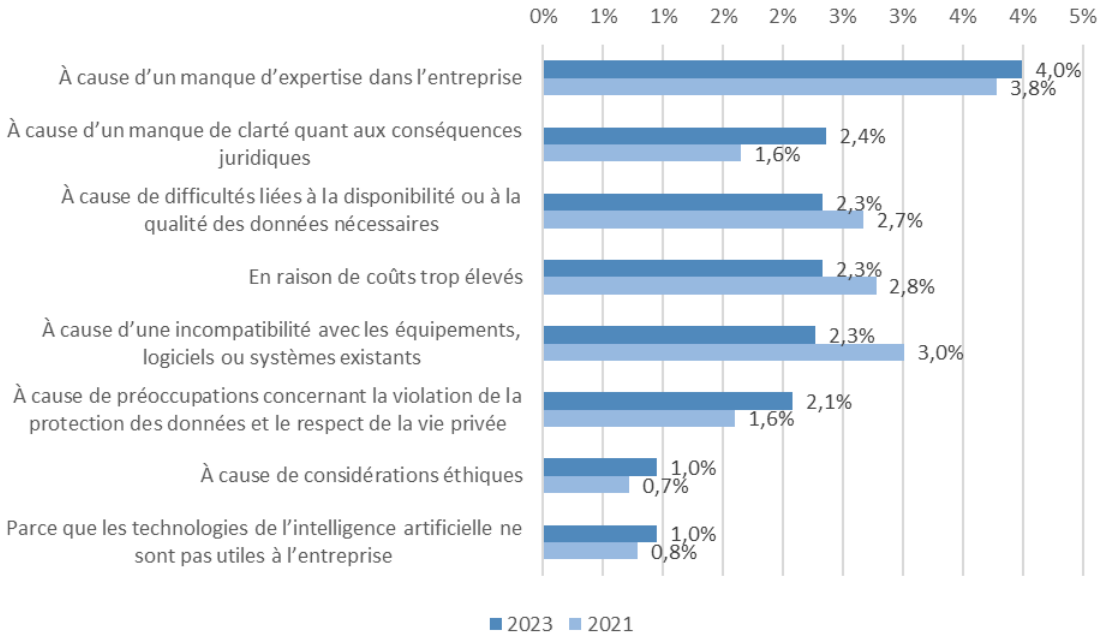
Cette variation dans la vitesse d'adoption souligne un défi commun à l'échelle de l'Union européenne. Toutefois, de manière générale, les entreprises belges se montrent globalement performantes en ce qui concerne l'intégration de ces technologies. La Belgique se distingue notamment en occupant la quatrième place dans l'UE pour l'adoption de l'IA par les entreprises, tous secteurs confondus[16].

Le manque d'expertise au sein d'une entreprise semble être le principal obstacle à l'utilisation de l'IA dans les entreprises en Belgique (Figure 13). Les coûts, le manque de clarté quant aux conséquences juridiques, l'incompatibilité avec l'infrastructure technologique existantes, ainsi que la disponibilité et la qualité des données nécessaires, figurent aussi parmi les obstacles majeurs.

[16] Source : The economic opportunity of AI in Belgium.

Figure 13. Raisons invoquées par les entreprises pour ne pas utiliser l'IA

% d'entreprises de 10 employés ou plus.



Source : Enquête Utilisation des TIC et de l'e-commerce dans les entreprises, Statbel.

L'analyse de l'évolution entre 2021 et 2023 révèle une tendance intéressante. Alors que les barrières techniques et les coûts paraissent régresser, suggérant que l'accessibilité à l'IA s'améliore — probablement sous l'influence de l'émergence de l'IA générative —, les préoccupations liées à la protection des données et du respect de la vie privée, des implications juridiques et des considérations éthiques se renforcent. Enfin, la perception de l'utilité de l'IA a légèrement fléchi, avec une petite hausse des entreprises considérant l'IA comme non pertinente pour leurs activités.

Ces obstacles sont de nature à ralentir l'adoption de l'IA. Le manque de compétences spécifiques pour exploiter pleinement le potentiel de l'IA générative peut constituer une barrière significative pour les PME. En outre, l'incertitude réglementaire entourant l'utilisation de l'IA générative peut augmenter les risques d'implémentation et les coûts de conformité, ce qui est particulièrement problématique pour les PME qui n'ont pas forcément les capacités juridiques internes pour naviguer dans ce paysage complexe.

2. L'économie de la donnée en Belgique

Principaux résultats

Entreprises de données

- Selon IDC, les entreprises de données sont les organisations qui sont directement impliquées dans la production, la livraison et/ou l'utilisation de données sous forme de produits numériques, de services et de technologies.
- Elles se déclinent en deux catégories : les entreprises fournisseuses de données, qui représentent le côté « offre » du marché des données et les entreprises utilisatrices de données, qui représentent le côté « demande » du marché des données

	Entreprises fournisseuses de données (Offre)	Entreprises utilisatrices de données (Demande)
Définition	▸ Entreprises ayant pour activité principale la production et la livraison de produits, de services et de technologies liés aux données numériques	▸ Entreprises qui génèrent, exploitent, collectent et analysent intensivement les données
Nombre (2023)	▸ 4.586 entreprises ▸ Part d'entreprises de ce type inférieure à la moyenne UE ▸ Croissance (2020-2023) : 37,3 % en ligne avec la croissance UE	▸ 16.120 entreprises ▸ Part d'entreprises de ce type qui s'aligne sur la moyenne UE ▸ Croissance (2020-2023) : 11,6 % légèrement en deçà de la croissance UE
Chiffre d'affaires (2023)	▸ 2.837 millions d'euros ▸ soit 0,3 % du CA total des entreprises belges ▸ Poids économique inférieur à celui des pays voisins et à la moyenne de l'UE ▸ CA montre une croissance encourageante mais inférieure à la moyenne UE	▸ N.D.
Profil	▸ Secteurs de l'information et communication (48,1 %) et services professionnels (39,9 %) ▸ Répartition sectorielle semblable à l'UE ▸ Principalement des entreprises comptant 1 à 249 employés	▸ Secteurs des services professionnels (38,1 %), construction (12,5 %), information et communication (9,3 %) et santé (6,2 %) ▸ La répartition sectorielle diffère par rapport à l'UE : plus d'entreprises dans les services professionnels et la construction, moins dans le transport, l'industrie manufacturière et la finance
Observations	▸ L'Allemagne et le Royaume-Uni se distinguent non seulement par une proportion beaucoup plus élevée d'entreprises fournisseuses et utilisatrices mais aussi par une plus forte croissance de cette proportion sur la période 2020 et 2023.	

Marché des données

- Le marché des données a atteint 2,7 milliards d'euros en Belgique en 2023.
- La trajectoire de croissance est en phase avec la moyenne européenne, mais affiche pour 2023 une hausse plus importante (+15,2 % contre 11,1 % pour l'UE).
- L'importance de la monétisation des données apparait comme un aspect émergent du marché, dans un contexte où les avancées dans l'IA générative incitent à un intérêt croissant pour l'acquisition de données de qualité.
- En 2022, les principaux secteurs contribuant au marché des données en Belgique étaient la finance, le secteur public, l'industrie manufacturière, et l'information & la communication, représentant plus de deux tiers de la valeur du marché.
- Certains secteurs en Belgique, comme les services professionnels et la finance, ont connu une croissance significative, bien que d'autres, comme l'agriculture et le marché intérieur, ont montré des taux de croissance plus modestes.
- Comparativement à l'UE, le secteur public détient une part plus importante en Belgique.

L'impact du marché des données sur l'économie belge

- L'économie des données en Belgique a connu une croissance notable, passant de 14,860 milliards d'euros en 2020 à 17,795 milliards d'euros en 2023, avec un taux annuel moyen de croissance de 6 %.
- L'importance de l'économie des données dans le PIB belge est en hausse, atteignant 4,2 % en 2023, ce qui la positionne dans la moyenne européenne et devant des pays comme la France, mais derrière l'Allemagne, les Pays-Bas, et le Royaume-Uni.
- La Belgique contribue à hauteur de 3,3 % à l'économie des données de l'UE-27, se plaçant légèrement en dessous de sa contribution au PIB de l'Union (3,5 %).
- Selon les prévisions, l'économie des données pourrait représenter jusqu'à 4,8 % du PIB belge d'ici 2025, soit un montant estimé à 21,1 milliards d'euros, bien que cette expansion apparaisse comme plus modeste en comparaison avec les voisins européens sur la période 2020-2025.
- Le secteur public est le plus impacté par l'économie des données, suivi par le secteur financier, les services professionnels, l'industrie manufacturière, et le secteur des transports. Notamment, le secteur de la santé et ceux de la construction et de l'agriculture montrent une part d'impact relativement faible.

Les retombées futures potentielles de l'IA générative

- L'IA générative est identifiée comme une avancée majeure avec le potentiel d'augmenter le PIB belge de 45 à 50 milliards d'euros annuellement, ce qui représenterait une hausse de 9 % sur une décennie.
- Un retard dans l'adoption de l'IA générative pourrait réduire considérablement les gains potentiels pour le PIB belge, mettant en évidence l'urgence d'intégrer cette technologie pour saisir pleinement ses avantages économiques.

2.1. Introduction

L'estimation de la valeur des données à un niveau agrégé est un véritable défi aujourd'hui, notamment en raison de la rapidité du changement technologique et, ironiquement, du manque de disponibilité des données relatives à l'utilisation des données et à leurs impacts.

Malgré l'omniprésence des données et l'avancement technologique permettant de mesurer leur volume, comme l'a fait le McKinsey Global Institute en quantifiant les flux de données transfrontaliers[17], l'attribution d'une valeur monétaire spécifique à ces données demeure complexe. Cette difficulté réside, comme on l'a vu, dans le fait que les données, en elles-mêmes, ne possèdent pas intrinsèquement de valeur sans être agrégées, analysées et mises en contexte. Leur véritable valeur émerge de leur interaction avec d'autres données et actifs, ainsi que de la manière dont les organisations les exploitent. Sans une compréhension approfondie de ces synergies et de l'utilisation concrète des données par les entreprises, estimer leur valeur économique globale s'avère un défi considérable.

Face à cette complexité, il existe tout de même des initiatives comme celles d'IDC pour la Commission européenne qui cherchent à proposer des mesures. Le rapport final de l'étude « European Data Market study 2021-2023 », publié en février 2024, s'inscrit dans cette démarche en tentant de mesurer la valeur du marché des données et son impact économique au sein de l'Union européenne et dans des pays spécifiques comme la Belgique.

Pour ce faire, l'étude modélise les agrégats suivants :

- Les entreprises de données, qui sont distinguées en deux catégories : les entreprises utilisatrices de données et les entreprises fournisseuses de données.
- La valeur du marché des données, qui se base sur les revenus des entreprises fournisseuses de données sur des produits et services liés aux données, et une estimation de la valeur de la monétisation des données.
- L'impact du marché des données sur l'économie belge.

Nous passons dans ce chapitre en revue ces estimations réalisées par IDC.

Notons que l'étude élabore également d'autres indicateurs relatifs au marché du travail, qui sont présentés dans la section sur les compétences.

2.2. Les entreprises de données

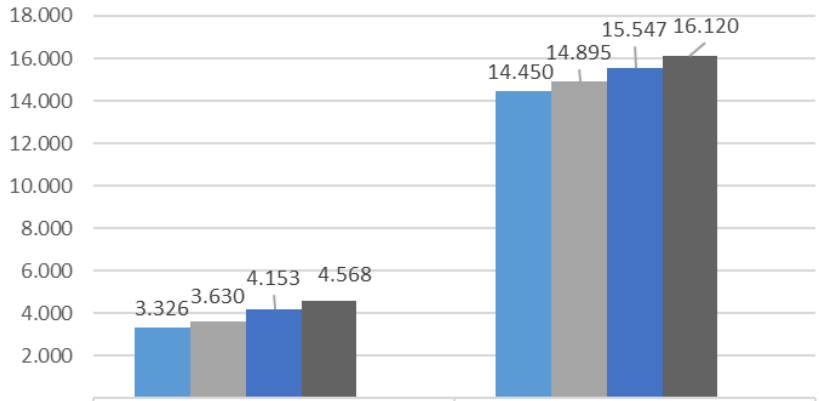
Les entreprises de données sont définies (IDC, 2024) comme des organisations qui sont directement impliquées dans la production, la livraison et/ou l'utilisation de données sous forme de produits numériques, de services et de technologies. Elles peuvent être à la fois des organisations fournisseuses de données et des utilisatrices de données :

- Les fournisseurs de données ont pour activité principale la production et la livraison de produits, de services et de technologies liés aux données numériques. Ils représentent le **côté offre** du marché des données.
- Les utilisatrices de données sont des organisations qui génèrent, exploitent, collectent et analysent intensivement des données numériques, et utilisent ce qu'elles apprennent pour améliorer leur entreprise. Elles représentent le **côté demande** du marché des données.

On retrouve significativement plus d'entreprises utilisatrices de données que d'entreprises fournisseuses de données, ce qui n'est pas surprenant étant donné que les entreprises fournisseuses de données sont souvent également des entreprises utilisatrices de données. En Belgique, on compte, en 2023, plus de 3,5 entreprises utilisatrices de données pour chaque entreprise fournisseuse de données alors que ce ratio pour l'ensemble de l'UE s'élève à 2,5.

[17] Estimer le volume des données et des flux de données a été rendu possible grâce aux technologies actuelles telles que le Cloud Computing. Le McKinsey Global Institute a estimé en 2019 le volume des flux de données transfrontaliers et son évolution, s'élevant à 704 téraoctets par seconde en 2017.

Figure 14. Évolution du nombre d'entreprises de données en Belgique



	Fournisseuses de données	Utilisatrices de données
2020	3.326	14.450
2021	3.630	14.895
2022	4.153	15.547
2023	4.568	16.120
Taux de croissance annuel moyen	11,2%	3,7%

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Entreprises fournisseuses de données

Le nombre d'entreprises fournisseuses de données enregistre une croissance constante dans tous les pays de l'Union européenne, ainsi que dans l'ensemble de la région (EU-27). Entre 2020 et 2023, ce nombre a augmenté de 37,3 % pour la Belgique, tandis que la croissance de l'ensemble de l'UE atteint 35,8 %.

Ces entreprises représentent toutefois une petite part du paysage économique. En Belgique, leur nombre était de 4.568 en 2023, que l'on retrouve principalement dans le secteur de l'information et de la communication (48,1 %) ainsi que dans les services professionnels (39,9 %)[18] et qui sont en grande majorité des entreprises comptant de 1 à 249 employés (99,7 %).

Lorsqu'on rapporte le nombre d'entreprises fournisseuses de données au total d'entreprises (à noter que ce total diffère du dénominateur utilisé pour calculer la proportion d'entreprises utilisatrices de données et concerne une sélection d'entreprises appartenant à des secteurs clés[19]), on observe des disparités marquées entre les pays de l'Union européenne. Certains pays se démarquent avec un taux nettement plus élevé que les autres, notamment le Royaume-Uni, qui affiche 16,9 % d'entreprises de ce type (Tableau 1). D'autres pays, tels que l'Estonie et la Grèce, présentent également des pourcentages élevés, respectivement 12,2 % et 10,4 %. En revanche, la France (1,0 %), les Pays-Bas (1,0 %) et l'Espagne (1,4 %) affichent des proportions nettement plus faibles. La Belgique, avec un taux de 1,4 %, se classe dans cette dernière catégorie.

Il est également intéressant de noter que l'évolution de ces taux varie d'un pays à l'autre entre 2020 et 2022. Si l'augmentation est présente partout, elle se révèle plus modeste en Belgique (+0.3) ou en France (+0,2) par rapport au Royaume-Uni (+3.2) ou à l'Allemagne (+0.6).

[18] Il s'agit de proportions très similaires à la moyenne de l'UE-27 (49,3 % et 39,3 %). Cette distribution sectorielle a été calculée sur base des chiffres de 2022 obtenus auprès d'Agoria.

[19] Le dénominateur utilisé pour calculer la proportion des entreprises fournisseuses de données est limité aux entreprises relevant des codes NACE suivants : A : Agriculture, sylviculture et pêche, C : Industrie manufacturière, D : Fourniture d'électricité, de gaz, de vapeur et d'air conditionné, E : Approvisionnement en eau, assainissement, gestion des déchets et activités de dépollution, G : Commerce de gros et de détail ; réparation de véhicules automobiles et de motocycles, H : Transports et entreposage, J : Information et communication, K : Activités financières et d'assurance, M : Activités professionnelles, scientifiques et techniques, P : Éducation, Q : Activités de santé humaine et d'action sociale.

Tableau 1. Entreprises fournisseuses de données

% du total d'entreprises appartenant à 12 secteurs[20].

	2020	2021	2022	2023	Croissance 2023/2020 (en points de %)
Belgique	1,2 %	1,3 %	1,4 %	1,5 %	+0.3
France	0,8 %	0,9 %	1,0 %	1,0 %	+0.2
Allemagne	2,2 %	2,4 %	2,6 %	2,8 %	+0.6
Pays-Bas	0,8 %	0,9 %	1,0 %	1,1 %	+0.2
Royaume-Uni	14,7 %	15,5 %	17,0 %	17,8 %	+3.2
Total UE-27	1,7 %	1,8 %	2,0 %	2,1 %	+0.5

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Les entreprises fournisseuses de données représentent une part de marché relativement modeste tant au niveau national qu'européen. En 2023, les 4.568 entreprises belges fournisseuses de données ont généré un chiffre d'affaires de 2.837 millions d'euros, ce qui représente une croissance de 19,2 % par rapport à 2020.

Notons par ailleurs que la croissance au niveau de l'UE était plus importante sur la même période (30,7 %) et donc, par conséquent, la Belgique a perdu du poids relatif au sein de l'économie de l'UE dans ce domaine. En 2023, elle représentait 3,06 %, ce qui est légèrement en-deçà du poids du PIB belge dans l'UE (3,5 %).

Tableau 2. Évolution des entreprises belges fournisseuses de données (2020-2023)

Selon plusieurs indicateurs-clé.

	2020	2021	2022	2023
Nombre	3.326	3.630	4.153	4.568
Chiffre d'affaires (millions d'euros)	2.381	2.603	2.735	2.837
Chiffre d'affaires (en % du CA total des entreprises belges)	0,25 %	0,27 %	0,29 %	0,30 %
Chiffre d'affaires (en % du CA européen des entreprises fournisseuses de data)	3,35 %	3,46 %	3,24 %	3,06 %

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

En outre, le poids de ces entreprises en pourcentage du chiffre d'affaires total des entreprises belges est demeuré relativement faible, passant de 0,25 % en 2020 à 0,30 % en 2022. Les entreprises fournisseurs de données ont un poids plus élevé chez les pays voisins et dans l'UE-27 dans son ensemble. En Allemagne et aux Pays-Bas, elles représentent respectivement 0,42 % et 0,39 % des entreprises tandis qu'au Royaume-Uni, il s'élève à 0,74 % (Tableau 3).

[20] Entreprises appartenant aux Codes NACE A, C, D, E, G ,H, J, K, M, P et Q.

Tableau 3. Évolution de la part du chiffre d'affaires réalisé par les entreprises fournisseuses de données

	2020	2021	2022	2023
Belgique	0,25 %	0,27 %	0,29 %	0,30 %
France	0,38 %	0,42 %	0,48 %	0,54 %
Allemagne	0,32 %	0,36 %	0,39 %	0,42 %
Pays-Bas	0,37 %	0,41 %	0,39 %	0,39 %
Royaume-Uni	0,59 %	0,62 %	0,71 %	0,74 %
Total EU-27	0,31 %	0,35 %	0,38 %	0,40 %

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

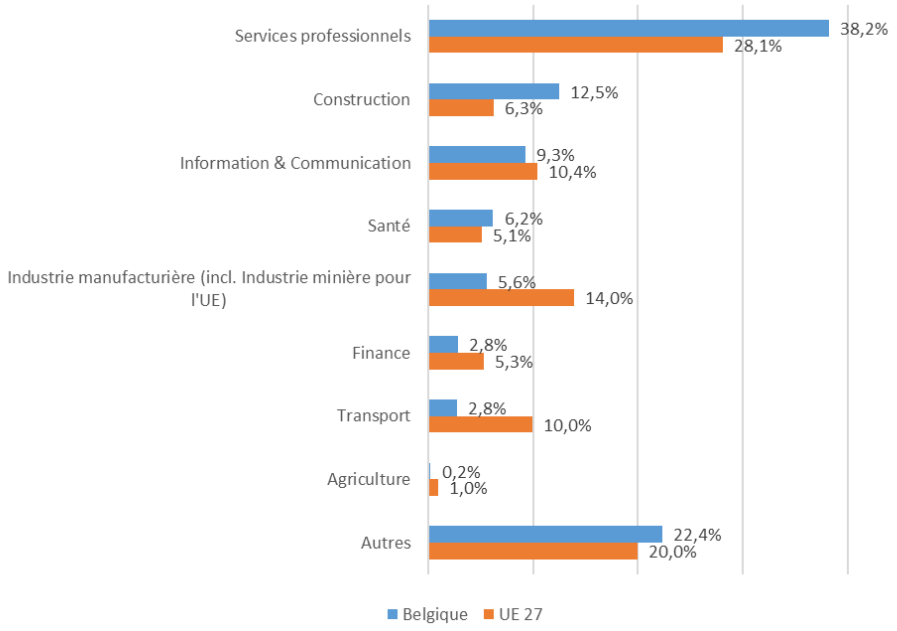
Cette tendance souligne que bien que le secteur des entreprises fournisseuses de données montre une croissance encourageante, il représente toujours une part marginale de l'économie belge.

Entreprises utilisatrices de données

Les entreprises utilisatrices de données sont plus nombreuses que les fournisseurs. En Belgique, on compte, en 2023, plus de 3,5 entreprises utilisatrices de données pour chaque entreprise fournisseuse de données. Entre 2020 et 2023, leur nombre est passé de 14.450 en 2020 à 16.120 en 2023, soit une augmentation de 11,6 %.

Ces entreprises se retrouvent, en 2022, principalement concentrées dans les services professionnels, pour 38,2 % d'entre elles, ce qui dépasse la moyenne de l'UE à 27 pays, qui est de 28,1 % (Figure 15). La Belgique affiche également une forte présence dans le secteur de la construction, avec 12,5 % contre seulement 6,3 % pour l'UE. En revanche, les secteurs de la finance, de la santé, de l'industrie manufacturière et du transport sont sous-représentés par rapport à la moyenne de l'UE.

Figure 15. Répartition des entreprises utilisatrices de données par secteur



Source : Élaboration d'Idea Consult sur base de données d'IDC (2024) et d'Agoria.

Note : La catégorie « Autres » regroupe les secteurs éducatifs, l'administration publique, le commerce de détail et de gros et les services d'utilité publique.

Malgré une croissance marquée, la proportion d'entreprises utilisatrices sur le total d'entreprises privées en Belgique est restée relativement stable entre 2020 et 2023 ; autour des 2,4 % (Tableau 4), ce qui indique que la croissance du nombre de ces entreprises est restée en phase avec celle du total d'entreprises. Une tendance similaire, c'est-à-dire une proportion relativement faible d'entreprises utilisatrices et relativement inchangée dans la période, peut être observée pour la France et les Pays-Bas.

En revanche, l'Allemagne et le Royaume-Uni se distinguent non seulement par une proportion plus élevée d'entreprises utilisatrices (4,8 % et 7,5 % en 2023) mais aussi par une plus forte croissance de cette proportion sur la période.

Tableau 4. Proportion des entreprises utilisatrices de données dans l'économie

	2020	2021	2022	2023
Belgique	2,3 %	2,3 %	2,4 %	2,4 %
France	1,8 %	1,8 %	1,8 %	1,9 %
Allemagne	4,3 %	4,5 %	4,6 %	4,8 %
Pays-Bas	2,0 %	1,9 %	2,0 %	2,0 %
Royaume-Uni	7,1 %	7,0 %	7,3 %	7,5 %
Total EU-27	2,1 %	2,1 %	2,2 %	2,3 %

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

En résumé, la tendance générale est à l'augmentation du nombre d'entreprises utilisatrices de données dans tous les pays ainsi que dans l'UE dans son ensemble, témoignant de l'importance croissante de l'économie des données. Cependant, la proportion par rapport au total des entreprises privées reste relativement stable en Belgique, ce qui suggère que les entreprises utilisatrices de données gagnent en importance économique sans nécessairement bouleverser radicalement la structure globale des entreprises dans ces pays, tandis que dans d'autre pays, tels que l'Allemagne et le Royaume-Uni, la culture de la donnée se diffuse davantage.

2.3. Le marché des données

Le marché des données est défini comme le lieu où les données numériques sont échangées en tant que « produits » ou « services » suite à l'élaboration de données brutes. Le marché des données « capte la valeur agrégée de la demande de données numériques sans mesurer les impacts directs, indirects ou induits des données dans l'économie dans son ensemble » (IDC, 2024).

La valeur du marché des données[21], a atteint dans l'UE-27 82 milliards d'euros en 2023, avec une augmentation de 11,1 % par rapport à l'année précédente. Le marché des données en Belgique s'estimait lui à 2,7 milliards d'euros.

[21] Pour réaliser ses estimations, IDC se base sur :

- Les revenus des entreprises de données. Les revenus des entreprises de données correspondent à la valeur agrégée de tous les produits et services liés aux données générées par les fournisseurs de données. Cet indicateur mesure donc les revenus des entreprises fournisseuses de données sur une série de produits et services liés aux données et précisées dans la méthodologie de la façon suivante : « Les parts respectives pour les dépenses en logiciels, matériel, services et monétisation des données sont issues d'enquêtes IDC portant sur les Big Data, les modèles et intentions de dépenses en TI sur le marché européen, ainsi qu'une enquête sur les entreprises fournisseurs de données et les entreprises utilisatrices de données dans les principaux États membres, complétées par l'expertise des analystes et l'harmonisation avec les prévisions d'IDC pour l'Europe et le monde concernant le marché de l'analyse commerciale et des Big Data. »

La valeur du marché des données n'est pas exactement égale aux revenus agrégés des entreprises européennes de données car elle inclut les importations (produits et services de données achetés sur le marché numérique mondial auprès de fournisseurs non basés en Europe) et exclut les exportations des entreprises européennes de données.

Le Tableau 5 révèle que le marché des données belge a suivi une trajectoire de croissance en phase avec la moyenne européenne, tout en affichant pour 2023 une hausse importante de sa valeur de marché (+15,2 %). Cette progression surclasse le taux de croissance de l'UE-27 (+11,1 %), soulignant une dynamique récente positive de l'économie des données en Belgique.

Tableau 5. Evolution du marché des données

En millions d'euros.

	2020	2021	2022	2023	Croissance annuelle moyenne	Croissance 2023/2022
Belgique	1.984	2.083	2.360	2.719	11,1 %	15,2 %
France	9.986	10.785	12.313	14.160	12,3 %	15,0 %
Allemagne	16.966	17.993	20.607	22.465	9,8 %	9,0 %
Pays-Bas	4.514	4.815	5.580	6.117	10,7 %	9,6 %
Royaume-Uni	17.845	19.204	22.062	24.503	11,1 %	11,1 %
Total EU-27	60.635	64.820	73.754	81.940	10,6 %	11,1 %

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Pour réaliser ces estimations, IDC se base non seulement sur les revenus des entreprises fournisseuses mais aussi sur une estimation de la valeur de monétisation des données, calculée séparément. En Belgique, les chiffres montrent une augmentation de la valeur de la monétisation des données de 220 millions d'euros en 2020 à 600 millions d'euros en 2023, avec ici aussi, un croissance accélération.

Encadré 5. L'importance croissante de la monétisation des données

La monétisation des données implique la vente de données entre organisations plutôt que la vente de logiciels, de matériel ou de services liés aux données. Les types de données vendues comprennent les données de localisation (informations sur la position des biens ou des personnes dans l'espace et le temps), les données d'achat (ce que les gens achètent), les données de performance des équipements (informations sur le fonctionnement des équipements), ainsi que les données d'activité des clients et les données personnelles, qui doivent souvent être agrégées et/ou anonymisées pour se conformer au RGPD (IDC, 2024).

Notons que les avancées récentes dans le domaine de l'IA générative suscitent un intérêt croissant de la part des grandes entreprises qui développent des modèles de langage pour l'acquisition de données de qualité. Par exemple, Tumblr et WordPress.com se prépareraient à vendre les données utilisateur à Midjourney et OpenAI[22], tandis que l'entreprise Photobucket discute avec des entreprises technologiques pour vendre 13 milliards de photos et vidéos pour entraîner des IA, avec

- Une estimation de la valeur de la monétisation des données qui est calculée séparément. La monétisation des données implique la vente de données entre organisations (plutôt que la vente de logiciels, de matériel ou de services liés aux données). IDC estime cette valeur sur base de recherches documentaires et des résultats d'une enquête ad hoc menée auprès des entreprises de données. L'enquête estime la taille du marché de la monétisation des données, en demandant aux répondants leur utilisation des marchés de données pour acheter et vendre des données, ainsi que le type de données vendues. Bien que ces données soient relativement faibles en raison de la nouveauté du marché et du manque de clarté parmi les répondants sur ce qui constitue la vente et l'achat de données, elles servent de guide pour un modèle de pénétration simple. Ce modèle utilise une part du marché total des données pour estimer la monétisation des données.

[22] https://www.404media.co/tumblr-and-wordpress-to-sell-users-data-to-train-ai-tools/?utm_source=www.therundown.ai&utm_medium=newsletter&utm_campaign=openai-strikes-huge-deal.

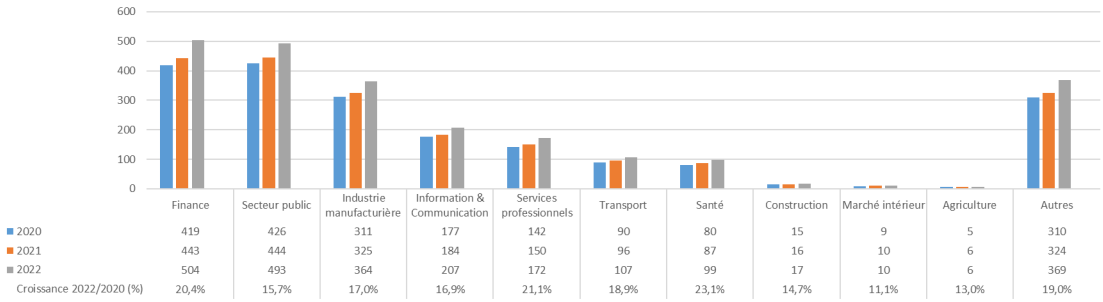
des tarifs variant de 5 cents à plus d'1 dollar par unité[23]. De son côté, Reddit a déclaré dans un document adressé à la Securities and Exchange Commission (l'organisme fédéral américain de réglementation des marchés financiers) que les publications de ses utilisateurs constituent « une source précieuse de données de conversation et de connaissances » qui a été et continuera d'être un mécanisme important pour l'entraînement de l'IA et des modèles de langage[24]. Ces actions s'inscrivent dans ce que Reuters qualifie de « course secrète des Big tech pour acquérir des données d'entraînement pour l'IA ».

Cette quête des données n'est pas sans soulever d'importants dilemmes éthiques, notamment concernant le respect de la vie privée et le consentement explicite des utilisateurs dont les données sont monétisées. La transparence quant à l'usage de ces données et leur sécurisation devient cruciale pour éviter les abus. En outre, ces pratiques interrogent sur l'équité de l'utilisation des données personnelles pour l'entraînement des IA, mettant en lumière le risque de renforcer des biais existants ou de créer de nouvelles formes d'inégalités.

L'année précédente, en 2022, les secteurs qui généraient la plus grande part de la valeur du marché des données en Belgique étaient les secteurs de la finance (€ 504 millions), le secteur public (€ 493 millions), l'industrie manufacturière (€ 364 millions) et le secteur de l'information et la communication (€ 207 millions). Ces quatre secteurs représentaient plus de deux tiers de la valeur liée au marché des données.

On retrouve ensuite les secteurs des services professionnels (€ 172 millions), du transport (€ 107 millions), de la santé (€ 99 millions) et plus loin dans la liste, avec une contribution plus marginale, les secteurs de la construction, du marché Intérieur et de l'agriculture (avec respectivement € 17, € 10, € 6 millions). Enfin, la catégorie Autres représente € 369 millions, soit 15,7 % de l'économie.

Figure 16. Évolution du marché de la donnée en Belgique, par secteur (2020 – 2022)



	Finance	Secteur public	Industrie manufacturière	Information & Communication	Services professionnels	Transport	Santé	Construction	Marché intérieur	Agriculture	Autres
2020	419	426	311	177	142	90	80	15	9	5	310
2021	443	444	325	184	150	96	87	16	10	6	324
2022	504	493	364	207	172	107	99	17	10	6	369
Croissance 2022/2020 (%)	20,4%	15,7%	17,0%	16,9%	21,1%	18,9%	23,1%	14,7%	11,1%	13,0%	19,0%

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Par comparaison, au sein de l'UE-27 (Figure 17. Répartition de la valeur du marché de la donnée par secteur en Belgique et dans l'UE (2022)Figure 17), la répartition de la valeur des données par secteur est semblable, avec toutefois quelques différences notables. La finance détient également une part importante (21,2 %), mais c'est l'Industrie manufacturière (incluant l'industrie minière) qui se retrouve en deuxième position, avec 20,7 % du marché, contre seulement 15,5 % en Belgique tandis que le secteur public en Belgique détient une part plus importante (21,0 %) comparée à l'UE (14,5 %). Cette différence marquée peut suggérer que la Belgique investit davantage dans la digitalisation et l'utilisation des données dans ses services publics par rapport à la moyenne de l'UE, reflétant un engagement plus fort envers la transformation numérique dans l'administration publique.

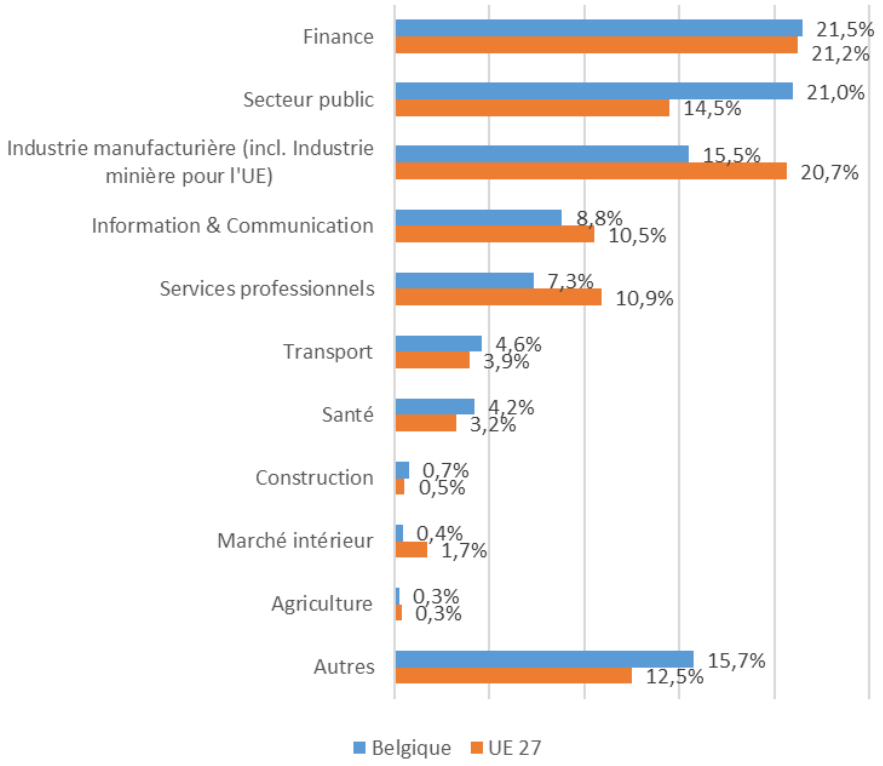
Les secteurs de l'information & communication et des services professionnels apparaissent quant à eux légèrement sous-représentés avec respectivement 8,8 % et 7,3 %, contre 10,5 % et 10,9 % dans l'UE, tandis que les secteurs du transport et de la santé en Belgique ont des parts similaires à celles

[23] https://www.reuters.com/technology/inside-big-techs-underground-race-buy-ai-training-data-2024-04-05/?utm_source=www.therundown.ai&utm_medium=newsletter&utm_campaign=big-tech-s-secret-race-for-ai-data.

[24] https://www.404media.co/reddit-we-are-in-the-early-stages-of-monetizing-our-user-base-2/.

de l'UE, bien que légèrement supérieures dans le cas du transport (4,6 % contre 3,9 %) et inférieures dans celui de la santé (4,2 % contre 3,2 %).

Figure 17. Répartition de la valeur du marché de la donnée par secteur en Belgique et dans l'UE (2022)



Source : Idea Consult sur base d'IDC (2024).

En ce qui concerne l'évolution par secteur, sur la période 2020 à 2022, le secteur de la santé affiche la plus forte croissance avec 23,1 %. Cette croissance est d'autant plus notable qu'elle s'inscrit dans un contexte où la santé est devenue un domaine clé, stimulé par les défis posés par la pandémie de COVID-19 et l'accent mis sur les données pour la prise de décision et l'innovation dans le domaine de la santé. La part du secteur de la santé reste toutefois modeste et ne représente en 2022 que 3,2 % du marché belge des données (Figure 18).

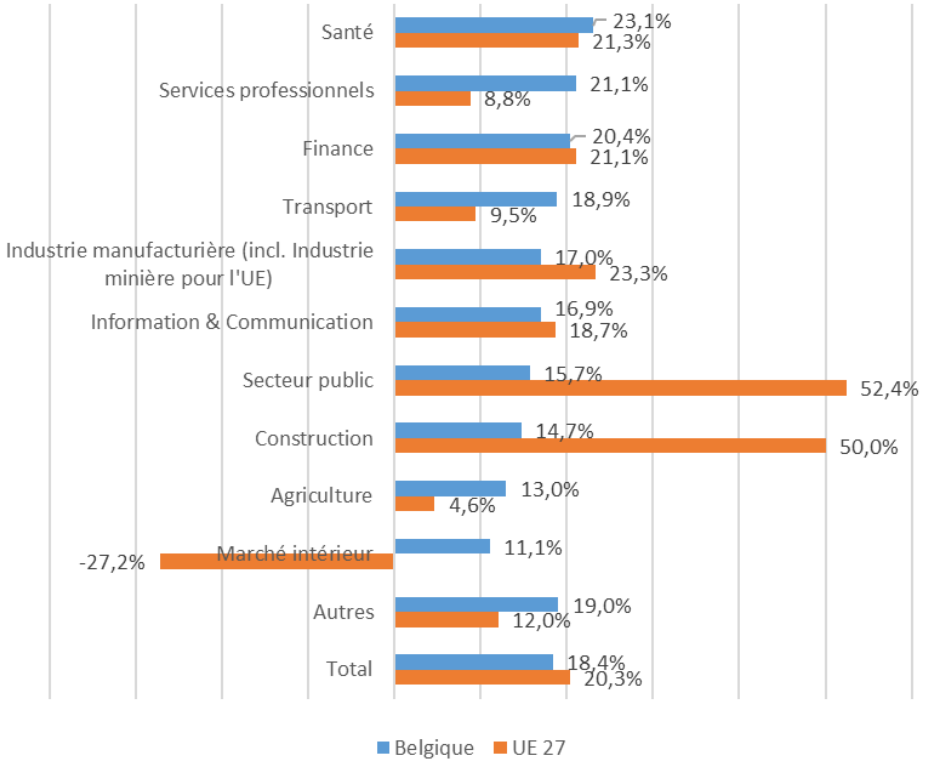
Les services professionnels et la finance sont également des secteurs en forte croissance, avec respectivement 21,1 % et 20,4% de croissance sur la période. Ces taux, bien qu'inférieurs à celui de la santé, sont significatifs, en particulier pour la finance qui part d'un volume déjà élevé. Les services professionnels, qui affichent une croissance robuste, représentent une part plus faible en Belgique (7,3 %) comparée à celle de l'UE (10,9 %), révélant ainsi une adoption rapide des données dans ce domaine.

D'autre secteurs tels que l'agriculture et le marché intérieur, bien que plus petits en termes de volume, ont connu des taux de croissance relativement modestes (13,0 % et 11,1 % respectivement). Cela suggère que ces domaines, bien qu'importants, n'ont pas connu le même dynamisme que d'autres secteurs dans le contexte de l'évolution du marché des données. Leur taux de croissance est toutefois bien au-dessus de la moyenne de l'UE, avec un point de divergence frappant en ce qui concerne le marché intérieur, qui enregistre une baisse considérable dans l'UE (-27,2 %).

Dans d'autres secteurs, la Belgique affiche des bons taux de croissance mais généralement inférieurs à ceux de l'UE-27, comme c'est le cas pour l'information & communication et l'industrie manufacturière, à l'exception du secteur du transport où la Belgique surpasse fortement l'UE (+18,9 % contre +9,5 %).

Enfin, dans certains secteurs, on constate dans l'UE-27 une croissance plus marquée. Les taux enregistrés pour le secteur public (52,4 %) et la construction (50,0 %), dépassent ainsi largement les taux observés en Belgique pour ces mêmes secteurs.

Figure 18. Évolution de la valeur du marché de la donnée par secteur en Belgique et dans l'UE (2020 - 2022)



Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

2.4. L'impact du marché des données sur l'économie belge

Selon l'étude d'IDC (2024), l'économie des données mesure les impacts globaux du marché des données sur l'économie dans son ensemble. Elle englobe la génération, la collecte, le stockage, le traitement, la distribution, l'analyse, l'élaboration, la livraison et l'exploitation des données rendues possibles par les technologies numériques. L'économie des données englobe un concept plus large que le seul marché des données, car elle prend en compte la valeur et la richesse générées dans l'économie dans son ensemble (et non seulement au sein des entreprises) grâce à l'exploitation des données[25].

[25] L'économie des données inclut trois ensembles d'impacts dans l'économie (IDC, 2024) :

1. **Les impacts directs** sont les effets initiaux et immédiats générés par les entreprises fournisseuses de données ; ils représentent l'activité potentiellement engendrée par toutes les entreprises actives dans la production de données. Les impacts directs quantitatifs seront alors mesurés comme les revenus provenant des produits et services de données vendus, c'est-à-dire la valeur du marché des données. IDC considère la valeur du marché des données comme une bonne approximation des impacts directs. Par conséquent, par souci de simplicité, les impacts directs coïncideront avec la valeur du marché des données.
2. **Les impacts indirects** sont les activités économiques générées le long de la chaîne d'approvisionnement de l'entreprise par les entreprises fournisseuses de données, en considérant les fournisseuses d'entrée et les clients des entreprises fournisseuses de données.
3. **Les impacts induits** incluent l'activité économique générée dans l'ensemble de l'économie comme un effet secondaire. Les dépenses additionnelles induites sont générées à la fois par de nouveaux travailleurs, qui reçoivent un nouveau salaire, et par l'augmentation des salaires des emplois existants. Cette dépense induit

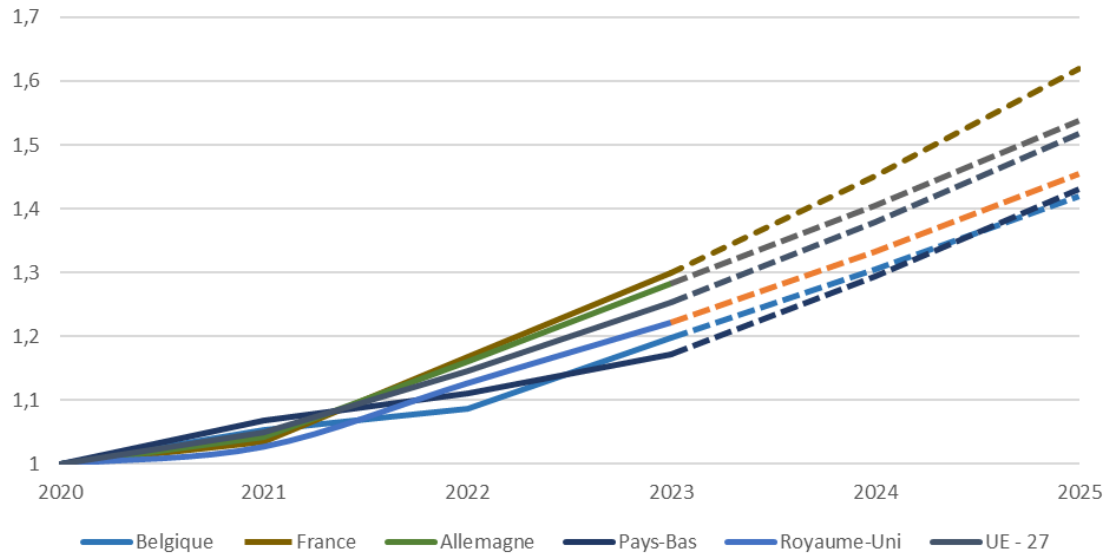
En Belgique, l'économie des données a connu une ascension marquée, atteignant 17,795 milliards d'euros en 2023, en hausse depuis les 14,860 milliards d'euros enregistrés trois ans plus tôt. Cette croissance, avec un taux annuel moyen de 6 %, témoigne de la vitalité du secteur. La Belgique contribue ainsi à hauteur de 3,3 % à l'économie des données de l'UE-27, un chiffre légèrement inférieur à sa contribution au PIB de l'Union, qui s'élève à 3,5 %.

Le poids de l'économie des données dans le PIB belge est également en augmentation, passant de 3,9 % en 2020 à 4,2 % en 2023, se plaçant ainsi dans la moyenne européenne. La Belgique devance la France, avec ses 3,1 %, mais se trouve derrière l'Allemagne, les Pays-Bas et le Royaume-Uni, qui affichent respectivement 4,9 %, 5,4 % et 6,9 %.

Selon les prévisions d'IDC, l'économie des données pourrait représenter jusqu'à 4,8 % du PIB belge d'ici 2025, soit 21,1 milliards d'euros. Cependant, comparativement à ses voisins et à la moyenne de l'UE, la Belgique pourrait avoir connu sur la période 2020-2025 une expansion plus modeste, si ces prévisions se réalisent.

Figure 19. Croissance réalisée et croissance attendue de l'économie de la donnée en Belgique

(2020 = 1).



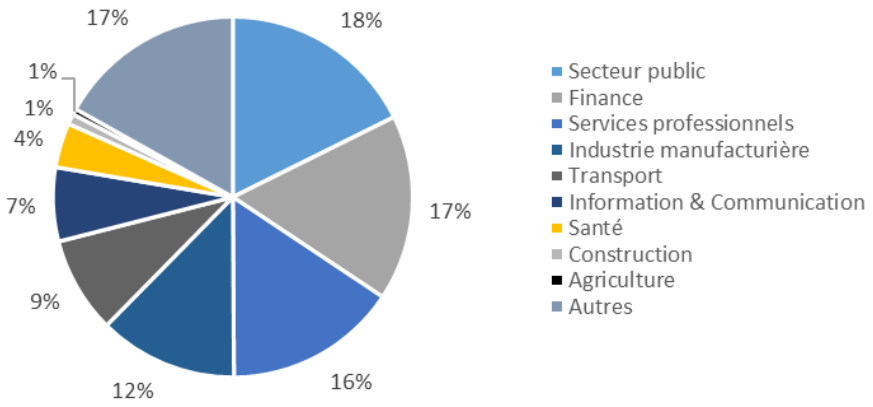
Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Ainsi, bien que la Belgique soit proche de la moyenne européenne en termes de part du PIB attribuable à l'économie de la donnée, sa croissance sur la période s'avère plus faible que celles des pays voisins.

Lorsqu'on regarde la répartition sectorielle des impacts globaux du marché des données, il apparait que le secteur public est le plus impacté, avec 18 % de la part totale. Suivent de près le secteur financier et les services professionnels, avec respectivement 17 % et 16 %. L'industrie manufacturière (12 %) et le secteur des transports (9 %) montrent également une utilisation significative des données. Plus surprenante peut-être est la part relativement faible allouée au secteur de la santé (4 %), qui semble avoir un impact moindre comparé aux autres secteurs. Les secteurs de la construction et de l'agriculture, chacun à 1 %, peuvent refléter un retard dans la numérisation et l'adoption de technologies avancées basées sur les données. Finalement, la catégorie « Autres » compte pour 17 %, un pourcentage non négligeable, suggérant que de nombreux autres secteurs tirent également parti des données.

la création de nouveaux revenus dans presque tous les secteurs de l'économie. La consommation supplémentaire soutiendra l'activité économique dans diverses industries telles que la vente au détail, les biens de consommation, les banques, le divertissement, etc.

Figure 20. L'économie de la donnée, par secteur



Source : Élaboration d'Idea Consult sur base des données d'Agoria.

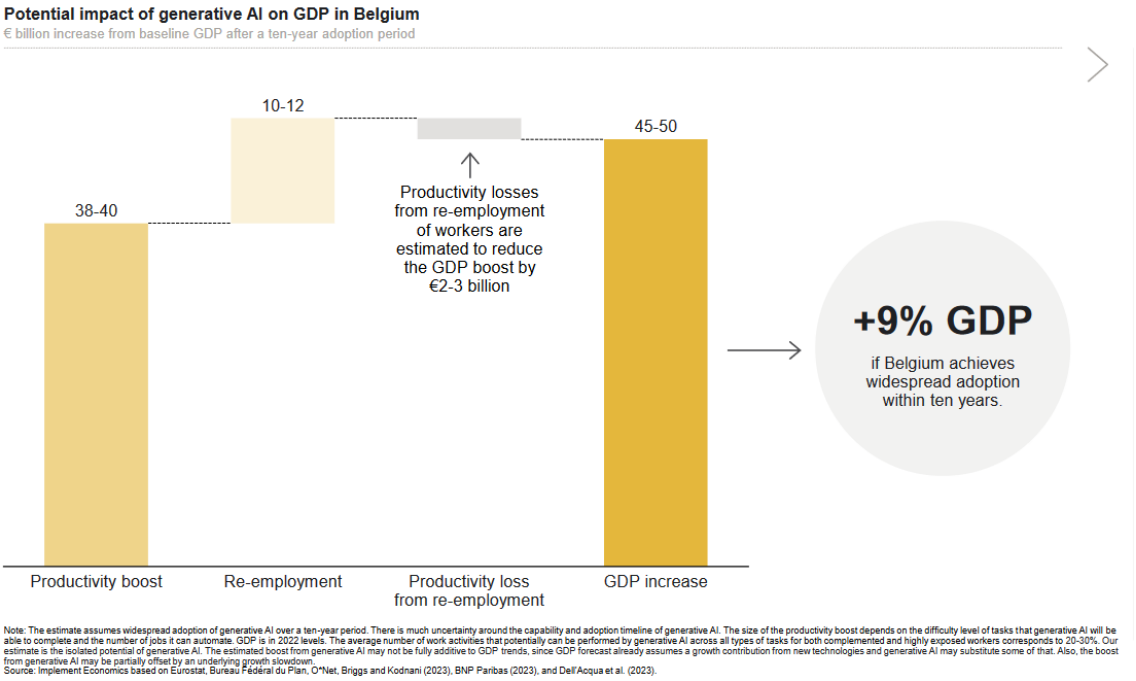
2.5. Les retombées futures potentielles de l'IA générative

L'opportunité économique et la croissance que représente l'intelligence artificielle (IA) pour la Belgique sont significatives et portent un potentiel de transformation majeure pour l'économie du pays. Selon le rapport « The economic opportunity of AI in Belgium »[26], l'IA générative, une avancée récente dans le domaine de l'intelligence artificielle, pourrait être le moteur d'une augmentation substantielle du PIB belge, estimée entre 45 et 50 milliards d'euros annuels. Cette croissance, représentant une hausse de 9 % sur une décennie, repose sur l'hypothèse d'une adoption généralisée de cette technologie à travers toutes les sphères de l'économie.

La Belgique, si elle saisit pleinement cette vague d'innovation IA, pourrait non seulement augmenter sa productivité mais aussi réinventer de nombreux secteurs d'activité grâce à l'automatisation et à la restructuration des tâches de travail. Les gains envisagés proviennent de plusieurs sources : une augmentation de la productivité grâce à l'assistance de l'IA générative dans les tâches quotidiennes, l'automatisation permettant de libérer du temps pour des activités à plus haute valeur ajoutée, et la réaffectation des ressources humaines vers des secteurs et des tâches où l'IA génère de nouveaux besoins et opportunités.

[26] https://cms.implementconsultinggroup.com/media/uploads/articles/2024/The-economic-opportunity-of-AI-in-Belgium/The-economic-opportunity-of-AI-in-Belgium.pdf.

Figure 21. Impact potentiel de l'IA générative sur le PIB belge



Source : The economic opportunity of AI in Belgium (2024).

Cependant, ce potentiel de croissance économique vient avec son lot de défis. Un retard dans l'adoption de l'IA générative pourrait considérablement diminuer les gains escomptés pour le PIB belge, selon le même rapport. Si la Belgique prend du retard sur le plan de l'adoption technologique, les gains potentiels pourraient se réduire drastiquement, passant de 9 % à seulement 2 % sur une période de dix ans.

3. Les compétences

Principaux résultats

Nombre de professionnels de la donnée :

- La Belgique affiche la plus faible proportion d'employés dans le domaine des données parmi ses pays voisins et se situe en dessous de la moyenne de l'UE-27.
- En 2023, elle occupe la 25e place sur 27 États membres, avec une part de professionnels des données représentant 3,1 % de l'ensemble de sa main-d'œuvre.
- Sur la période 2020 à 2023, la Belgique n'a pas enregistré de croissance significative de la part d'emploi des professionnels des données, contrairement à ses pays voisins et à la moyenne de l'UE-27, qui est en augmentation. Elle se positionne ainsi en dernière place dans l'UE en termes de croissance du nombre de professionnels des données.

Intensité des professionnels des données dans les entreprises

- La Belgique enregistre l'évolution la plus défavorable de l'UE concernant l'indicateur d'intensité des professionnels des données dans les entreprises. Cet indicateur a diminué de 9,4 à 8,7 entre 2020 et 2023, marquant une baisse de 6,8 %, alors que la tendance dans l'UE 27 et chez les pays voisins est à l'augmentation.
- Cette situation pourrait indiquer une réduction de la capacité d'analyse et de gestion des données au sein des entreprises belges, ce qui est préoccupant dans un contexte où l'usage des données est de plus en plus crucial pour la compétitivité et l'innovation.

Écart de compétences (pas de données disponibles pour 2023) :

- Entre 2019 et 2022, tant la demande que l'offre en professionnels de la donnée ont annuellement crû en Belgique.
- L'offre a augmenté de 118.500 à plus de 131.000, tandis que la demande est passée de 122.000 à 138.500.
- En 2022, les entreprises belges, tant fournisseuses qu'utilisatrices de données, avaient besoin d'environ 7.000 travailleurs de données pour combler le déficit de compétences en données.
- Contrairement à 2021, cet écart a diminué, une tendance non observée chez les voisins de la Belgique ni dans l'UE
- La réduction de cet écart en Belgique est principalement attribuable à une augmentation relativement modeste de la demande de professionnels de la donnée (+0,7 %), tandis que les pays voisins et l'UE ont connu une croissance plus significative. Cette situation interroge sur les particularités du marché du travail belge et sur la nécessité d'identifier les possibles obstacles à la demande de spécialistes de la donnée.

Gender gap:

- En 2023, dans tous les pays de l'OCDE sauf la Grèce, la proportion de jeunes hommes sachant programmer est supérieure à celle des femmes. La Belgique se distingue par un des plus faibles pourcentages de jeunes femmes sachant programmer (4,7 %), bien en dessous de la moyenne de l'OCDE de 9,9 %.
- En Belgique, on compte 3,7 fois plus de jeunes hommes sachant programmer que de jeunes femmes, un ratio supérieur à celui des pays voisins (Allemagne : 2,7 ; Pays-Bas : 2,6 ; France : 2).

- Dans le secteur de l'intelligence artificielle, la Belgique avait en 2018-2019 seulement 33 % de femmes parmi les travailleurs en IA, contre 53 % de femmes parmi les diplômés de niveau équivalent. La présence féminine dans ce secteur variait au sein de l'OCDE, de 46 % en Lettonie à 30 % aux Pays-Bas.
- L'écart de genre en programmation et dans le secteur l'IA en Belgique souligne l'importance de mettre en place des initiatives fortes pour encourager les jeunes femmes à s'engager dans ces domaines.

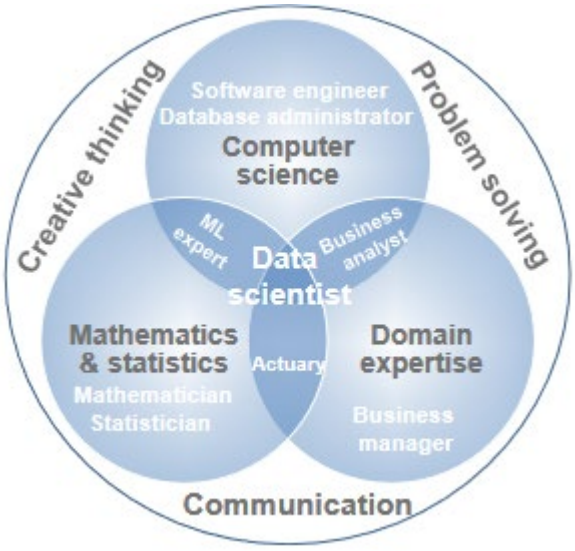
3.1. Les compétences en data

Les nombreux débats sur la définition des compétences nécessaires aux spécialistes des données mettent en évidence des exigences en constante évolution dans ce domaine, une tendance était déjà soulignée par une étude de l'OCDE de 2015[27].

L'étude soulignait l'importance croissante de la polyvalence des compétences pour gérer le cycle complet de la donnée, de la collecte à l'analyse, jusqu'à son utilisation stratégique. Cette polyvalence est essentielle pour tirer pleinement parti des flux de données et de l'automatisation avancée, lesquels sont rendus possibles par les progrès technologiques et l'interconnexion croissante des systèmes informatiques.

Comme le notait l'étude, les spécialistes des données doivent posséder une palette étendue de compétences techniques, comprenant non seulement le génie logiciel, la gestion de bases de données et l'apprentissage automatique (ML), mais également des connaissances approfondies en statistiques et en expertise sectorielle. À ces compétences techniques s'ajoutent des « soft skills » comme la communication, la créativité et la capacité à résoudre des problèmes complexes.

Figure 22. Le mix de compétences et aptitudes du spécialiste des données



Source : OECD (2015)[28].

Cette étude positionnait déjà le Data scientist à l'intersection de plusieurs disciplines : l'informatique, les mathématiques et les statistiques et un certain domaine d'expertise (Figure 22).

[27] OCDE (2015), « Skills and employment in a data-driven economy », dans Data-Driven Innovation : Big Data for Growth and Well-Being, Éditions OCDE, Paris, https://doi.org/10.1787/9789264229358-10-en.

[28] OCDE (2015), « Skills and employment in a data-driven economy », dans Data-Driven Innovation : Big Data for Growth and Well-Being, Éditions OCDE, Paris, https://doi.org/10.1787/9789264229358-10-en.

Parmi les profils cités, tels que l'ingénieur logiciel, l'administrateur de bases de données, l'actuaire, le statisticien, l'expert en *machine learning* et le *Business analyst*, une myriade de nouvelles spécialisations ont émergé ces dernières années dans les intersections de ces disciplines, avec une évolution qui tend vers une spécialisation accrue.

Les ***Data Engineers*** et ***Data Architects*** jouent un rôle crucial dans la conception, le déploiement et la maintenance des infrastructures nécessaires au stockage et au traitement des données. Ils travaillent en étroite collaboration pour assurer que les données soient non seulement stockées de manière sécurisée mais aussi facilement accessibles pour analyse. Leurs efforts sont complétés par des spécialistes tels que les ***Architectes Cloud*** **et** ***Cloud Devops***, qui développent des solutions cloud pour optimiser l'élasticité et la scalabilité des données.

Les ***Data Analysts*** transforment les données brutes en *reporting*, aidant les organisations à la prise de décision. Ces profils peuvent se spécialiser davantage dans les domaines comme la BI (Business Intelligence) où des ***BI Analysts*** **et** ***BI Developers*** utilisent des données ou développent des solutions digitales, des tableaux de bords interactifs composés de KPI, pour informer la stratégie et la planification. Enfin, les métiers liés à la visualisation de données comme les ***Data Visualization Specialists*** **ou encore les** ***Data Designer*** créent des représentations visuelles des données pour faciliter leur compréhension.

Avec les changements constants du paysage technologique et commercial, des rôles spécialisés voient le jour, tels que le MLOps Engineer, qui combine l'apprentissage automatique avec les pratiques de développement continu de DevOps, facilitant le passage des algorithmes aux systèmes de production[29].

Avec la diversification des fonctions et la valorisation des données comme actif stratégique, d'autres rôles prennent de l'importance pour répondre aux enjeux de gestion et stratégiques. La gouvernance des données, incarnée par des profils tels que le *data manager* ou le *Chief Data Officer (CDO),* devient cruciale pour orienter et piloter la stratégie de gestion des données dans les organisations. Le rôle plus récent de *data steward*[30] souligne l'importance de la qualité et de l'organisation des données. Enfin, des postes comme celui de *Data Protection Officer* deviennent cruciaux pour assurer la conformité des pratiques d'une organisation avec des réglementations telles que le RGPD.

3.2. Les compétences nécessaires à l'ère de l'IA

Si le traitement de données et la création de modèles de machine learning et de deep learning nécessitent des compétences techniques pointues, ce n'est pas nécessairement le cas lorsqu'il s'agit d'intégrer l'IA dans l'entreprise.

En effet, l'intégration de technologies d'intelligence artificielle dans les entreprises ne semble pas toujours nécessiter des changements majeurs dans les compétences requises. Selon une étude de l'OCDE menée par Lane, Williams et Broecke (2023), près de la moitié des employeurs ayant intégré l'IA dans les secteurs de la finance et de la fabrication n'ont pas constaté de changements notables dans leurs besoins en compétences.

Cette stabilité peut s'expliquer par le fait que l'impact de l'IA sur les tâches des travailleurs reste limité, que son adoption réorganise parfois les tâches existantes plutôt que d'en créer de nouvelles, et que les compétences numériques de base nécessaires sont souvent déjà présentes chez les employés. De plus, les entreprises qui adoptent l'IA sont généralement celles dont la main-d'œuvre possède déjà les compétences requises, ce qui explique pourquoi elles ne signalent pas de changements significatifs dans leurs besoins en compétences.

---

[29] La MLOps est composé du terme machine learning, ou et de la pratique de développement continu de devops dans le domaine des logiciels. Les modèles d'apprentissage automatique sont testés et développés dans des systèmes expérimentaux isolés. Lorsqu'un algorithme est prêt à être lancé, le MLOps est pratiqué entre les Data Scientist, les Devops et les Data Engineers en apprentissage automatique pour faire passer l'algorithme aux systèmes de production.

[30] Responsables de la gestion de la qualité des données, assurant que les données à travers l'organisation sont exactes, disponibles et utilisées de manière responsable.

Si dans de nombreuses entreprises, l'introduction de l'intelligence artificielle ne semble pas entraîner de changements significatifs dans les compétences requises, dans une part significative des entreprises, la mise en œuvre de l'IA est associée à un besoin de compétences plus élevées et plus larges, entrainant une demande accrue en compétences numériques, analytiques et en *soft skills* (Milanez, 2023[4]).

À l'ère de l'intelligence artificielle, il est donc crucial de distinguer les besoins en compétences nécessaires au développement et à la maintenance de systèmes d'IA, des compétences nécessaires à l'interaction avec les applications IA.

Pour développer et entretenir les systèmes d'IA, des compétences spécialisées en IA, en science des données, ainsi que des compétences cognitives et transversales sont nécessaires. L'acquisition de ces compétences spécialisées demande une combinaison d'une formation académique avancée et d'une expérience pratique significative. L'éducation initiale joue un rôle crucial, avec un nombre considérable de professionnels de l'IA possédant un diplôme universitaire. En outre, l'apprentissage sur le terrain, sous forme de stages ou d'apprentissage informel au sein d'équipes de recherche ou de processus de développement d'IA dans les entreprises, revêt une importance capitale.

Quant à l'utilisation et à l'interaction avec les applications d'IA, elles requièrent une connaissance de base de l'IA et des compétences numériques de base, ainsi que des compétences cognitives et transversales. Cette connaissance de base de l'IA, souvent désignée sous le terme de littératie en IA, est un concept récent qui suscite un intérêt croissant dans la littérature sur l'apprentissage des adultes. Long et Magerko (2020)[31] la définissent comme un ensemble de compétences permettant aux individus d'évaluer de manière critique les technologies d'IA, de collaborer efficacement avec elles et de les utiliser comme outils dans divers contextes. Ng et al. (2021)[32] insistent sur le fait que cette alphabétisation ne se limite pas aux compétences requises pour développer des modèles d'IA, mais englobe également la compréhension, l'utilisation, le suivi et la réflexion critique sur les applications d'IA.

Les compétences numériques de base et une connaissance élémentaire de l'IA sont souvent nécessaires pour utiliser les applications d'IA, tandis que les compétences analytiques et les *soft skills* deviennent plus importantes en raison de l'automatisation des tâches simples, qui libère du temps pour des tâches plus complexes et interactives. Des constatations similaires sont rapportées par Lane, Williams et Broecke (2023) : l'IA accroît principalement l'importance de compétences telles que la créativité et la communication au sein de l'entreprise, ainsi que le besoin de travailleurs hautement qualifiés.

[31] Long, D. and B. Magerko (2020), « What is AI Literacy? Competencies and Design Considerations », Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

[32] Ng, D. et al. (2021), « AI Literacy: Definition, Teaching, Evaluation and Ethical Issues », Proceedings of the Association for Information Science and Technology, Vol. 58/1, pp. 504-509, https://doi.org/10.1002/pra2.487.

Tableau sur les compétences nécessaires à l'ère de l'intelligence artificielle

Table 5.1. **Skill needs in the age of AI**

	Type of skill	Examples
Skills to develop and maintain AI systems	Specialised AI skills	General knowledge of AI (such as Machine Learning) Specific knowledge of AI models ("decision trees", "deep learning", "neural network", "random forest", etc), AI tools ("tensorflow", "pytorch", etc) and AI software ("java", "gradle", "galaxy cluster", etc).
	Data science skills	Data analysis Software Programming languages, in particular Python Big data Data visualisation Cloud computing
	Other cognitive skills	Creative problem solving
	Transversal skills	Social skills Management skills
Skills to adopt, use and interact with AI applications	Elementary AI knowledge	Principles of machine learning
	Digital skills	Ability to use a computer or a smartphone
	Other cognitive skills	Analytical skills Problem-solving Critical thinking Judgement
	Transversal skills	Creativity Communication Teamwork Multitasking

Source: Authors' elaborations on Alekseeva et al. (2021[12]), "The demand for AI skills in the labor market", https://www.doi.org/10.1016/j.labeco.2021.102002, Lane, Williams and Broecke (2023[2]), "The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers", https://www.doi.org/10.1787/ea0a0fe1-en; Manca (2023[13]), "Six questions about the demand for artificial intelligence skills in labour markets", https://www.doi.org/10.1787/ac1bebf0-en, Milanez (2023[4]), "The impact of AI on the workplace: Evidence from OECD case studies of AI implementation", https://www.doi.org/10.1787/2247ce58-en; OECD.AI (2023[14]), *Visualisations powered by JSI using data from www.adzuna.co.uk*, https://oecd.ai/; Squicciarini and Nachtigall (2021[15]), "Demand for AI skills in jobs: Evidence from online job postings", https://www.doi.org/10.1787/3ed32d94-en.

Source : OECD Employment Outlook 2023 : Artificial Intelligence and the Labour Market | OECD iLibrary (oecd-ilibrary.org).

3.3. Le marché du travail dans le secteur des données

Dans cette section, notre objectif est d'analyser le marché du travail belge dans le secteur des données, en nous appuyant sur les données fournies par l'étude d'IDC (2024). Cette étude adopte une définition précise des professionnels de la donnée, considérés comme les travailleurs qui collectent, stockent, gèrent et/ou analysent, interprètent et visualisent des données comme activité principale ou partie intégrante de leur activité pertinente. Les professionnels des données doivent être compétents dans l'utilisation de données structurées et non structurées, être capables de travailler avec une énorme quantité de données et être familiers avec les technologies de bases de données émergentes.

IDC distingue trois rôles joués par les utilisateurs de données : les professionnels techniques des données, les professionnels métier des données et les consommateurs de données.

- Les professionnels techniques des données sont des spécialistes de la collecte, du stockage, de la gestion, de la modélisation et de l'assurance qualité des données, ainsi que de l'intégration de différentes sources de données pour assurer la cohérence, l'exactitude et la qualité des données. Un professionnel technique des données peut, étant donné la question qui doit être répondue, s'assurer que la chaîne d'approvisionnement en données est fournie et qu'elle est exacte.
- Les professionnels métier des données ont pour tâche principale ou significative l'exécution d'analyses prédictives, d'analyses qualitatives, de modélisation de données, d'extraction de données et de synthèses de données dans le but de créer de nouvelles idées et connaissances à partir des données disponibles. Ils possèdent une compréhension approfondie de l'industrie et/ou du processus et peuvent mettre l'analyse de données dans son contexte et la relier aux tendances existantes dans l'industrie ou la ligne de métier dans laquelle ils se trouvent. Ils confient généralement la collecte, la gestion et la qualité des données au professionnel technique des données, mais, en utilisant des outils d'analyse tels que Excel, Tableau et Power BI, ils peuvent résumer de grandes quantités de données et visualiser et présenter des tendances et des idées à un public plus large de parties prenantes clés de l'entreprise afin de conduire le processus décisionnel stratégique au sein de l'organisation. Les scientifiques des données (*Data Scientist*) résident dans cette catégorie.
- Les consommateurs de données sont des employés et des gestionnaires de produits, de processus, de ressources humaines, d'actifs ou de départements responsables de la conduite du changement ou du maintien d'une position où la prise de décision dépend fortement de l'approvisionnement en données et en idées basées sur de grandes quantités de données. Ils travaillent directement avec les données seulement une partie du temps. Ce sont des décideurs ou des parties prenantes dans un processus de décision où les données et les idées fournies déterminent la qualité des décisions prises. Un consommateur de données guide l'entreprise en fonction des données et des idées fournies par la chaîne d'approvisionnement en données.

Les estimations fournies[33] dans cette section sur les professionnels des données ne concernent que les deux premières catégories, c'est-à-dire les professionnels techniques des données et les professionnels métier des données.

L'évolution de l'emploi en Belgique

Le tableau ci-dessous représente la part de l'emploi des professionnels des données en pourcentage de l'emploi total pour les années 2020 à 2023 dans plusieurs pays, y compris la Belgique, ainsi que la moyenne pour les 27 pays de l'Union européenne (UE-27). Ce tableau révèle deux éléments cruciaux : la position défavorable de la Belgique par rapport à d'autres pays et l'évolution stagnante de cette proportion d'emploi au cours de la période.

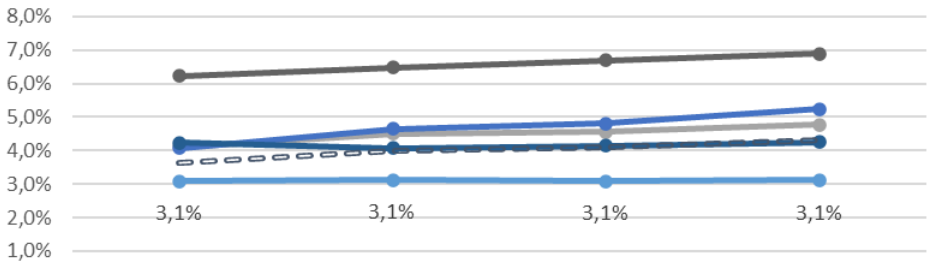
[33] Les dernières estimations réalisées par IDC se basent sur les sources suivantes :

- Eurostat Business Demographic Statistics Sep 2023
- Eurostat Annual Structural Business Statistics Sep 2023
- IDC Worldwide Black Book (standard edition) Sep 2023
- ILOSTAT statistics and databases Oct 2023

En effet, en comparaison avec ses pays voisins, la Belgique affiche la plus faible proportion d'employés travaillant dans le domaine des données. En 2023, la Belgique comptait 140.790 professionnels des données, représentant 3,1 % de l'ensemble de sa main-d'œuvre. Cela place la Belgique en dessous de la moyenne de l'UE-27. Plus précisément, elle se situe à la 25e place parmi les 27 États membres, avec seulement la Roumanie (3 %) et la Grèce (2,1 %) affichant des proportions inférieures.

Figure 23. Part de l'emploi des professionnels de la donnée

En pourcentage de l'emploi total.



	2020	2021	2022	2023
Belgique	3,1%	3,1%	3,1%	3,1%
France	4,1%	4,5%	4,6%	4,8%
Allemagne	4,1%	4,7%	4,8%	5,2%
Pays-Bas	4,2%	4,1%	4,1%	4,2%
Royaume-Uni	6,2%	6,5%	6,7%	6,9%
EU27	3,6%	4,0%	4,1%	4,3%

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

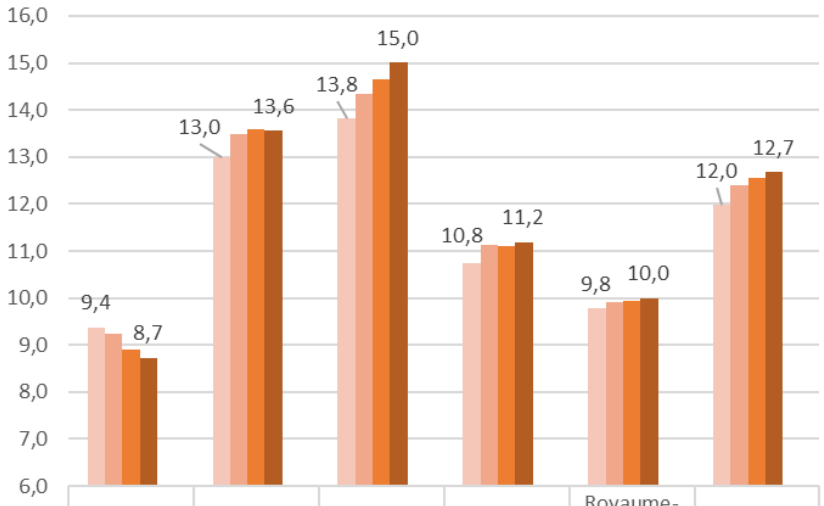
En termes d'évolution de la part d'emploi des professionnels des données, la Belgique n'enregistre pas de variation marquée sur la période d'observation. En revanche, les pays voisins, la France, l'Allemagne, les Pays-Bas et le Royaume-Uni ont tous augmenté leur part d'emploi des professionnels des données au fil du temps. La moyenne de l'UE-27 est également en augmentation, passant de 3,6 % en 2020 à 4,3 % en 2023. Cette évolution montre que la Belgique n'a pas suivi la tendance à la hausse observée dans d'autres pays européens. En outre, la Belgique se positionne en dernière place de tous les pays de l'UE des 27 en termes de croissance du nombre de professionnels de la donnée au sein de l'Union européenne entre 2020 et 2023 (+3,9 % contre +17,8 % pour l'UE-27).

Une autre mesure importante consiste à comparer le nombre de professionnels spécialisés dans les données au nombre d'entreprises utilisatrices de données. On dérive ainsi une mesure d'intensité des professionnels des données dans les entreprises concernées.

L'évolution de cet indicateur montre des tendances contrastées au sein de l'Union européenne (Figure 24). La Belgique enregistre l'évolution la plus défavorable parmi tous les pays de l'Union européenne dans la période observée. Entre 2020 et 2023, cet indicateur a diminué de 9,4 à 8,7, représentant une baisse de 6,8 %.

En revanche, dans les pays voisins, ainsi que dans tous les autres pays de l'Union européenne (hormis l'Autriche et la Pologne), cet indicateur est en augmentation au cours de la même période. En France, par exemple, il est passé de 13,0 à 13,5, soit une augmentation de 4 %. En Allemagne, cette augmentation est encore plus marquée, avec une croissance de 9,0 %. Pour l'ensemble de l'UE-27, l'indicateur est passé de 12,0 à 12,7, ce qui représente une augmentation de 6 %.

Figure 24. Évolution du nombre de professionnels de la donnée par entreprise utilisatrice



	Belgique	France	Allemagne	Pays-Bas	Royaume-Uni	EU27
2020	9,4	13,0	13,8	10,8	9,8	12,0
2021	9,2	13,5	14,3	11,1	9,9	12,4
2022	8,9	13,6	14,6	11,1	9,9	12,5
2023	8,7	13,6	15,0	11,2	10,0	12,7
Taux de croissance 2023/2020	-6,8%	4,3%	8,6%	4,1%	2,3%	5,8%

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Dans l'ensemble, l'Union européenne a donc augmenté l'intensité d'utilisation de professionnels des données. La Belgique, en revanche, semble prendre du retard par rapport à ses voisins en termes d'emploi de professionnels des données et enregistre une croissance négative. Cette situation pourrait refléter une réduction de la capacité d'analyse et de gestion des données au sein des entreprises belges, ce qui est inquiétant dans la mesure où le nombre d'entreprises utilisatrices des données devrait connaitre une accélération dans les années à venir. Si ces réductions ne sont pas compensées par des gains de productivité, cela pourrait avoir des conséquences négatives sur la compétitivité des entreprises et leur capacité à innover.

Écart entre l'offre et la demande sur le marché du travail

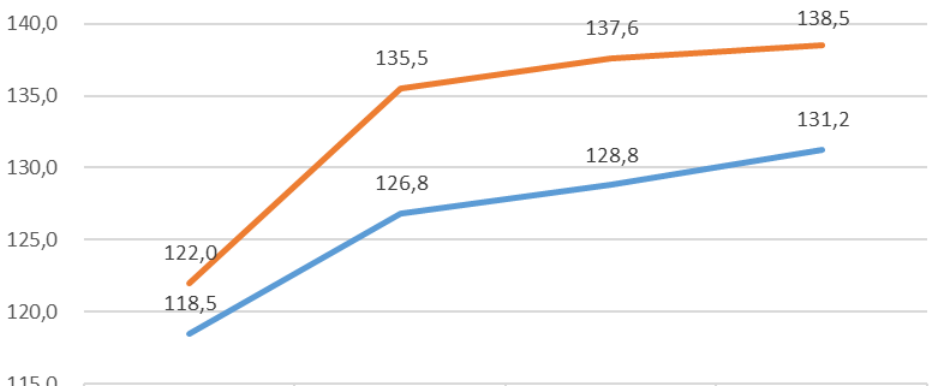
Les données présentées dans le graphique ci-dessous illustrent l'évolution de l'offre et de la demande de professionnels des données de 2019 à 2022, ainsi que l'écart entre ces deux mesures, exprimé en milliers d'unités.

On peut y observer tout d'abord que le nombre de professionnels des données ne cesse de croître. Entre 2019 et 2022, à la fois l'offre et la demande en termes de professionnels des données a connu une forte augmentation. L'offre est passée de 118.500 en 2019 à plus de 131.000 en 2022, tandis que la demande est passée de 122.000 à 138.500 pendant la même période.

L'écart entre ces deux indicateurs s'est d'abord fortement creusé entre 2019 et 2020, pour se maintenir l'année suivante autour des 6,4 %. Toutefois, en 2022, l'écart observé en Belgique diminue et se rapproche des niveaux constatés chez ses homologues européens (à l'exception de l'Allemagne qui affiche un écart plus élevé -6,7 %). En 2022, les entreprises belges, tant fournisseuses qu'utilisatrices de données avaient besoin d'environ 7.000 travailleurs de données pour combler le déficit de compétences en données.

Figure 25. Évolution de l'offre et de la demande de professionnels de la donnée en Belgique

En milliers d'individus.



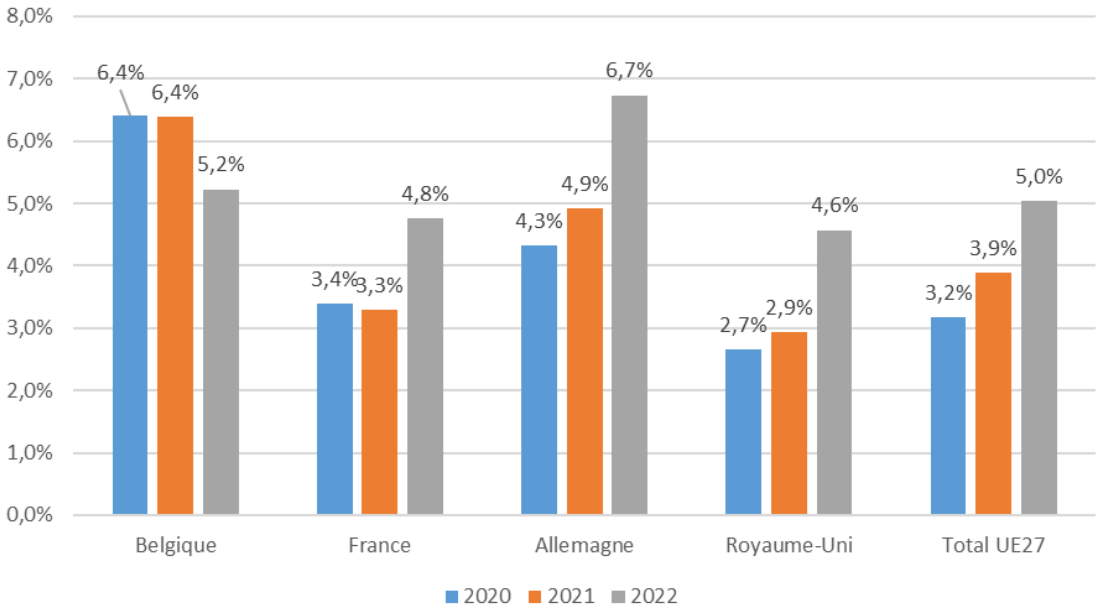
	2019	2020	2021	2022
Offre	118,5	126,8	128,8	131,2
Demande	122,0	135,5	137,6	138,5
Déficit de compétences	2,9%	6,4%	6,4%	5,2%

Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Bien que la diminution de l'écart en déficit de compétences observée en 2022 puisse être considérée comme une évolution positive, il est crucial d'examiner attentivement les raisons sous-jacentes à cette diminution. En comparaison avec les pays voisins et la moyenne de l'UE, qui ont tous connu une augmentation de l'écart, la baisse observée en Belgique en 2022 s'explique principalement par une croissance annuelle relativement faible de la demande des entreprises en professionnels de la donnée (+0,7 %), tandis que cette croissance est significativement plus élevée dans les pays voisins, variant entre 3,9 % et 6,6 %, et dans l'ensemble de l'UE (5 %).

Figure 26. Écart entre la demande et l'offre de professionnels de la donnée

En % de la demande totale.



Source : Élaboration d'Idea Consult sur base de données d'IDC (2024).

Cette différence notable dans la croissance de la demande souligne la nécessité d'explorer plus profondément les dynamiques du marché du travail spécifiques à la Belgique afin d'identifier des possibles freins à la demande en professionnels de la donnée.

3.4. Le gender gap

Cet article examine deux indicateurs clés pour rendre compte des inégalités de genre dans les domaines de la data : la proportion de jeunes femmes sachant programmer et la représentation des femmes dans le secteur de l'IA.

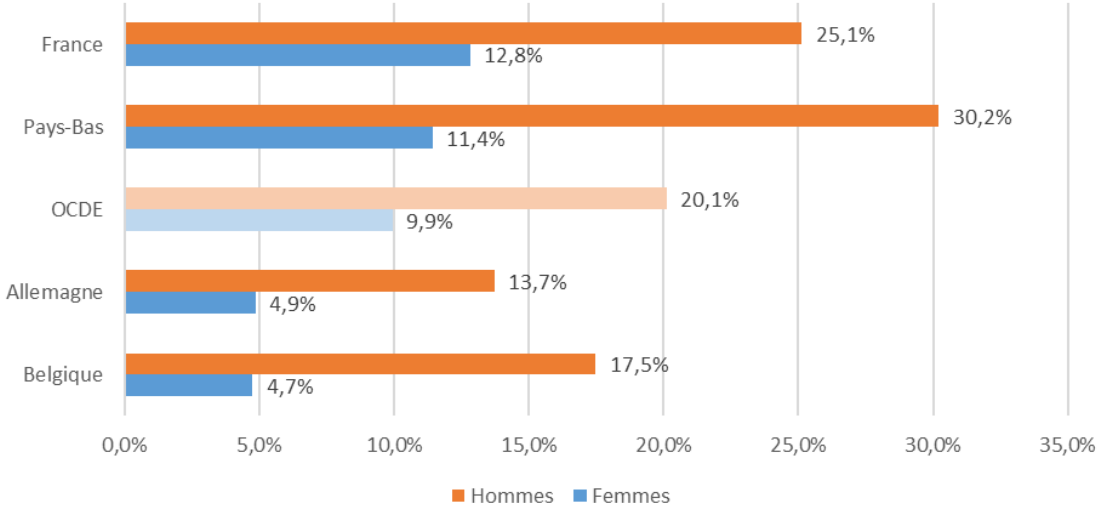
Sur le premier indicateur, le premier constat est qu'en 2023, dans tous les pays de l'OCDE sauf la Grèce, la proportion de jeunes hommes sachant programmer est supérieure à celle des femmes. Le Belgique ne fait pas exception et se situe même parmi les pays ayant les plus faibles pourcentages de jeunes femmes (âgées de 16 à 24 ans) sachant programmer, avec seulement 4,7 % en 2023 (voir Figure 27), bien en dessous de la moyenne de l'OCDE de 9,9 %. Seules la Turquie et l'Irlande enregistrent des pourcentages inférieurs.

Les Pays-Bas et la France présentent en revanche des pourcentages beaucoup plus élevés, avec respectivement 11,4 % et 12,8 % des jeunes femmes sachant programmer. L'Allemagne, avec un pourcentage de 4,9 %, est à peine au-dessus de la Belgique. Toutefois, en raison d'un pourcentage d'hommes sachant programmer significativement plus faible qu'en Belgique (13,7 %), son écart de genre s'en trouve réduit.

Ainsi, la Belgique compte 3,7 fois plus de jeunes hommes sachant programmer que de jeunes femmes, tandis que ce ratio est inférieur chez les pays voisins (2,7 pour l'Allemagne , 2,6 pour les Pays-Bas, 2 pour la France). Si l'OCDE obtient en moyenne un indicateur de 2, c'est notamment en raison de toute une série de pays qui ont réussi à réduire significativement leur écart de genre, comme le Chili, la Corée du Sud, le Mexique, la Colombie et la Grèce.

Figure 27. Écart de genre en termes des jeunes sachant programmer (2023)

% des jeunes de 16 à 24 ans.



Source : OECD Going Digital Toolkit : https://goingdigital.oecd.org/fr/indicator/54.

Dans le secteur de l'intelligence artificielle, la présence masculine est également plus élevée que celle des femmes, comparativement à la population active diplômée de l'enseignement supérieur. Selon une étude de l'OCDE[34], en 2018-2019, seulement 33 % des travailleurs en intelligence artificielle en

[34] Green, A. et L. Lamby (2023), « The supply, demand and characteristics of the AI workforce across OECD countries », Documents de travail de l'OCDE sur les questions sociales, l'emploi et les migrations, n° 287,

Belgique étaient des femmes, contre 53 % parmi les diplômés de niveau équivalent. À l'échelle de l'OCDE, la représentation féminine dans ce secteur variait de 46 % en Lettonie à 30 % aux Pays-Bas. Aucun pays ne présentait une majorité féminine dans la main-d'œuvre en IA, et aucune proportion équivalente à celle de la population active diplômée n'était observée[35] (OCDE, p.30).

L'écart de genre observé dans les compétences en programmation chez les jeunes et dans le secteur de l'IA est présent dans la plupart des pays, mais il semble particulièrement prononcé en Belgique. Cette situation appelle à mettre en place des initiatives volontaristes pour encourager les jeunes femmes à s'engager dans des études et des carrières dans ces domaines. Des programmes éducatifs ciblés, des campagnes de sensibilisation pour déconstruire les stéréotypes de genre, et des systèmes de soutien renforcés pourraient aider à réduire ces disparités.

Éditions OCDE, Paris, https://doi.org/10.1787/bb17314a-en. Disponible ici : https://www.oecd-ilibrary.org/docserver/bb17314a-en.pdf?expires=1708459436&id=id&accname=guest&checksum=58411FCC2C99503A921A401F99EABBA3.

[35] Ibid, p.30.

4. Culture de la donnée

Principaux résultats

CULTURE ET PARTAGE DE DONNEES

Open Data pour les statistiques publiques

- La Belgique se classe 72e en matière d'Open Data sur 194 pays et accuse du retard par rapport aux pays voisins.
- Ce score se compose d'un score de couverture de 53 % et d'ouverture de 61 %, révélant un potentiel d'amélioration significatif.
- Le score est particulièrement bas (34) dans les **statistiques sociales** avec des lacunes dans les domaines de l'éducation, la santé, et la justice marqués par l'absence de données cruciales telles que le nombre d'établissements scolaires, le taux de vaccination, et la prévalence de la sous-nutrition.
- En contraste, le domaine des **statistiques économiques** présente une meilleure performance pour la Belgique, avec un score de 84. Cependant, le marché du travail montre des résultats moins satisfaisants.
- Avec un score global de 55, **l'environnement** représente un domaine d'une performance intermédiaire pour la Belgique. Malgré une couverture et ouverture moyennes, certains domaines comme la pollution et l'environnement bâti révèlent des lacunes importantes, tandis que la catégorie de l'énergie se distingue positivement.
- Sur 22 catégories de données analysées, seules 6 surpassent le score de 80 % en termes d'ouverture et aucune n'atteint 100 %, soulignant la marge de progression existante pour intégrer pleinement les standards d'Open Data dans la culture et les pratiques courantes.
- Recommandations d'ODIN pour la Belgique:
  - Identifier et combler les lacunes de données.
  - Publier plus de données désagrégées par sexe.
  - Utiliser des formats de données non propriétaires.
  - Améliorer les métadonnées.
  - Fournir plus de données historiques et récentes.
  - Offrir des données plus détaillées au niveau sous-national.

Ouverture de données au sein des entreprises

- Le partage de données non personnelles par les entreprises reste peu courant, tant au niveau national qu'européen.
- Peu d'entreprises partagent des données avec des ONG (12 %), des entités commerciales (12 %), ou des clients (22 %).
- Le partage avec des institutions publiques ou gouvernementales est un peu plus fréquent (27 %).

Préoccupations citoyennes (sondage sur la population flamande)

- Les citoyens expriment clairement leurs préoccupations concernant la protection de leur vie privée en ligne et l'usage qui est fait de leurs données en ligne.
  - 89 % perçoivent internet comme une menace pour leur vie privée, une inquiétude en hausse.
  - 79 % sont frustrés par l'opacité des entreprises quant à l'usage des données personnelles.
  - 66 % se déclarent incertains quant à l'utilisation qui est faite de leurs données
  - Près de 60 % se sentent impuissants à protéger leurs données personnelles.
  - 98 % prennent des mesures pour sécuriser leur vie privée, comme limiter l'accès à leurs données et utiliser des outils de protection.
- Une atmosphère de scepticisme règne et les citoyens sont prudents dans le partage de leurs données. **Seuls 6 % de la population se disent déjà à l'aise avec l'idée de partager leurs données.**
- **Cette situation** met en lumière le besoin urgent de solutions tangibles et de mesures protectrices. 96 % soutiennent des sanctions contre les abus de gestion des données personnelles de la part des entités qui les gèrent.
- Un intérêt pour le partage des données reste présent, mais est influencé par plusieurs facteurs clés :
  - Préférence pour partager des données en échange d'avantages tangibles tels que des économies de temps, l'accès à des services gratuits, ou des recommandations sur mesure, bien que l'intérêt diminue avec l'âge.
  - Plus grande disposition à partager des informations non sensibles (âge, sexe, opinions) que des données sensibles (santé, finances, localisation).
- Pour augmenter la confiance dans le partage de données, des besoins en termes de **transparence et de sécurité** sont exprimés. 63 % des sondés souhaitent une plus grande clarté sur l'usage prévu de leurs données par les organisations, tandis que la moitié attend des instructions détaillées pour un partage sécurisé.
- Toutefois, 15 % estiment qu'aucune mesure ne les rendrait plus confortables avec le partage de leurs données personnelles.
- La confiance des citoyens dans le traitement de leurs données personnelles varie grandement d'une entité à l'autre. Les hôpitaux et fournisseurs de soins de santé bénéficient de la plus grande confiance (>80 %), tandis que la confiance pour les télécommunications (39 %), les boutiques en ligne (30 %), et surtout les médias sociaux et les start-ups (5-8 %) est nettement plus basse.
- L'année 2023 a enregistré une baisse de cette confiance auprès de presque toutes les entités considérées, à l'exception des hôpitaux et fournisseurs de soins de santé.
- La majorité est plus encline à partager des données avec des entreprises jouissant d'une bonne réputation ou avec lesquelles ils ont eu par le passé des expériences positives.

4.1. L'ouverture des statistiques publiques

L'Inventaire des Données Ouvertes (ODIN) 2022/2023 offre une évaluation basée sur 65 indicateurs représentatifs répartis dans 22 catégories de données, visant à fournir un échantillon indicatif des types fondamentaux de données produites par les pays.

L'ODIN examine l'offre statistique d'un pays sous deux aspects principaux : la complétude de cette offre et sa conformité aux normes internationales d'ouverture, aboutissant à l'attribution de deux scores distincts, soit le score de couverture et le score d'ouverture.

Encadré 6. Que prend en compte le score d'ouverture d'ODIN ?

Les scores d'ouverture sont calculés sur base de cinq éléments d'ouverture :

Lisibilité par machine

Cet élément d'ouverture mesure si les données sont disponibles dans des formats lisibles par machine. Les formats de fichiers lisibles par machine permettent aux utilisateurs de traiter facilement les données à l'aide d'un ordinateur. Les formats communs lisibles par machine incluent les fichiers XLS, XLSX, CSV et JSON.

Format non propriétaire

Cet élément d'ouverture mesure si les données sont disponibles dans des formats non propriétaires. Les formats de fichiers non propriétaires sont importants car ils permettent aux utilisateurs d'accéder aux données sans nécessiter l'utilisation d'un logiciel propriétaire coûteux qui pourrait empêcher certains utilisateurs d'accéder aux données. Les formats non propriétaires courants incluent les fichiers PDF, HTML, XLSX, DOCX, CSV et JSON.

Disponibilité des métadonnées

Cet élément d'ouverture mesure si les métadonnées sont disponibles pour les indicateurs publiés. Les métadonnées doivent se situer dans ou à proximité du fichier de données ou dans une section des métadonnées désignée sur le site web. ODIN recherche trois aspects des métadonnées : (1) la définition de l'indicateur ; (2) la date de mise en ligne ; et (3) l'agence source.

Options de téléchargement

Cet élément d'ouverture mesure si des options de téléchargement sont disponibles. ODIN recherche trois options de téléchargement : (1) téléchargement en masse (au niveau de l'indicateur), (2) API, et (3) téléchargement sélectif par l'utilisateur (téléchargements personnalisés). Les options 2 et 3 sont interchangeables aux fins de l'évaluation.

Licence des données/Conditions d'utilisation

Cet élément d'ouverture mesure si les données sont mises à disposition sous une licence ouverte. Les licences ouvertes doivent permettre l'utilisation, la réutilisation et le partage ou l'adaptation des données pour un usage commercial et non commercial sans aucune obligation autre que l'attribution, selon la Définition de l'Open. Les licences interdisant l'utilisation commerciale ou ayant deux stipulations supplémentaires ou plus sont classées comme « Non Ouvertes ». Les licences avec pas plus d'une stipulation supplémentaire sont classées comme « Quelques Restrictions ». Les licences qui n'explicitent pas clairement toutes les utilisations autorisées selon la Définition de l'Open et qui n'incluent pas de langage restrictif sont également classées comme « Quelques Restrictions ».

Source : https://odin.opendatawatch.com/.

En 2022, la Belgique occupe la 72e place avec un score global de 57, combinant un score de couverture de 53 et d'ouverture de 61. Ce résultat, inférieur à celui de ses voisins, souligne un fort potentiel d'amélioration dans l'accès et la disponibilité des données.

Tableau 6. Classement de la Belgique en termes de couverture et ouverture de statistiques publiques

Pays	Score Global	Score de Couverture	Score d'ouverture	Statistiques Sociales	Statistiques Économiques	Statistiques Environnementales
Belgique	57	53	61	34	84	55
France	68	59	77	63	81	62
Allemagne	79	72	85	70	88	80
Pays-Bas	80	67	91	77	94	68
Royaume-Uni	63	53	71	62	66	60

Source : https://odin.opendatawatch.com/Report/countryProfileUpdated/BEL?year=2022.

Les catégories de données analysées peuvent être divisées en trois groupes principaux : Statistiques Sociales, Statistiques Économiques et Environnement.

D'après ces indicateurs, la Belgique présente de fortes carences dans les Statistiques Sociales où elle n'obtient qu'un score de 34. Ces carences concernent surtout les secteurs de l'éducation, de la santé, et de la justice, avec l'absence de données clés comme le nombre d'établissements scolaires, le taux de vaccination, et la prévalence de la sous-nutrition.

Concernant les Statistiques Économiques, la Belgique se distingue par une meilleure performance (84). Cette catégorie témoigne d'une couverture et d'une ouverture relativement bonnes des données, particulièrement dans les domaines des comptes nationaux, de la finance gouvernementale, du commerce international et de la balance des paiements, où la Belgique atteint des scores élevés. Toutefois, la catégorie « marché du travail » montre des performances moins bonnes.

Dans le domaine de l'Environnement, la Belgique affiche un score intermédiaire, avec un score global de 55 pour l'Environnement, qui combine une couverture de 50 et une ouverture de 60. Les catégories telles que l'usage des terres agricoles et l'utilisation des ressources montrent des scores modérés, tandis que la catégorie de l'énergie se distingue par un score plus élevé, reflétant une meilleure couverture et ouverture. Cependant, des domaines critiques comme la pollution et l'environnement bâti révèlent des lacunes significatives, avec ce dernier enregistrant un score particulièrement bas.

Pour adresser ces défis et améliorer sa position globale en matière d'Open Data, plusieurs recommandations sont formulées par l'organisation à l'attention de la Belgique :

- Combler les lacunes de données : identifier et prioriser la collecte dans les domaines où les données sont insuffisantes ou manquantes.
- Publier davantage de données désagrégées par sexe : afin d'offrir une meilleure analyse et compréhension des disparités et inégalités de genre.
- Publier plus de données dans des formats non-propriétaires : faciliter l'accès et l'utilisation des données en choisissant des formats ouverts et facilement exploitables.
- Publier des métadonnées plus complètes : améliorer la documentation accompagnant les données pour en faciliter la compréhension, l'utilisation et la réutilisation.
- Publier davantage de données historiques : fournir un contexte et permettre des analyses temporelles en rendant disponibles des séries de données sur de longues périodes.
- Publier plus de données récentes : garantir la pertinence et l'utilité des données en assurant leur actualité.
- Publier plus de données au niveau sous-national : offrir une granularité plus fine des données pour permettre des analyses régionales et locales approfondies.

La comparaison avec les pays voisins révèle un potentiel d'amélioration notable pour aligner la Belgique sur les meilleures pratiques d'Open Data. Par ailleurs, parmi les 22 catégories de données analysées, il est intéressant de noter que seules 6 d'entre elles dépassent le score de 80 % et qu'aucune n'atteint un taux d'ouverture de 100 %, ce qui met en évidence le fait que l'adoption de

l'Open Data reste une pratique encore marginale, dont les standards ne sont pas fortement intégrés dans la culture actuelle.

4.2. Partage de données par les entreprises

Dans l'ensemble, le partage de données non personnelles avec des tiers n'est pas une pratique courante parmi les entreprises qui stockent des données en Belgique, tout comme au sein de l'UE dans son ensemble. Elles sont peu nombreuses à déclarer fournir des données à des organisations à but non lucratif (12 % contre 9 % dans l'UE), à des entités commerciales (12 % contre 12 % dans l'UE) ou à des individus tels que des clients (22 % contre 18 % dans l'UE). Bien qu'une majorité n'échange pas de données avec des institutions publiques ou gouvernementales, près d'un tiers (27 % contre 31 % dans l'UE) déclarent le faire, ce qui en fait la catégorie la plus courante de tiers avec lesquels les entreprises partagent leurs données[36].

4.3. Préoccupations citoyennes face au partage de leurs données

Dans un contexte où les entreprises et les organisations exploitent de manière croissante les données personnelles, la digitalisation massive accentue la récolte d'informations à une échelle sans précédent, et notamment les données à caractère personnel. Ceci se traduit par une pression considérable sur la vie privée en ligne, comme en témoignent les préoccupations profondes exprimées au sein de la population.

Selon l'enquête menée dans le cadre du SOLIDMonitor 2023[37], un constat transgénérationnel émerge quant aux dangers que représente internet pour la vie privée, avec une prise de conscience générale des risques de violations de la vie privée liées à l'internet manifestée par des pourcentages élevés à travers toutes les tranches d'âge. 89 % des Flamands sont conscientisés aux problématiques de confidentialité spécifiques au monde numérique et estiment qu'internet pose de graves problèmes de protection de la vie privée. Ce chiffre est en augmentation puisqu'il était de 83 % en 2022. Par ailleurs, 65% expriment une préoccupation quant à leur vie privée en ligne.

Cette préoccupation pour la sécurité et la confidentialité des données personnelles débouche sur un consensus très large en termes d'exigence d'une protection légale, d'une transparence accrue de la part des entreprises, et d'une régulation stricte par des entités gouvernementales. En effet, le manque de clarté de la part des entreprises concernant l'utilisation des données personnelles agace 79 % des répondants, affichant là aussi une hausse de 6 %. Un pourcentage encore plus élevé (84 %) souhaite savoir qui détient quelles données personnelles à leur sujet. 68 % des Flamands se sentent désarmés face au contrôle de leurs informations personnelles, suggérant un sentiment d'impuissance face à la gestion de leurs données (contre 72 % en 2022). Les Flamands appellent à une responsabilisation des organisations, en ligne avec les régulations européennes de protection des données. Un impressionnant 92 % estime que les entreprises doivent être transparentes sur l'exploitation des données personnelles, tandis que 93 % réclament une protection légale et un contrôle sur l'utilisation de ces données. Enfin, le soutien quasi unanime (96 %) pour des sanctions en cas d'abus de données souligne l'importance accordée à la responsabilisation des entreprises et à la protection des droits individuels dans l'espace numérique.

La méfiance et le sentiment d'impuissance prévalent, avec 66 % des sondés qui se déclarent incertains quant à l'utilisation qui est faite de leurs données par les services en ligne. Près de 60 % se sentent incapables de protéger efficacement leurs données personnelles en ligne, tandis que la moitié est convaincue qu'il y a peu à faire contre les abus de leurs données personnelles. Une attitude de résignation s'observe chez près d'un quart des Flamands, avec 25 % réticents à consacrer du temps à la protection de leurs données et 22 % convaincus de l'inutilité de leurs efforts, bien que l'étude relève une diminution des personnes ayant perdu foi en la protection de leurs données en ligne (de 26 à 23 %).

[36] Source : Survey of Businesses on the Data Economy.

[37] SOLIDMonitor : Privacy, persoonlijke data & datakluizen, editie 2 -2023.

Cependant, une grande majorité des Flamands adoptent des mesures pour protéger leur vie privée (98 %). 72 % limitent l'accès à leurs données, 70 % évitent certains téléchargements d'applications, et environ une personne sur deux s'efforce de minimiser l'acceptation de cookies ou de les supprimer régulièrement. Les logiciels de sécurité anti-pistage et les changements fréquents de mots de passe sont également courants (40 % et 44 %, respectivement), mais les outils avancés comme les gestionnaires de mots de passe (20 %) ou les VPN (21 %) restent moins utilisés.

Malgré les inquiétudes existantes concernant la confidentialité et la sécurité des données, un désir de partage de données émerge sous certaines conditions. Par exemple, pour des avantages perçus, tels que des gains de temps (26 %), l'accès à des services gratuits (23 %), ou des recommandations personnalisées (18 %). Notons que la volonté de partager des données pour ces différentes raisons diminue fortement avec l'âge (Tableau 7 : Approche personnelle du partage de données, selon l'âge), traduisant une méfiance chez les générations plus âgées.

Tableau 7. Approche personnelle du partage de données, selon l'âge

Tranche d'âge	**18-24**	**25-34**	**35-44**	**45-54**	**55-64**	**65+**	**Total 2022**	**Total**
Partager facilement vos données en échange de services ou de produits gratuits	39 %	37 %	33 %	21 %	16 %	10 %	24 %	23 %
Ne pas s'inquiéter de la manière dont vos données sont collectées et utilisées	45 %	29 %	22 %	20 %	15 %	15 %	22 %	22 %
Partager facilement vos données pour gagner du temps	45 %	42 %	34 %	26 %	14 %	11 %	26 %	26 %
Partager facilement vos données pour recevoir des recommandations personnalisées	25 %	26 %	19 %	19 %	13 %	11 %	21 %	18 %
La transparence concernant la collecte et l'utilisation de vos données personnelles est importante	70 %	79 %	82 %	79 %	83 %	85 %	/	81 %
Partager facilement vos données avec des entreprises que vous connaissez déjà	74 %	74 %	68 %	64 %	54 %	58 %	67 %	64 %
La réputation d'une entreprise influence le partage de vos données personnelles	77 %	73 %	68 %	63 %	59 %	64 %	61 %	66 %
Il doit être clair quels avantages vous obtenez en partageant vos données personnelles	69 %	70 %	73 %	67 %	63 %	61 %	57 %	66 %

Réponses à la questions : Dans quelle mesure êtes-vous d'accord avec les affirmations suivantes ?

Source : SOLIDMonitor : Privacy, persoonlijke data & datakluizen, editie 2 -2023

Une majorité significative de Flamands montre une plus grande propension à divulguer leurs informations à des entreprises ou des organisations jouissant d'une bonne réputation (66 %), ou avec lesquelles ils ont eu par le passé des expériences positives (64 %). Ces proportions ont connu une augmentation notable par rapport à l'année 2022 (respectivement de 61 % et 58 %), ce qui peut refléter une confiance grandissante envers les systèmes d'évaluation et les avis en ligne, ainsi que la

capacité des plateformes et partenaires numériques à établir un climat de confiance avec leur clientèle.

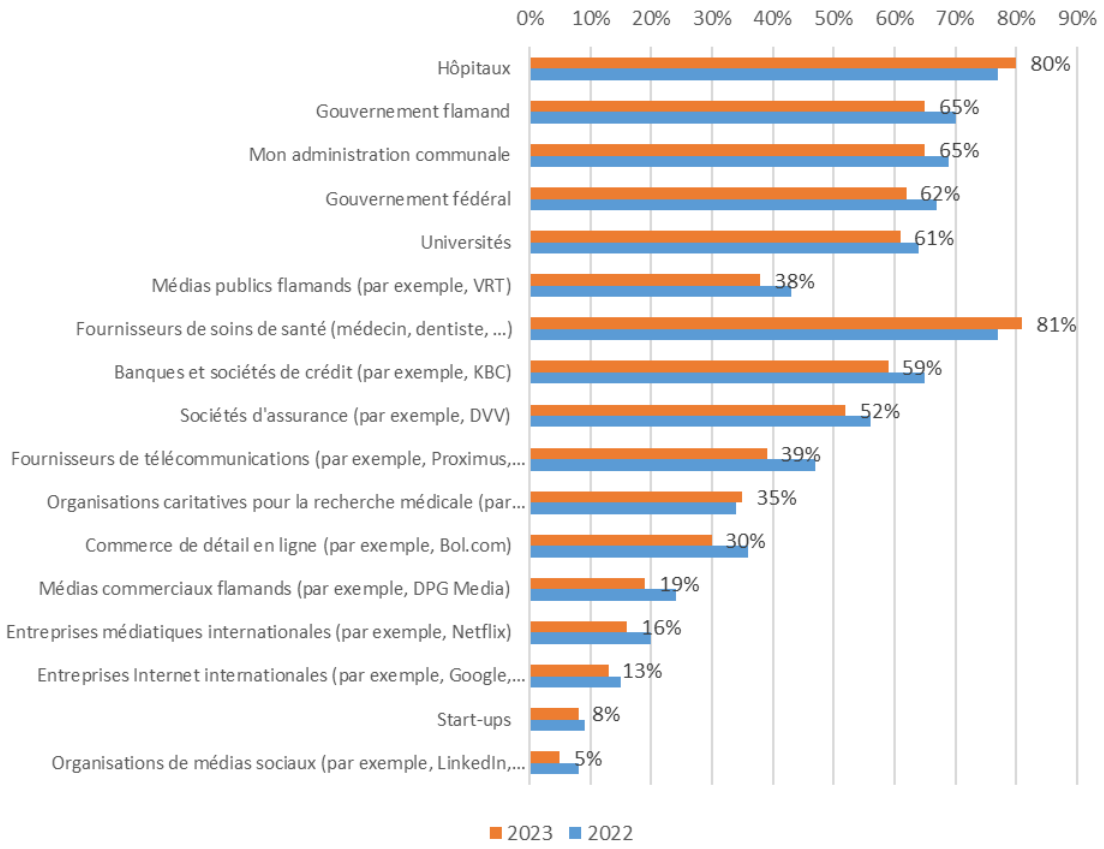
Toutefois, un scepticisme marqué persiste à l'égard des plateformes en ligne. 30 % des personnes sondées estiment qu'elles ne sont pas dignes de confiance, ce qui représente une augmentation de 3 % par rapport à 2022. Une majorité écrasante de Flamands, soit 83 %, perçoit ces compagnies comme agissant principalement dans leur propre intérêt et 79 % des personnes sondées sont d'avis que l'objectif principal de ces entreprises est de générer des profits à partir des données utilisateurs.

L'enquête de SOLIDMonitor révèle également qu'en 2023 (Figure 28) la confiance des citoyens dans la gestion de leurs données personnelles varie considérablement selon les secteurs et organisations :

- Les soins de santé, y compris les hôpitaux et les fournisseurs de soins, jouissent du niveau de confiance le plus élevé avec plus de 80 %, reflétant une confiance dans la confidentialité et la gestion éthique des données de santé. Ils sont suivis par les différents niveaux de gouvernement qui enregistrent une confiance modérée de 62 % (gouvernement fédéral) à 65 % (gouvernement flamand et administrations communales). Les universités se situent légèrement en dessous avec 61 %.
- Cependant, la confiance décline lorsque nous regardons vers le secteur financier et les sociétés d'assurance, où la confiance est de 59 % et 52 % respectivement, indiquant une vigilance quant à la sécurité et à la confidentialité financière. La prudence est encore plus marquée dans le secteur des médias publics flamands et des organisations caritatives pour la recherche médicale, avec des scores de 38 % et 35 % respectivement. Les fournisseurs de télécommunications (39 %), le commerce de détail en ligne (30 %), et les médias commerciaux flamands (19 %) reçoivent une confiance encore plus limitée.
- Enfin, les grandes plateformes numériques, ainsi que les médias sociaux privés et les start-ups ferment la marche avec moins de 20 % de confiance, révélant une grande prudence parmi le public vis-à-vis de ces entités dans le domaine de la vie privée. Les entreprises médiatiques internationales comme Netflix sont considérées comme légèrement plus dignes de confiance (16 %) que les géants technologiques internationaux tels que Google, Microsoft et Amazon, qui ne rassemblent que 13 % de la confiance. Enfin, la confiance est au plus bas pour les start-ups (8 %) et les plateformes de médias sociaux (5 %).

Globalement, l'année 2023 a été témoin d'une baisse générale de la confiance dans presque tous les secteurs et types d'organisations par rapport à 2022. Alors que les institutions de santé continuent de bénéficier d'une forte confiance, une méfiance accrue se manifeste envers les entités gouvernementales et les entreprises privées, notamment celles opérant dans les médias et la technologie.

Figure 28. Confiance accordée dans les secteurs/organisations dans le traitement des données personnelles



Réponses à la question : Dans quelle mesure faites-vous confiance aux institutions suivantes pour le traitement de vos données personnelles ?

Source : SOLIDMonitor : Privacy, persoonlijke data & datakluizen, editie 2 -2023.

L'étude aborde par ailleurs la question de la confiance accordée au partage selon le type de données partagées. Ainsi, une grande majorité des Flamands se sent à l'aise de partager des informations telles que leur tranche d'âge et leur sexe (86 %), leurs opinions sur des produits et services (82 %) ainsi que leurs centres d'intérêt (69 %), tandis que les données sur les achats effectués (36 %) et les informations de localisation (14 %) obtiennent des scores significativement plus bas que l'année précédente. De surcroît, certains types de données sont clairement considérés comme plus sensibles. Par exemple, seulement 9 % se sentent à l'aise de partager des messages, des photos et des contacts sur les réseaux sociaux avec des entreprises. Il n'est donc pas surprenant que les données de santé (15 %), les données financières (6 %) et les informations de carte de crédit (4 %) soient particulièrement mal notées.

L'étude s'est aussi penchée sur les éléments susceptibles d'accroître la confiance des Flamands dans le partage de leurs données avec des organisations. Parmi les éléments identifiés, la clarification de l'usage prévu des données personnelles par les organisations se détache nettement, recueillant l'adhésion de 63 % des participants. Cet aspect révèle une demande pressante pour plus de transparence et d'imputabilité de la part des entités collectrices. De façon notable, près de la moitié des sondés (47 %) souhaitent recevoir des instructions claires et détaillées pour partager leurs données de manière sécurisée.

En revanche, la simplification du processus de partage des données personnelles ne séduit que 22 % des répondants, suggérant ainsi que la complexité n'est pas perçue comme le principal frein à la confiance dans l'échange d'informations. Cette perspective peut indiquer que les individus accordent davantage d'importance à la sécurité et à la clarté des procédures qu'à leur simplicité. D'autre part, le

recours à une assistance personnalisée ou à des formations locales en matière de sécurité des données semble moins convaincant, avec respectivement 15 % et 20 % de réactions favorables.

En outre, il est notable que 15 % des participants estiment qu'aucune mesure ne les rendrait plus confortables avec le partage de leurs données personnelles. Cette attitude révèle potentiellement une méfiance profonde ou un mécontentement vis-à-vis des dispositifs de protection actuels. En contraste, seule une minorité de 6 % se dit déjà à l'aise avec l'idée de partager leurs informations.

5. Infrastructure

Principaux résultats

Couverture des réseaux à très haute capacité (VHCN) :

- 78 % des ménages couverts en 2022, dépassant la moyenne européenne de 73 %.
- 88,4 % des PME ont accès à une connexion à large bande rapide (au moins 30 Mbps) et 66,2 % des entreprises disposent d'une connexion ultra-rapide (≥100 Mbps).

Déploiement de la fibre optique :

- Seulement 17 % de couverture, en dessous de la moyenne européenne de 56 %.
- Objectif de l'opérateur historique d'atteindre 50 % des foyers d'ici 2025.

Connexions internet à très haut débit :

- 3,1 % des ménages bénéficient de connexions d'au moins 1 Gbps.
- 10,2 % des entreprises accèdent à des débits d'au moins 1 Gbps.

Déploiement et adoption de la 5G :

- Augmentation de la couverture 5G de 4 % en 2021 à 30 % en 2022, mais la Belgique reste loin derrière la moyenne de l'UE de 81 %.

En matière d'infrastructures numériques, la Belgique affiche une image contrastée. D'une part, le pays se distingue par une bonne couverture des réseaux à très haute capacité (VHCN[38]), où elle enregistre une couverture de 78 % des ménages en 2022, un chiffre supérieur à la moyenne européenne de 73 %. Ainsi, environ 88,4 % des PME bénéficient d'une connexion à large bande rapide (au moins 30 Mbps), surpassant légèrement la moyenne européenne de 83,8 %. De plus, 66,2 % des entreprises disposent d'une connexion ultra-rapide (≥100 Mbps).

En revanche, la Belgique est à la traîne en ce qui concerne le déploiement de la fibre optique à domicile, avec une couverture de seulement 17 %, loin derrière la moyenne européenne de 56 %. Toutefois, la perspective concernant la fibre optique est optimiste, l'opérateur historique prévoyant que d'ici 2025, la moitié des foyers belges seront raccordés à la fibre.

L'adoption des connexions internet à très haut débit (au moins 1 Gbps) reste également limitée. Seulement 3,1 % des ménages bénéficient de telles connexions, contre une moyenne de 13,8 % au niveau européen. Pour les entreprises, la situation est similaire, avec 10,2 % d'entre elles accédant à des débits d'au moins 1 Gbps, plaçant la Belgique légèrement en dessous de la moyenne de l'UE de 11,4 %.

Enfin, la Belgique accuse un certain retard dans le déploiement et l'adoption de la 5G. Malgré des progrès importants réalisés en 2022, avec une augmentation de la couverture 5G, qui est passée de 4 % à 30 % en un an, la Belgique reste loin derrière la moyenne de l'UE (81 %) et de l'objectif de la Décennie numérique 2030.

[38] Selon la définition du code des communications électroniques européen, un réseau à très haute capacité ou VHCN est un réseau de communications électroniques qui :

- est entièrement composé d'éléments de fibre optique au moins jusqu'au point de distribution ;
- ou est capable d'offrir, dans des conditions d'heures de pointe habituelles, une performance du réseau comparable (par ex. en termes de largeur de bande, de latence, etc.).

6. Réglementations

Principaux résultats

Transpositions des réglementations européennes en loi belge

DGA

- Désignation en mai 2024 des autorités compétentes (SPF BOSA et SPF Economie).
- Le SPF BOSA a été désigné en mai 2024 comme intégrateur de services fédéral
- Des candidats sont prêts à solliciter une accréditation en tant qu'intermédiaires de données au sens du DGA.

DATA ACT

- Organisation, notamment par Agoria, de groupes de travail pour la préparation à l'application effective au 12 septembre 2025

Développer un cadre juridique des données

Dans la dernière décennie, l'UE s'est engagée dans une dynamique législative pour créer un écosystème numérique qui protège les données tout en favorisant l'innovation et la croissance économique. La Commission européenne a mis en place un cadre juridique (voir Box 2 pour un aperçu des réformes législatives) pour favoriser l'adoption et le développement des technologies à base de données, tout en veillant à leur alignement avec les normes éthiques et de protection des données européennes. Les règlements sont directement applicables dans les 27 États membres de l'Union européenne tandis que les directives doivent être transposées dans des lois nationales.

Un des piliers de cette réforme législative est le règlement général sur la protection des données (RGDP) relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données. Il est applicable dans les 27 États membres à compter du 25 mai 2018.

En outre, les pays européens appliquent le règlement européen relatif à la libre circulation des données à caractère non personnel qui vise à faciliter des nouveaux modèles commerciaux et à améliorer les services des administrations publiques. Ensemble avec le Data Act, le Data Governance Act, le Digital Markets Act et Digital Service Act, ce règlement définit le cadre juridique de l'économie complétive du numérique

Selon la Data Governance Act, chaque État membre est tenu de désigner un ou plusieurs organismes compétents pour s'acquitter des tâches énoncées dans l'article 7, ainsi qu'une ou plusieurs autorités compétentes chargées de réaliser les tâches relatives à la procédure de notification pour les services d'intermédiation de données (article 13). De plus, chaque État membre doit désigner un ou plusieurs organes compétents responsables du registre public national des organisations altruistes reconnues en matière de données (article 23).

Comme le marché des données touche tous les secteurs, le développement de législations adaptées en fonction des spécificités des secteurs est laborieux. En attendant d'avoir une législation par secteur et domaine d'application, plusieurs pays misent sur les **bacs à sable réglementaires**. Ce sont des outils réglementaires permettant à des entreprises sélectionnées de tester et d'expérimenter des produits ou services innovants dans un cadre réglementaire assoupli sous la supervision d'un régulateur pendant une période limitée. De plus, le AI Act encourage les États membres à établir des bacs à sable en matière d'intelligence artificielle pour faciliter le développement et les tests de systèmes d'IA innovants avant qu'ils ne soient commercialisés ou mis en service.

Réformes législatives européennes en matière de gouvernance de données

Afin d'atteindre les objectifs de la Décennie numérique de l'Europe et de promouvoir la construction d'une économie européenne de la donnée respectueuse des libertés et droits fondamentaux, la Commission a adopté plusieurs **propositions de règlement européen** sur les données :

Data Governance Act

L'un des piliers de la stratégie européenne pour les données est le **règlement européen sur la gouvernance des données**[39]. Ce dernier a été publié au Journal officiel de l'Union européenne le 3 juin 2022, est entré en vigueur le 23 juin 2022 et s'applique depuis septembre 2023.

Le règlement européen sur la gouvernance des données vise à accroître la disponibilité des données, les possibilités de les réutiliser et la confiance dans le partage des données. Ce règlement comporte :

- un encadrement ainsi qu'une assistance technique et juridique facilitant **la réutilisation de certaines catégories de données protégées du secteur public** (informations commerciales confidentielles, propriété intellectuelle, données personnelles) ;
- une **certification obligatoire** pour les fournisseurs de services d'intermédiation de données ;
- une **certification facultative** pour les organismes pratiquant l'altruisme en matière de données.

Data Act

Le 22 décembre 2022, la Commission a publié le **règlement européen sur les données**[40] fixant des règles harmonisées pour l'équité de l'accès aux données et de l'utilisation des données.

À ce titre, le Data Act a pour objectifs de :

- faciliter le **partage entre entreprises (B2B) et avec le consommateur (B2C) des données**, en fixant notamment une obligation de rendre accessibles les données générées par l'utilisation des objets connectés et services connexes, en contrepartie d'une compensation juste et équitable ;
- permettre **l'utilisation des données** détenues par les entreprises et, sous réserve de justifier d'un besoin exceptionnel, par les organismes publics des États membres et les institutions, agences ou organes de l'Union ;
- faciliter **le changement de fournisseur de services** de traitement de données (cloud et edge computing) par l'encadrement des relations contractuelles entre les fournisseurs de services et les consommateurs, et notamment par la suppression progressive des frais liés au changement pour le consommateur ;
- prévoir l'élaboration de **normes d'interopérabilité** pour les données et leurs réutilisations entre les secteurs ;
- mettre en place des **garanties contre les accès illicites de gouvernements de pays tiers** aux données non-personnelles contenues dans le cloud.

AI Act

Le 21 avril 2021, la Commission européenne a publié sa proposition de **législation sur l'intelligence artificielle**[41], consacrant dans le droit de l'Union une définition des systèmes d'IA qui soit neutre d'un point de vue technologique et adoptant un ensemble de règles distinctes qui suivent une approche fondée sur le risque.

[39] https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=celex:32022R0868.

[40] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202302854.

[41] https://eur-lex.europa.eu/legal-content/FR/TXT/?qid=1623335154975&uri=CELEX:52021PC0206.

General Data Protection Regulation

Le **règlement général sur la protection des données**[42] (RGDP) renforce les droits existants, octroie de nouveaux droits et accorde aux personnes un meilleur contrôle sur les données à caractère personnel les concernant. Les points clés de cette législation comprennent[43] :

- Un **accès facilité aux données personnelles** d'un individu. Cela inclut de fournir davantage d'informations sur la manière dont ces données sont traitées et de veiller à ce que ces informations soient disponibles de manière claire et compréhensible.
- Un nouveau **droit à la portabilité des données**. Il facilite le transfert de données à caractère personnel entre prestataires de services.
- Un **droit d'effacement plus clair** (droit à l'oubli). Lorsqu'une personne ne souhaite plus que ses données soient traitées et qu'il n'existe pas de motif légitime de les conserver, les données seront effacées.
- Le **droit de savoir** quand des données à caractère personnel ont été violées. Les entreprises et les organisations doivent informer l'autorité de surveillance compétente en matière de protection des données ainsi que, en cas de violation grave des données, les personnes concernées.

Free flow of non-personal data

Le **règlement relatif à la libre circulation des données à caractère non personnel**[44] vise à éliminer les obstacles à la libre circulation des données à caractère non personnel entre les différents pays de l'UE et les systèmes informatiques en Europe.

Le règlement garantit :

- La **libre circulation des données à caractère non personnel** au-delà des frontières : chaque organisation devrait être en mesure de stocker et de traiter des données n'importe où dans l'UE ;
- La **disponibilité des données à des fins de contrôle réglementaire** : les autorités publiques conserveront l'accès aux données, même lorsqu'elles sont situées dans un autre pays de l'UE ou lorsqu'elles sont stockées ou traitées dans le cloud ;
- Un **portage de données plus facile** entre les fournisseurs de services cloud pour les utilisateurs professionnels.

Open data directive

La **directive sur les données ouvertes et la réutilisation des informations du secteur public**[45] (directive sur les données ouvertes) prévoit des règles communes pour un marché européen des données détenues par les pouvoirs publics. En bref, la philosophie de la législation sur les données ouvertes est de rendre les documents existants des organismes du secteur public disponibles afin qu'ils puissent être utilisés dans un cadre commercial ou non commercial.

Adopter des lignes directrices d'éthique

La pratique de collection, analyse et stockage de données – en particulier celles à caractère personnel – représente un fort enjeu d'éthique. La Commission a pris plusieurs mesures afin de veiller à ce que l'avenir du numérique et des données en Europe repose sur les valeurs européennes et les droits fondamentaux. En avril 2019 le groupe d'experts de haut niveau sur l'intelligence artificielle a présenté des lignes directrices d'éthique pour une IA digne de confiance[46]. Le rapport qui forme la base pour ces lignes directrices propose une approche centrée sur l'humain en matière d'IA et énumère sept exigences clés auxquelles les systèmes d'IA devraient satisfaire afin d'être dignes de confiance : 1) action humaine et contrôle humain, 2) robustesse technique et sécurité, 3) respect de

[42] https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=celex%3A32016R0679.

[43] https://eur-lex.europa.eu/FR/legal-content/summary/general-data-protection-regulation-gdpr.html.

[44] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1807.

[45] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024.

[46] https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai.

la vie privée et gouvernance des données, 4) transparence, 5) diversité, non-discrimination et équité, 6) bien-être sociétal et environnemental, et 7) responsabilité. En 2021, la Direction Générale de la recherche et de l'innovation a publié un rapport sur l'éthique de protection des données. Ce document présente un aperçu des méthodes à suivre pour garantir une meilleure protection des données suivant le règlement général sur la protection des données. Enfin, en octobre 2022, la Commission a publié des lignes directrices d'éthique pour l'usage de l'IA et des données dans l'enseignement[47]. Les recommandations ciblent principalement les enseignants de l'enseignement primaire et secondaire.

DGA

La Belgique a désigné les autorités compétentes au regard du Data Governance Act, par la loi du 15.05.2024[48] :

- Comme **autorité compétente** au sens de l'article 7 : **SPF BOSA**
  - Le SPF BOSA a également été désigné en tant qu'intégrateur de services fédéral par la loi du 13.05.2024[49] ;
- Comme autorité compétente pour l'intermédiation des données au sens de l'article 13 : Service public fédéral Economie, P.M.E., Classes moyennes et Energie
- Comme autorité compétente pour l'altruisme des données au sens de l'article 23 :
  - Service public fédéral Economie, P.M.E., Classes moyennes et Energie

Le SPF Economie doit faire appel à l'expertise des autorités sectorielles compétentes, en particulier l'Autorité de protection des données, l'Autorité belge de la concurrence et le Centre de Cybersécurité Belgique, lorsque des questions qui relèvent de leurs compétences sont posées (par exemple protection des données, concurrence, sécurité).

« Pour encourager la réutilisation à des fins non commerciales, comme pour la recherche scientifique, ainsi que par les PME et les jeunes pousses, la redevance sera réduite ou supprimée. Le Roi détermine le montant d'une redevance réduite ou la gratuité de certaines données, notamment en vue d'une réutilisation par les PME et les jeunes pousses, les organisations de la société civile, les organismes exerçant une activité de recherche et les établissements d'enseignement. »

« Le Service public fédéral Stratégie et Appui (ci-après « SPF BOSA ») sera désigné dans les dispositions modifiant la loi du 15 août 2012 portant création et organisation d'un intégrateur de services fédéral comme organisme compétent (conformément à l'article 7, paragraphe 1, du règlement 2022/868). Dans ce rôle il fournira du soutien technique comme prévu dans le règlement 2022/868 aux organismes fédéraux du secteur public. En tant qu'intégrateur de services fédéral, le SPF BOSA est celui qui permet déjà l'échange de nombreuses données de manière sécurisée au niveau fédéral. Grâce à cette expérience et expertise, le SPF BOSA est en mesure d'assumer le rôle d'organisme compétent pour fournir du soutien technique aux organismes du secteur public lors de la mise à disposition des données. »

Par ailleurs, le réseau fédéral d'experts en données (FEDAX) organisé par BOSA, a pour fonction de coordonner les actions des Chief Data Officer (CDO) des différents services publics fédéraux (SPF).

DATA ACT

Le Data Act est entré en vigueur le 11 janvier 2024. Une période de transition permettant aux acteurs concernés de se mettre en règle se déroule jusqu'au 12 septembre 2025 pour la plupart des dispositions.

---

[47] https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en.

[48] Loi du 15/05/2024 mettant en oeuvre le règlement (UE) 2022/868 du Parlement européen et du Conseil du 30 mai 2022 portant sur la gouvernance européenne des données et modifiant le règlement (UE) 2018/1724, Publication au Moniteur le 07/06/2024 https://www.lachambre.be/FLWB/PDF/55/3934/55K3934001.pdf.

[49] Loi modifiant la loi du 15 août 2012 relative à la création et à l'organisation d'un intégrateur de services fédéral. (Publication au Moniteur le 31/05/2024) https://www.lachambre.be/FLWB/PDF/55/3961/55K3961001.pdf.

Le Workgroup Data de Agoria accompagne ses membres dans leur préparation à cette application du Data Act.

Partie 2. Analyse comparative des stratégies nationales en France, aux Pays-Bas, en Allemagne et au Royaume-Uni

1. Stratégies nationales de 4 pays européens

L'objectif de ce rapport d'étude est de réaliser un **rapport synthétique** qui analyse de manière comparative les éléments clés des stratégies nationales en matière de données dans quatre pays spécifiques : la France, les Pays-Bas, l'Allemagne et le Royaume-Uni. Le but est de mettre en évidence les différences, les similitudes et les tendances générales au sein de ces pays en ce qui concerne leurs approches des données. Cette analyse servira de base pour éclairer la formulation de recommandations pour la situation belge.

Plusieurs pays, dont l'Allemagne, les Pays-Bas et le Royaume-Uni, ont adopté des stratégies qui se concentrent spécifiquement sur la gouvernance des données et des politiques favorisant le bon fonctionnement du marché des données. Cependant, certains pays comme la France et les Pays-Bas, intègrent les recommandations et actions en matière de données dans d'autres stratégies, telles que celles liés au développement de l'intelligence artificielle et du numérique.

1.1. Allemagne

Le gouvernement fédéral d'Allemagne a publié **une stratégie nationale en matière de données**[50] en janvier 2021. Elle vise à accroître considérablement le développement et l'utilisation des données innovantes et responsables dans l'industrie, la recherche scientifique, la société civile et l'administration publique. Elle prévoit des actions fondées sur les valeurs européennes pour garantir la participation équitable des parties prenantes concernées et lutter contre l'utilisation abusive des données.

La stratégie couvre quatre domaines d'actions :

- Améliorer l'infrastructure des données et sécuriser l'accès aux données au niveau des infrastructures ;
- Encourager une **utilisation responsable des données** et exploiter le potentiel d'innovation ;
- Renforcer **les compétences en matière de données** et établir une nouvelle culture des données en Allemagne ;
- Créer une **nouvelle culture des données** afin que le gouvernement fédéral puisse devenir un leader mondial du marché des données.

En annexe, la stratégie présente une **liste de toutes les mesures proposées** en matière de données ainsi que le département responsable pour le développement et le suivi.

Aux actions mentionnées dans cette stratégie s'ajoutent celles d'autres **stratégies dans des domaines complémentaires** telles que l'intelligence artificielle et le numérique en général. Citons notamment la stratégie nationale en matière d'intelligence artificielle en 2018[51] (et sa mise à jour en 2020[52]), la stratégie en matière de données ouvertes[53], et la stratégie nationale du numérique publiée en août 2022[54].

---

[50] https://www.bundesregierung.de/breg-en/service/information-material-issued-by-the-federal-government/data-strategy-of-the-federal-german-government-1950612.

[51] https://www.ki-strategie-deutschland.de/home.html.

[52] https://www.ki-strategie-deutschland.de/files/downloads/Fortschreibung_KI-Strategie_engl.pdf.

[53] https://www.bundesregierung.de/breg-en/service/archive/open-data-strategy-1940558.

[54] https://digitalstrategie-deutschland.de/.

Plus récemment, deux mises à jour ont été publiées par le gouvernement fédéral. En août 2023, le gouvernement a publié une **stratégie sur le progrès grâce à l'utilisation des données**[55]. Elle reprend les mêmes thématiques que la stratégie en matière de données publiée en 2021. Elle vise à promouvoir une utilisation responsable et efficace des données et met l'accent sur le rôle des données dans l'innovation, la transparence et le progrès sociétal, en préconisant la culture des données, l'accessibilité et le contrôle. La stratégie soutient la recherche, une main-d'œuvre productive, une économie numérique et des objectifs de durabilité tout en abordant les questions de qualité et de diversité des données, y compris les lacunes en matière de données liées au sexe et à l'âge.

En outre, en novembre 2023, un nouveau **plan d'action pour l'intelligence artificielle**[56] a été établi par le ministère fédéral de l'Éducation et de la Recherche. Le plan d'action mentionne 50 actions en cours axées sur la recherche, le développement de compétences et de la construction d'infrastructures dans le cadre de l'intelligence artificielle. Le plan introduit au moins 20 initiatives supplémentaires, reflétant l'engagement d'investir plus de 1,6 milliard d'euros dans l'IA au cours de la période législative actuelle. Entre autres, le plan d'action a pour objectif de contribuer au discours actuel sur les risques de l'IA et des cadres juridiques nécessaires et vise à concrétiser les mesures de la stratégie nationale pour l'IA. En même temps, le plan d'action renforce le niveau de coopération des parties prenantes et contribue à intégrer ses activités dans le contexte européen.

1.2. Royaume-Uni

Le Royaume-Uni a publié une **stratégie nationale en matière de données**[57] en décembre 2020. La stratégie constitue un cadre pour les mesures que le gouvernement prévoit afin de promouvoir la compétitivité du Royaume-Uni dans le marché des données.

La stratégie est basée sur **quatre piliers** qui permettent de garantir une meilleure utilisation des données au Royaume-Uni :

- **Créer des bases de données** : garantir une meilleure qualité des données, tout en s'assurant qu'elles soient disponibles, accessibles, interopérables, réutilisables et enregistrées dans des formats standardisés sur des systèmes modernes ;
- **Améliorer les compétences en matière de données** : fournir les compétences nécessaires en matière de données à travers le système éducatif, mais aussi veiller à ce que les gens puissent continuer à développer les compétences en matière de données dont ils ont besoin tout au long de leur vie ;
- **Encourager la disponibilité des données** : encourager une meilleure coordination, l'accès et le partage de données entre les organisations des secteurs public, privé et tiers, et assurer des protections appropriées pour le flux de données à l'échelle internationale ;
- **Évoluer vers un traitement responsable des données** : veiller à ce que les données soient utilisées de manière responsable, équitable et durable, tout en soutenant l'innovation et la recherche.

À partir de ces piliers, la stratégie identifie **cinq missions d'actions prioritaires** permettant d'exploiter pleinement les opportunités qu'offrent les données :

- Exploiter **la valeur des données** dans l'ensemble de l'économie : définir des actions afin de rendre les données utilisables, accessibles et disponibles dans l'ensemble de l'économie, tout en protégeant les droits des personnes en matière de données et la propriété intellectuelle des entreprises privées ;
- Assurer **un régime de données de confiance** et favorable à la croissance : définir des actions qui favorisent une concurrence et une culture d'innovation dynamique, qui renforcent la confiance tout en maintenant des normes élevées en matière de protection des données sans créer d'obstacles inutiles à l'utilisation des données ;

[55] https://www.bmwk.de/Redaktion/DE/Publikationen/Digitale-Welt/fortschritt-durch-datennutzung.html.

[56] https://www.bmbf.de/bmbf/de/forschung/digitale-wirtschaft-und-gesellschaft/kuenstliche-intelligenz/ki-aktionsplan.html.

[57] https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy.

- Accroître **l'efficacité et la qualité des services publics** : définir des actions qui permettent d'améliorer la création, l'utilisation et le partage de données dans les services publics. Ceci inclut le développement d'une infrastructure adéquate, de normes d'interopérabilité et de compétences au sein du secteur public ;
- Assurer **la sécurité et la résilience de l'infrastructure** sur laquelle reposent les données : créer une infrastructure stable qui garantisse la protection des données et qui soit résiliente face aux risques potentiels ;
- **Défendre le flux international de données** : définir des actions qui favorisent le flux d'informations à travers les frontières afin d'alimenter les opérations commerciales mondiales, les chaînes d'approvisionnement et le commerce.

Le gouvernement du Royaume-Uni a mis en place plusieurs mesures afin de veiller à ce que la stratégie réponde à ses objectifs. Premièrement, à la fin de 2020, le gouvernement a organisé une **consultation publique** de la stratégie nationale en matière de données. L'engagement des parties prenantes ajoute de la valeur au processus d'élaboration des lignes directrices et agit comme un catalyseur de confiance. Une collaboration étroite entre un large éventail de secteurs et d'organisations est également essentielle compte tenu de la nature transversale des données. Deuxièmement, le gouvernement a mis en place un **processus de suivi et d'évaluation** qui permet de surveiller l'exécution de la stratégie et de s'assurer qu'elle atteigne les résultats escomptés. Ainsi, la mise en œuvre et la responsabilité de chaque mission et action prioritaire de la stratégie sont attribuées à un département spécifique du gouvernement (voir en annexe de la stratégie). De plus, en septembre 2021, le gouvernement du Royaume-Uni a publié un document d'orientation présentant un cadre de suivi et d'évaluation de la stratégie nationale de données. Ce document permet de suivre l'exécution des interventions gouvernementales, d'évaluer leur efficacité et de planifier d'autres interventions à l'avenir. Pour chaque mission, le document présente également une liste préliminaire d'indicateurs économiques afin d'évaluer le progrès de manière quantitative.

1.3. Pays-Bas

Le gouvernement des Pays-Bas a défini les priorités politiques et les actions pertinentes en matière de données dans plusieurs stratégies, à savoir une stratégie du numérique[58], une stratégie nationale en matière de données inter-administratives[59] et un plan d'action pour l'intelligence artificielle[60].

Le gouvernement des Pays-Bas a publié une **stratégie du numérique** en avril 2021, afin de présenter ses ambitions et objectifs pour réussir la transformation digitale aux Pays-Bas. Les **principaux domaines d'actions** de cette stratégie comprennent :

- Promouvoir le **développement, le partage et l'accès aux donnée**s ;
- Renforcer les **compétences du numérique** dans le système d'éducation et le marché du travail ;
- Soutenir le développement de centres de données et améliorer l'**infrastructure numérique** (réseau internet 5G) ;
- Accroitre la **résilience numérique** à l'aide d'une cybersécurité renforcée ;
- Stimuler l'utilisation du numérique dans le **service public**.

La stratégie mentionne également l'importance du numérique pour relever les défis sociétaux, en particulier dans les **secteurs clés** comme l'agriculture, la santé, l'éducation, la recherche, l'industrie intelligente, l'énergie, la mobilité et le secteur public.

Axé sur ce dernier secteur, le gouvernement des Pays-Bas a lancé une **stratégie en matière de données inter-administratives** en octobre 2021. Elle vise à accroitre l'utilisation et le partage de données dans le service public en créant un cadre juridique adéquat, un système de données fédéré sécurisé (cependant laissant la collecte, le stockage et la gestion des données primaires aux détenteurs de sources décentralisées). Elle mise également sur des études de cas avant de développer certaines mesures à l'échelle nationale.

---

[58] https://open.overheid.nl/documenten/ronl-f22bf801-e7f1-4561-ae63-8c9708a31e84/pdf.

[59] https://open.overheid.nl/documenten/ronl-bf2acf54-ad5f-4f32-afe2-0904a1d8e700/pdf.

[60] https://open.overheid.nl/documenten/ronl-e14cdcee-690c-4995-9870-fa4141319d6f/pdf.

Enfin, le **plan d'action pour l'intelligence artificielle** publié en octobre 2019 constitue un important pilier du marché des données. La stratégie mentionne l'importance d'améliorer l'accès aux données de qualité et d'investir dans des infrastructures numériques de pointe.

1.4. France

La France n'a pas de stratégie spécifique en matière de données, mais mentionne plusieurs objectifs clés dans sa **stratégie nationale pour l'intelligence artificielle**[61]. Lancée en 2018, la stratégie met en avant trois axes pour renforcer la position de la France dans ce secteur :

- Développer un écosystème de talents ;
- Diffuser l'IA et la donnée dans l'économie et dans l'administration ;
- Promouvoir un modèle éthique équilibré entre innovation et protection des droits fondamentaux.

Reconnaissant que la production et le partage de données jouent un rôle moteur dans l'essor de l'IA, plusieurs actions sont proposées afin de **renforcer les infrastructures de données** dans l'économie et la société pour alimenter l'innovation en IA. Ainsi, le programme national de recherche en intelligence artificielle (PNRIA) prévoit l'accès à des infrastructures numériques de pointe afin d'accroître les capacités de calcul disponibles. Le gouvernement a également contribué au financement du supercalculateur Jean Zay[62] acquis en 2019.

En novembre 2021, le gouvernement a publié la **2e phase de la stratégie nationale pour l'intelligence artificielle**[63] visant à conquérir les talents et transformer le potentiel scientifique en succès économiques. En matière de données, la stratégie mentionne les objectifs et résultats suivants :

- Accroitre les **compétences en matière de données** par la 1) création d'un portail national de recensement des formations à la science des données, à l'IA et à la robotique, 2) la formation continue à la science des données, à l'IA et à la robotique ;
- Faciliter **l'accès et le partage de données** : par exemple, 4 projets majeurs de mutualisation des données, inter ou intra-filières soutenus depuis 2019 dans les secteurs de l'agriculture (AgDataHub), de la logistique (IACargo), du traitement automatique de la voix (VoiceLab), accompagnés d'une plateforme générique d'IA as-a-service ;
- **Accompagner les PME dans leur transformation** dans l'économie des données et de l'IA et à donner davantage de visibilité à l'écosystème des acteurs émergents.

Concernant les compétences, le gouvernement français a également publié un **rapport d'évaluation des besoins en compétences et expertises en matière de données** en septembre 2021. Le rapport rappelle en premier lieu les enjeux de la politique de la donnée et les mesures gouvernementales, et dresse un état des lieux des compétences data et un panorama des différents métiers de la donnée. En conclusion, le rapport met en lumière la nécessité de développer une gestion dynamique des compétences avec une attention particulière portée à la formation continue, à l'évaluation et au suivi des compétences.

---

[61] https://www.entreprises.gouv.fr/fr/numerique/enjeux/la-strategie-nationale-pour-l-ia

[62] http://www.idris.fr/jean-zay/jean-zay-presentation.html.

[63] https://minefi.hosting.augure.com/Augure_Minefi/r/ContenuEnLigne/Download?id=334FD34F-7844-497E-9551-79EDFF3B2EEF&filename=1645%20-%20DP%20-%20Strat%C3%A9gie%20Nationale%20pour%20l%27IA%202%C3%A8me%20phase.pdf.

Tableau 8. Aperçu des stratégies en matière de données

Pays	Stratégie	Année
Allemagne	Stratégie sur le progrès grâce à l'utilisation des données	2023
	Plan d'action pour l'intelligence artificielle	2023
	Stratégie nationale du numérique	2022
	Stratégie en matière de données ouvertes	2021
	Stratégie nationale en matière de données du gouvernement fédéral	2021
	Stratégie nationale pour l'intelligence artificielle (mise à jour)	2018 (2020)
Royaume-Uni	Feuille de route pour le numérique et les données (mise à jour)	2022 (2023)
	Stratégie nationale du numérique	2022
	Stratégie nationale pour l'intelligence artificielle	2021
	Stratégie nationale en matière de données	2020
Pays-Bas	Stratégie nationale en matière de données inter-administratives	2021
	Stratégie du numérique (mise à jour)	2021 (2022)
	Stratégie pour le partage de données entre les entreprises	2019
	Plan d'action pour l'intelligence artificielle	2019
France	Évaluation des besoins de l'État en compétences et expertises en matière de données	2021
	Stratégie nationale pour l'intelligence artificielle (2e phase)	2018 (2021)

Note : les stratégies dans le domaine des données en particulier sont marquées en gras.

2. Analyse comparative des 4 stratégies

En règle générale, les lignes directrices et les principales initiatives proposées dans les stratégies nationales sont comparables à celles des stratégies européennes dans le domaine du numérique et des données.

En février 2020, la Commission européenne a lancé sa **stratégie numérique pour l'Europe**[64] qui porte sur cinq ans. Elle est axée sur trois objectifs décisifs : la technologie au service des personnes, une économie juste et compétitive, ainsi qu'une société durable et démocratique. Ainsi, elle constitue – avec le **Green Deal**[65] – le double défi de la transformation de l'économie européenne vers une économie et société durable du futur. Elle contient une série d'initiatives thématiques et législatives dans le domaine du numérique.

Stratégie européenne pour les données

En Europe, les données sont au cœur de la transformation numérique. Ainsi, l'écosystème des technologies numériques exploite et produit des volumes considérables de données, qui sont devenues une importante source de valeur économique et sociale.

Par conséquent, le deuxième grand axe sur lequel s'appuie la stratégie numérique pour l'Europe est **le partage de données**. Tout en fondant l'innovation sur les données, l'Union ambitionne de maintenir un équilibre entre la libre circulation des données et la préservation des normes relatives à la vie privée, à la sécurité, à la sûreté et à l'éthique.

À cette fin, une **stratégie européenne pour les données** a été publiée en février 2020, parallèlement à la stratégie numérique. Elle entend établir un véritable espace européen des données, un marché unique des données, pour mobiliser les données inutilisées, en autorisant leur libre circulation dans

[64] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age_fr.

[65] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_fr.

l'Union et entre les secteurs, au bénéfice des entreprises, des chercheurs et des administrations publiques. Les actions reposent sur quatre piliers :

1. Un cadre horizontal de **gouvernance pour l'accès aux données** et leur utilisation ;
2. Catalyseurs : investissements dans les données et **renforcement des capacités et des infrastructures européennes** pour l'hébergement, le traitement et l'utilisation des données, tout en renforçant l'interopérabilité ;
3. Compétences : donner à chacun les moyens d'agir, **investir dans les compétences** et dans les PME ;
4. Création **d'espaces européens communs des données** dans des secteurs stratégiques et des domaines d'intérêt public.

Actions et stratégies afin de soutenir le développement des technologies numériques et des données

La stratégie numérique a été complétée, en 2021, par la communication stratégique **Décennie numérique de l'Europe**[66] et le plan d'action **Boussole Numérique : l'Europe balise la décennie numérique**[67], qui porte sur 10 ans et traduit les ambitions de l'Union à l'horizon 2030 en objectifs concrets. Le plan d'action comporte quatre objectifs en matière de numérique à atteindre d'ici 2030[68] : améliorer les compétences numériques des citoyens et le taux d'employés dans le domaine des technologies de l'information et de la communication, soutenir la transformation numérique des entreprises, développer des infrastructures numériques sûres et durables et renforcer l'utilisation des technologies numériques dans les services publics.

Tout au long de cette décennie, la Commission a également développé des stratégies afin de soutenir la croissance des services et marchés numériques en Europe dans des secteurs clés, tels que :

- Intelligence artificielle (IA) : en 2021 la Commission a mis à jour le **plan coordonné sur l'IA**[69] qui vise à accélérer les investissements dans l'IA, à permettre la mise en œuvre rapide des stratégies nationales en matière d'IA et à aligner les politiques concernées à l'échelle de l'Union européenne ;
- Chaine de blocs : L'UE a développé une stratégie dans cette technologie afin d'être **chef de file dans le domaine de la chaine de blocs**[70]. En matière de données, il est important de s'assurer que la technologie de chaine de blocs soit compatible avec la forte réglementation européenne concernant la protection des données et de la vie privée ;
- Cybersécurité : en 2020, la Commission a publié une **stratégie de cybersécurité de l'UE**[71], afin de renforcer la résilience collective de l'Europe face aux cybermenaces et de promouvoir l'UE dans le cyberespace mondial. Entre autres, cette stratégie vise à améliorer la sécurité des réseaux (5G et prochainement 6G) et de l'approvisionnement numérique en général (les données et le nuage).

Chacune de ces stratégies prévoit des recommandations et des actions concrètes afin de **garantir le bon fonctionnement des marchés de données** et des systèmes basés sur les données. Les principaux leviers d'actions mentionnés dans ces stratégies se rapportent à :

- L'amélioration de la disponibilité, le stockage, l'échange, le partage et l'utilisation des données ;
- La **protection des données** et la sécurité des services ;
- L'amélioration des **compétences numériques** des citoyens en général, des générations à venir (grâce à une politique d'intégration du numérique en enseignement et une formation initiale et continue des enseignants) et de spécialistes du numérique (par des formations permanentes adaptées au besoin du marché) ;

---

[66] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_fr.

[67] https://eur-lex.europa.eu/legal-content/fr/TXT/?uri=CELEX:52021DC0118.

[68] Voir https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_fr pour les objectifs concrets à atteindre.

[69] https://digital-strategy.ec.europa.eu/fr/library/coordinated-plan-artificial-intelligence-2021-review.

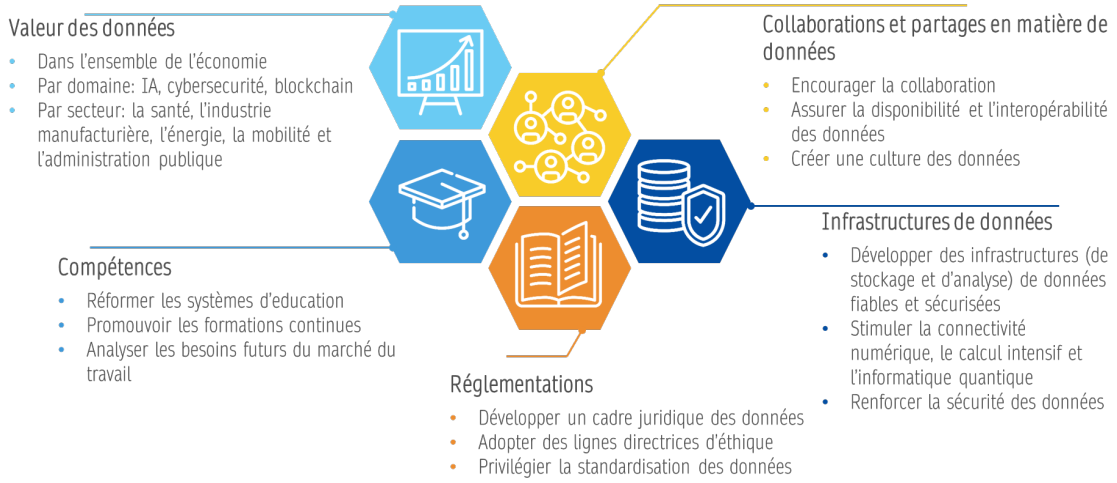
[70] https://digital-strategy.ec.europa.eu/fr/policies/blockchain-strategy.

[71] https://ec.europa.eu/commission/presscorner/detail/fr/ip_20_2391.

- Le renforcement des capacités de l'UE en matière d'**infrastructures de données** (par exemple, les initiatives visant le calcul haute performance[72], les technologies quantiques[73] et le cloud computing[74] et le réseau de connectivité 5G et 6G[75]) ;
- Le développement de **réformes législatives**, des exigences clés des **lignes directrices en matière d'éthique** sur l'utilisation et la gouvernance de données et participation au développement de **normes internationales** de données.

Tenant compte des lignes directrices proposées dans les stratégies européennes et nationales présentées ci-dessus, nous proposons de structurer l'analyse comparative des quatre pays le long des dimensions suivantes : compétences, valeur des données, collaborations en matière de données, infrastructures de données et réglementations (voir Figure 29).

Figure 29. Dimensions prises en compte dans l'analyse comparative des stratégies nationales en matière de données



2.1. Compétences

Les marchés professionnels des données ne cessent d'augmenter. Selon des projections récentes de la Commission européenne, le nombre de professionnels des données passera de 5,7 millions en 2018 à 10,9 millions en 2025[76]. En outre, l'UE-27 accuserait un déficit potentiel de 552.000 positions non pourvues d'ici 2030 causé par un déséquilibre entre la demande et l'offre de compétences en matière de données en Europe[77]. Pour le Royaume-Uni, un rapport commandé par le Département du numérique, de la culture, des médias et du sport (DCMS) prédit à base d'une analyse de 9,4 millions d'annonces d'emplois en ligne que la croissance de demandes en compétences d'analyse de données sera plus rapide que d'autres compétences du numérique au cours des cinq prochaines années. Afin de combler un déficit potentiel et d'encourager le développement de l'industrie des données, les quatre pays de cette analyse comparative prévoient des réformes visant à renforcer la littératie et les compétences en matière de données (voir Box 1 pour les définitions mentionnées dans les stratégies nationales).

[72] https://digital-strategy.ec.europa.eu/en/policies/high-performance-computing.

[73] https://digital-strategy.ec.europa.eu/fr/policies/quantum.

[74] https://digital-strategy.ec.europa.eu/fr/policies/cloud-computing.

[75] https://digital-strategy.ec.europa.eu/fr/policies/5g.

[76] https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_fr.

[77] https://ec.europa.eu/newsroom/dae/redirection/document/93914.

Les actions prévues dans les stratégies comportent entre autres :

- Réformer le système d'éducation afin d'accroitre la connaissance des données ;
- Promouvoir les formations continues ;
- Analyser les besoins futurs du marché du travail.

Box 1 : Définition des compétences en matière de données dans les stratégies

Selon la **stratégie fédérale allemande** les compétences en matière de données ou la connaissance des données peuvent être décrites comme la capacité de traiter les données de manière compétente et axée sur la valeur en termes techniques, économiques, éthiques et juridiques. Les compétences en matière de données sont essentielles au processus de numérisation et constituent les fondements de celui-ci.

La **stratégie du Royaume-Uni** utilise le terme des compétences en matière de données pour définir l'ensemble des compétences de base, techniques, et de gouvernance nécessaires pour maximiser l'utilité des données. Les compétences techniques requises vont de la programmation, de la visualisation des données, de l'analyse et de la gestion des bases de données aux compétences de base telles que la résolution de problèmes, la gestion de projets et la communication.

En **France**, le rapport d'évaluation des besoins de l'État en compétences et expertises en matière de données ne mentionne pas de définition des compétences, mais prévoit une définition des métiers de la donnée. Le rapport définit comme étant « experts de la donnée » les personnes dont les données sont l'objet essentiel du travail et non seulement un moyen de le réaliser. Cela inclut les métiers qui assurent le traitement informatique des données (et non seulement la collecte par une interface) et ceux qui exploitent les données pour produire des analyses, des recommandations, des prédictions.

Les **stratégies néerlandaises** suivent la même approche que celle de la France et définissent les compétences de plusieurs métiers de la donnée. L'ingénieur de la donnée est défini comme un spécialiste capable de collecter, éditer et stocker de grandes quantités de données de manière efficace. Le scientifique des données est défini comme un chercheur utilisant des méthodes, des processus et des systèmes d'analyse pour traiter de vastes ensembles de données et les interpréter.

Réformer les systèmes d'éducation

Une première étape pour garantir de meilleures compétences en matière de données des générations à venir est de **réformer le système d'éducation**. Les stratégies mentionnent la nécessité de renforcer les compétences de données dans l'enseignement scolaire dans son ensemble. Ainsi, la stratégie allemande et celle du Royaume-Uni prévoient l'intégration de programmes d'acquisition de compétences en matière de données dès le plus jeune âge ainsi qu'une formation de base et continue des enseignants en la matière.

Dans le contexte de l'enseignement universitaire, la stratégie allemande met en place un programme de doctorat dans le domaine des sciences de données. La stratégie du Royaume-Uni mentionne également des réformes de l'enseignement technique supérieur qui répondent aux besoins des employeurs, notamment en matière de données. En outre, en juin 2020, le DCMS et l'office pour l'IA a mis en place un programme[78] pour soutenir la conversion de diplômes en science des données et en intelligence artificielle.

Aux Pays-Bas, des réformes formelles de l'éducation sont prévues afin d'accroître la culture du numérique dans l'enseignement primaire et secondaire et l'enseignement supérieur à travers le programme national de formation en science des données[79]. Le gouvernement a également prévu un

[78] https://www.gov.uk/government/news/2500-new-places-on-artificial-intelligence-and-data-science-conversion-courses-now-open-to-applicants.

[79] https://www.cbs.nl/en-gb/corporate/2018/36/rijks-data-science-programme-kicks-off.

budget de 80 millions d'euros pour développer un laboratoire national d'éducation de l'intelligence artificielle[80].

Le gouvernement français mise principalement sur le renforcement de compétences dans le domaine de l'intelligence artificielle et du numérique. Dans le cadre de la stratégie d'accélération pour l'IA[81] et du Plan France 2030[82], le gouvernement déploie un plan de formation en deux volets sur une période de 5 ans pour un montant total de plus de 700 millions d'euros. Le premier volet (avec un budget de 500 millions d'euros) sert à développer des formations d'excellence de l'IA au niveau master et doctorat. Ce volet sert également comme pôle d'attractivité afin d'attirer les meilleurs potentiels étrangers vers cette école française de l'IA. Un deuxième volet (avec un budget de 276 millions d'euros) prévoit la massification de l'offre nationale de formation de l'IA dans l'enseignement supérieur. L'objectif est de former et de financer au moins 2.000 étudiants (niveau bachelor et diplôme université technique), 1.500 étudiants (niveau master), et 200 thèses supplémentaires par an. Le gouvernement français prévoit également la création d'un portail national de recensement des formations à la science des données, à l'IA et à la robotique.

Les stratégies de la France et du Royaume-Uni mentionnent également l'importance de suivre une approche multidisciplinaire et de favoriser une pluralité d'enseignements de la science de données. Ainsi ils favorisent l'intégration de cours de données dans une multitude de disciplines (par exemple l'IA dans les sciences humaines) et non seulement dans le cursus d'ingénieur.

Promouvoir les formations continues

De plus en plus d'entreprises ont besoin d'employés avec des compétences en matière de données. La pénurie de personnel qualifié dans ce domaine est un obstacle essentiel pour les entreprises, et principalement pour les PME.

La stratégie allemande prévoit plusieurs mesures afin d'accompagner les entreprises dans la formation de leurs employés dans le domaine des données. Des initiatives comme PME Digital[83] et Go Data facilitent la transformation numérique des PME et encourage leur participation à l'économie des données. La première initiative est axée sur des centres d'excellence situés dans toute l'Allemagne qui constituent un réseau d'échange d'expériences et d'exemples pratiques. Un projet de soutien éducatif durable et de conseils est également en cours de préparation. Le soutien éducatif se fera par l'intermédiaire de mentors internes de l'entreprise pour la formation continue, par exemple, par l'intermédiaire de 200 représentants d'entreprises. À l'échelle régionale, le gouvernement prévoit des centres de compétences pour la recherche sur le travail. Ces centres serviront à promouvoir l'utilisation de méthodes et d'outils basés sur l'intelligence artificielle dans les entreprises. En outre, le gouvernement prévoit des mesures de sensibilisation à travers une campagne nationale d'éducation numérique (par la création d'un espace de données sur les compétences), et plusieurs plateformes d'apprentissage (comprenant entre autres un portail central d'accès en ligne pour la formation continue).

Dans le cadre de la massification de l'offre nationale de formation de l'IA (voir ci-dessus dans la section Réformes du système d'éducation) le gouvernement français prévoit une formation continue à la science des données, à l'IA et à la robotique qui sera conduite dans le secteur de la formation professionnelle et de la préparation de l'économie française aux métiers numériques automatisés du futur.

Aux Pays-Bas, les formations continues sont encouragées par le programme STAP[84] — un investissement de 200 millions d'euros pour créer des possibilités de formation dans le domaine de l'IA et des compétences numériques pour les citoyens — et par un programme de subsides[85] pour des

---

[80] https://www.rijksoverheid.nl/documenten/rapporten/2020/09/30/verkenning-naar-het-nederlandse-onderwijslab-artificiele-intelligentie.

[81] Stratégie d'accélération « Intelligence artificielle » | entreprises.gouv.fr.

[82] France 2030 | gouvernement.fr.

[83] https://www.mittelstand-digital.de/MD/Redaktion/DE/Publikationen/mittelstand-digital-flyer-englisch.pdf?__blob=publicationFile&v=3.

[84] https://www.rijksoverheid.nl/actueel/nieuws/2019/06/03/persoonlijk-ontwikkelbudget-voor-iedereen.

[85] https://www.rijksoverheid.nl/onderwerpen/leven-lang-ontwikkelen.

formations continues, notamment dans le domaine du numérique. Le groupe de travail thématique sur les ressources humaines de la Coalition sur l'intelligence artificielle des Pays-Bas (NL AIC)[86] a également développé des cours d'apprentissage en ligne, tels que des cours nationaux consacré à l'IA dans les secteurs des soins et santé[87] et agroalimentaire[88].

La stratégie du Royaume-Uni mentionne le développement d'un portail en ligne qui soutiendra l'accès des entreprises à la formation aux compétences en matière de données, en aidant les PME à afficher des supports de formation en ligne adaptés à leurs capacités et ambitions en matière de science des données techniques.

La plupart des stratégies ne se focalisent pas uniquement à combler la pénurie de compétences de données dans l'industrie, mais prévoient également des mesures d'encouragements pour les formations continues dans l'administration publique et les centres de recherche. Ainsi le gouvernement du Royaume-Uni prévoit d'examiner les besoins des administrations publiques, de faciliter l'accès aux données et de mettre à disposition des formations sur les données à tous les fonctionnaires. De plus, par l'intermédiaire du campus Data Science[89] de l'office national des statistiques (ONS), la stratégie du Royaume-Uni prévoit de répondre aux besoins de compétences du gouvernement. Des mesures similaires sont prévues dans les stratégies des autres pays de cette analyse comparative. Contrairement aux autres stratégies, le gouvernement allemand prévoit des mesures spécifiques afin d'améliorer les compétences en matière de données dans les organisations civiles. La stratégie mentionne plusieurs programmes de financements pour aider les organisations civiles qui enseignent les compétences en matière de données aux citoyens allemands, ainsi que la création d'une plate-forme intégrée pour la fourniture de données dans les secteurs à but non lucratif et ceux pour le bien commun.

Analyser les besoins futurs du marché du travail

Les stratégies de l'Allemagne et du Royaume-Uni mentionnent l'importance de suivre les évolutions des exigences dans l'industrie et des besoins futurs du marché de travail. Pour ce faire, le gouvernement du Royaume-Uni a mis en place la Data Skills Taskforce **afin** d'agir comme un forum de partage des connaissances et des meilleures pratiques entre les principaux participants de l'industrie et de l'enseignement supérieur, et de promouvoir les compétences et l'analyse des données pour les besoins du futur. Elle compte un large éventail de membres de l'industrie, du milieu universitaire, des centres de recherche et du gouvernement. De plus, le gouvernement du Royaume-Uni a publié une stratégie[90] en 2021 qui estime le déficit de compétences en matière de données pour les années à venir.

---

[86] https://nlaic.com/en/.

[87] https://zorg.ai-cursus.nl/home.

[88] https://agrifood.ai-cursus.nl/home.

[89] https://www.ons.gov.uk/aboutus/whatwedo/datasciencecampus.

[90] https://www.gov.uk/government/publications/quantifying-the-uk-data-skills-gap/quantifying-the-uk-data-skills-gap-full-report.

Tableau 9. Aperçu des actions mentionnées dans les stratégies pour améliorer les compétences en matière de données

Actions	Pays
Réformer le système d'éducation	
Développer des formations pour la mise à niveau des compétences en matière de données dans l'enseignement scolaire primaire et secondaire	DE, UK, FR, NL
Réformer l'enseignement technique supérieur et universitaire orienté vers des disciplines de sciences de données (masters et doctorats)	DE, UK, FR, NL
Prévoir une formation de base et continue des enseignants en matière de données	DE, UK
Valoriser les formations multidisciplinaires en intégrant la science des données, de l'IA et de la robotique dans un maximum de disciplines	UK, FR
Créer un portail national de recensement des formations à la science des données, à l'IA et à la robotique	FR
Promouvoir les formations continues	
Encourager des formations continues en matière de données dans l'industrie	DE, UK, FR, NL
Favoriser le développement de plateformes en ligne d'apprentissage de compétences en matière de données	DE, UK, FR
Stimuler les formations continues dans l'administration publique et les centres de recherche	DE, UK, FR, NL
Améliorer les compétences en matière de données dans les organisations civiles	DE
Analyser les besoins futurs du marché du travail	
Estimer le déficit de compétences en matière de données pour mieux répondre aux besoins futurs du marché de travail	DE, UK

2.2. Valeur des données

L'Allemagne mentionne plusieurs programmes de financement et de **soutien pour la recherche et l'innovation de l'économie des données**[91]. Une des sections du centre de données de recherche national finance la recherche les aspects éthiques, juridiques et sociaux de la numérisation, des mégadonnées et de l'intelligence artificielle. Un autre programme nommé « Smart Data Management » finance des projets qui développent de nouveaux produits ou service à base de données. D'autres programmes se concentrent sur l'avenir de la création de valeur dans l'économie des plateformes et de données, ainsi que de l'industrie 4.0. En collaboration avec la Commission européenne, le gouvernement allemand promeut l'utilisation responsable des données et la création de valeur en Afrique en créant des espaces de données africains-européens. Le gouvernement prévoit également un plan d'action qui vise à stimuler une culture de partage et de réutilisation de données pour la recherche.

En France, la deuxième phase de la stratégie nationale pour l'intelligence artificielle prévoit un budget de 134 millions d'euros pour **stimuler la recherche scientifique en IA**. Les priorités de la recherche sont l'IA embarquée (composants et architectures nanoélectroniques, couches logicielles et interfaces composants/logiciels) et l'IA frugale (en données / en puissance de calcul / efficience énergétique), l'IA décentralisée (complémentarités et alternatives entre les architectures décentralisées et le cloud pour l'IA) et l'IA de confiance dans les systèmes critiques et complexes. Une partie de la recherche sera dédiée à l'approfondissement des fondements mathématiques de l'IA.

[91] Le site https://germandigitaltechnologies.de/digital-landscape/research-and-innovation-programmes/ donne un aperçu des programmes de recherche et d'innovation pour promouvoir l'économie du numérique et des données.

Aux Pays-Bas, l'université d'Amsterdam et la NL AIC ont créé un **centre d'innovation pour l'IA**[92]. Financé par ses différents partenaires du gouvernement, de centres de recherche et industriels, ce centre stimule les connaissances en matière de l'IA et du numérique. En outre, le Conseil néerlandais de la recherche soutient plusieurs programmes de recherche sur l'application de la numérisation dans la recherche scientifique, l'IA et l'analyse de mégadonnées. Citons également les programmes Small Innovation Research (SBIR) et le programme de recherche « Digital Society » mentionnés dans la stratégie du numérique. Le SBIR[93] est un programme qui encourage la recherche et l'innovation dans les PME.

Le Royaume-Uni s'appuie des initiatives telles que le plan pour la croissance[94] et la feuille de route pour la recherche et le développement[95] qui établissent un cadre pour libérer la valeur des données et stimuler la croissance à la fois dans le secteur numérique et dans l'ensemble de l'économie. Par ailleurs, le Royaume-Uni envisage des investissements stratégiques dans ce domaine. Dans le cadre de la Révision des dépenses de 2021[96], le gouvernement s'engage à injecter un supplément de 8 milliards de livres sterling dans la transformation numérique, des données et des technologies d'ici 2025.

Stimuler des secteurs prioritaires

La majorité des stratégies nationales mentionnent plusieurs secteurs prioritaires pour stimuler l'économie, tels que **la santé, les transports, l'énergie, l'industrie manufacturière et l'administration publique**. Ces secteurs génèrent une grande quantité de données, qui sont susceptibles d'être utilisées pour créer de nouveaux produits, ou services. Ainsi ils ont un potentiel élevé de création de valeur et d'innovation. Le Royaume-Uni mentionne également les secteurs des finances et des communications. Les gouvernements des Pays-Bas et du Royaume-Uni ont aussi créé des **groupes de travail sectoriels** sur les données intelligentes, qui ont pour objectif de coordonner et accélérer les initiatives existantes dans ces secteurs prioritaires.

Ci-dessous nous mentionnons des exemples d'initiatives clés présentés dans les stratégies des quatre pays de cette analyse comparative.

Administration publique

Le gouvernement du Royaume-Uni a développé des recommandations à suivre concernant l'utilisation, l'accès, le couplage et la sécurité des données dans l'administration publique. Ces recommandations créent un **cadre de gouvernance des données** au sein des services du secteur public et offrent plus de transparence sur l'utilisation de la prise de décision assistée par algorithme dans l'ensemble du gouvernement. Le Royaume-Uni prévoit aussi la création d'un **audit des inventaires de données** et de développer une plateforme de données intégrée pour le gouvernement. Il s'agira d'un environnement collaboratif numérique qui aidera le gouvernement à libérer le potentiel des données. Finalement, le Royaume-Uni a mis en place le système One Login qui donne accès à tous les services publics de la plateforme Gov.uk.

Pour mieux cibler les initiatives et comprendre le déploiement de l'économie des données dans l'administration publique, le Royaume-Uni a développé un **cadre d'évaluation de la maturité des données** (le Data Maturity Assessment ou DMA)[97]. Il s'agit d'un moyen de comprendre et d'identifier les forces et les faiblesses de l'écosystème de données d'une organisation. Il est mesuré et évalué à l'aide du cadre d'auto-évaluation de la maturité des données comprenant plusieurs thèmes présentés dans le Tableau 10, Tableau 11 et Tableau 12. Les thèmes et les descripteurs de maturité qui composent le DMA sont pertinents pour le contexte gouvernemental et s'appliquent à une large gamme d'organisations du secteur public de différentes tailles, structures et complexités. Pour chaque

[92] https://icai.ai/.

[93] https://www.rvo.nl/subsidies-financiering/sbir/sbir-innovatie-opdracht.

[94] https://www.gov.uk/government/publications/build-back-better-our-plan-for-growth.

[95] https://www.gov.uk/government/publications/uk-research-and-development-roadmap.

[96] https://www.gov.uk/government/publications/autumn-budget-and-spending-review-2021-documents.

[97] https://assets.publishing.service.gov.uk/media/64184bccd3bf7f79d9675dbd/Data_Maturity_Assessment_for_Government_-_FINAL_PDF.pdf.

thème, les organisations peuvent se situer sur une échelle de cinq niveaux de maturité. Cette évaluation leur permet de déterminer les actions à mettre en œuvre pour améliorer leur maturité des données dans chaque thème.

Les Pays-Bas sont également très avancés dans ce domaine avec la **stratégie nationale en matière de données inter-administratives**. Elle présente des mesures importantes pour améliorer les services et la gestion de l'information, entre autres par la construction d'un système de données fédéré et d'un point d'accès commun des données. La coalition de l'IA a également créé un groupe de travail afin de promouvoir l'optimisation des services publics à l'aide de l'intelligence artificielle.

Le gouvernement allemand envisage également l'utilisation de l'IA et des **grands modèles linguistiques (LLM)** tels que le chatbot ChatGPT dans le secteur public, tout en respectant les normes de protection des données. Pour cela, plusieurs acteurs, tels que le Centre Consultatif pour l'Intelligence Artificielle dans l'Administration Publique (BeKI), le Commissaire fédéral à la protection des données et à la liberté de l'information (BfDI) et les laboratoires de données collaborent pour identifier les besoins et les exigences des LLM, notamment en matière de sécurité des données, de souveraineté numérique et d'applications open source.

En France, le service public de la donnée vise à mettre à disposition, en vue de faciliter leur réutilisation, les bases de données de référence qui présentent le plus fort impact économique et social.

Tableau 10. Aperçu des thèmes et concepts inclus dans le Data Maturity Assessment

Être en contact avec les autres
Comprendre les besoins des utilisateurs : établir une communication ouverte et une collaboration entre les équipes de données et les utilisateurs afin de comprendre leurs besoins et leurs attentes.
Collaboration : encourager et faciliter la collaboration entre les équipes de données et les autres parties prenantes afin de tirer parti des expertises et des ressources partagées.
Partage des données : mettre en place des processus sécurisés et responsables pour le partage et l'échange de données entre les différentes parties prenantes.
Réseaux et communautés de pratique : encourager la participation aux réseaux et aux communautés de pratique pour partager des connaissances, des expériences et des meilleures pratiques en matière de gestion des données.
Disposer des bonnes compétences en matière de données
Littératie des données des employés : mettre en place des formations et des ressources pour améliorer la compréhension et l'utilisation des données parmi tous les membres du personnel.
Compétences en matière de données dans les rôles de leadership : encourager et soutenir les leaders à développer des compétences en analyse de données pour prendre des décisions éclairées.
Développement de spécialistes des données : investir dans la formation et le recrutement de spécialistes des données qualifiés pour gérer et exploiter les données efficacement.
Développement d'opportunités de perfectionnement des compétences en matière de données : offrir des programmes de formation, des ateliers et des mentorats pour aider les employés à développer leurs compétences en analyse de données.

Tableau 11. Aperçu des thèmes et concepts inclus dans le Data Maturity Assessment (suite)

Disposer des systèmes et outils adéquats
Infrastructure : établir une infrastructure technologique solide pour stocker, gérer et analyser efficacement les données.
Stockage : disposer d'un stockage sécurisé et fiable pour gérer les volumes croissants de données.
Outils/Technologie : investir dans des outils et des technologies de pointe pour analyser et visualiser les données efficacement.

Investissements dans les systèmes de données : allouer des ressources adéquates pour l'acquisition et la maintenance des systèmes et outils de données.
Prendre des décisions avec les données
Prise de décision opérationnelle et stratégique basée sur les données : faire de la data une source de référence pour les décisions prises au quotidien et pour établir des stratégies à long terme.
Compréhension des utilisateurs : collecter et analyser des données sur les utilisateurs pour mieux comprendre leurs besoins, leurs attentes et leurs comportements.
Surveillance des performances : suivre et analyser les indicateurs clés de performance (KPI) pour évaluer l'efficacité des opérations et identifier les domaines à améliorer.
Connaître les données dont vous disposez
Registres d'actifs / catalogues de données : mettre en place des registres d'actifs ou des catalogues de données pour suivre et gérer efficacement les données.
Suppression de données : disposer de processus clairs et sécurisés pour la suppression et la mise au rebut des données obsolètes ou inutiles.
Métadonnées : générer et gérer correctement les métadonnées afin de fournir des informations contextuelles et de faciliter la recherche et l'accès aux données.
Gérer et utiliser les données de manière éthique
Transparence : veiller à ce que les pratiques de collecte, d'utilisation et de stockage des données soient transparentes et compréhensibles pour toutes les parties prenantes.
Atténuation des biais : identifier et atténuer les biais potentiels dans les données et leur traitement pour garantir une prise de décision équitable et non discriminatoire.
Inclusion : promouvoir une utilisation inclusive des données qui respecte la diversité des populations et ne marginalise aucun groupe.
Surveillance et contrôle : établir des mécanismes de surveillance et de contrôle pour garantir que les données sont utilisées de manière responsable et conformément aux normes éthiques en vigueur.
Gouvernance des données
Normes de données : établir des normes communes pour la collecte, la gestion et l'utilisation des données afin d'assurer leur cohérence et leur qualité.
Qualité des données : identifier, corriger et prévenir les erreurs et les anomalies dans les données afin de garantir leur fiabilité et leur utilité.

Tableau 12. Aperçu des thèmes et concepts inclus dans le Data Maturity Assessment (suite)

Transfert de données : concevoir et utiliser des pipelines de données pour automatiser le transfert, le nettoyage et le traitement des données, en garantissant leur flux continu et leur disponibilité.
Suppression de données : disposer de processus clairs et sécurisés pour la suppression et la mise au rebut des données obsolètes ou inutiles afin de protéger la vie privée et de réduire la charge de stockage.
Définir une direction en matière de données
Principes de données : établir des principes clairs pour la collecte, l'utilisation et la gestion des données afin de guider les décisions et les comportements de l'organisation.
Politique de données : mettre en place une politique de données complète et détaillée qui définit les règles et les procédures régissant l'utilisation des données.
Stratégie de données : développer une stratégie de données globale qui définit les objectifs, les priorités et les initiatives en matière de données.
Protéger vos données
Sécurité des données : mettre en œuvre des mesures de sécurité robustes pour protéger les données contre les cybermenaces et les accès non autorisés.

Protection des données : respecter les réglementations en matière de protection des données et garantir l'intégrité et la confidentialité des données.
Récupération en cas de sinistre : disposer de plans de récupération en cas de sinistre pour protéger les données contre les pertes dues aux pannes d'équipement, aux catastrophes naturelles ou aux erreurs humaines.
Assumer la responsabilité des données
Gouvernance des données : établir un cadre de gouvernance des données clair et structuré pour définir les rôles, les responsabilités et les processus liés à la gestion des données.
Propriété des actifs et des processus : déterminer qui est responsable de la propriété et de la gestion des actifs et des processus liés aux données.
Surveillance de l'adhésion aux politiques et procédures : mettre en place des mécanismes de surveillance pour vérifier que les politiques et procédures en matière de données sont respectées.

Source : Adapté du Data Maturity Assessment https://www.gov.uk/government/collections/data-maturity-assessment-for-government.

Secteur de l'industrie manufacturière

L'importance des données dans l'industrie manufacturière réside dans leur capacité à transformer fondamentalement les processus de production et à stimuler l'innovation. Dans ce contexte, l'initiative **Manufacturing X**[98] a été conçue en Allemagne pour établir un **écosystème de données** permettant un échange de données fiable entre les entreprises, basé sur des normes ouvertes, et offrant aux entreprises une souveraineté numérique. Manufacturing X implémente la création de l'espace de données Industrie 4.0 mentionné dans la stratégie allemande « DataSpace Industrie 4.0 »[99].

Au Royaume-Uni, le **programme d'innovation** Made Smarter[100], cofinancé par des investissements privés et publics, favorise la transformation de l'industrie manufacturière britannique en intégrant de nouvelles technologies numériques et de données telles que l'intelligence artificielle et la réalité virtuelle. Le programme se focalise sur quatre activités : 1) encourager un leadership plus ambitieux, informé et ciblé pour la numérisation, 2) plaider en faveur d'une adoption plus large de la technologie dans les chaînes d'approvisionnement, 3) appeler à une accélération de l'innovation, à la création de nouvelles technologies, d'entreprises et de flux de valeur et 4) mettre l'accent sur l'importance de former les effectifs et d'identifier les compétences futures nécessaires à la numérisation.

En France, citons le **plan d'investissement** France 2030[101], dont l'objectif est de créer de nouvelles filières industrielles et technologiques afin d'accompagner les transitions écologiques et numériques. Doté de 54 milliards d'euros sur une période de 5 ans, ce plan vise à renforcer la compétitivité industrielle et à promouvoir les technologies d'avenir.

Enfin, la stratégie néerlandaise mentionne le **programme de soutien** « Smart industry »[102], visant à accompagner les entreprises dans leur transformation numérique. Ce programme apporte son soutien aux entreprises à travers des subventions, des consultations et des formations professionnelles.

Secteur de l'énergie

Dans le secteur de l'énergie, le Royaume-Uni vise à élaborer des recommandations du groupe de travail sur les données énergétiques et à proposer un **programme de modernisation de l'accès aux données énergétiques**. La stratégie allemande prévoit un financement stratégique pour le développement de centres de compétences dans les technologies de l'information et communication vertes ainsi que de super ordinateurs et de données pour l'analyse dans ce domaine.

---

[98] https://www.plattform-i40.de/IP/Navigation/EN/Manufacturing-X/Manufacturing-X.html.

[99] https://www.plattform-i40.de/IP/Redaktion/EN/Downloads/Publikation/PositionPaper-DataSpace.html.

[100] https://www.madesmarter.uk/.

[101] https://www.economie.gouv.fr/france-2030.

[102] https://smartindustry.nl/.

Le groupe de travail sur l'énergie et la durabilité de la Coalition néerlandaise pour l'intelligence artificielle a publié un document de position « L'IA comme accélérateur de la transition énergétique, opportunités pour un système énergétique sans CO2 »[103]. Le document définit les opportunités et les défis de l'IA dans la transition énergétique et fournit des recommandations pour la future utilisation de l'IA dans ce domaine.

Dans sa stratégie de l'IA, le gouvernement français prévoit un budget de 120 millions d'euros (dont 80 millions de financement public) pour accélérer le déploiement de l'IA au service de la transition écologique (par exemple pour la prévision de la production d'énergies renouvelables, la modélisation des choix de rénovation des bâtiments, et l'agriculture de précision).

Secteur de la mobilité

Dans le secteur de la mobilité, l'Allemagne a créé plusieurs **portails de données** tels que :

- mCloud[104] : une plateforme de données ouvertes offrant un accès gratuit à des données issues des secteurs de la mobilité, de l'espace et de la météorologie. La base de données est constamment mise à jour par le gouvernement fédéral allemand avec des données brutes et prend également en compte les données fournies par des fournisseurs privés du secteur de la mobilité. Elle s'adresse principalement aux utilisateurs du gouvernement, de la recherche et du secteur privé ;
- Mobility Data Marketplace (MDM)[105] : cette plateforme B2B offre aux fournisseurs et utilisateurs de données de mobilité un lieu pour partager, rechercher et s'abonner à des données en ligne pertinentes pour la circulation. La plateforme transmet les données fournies par les fournisseurs de données sans les modifier aux clients de données. Avec ses normes définies pour l'échange de données, la MDM est le plus grand volume d'informations sur la circulation routière, les embouteillages, les travaux routiers, les options de mobilité, les parkings et plus encore dans le pays.

Ces bases de données seront graduellement intégrées dans la nouvelle plateforme Mobilithek[106] afin d'être finalisée au milieu de l'année 2024.

Une autre initiative allemande est celle du **Mobility Data Space (MDS)**[107], une place de marché des données où des partenaires égaux du secteur de la mobilité peuvent échanger des données. Le MDS rassemble des entreprises, des organisations et des institutions et crée un cadre de coopération de confiance entre celles qui souhaitent monétiser leurs bases de données et celles qui ont besoin de données pour des solutions de mobilité innovantes.

Citons également l'initiative **Catena X**[108], qui a pour objectif de construire un échange de données collaboratif à l'échelle de l'industrie dans toute la chaîne de valeur mondiale de l'automobile. Rassemblant des entreprises, des acteurs clés de l'économie de la donnée ainsi que de l'édition de logiciels et des cas d'utilisation du réseau automobile, le hub vise à soutenir l'expansion du réseau mondial.

Dans sa stratégie de l'IA, le gouvernement français mentionne également un financement de 120 millions d'euros afin d'accélérer le déploiement d'une IA frugale dans des domaines clés de la transition écologique, notamment dans l'optimisation des réseaux de transports.

En décembre 2020, le ministère des infrastructures des Pays-Bas a publié un rapport intitulé « Smart Mobility, Dutch Reality »[109], qui prévoit des initiatives pour stimuler l'accès et le partage des données dans le secteur de la mobilité. Le rapport mentionne la nécessité d'améliorer l'infrastructure routière aux besoins du futur et d'améliorer la collaboration dans ce secteur. Ceci se fait entre autres à l'aide

---

[103] https://nlaic.com/wp-content/uploads/2021/09/Position_paper_AI_as_an_accelerator_for_the_energy_transition.pdf.

[104] https://www.mcloud.de/.

[105] https://www.mobility-data-space.de/en/the-mdm-in-the-mobility-data-space.html.

[106] https://mobilithek.info/.

[107] https://mobility-dataspace.eu/.

[108] https://catena-x.net/en/.

[109] https://www.rijksoverheid.nl/documenten/rapporten/2020/11/30/brochure-smart-mobility-in-beweging.

d'une **plateforme de collaboration en ligne**[110] pour les professionnels du secteur public et privé actifs dans les domaines de l'urbanisation intelligente et durable et de la mobilité.

En octobre 2023, le gouvernement britannique a publié un rapport sur **l'utilisation des données de localisation**[111] pour faciliter la mobilité connectée et automatisée. Le document aborde les défis et les opportunités liés à l'utilisation de ces données dans le cadre du développement de solutions de transport intelligentes et durables. Il met en lumière l'importance des données de localisation pour soutenir des initiatives telles que les véhicules autonomes, les services de transport partagé et d'autres innovations dans le domaine de la mobilité.

Secteur *de la santé*

Dans le cadre du programme de financement « Innovations numériques pour l'amélioration des soins de santé centrés sur le patient dans le système de santé », le gouvernement fédéral allemand a financé 22 projets pendant une durée maximale de 36 mois entre 2020 et 2023, pour un montant total de financement d'environ 50 millions d'euros. Les projets étudient les avantages des capteurs intelligents, de l'utilisation intelligente des données, des systèmes de support décisionnel intelligents et de la communication intelligente dans des scénarios d'application cliniquement pertinents. Les projets contribuent à démontrer les avantages potentiels des méthodes d'IA en combinaison avec l'utilisation et l'évaluation de grandes quantités de données pour les soins de santé.

En outre, la stratégie allemande mentionne d'autres initiatives dans le domaine de la santé telles que des approches numériques pour les prestations des soins de santé, des bases de données pour la santé, un centre de traitement de données de l'Institut allemand de documentation et d'informations médicales (DIMDI) et des projets pilotes servant à mesurer l'efficacité du système de santé allemand.

Au niveau législatif, le gouvernement allemand a récemment instauré de nouvelles réglementations concernant l'accès et l'utilisation des données dans le domaine de la santé. Ces règles comprennent une loi sur l'utilisation des données de santé, une loi sur les données de recherche et une loi sur les registres. La loi sur l'utilisation des données de santé établit un point central d'accès et de coordination des données, permettant ainsi aux chercheurs d'accéder à des données provenant de diverses sources telles que les registres du cancer ou les données des assurances maladie. De même, la loi sur les données de recherche a pour objectif d'améliorer l'accès aux données de traitement à des fins de recherche scientifique, tout en établissant un cadre uniforme pour leur diffusion, leur stockage et leur protection. Enfin, la loi sur les registres vise à renforcer les registres médicaux en introduisant un répertoire des registres qui offre une vue d'ensemble de l'inventaire des données, ainsi que de leur qualité et de leur disponibilité.

Le gouvernement des Pays-Bas a également lancé des initiatives pour mieux comprendre l'utilisation des applications basées sur l'IA dans le secteur des soins de santé. Une des dernières initiatives concerne le Centre d'expertise pour les algorithmes de santé. Il a lancé des algorithmes à base d'IA évolutifs dans quatre hôpitaux pour prédire si un patient sera admis à la salle d'attente de l'urgence ou autorisé à rentrer chez lui.

Au Royaume-Uni, l'Agence de Sécurité Sanitaire a publié une stratégie en matière de données[112] en septembre 2023. La stratégie de données établit 5 principes fondamentaux qui fixent les normes pour la gestion des données et la génération d'informations à travers l'agence : la responsabilité, l'accessibilité, la fiabilité, l'efficacité, et la collaboration.

En France, mentionnons la stratégie nationale du numérique de la e-santé en vigueur depuis 2019. Elle a pour but d'accélérer la sécurisation des échanges et du partage des données. Citons également

---

[110] https://dutchmobilityinnovations.com/landing-feb.

[111] https://www.gov.uk/government/publications/finding-the-way-forward-location-data-to-enable-connected-and-automated-mobility/finding-the-way-forward-location-data-to-enable-connected-and-automated-mobility.

[112] https://www.gov.uk/government/publications/ukhsa-data-strategy/ukhsa-data-strategy#strategy-governance.

les plateformes en ligne DREES[113] et Data.gouv[114] donnant accès à des bases de données ouvertes dans le domaine de la santé publique.

2.3. Collaborations et partages en matière de données

Encourager la collaboration

Les stratégies mettent en avant un large éventail d'initiatives visant à favoriser les collaborations entre les entreprises, les universités, les centres de recherche publics et le gouvernement. L'objectif de ces actions est d'encourager le développement de projets multidisciplinaires de recherche et d'innovation de pointe, d'exploiter pleinement les synergies des participants et de favoriser les transferts de connaissances.

En Allemagne plusieurs plateformes collaboratives dans le domaine du numérique et de l'IA ont été créées. Ceci inclut par exemple le Digital Hub Initiative[115] et la plateforme Industrie 4.0[116], deux plateformes qui stimulent les collaborations et les échanges d'expertise dans plusieurs domaines du numérique. La deuxième plateforme vise à créer un écosystème du numérique ouvert et flexible en misant sur trois piliers fondamentaux : l'autonomie, l'interopérabilité et la durabilité. Citons également la Plateforme Lernende Systeme[117], un réseau d'experts dans le domaine de l'IA qui stimule le dialogue social et interdisciplinaire. Dans le cadre de cette initiative, le gouvernement fédéral allemand a lancé une carte de l'IA[118] qui permet de découvrir les applications et projets innovants sur l'IA et d'identifier les différentes entités actives dans ce domaine (classifiées par type d'institutions et secteur de recherche).

Les Pays-Bas ont mis en place des **partenariats public-privé (PPP)** afin de promouvoir la collaboration dans le domaine du numérique et des données. Des initiatives telles que Commit2Data[119] et VWData[120], des programmes nationaux de recherche et d'innovation à base d'un PPP, stimulent davantage l'usage des données massives. Similaire à la carte de l'IA en Allemagne, les Pays-Bas ont créé une carte de bancs de test[121] dans plusieurs domaines du numérique. Cette carte donne un aperçu des partenariats public-privé dans les domaines de l'IA, la connectivité, l'exploitation de données pour relever les défis sociaux et stimuler la croissance économique, le gouvernement numérique, la résilience numérique, les compétences numériques et l'inclusion.

Pour sa part, le Royaume-Uni mentionne l'enjeux des différents instituts et centres de recherche et les conseils tels que l'institut Alan Turing, le centre national d'innovation pour les données, l'Open Data Institute, le conseil de l'IA et de la cyber sécurité, et la Data Skills Taskforce pour faciliter de nouvelles et meilleures collaborations entre l'industrie, le secteur public, les universités et les instituts.

Dans sa stratégie, la France mentionne la création de **4 instituts interdisciplinaires de l'IA**[122] (MIAI à Grenoble, 3IA Côte d'Azur à Nice, PRAIRIE à Paris, ANITI à Toulouse). Ces instituts ont un mode de financement très incitatif à la recherche partenariale et collaborative. Tous les centres ont noué des partenariats avec des dizaines d'entreprises de toute taille.

La **création d'espaces de données** est un autre élément clé pour encourager la collaboration. Ils fournissent aux utilisateurs des lieux de confiance partagés dans lesquels les données peuvent être fournies, analysées et gérées en collaboration. Les espaces de données sont créés à l'échelle européenne et nationale. En décembre 2023, la Commission a attribué un contrat de 41 millions

---

[113] https://data.drees.solidarites-sante.gouv.fr/pages/accueil/.

[114] https://www.data.gouv.fr/fr/organizations/sante-publique-france/#/presentation.

[115] https://www.de-hub.de/en/.

[116] https://www.plattform-i40.de/IP/Navigation/EN/Industrie40/Vision/vision.html.

[117] https://www.plattform-lernende-systeme.de/home-en.html.

[118] https://www.plattform-lernende-systeme.de/map-on-ai-map.html?FIT=1.

[119] https://commit2data.nl/en/.

[120] https://www.nwo.nl/en/projects/40017605.

[121] https://www.nederlanddigitaal.nl/initiatieven/kaart-met-proeftuinen-in-nederland.

[122] https://instituts-3ia.fr/.

d'euros pour le développement d'espaces européens communs de données[123]. Le consortium – composé d'entités belges, italiennes, néerlandaises et allemandes - développera « **Simpl** », une plateforme de logiciels intermédiaires sécurisés qui soutiendra l'accès aux données et l'interopérabilité entre les espaces européens des données. Des exemples d'espaces de données nationaux sont repris dans la section 2.2.

Il est également important d'éliminer tout obstacle qui puisse entraver le **flux international de données**. Dans ce contexte, la stratégie du Royaume-Uni mentionne les négociations en cours avec l'UE, les États-Unis, le Japon, l'Australie et la Nouvelle-Zélande pour créer des partenariats internationaux et stimuler la croissance économique grâce au flux transfrontière des données.

Au niveau international, tous les pays de cette analyse comparative mentionnent leur participation au **Partenariat mondial sur l'IA (GPAI)**[124], une initiative internationale visant à stimuler un développement et une utilisation responsables de l'IA dans le respect des droits de l'homme, de l'inclusion, de la diversité, de l'innovation et de la croissance économique. Le GPAI collabore avec des partenaires et des organisations internationales pour réunir des experts de l'industrie, de la société civile, des gouvernements et du monde universitaire. Un des groupes de travail de cette initiative se concentre sur la gouvernance de données afin d'analyser et de promouvoir l'utilisation, l'accès et le partage de données dans le domaine de l'IA.

Assurer la disponibilité et l'interopérabilité des données

La disponibilité de données de qualité et l'assurance de cette disponibilité sont des éléments importants pour stimuler l'économie des données. Dans ce contexte, le Royaume-Uni s'engage à établir un cadre stratégique plus clair afin de garantir un **meilleur accès aux données et une plus grande disponibilité et qualité des données** dans l'ensemble de l'économie afin de soutenir sa croissance et l'innovation. Ceci inclut la création d'une équipe centrale d'experts[125] capable d'assurer une interprétation cohérente du régime juridique autour du partage des données, de prévoir un cadre de qualité des données ainsi que d'évaluer la maturité des données pour l'administration publique. L'Allemagne s'engage à améliorer les conditions cadres pour les données à caractère personnel grâce à une meilleure protection des données d'importance nationale et régionale dans des domaines tels que la santé, les télécommunications[126] (voir également la section 2.5 sur les réglementations).

Le service cloud favorise également un accès amélioré aux données. Cette technologie permet une gestion plus efficace et flexible des informations et son accessibilité à distance facilite le partage et la collaboration. Dans ce contexte, le gouvernement français a lancé une stratégie d'accélération cloud[127] dotée de 1,8 milliard d'euros, dont 667 millions d'euros de financement public, 680 millions d'euros de cofinancements privés et 444 millions d'euros de financements européens.

Dans le **cadre de l'interopérabilité**, les quatre stratégies mentionnent le besoin de développer des normes universelles de qualité des données, de métadonnées et d'interprétabilité des données dans les différents centres de données et les plateformes en ligne. L'Allemagne s'engage par exemple à assurer l'interopérabilité avec les données de l'EOSC[128]. Le Royaume-Uni soutient également le développement de mesures d'interopérabilité et de cohérence lors du processus de publication de données ouvertes. Les Pays-Bas ont développé une stratégie pour le partage de données entre les entreprises[129] publiée en 2019. Le gouvernement se guide par trois principes fondamentaux. Premièrement, le partage de données se fait de préférence sur une base volontaire car les parties actives sur le marché sont souvent les mieux placées pour reconnaître et exploiter les opportunités

---

[123] https://digital-strategy.ec.europa.eu/en/news/commission-awards-eu41-million-contract-develop-infrastructure-common-european-data-spaces-0.

[124] https://gpai.ai/.

[125] D'autres stratégies comme celles de l'Espagne et l'Estonie ont mis en place des directeurs des données dans les administrations publiques afin de coordonner la disponibilité et l'interopérabilité des données dans l'ensemble de l'économie.

[126] https://gesetz-ttdsg.de/.

[127] https://www.entreprises.gouv.fr/fr/numerique/strategie-d-acceleration-cloud#industry.

[128] https://digital-strategy.ec.europa.eu/fr/policies/open-science-cloud.

[129] https://zoek.officielebekendmakingen.nl/blg-873881.pdf.

de partage de données. Deuxièmement, si nécessaire, le partage de données se fait de manière obligatoire (dans le cas d'un intérêt public par exemple). Finalement, les citoyens et les entreprises gardent le contrôle sur les données. Le partage de données, qu'il soit volontaire ou obligatoire, doit se faire de manière responsable, c'est-à-dire en respectant les droits et les intérêts des parties concernées en matière de confidentialité et d'informations commerciales confidentielles.

Une autre mesure pour améliorer l'interopérabilité et le partage des données est **l'anonymisation et la protection des données**. Dans ce domaine, l'Allemagne prévoit d'établir un réseau de recherche sur l'anonymisation, de financer la recherche sur les mesures de protection de la vie privée et les garanties de confidentialité et de mettre en place réseau de laboratoires d'essais et de certification qui effectuent des essais techniques pour déterminer la conformité des produits et services fondés sur la protection des données.

En règle générale, les quatre stratégies **adhèrent aux principes FAIR** (Findable, Accessible, Interoperable, Reusable) qui décrivent comment les données doivent être organisées pour être plus facilement accessibles, comprises, échangeables et réutilisables.

Créer une culture de données

Afin de promouvoir l'accès aux données et la collaboration de plusieurs parties prenantes, il est important de créer une culture de confiance. À cet effet, plusieurs pays envisagent le développement de fiducies de données. Ce sont des structures légales qui permettent à un fiduciaire d'assurer la gestion responsable d'une base de données au nom d'un bénéficiaire. Dans sa stratégie, le gouvernement allemand prévoit une campagne de financement de modèles de fiduciaires de données, ainsi que le développement d'un système d'accréditation ou de certification des fiduciaires en vertu de la loi sur la gouvernance des données.

L'Allemagne et le Royaume-Uni misent également sur plus d'équité et de transparence en renforçant les intérêts des citoyens dans l'économie des données. La stratégie du Royaume-Uni mentionne par exemple qu'une meilleure utilisation des données permettrait aux organismes de bienfaisance de concentrer leurs ressources sur la protection des parties les plus vulnérables de la société. L'Allemagne prévoit d'analyser comment la discrimination à l'égard des personnes résultant de l'analyse des données agrégées ou anonymisées peut être évitée. Un autre volet se concentre sur le rôle des technologies qui peuvent améliorer la protection de la vie privée et la confiance des consommateurs. L'Allemagne mentionne par exemple le projet CheckMyVA[130] qui a pour objectif de renforcer la souveraineté des données des utilisateurs d'assistants vocaux. Finalement, au niveau national, le gouvernement a révisé un règlement sur les systèmes de gestion de l'information à caractère personnel (PIMS) qui permet aux consommateurs d'exercer plus facilement leurs droits en matière de protection des données.

En France, la loi Lemaire vise à instaurer un service public des données digne de confiance en encourageant l'innovation et en bâtissant un cadre de confiance qui garantisse les droits des usagers tout en protégeant leurs données personnelles.

Une autre manière de créer une culture de données est la mise en place d'une **communauté de bonnes pratiques**. Avec la création du Centre de Confiance Numérique[131], le ministère de l'Économie des Pays-Bas s'engage à promouvoir la sécurité numérique des entreprises et à favoriser un échange d'informations efficace entre le gouvernement et le secteur privé en matière de cybersécurité. Le centre vise à aider les entrepreneurs à améliorer leur autonomie dans le domaine de la résilience numérique. Une résilience numérique accrue permet de garantir la sécurité, la confidentialité, la fiabilité et la transparence des données, ce qui contribue à accroître la confiance des utilisateurs dans leur utilisation. La stratégie du Royaume-Uni envisage une approche similaire par la mise en place d'une communauté de bonnes pratiques en matière de gestion des données, portée par des projets phares.

[130] https://www.fit.fraunhofer.de/en/business-areas/human-centered-engineering-and-design/HCED_Research/CheckMyVA.html.

[131] https://www.digitaltrustcenter.nl/.

Tableau 13. Aperçu des actions mentionnées dans les stratégies pour améliorer les compétences en matière de données

Actions	Pays
Encourager la collaboration	
Promouvoir les plateformes collaboratives dans le domaine du numérique et de l'IA	DE, UK, FR, NL
Stimuler des partenariats public-privé (PPP) dans le domaine du numérique et des données	DE, UK, FR, NL
Créer ou participer à des espaces communs de données à l'échelle européenne et nationale	DE, FR, NL
Mettre en place des partenariats internationaux pour stimuler la croissance économique grâce au flux transfrontalier des données.	DE, UK, FR, NL
Participer au Partenariat mondial sur l'IA (GPAI)	DE, UK, FR, NL
Assurer la disponibilité et l'interopérabilité des données	
Créer une équipe centrale d'experts (ou des directeurs des données) pour garantir la disponibilité et l'interopérabilité des données	DE, UK, FR, NL
Développer des normes universelles de qualité des données, de métadonnées et d'interprétabilité des données	DE, UK, FR, NL
Assurer l'anonymisation et la protection des données pour garantir une meilleure interopérabilité des données	DE, UK, FR, NL
Adhérer aux principes FAIR (Findable, Accessible, Interoperable, Reusable) des données	DE, UK, FR, NL
Créer une culture de données	
Développer des fiducies de données et des certifications de fiduciaires	DE, UK
Renforcer les intérêts des citoyens dans l'économie des données et combattre la discrimination	DE, UK, FR, NL
Mettre en place une communauté de bonnes pratiques portée par des projets phares ou des centres d'informations	UK, NL

2.4. Infrastructures de données

Développer des infrastructures (de stockage et d'analyse) de données fiables et sécurisées

L'un des piliers fondamentaux pour le bon fonctionnement et le développement d'une économie des données est la **création de centres sécurisés et durables axés sur le stockage et l'analyse de données**. Dans ce contexte, tous les pays de cette analyse comparative ont adopté une approche centralisée en créant des centres de données : le centre de données de recherche national (NFDI)[132] en Allemagne, le CASD Secure Data Hub[133] en France, l'Open Data Institute[134] au Royaume-Uni et le Centre of Excellence for Data Sharing and Cloud (CoE-DSC)[135] aux Pays-Bas.

Le NFDI donne accès à des stocks de données scientifiques et de recherche locaux grâce à une infrastructure cloud. Cette approche facilite le partage et l'interopérabilité de données dans plusieurs disciplines telles que les sciences de l'ingénieur, les sciences humaines et sociales, les sciences de la vie, et les sciences naturelles. Plusieurs groupes de travail ont été mis en place afin de garantir un bon fonctionnement de l'infrastructure. Entre autres, ces groupes se focalisent sur l'intégration des

[132] https://www.nfdi.de/?lang=en.

[133] https://www.casd.eu/le-centre-dacces-securise-aux-donnees-casd/le-casd/.

[134] https://theodi.org/.

[135] https://coe-dsc.nl/.

données, la planification de gestion des données, la gestion des identités et des accès, et la sécurité des infrastructures.

Le centre de données CASD en France met à disposition des données de l'INSEE, des ministères de la Justice, de l'Éducation nationale, de l'Agriculture et de l'alimentation, de l'Économie et des Finances. Ce centre prévoit aussi l'accès aux données dans le domaine de la santé et du privé (des entreprises mettant leurs données en accès externe à travers le centre CASD dans le cadre de collaboration avec des chercheurs, des start-ups ou des consultants). Dans le Royaume-Uni, l'ODI collabore avec le gouvernement et des entreprises pour construire une infrastructure de données, des connaissances et des stratégies qui crée la confiance dans les données et les pratiques de données. Finalement, le centre de données aux Pays-Bas accompagne les organisations à établir des mécanismes de confiance pour un partage de données sécurisé et fiable, leur permettant ainsi de saisir les opportunités commerciales offertes par cette pratique transformatrice. Sa mission est de libérer la valeur réelle du partage de données en soutenant les organisations dans la réalisation d'espaces de données évolutifs.

Favoriser l'accès aux données ouvertes en créant des **plateformes en ligne** est une des priorités de chacun des pays analysés dans cette étude comparative. En France par exemple, généraliser l'accès ouvert aux publications est un des axes mentionnés dans le deuxième plan national pour la science ouverte 2021-2024[136]. Le plan structure les actions en faveur de l'ouverture ou du partage des données à travers la création de la plateforme Recherche Data Gouv[137]. Ces initiatives nationales font partie d'une initiative européenne nommée European Open Science Cloud (EOSC)[138]. L'objectif de l'EOSC est de créer un environnement fiable et interconnecté à l'échelle européenne dans lequel les résultats de la recherche (publications, données, logiciels) peuvent être stockés et partagés.

Une autre initiative importante à l'échelle européenne est celle de **Gaia-X**[139]. Gaia-X vise à créer une infrastructure fédérée de données ouvertes basée sur les valeurs européennes en matière de données et de souveraineté numérique. La mission est de concevoir et de mettre en œuvre une architecture de partage de données qui consiste à créer des normes communes pour le partage de données, de meilleures pratiques, et de mécanismes de gouvernance. Gaia-X constitue également une fédération d'infrastructures et de services de données en nuage ancrée dans l'UE, à laquelle les 27 États membres se sont engagés[140].

Stimuler la connectivité numérique, le calcul intensif et l'informatique quantique

Les gouvernements mettent en place plusieurs initiatives servant à **améliorer la qualité et la capacité des infrastructures numériques**, indispensables pour les ressources d'analyse croissantes des nouvelles technologies de pointe telles que l'IA et le blockchain. Entre autres, les stratégies des pays analysés mentionnent leur soutien à l'accès au stockage sécurisé des données, au calcul intensif, aux logiciels de nouvelle génération pour l'analyse des données ainsi que le déploiement des **réseaux 5G et 6G** à haut débit.

Citons la collaboration du gouvernement fédéral allemand avec les États fédérés pour accélérer l'expansion du Gauss Centre for Supercomputing (GCS)[141] à la capacité exaflopique en tenant compte des besoins futurs d'analyses de données. Ce centre combine les capacités de calculs intensifs de trois autres institutions installées à Stuttgart, Jülich et Leibniz. Ensemble elles forment **la plus puissante infrastructure de calculs intensifs en Europe** et contribuent au déploiement d'activités industrielles et de recherche dans diverses disciplines. En 2020, la France a inauguré Jean Zay[142], un supercalculateur qui se situe dans le top 500 des superordinateurs les plus puissants au monde. Le Royaume-Uni mentionne ces différents supercalculateurs sur le site de HPC-UK[143], tandis que les

[136] https://www.ouvrirlascience.fr/deuxieme-plan-national-pour-la-science-ouverte/.

[137] https://recherche.data.gouv.fr/fr.

[138] https://eosc.eu/eosc-about/.

[139] https://gaia-x.eu/.

[140] https://digital-strategy.ec.europa.eu/en/news/towards-next-generation-cloud-europe.

[141] https://www.gauss-centre.eu/.

[142] https://www.cnrs.fr/fr/presse/jean-zay-le-supercalculateur-le-plus-puissant-de-france-pour-la-recherche.

[143] https://www.hpc-uk.ac.uk/facilities/.

Pays-Bas ont développé le supercalculateur Snellius[144]. Certains de ces centres de calculs intensifs font partie de l'entreprise commun européenne EuroHPC[145] qui sert à créer un écosystème de supercalculateurs en Europe.

Dans le domaine de **l'informatique quantique**, mentionnons l'initiative européenne Quantum Technologies[146] qui vise à soutenir le développement de systèmes d'informatique quantique de performance compétitive à l'échelle mondiale. Un des projets phares de cette initiative concerne OpenSuperQPlus[147], un système quantique basé dans le centre de recherche de Jülich en Allemagne. À l'échelle nationale, l'Allemagne prévoit également de préparer une feuille de route dédiée à l'informatique quantique et un réseau pilote de communication quantique dans le cadre du projet QuNET[148].

Renforcer la sécurité des données

Un autre aspect important de l'infrastructure est **d'assurer sa sécurité et sa résilience**. Dans le contexte de la libre circulation des données à caractère non personnel, le gouvernement allemand contribue au renforcement des droits des entreprises et des consommateurs à l'égard des services d'informatique en nuage en s'assurant des dispositions de stockage sécurisé.

Plusieurs pays européens ont également transposé la directive en matière de réseaux de systèmes d'information (SRI 2)[149] dans des lois nationales. En Allemagne, depuis 2021, la loi sur la sécurité informatique 2.0 renforce le niveau de sécurité des entreprises critiques désignées. Ces entités doivent détecter les attaques, signaler les incidents pertinents et fournir des informations à l'agence fédérale allemande qui gère la sécurité des technologies de l'information.

En France, les grands acteurs essentiels au bon fonctionnement de l'État (les OIV – Organisme d'importance vitale) sont aussi obligés de se soumettre à des mesures de cybersécurité et de signaler leurs incidents de sécurité à l'ANSSI (Agence nationale de la sécurité des systèmes d'information), l'autorité nationale en matière de sécurité numérique. Du côté des Pays-Bas, le gouvernement a mis en place un outil d'auto-évaluation. Les entreprises peuvent l'utiliser pour déterminer elles-mêmes si les dispositions de la loi sur la sécurité s'appliquent à elles.

Plusieurs pays ont également créé des **centres de cybersécurité**. Pour sa part, le Royaume-Uni a créé un centre national de cybersécurité (NCSC)[150] en 2016. Le NCSC fournit un point de contact unique pour les PME, les grandes organisations, les organismes gouvernementaux, le grand public et les ministères. Le centre collabore avec les forces de l'ordre, la défense, les agences de renseignement et de la sécurité du Royaume-Uni et d'autres partenaires internationaux. Des centres ayant les mêmes objectifs ont aussi été créés en Allemagne, France et aux Pays-Bas.[151]

Enfin, la plupart des pays mentionnent la nécessité de mettre à jour les **stratégies et programmes de cybersécurité et de sécurité informatique**, notamment pour le traitement de données.

---

[144] https://www.surf.nl/en/services/snellius-the-national-supercomputer.

[145] https://eurohpc-ju.europa.eu/index_en.

[146] https://digital-strategy.ec.europa.eu/en/policies/quantum-technologies-flagship.

[147] https://opensuperqplus.eu/.

[148] https://qunet-initiative.de/.

[149] https://digital-strategy.ec.europa.eu/fr/policies/nis2-directive.

[150] https://www.ncsc.gov.uk/.

[151] Allemagne : https://www.bmi.bund.de/EN/topics/it-internet-policy/ncc-de/ncc-de-node.html, France : https://cyber.gouv.fr/ et Pays-Bas : https://www.ncsc.nl/.

Tableau 14. Aperçu des actions mentionnées dans les stratégies pour créer des infrastructures sécurisées en matière de données

Actions	Pays
Développer des infrastructures (de stockage et d'analyse) de données fiables et sécurisées	
Développer des centres de données à l'échelle nationale de forme centralisée ou décentralisée	DE, UK, FR, NL
Créer des plateformes en ligne nationales afin de promouvoir une culture des données ouvertes	DE, UK, FR, NL
Participer aux initiatives européennes telles que le nuage européen pour la science ouverte (EOSC) et l'infrastructure fédérée de données ouvertes (Gaia-X)	DE, FR, NL
Stimuler la connectivité numérique, le calcul intensif et l'informatique quantique	
Déployer des réseaux informatiques de haut débit (5G, 6G)	DE, UK, FR, NL
Développer des centres de calculs intensifs à l'échelle nationale	DE, UK, FR, NL
Participer à l'initiative européenne de l'entreprise commune Euro HPC qui encourage le développement de centres de calculs intensifs	DE, FR, NL
Prévoir des programmes de soutien nationaux et européens dédiés à l'informatique quantique	DE, UK, FR, NL
Renforcer la sécurité des données	
Transposer la directive SRI 2 en matière de cybersécurité dans des lois nationales	DE, FR, NL
Créer un centre national de cybersécurité	DE, UK, FR, NL
Mettre à jour les stratégies et programmes de cybersécurité et de sécurité informatique	DE, UK, FR, NL

2.5. Réglementations

Développer un cadre juridique des données

Dans la dernière décennie, l'UE s'est engagée dans une dynamique législative pour créer un écosystème numérique qui protège les données tout en favorisant l'innovation et la croissance économique. La Commission européenne a mis en place un cadre juridique (voir Box 2 pour un aperçu des réformes législatives) pour favoriser l'adoption et le développement des technologies à base de données, tout en veillant à leur alignement avec les normes éthiques et de protection des données européennes. Les règlements sont directement applicables dans les 27 États membres de l'Union européenne tandis que les directives doivent être transposées dans des lois nationales.

Un des piliers de cette réforme législative est le règlement général sur la protection des données (RGDP) relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données. Il est applicable dans les 27 États membres à compter du 25 mai 2018. Le Royaume-Uni pour sa part a adopté le Digital Economy Act[152] en 2017. Cette loi introduit un cadre pour le partage de données personnelles dans le secteur public et a pour objectif de créer plus de cohérence dans la manière dont le secteur public partage les données personnelles, d'améliorer les services publics grâce à une meilleure utilisation des données et de garantir la confidentialité des données.

En outre, les pays européens appliquent le règlement européen relatif à la libre circulation des données à caractère non personnel qui vise à faciliter des nouveaux modèles commerciaux et à améliorer les services des administrations publiques. Ensemble avec le Data Act, le Data Governance Act, le Digital Markets Act et Digital Service Act, ce règlement définit le cadre juridique de l'économie

[152] https://www.legislation.gov.uk/ukpga/2017/30/contents/enacted.

complétive du numérique. À l'échelle nationale, le gouvernement allemand a adopté le German Competition Act[153] pour prévoir une réglementation conforme aux exigences de l'industrie en matière de numérisation. La stratégie allemande prévoit de lancer un projet de recherche afin d'étudier les conditions-cadres liées au droit de la concurrence pour les marchés des données.

Selon la Data Governance Act, chaque État membre est tenu de désigner un ou plusieurs organismes compétents pour s'acquitter des tâches énoncées dans l'article 7, ainsi qu'une ou plusieurs autorités compétentes chargées de réaliser les tâches relatives à la procédure de notification pour les services d'intermédiation de données (article 13). De plus, chaque État membre doit désigner un ou plusieurs organes compétents responsables du registre public national des organisations altruistes reconnues en matière de données (article 23). Selon un document de la Commission européenne[154], les Pays-Bas ont désigné le Bureau Central des Statistiques comme l'organisme compétent pour l'article 7, tandis que l'Autorité néerlandaise de la concurrence et des marchés a été désignée pour les articles 13 et 23. En Allemagne, c'est l'Agence fédérale allemande des réseaux qui a été chargée des articles 13 et 23, mais il manque encore un organisme compétent pour l'article 7. Pour sa part, la France n'a pas encore désigné d'autorité compétente pour l'un des articles.

Finalement, plusieurs pays mentionnent aussi la nécessité d'harmoniser les fondements juridiques relatifs à la protection et à l'utilisation des données au niveau régional et national. De plus, les Pays-Bas conseillent de développer une **législation sectorielle** comme proposé dans l'acte sur la gouvernance des données. Certains secteurs de l'économie sont déjà réglementés par des volets sectoriels du droit de l'Union qui comprennent des règles relatives au partage de données ou à l'accès aux données, au niveau transfrontalier ou à l'échelle de l'Union. Par exemple, la directive 2011/24/UE dans le contexte de l'espace européen des données de santé, et la législation pertinente en matière de transports, notamment la directive 2010/40/UE, le règlement (UE) 2019/1239 et le règlement (UE) 2020/1056, dans le contexte de l'espace européen des données relatives à la mobilité. Dans ce dernier secteur, certains pays ont adopté des règlements nationaux afin de faciliter les essais routiers publics impliquant des véhicules autonomes sous certaines conditions. Aux Pays-Bas, une loi qui permet l'expérimentation avec des voitures autonomes sur la voie publique est en vigueur depuis 2019. En France, les voitures autonomes de niveau 3 (semi-autonomes, avec la présence d'un conducteur) sont autorisées à circuler librement sur le territoire national. L'Autonomous Driving Act en Allemagne permet déjà la circulation de véhicules autonomes sans conducteur à condition qu'un humain puisse intervenir à distance. Finalement, au Royaume-Uni, la circulation de voitures sans conducteurs serait envisagée pour 2026.

Comme le marché des données touche tous les secteurs, le développement de législations adaptées en fonction des spécificités des secteurs est laborieux. En attendant d'avoir une législation par secteur et domaine d'application, plusieurs pays misent sur les **bacs à sable réglementaires**. Ce sont des outils réglementaires permettant à des entreprises sélectionnées de tester et d'expérimenter des produits ou services innovants dans un cadre réglementaire assoupli sous la supervision d'un régulateur pendant une période limitée. Le Royaume-Uni prévoit des bacs à sable réglementaires, dont certains dans le domaine du numérique et des données, à travers l'Information Commissioner's Office (ICO) et le Financial Conduct Authority (FCA). L'Allemagne a également développé des bacs à sable réglementaires et des bancs de test[155], notamment dans le domaine de la mobilité. De plus, l'AI Act encourage les États membres à établir des bacs à sable en matière d'intelligence artificielle pour faciliter le développement et les tests de systèmes d'IA innovants avant qu'ils ne soient commercialisés ou mis en service.

[153] https://www.gesetze-im-internet.de/englisch_gwb/.

[154] Les informations contenues dans le document ont été consultées le 15 février 2024 : https://ec.europa.eu/newsroom/dae/redirection/document/98966.

[155] https://bmdv.bund.de/EN/Topics/Digital-Matters/Digital-Test-Beds/digital-test-beds.html.

Box 2 : Réformes législatives européennes en matière de gouvernance de données

Afin d'atteindre les objectifs de la Décennie numérique de l'Europe et de promouvoir la construction d'une économie européenne de la donnée respectueuse des libertés et droits fondamentaux, la Commission a adopté plusieurs **propositions de règlement européen** sur les données :

Data Governance Act

L'un des piliers de la stratégie européenne pour les données est le **règlement européen sur la gouvernance des données**[156]. Ce dernier a été publié au Journal officiel de l'Union européenne le 3 juin 2022, est entré en vigueur le 23 juin 2022 et s'applique depuis septembre 2023.

Le règlement européen sur la gouvernance des données vise à accroître la disponibilité des données, les possibilités de les réutiliser et la confiance dans le partage des données. Ce règlement comporte :

- un encadrement ainsi qu'une assistance technique et juridique facilitant **la réutilisation de certaines catégories de données protégées du secteur public** (informations commerciales confidentielles, propriété intellectuelle, données personnelles) ;
- une **certification obligatoire** pour les fournisseurs de services d'intermédiation de données ;
- une **certification facultative** pour les organismes pratiquant l'altruisme en matière de données.

Data Act

Le 22 décembre 2022, la Commission a publié le **règlement européen sur les données**[157] fixant des règles harmonisées pour l'équité de l'accès aux données et de l'utilisation des données.

À ce titre, le Data Act a pour objectifs de :

- faciliter le **partage entre entreprises (B2B) et avec le consommateur (B2C) des données**, en fixant notamment une obligation de rendre accessibles les données générées par l'utilisation des objets connectés et services connexes, en contrepartie d'une compensation juste et équitable ;
- permettre **l'utilisation des données** détenues par les entreprises et, sous réserve de justifier d'un besoin exceptionnel, par les organismes publics des États membres et les institutions, agences ou organes de l'Union ;
- faciliter **le changement de fournisseur de services** de traitement de données (cloud et edge computing) par l'encadrement des relations contractuelles entre les fournisseurs de services et les consommateurs, et notamment par la suppression progressive des frais liés au changement pour le consommateur ;
- prévoir l'élaboration de **normes d'interopérabilité** pour les données et leurs réutilisations entre les secteurs ;
- mettre en place des **garanties contre les accès illicites de gouvernements de pays tiers** aux données non personnelles contenues dans le cloud.

AI Act

Le 21 avril 2021, la Commission européenne a publié sa proposition de **législation sur l'intelligence artificielle**[158], consacrant dans le droit de l'Union une définition des systèmes d'IA qui soit neutre d'un point de vue technologique et adoptant un ensemble de règles distinctes qui suivent une approche fondée sur le risque.

[156] https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=celex:32022R0868.

[157] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202302854.

[158] https://eur-lex.europa.eu/legal-content/FR/TXT/?qid=1623335154975&uri=CELEX:52021PC0206.

General Data Protection Regulation

Le **règlement général sur la protection des données**[159] (RGDP) renforce les droits existants, octroie de nouveaux droits et accorde aux personnes un meilleur contrôle sur les données à caractère personnel les concernant. Les points clés de cette législation comprennent[160] :

- Un **accès facilité aux données personnelles** d'un individu. Cela inclut de fournir davantage d'informations sur la manière dont ces données sont traitées et de veiller à ce que ces informations soient disponibles de manière claire et compréhensible.
- Un nouveau **droit à la portabilité des données**. Il facilite le transfert de données à caractère personnel entre prestataires de services.
- Un **droit d'effacement plus clair** (droit à l'oubli). Lorsqu'une personne ne souhaite plus que ses données soient traitées et qu'il n'existe pas de motif légitime de les conserver, les données seront effacées.
- Le **droit de savoir** quand des données à caractère personnel ont été violées. Les entreprises et les organisations doivent informer l'autorité de surveillance compétente en matière de protection des données ainsi que, en cas de violation grave des données, les personnes concernées.

Free flow of non-personal data

Le **règlement relatif à la libre circulation des données à caractère non personnel**[161] vise à éliminer les obstacles à la libre circulation des données à caractère non personnel entre les différents pays de l'UE et les systèmes informatiques en Europe.

Le règlement garantit :

- La **libre circulation des données à caractère non personnel** au-delà des frontières : chaque organisation devrait être en mesure de stocker et de traiter des données n'importe où dans l'UE ;
- La **disponibilité des données à des fins de contrôle réglementaire** : les autorités publiques conserveront l'accès aux données, même lorsqu'elles sont situées dans un autre pays de l'UE ou lorsqu'elles sont stockées ou traitées dans le cloud ;
- Un **portage de données plus facile** entre les fournisseurs de services cloud pour les utilisateurs professionnels.

Open data directive

La **directive sur les données ouvertes et la réutilisation des informations du secteur public**[162] (directive sur les données ouvertes) prévoit des règles communes pour un marché européen des données détenues par les pouvoirs publics. En bref, la philosophie de la législation sur les données ouvertes est de rendre les documents existants des organismes du secteur public disponibles afin qu'ils puissent être utilisés dans un cadre commercial ou non commercial.

[159] https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=celex%3A32016R0679.

[160] https://eur-lex.europa.eu/FR/legal-content/summary/general-data-protection-regulation-gdpr.html.

[161] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32018R1807.

[162] https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024.

Adopter des lignes directrices d'éthique

La pratique de collection, analyse et stockage de données – en particulier celles à caractère personnel – représente un fort enjeu d'éthique. La Commission a pris plusieurs mesures afin de veiller à ce que l'avenir du numérique et des données en Europe repose sur les valeurs européennes et les droits fondamentaux.

En avril 2019, le groupe d'experts de haut niveau sur l'intelligence artificielle a présenté des lignes directrices d'éthique pour une IA digne de confiance[163]. Le rapport qui forme la base pour ces lignes directrices propose une approche centrée sur l'humain en matière d'IA et énumère sept exigences clés auxquelles les systèmes d'IA devraient satisfaire afin d'être dignes de confiance : 1) action humaine et contrôle humain, 2) robustesse technique et sécurité, 3) respect de la vie privée et gouvernance des données, 4) transparence, 5) diversité, non-discrimination et équité, 6) bien-être sociétal et environnemental, et 7) responsabilité.

En 2021, la Direction Générale de la recherche et de l'innovation a publié un rapport sur l'éthique de protection des données. Ce document présente un aperçu des méthodes à suivre pour garantir une meilleure protection des données suivant le règlement général sur la protection des données. Enfin, en octobre 2022, la Commission a publié des lignes directrices d'éthique pour l'usage de l'IA et des données dans l'enseignement[164]. Les recommandations ciblent principalement les enseignants de l'enseignement primaire et secondaire.

Au niveau national, l'Allemagne a créé une **commission d'éthique de données**[165] afin de développer des lignes directrices dans ce domaine et de promouvoir un dialogue ouvert et inclusif sur des aspects d'éthique, techniques, légaux, et sociologiques de l'utilisation des données. La commission préconise une approche « éthique dès sa conception » en incorporant les principes éthiques à travers toutes les étapes de développement d'une application ou d'un produit basé sur des technologies des données, de l'IA ou du numérique en général. Ainsi, la commission présente des lignes directrices d'éthique[166] pour les véhicules autonomes visant à garantir la transparence, la vérifiabilité et la prévisibilité des systèmes d'IA. Une commission d'étude sur l'IA a également été mise en place. Dans son rapport final[167], la commission d'étude formule des recommandations d'actions : elle appelle, entre autres, au développement de régimes réglementaires sectoriels pour l'IA, tout en garantissant les principes de proportionnalité et de responsabilité.

Début 2020, un **comité national pilote d'éthique numérique** (CNPEN) a été créé en France pour aborder de manière globale les enjeux d'éthique du numérique. Au cours de la phase pilote, il a été chargé de 3 études initiales : le véhicule autonome[168] (avis paru en mai 2021), les agents conversationnels[169] (avis paru en novembre 2021) et le diagnostic médical et l'intelligence artificielle[170] (avis paru en janvier 2023). Plus récemment d'autres avis sur les enjeux d'éthique ont été publiés, notamment sur les plateformes de données de santé et des systèmes d'intelligence artificielle générative.

Le Royaume-Uni a plusieurs centres qui contribuent à l'éthique des données. Premièrement, le Centre du numérique et des données a publié un Cadre d'éthique des données[171] pour le secteur public. Dans sa stratégie des données, le gouvernement du Royaume-Uni s'engage à aider les spécialistes des

---

[163] https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai.

[164] https://op.europa.eu/en/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en.

[165] https://www.bmi.bund.de/EN/topics/it-internet-policy/data-ethics-commission/data-ethics-commission-node.html.

[166] https://bmdv.bund.de/SharedDocs/EN/publications/report-ethics-commission.html.

[167] https://www.bundestag.de/resource/blob/804184/f31eb697deef36fc271c0587e85e5b19/Kurzfassung-des-Gesamtberichts-englische-Uebersetzung-data.pdf.

[168] https://www.ccne-ethique.fr/publications/cnpen-le-vehicule-autonome-enjeux-dethique?taxo=56.

[169] https://www.ccne-ethique.fr/publications/cnpen-agents-conversationnels-enjeux-dethique?taxo=56.

[170] https://www.ccne-ethique.fr/publications/avis-141-du-ccne-et-4-du-cnpen-diagnostic-medical-et-intelligence-artificielle-enjeux?taxo=56.

[171] https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020.

données à développer des capacités durables en matière d'utilisation éthique des données en suivant les recommandations de ce cadre. Il s'engage aussi à diffuser les connaissances, les ressources et les études de cas à travers cette communauté de l'éthique des données. Deuxièmement, le Centre pour les données éthiques et l'innovation[172] développe des cadres de gouvernance, des orientations et des normes d'éthique qui permettent aux organisations d'utiliser l'IA et les données de manière fiable et responsable.

Dans leur stratégie du numérique, les Pays-Bas mentionnent plusieurs initiatives créant un cadre éthique. Les laboratoires ELSA (Ethical, Legal and Societal Aspects)[173] ont pour objectif de créer des cadres d'éthique pratique et des lignes directrices pour le développement de l'IA. Le ministère de l'Intérieur a également développé une boite à outils[174] pour soutenir les innovations éthiquement responsables. La boite à outils prévoit sept principes fondamentaux qui aident les gouvernements à innover sur le plan éthique. Des modèles, méthodologies ou des lignes directrices sont proposés pour chaque principe afin de guider les gouvernements dans le processus d'innovation.

Privilégier la standardisation des données

L'importante croissance des données ouvertes nécessite une harmonisation des pratiques afin de privilégier les formats standards et faciliter le partage et la réutilisation à long terme des données. Dans ce contexte, le gouvernement allemand propose plusieurs initiatives. En 2022, l'institut de standardisation allemand (DIN) et la commission allemande des technologies d'informations (DKE), en collaboration avec le ministère fédéral de l'Économie et de l'Action climatique (BMWK), ont publié la deuxième édition de la feuille de route allemande sur la standardisation de l'IA[175]. La feuille de route présente six recommandations d'action globales et plus de 100 besoins en matière de standardisation issus d'une participation active de plus de 570 experts de l'industrie, de la science, de la société civile et de la politique. Les résultats sont présentés par thème, tel que la sécurité, la mobilité, l'énergie et la médecine. Le gouvernement allemand propose également le financement d'experts, en particulier de PME et de start-ups afin de soutenir leur participation aux processus de standardisations internationaux.

Le ministère de l'Intérieur des Pays-Bas a mis en place une commission de conseil nommé Forum Standardisatie[176] qui regroupe des experts du gouvernement, du secteur privé et de la science. La plateforme en ligne présente les standards à suivre par le secteur public pour les structures de données, le développement de site internet, de protection des données et de protocole d'authentification.

En France, l'Observatoire Data Publica a publié un cahier sur la standardisation des données ouvertes en 2022. Ce cahier s'appuie sur plusieurs mois d'enquête auprès de concepteurs et de ré-utilisateurs de standards, et fait ressortir l'importance de la standardisation. Ce cahier retrace les origines de la standardisation des données, présente des études de cas (sur la transparence et la mobilité entre autres) et propose des pistes de réflexion et d'actions sur les défis de la conception et de la mise en œuvre de standards.

Enfin, le Royaume-Uni dévoile une stratégie de données[177] à travers l'office national des statistiques. Dans sa stratégie, l'office national des statistiques présente entre autres des principes de standardisation et de gouvernance de données. La gouvernance de données traite de sujets comme la protection, l'éthique, la qualité, le partage et le moissonnage de données. Les standards se rapportent aux formats des données et métadonnées et aux standards de partage et couplage de données.

---

[172] https://www.gov.uk/government/organisations/centre-for-data-ethics-and-innovation.

[173] https://nlaic.com/en/news/all-the-signs-are-in-place-for-elsa-labs-and-human-centred-ai/.

[174] https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/nieuwe-technologieen-data-en-ethiek/publieke-waarden/toolbox-voor-ethisch-verantwoorde-innovatie/over-de-toolbox-voor-ethisch-verantwoorde-innovatie/.

[175] https://www.din.de/en/innovation-and-research/artificial-intelligence/ai-roadmap.

[176] https://www.forumstandaardisatie.nl/.

[177] https://www.ons.gov.uk/aboutus/transparencyandgovernance/datastrategy/datastandards.

Tableau 15. Aperçu des actions mentionnées dans les stratégies pour développer une réglementation en matière de données

Actions	Pays
Développer un cadre juridique des données	
Soutenir et appliquer les réformes législatives européennes (règlements et directives) dans l'économie des données et du numérique	DE, UK, FR, NL
Appliquer une législation pour la protection des données à caractère personnel et soutenir la libre circulation de ces données	DE, UK, FR, NL
Appliquer une législation à la libre circulation des données à caractère non personnel qui vise à faciliter des nouveaux modèles commerciaux et à améliorer les services des administrations publiques	DE, FR, NL
Harmoniser les fondements juridiques relatifs à la protection et l'utilisation des données au niveau régional et national	DE, UK, FR, NL
Développer une législation sectorielle comme proposé dans l'acte sur la gouvernance des données	DE, FR, NL
Prévoir des bacs à sable réglementaires dans les domaines des données et du numérique adaptés aux spécificités des secteurs (santé, mobilité, énergie...)	DE, UK, FR, NL
Adopter des lignes directrices d'éthique	
Créer une commission ou un comité national d'éthique de données	DE, UK, FR, NL
Développer des lignes directrices d'éthique pour une économie des données et de l'IA digne de confiance	DE, UK, FR, NL
Prévoir un cadre d'éthique des données pour le secteur public	DE, UK, FR, NL
Privilégier la standardisation des données	
Développer des plateformes en ligne ou une stratégie des données présentant les standards à suivre, notamment pour le secteur public	FR, UK, NL
Publier une feuille de route sur la standardisation de l'IA avec des recommandations concrètes	DE
Participer aux processus de standardisations internationaux	DE, UK, FR, NL

3. Entretiens parties prenantes

Plusieurs entretiens ont été menés avec une sélection représentative d'acteurs clés du domaine des données. Ces entrevues ont permis d'explorer en profondeur diverses perspectives sur les sujets abordés dans ce rapport, notamment les compétences en matière de données, la valeur des données, la collaboration, l'infrastructure et la réglementation. L'objectif de ces discussions est de mieux comprendre le rôle du gouvernement et d'identifier les forces et les faiblesses de l'économie des données dans les pays de cette étude comparative. Cette approche a permis d'identifier les secteurs et domaines spécifiques de l'économie des données qui nécessitent un soutien particulier de la part du gouvernement.

3.1. Pays-Bas

Pour les Pays-Bas, Jaap Kerr, Simon de Jong et Geertjan Holtrop ont été interviewés. Ils travaillent tous en tant qu'agents de politique économique numérique au ministère des Affaires économiques et de la Politique climatique aux Pays-Bas. Jaap Kerr est principalement impliqué dans les négociations et la mise en œuvre des politiques européennes relatives aux données, notamment la décennie numérique de l'Europe, la gouvernance des données, et les espaces de données dans le cloud. Geert Holtrop a travaillé sur des questions telles que la régulation du libre flux des données, la gouvernance

des données et le partage des données entre entreprises. Il est également engagé dans l'élaboration d'une nouvelle stratégie sur l'économie des données, en mettant l'accent sur la standardisation de l'interopérabilité des données et sur les espaces de données, notamment en lien avec le portefeuille européen de l'identité numérique. Enfin, Simon de Jong travaille sur la mise en œuvre des négociations sur les données, la normalisation de l'interopérabilité des données et les espaces de données sectoriels visant à faciliter le partage de données entre entreprises.

Les personnes interrogées ont relevé différentes opportunités et défis en matière de politiques de données aux Pays-Bas, comme indiqué dans le Tableau 16.

Tableau 16. Opportunités et défis de l'économie de la donnée aux Pays-Bas

Opportunités	Défis
Les Pays-Bas ont mis en place de nombreuses initiatives de partage de données à travers des partenariats public-privé dans divers secteurs, comme l'agriculture et la logistique.	Il existe un défi relatif à la littératie et aux compétences en matière de données, notamment chez les PME, ce qui pourrait limiter leur participation à l'économie des données.
Les Pays-Bas disposent d'une excellente infrastructure de données et de télécommunications, ce qui est crucial pour le partage et l'exploitation efficace des données.	Les entreprises présentent différents niveaux de maturité en matière de transformation numérique et de données. L'un des défis consiste à fournir un soutien adapté et pratique aux entreprises à différents stades de maturité.
Les Pays-Bas et la Belgique partagent une culture similaire, ce qui constitue une opportunité pour une collaboration transfrontalière plus étroite en matière de partage de données.	Bien que plusieurs initiatives aient rencontré du succès, elles sont demeurées à une échelle restreinte. Le gouvernement pourrait jouer un rôle crucial dans l'expansion de ces initiatives pour qu'elles touchent un public plus large.

L'entretien révèle que les Pays-Bas ont établi une base solide pour le partage et l'exploitation des données grâce à des partenariats public-privé et des infrastructures robustes. Sur ce dernier aspect, Jaap Kerr a mis en avant la nécessité de promouvoir le développement de l'informatique quantique et de l'intelligence artificielle générative[178]. La discussion a également mis en évidence l'importance des partenariats, en particulier au niveau transfrontalier entre la Belgique et les Pays-Bas.

En ce qui concerne la réglementation, les Pays-Bas s'appuient principalement sur la législation de l'UE. Bien que les règles s'appliquent directement, des ajustements nationaux sont nécessaires, tels que la désignation d'autorités compétentes. Dans d'autres domaines, cependant, certains défis persistent, notamment en ce qui concerne les compétences en matière de données et les disparités de maturité des entreprises dans leur participation à l'économie des données. Certaines initiatives ont été mises en place pour améliorer les compétences et faciliter la transition des entreprises grâce à des lieux de travail axés sur les données, où des scientifiques des données peuvent soutenir les entreprises dans leur utilisation des données. Toutefois, ces initiatives demeurent souvent limitées et de petite envergure. Bien que les initiatives pilotes soient importantes, il est également crucial de les développer et de les déployer à une plus grande échelle. De plus, les intervenants externes plaident en faveur d'une meilleure connectivité entre les initiatives, ce qui peut être réalisé grâce à une collaboration plus étroite entre les ministères.

Au niveau sectoriel, les intervenants externes identifient les secteurs d'excellence où les entreprises néerlandaises et les centres de recherche se distinguent à l'échelle mondiale. Ces secteurs stratégiques[179] englobent l'agroalimentaire, la chimie, les industries créatives, l'énergie, les technologies de pointe, les sciences de la vie et de la santé, la logistique, l'horticulture, les matériaux de base, le domaine de l'eau et maritime, ainsi que les technologies de l'information et de la communication. Ils revêtent également une importance significative dans le contexte des données. Ainsi, la valorisation des données dans ces secteurs est particulièrement soutenue par le Fonds de

[178] Consultez la vision du gouvernement concernant l'intelligence artificielle générative : https://www.government.nl/documents/parliamentary-documents/2024/01/17/government-wide-vision-on-generative-ai-of-the-netherlands.

[179] https://www.rvo.nl/onderwerpen/innovatie-en-topsectoren/topsectoren.

Croissance[180], qui dispose d'un budget de 20 milliards d'euros pour la période 2021-2025. Une autre initiative pertinente, surtout dans le domaine de la santé, est celle de la fondation Go Fair[181], qui promeut les principes de disponibilité, d'accessibilité, d'interopérabilité et de réutilisabilité des données.

En conclusion, les intervenants soulignent la nécessité cruciale de développer des indicateurs pertinents pour évaluer l'efficacité des politiques de données nationales. Ils mettent également en avant l'importance des investissements visant à garantir l'adaptabilité et la croissance durable des initiatives, favorisant ainsi le développement de l'économie des données aux Pays-Bas. De plus, ils insistent sur l'importance de la cohérence entre les stratégies nationales et européennes en matière de données afin d'assurer une approche harmonisée et efficace dans ce domaine.

180 https://www.nationaalgroeifonds.nl/#:~:text=Met%20het%20Nationaal%20Groeifonds%20investeert,groei%20met%20positieve%20maatschappelijke%20effecten.

181 https://www.forumstandaardisatie.nl/vergaderingen/2021/onderzoek-naar-de-stimulering-van-fair-data-principes-bij-de-overheid.

Partie 3. Analyse SWOT et consultation des parties prenantes belges

Afin d'établir la proposition de stratégie fédérale de la donnée, une analyse SWOT a été réalisée et un ensemble de parties prenantes ont été consultées au travers d'ateliers collaboratifs.

Analyse SWOT

L'analyse SWOT présentée ci-dessous met en lumière les principaux atouts, faiblesses, opportunités et menaces de l'économie belge des données, en se concentrant sur les volets infrastructure, compétences, collaboration et valeur des données.

Pour établir cette matrice SWOT, une démarche en trois temps a été réalisée :

5. Extraction des principaux résultats du rapport sur l'économie belge des données
6. Interviews de parties prenantes ciblées belges spécialistes des données
7. Enquête en ligne permettant de valider et compléter les éléments récoltés lors des 2 premières étapes.

Forces

Valeur des données

Digitalisation élevée des grandes entreprises : le niveau élevé de digitalisation des grandes entreprises en Belgique améliore la qualité et la disponibilité des données pour les analyses.

Utilisation des données par les entreprises : les entreprises belges utilisent efficacement les données pour améliorer leurs processus et produits.

Compétences

Formations en STEM : les formations en sciences, technologies, ingénierie et mathématiques (STEM) de haute qualité en Belgique créent un vivier de talents compétents pour travailler avec les données.

Programmes de formation diversifiés : de nombreux programmes de formation impliquant les données sont disponibles, améliorant les compétences des travailleurs.

Collaborations et partage de données

Interopérabilité et normes : les discussions et initiatives sur l'interopérabilité et les normes de partage de données facilitent la collaboration entre acteurs.

Initiatives régionales et nationales : de nombreuses initiatives existent pour promouvoir la collaboration et le partage de données, tant au niveau régional que national.

Infrastructure de données

Accès internet à haut débit : une grande proportion de ménages belges (78 %) disposent d'un accès internet à haut débit, facilitant l'utilisation des services basés sur les données.

Faiblesses

Valeur des données

Quantité et qualité des données : la quantité et la qualité des données partagées restent insuffisantes, limitant ainsi leur utilisation efficace.

Investissements insuffisants : trop peu d'investissements sont réalisés dans les infrastructures de données, ce qui freine le développement de l'économie des données.

Compétences

Manque d'expertise : les PME manquent souvent d'expertise en matière de données, ce qui limite leur capacité à exploiter pleinement les données disponibles.

Manque de collaboration entre les organismes de formation au niveau national : le Rapport Digital Decade fait ressortir une faiblesse liée à un manque d'alignement des acteurs encadrant les formations digitales en Belgique

Collaborations et partage de données

Faible collaboration : la collaboration autour des données entre les différents acteurs est encore trop limitée, réduisant les synergies potentielles.

Infrastructure de données

Investissements dans les infrastructures : les investissements dans les infrastructures de données sont insuffisants, ce qui limite la capacité à traiter et à analyser les données de manière efficace.

Opportunités

Valeur des données

Augmentation de la digitalisation des PME : la croissance continue de la digitalisation des PME offre de nouvelles opportunités pour l'analyse et l'utilisation des données.

Réglementations européennes favorables : les réglementations européennes en faveur de la protection et de l'utilisation des données créent un environnement propice à l'innovation.

Compétences

Formation continue : le développement de programmes de formation continue peut améliorer les compétences des travailleurs et répondre à la demande croissante de spécialistes des données.

Collaborations et partage de données

Initiatives de collaboration : les initiatives visant à renforcer la collaboration entre acteurs publics et privés peuvent améliorer l'utilisation des données et favoriser l'innovation.

Infrastructure de données

Investissements dans les infrastructures : davantage d'investissements dans les infrastructures de données peuvent catalyser l'innovation et la croissance économique.

Menaces

Valeur des données

Concentration du marché des données : la concentration du marché des données entre les mains de quelques acteurs majeurs peut limiter la concurrence et l'innovation.

Cybercriminalité : la cybercriminalité pose une menace significative à la sécurité des données, pouvant entraîner des pertes financières et de réputation.

Compétences

Pénurie de compétences : la pénurie de professionnels qualifiés dans le domaine des données peut freiner le développement de l'économie des données.

Collaborations et partage de données

Manque d'interopérabilité : le manque d'interopérabilité entre les systèmes de données peut entraver la collaboration et l'efficacité.

Infrastructure de données

Risques de cybersécurité : les risques liés à la cybersécurité constituent une menace majeure pour les infrastructures de données.

Besoins en ressources : les besoins croissants en électricité et en eau pour les data centers peuvent poser des défis environnementaux et économiques.

Consultation des parties prenantes

Afin de confronter les premières conclusions concernant la proposition de stratégie fédérale de la donnée à la réalité de terrain, plusieurs ateliers incluant des acteurs sectoriels et des parties prenantes impliquées dans l'analyse des données belges ont été organisés et ont permis de faire ressortir des objectifs prioritaires à insérer dans le volet stratégique de l'étude et les actions devant figurer dans son volet opérationnel.

Cinq ateliers ont été mis en place afin de permettre à des acteurs représentatifs de leur secteur d'identifier collectivement les éléments clés qui devraient figurer dans la stratégie fédérale de la donnée :

- Quatre ateliers sectoriels :
  - Énergie
  - Mobilité
  - Santé
  - Services publics
- Un atelier de conclusion permettant de valider les éléments transversaux aux différents secteurs. Le tableau suivant présente les éléments clés qui sont ressortis du travail de consultation.

Tableau présentant les éléments clés transversaux issus des ateliers impliquant les parties prenantes

Eléments clés transversaux

COLLABORATION ET PARTAGE	Réglementation	INFRASTRUCTURE	VALEUR (F.A.I.R.)	COMPÉTENCES
Chief BE Data officer - facilitateur/coordinateur de toutes les initiatives relatives aux données de BE afin de stimuler le partage - Habilitation à faire respecter les normes et le partage des données essentielles - Suivi des actions - Financements (incitation)	**Renforcer le rôle de régulateur** de BE : Traduire, adapter et aligner rapidement toutes les réglementations (UE-BE-locales) afin de créer une sécurité juridique et de débloquer l'action.	**Plateformes centrales de partage de données** offrir un accès facile, sûr, sécurisé et puissant aux données (pour tous)	**Normes de données et référentiels**	**Talents Data/AI** des programmes de développement pour
Positionner BE dans les initiatives de l'UE (secteurs stratégiques)	**IMPOSEZ** partage et normes			**Maîtrise des données/de l'intelligence artificielle** programme
Booster les Data Spaces en tant que lieu de réalisation de cas d'utilisation en étroite collaboration **avec les** acteurs motivés de l'**industrie**	**LA NORMALISATION des données Gouvernance** clarifier les rôles (et simplifier)			
Plus de partage de données (plus facile, plus riche, ...) à partir des données disponibles	**Stratégie et objectifs clairs** (définition des priorités)			

Partie 4. Une proposition de stratégie fédérale pour la Belgique

La proposition de stratégie fédérale de la donnée s'appuie sur les différents éléments constitués en aval, à savoir le rapport sur l'économie belge des données, la comparaison des stratégies nationales des voisins de la Belgique, l'analyse SWOT validée par les parties prenantes et les ateliers collaboratifs.

1. Volet stratégique

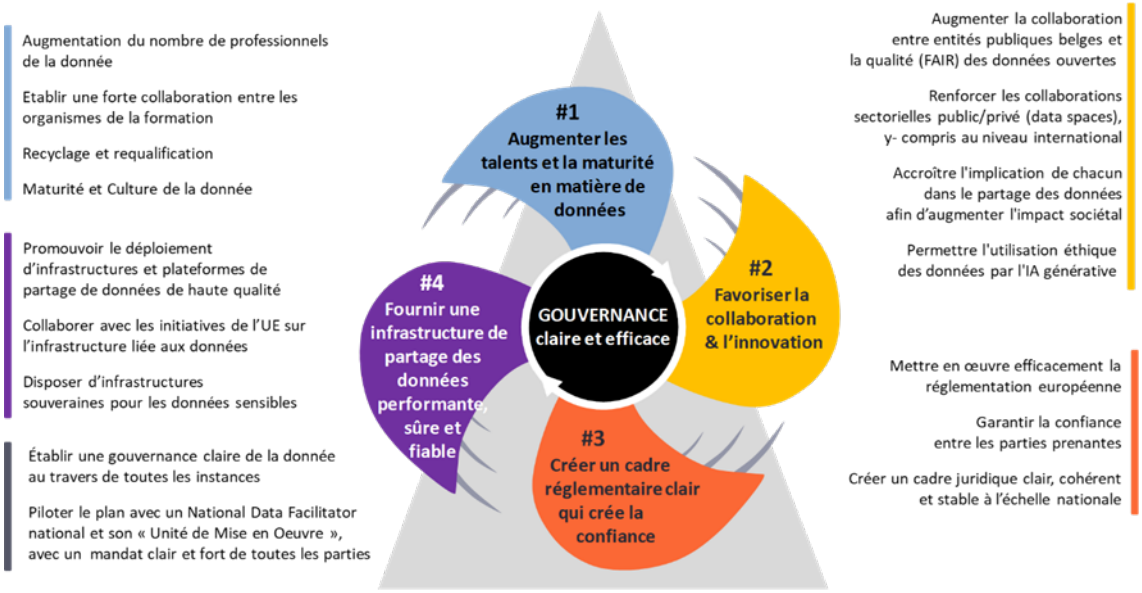
Objectif général : ***« Belgium 2030, towards a Data-Driven Economy »***

En 2030, la Belgique fait partie du TOP 5 des États européens en termes de croissance de la part de l'économie de la donnée par rapport à son PIB[182].

Le poids de l'économie des données dans le PIB belge était passé de 3,9 % en 2020 à 4,2 % en 2023, la Belgique se plaçant dans la moyenne européenne en devançant la France, avec ses 3,1 %, mais se trouvant derrière l'Allemagne, les Pays-Bas et le Royaume-Uni, qui affichaient respectivement 4,9 %, 5,4 % et 6,9 %. Les prévisions de croissance pour la Belgique sur cet indicateur étant inférieures à celles de ses voisins, il est important d'agir sur cet indicateur et d'infléchir la tendance future, sous peine de perte de compétitivité.

Afin d'activer, d'ici 2030, l'économie belge de la donnée et, au-delà, de générer un impact positif sur l'économie belge dans son ensemble et la compétitivité du territoire par rapport à ses voisins, un objectif transversal et quatre objectifs stratégiques ont été définis :

Figure 30. Illustration de l'objectif général et des 4 objectifs stratégiques



[182] Cfr supra 2.4. L'impact du marché des données sur l'économie belge.

Objectif transversal : délivrer des effets concrets grâce à une organisation efficace portée par une gouvernance centrale forte orientée résultats

L'urgence liée à l'accélération internationale de l'innovation autour de l'économie de la donnée, le caractère transversal de l'économie des données et la multiplicité des chantiers exigent de structurer une stratégie nationale efficace et coordonnée pour obtenir rapidement des résultats concrets, structurants, visibles et inspirants pour lancer la Belgique sur une trajectoire vertueuse, dans le haut du peloton européen.

Une gouvernance adaptée et un mandat clair

Afin de maximiser l'efficacité collective et l'impact sur tout le territoire, d'éviter toute lenteur ou difficulté pour des questions de coordination ou de prises de décisions, par ou entre certaines parties prenantes, il est important que l'opérationnalisation soit confiée à une structure centrale disposant d'un mandat clair. Ce mandat doit lui être fourni par les principales parties prenantes que représentent l'état fédéral, les entités fédérées et les partenaires sociaux et la société civile.

Un réseau élargi d'acteurs relayant la culture de la donnée et de la collaboration

La stratégie devrait idéalement bénéficier du support d'un réseau d'acteurs de terrain (entreprises, fédérations, associations, société civile, universités, économie sociale, etc.) qui seraient les ambassadeurs de la stratégie et partageraient la vision et l'ambition.

La question de la diffusion, avec l'aide de ce réseau, d'une culture de la donnée au travers de tous les pans de la société sera un élément-clé du succès de la stratégie.

Objectif stratégique #1 : augmenter le nombre de talents au service de l'économie de la donnée

Les talents sont au cœur de la transformation de notre société par les données. Au cours des trois dernières années, la Belgique a vu le nombre de ses talents actifs dans les métiers de la donnée croître moins vite que ses voisins. La moyenne de l'UE-27 concernant la proportion d'employés actifs dans les métiers de la donnée en comparaison avec l'emploi total est en augmentation, passant de 3,6 % en 2020 à 4,3 % en 2023. Dans la même période, la Belgique n'a pas suivi la tendance à la hausse observée dans d'autres pays européens[183]. En outre, la Belgique se positionne en dernière place de tous les pays de l'UE des 27 en termes de croissance du nombre de professionnels de la donnée au sein de l'Union européenne entre 2020 et 2023 (+3,9 % contre +17,8 % pour l'UE-27).

Par ailleurs, la Belgique enregistre l'évolution la plus défavorable de l'UE concernant l'indicateur d'intensité des professionnels des données dans les entreprises.

En 2022, les entreprises belges, tant fournisseuses qu'utilisatrices de données, avaient besoin d'environ 7.000 travailleurs de la donnée pour combler le déficit de compétences en données !

À l'heure du déploiement de l'intelligence artificielle générative (qui ne peut exister sans données) qui va bouleverser profondément les modèles économiques et l'organisation du travail au sein de toutes les organisations, l'activation des compétences liées à la donnée s'avère cruciale. Il importe, d'une part, non seulement d'encourager les entreprises et pouvoirs publics à attirer, former et retenir ces talents du monde de la donnée, mais aussi de les aider à y parvenir en faisant de la Belgique un pays de travail attractif pour ces profils. Il faut aussi aider nos organisations à doter l'ensemble de leurs collaborateurs de compétences nouvelles liées à la donnée. D'autre part, comme le recommande l'annexe du rapport Digital Decade 2030 portant sur la Belgique[184], il convient d'aligner dans des démarches coordonnées les actions des acteurs de la sensibilisation - depuis le plus jeune âge et tout au long de la carrière - et de la formation sur tout le territoire.

---

[183] Cfr supra, 3 Les compétences.

[184] Report on the state of the Digital Decade 2023- Annex Belgium https://ec.europa.eu/newsroom/dae/redirection/document/98647.

Résultats-clés attendus :

- Augmentation de la proportion d'employés travaillant dans le domaine des données afin de placer la Belgique dans la moyenne européenne. Concrètement, cela implique que la croissance des emplois liés à la donnée doit être supérieure à la croissance de l'emploi total.
  - L'augmentation pourra être obtenue tant par le recrutement et l'attraction de nouveaux professionnels spécialisés dans la donnée que par la formation (upskilling ou reskilling) de professionnels auxquels des compétences liées à la donnée seront adjointes aux compétences de base.
  - L'augmentation pourra être mesurée en s'appuyant sur l'indicateur d'IDC identifiant la part de l'emploi des professionnels de la donnée en pourcentage de l'emploi total.
- Augmentation du nombre de professionnels de la donnée :
  - L'augmentation pourra être mesurée en s'appuyant sur l'indicateur d'IDC identifiant l'évolution du nombre de professionnels de la donnée par entreprise utilisatrice.

Objectif stratégique #2 : faire de la Belgique un des territoires européens les plus dynamiques en termes de collaboration et d'innovation entre structures publiques et privées autour de la donnée

Capitaliser sur l'existant et corriger les points faibles connus

La Belgique bénéficie déjà d'un ensemble d'initiatives liées à l'ouverture et à la collaboration autour de la donnée et pourrait occuper une place de leader dans ce domaine au niveau européen. Il conviendra bien sûr de capitaliser sur ces initiatives déjà engagées.

Au niveau de l'ouverture des données, malgré des initiatives prises au niveau des entités fédérées dans le cadre de la transposition en loi belge de la directive sur l'ouverture des données, l'Inventaire des Données Ouvertes (ODIN)[185] classe la Belgique à la 72e place sur 194 pays. L'ODIN a fait six recommandations à la Belgique afin d'identifier et combler les lacunes de données, de publier plus de données désagrégées par sexe, d'utiliser des formats de données non propriétaires, d'améliorer les métadonnées, de fournir plus de données historiques et récentes et d'offrir des données plus détaillées au niveau sous-national. Ces recommandations restent pleinement d'actualité et doivent être suivies car elles permettent d'atteindre l'objectif d'augmentation de la collaboration autour de la donnée grâce à la mise à disposition des données selon le principe « FAIR » (Findable, Accessible, Interoperable, Reuseable)

Il convient d'**assurer la disponibilité et l'interopérabilité des données afin de permettre aux entreprises et aux décideurs politiques de prendre des décisions et d'établir des stratégies en s'appuyant sur ces données, notamment au moyen de l'intelligence artificielle. Cette interopérabilité doit également permettre aux acteurs de la recherche et de l'innovation de disposer, dans le cadre éthique proposé par l'Europe, de données d'entrainement pour leurs algorithmes et autres modèles de langage et éviter d'être distancés par les grands acteurs de l'intelligence artificielle générative.**

Dans le **cadre de l'interopérabilité**, à l'instar des stratégies développées dans les pays voisins[186], il importe de développer des normes universelles de qualité des données, de métadonnées et d'interprétabilité des données dans les différents centres de données et les plateformes en ligne.

Il pourrait être intéressant de s'inspirer des trois principes semblant conduire les stratégies de nos pays voisins :

- premièrement, le partage de données se fait de préférence sur une base volontaire car les parties actives sur le marché sont souvent les mieux placées pour reconnaître et exploiter les opportunités de partage de données ;

[185] Tableau 6. Classement de la Belgique en termes de couverture et ouverture de statistiques publiques.

[186] Cfr supra 2.3. Collaborations et partages en matière de données.

- deuxièmement, si nécessaire, le partage de données se fait de manière obligatoire (dans le cas d'un intérêt public par exemple) ;
- troisièmement, les citoyens et les entreprises gardent le contrôle sur les données. Le partage de données, qu'il soit volontaire ou obligatoire, doit se faire de manière responsable, c'est-à-dire en respectant les droits et les intérêts des parties concernées en matière de confidentialité et d'informations commerciales confidentielles.

Capitaliser sur les initiatives existantes sur le territoire et encourager la collaboration public-privé

Promouvoir les plateformes collaboratives dans le domaine du numérique et de l'IA

Afin de bénéficier des retombées positives de l'IA générative[187], il importera de pouvoir entraîner des modèles de langages, notamment disponibles en open source (ex. Llama), sur des données spécifiques contextuellement[188], pour lesquelles la collaboration sera cruciale.

Tout d'abord, il conviendra de définir les thématiques prioritaires sur lesquelles la Belgique souhaite concentrer ses efforts et collaborer. Ensuite, on pourra passer à l'opérationnalisation en s'appuyant notamment sur les « building blocks » proposés par l'Union européenne au travers notamment de la stratégie de données européennes et son support aux espaces de données et sur les initiatives présentes sur le territoire.

Stimuler et financer des partenariats public-privé (PPP) dans le domaine du numérique et des données

La capacité à faire collaborer les acteurs privés et publics autour des données est un facteur très important pour améliorer la génération de valeur à partir des données. Les pays voisins de la Belgique ont déjà mis en œuvre des partenariats publics-privés qui stimulent l'usage des big data[189]. Au niveau belge, plusieurs initiatives ont déjà émergé permettant la collaboration autour des données, comme, par exemple, les living labs (proeftuins) flamands autour des données dans le secteur de l'Industry 4.0, des médias ou de l'énergie ou en Wallonie des collaborations autour des Data Challenges. Les démarches entreprises autour du protocole SOLID en région flamande (SolidLab, Athumi, etc.) méritent également d'être analysées et, le cas échéant, servir de base à certaines plateformes de collaboration.

Stimuler la participation aux espaces communs de données à l'échelle européenne et nationale

La montée en maturité des espaces communs de données (European Common Data Spaces[190]) est une opportunité pour la Belgique. L'action de la Belgian Data Spaces Alliance[191] (BDSA), qui regroupe des acteurs tant publics que privés et académiques de tout le pays, est un socle intéressant pour structurer la démarche de manière collaborative. La structuration des acteurs en groupes de travail thématiques permet déjà de cadrer les actions en s'alignant sur les thématiques prioritaires de l'Europe.

Au niveau de la santé par exemple, la mise en place récente de l'European Health Data Space (EHDS), auquel la Belgique collabore[192], est un exemple très prometteur de canevas de collaboration pour les acteurs de la santé en assurant notamment l'interopérabilité et les échanges de données dans un cadre éthique et sécurisé.

---

[187] Cfr supra PARTIE 1, 2/ Les retombées futures potentielles de l'IA générative.

[188] https://cms.implementconsultinggroup.com/media/uploads/articles/2024/The-economic-opportunity-of-AI-in-Belgium/The-economic-opportunity-of-AI-in-Belgium.pdf p. 8 : Leveraging the full potential of AI will require further research, development and innovation.

[189] Cfr supra 2.3/ Collaborations et partages en matière de données.

[190] https://digital-strategy.ec.europa.eu/fr/policies/data-spaces.

[191] https://www.dataspacesalliance.be/.

[192] Par exemple au travers du Pilote EHDS2 portant sur les usages secondaires des données via Sciensano https://www.sciensano.be/en/projects/european-health-data-space-pilot-secondary-use-health-data.

Par ailleurs, la présence notable d'acteurs belges au niveau des initiatives européennes (notamment KULeuven au niveau du Data Space Support Centre et Eviden au niveau du middleware[193]) est une source d'inspiration pour l'intégration d'autres acteurs dans la dynamique.

Résultats-clés attendus :

- Amélioration de la disponibilité et de l'interopérabilité des données au niveau belge de manière transversale (interrégional/intergouvernemental et intersectoriel)

  Cette amélioration pourra être mesurée par l'inventaire ODIN.
- Au sein des secteurs prioritaires, mise en œuvre de partenariats public-privé de grande ampleur pour l'implémentation des solutions de collaboration autour des données
- European Common Data Spaces :
  - Implication belge dans les dynamiques européennes permettant la constitution des Data Spaces liés aux thématiques prioritaires.
  - Implémentation progressive, au sein des écosystèmes belges, des building blocks mis à disposition par les structures européennes de support aux Data Spaces.

Objectif stratégique #3 : offrir aux parties prenantes belges un cadre réglementaire clair et stable autour de la donnée

Après le règlement général sur la protection des données (RGPD), l'Europe a complété, dans le cadre de la Décennie Numérique, son arsenal juridique encadrant les données avec plusieurs règlements tels que l'Open Data Directive, le règlement européen 2018/1807 sur le libre flux des données à caractère non personnel, le Data Governance Act, le Digital Markets Act, le Digital Service Act et le Data Act.

Afin de faciliter l'exploitation des données et en extraire de la valeur, il importe que les États européens transposent et/ou adaptent leur législation afin de permettre l'implémentation des règlements sur le territoire.

La consultation des parties prenantes a fait ressortir le besoin d'une lisibilité des règles d'implémentation des règlements et d'une rapidité dans la transposition des nouveaux règlements européens. La durée très longue de l'implémentation du RGPD est le parfait contre-exemple de ce que souhaitent les acteurs de l'économie des données.

Deux lois récemment votées au Parlement fédéral[194] ont désigné, en lien avec le DGA, les SPF BOSA et SPF Economie respectivement comme autorités compétentes pour l'application du DGA et pour l'intermédiation des données et l'altruisme des données. Par ailleurs, le SPF BOSA est désigné comme intégrateur de services fédéral.

Par ces désignations, le législateur a confié la responsabilité au SPF Economie de jouer un rôle de contrôle de la bonne application des règles et de consultation des autorités sectorielles compétentes en particulier l'Autorité de protection des données, l'Autorité belge de la concurrence et le Centre de Cybersécurité Belgique, lorsque des questions qui relèvent de leurs compétences sont posées (par exemple protection des données, concurrence, sécurité). Le SPF BOSA a lui pour rôle, en tant qu'intégrateur fédéral, d'aider les autorités fédérales à remplir leurs obligations envers les citoyens et les entreprises et de permettre l'échange de données via les identités numériques, notamment au travers de la certification des attributs d'identité.

Ce cadre est précieux pour construire la stratégie d'activation de l'économie par la donnée. Il conviendra d'opérationnaliser le dispositif permettant aux citoyens, aux organismes sans but lucratif et aux entreprises d'identifier la procédure à suivre pour s'insérer dans le cadre du DGA.

---

[193] https://eviden.com/fr-fr/actualites/news/eviden-dirige-le-consortium-qui-fournira-a-la-commission-europeenne-le-middleware-simpl-open-pour-les-espaces-de-donnees-oss/.

[194] Cfr supra 0 DGA.

Rassurer les acteurs des initiatives collaboratives en s'appuyant sur la confiance dans le processus

La collaboration d'acteurs économiques dans des démarches d'innovation avec d'autres acteurs, par exemple dans le cadre des Data Spaces, nécessite une confiance vis-à-vis des autres partenaires mais aussi dans le processus mis en place.

Accompagner la Belgique dans le processus de négociation avec l'Europe pour les futures stratégies digitales européennes liées au sujet des données

Les stratégies digitales européennes actuelles (Digital Europe Programme (DEP), Digital Education Action Plan (DEAP)) se termineront en 2027 et les travaux de préparation des programmes et plans d'actions européens liés au digital et donc à la donnée seront initiés probablement en 2025 et 2026. À ce niveau, il est important que la Belgique puisse définir ses priorités de manière la plus coordonnée possible, par exemple au niveau de l'éducation au digital, afin que la voix belge soit entendue et ses priorités intégrées au processus.

Résultats-clés attendus :

- Implémentation effective et efficace des lois liées à la DGA :
  - Désignation d'intermédiaires belges de données dans la base de données européenne
  - Inscription de structures altruiste des données dans la base de données européenne dédiée

  Il importera, au-delà de ces inscriptions, de monitorer de quelle manière ces intermédiaires entrainent la circulation de la donnée entre les acteurs belges.
- Implémentation effective et efficace du DATA ACT dans les entreprises
  - Adhésion des entreprises aux principes du DATA ACT en matière de libération de la donnée et de respect des obligations visant à faciliter l'accès et le partage des données par les entreprises
  - Réutilisation effective des données libérées au bénéfice de l'économie
- Intégration des priorités belges dans les futurs programmes et plan d'actions européens après 2027

Objectif stratégique #4 : encourager la mise en place d'infrastructures de partage de la donnée sécurisée et performante

En particulier pour ses secteurs prioritaires, la Belgique doit se soucier de mettre en œuvre une stratégie garantissant la disponibilité, la sécurité du stockage, la rapidité de la circulation et du traitement des données.

La couche d'infrastructure comprenant les data centers, les réseaux, les machines connectées et l'IOT, etc. doit faire l'objet d'une attention particulière dans les objectifs d'une stratégie nationale de la donnée.

Dans un contexte géopolitique tendu tant au niveau de l'ingérence (cybersécurité) et de la disponibilité des composants clés (microprocesseurs, GPU, etc.) qu'au niveau des impacts générés par exemple par les data centers (consommation énergétique, consommation en eau, emprise sur les sols, etc.), il importera de mettre en œuvre un plan coordonné afin de permettre à l'économie de données de s'appuyer sur des bases solides.

Résultats-clés attendus :

- Disponibilité des couches d'infrastructure permettant la circulation des données en de bonnes conditions
- Identification des infrastructures de données prioritaires nécessitant une approche spécifique en termes de cybersécurité et de souveraineté
- Mise en place effective des infrastructures répondant à l'évolution de la consommation et production des données

2. Volet opérationnel

Préambule

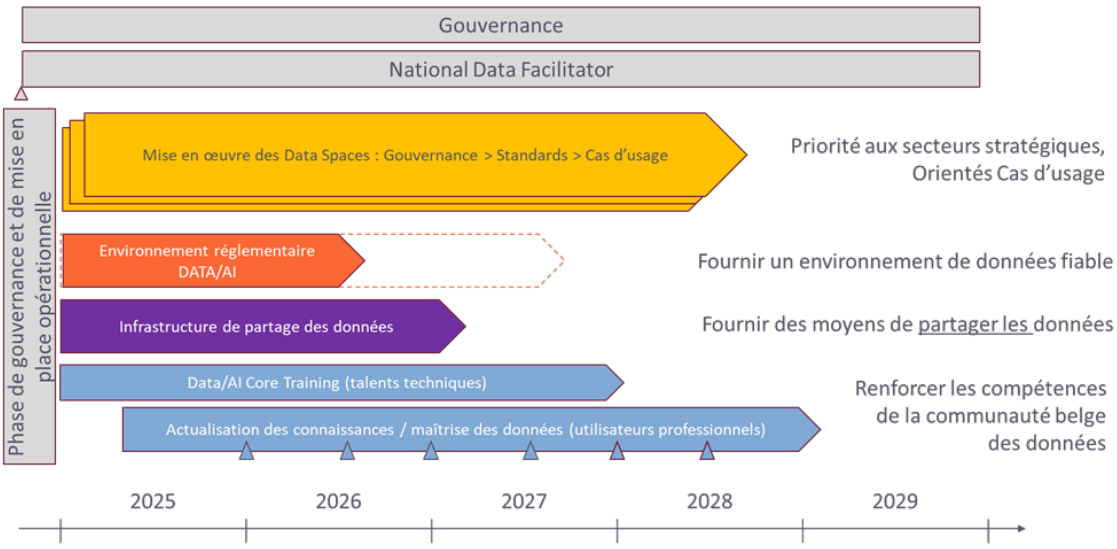
Les objectifs stratégiques tels qu'identifiés au paragraphe précédent requièrent la mise en place d'un plan d'actions centré sur les 5 points suivants :

- Mise en place d'une Delivery Unit pilotée par un National Data Facilitator
- Chantiers liés à la collaboration autour de la donnée
- Chantier de réglementation
- Chantiers pour la mise en place d'une infrastructure de données
- Chantiers « Formation » liés à la culture de la donnée et la formation à la donnée

Pour chacun des points, un horizon de réalisation est proposé. Il convient de garder à l'esprit que le monde des données évolue très vite et que tout retard sur la mise en place de solutions peut se traduire par une perte d'opportunités et de positionnement pour la Belgique.

La démarche se doit donc d'être réfléchie mais suivre un cadencement volontariste et miser sur le parallélisme des actions des intervenants.

31. Feuille de route globale du plan d'actions



2.1. National Data Facilitator (NDF)

Objectif

Les objectifs stratégiques à atteindre nécessitent la coopération de nombreux acteurs à la fois publics et privés. Cependant, de par la nature soit concurrentielle des activités ou plus simplement la diversité des objectifs et responsabilités de chacun, la mise en place de cette stratégie globale peut s'avérer compliquée.

De fait il apparait comme nécessaire de mettre en place un mécanisme de coordination afin de faire progresser les sujets identifiés.

Cette coordination se doit d'être neutre (pour éviter les conflits d'intérêts entre les différentes parties) tout en ayant des moyens d'actions au niveau fédéral.

Mission

Le National Data Facilitator (NDF) belge a pour mission de décliner les objectifs stratégiques établis pour la Belgique en plans d'action concrets avec un horizon de réalisation défini.

La mesure objective de la mission est celle de l'impact économique induit.

Il n'a pas de mission de réalisation à proprement parler mais un travail de facilitation/coordination des différents acteurs afin d'aboutir à un résultat précis dans un délai donné.

Le mandat conféré par les instances fédérales, en bonne intelligence avec les entités fédérées, lui permet de contacter les organismes, administrations et sociétés qu'il juge pertinents pour la réalisation de sa mission.

Il peut ainsi provoquer et suivre les groupes de travail représentatifs de chacun des verticaux considérés comme stratégiques (énergie, transport, santé, industrie, secteur public).

Il est le SPOC (single point of contact) entre les entités européennes, les entités fédérales et les sociétés belges pour les matières dont il a reçu le mandat.

Il assure la cohérence des travaux entre les différents chantiers. De par son rôle de SPOC, il facilite également la promotion des travaux belges au niveau européen et international.

Gouvernance

Le National Data Facilitator travaille sous les directives d'un groupe de suivi de niveau national qui pourra comprendre des entités politiques et économiques.

Le groupe de travail agira en tant qu'entité de contrôle. Le NDF fait valider à ce groupe ses actions et indique les progrès effectués.

Le groupe lui apporte le soutien nécessaire auprès des différents interlocuteurs (nationaux ou internationaux) et facilite la mise en place des budgets nécessaires pour la bonne réalisation de ses missions.

Comité consultatif

Le National Data Facilitator s'appuie sur un comité consultatif pour identifier, préparer et suivre les chantiers dont il a la responsabilité.

Le comité à un rôle consultatif et non pas coercitif.

Le comité peut contacter le groupe de suivi national (« Triumvirat »)

Le National Data Facilitator propose ses orientations, décrit ses progrès et difficultés.

Le comité de support :

- le conseille sur les choix ;
- l'aide à argumenter/justifier les choix ;
- l'aide à identifier les leviers nécessaires pour faire avancer sa mission.

Le comité consultatif est composé de :

- parties prenantes fédérales et fédérées ;
- parties prenantes privées ;
- parties prenantes académiques.

Moyens

Le National Data Facilitator peut engager des moyens financiers et humains pour l'atteinte de ses objectifs.

Il veille cependant à la bonne utilisation de ceux-ci et est garant de leur contribution à l'atteinte des objectifs stratégiques tels qu'identifiés précédemment.

Afin de garantir une exécution effective des actions, le National Data Facilitator sera idéalement entouré d'une équipe structurée sous la forme d'une *Delivery Unit*[195], à l'instar de celles mises en place à Genk (SALK[196]) ou Charleroi (CATCH[197]) suite aux fermetures respectives des usines Ford et Caterpillar. L'avantage d'une Delivery Unit est son caractère par définition limité dans le temps (typiquement 3 ans), son mode de management basé sur la réalisation concrète d'objectifs et sur le mandat clair reçu par la Delivery Unit pour avancer.

Dans cette équipe, le National Data Facilitator pourra organiser les travaux de manière globale et bénéficier du support de Delivery Officers, des spécialistes à même de porter les chantiers spécifiques avec les compétences adéquates.

Suivi des chantiers

Le National Data Facilitator est responsable des chantiers dont il assure la coordination.

À ce titre, il :

- identifie et fait valider les chantiers à lancer et leur séquencement ;
- identifie et réunit les personnes et entités devant y prendre part et notamment le/la responsable de chaque chantier ;
- valide le périmètre des travaux ainsi que le calendrier de livraison souhaitable (horizon) ;
- assure le suivi régulier ;
- mène toute action nécessaire en cas d'écart de trajectoire.

2.2. Chantiers liés à la collaboration et la standardisation de la donnée

Objectif

La collaboration autour de la donnée entre les acteurs publics et privés s'appuie sur une implémentation des Data Spaces. Le premier chantier portera sur les actions liées à l'accompagnement de la mise en place des Data Spaces. Le second chantier consiste en la mise en place de standards pour les données. Le troisième chantier porte sur la définition de Use Cases et leur implémentation effective.

Premier Chantier : l'implémentation des Data Spaces

Les Data Spaces[198] sont un cadre de collaboration autour de la donnée qui englobent, pour un secteur, des éléments structurants suivants :

- un pilotage des démarches par les parties prenantes
- une gouvernance de la donnée au niveau du secteur
- une définition des outils et des regroupements et du partage des données
- une intégration avec des données publiques ouvertes et personnelles

---

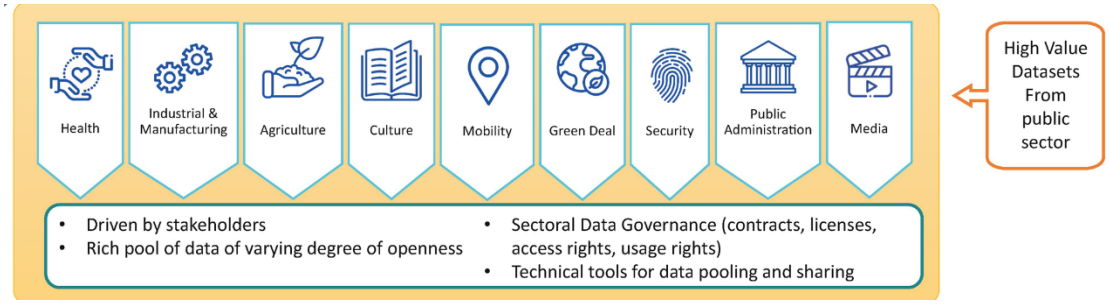
[195] Pour une définition de la Delivery Unit, par exemple : https://www2.deloitte.com/content/dam/Deloitte/us/Documents/public-sector/us-fed-what-is-a-delivery-unit.pdf.

[196] https://www.limburg.be/Limburg/landbouw/Kennis,-verduurzaming-en-innovatie-in-de-landbouw/SALK-projecten.html.

[197] https://charleroi-metropole.be/fr/2021/10/08/catch-turbo-axes-prioritaires-pour-la-transition-socio-economique-de-charleroi-metropole/.

[198] Voir les Common European Data Spaces : https://digital-strategy.ec.europa.eu/en/policies/data-spaces.

Figure 32. Common European Data Spaces



Source : https://link.springer.com/chapter/10.1007/978-3-030-98636-0_16/figures/1#ref-CR5.

Dans son rôle de facilitation et d'activation de la dynamique autour des données, le NDF aura pour but de faciliter le lancement de la mise en œuvre des Data Spaces.

Plusieurs actions seront définies dans cette mission :

1. Priorisation des Data Spaces

   Le NDF proposera au Comité de suivi de définir les priorités au niveau des secteurs pour lesquels il convient d'avancer.

   Les critères pris en compte pour cette priorisation sont :

   - La maturité de la gestion des données du secteur en Belgique et la volonté des acteurs à collaborer au Data Space de leur secteur ;
   - L'identification des chefs de file pour le pilotage du Data Space dans ce secteur ;
   - La qualité et la quantité des sets de données publiques (High Value datasets) utiles à ce secteur ;
   - La maturité du Data Space européen dédié à ce secteur (normes, outils d'interopérabilité) ;
   - L'impact économique et sociétal global potentiel de ce Data Space sectoriel.

2. Accompagnement de la gouvernance sectorielle (avec BDSA notamment) des Data Spaces

   Le NDF amènera son support à la mise en place de la gouvernance sectorielle des Data Spaces et en particulier s'assurera des éléments suivants :

   - Intégration des acteurs publics et privés pertinents dans les démarches
     - Ouverture des données et confiance
     - Collaboration inter-entités publiques
   - Mise en place de la gouvernance
   - Identification des données disponibles
   - Identification des besoins en infra et outils et adaptation éventuelle des normes

3. Soutien aux activités des Data Spaces après leur lancement

   - Financement, appels européens, collaboration avec les parties prenantes tierces, liens vers l'Europe
   - Évaluation de l'avancement et de l'impact

Deuxième chantier : standardisation

Il apparait qu'au sein d'un même secteur, plusieurs limites peuvent apparaitre sur l'échange des données :

- Format
- Signification
- Granularité
- Etc.

Ces points sont autant d'obstacles au partage de l'information et au développement des cas d'usages associés.

Les chantiers doivent mettre en avant un large éventail d'initiatives visant à favoriser les collaborations entre les entreprises, les universités, les centres de recherche publics et le gouvernement. L'objectif de ces actions est d'encourager le développement de projets multidisciplinaires de recherche et d'innovation de pointe et d'exploiter pleinement les synergies des participants et de favoriser les transferts de connaissances. L'objectif des chantiers de standardisation est dès lors de permettre l'émergence de standards pour les domaines concernés.

Ces standards porteront sur :

- Le modèle de données (concepts, granularité, etc.)
- Les mécanismes d'interrogation et d'échange des informations

De manière générale, il convient de s'inscrire dans le cadre d'une stratégie qui **adhère aux principes FAIR** (Findable, Accessible, Interoperable, Reusable) qui décrivent comment les données doivent être organisées pour être plus facilement accessibles, comprises, échangeables et réutilisables.

Dans l'idéal, il s'agira de standards en ligne avec les initiatives européennes ou internationales.

Il convient d'ailleurs de noter que l'objectif n'est pas nécessairement de créer des standards entièrement nouveaux. Des éléments peuvent exister :

- à partir de normes et standards européens
- au niveau des domaines

Les chantiers pourront également fonctionner sur le principe d'un mode agile avec des itérations successives afin d'accompagner la maturation des Data Spaces sectoriels correspondants.

La gouvernance de ce type de projet pourra suivre la méthodologie mise en place par exemple sur le projet ODETTE (www.odette.org) qui vise à définir des standards d'interopérabilité dans le monde de l'automobile et en particulier en Europe, les constructeurs partageant une base de fournisseurs et des intérêts communs.

On retrouvera comme points notables :

- La mise en place de différents comités suivant les aspects métiers et un comité technique chargé des considérations purement informatiques.
  Cette partie est à prendre en charge par le National Data Facilitator afin d'éviter la mise en place de normes disparates qui nuiraient à l'efficacité globale du projet. On pourra envisager l'adoption de normes de type SOLID (https://solidproject.org/) comme base technique pour l'ensemble des chantiers liés à l'infrastructure.
- La participation des différentes sociétés du consortium via des représentants chargés de trouver le meilleur compromis pour tous.
- Une gouvernance assurant la synchronisation des travaux.

Invitation des partenaires

L'administration fédérale belge est partie prenante dans la mise en place de standards mais doit s'appuyer sur les experts d'un domaine pour définir les standards.

De fait, différents acteurs sont appelés à collaborer pour la création des standards d'échanges requis à savoir :

- Les administrations et sociétés publiques
- Organismes universitaires
- Sociétés privées

Le National Data Facilitator a pour rôle :

- D'identifier les intervenants
- De valider l'utilité des projets et donc l'éventuelle pertinence de leur financement
- De définir le planning des travaux et de s'assurer que celui-ci est suivi

Troisième chantier : définition des Use Cases

Objectif

Le NDF, au-delà de ses prérogatives d'application (« empowerment »), a également un rôle de facilitateur auprès des différents acteurs de la data en Belgique. Il s'assurera que l'implémentation des Data Spaces soit effective grâce à l'identification de Use Cases concrets et exemplatifs.

Ces acteurs impliqués dans ces Use Cases comprennent bien évidemment les acteurs du domaine privé avec lesquels des structures de partenariat devront être mises en place. Par ailleurs, des fonds européens peuvent être identifiés pour les projets liés à ces Use Cases et il conviendra de structurer ces relations.

Partenariat Public Privé (PPP)

Les chantiers (régulation, standardisation, infrastructure...) ne doivent pas nécessairement être portés entièrement par le secteur public.

Certaines fonctions peuvent être portées par des acteurs privés pour :

- la définition de standards propres à un secteur
- la mise à disposition d'infrastructures liées à la fourniture de données.

Cette charge pourra être portée conjointement par le secteur public et privé avec les mécanismes de compensation/incitation/sanction adéquates.

L'objectif du projet étant le développement du secteur de la data en Belgique et à travers elle l'ensemble de l'économie du Royaume, une attention particulière pourra être portée à l'émergence de jeunes pousses (« start-ups ») à travers :

- des financements
- une facilité d'accès à des crédits
- tout autre mécanisme facilitant leur financement

Effet de levier européen

Les financements européens sont à traiter avec le plus grand soin afin de permettre l'avancement des projets.

Le National Data Facilitator suivra donc les différents appels de l'UE afin d'y répondre.

Actions de supports pour les chantiers

L'implication des différents acteurs est difficilement concevable sur la base du volontariat pur sauf à risquer de :

- servir les intérêts particuliers des participants
- se priver de contributions intéressantes

Le National Data Facilitator se doit de valider avec les acteurs identifiés les projets, actions à mener par ceux-ci. Dans ce cadre, il s'assure de la pertinence des projets et valide la gouvernance de ceux-ci et ce, de manière homogène quel que soit le domaine concerné.

En tant que facilitateur, le National Data Facilitator peut :

- Prendre en charge l'organisation des réunions (notamment pour les aspects logistiques)
- Faire acter au travers de PV les contributions aux groupes de travail
- Faciliter l'accès à certaines ressources/personnes et notamment pour la promotion/validation/interaction avec les entités internationales.

Identification & construction des outils communs

Les groupes de travail pourront s'appuyer sur des outils et mécanismes communs à la Belgique.

On pensera notamment à :

- Itsme, NIS, etc.
- Architecture de type SOLID
- Etc.

Le National Data Facilitator prendra en compte les besoins exprimés par les groupes de travail et cherchera la mutualisation des moyens et solutions afin d'obtenir un résultat cohérent.

Horizon

Les différents groupes de travail pour les domaines identifiés comme stratégiques devront être mis en place dès que possible.

Il conviendra de cibler une durée de réalisation d'environ un an.

2.3. Chantiers de réglementation

Objectif

Le développement d'un écosystème performant suppose qu'un cadre réglementaire à la fois stable et de confiance soit mis en place en Belgique.

Ces travaux s'inscrivent dans la lignée de ce qui est fait au niveau européen.

Les éléments d'implémentation des lois engagés sont actuellement :

DGA

- Désignation en mai 2024 des autorités et organismes compétents (SPF BOSA et SPF Economie).
- Le SPF BOSA a été désigné en mai 2024 comme intégrateur de services fédéral.
- Le SPF Economie a été désigné en mai 2024 comme autorité compétente pour les intermédiaires de données et les organisations altruistes de données. Des candidats sont prêts à solliciter une accréditation en tant qu'intermédiaires de données au sens du DGA.

Alignement sur les normes européennes

Il est rappelé que dans la dernière décennie, l'UE s'est engagée dans une dynamique législative pour créer un écosystème numérique qui protège les données tout en favorisant l'innovation et la croissance économique. La Commission européenne a mis en place un cadre juridique pour favoriser l'adoption et le développement des technologies à base de données, tout en veillant à leur alignement avec les normes éthiques et de protection des données européennes. Les pays européens appliquent le règlement européen relatif à la libre circulation des données à caractère non personnel qui vise à faciliter des nouveaux modèles commerciaux et à améliorer les services des administrations publiques. Ensemble avec le Data Act, le Data Governance Act, le Digital Markets Act et Digital Service Act, ce règlement définit le cadre juridique de l'économie complétive du numérique. Les bacs à sable réglementaires (Regulatory Sandboxes) tels que définis notamment dans le cadre de l'AI Act et déjà utilisés dans les pays voisins[199] seront implémentés lorsque cela permettra une facilitation des actions.

Le National Data Facilitator s'assure :

- De l'identification des éventuels freins à l'implémentation rapide et efficace du cadre juridique européen
- De la mise en œuvre d'actions permettant de lever les freins à l'implémentation des directives européennes aux spécificités du marché Belge

[199] Cfr supra Partie 2 - 2.5 Réglementations.

Horizon

La publication des normes est nécessaire à la mise en place d'un contexte de confiance propre au développement de l'écosystème.

Il conviendra donc de s'assurer que la régulation soit validée dès la mise en place des premières infrastructures d'échanges de données et suive ainsi ce planning.

L'horizon à considérer est donc d'un à deux ans maximum. Ce rythme rapide est nécessaire car préalable à la mise en place des actions suivantes.

2.4. Chantiers Infrastructure

Objectif

L'un des piliers fondamentaux pour le bon fonctionnement et le développement d'une économie des données est la création de centres sécurisés et durables axés sur le stockage et l'analyse de données. Dans ce contexte, tous les pays étudiés dans le présent rapport ont adopté une approche centralisée en créant des centres de données.

De fait, le gouvernement fédéral se doit d'agir en tant que facilitateur pour la création de telles infrastructures et ainsi permettre l'émergence d'une économie de la donnée et ses cas d'applications.

Modalités d'implémentation

La mise à disposition des informations au service de la communauté suppose un moyen unifié d'accès à l'information. Cependant, différentes technologies de type DataHub existent et peuvent s'appuyer à la fois sur des infrastructures techniques publiques et privées avec chacune leurs contraintes spécifiques.

Infrastructure publique

Une initiative importante à l'échelle européenne est celle de Gaia-X. Gaia-X vise à créer une infrastructure fédérée de données ouvertes basée sur les valeurs européennes en matière de données et de souveraineté numérique. La mission est de concevoir et de mettre en œuvre une architecture de partage de données qui consiste à créer des normes communes pour le partage de données, de meilleures pratiques, et de mécanismes de gouvernance. Gaia-X constitue également une fédération d'infrastructures et de services de données en nuage ancrée dans l'UE, à laquelle les 27 États membres se sont engagés.

Le gouvernement fédéral se doit d'avancer dans le cadre de cette initiative et si possible être parmi les plus rapides pour l'implémentation de telles solutions.

Infrastructure déléguée

Un accélérateur pour cette mise à disposition des informations peut être de se baser sur un partenariat public/privé.

Ainsi certaines sociétés majeures des secteurs envisagés disposent de quantité d'informations utiles pour d'autres partenaires/cas d'utilisation.

Deux éléments doivent alors pris en compte dans le cadre de cette réflexion :

- La duplication des informations vers un hub public n'est pas nécessairement une solution garantissant la fraicheur ni la complétude des données.
- Mettre à disposition les informations en 24/7 nécessite une infrastructure répondant à ces contraintes qui n'est pas gratuite.

Il en résulte le besoin de mettre en place des accords de type « délégation de service publique », impliquant à la fois un engagement des sociétés mais aussi une compensation pour l'usage des infrastructures mises à disposition par les sociétés privées.

Incitation/Sanction

La mise à disposition des données est un projet fédéral avec une ambition forte pour l'économie du royaume. L'implication des différents acteurs, la complétude et la fiabilité des informations sont essentielles pour le succès de l'initiative.

De fait un double mécanisme incitation/sanction devra être mis en place.

- Incitation : celles-ci pourront prendre différentes formes financières telles que crédits d'impôts, allègement de charge sur les personnels travaillant sur les projets ou toute autre solution motivant les sociétés à contribuer activement au projet.
  Les incitations devront être proportionnelles à la valeur ajoutée (de manière à éviter toute dérive) et reprendre ainsi les critères de valorisation de la donnée, à savoir la capacité de celles-ci à créer de la valeur par recombinaison. C'est cette singularité des données, capables d'être agrégées, réutilisées et partagées sans s'épuiser qui est en jeu ici mais il convient de favoriser les données les plus pertinentes aux autres utilisateurs.
- Sanctions. Il est rappelé que les données sont excluables, ce qui signifie qu'un agent économique peut empêcher sa consommation par un autre simplement en les gardant cachées (Buchanan, 1965). Toutefois, cela ne s'applique qu'aux données privées. Une fois qu'une donnée devient publiquement disponible, elle devient un bien public. Il convient donc de mettre en place des mécanismes de sanctions pour les entreprises ou organismes disposant d'information d'intérêt général mais cherchant à les conserver à des fins de contrôle de marché ou d'intérêt économique particulier.
  Le rôle du National Data Facilitator aura pour mission d'identifier ces données privatives et d'évaluer l'impact de leur rétention sur l'écosystème belge afin de quantifier les dommages et réparation induits.

Horizon

La mise en place des DataHub au niveau belge nécessite la collaboration d'un grand nombre d'acteurs. Il conviendra de :

- S'appuyer sur les travaux européens pour la définition de la stratégie et des règles d'implémentations
- Travailler avec les différents acteurs des domaines stratégiques pour identifier et planifier les travaux d'ouverture et de partage de l'information

La première étape (définition des techniques, outils, etc.) doit venir en préalable du déploiement par les partenaires pour lesquels les actions peuvent être menées en parallèle.

Il conviendra de valider sur certains pilotes le bon fonctionnement avant la généralisation l'ensemble des écosystèmes.

Cependant, il est souhaitable que cette mise en place soit réalisée sur une période d'environ 2 ans.

2.5. Chantiers Formation

Objectif

Il est rappelé que la croissance de demandes en compétences d'analyse de données sera plus rapide que d'autres compétences du numérique au cours des cinq prochaines années. Afin de combler un déficit potentiel et d'encourager le développement de l'industrie des données, les pays étudiés dans la présente analyse comparative prévoient des réformes visant à renforcer la littératie et les compétences en matière de données.

Pour rappel

- La Belgique affiche la plus faible proportion d'employés dans le domaine des données parmi ses pays voisins et se situe en dessous de la moyenne de l'UE-27.
- En 2023, elle occupe la 25e place sur 27 États membres, avec une part de professionnels des données représentant 3,1 % de l'ensemble de sa main-d'œuvre.

- Sur la période 2020 à 2023, la Belgique n'a pas enregistré de croissance significative de la part d'emploi des professionnels des données, contrairement à ses pays voisins et à la moyenne de l'UE-27, qui est en augmentation. Elle se positionne ainsi en dernière place dans l'UE en termes de croissance du nombre de professionnels des données

Les actions prévues dans les stratégies des pays étudiés comportent entre autres :

- Réformer le système d'éducation afin d'accroitre la connaissance des données ;
- Promouvoir les formations continues ;
- Analyser les besoins futurs du marché du travail.

Naturellement, le rôle du National Data Facilitator doit prendre en compte cette dimension afin de fournir à l'industrie de la donnée belge les compétences dont elle a besoin et ce, pour différents types de formation :

- Initiale
- Continue
- À destination du grand public (« data literacy »)

Coordination avec les organismes de formation

Le rôle de National Data Facilitator est également d'identifier les compétences nécessaires pour les différents postes à occuper.

À l'image du Royaume-Uni, le NDF a pour mission de mettre en place **une Data Skills Taskforce** afin d'agir comme un forum de partage des connaissances et des meilleures pratiques entre les principaux participants de l'industrie (et en l'occurrence sur les domaines considérés comme stratégiques) et de l'enseignement supérieur, et de promouvoir les compétences et l'analyse des données pour les besoins du futur.

Cette TaskForce fournira :

- Une matrice des compétences requises par métier et niveau d'engagement (de la simple « data literacy » au doctorant)
- Les syllabus requis
- Une évaluation du nombre de personnes à former avec les projections adéquates
- Une identification des centres d'excellences propres à travailler sur les innovations dont la Belgique aura besoin dans les années à venir.
  Il conviendra à ce titre de viser une efficacité opérationnelle en évitant un éparpillement des ressources comme en Allemagne où des centres d'excellence constituent un réseau d'échange d'expériences et d'exemples pratiques.

Mise en place des cursus

Formation initiale

Le contenu des programmes identifiés par la TaskForce seront pris en compte par les organismes de formation dépendant du gouvernement belge (universités).

La TaskForce, sous la gouvernance du National Data Facilitator, devra communiquer l'avancement de ses travaux à intervalle régulier afin de permettre :

- Le retour/les commentaires des organismes
- La prise en compte des objectifs afin de préparer les rentrées universitaires.

Formation continue

La mise en place des standards et datahub nécessite une montée en compétence des équipes des entreprises des domaines concernés.

Cette montée en compétence est réalisée par :

- Des organismes de formation s'alignant sur les cursus identifiés par la TaskForce

- Les organismes de formation initiale ouvrant des sessions particulières de formation (à l'image des programmes de type « Executive Training » de certaines écoles en France).

Sensibilitation et litéracie du grand public

La Belgique pourra suivre la stratégie du gouvernement allemand qui prévoit des mesures spécifiques afin d'améliorer les compétences en matière de données dans les organisations civiles. La stratégie mentionne plusieurs programmes de financements pour aider les organisations civiles qui enseignent les compétences en matière de données aux citoyens allemands, ainsi que la création d'une plate-forme intégrée pour la fourniture de données dans les secteurs à but non lucratif et ceux pour le bien commun.

Par ailleurs, il conviendra de prévoir l'intégration de programmes d'acquisition de compétences en matière de données dès le plus jeune âge ainsi qu'une formation de base et continue des enseignants en la matière, comme prévu dans d'autres pays européens.

De manière plus générale, le National Data Facilitator aura pour mission d'organiser une communication auprès du grand public afin d'expliciter pour tous les bénéfices du partage de données : la fourniture d'information pour la recherche médicale ou bien encore l'optimisation des transports en vue d'une réduction de l'empreinte carbone pourront être utilisés pour traduire en gains concrets la démarche.

Horizon

La formation d'un individu se place sur un temps relativement long. Il en résulte qu'afin que les personnes formées puissent contribuer à l'atteinte des objectifs en 2030, il conviendra de suivre les ordres de grandeur suivants :

- Travaux de la Taskforce et identification des centres d'excellence < 1 an
- Mise en place des programmes par les organismes de formation au cours de l'année suivante
- Arrivée sur le marché du travail des personnes en formation initiale dans les trois années qui suivront les travaux de la TaskForce avec une cible à l'horizon Juin 2028.

Partie 5. Conclusions et recommandations

Conclusions

La stratégie belge pour l'économie des données vise à renforcer la compétitivité du pays dans un secteur en pleine croissance et à favoriser l'innovation, la création de valeur et le bien-être social. Cette stratégie s'appuie sur une étude d'analyse de la situation actuelle et des perspectives futures de l'économie des données en Belgique, en se basant sur les sources suivantes :

- Une analyse de l'économie belge des données, qui a porté sur la valeur des données dans les organisations et pour l'économie belge, ainsi que sur les domaines des compétences, de la culture de la donnée, des infrastructures et des réglementations de l'économie belge des données.
- Une comparaison avec les pays voisins, la France, l'Allemagne, les Pays-Bas et le Royaume-Uni, qui ont également élaboré ou sont en train d'élaborer leurs propres stratégies liées aux données. La comparaison a permis de mettre en évidence un ensemble d'actions et d'initiatives au sein de ces pays qui peuvent inspirer une stratégie belge en matière de données.
- Une consultation des parties prenantes au travers d'interviews et d'une enquête en ligne qui a permis de recueillir les besoins, les attentes, les préoccupations et les suggestions des parties prenantes, ainsi que de valider la matrice SWOT.

Les principaux enseignements tirés de l'étude sont les suivants :

- La Belgique dispose de nombreux atouts pour s'appuyer sur les données afin de renforcer sa compétitivité à l'international.
- La Belgique fait face également à plusieurs défis, tels que le manque de coordination entre les niveaux de pouvoir, la complexité de sa structure institutionnelle et le déficit de compétences et de talents.
- La Belgique a des opportunités à saisir, telles que le développement du marché unique européen des données, la croissance de la demande de solutions basées sur les données, l'émergence de nouvelles technologies et de nouveaux modèles d'affaires, la création de synergies entre les secteurs et les régions.
- La Belgique doit également faire face à des menaces, telles que la concurrence internationale, la dépendance vis-à-vis des acteurs étrangers, les risques liés à la sécurité, à la vie privée et à l'éthique des données, les inégalités sociales et territoriales, et les impacts environnementaux de la numérisation.

Sur la base de ces enseignements, la stratégie belge pour l'économie des données définit les priorités et les actions principales qui seront conduites.

Ces priorités et actions sont regroupées en cinq axes stratégiques :

1. Renforcer la coordination et la coopération entre les niveaux de pouvoir et les acteurs de l'écosystème des données, en s'appuyant sur une gouvernance nationale pour l'économie des données, en implémentant de manière cohérente les cadres juridiques et réglementaires, en soutenant les initiatives transversales entre structures institutionnelles et les partenariats public-privé.
2. Développer les compétences et les talents nécessaires pour l'économie des données, en renforçant l'éducation et la formation initiale et continue, en attirant et en retenant les profils qualifiés, et en favorisant la diversité et l'inclusion dans le secteur des données.
3. Stimuler l'innovation et la création de valeur à partir des données, en soutenant la recherche et le développement, en facilitant l'accès et le partage des données, en accompagnant la transformation numérique des entreprises, et en promouvant l'entrepreneuriat et l'écosystème des start-ups.
4. Renforcer la confiance et la responsabilité dans l'économie des données, en garantissant la protection des données personnelles, en assurant la sécurité et la résilience des systèmes d'information, en appliquant les principes éthiques et les valeurs européennes, et en impliquant les citoyens dans la gouvernance des données.
5. Mettre en place une infrastructure adaptée aux ambitions liées à cette stratégie.

La mise en œuvre de la stratégie belge pour l'économie des données, sous la houlette du National Data Facilitator, requiert l'engagement et la mobilisation de toutes les parties prenantes, ainsi qu'un suivi et une évaluation réguliers des progrès réalisés. La stratégie belge pour l'économie des données sera adaptée aux évolutions du contexte et aux retours d'expérience. La stratégie belge pour l'économie des données vise à être un catalyseur pour le développement d'un secteur dynamique, compétitif, responsable et inclusif, au service de la croissance économique et du bien-être social du pays.

L'impact recherché par cette stratégie est d'augmenter, pour la période 2025-2030, le poids de l'économie des données dans le PIB belge. Celui-ci était passé de 3,9 % en 2020 à 4,2 % en 2023, la Belgique se plaçant dans la moyenne européenne en devançant la France, avec ses 3,1 %, mais se trouvant derrière l'Allemagne, les Pays-Bas et le Royaume-Uni, qui affichaient respectivement 4,9 %, 5,4 % et 6,9 %.

Les prévisions de croissance pour la Belgique sur cet indicateur étant inférieures à celles de ses voisins, il est important d'agir sur cet indicateur et d'infléchir la tendance future, sous peine de perte de compétitivité.

Recommandations

La mise en œuvre de la stratégie belge pour l'économie des données, sous la houlette du National Data Facilitator et sa Delivery Unit, requiert l'engagement et la mobilisation de toutes les parties prenantes, à commencer par celles des secteurs stratégiques, ainsi qu'un suivi et une évaluation réguliers des progrès réalisés. La stratégie belge pour l'économie des données sera adaptée aux évolutions du contexte et aux retours d'expérience. La stratégie belge pour l'économie des données vise à être un catalyseur pour le développement d'un secteur dynamique, compétitif, responsable et inclusif, au service de la croissance économique et du bien-être social du pays.

Pour atteindre l'ambition d'augmentation du poids de l'économie des données dans le PIB, il sera nécessaire de soutenir une dynamique proactive et axée sur la collaboration entre les parties prenantes institutionnelles et avec le privé.

Une condition préalable à la réussite de cette démarche consiste à intégrer dans la stratégie les différents niveaux de pouvoirs belges en mettant en avant l'intérêt commun, le caractère gagnant-gagnant d'une participation des entités fédérées à la dynamique nationale. En effet, au niveau des secteurs de la santé, de la mobilité et de l'énergie par exemple, les défis dépassent largement les frontières régionales et une coordination efficace entre tous les niveaux de pouvoir et entités (y compris les régulateurs) est primordiale pour y faire face avec efficacité, dans le respect des prérogatives de chacun. Une première recommandation porterait donc sur l'encadrement de la gouvernance du projet par un organe incluant les différentes parties prenantes institutionnelles fédérales et appartenant aux entités fédérées et la définition d'un mandat clair au National Data Facilitator, lequel rendra des comptes à cet organe de gouvernance. Il importera que les gouvernements s'engagent sur les objectifs que le National Data Facilitator et sa Delivery Unit devront atteindre.

Par ailleurs, une forte collaboration entre acteurs publics et privés sera nécessaire afin de garantir un alignement des visions et des standards ainsi qu'une pleine adoption aux démarches initiées. Pour y arriver, la Belgique devrait soutenir des dynamiques centrées sur la réalisation de cas d'usages concrets basés sur le partage de la donnée, dans lesquelles des acteurs désireux d'avancer s'engageront pleinement, à commencer par les secteurs stratégiques pour le développement de l'économie de Belgique. Nous recommandons que ceci se concrétise tout particulièrement via les Data Spaces qui s'inscrivent totalement dans cette dynamique collaborative orientée résultat et impact.

La Belgique accueille sur son territoire un vivier important de chercheurs et d'experts, tant académiques que privés, qui ont une reconnaissance internationale et dont les compétences pourraient être mises au service de la stratégie nationale de la donnée. La troisième recommandation porterait sur l'intégration des instituts de recherche spécialisés dans la donnée dans la gouvernance stratégique d'une part et sur la mise en œuvre d'un Comité d'experts, éventuellement sous le patronage des organes fédéraux de politique scientifique, d'autre part.

Enfin, il importera de doter cette stratégie des moyens de son ambition en permettant la mise en place d'une équipe de soutien (Delivery Unit) au National Data Facilitator et en finançant les actions de la stratégie nationale de la donnée. La **quatrième recommandation** porte ainsi sur la réservation d'un **budget permettant de conduire cette stratégie sur la période de la prochaine législature** en lançant les démarches dès 2025, après une mise en place rapide de la structure opérationnelle et de la gouvernance du projet.

SPF Economie, P.M.E., Classes moyennes et Energie

Rue du Progrès 50
1210 Bruxelles
N° d'entreprise : 0314.595.348
economie.fgov.be