



Berne, le 2 décembre 2022

Stratégie de la Confédération en matière de science des données

Bases communes, compétences et objectifs de l'administration fédérale en matière d'utilisation de la science des données

Table des matières

Résumé	3
1 Introduction	5
2 Pertinence, objectif et approche de la stratégie.....	7
2.1 Pertinence de la stratégie.....	7
2.2 Objectif de la stratégie.....	7
2.3 Approche de la stratégie	8
3 Fondement: langage commun et compréhension partagée	9
3.1 Importance du langage commun	9
3.2 Définition de la science des données	9
3.3 Importance de la science des données pour l'administration fédérale	10
3.3.1 Science des données centrée sur l'être humain («human-centric») et digne de confiance («trustworthy»)	10
3.3.2 Science des données pour le bien commun («public good» et «public policy»).....	11
4 Conception commune de la science des données au sein de l'administration fédérale	15
4.1 Vision	15
4.2 Principes fondamentaux.....	15
4.3 Ambition	16
5 Lignes directrices, leviers et mesures	17
5.1 Lignes directrices	17
5.1.1 Établir la confiance dans l'aide à la décision basée sur les données	17
5.1.2 Développer la conscience et les compétences.....	17
5.1.3 Améliorer l'accessibilité et la disponibilité techniques	18
5.1.4 Tirer profit des synergies et relever les défis ensemble	18
5.2 Leviers et mesures	18
6 Objectif «écosystème transversal de la science des données»	23
6.1 Vue d'ensemble.....	23
6.2 Description des facilitateurs de l'écosystème	23
6.2.1 Culture en matière de science des données	23
6.2.2 Talents et développement du personnel	23
6.2.3 Collaboration transversale.....	24
6.2.4 Outils et plateformes en science des données	25
6.2.5 Gestion et infrastructure des données.....	26

Résumé

Au fil de l'évolution technologique, divers problèmes sont devenus plus complexes, moins structurés et plus riches en données, mais aussi plus éphémères. La résolution de ces problèmes par les sphères politiques et administratives devient d'autant plus exigeante. La science des données s'avère être un outil approprié tant pour relever ces défis au niveau opérationnel que pour soutenir les décisions politiques. Elle aide à mieux comprendre les problèmes et les solutions possibles dans un tel environnement non structuré, riche en données et éphémère. Ce faisant, la science des données complète les instruments existants de recherche de solutions et d'aide à la décision utilisés par les sphères politiques et administratives pour que celles-ci soient en mesure de remplir leurs tâches légales de manière aussi efficace qu'efficiente au bénéfice de la communauté.

L'objectif de développer des compétences et de graduellement concevoir une nouvelle discipline à cet effet a été fixé dans la «stratégie numérique de la Confédération 2020-2023» (initiative stratégique «Science des données» IS 6). La présente stratégie de la Confédération en matière de science des données (DSStB) répond à cette demande. La DSStB doit démontrer de quelle manière l'administration fédérale pourra utiliser la science des données, quels seront alors ses objectifs et quelles mesures seront nécessaires en vue d'une utilisation ciblée et coordonnée.

Il est essentiel de définir un langage commun et des notions centrales dans la perspective de développer la science des données au sein de l'administration fédérale. La DSStB définit par conséquent la notion de «science des données» et souligne l'importance de la science des données pour le bien commun dans le contexte du processus d'élaboration des politiques («policy making»). Cette approche doit contribuer à la conception commune que l'administration fédérale se fait de la science des données, une conception essentielle pour le développement de toutes les activités de science des données au sein de cette dernière. Dans le cadre de la DSStB, cette conception commune se fonde sur une vision, des principes fondamentaux et une ambition, définis comme suit:

Vision: «Une science des données centrée sur l'humain et digne de confiance qui soutient le bien commun ainsi que les politiques publiques».

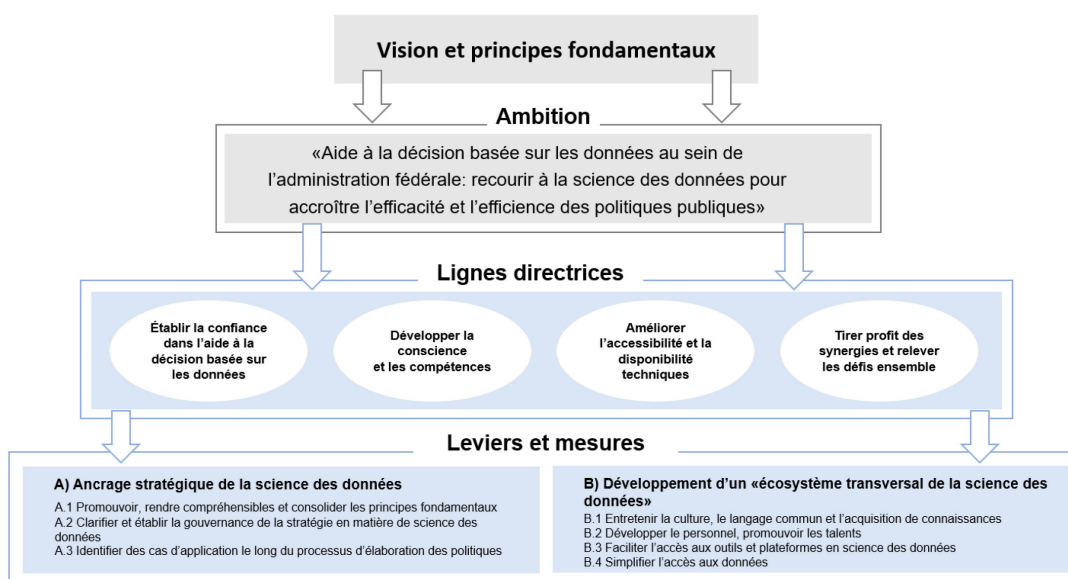
Principes fondamentaux: sécurité de l'information, protection des données et de l'information, sécurité des données, gouvernance des données, non-discrimination, explicabilité, traçabilité, transparence, reproductibilité, neutralité, objectivité et traitement éthique des données et des résultats.

Ambition: «Aide à la décision basée sur les données au sein de l'administration fédérale: recourir à la science des données pour accroître l'efficacité et l'efficience des politiques publiques».

Des lignes directrices, leviers et mesures stratégiques ont été formulés en vue de la mise en œuvre de la vision, des principes fondamentaux et de l'ambition. Le premier levier – ainsi que les mesures concrètes correspondantes – vise à l'ancrage stratégique de la science des données au sein de l'administration fédérale. Le second a pour but

de promouvoir l'«**écosystème transversal de la science des données**». Cet écosystème énonce le principe directeur de l'administration fédérale sous forme de système coordonné de valeur ajoutée. L'écosystème consolide l'ensemble des innovations, améliorations et champs d'application en lien avec la science des données. Il a aussi pour but d'optimiser les avantages que l'on peut en tirer via cinq facilitateurs centraux. Ces facilitateurs couvrent les domaines: talents et développement du personnel, culture, collaboration transversale, outils et plateformes en science des données, ainsi que gestion et infrastructure des données.

La vue d'ensemble ci-dessous expose les fondements de la DSStB et les mesures en découlant:



1 Introduction

Champ d'action prioritaire en matière de transformation numérique, la science des données présente un potentiel considérable pour l'administration publique. La science des données n'est plus vraiment une *terra incognita* au sein de l'administration fédérale: en effet, nombre de ses unités utilisent déjà des applications de cette discipline. La science des données trouve application dans tous les domaines, de l'épidémiologie à l'aménagement du territoire en passant par la détection des fraudes. Les exemples d'application couvrent d'ailleurs toute l'étendue des activités de l'administration fédérale.

L'administration fédérale a pris conscience de l'importance d'orienter stratégiquement et de coordonner le recours à la science des données. L'objectif correspondant est formulé dans la «Stratégie numérique de la Confédération 2020-2023»¹, sous l'initiative stratégique «Science des données» (IS 6) menée par l'Office fédéral de la statistique (OFS) et la Chancellerie fédérale (ChF / TNI): *«Informations et données essentielles à l'accomplissement des tâches administratives: bâtir des compétences en matière de science des données, et en faire une nouvelle discipline spécialisée»*. L'initiative vise notamment à élaborer une «stratégie de la Confédération en matière de science des données» (DSStB) ainsi qu'un «concept de formation en science des données», deux facteurs qui améliorent grandement les conditions-cadres au profit de la transformation numérique de l'administration fédérale.

En mai 2020, le Conseil fédéral a confié à l'OFS le mandat de mettre en place un Centre de compétences en science des données (DSCC)². Depuis janvier 2021, le DSCC met son savoir-faire dans le domaine de la science des données au service des administrations suisses (Confédération, cantons et communes), selon le principe de la «science des données en tant que service». Souhaitant en outre promouvoir l'intelligence artificielle (IA), le Conseil fédéral a décidé fin août 2021 de créer un «Réseau de compétences en intelligence artificielle» (CNAI)³. Le secrétariat en charge de ce réseau est rattaché à l'OFS. Opérationnel depuis juillet 2022, il propose aujourd'hui des offres de base. Plusieurs projets en IA sont actuellement en cours dans diverses unités de l'administration⁴. Quel que soit le département, le recours à l'IA présente notamment un potentiel intéressant pour l'automatisation et la rationalisation de tâches à caractère répétitif impliquant des volumes de données considérables. Toutefois, les méthodes issues de l'apprentissage automatique (machine learning) et de l'intelligence artificielle ne représentent qu'une partie des méthodes provenant de la science des données et utilisées dans le processus de résolution de problèmes lors de l'analyse de données de masse pour des domaines définis⁵.

La présente DSStB répond au mandat de l'initiative stratégique «Science des données» (IS 6) de la «Stratégie numérique de la Confédération 2020-2023». Elle soutient et encourage les projets en cours, qu'il s'agisse de projets interdépartementaux ou de

¹ [Stratégie numérique de la Confédération 2020-2023](#)

² [Centre de compétences en science des données \(DSCC\)](#)

³ [Réseau de compétences en intelligence artificielle \(CNAI\)](#)

⁴ La [banque de données des projets du CNAI](#) offre en toute transparence un aperçu des projets recourant à l'IA au sein de l'administration fédérale.

⁵ Les domaines dans lesquels l'administration fédérale voit un grand potentiel pour l'IA sont les suivants: reconnaissance textuelle («text recognition»), reconnaissance vocale («speech recognition»), reconnaissance d'images («image recognition»), détection de fraudes («fraud detection»), contrôles de plausibilité et validation des données («plausibility checks and data validation») et maintenance prévisionnelle («predictive maintenance»).

projets d'une seule unité administrative. La DSStB montre comment et dans quel but l'administration fédérale souhaite utiliser dorénavant la science des données, quels sont ses objectifs en la matière ainsi que les mesures nécessaires en vue d'une utilisation ciblée et coordonnée. Des mesures concrètes permettront également d'aborder les deux autres objectifs clés de l'IS 6 «Coordination et soutien DSCC et CNAI» et «Concept de formation en science des données».

La conception d'espaces de données⁶ tout comme le développement des données publiques ouvertes (Open Government Data - OGD)⁷ ne font pas partie de la stratégie DSStB. Certes, les applications de la science des données recourent aux données mises à disposition, par exemple dans des espaces de données ou via les OGD, mais elles ne déterminent aucunement leur conception. Dans la même optique, la DSStB ne couvre pas non plus la politique des données, qui définit le cadre organisationnel et technique de la gestion et de l'infrastructure des données, dans lesquelles évoluent les applications de la science des données.

La présente publication est structurée comme suit:

- Le chapitre 2 décrit la pertinence, l'objectif et l'approche de la DSStB. Ce faisant, il en explique et dépeint la logique.
- Le chapitre 3 définit le processus de la science des données en tant que discipline tout en la plaçant dans le contexte des politiques publiques.
- Le chapitre 4 présente la conception commune de la science des données au sein de l'administration fédérale. Cette conception commune, qui comprend une vision, des principes fondamentaux et une ambition, forme la base de la DSStB.
- Quant au chapitre 5, il décrit les lignes directrices et mesures permettant de mettre en œuvre la vision, les principes fondamentaux et l'ambition.
- L'interaction entre différents facteurs est essentielle pour veiller à ce que la science des données soit utilisée de manière coordonnée et ciblée. Le chapitre 6 présente donc l'objectif «écosystème transversal de la science des données».

⁶ Citons le communiqué du Conseil fédéral concernant le rapport [«Promouvoir des espaces de données fiables et l'autodétermination numérique»](#).

⁷ Site de l'OFS dévolu aux [OGD](#)

2 Pertinence, objectif et approche de la stratégie

L'illustration 1 fournit un aperçu de la DSStB en précisant sa pertinence, son objectif et son approche.

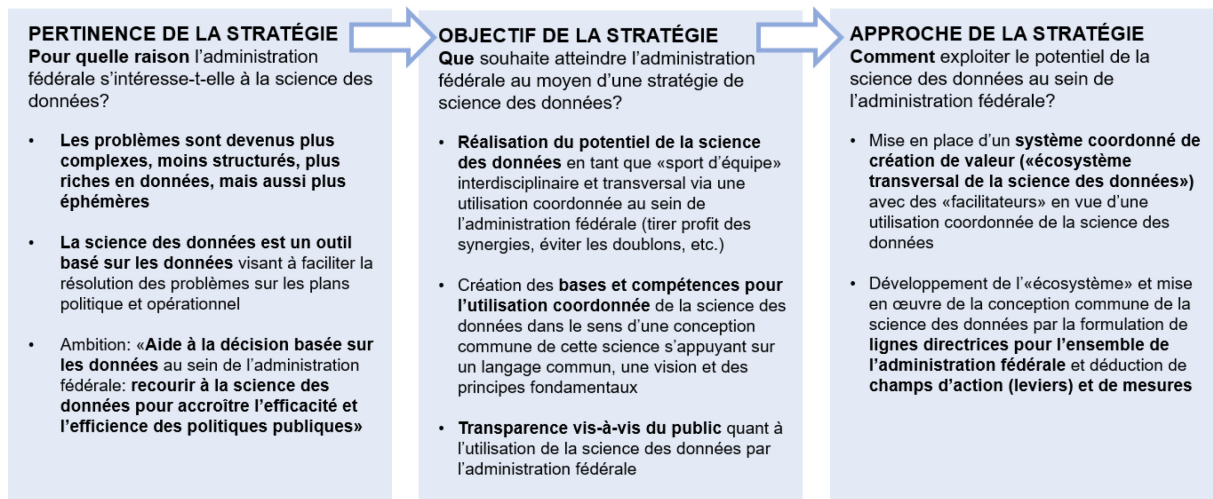


Illustration 1: Pertinence, objectif et approche de la DSStB (source: propre représentation).

2.1 Pertinence de la stratégie

Au fil de l'évolution technologique, divers problèmes sont devenus plus complexes, moins structurés et plus riches en données, mais aussi plus éphémères. La résolution de ces problèmes par les sphères politiques et administratives devient d'autant plus exigeante. La science des données s'avère être un outil approprié tant pour relever ces défis au niveau opérationnel que pour soutenir les décisions politiques. Elle aide à mieux comprendre les problèmes et les solutions possibles dans un tel environnement non structuré, riche en données et éphémère. Ce faisant, la science des données complète les instruments existants de recherche de solutions et d'aide à la décision utilisés par les sphères politiques et administratives pour que celles-ci soient en mesure de remplir leurs tâches légales de manière aussi efficace qu'efficiente au bénéfice de la communauté, au XXI^e siècle également. Dans ce contexte, l'ambition de la DSStB se formule comme suit:

«Aide à la décision basée sur les données au sein de l'administration fédérale: recourir à la science des données pour accroître l'efficacité et l'efficience des politiques publiques».

2.2 Objectif de la stratégie

La DSStB crée les bases permettant d'exploiter le potentiel de la science des données pour l'ensemble des unités administratives de la Confédération. Elle encourage tant le recours à la science des données que la mise en place des fondements et compétences organisationnels, éthiques, légaux et techniques indispensables au sein de l'administration fédérale. Il s'agit d'atteindre une utilisation coordonnée de la science des données dans le but d'exploiter les synergies, d'éviter les doublons et d'établir des principes communs. Avec la DSStB, l'administration fédérale indique en outre en

toute transparence de quelle manière elle utilise la science des données. Compte tenu des préoccupations de la population quant à l'utilisation de ses données par l'État, l'administration fédérale se doit d'être exemplaire dans l'application d'une science des données centrée sur l'être humain et digne de confiance (voir chapitre 3.3.1). Généraliser le recours à la science des données ne sera possible que si chacun au sein de l'administration fédérale considère son utilisation comme un «sport d'équipe» interdisciplinaire et transversal.

2.3 Approche de la stratégie

Deux aspects sont essentiels pour le développement et la mise en œuvre de la stratégie. Tout d'abord, une utilisation coordonnée et ciblée de la science des données exige l'interaction de différents facteurs. C'est pourquoi la DSStB vise la création et la maintenance d'un système coordonné de création de valeur. Cinq facilitateurs («rouages») forment l'«écosystème transversal de la science des données» (voir chapitre 6) que vient compléter l'ancrage stratégique de la science des données. Ces deux éléments forment conjointement les leviers centraux de la DSStB, sur la base desquels il est possible de prévoir des mesures concrètes. Mesures qui seront mises en œuvre sous la forme de mandats subséquents.

Ensuite, la DSStB constitue un cadre et définit des bases communes, sans pour autant définir des objectifs d'efficacité et de mise en œuvre pour les différentes unités administratives. Ce faisant, la stratégie tient compte des différents niveaux d'ambition et de maturité au sein de celles-ci. En effet, elle laisse à chaque unité de l'administration fédérale la liberté d'opter pour sa propre voie de développement en matière de science des données, en fonction de ses possibilités. Reste que la stratégie exige cependant que les unités administratives se penchent activement sur le sujet de la science des données et définissent leur propre voie de développement.

3 Fondement: langage commun et compréhension partagée

3.1 Importance du langage commun

La création d'un langage commun doit être vue comme un élément central de la DSStB. Une compréhension partagée des concepts, et le langage commun qui en résulte, permettent un échange actif d'expériences et de connaissances tout en simplifiant la communication en termes d'idées, de projets et de prestations. En effet, si une partie de l'administration fédérale dispose d'une certaine expérience en matière de science des données, l'application pratique de ce domaine est très hétérogène.

Dans la perspective de développer la science des données au sein de l'administration fédérale – un projet ambitieux –, il est essentiel de miser dès le début sur un langage commun et des notions clés clairement définies afin d'éviter tout malentendu. La signification pour l'administration fédérale de la science des données et de ses principales notions est présentée ci-après.

3.2 Définition de la science des données

La science des données (data science) est la science interdisciplinaire de l'apprentissage à partir des données. Elle vise à tirer des enseignements des données pour pouvoir prendre des décisions fondées. Elle se concentre sur le processus de résolution de problèmes et sur un processus d'amélioration continue dans le but de résoudre des problèmes complexes, impliquant de grandes quantités de données dans un environnement non structuré, grâce à des méthodes (tirées des statistiques avancées, de l'apprentissage automatique et du domaine de l'intelligence artificielle, p. ex.), des techniques et des pratiques. La science des données couvre l'entier du processus de la formulation du problème à la communication et à la mise à disposition des résultats, en passant par la saisie, la sélection, la préparation, l'analyse des données, et l'évaluation et l'interprétation des résultats. Comme le montre l'illustration 2, le processus inhérent à la science des données est un cycle d'amélioration continue. Si nécessaire, il est possible de relancer le cycle sur la base des connaissances nouvellement acquises. À quoi s'ajoute que la science des données doit être considérée comme un «sport d'équipe» interdisciplinaire et transversal.

La science des données et la statistique sont deux domaines étroitement liés. Outre la statistique, la science des données tire encore parti d'autres disciplines, notamment la gestion des données, l'ingénierie, l'apprentissage automatique, l'IA, les algorithmes, l'optimisation et la visualisation des données. Dans le but d'uniformiser la terminologie du domaine de l'intelligence artificielle pour toute l'administration fédérale, le CNAI donne la définition suivante du terme «science des données»: «on dit parfois que la science des données est l'enfant de la statistique et de l'informatique. En filant cette métaphore, on peut conclure que la science des données a hérité quelque chose (dans

l'idéal le meilleur) de ses deux parents, pour devenir une unité distincte. Son centre d'intérêt du moins la distingue de ses parents.⁸»

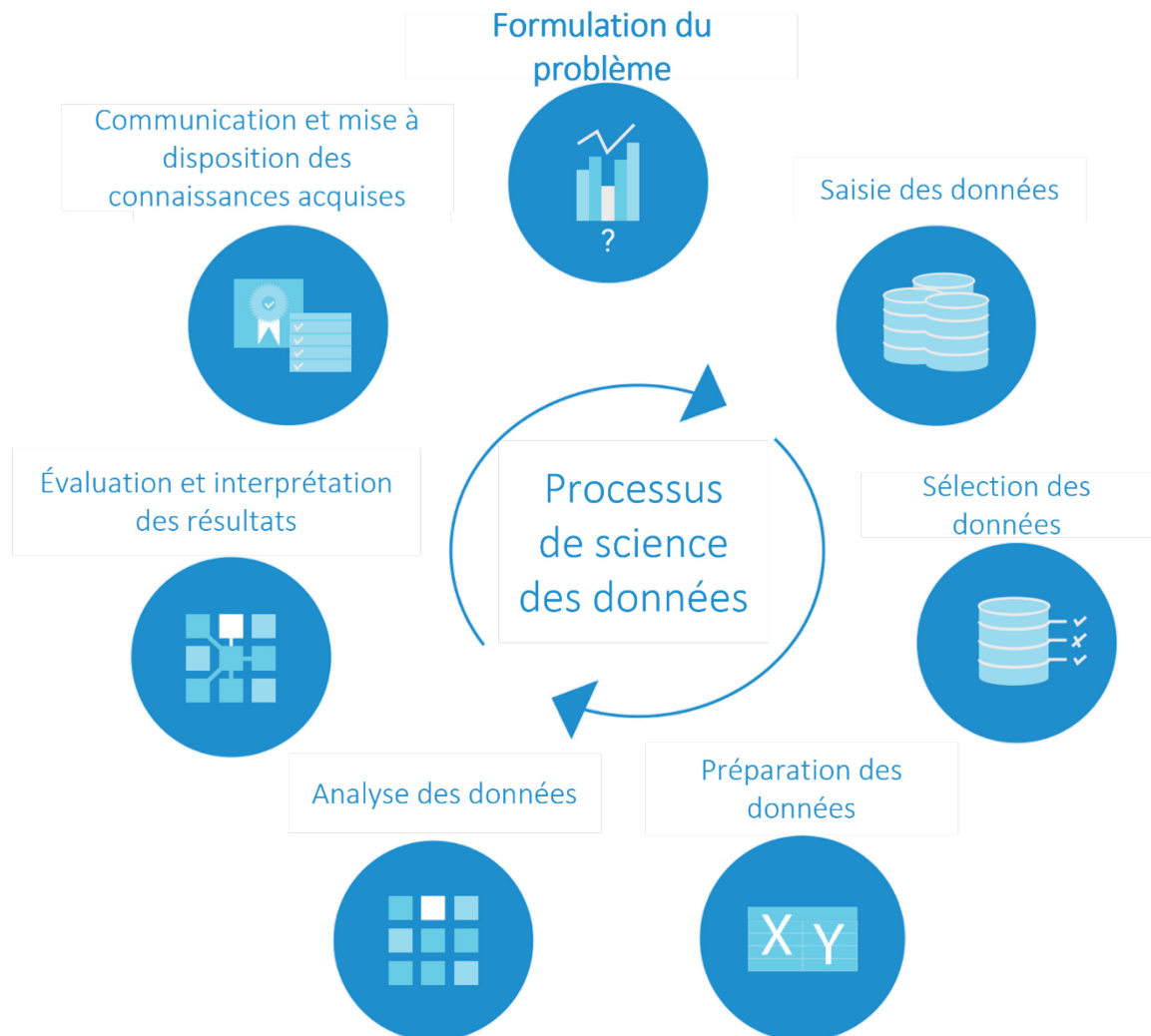


Illustration 2: Processus interdisciplinaire de résolution de problèmes de la science des données (source: propre représentation élaborée sur la base des données du [Centre de compétences en science des données \(DSCC\)](#)).

3.3 Importance de la science des données pour l'administration fédérale

3.3.1 Science des données centrée sur l'être humain («human-centric») et digne de confiance («trustworthy»)

Toute application de la science des données au sein de l'administration fédérale est centrée sur l'être humain. Le recours à la science des données doit tenir compte des besoins actuels des personnes et les aider à accomplir leurs tâches. Il va de soi que l'humain doit garder le contrôle en toute circonstance. Le but est donc de développer

⁸ [Terminologie du CNAI](#)

des applications en science des données qui respectent les valeurs éthiques, juridiques, démocratiques et culturelles de notre pays. Cette approche axée sur le facteur humain doit inspirer confiance dans l'application de méthodes, de techniques et de pratiques de science des données et dans les connaissances acquises de la sorte. Elle met aussi en valeur le potentiel élevé en matière de nouvelles applications et de modèles commerciaux novateurs.

Si l'administration fédérale a identifié les opportunités offertes par une prise de décision davantage basée sur les données, elle prend très au sérieux les défis et risques inhérents à cette méthode. Il s'agit en particulier de garantir la protection des droits fondamentaux lors de tout recours à une application de la science des données (notamment la non-discrimination et la protection des données). Il est par conséquent de la responsabilité de l'administration fédérale d'assurer la traçabilité et la transparence de toute application de la science des données. À défaut de quoi le respect des principes démocratiques et des exigences légales risque de ne pas pouvoir être garanti dans le cadre de la prise de décision basée sur les données. Une telle situation pourrait voir la confiance de la population, des entreprises et des institutions dans les politiques publiques diminuer.

Il incombe à l'État – en collaboration avec les domaines de la science et de la recherche ainsi qu'avec l'économie – de réglementer l'utilisation croissante de la science des données et de relever les défis en découlant. Ce faisant, il s'agit de prendre en compte de manière proactive les préoccupations de la population, de l'économie et de la société civile. En effet, on ne peut imposer la confiance, on ne peut que l'inspirer par les actes. Dans cette optique, l'administration fédérale doit définir des principes fondamentaux clairs, assurer leur respect mais aussi renforcer la capacité des citoyens de qualifier si les services de l'État appliquent la science des données d'une manière digne de confiance.

3.3.2 Science des données pour le bien commun («public good» et «public policy»)

Les politiques publiques servent le bien commun. Il en va de même du recours à la science des données. En se basant sur la définition de la «UK Statistics Authority»⁹, la science des données contribue au bien commun («public good»¹⁰) selon quatre aspects. Tout d'abord, elle fournit une base de connaissances qui soutient tant la prise de décision basée sur les données en politique que la fourniture de services publics. Ce point inclut d'autres décisions pour le bénéfice de l'économie, de la société ou de la qualité de vie de la population. Ensuite, la science des données permet de reproduire, de valider, de contester ou de vérifier les solutions existantes. Troisièmement, le

⁹ [Science des données pour le bien commun du point de vue de la «UK Statistics Authority»](#)

¹⁰ En économie (<https://www.investopedia.com/terms/p/public-good.asp>), on définit les biens publics («public good») comme étant des «biens et services mis à la disposition de l'ensemble des membres de la société. En règle générale, ces services sont gérés par un gouvernement et financés collectivement par le biais d'impôts. Quelques exemples de biens publics: la poursuite pénale, la défense nationale ou encore l'État de droit. Des biens fondamentaux, tels que l'accès à l'air pur et à l'eau potable, font également partie des biens publics».

recours à la science des données améliore la qualité, la couverture ainsi que la présentation des applications existantes. Pour terminer, l'amélioration des connaissances permet de mieux appréhender les tendances et événements sociaux ou économiques.

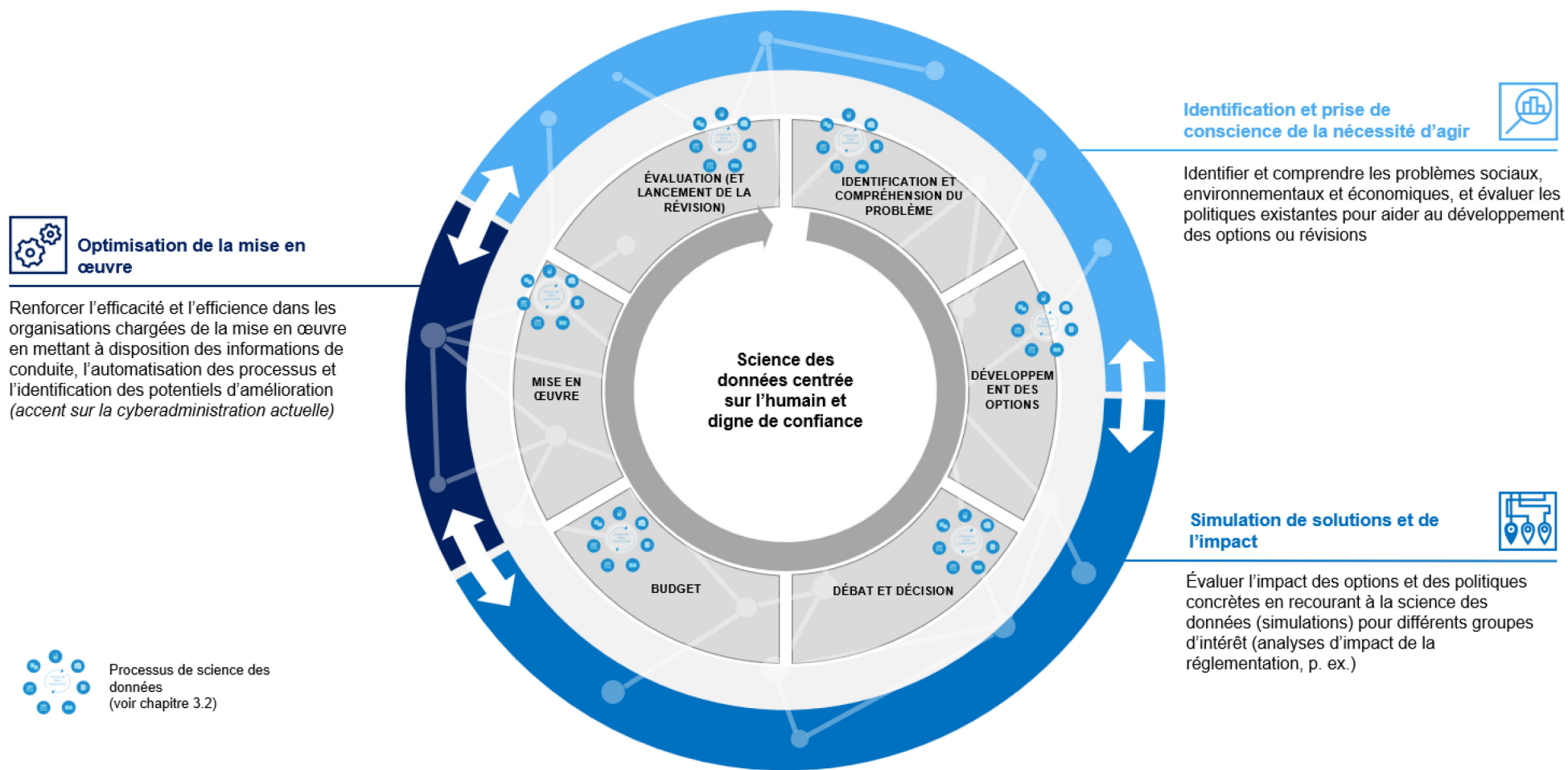


Illustration 3: La science des données tout au long du processus d'élaboration des politiques («policy making») (*source: propre représentation*).

Concrètement, le processus d'élaboration des politiques («policy making») illustre l'apport de la science des données au bien commun. Il décrit (de manière simplifiée) le cycle comprenant l'élaboration, la mise en œuvre et l'évaluation des politiques. La science des données peut soutenir de manière déterminante l'ensemble de ces activités en complétant les instruments existants d'élaboration et de mise en œuvre des politiques. L'illustration 3 montre comment les applications en science des données peuvent être intégrées au processus des politiques publiques («public policy»¹¹). Cette illustration montre clairement la pertinence de la science des données, pour l'ensemble des départements, mais aussi pour la Chancellerie fédérale (pour les demandes des citoyens, les questions parlementaires, l'identification de problèmes, l'évaluation des mesures politiques, p. ex.).

Pour prendre une décision politique et la mettre en œuvre, il faut commencer par identifier le problème et déterminer les mesures à prendre. Analyser et évaluer les données pertinentes permettra de mieux cerner la problématique ainsi que les relations et interdépendances opportunes. Lors de l'étape suivante, à savoir l'identification des solutions potentielles (modifier la législation, établir un concept de gestion du trafic, p. ex.), il est essentiel de mieux connaître le problème. C'est là que peut intervenir une simulation basée sur des données afin de restreindre le choix des options. Dans la suite du processus, les connaissances acquises soutiennent le débat politique tout en développant ou précisant les options possibles et permettent, au final, de prendre des décisions fondées. Une fois la décision prise, il est également possible de procéder à une simulation pour analyser l'affectation des ressources ainsi que l'impact possible. Pour terminer, lors de la mise en œuvre de la décision, le traitement scientifique de l'information ou l'automatisation des processus peuvent améliorer l'application. Dans le cadre de l'évaluation et d'éventuelles révisions, l'analyse des données collectées permet de tirer des conclusions sur le succès de la politique conduite afin, dans une optique d'amélioration continue, d'affiner la compréhension du problème. Pour résumer, la science des données renforce tant l'efficacité que l'efficience tout au long du processus d'élaboration des politiques, conformément à l'ambition de la DSStB.

¹¹ Les politiques publiques («public policy») désignent l'ensemble des décisions, objectifs et activités de toutes les personnes impliquées dans le processus politique. Il s'agit donc d'une discipline interdisciplinaire des sciences de l'État qui traite de questions politiques, économiques et sociétales.

4 Conception commune de la science des données au sein de l'administration fédérale

En se basant sur l'importance de la science des données (chapitre 3.3), il s'agit de définir une conception commune de la science des données qui sera déterminante pour l'ensemble des activités de l'administration fédérale en matière de science des données. Cette conception comporte une vision, des principes fondamentaux et une ambition.

4.1 Vision

Elle se définit comme suit:

«Une science des données centrée sur l'humain et digne de confiance qui soutient le bien commun ainsi que les politiques publiques»

(«human-centric and trustworthy data science for public good and public policy»).

La science des données applique des méthodes, des techniques et des pratiques appropriées afin d'améliorer la compréhension et de promouvoir l'aide à la décision basée sur les données. Ce faisant, on crée une valeur ajoutée pour le bien commun et les politiques publiques tout en mettant en œuvre le potentiel de la science des données pour une transformation numérique efficiente de l'administration fédérale. Cette dernière coordonne le recours à la science des données pour l'ensemble des unités administratives et avec les autres niveaux de l'État dans le but de tirer profit des synergies et de développer des compétences et des solutions communes. Dans le travail quotidien, les principes fondamentaux sont appliqués de manière conséquente.

4.2 Principes fondamentaux

Les principes fondamentaux sont essentiels pour une science des données centrée sur l'humain et digne de confiance. Ils servent de ligne directrice pour l'application de la science des données et la mise en œuvre de la vision et de l'ambition. Les principes fondamentaux de la DSSStB correspondent aux [«valeurs fondamentales du DSCC»](#) existantes pour la fourniture de prestations en matière de science des données dans le but de créer une valeur ajoutée durable. Ils reposent notamment sur les [«Fundamental Principles of Official Statistics»](#) de l'ONU et sur le [«Code de bonnes pratiques de la statistique européenne»](#) de l'UE, qui s'en inspire, également valable pour la Suisse¹². Si les principes fondamentaux sont coordonnés avec des principes existants et reconnus au niveau international, il s'agit néanmoins de clairement les définir pour l'administration fédérale tout en les étayant par un contenu correspondant¹³ (voir mesure A.1 au chapitre 5.2). Les principes fondamentaux:

- sécurité de l'information,
- protection des données et de l'information,

¹² Ces principes fondamentaux font partie intégrante de la [gestion de la qualité de l'OFS](#).

¹³ Inspiré p. ex. du [«Guide for Ethical Data Science»](#) de la «Royal Statistical Society» (RSS) et du «Institute and Faculty of Actuaries» (IFoA).

- sécurité des données,
- gouvernance des données,
- non-discrimination,
- explicabilité,
- traçabilité,
- transparence,
- reproductibilité,
- neutralité,
- objectivité,
- traitement éthique des données et des résultats.

Les résultats de chaque projet doivent par exemple être documentés et mis à disposition de manière transparente (pour autant que les dispositions sur la protection des données et les autres prescriptions légales le permettent)¹⁴. Chaque unité administrative est responsable de la pertinence et de la mise en œuvre des principes fondamentaux. La pondération de ces principes peut donc varier en fonction de l'application de la science des données. Sur la base d'un code de conduite (restant à élaborer, voir chapitre 5.2), les principes fondamentaux seront concrétisés en fonction des besoins de chaque unité administrative. En outre, la DSStB doit contribuer à assurer le respect des principes fondamentaux à plus long terme¹⁵.

4.3 Ambition

Déjà mentionnée plus haut, l'ambition met en œuvre la vision et concrétise la façon dont l'administration fédérale souhaite contribuer au bien commun par le recours à une science des données centrée sur l'être humain et digne de confiance. Lors de la prise de décision, l'objectif est de se fonder davantage sur les connaissances acquises via des processus en science des données. Une approche qui aidera les unités administratives à gérer des problèmes qui ont tendance à gagner en complexité. Bien sûr, au cœur de la résolution de problèmes au moyen de la science des données, c'est toujours l'être humain que l'on retrouve. L'ambition peut être définie comme suit:

«Aide à la décision basée sur les données au sein de l'administration fédérale: recourir à la science des données pour accroître l'efficacité et l'efficience des politiques publiques».

En mettant en évidence des lignes directrices, des leviers et des mesures, le chapitre suivant présente comment l'administration fédérale mettra en œuvre la conception commune qu'elle se fait de la science des données.

¹⁴ Dans l'optique de l'opérationnalisation d'applications et de projets en matière de science des données, on pourra p. ex. s'inspirer de «best practices» telles que les [«Reproducible Analytical Pipelines»](#) (RAP) ou du cadre [«Reproducible Machine Learning»](#) développé par «Statistics Canada».

¹⁵ Il sera nécessaire d'assurer la gouvernance de ces valeurs à plus long terme, p. ex. en créant un «label» du genre «Développé sur la base d'une science des données centrée sur l'être humain et digne de confiance!» qui mettra surtout l'accent sur l'utilisation éthique des données et des résultats («ethics») et sur la confiance du public («public trust»). Cela sera harmonisé et coordonné avec les prestations de base en science des données du DSCC (à savoir le «Développement de normes de qualité, de directives pour le respect de la protection des données et d'infrastructures de base (sandboxes) pour les applications utilisées en science des données») et le pilotage externe du DSCC au sein de l'administration fédérale.

5 Lignes directrices, leviers et mesures

Des lignes directrices, des leviers et des mesures concrètes ont été formulés pour mettre en œuvre la vision, les principes fondamentaux et l'ambition. L'illustration 4 en fournit une vue d'ensemble.

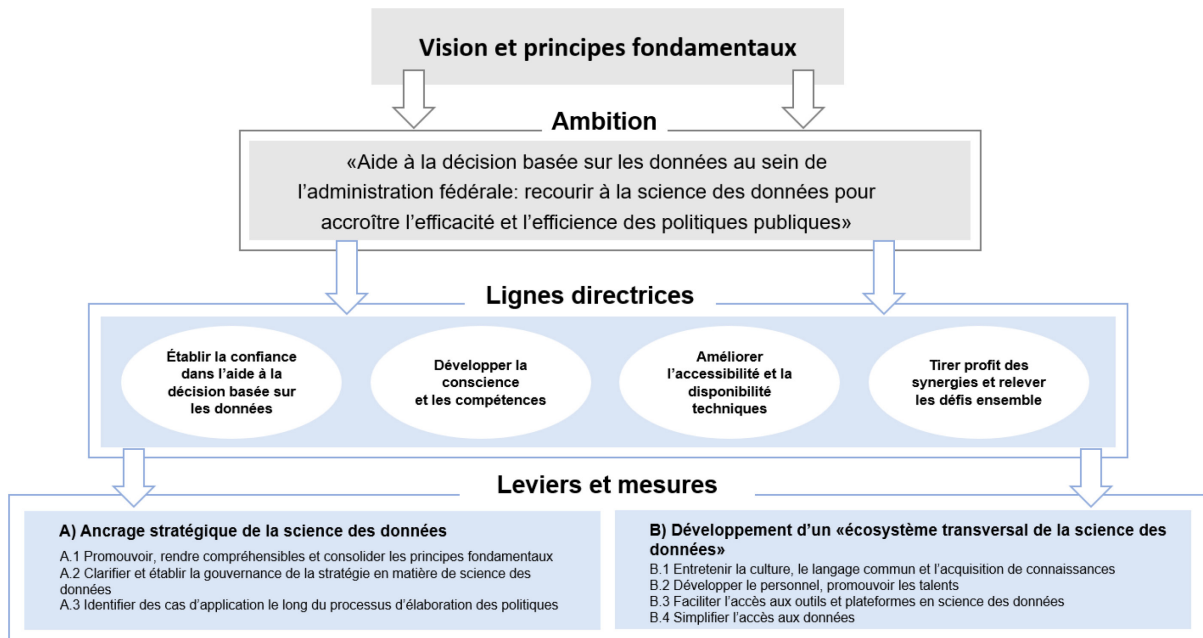


Illustration 4: Structure et hiérarchie de l'ambition, des lignes directrices, des leviers et des mesures (source: propre représentation).

5.1 Lignes directrices

Équivalentes et interdépendantes, les lignes directrices présentent ce que la DSStB doit permettre d'atteindre.

5.1.1 Établir la confiance dans l'aide à la décision basée sur les données

La confiance des acteurs, qu'ils soient impliqués directement ou indirectement, est un facteur central pour façonner la manière dont l'administration fédérale conçoit l'utilisation à des fins de politiques publiques d'une science des données centrée sur l'être humain, digne de confiance et orientée vers le bien commun. C'est précisément la DSStB qui permettra de créer la confiance dans l'application de méthodes, de techniques et de pratiques en science des données, parmi la population comme au sein de l'administration fédérale.

5.1.2 Développer la conscience et les compétences

Pour exploiter le potentiel de la science des données, il faut commencer par l'identifier. La DSStB vise à faire prendre conscience aux acteurs de l'administration fédérale de l'intérêt et de la nécessité d'approfondir leurs connaissances de la science des données. Cela implique de plus amples échanges tant sur les méthodes, pratiques et tech-

niques en science des données que sur les résultats obtenus. Il s'agit donc de promouvoir davantage encore les connaissances et compétences au sein de l'administration fédérale et de les consolider de manière transversale.

5.1.3 Améliorer l'accessibilité et la disponibilité techniques

L'accès simple, rapide et sécurisé aux outils et plateformes de science des données doit permettre aux acteurs d'aisément faire leurs propres expériences, d'explorer des solutions potentielles à leurs problèmes en science des données et de se servir d'applications issues de la science des données. De plus, la gestion et l'infrastructure des données constituent un élément fondamental pour la mise en œuvre durable du processus de résolution de problèmes par la science des données. Mettre en œuvre cette science avec une petite quantité de données est possible, à condition qu'elles soient de qualité. Reste que l'accès aux données doit être facilité conformément aux dispositions légales, tant sur le plan technique que sur le plan organisationnel.

5.1.4 Tirer profit des synergies et relever les défis ensemble

La DSStB contribue à tirer profit des synergies au sein de l'administration fédérale, entre les différents niveaux de l'État, avec les domaines de la science et de la recherche, avec l'économie et même avec d'autres États. Pour le succès de la DSStB, cette collaboration transversale dans le domaine de la science des données est essentielle: en effet, c'est elle qui permet de gérer conjointement la complexité des problèmes actuels. Les bases de connaissances et les compétences en science des données des différents acteurs permettront d'élaborer des solutions efficaces.

5.2 Leviers et mesures

Afin de mettre en œuvre les lignes directrices susmentionnées, la DSStB prévoit d'actionner les deux leviers: «A) l'ancrage stratégique de la science des données» et «B) le développement d'un «écosystème transversal de la science des données»», mesures concrètes comprises.

A) Ancrage stratégique de la science des données

L'ancrage stratégique du développement et de la planification de projets en science des données au niveau de direction de chaque unité administrative est primordial pour mener à terme la mise en œuvre de la DSStB. Cet ancrage constitue donc le premier levier et se traduit par trois mesures concrètes principales.

A.1 Promouvoir, rendre compréhensibles et consolider les principes fondamentaux

Édicter un code de bonnes pratiques pour une science des données centrée sur l'être humain et digne de confiance, valable pour l'ensemble des équipes de science des données et des décideurs de la Confédération, permettra de promouvoir, de rendre compréhensibles et de consolider les principes fondamentaux. Ce code pourra être compris comme une aide à l'orientation pour respecter les principes fondamentaux lors de tout projet en science des données. C'est dans cette optique que s'inscrivent la création et la diffusion d'un «serment en matière de science des données (data

science oath)»¹⁶ ou d'un «code de conduite (data science code of conduct)»¹⁷ pour tout membre d'une équipe de science des données. Le document «*Intelligence artificielle*» – *lignes directrices pour la Confédération*» du Conseil fédéral¹⁸ et le «*Code de conduite pour la création et l'exploitation d'espaces de données fiables*», élaboré par le DFAE et le DETEC,¹⁹ peuvent également servir d'inspiration.

A.2 Clarifier et établir la gouvernance de la stratégie en matière de science des données

Il importe de définir la gouvernance appropriée durant toutes les étapes de la DSStB: élaboration, mise en œuvre, suivi et développement. Ce faisant, il s'agira de déterminer qui assume quelles responsabilités, de régler le financement, etc. La définition des rôles comprend aussi la coordination du paysage des organes dans le but de promouvoir la collaboration transversale. La gouvernance de la DSStB inclut les domaines de la science et de la recherche, l'économie et la coopération internationale, tout en tenant compte des organes existants (p. ex. le «groupe d'experts DSCC») afin de tirer profit des synergies. Il n'est pas prévu de créer d'autre paysage des organes. La gouvernance met en outre en place la coordination et l'organisation nécessaires ainsi que les processus pertinents pour faire entendre les exigences du domaine de la science des données dans les organes de la politique en matière de données de la Confédération. Et ce en tenant compte des délimitations claires entre les différentes initiatives stratégiques de l'administration fédérale.

A.3 Identifier des cas d'application le long du processus d'élaboration des politiques

La science des données peut être appliquée à chaque étape du processus d'élaboration des politiques («policy making»), dans le but de soutenir et de promouvoir davantage encore les décisions basées sur les données. Aujourd'hui toutefois, la science des données est principalement utilisée dans le cadre de la mise en œuvre. Afin d'établir la collaboration transversale et de soutenir l'ensemble du processus d'élaboration des politiques, il faut identifier des cas d'application transversaux au sein de l'administration fédérale et les promouvoir activement. Ce faisant, il s'agit d'évaluer les bases légales de cas en cas et, pour autant que nécessaire, de formuler des recommandations.

B) Développement d'un «écosystème transversal de la science des données»

L'application efficace et efficiente de la science des données, telle que la conçoit l'administration fédérale, exige un cadre approprié. Il faut en effet réunir les conditions nécessaires – juridiques, éthiques, techniques et organisationnelles – mais aussi adapter certains aspects de la culture organisationnelle. Dans leurs interactions, ces facteurs forment un «écosystème transversal de la science des données». Cet écosystème constitue la base permettant de consolider les innovations, les améliorations et les

¹⁶ À l'image d'un «[serment d'Hippocrate](#)» pour la science des données.

¹⁷ Comme par exemple le «[Guide for Ethical Data Science](#)» de la «Royal Statistical Society» (RSS) et de l'«Institute and Faculty of Actuaries» (IFoA).

¹⁸ <https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-81319.html>

¹⁹ <https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-87780.html>

champs d'application en lien avec la science des données et d'en tirer le meilleur parti. Un tel écosystème est en outre indispensable pour mettre en œuvre la vision, les principes fondamentaux et l'ambition de la DSStB. L'illustration 5 montre l'écosystème avec ses cinq facilitateurs («rouages»). Le chapitre 6 décrit en détail l'objectif de l'«écosystème transversal de la science des données».

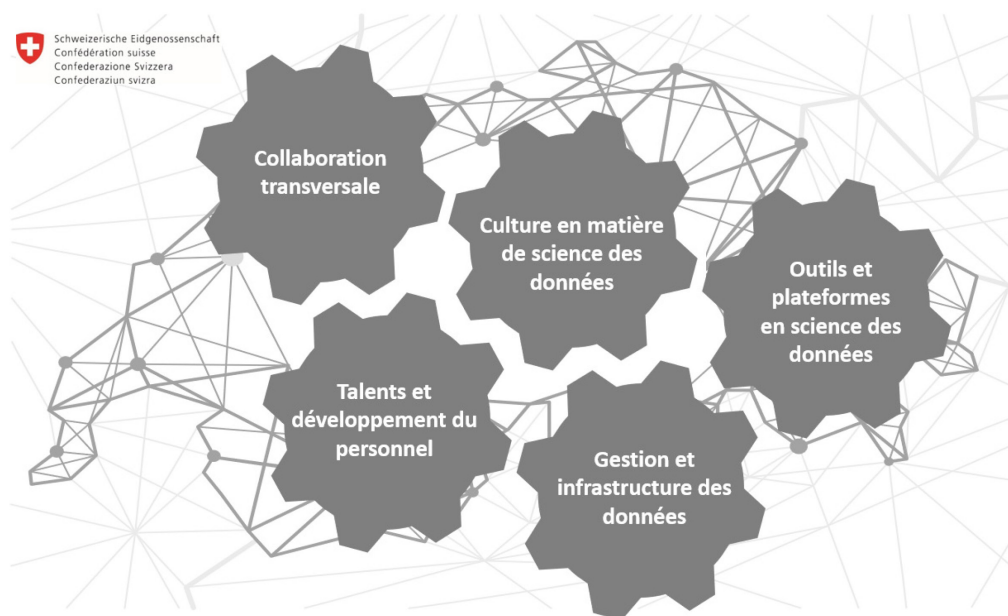


Illustration 5: Objectif de l'«écosystème transversal de la science des données» avec ses facilitateurs («rouages») (source: propre représentation).

Le développement de cet écosystème constitue le deuxième levier de la DSStB. Les mesures concrètes correspondantes sont décrites ci-dessous:

B.1 Entretenir la culture, le langage commun et l'acquisition de connaissances

Entretenir une culture en matière de science des données, un langage commun et l'acquisition collaborative des connaissances constituent des éléments centraux de la DSStB. Raison pour laquelle il ne sera créé qu'une seule communauté de pratique en science des données («Community of Practice», CoP). La CoP est un groupe de personnes qui doivent relever des défis similaires, prêtes à apprendre les unes des autres. La CoP doit favoriser les partages d'expériences entre les unités administratives de la Confédération tout en fournissant une vue d'ensemble des projets en science des données de l'administration fédérale. Elle permet à toutes les parties prenantes d'échanger régulièrement sur leurs sujets, même au plan international. La CoP est ouverte à toute personne montrant de l'intérêt pour la science des données et prête à faire profiter les autres de son savoir. La communauté se caractérise par la passion, l'engagement et les connaissances de ses membres. La CoP se réunit régulièrement pour échanger des informations, développer le savoir-faire et contribuer activement à élargir les connaissances générales dans ce domaine. Afin d'éviter les doublons, la création de la CoP se fait en collaboration et en coordination avec la communauté externe à l'OFS

«Data Science for Public Good» du DSCC²⁰ et avec la «CoP IA» intégrée du CNAI²¹. La CoP se concentre sur le développement des prestations de base du DSCC en matière de directives pour l'utilisation d'outils et de plateformes en science des données. Elle vise en outre l'établissement d'un «catalogue» de ces outils et plateformes disponibles au sein de l'administration fédérale.

B.2 Développer le personnel, promouvoir les talents

Promouvoir les talents et développer le personnel est essentiel pour renforcer les capacités en matière de science des données au sein de l'administration fédérale. À cet effet, on commencera par mettre en place un programme coordonné de formation et de perfectionnement («upskilling / reskilling») pour l'application de la science des données au sein de l'administration fédérale, sur la base des expériences acquises²². Le programme doit renforcer les connaissances des cadres en matière d'application de la science des données. En combinaison avec le programme de formation et de perfectionnement, il s'agit de recruter des talents externes selon des profils clairement définis. Pour ce faire, l'administration fédérale doit se positionner en tant qu'employeur attractif dans le domaine de la science des données. Afin de promouvoir un développement durable du personnel, il convient de développer des profils professionnels clairement définis dans le domaine de la science des données.

B.3 Faciliter l'accès aux outils et plateformes en science des données

Pour appliquer les méthodes, pratiques et techniques en science des données dans un proche avenir, l'accès aux outils et plateformes correspondants doit être facilité autant que possible (dans le respect de la protection des données). À quoi s'ajoute qu'il faudra coordonner l'utilisation commune de ces outils et plateformes. Pour ce faire, il convient d'établir un produit standard (qui comprend l'analyse des besoins, l'architecture, les rôles d'exploitation, les exigences en matière de sécurité, l'archivage, la performance et la disponibilité) pour une «plateforme collaborative en science des données», compte tenu des systèmes techniques existants. Sachant que les outils et plateformes évoluent très vite dans ce domaine, il s'agit d'assurer une acquisition du produit standard aussi dynamique que flexible.

De même, il s'agit de sensibiliser les unités administratives ainsi que les collaborateurs à l'utilisation de ces outils et plateformes en publiant des directives (principes de base et dispositions juridiques). Des directives qu'il s'agira de coordonner avec les prestations de base correspondantes du DSCC, puis de développer. Un «catalogue» des outils et plateformes en science des données disponibles au sein de l'administration fédérale devra en outre être établi. Ce catalogue évaluera l'applicabilité de ces solutions tout en énumérant les défis liés à leur utilisation, compte tenu des expériences

²⁰ [Communauté du DSCC «Science des données pour le bien commun»](#)

²¹ [«CoP intelligence artificielle» du CNAI](#)

²² Depuis début 2022, le «Concept de formation en science des données», élaboré et validé par l'OFS, permet aux collaborateurs de cet office d'acquérir et de développer des compétences et aptitudes en science des données. Ce plan de formation définit différents groupes cibles, étapes de formation (sensibilisation, introduction, approfondissement, spécialisation) et parcours d'apprentissage. Reposant sur une collaboration avec l'«EPFL Extension School», le cursus est composé de différents modules et modalités de formation. En 2021, l'OFS et l'OPPER ont conduit de premiers entretiens en vue d'étendre cette offre de formation à l'ensemble de l'administration fédérale. D'autres entretiens suivront.

faites par l'administration fédérale. Il fournit enfin une vue d'ensemble des outils et plateformes, permet de s'y référer rapidement, garantit que ces derniers répondent à des normes de qualité définies tout en précisant dans quelle mesure ils favorisent durablement les principes fondamentaux. En fonction des besoins et exigences découlant des problématiques en matière de science des données, le catalogue couvrira autant les outils «stand-alone», les plates-formes que les solutions «hyperscale».

B.4 Simplifier l'accès aux données

La disponibilité de données pertinentes renforce considérablement l'utilité de la science des données. Au vu des quantités de données collectées par les différentes unités administratives de l'administration fédérale, il existe un grand potentiel à cet égard au sein de cette dernière²³. Pour tirer profit de ce potentiel, il convient de développer des méthodes, pratiques et techniques en science des données qui permettent un partage des données efficace, dans le respect de la protection des données et en application du principe de minimisation des données. Cette approche permet de générer des informations à partir de données relatives à des individus, sans qu'il ne soit nécessaire de transmettre les données des personnes en question. Citons pour exemples la création et l'utilisation de données synthétiques ainsi que l'application de la science des données pour préserver la sphère privée²⁴. Ce faisant, il s'agit d'évaluer les bases légales de cas en cas et, pour autant que nécessaire, de formuler des recommandations.

Les deux leviers «A) ancrage stratégique de la science des données» et «B) développement d'un «écosystème transversal de la science des données»» permettent de mettre en œuvre la conception que se fait l'administration fédérale de la science des données en son sein. Pour davantage de clarté, le chapitre suivant décrit en détail l'objectif de l'écosystème qui constitue la pierre angulaire de la mise en œuvre de cette conception commune de la science des données.

²³ Voir p. ex. la mesure 4 du rapport du Conseil fédéral concernant l'amélioration de la gestion des données dans le domaine de la santé: <https://www.admin.ch/gov/fr/accueil/documentation/communiques.msg-id-86762.html>

²⁴ [Privacy Preserving Techniques](#)

6 Objectif «écosystème transversal de la science des données»

6.1 Vue d'ensemble

Comme expliqué au chapitre 5.2, cinq facilitateurs («rouages») centraux entraînent l'«écosystème transversal de la science des données». Comme ils interagissent en permanence, ces facilitateurs ne peuvent fonctionner et se développer de manière optimale que s'ils coopèrent de façon harmonisée. Sans compter que cette interaction harmonieuse favorise la création et la maintenance d'un système coordonné de création de valeur.

Les différents facilitateurs («rouages») de l'«écosystème transversal de la science des données» sont détaillés ci-dessous.

6.2 Description des facilitateurs de l'écosystème

6.2.1 Culture en matière de science des données

Pour un «écosystème transversal de la science des données» fonctionnel, il faut tout d'abord établir une culture en science des données qui favorise une utilisation centrée sur l'être humain et digne de confiance. Les principes fondamentaux définis (au chapitre 4.2) constituent les piliers de cette culture et doivent être mis en œuvre et garantis par cette dernière. À quoi s'ajoute que le respect des principes fondamentaux favorise la confiance de la population dans les projets en science des données de l'administration fédérale.

La culture recherchée compte deux niveaux. Premièrement, la culture de conduite détermine la manière d'aborder les projets de science des données. Ce faisant, la direction doit créer l'espace nécessaire permettant d'encourager l'organisation, dans la mesure du possible, à expérimenter et à utiliser la science des données de sorte à soutenir l'apprentissage via des projets concrets. Il faudra en outre que chacun prenne conscience de la valeur ajoutée de l'apprentissage à partir des données, mais aussi de l'importance de la qualité des données (autrement dit, savoir évaluer ces données en fonction de leur aptitude à répondre à un but précis). Par conséquent, sera tout aussi importante la volonté d'investir dans une collecte de données correcte et aussi complète que possible, conformément aux principes fondamentaux définis. Le second niveau est la culture de travail, essentiellement marquée par les collaborateurs qui mettent en œuvre des applications en science des données. Lors de la mise en œuvre de la culture de travail, il faut veiller à ce que la valeur ajoutée se rapporte à un État centré sur l'être humain et digne de confiance, qu'il règne un climat d'entreprise renforçant l'autonomie des collaborateurs et favorisant l'innovation et l'amélioration continue. Pour que l'écosystème fonctionne avec succès, la culture doit avant tout permettre de mieux prendre conscience de l'utilité de la science des données.

6.2.2 Talents et développement du personnel

L'administration fédérale, qui a besoin de ses collaborateurs pour mener à bien ce projet, doit faire face à un défi de taille: s'entourer de collaborateurs qui disposent des

connaissances nécessaires pour mettre en œuvre des projets en science des données. Elle doit veiller à disposer des bons spécialistes en proposant une formation et un perfectionnement internes ou externes («upskilling / reskilling») ou en les recrutant au sein de l'administration fédérale. Parallèlement, elle doit renforcer les connaissances en matière de leadership en vue de l'application de la science des données.

Il y a une forte demande pour les talents en science des données sur le marché du travail. Le recrutement de ces talents sera d'autant plus facile si l'administration est perçue comme un employeur attrayant dans le domaine de la science des données. Pour favoriser cette attractivité, il faut définir des profils professionnels présentant des parcours de formation et de carrière clairs et adopter une communication cohérente sur les multiples champs d'activité du domaine de la science des données au sein des unités administratives. Des profils professionnels harmonisés en science des données facilitent la collaboration et apportent davantage de clarté.

6.2.3 Collaboration transversale

La science des données est un «sport d'équipe» interdisciplinaire et transversal. Son succès dépend dans une large mesure de la collaboration entre spécialistes de différentes disciplines, organisations, secteurs et pays, apportant chacun des bases de connaissances différentes qui se complètent en termes de compétences en science des données. Le partage continu du savoir-faire et des projets communs permet d'acquérir de précieuses connaissances et de recourir aux applications existantes pour les utiliser dans d'autres domaines. La collaboration se déroule sur trois niveaux.

Parmi les administrations suisses: les projets interdépartementaux ou impliquant plusieurs niveaux de l'État permettent de tirer profit des synergies au sein de l'administration fédérale et entre les niveaux fédéraux. Dans l'idéal, les facilitateurs («rouages») de l'écosystème forment un cadre qui peut également être utilisé par les cantons et les communes, de préférence en intégrant les plateformes existantes, p. ex. l'Administration numérique suisse (ADS). L'objectif est ici de permettre à toute unité administrative de résoudre elle-même ses problèmes en recourant à des applications en science des données. Il ne s'agit en aucun cas de centraliser les compétences, mais bien d'encourager chaque unité administrative à utiliser la science des données dans le cadre de ses propres ambitions et possibilités. Afin de promouvoir la coopération dans le domaine de la science des données, il convient de plus d'élaborer des dispositions légales et d'évaluer les besoins d'adaptation.

Avec la science, la recherche et l'économie: une collaboration étroite entre l'administration, la science, la recherche et l'économie permet d'ancrer à brève échéance la recherche fondamentale en science des données et de partager entre les acteurs les solutions novatrices issues de la science des données. Un excellent exemple en est la collaboration entre le «Swiss Data Science Center» (SDSC) exploité conjointement par les deux écoles polytechniques fédérales et le DSCC, dans le but d'utiliser la science des données et l'IA pour le bien commun de la Suisse²⁵. De même, le programme de

²⁵ [Communiqué de presse de l'OFS: «Science des données et intelligence artificielle pour le bien commun»](#)

recherche «Data Science» d'armasuisse Science et technologies S+T recourt à un réseau d'experts pour assurer la disponibilité des compétences technologiques nécessaires dans les unités opérationnelles de l'armée, du Service de renseignement de la Confédération, de l'Office fédéral du sport et de l'Office fédéral de la police²⁶. Sur la base de telles collaborations, l'association des capacités méthodologiques des chercheurs et des connaissances spécifiques de l'administration fédérale permet d'obtenir de précieuses connaissances à partir des données disponibles. Une collaboration encore plus étroite avec des unités administratives décentralisées et des entreprises privées pourrait créer un espace pour apprendre les uns des autres, s'inspirer réciproquement et exploiter ensemble de nouveaux cas d'application.

Dans les échanges internationaux: la collaboration internationale est elle aussi essentielle pour le développement durable d'un écosystème en science des données pour l'administration fédérale. Acteur ouvert et neutre, la Suisse peut offrir un lieu où tant les chercheurs que les entreprises collaboreront au niveau international²⁷. Dans le domaine de la science des données, l'administration fédérale participe activement au développement de normes et d'initiatives internationales et européennes sur les plans technique et réglementaire. C'est dans cette optique qu'il faut par exemple considérer la prise de connaissance du rapport «IA et cadre réglementaire international»²⁸ du Conseil fédéral. Les mesures qu'il contient invitent la Suisse en tant qu'État à participer au niveau international à la définition du cadre réglementaire international en matière d'intelligence artificielle, et à contribuer ainsi à une réglementation modérée. En parallèle, certaines unités administratives doivent aussi jouer ce rôle de leader lors de projets internationaux. La coordination au sein de l'administration fédérale est essentielle pour tirer profit du potentiel de la collaboration internationale.

6.2.4 Outils et plateformes en science des données

La mise en œuvre des méthodes, pratiques et techniques en science des données nécessite des outils et des plateformes appropriés. Il s'agit notamment d'outils et de plateformes qui peuvent p. ex. être utilisés pour la visualisation de données, le contrôle de plausibilité de données, le développement collaboratif et l'utilisation d'algorithmes en science des données, la reconnaissance automatisée de textes, de langues et d'images. Le choix de tels outils et plateformes est gigantesque et ne cesse de croître²⁹, à tel point qu'il est difficile de garder une vue d'ensemble et de distinguer quelle application est la plus adaptée pour la résolution d'un problème scientifique particulier. Chaque unité administrative doit être en mesure d'identifier et d'utiliser l'outil «idéal» pour la résolution de sa problématique en matière de science des données. Outre la disponibilité des outils et des plateformes, il convient de garantir leur utilisation en conformité avec les principes fondamentaux et les dispositions légales (en particulier en matière de protection des données et de l'information).

²⁶ [«Programme de recherche 3b: Data Science» d'armasuisse Science et technologies S+T](#)

²⁷ Voir p. ex. la [«Stratégie de politique extérieure numérique 2021-2024»](#).

²⁸ [Rapport «Intelligence artificielle et réglementation internationale»](#)

²⁹ Voir p. ex. [«Machine Learning, Artificial Intelligence and Data \(MAD\) Landscape»](#).

6.2.5 Gestion et infrastructure des données

La gestion et l'infrastructure des données constituent le fondement d'une mise en œuvre optimale du processus de résolution de problèmes par la science des données. Recourir à cette science est possible même si les données sont peu nombreuses, à condition qu'elles soient de qualité. Cela ne nécessite pas forcément une grande infrastructure de données. Toutefois, pour exploiter de manière efficace, efficiente et durable tout le potentiel de la science des données et pour faire évoluer les applications, il est nécessaire de disposer d'une infrastructure et d'une architecture des données permettant une coordination commune de l'acquisition, de la sécurisation, de la rectification, de l'agrégation, de la mise à disposition et de l'utilisation des données. Tout cela, bien entendu, en tenant compte des exigences réglementaires et légales en matière de protection et de sécurité des données et de l'information. Une telle infrastructure de données comprend des éléments juridiques, techniques et organisationnels qui permettent l'élaboration, le traitement et la transmission de données tout au long de leur cycle de vie.

L'initiative stratégique clé IS 3 de la stratégie de numérisation de la Confédération poursuit l'objectif de «*mettre en place des architectures de prestations pour les services numériques en réseau, en application du principe once-only*»³⁰. Ce facilitateur («rouage») de l'écosystème de la science des données est donc couvert par cette initiative stratégique. Reste qu'il faudra améliorer la coordination et l'harmonisation.

³⁰ Voir les domaines d'action correspondants de l'IS 3 dans la [«Stratégie numérique de la Confédération 2020-2023»](#).