



# Terminologie

## Réseau de compétences en intelligence artificielle (CNAI)

Auteur	Secrétariat CNAI
Version	2.0
Date	21.12.2023

## TABLE DES MATIÈRES

<b>1. OBJECTIF DU PRÉSENT DOCUMENT ET DOMAINE D'APPLICATION.....</b>	<b>3</b>
<b>2. TERMINOLOGIE RELATIVE AUX (SOURCES) DE DONNÉES .....</b>	<b>4</b>
<b>3. TERMINOLOGIE PROPRE À L'«APPRENTISSAGE À PARTIR DE (SOURCES) DE DONNÉES».....</b>	<b>6</b>
<b>4. TERMINOLOGIE ACCESSOIRE.....</b>	<b>11</b>
<b>5. PRINCIPES FONDAMENTAUX POUR UNE SCIENCE DES DONNÉES (ET UNE IA) CENTRÉE SUR L'ÊTRE HUMAIN ET DIGNE DE CONFIANCE.....</b>	<b>12</b>



# 1. Objectif du présent document et domaine d'application

La création d'un réseau de compétences en intelligence artificielle (*Competence Network for Artificial Intelligence, CNAI*) vise à promouvoir durablement l'utilisation de l'intelligence artificielle et d'autres nouvelles technologies au sein de l'administration fédérale et au-delà.

L'un des objectifs du CNAI consiste à faire adopter une terminologie uniformisée afin d'instaurer une compréhension commune, au niveau de l'administration fédérale, des notions ayant trait à l'intelligence artificielle et aux nouvelles technologies.

Parler une même langue ne peut en effet que faciliter l'échange actif d'expériences et de savoir-faire au sein du CNAI. De plus, il va sans dire que cela simplifiera la communication, tant pour parler d'idées nouvelles que de projets et de services dans le domaine de l'intelligence artificielle.

La gestion de la terminologie qui incombe au secrétariat du CNAI représente dès lors l'une des pierres angulaires du bon fonctionnement du CNAI en général et de son secrétariat en particulier.

Cette gestion terminologique se fonde sur des documents disponibles à l'OFS et auprès de son centre de compétences en science des données (*Data Science Competence Center, DSCC*) et sur les documents élaborés suite à la décision du Conseil fédéral de mettre sur pied un réseau de compétences en intelligence artificielle en le dotant d'un secrétariat.

Une fois adoptées, les définitions des termes et des expressions retenues ne devraient être complétées que si ces ajouts s'imposent pour distinguer les technologies définies de technologies nouvelles. Le développement du glossaire sera assuré de manière ciblée en collaboration avec les nœuds du réseau et devrait illustrer le domaine dans lequel s'inscrit le CNAI.

Le présent document est la propriété du secrétariat CNAI.

Ce document fera l'objet de mises à jour périodiques et c'est sa version la plus récente (les références étant le numéro de version et la date) qui est valable.

## Convention typographique

Aussi longtemps qu'un terme n'a pas été explicitement défini, il est écrit entre guillemets. Exemple: «science des données».



## 2. Terminologie relative aux (sources) de données

On distingue en général **trois types de sources de données**:

- les sources de données internes à l'administration («données statistiques»<sup>1</sup>, données administratives, données officielles, géodonnées de la Confédération, *Open Government Data* ou «données publiques ouvertes» [OGD]<sup>2</sup>, p. ex.);
- les sources de données externes à l'administration (données ouvertes, p. ex.);
- les systèmes cyberphysiques (telles les données provenant de capteurs mis en réseau et appartenant à l'«internet des objets» [IdO]), qui peuvent constituer une source de données aussi bien interne qu'externe à l'administration.

L'utilisation intégrée de données provenant des sources de données internes ou externes à l'administration ainsi que des systèmes cyberphysiques engendre une (r)évolution des données qui offre d'énormes possibilités nouvelles pour les responsables politiques. C'est précisément pour exploiter ces possibilités à l'aide de méthodes scientifiques spécifiques que la «science des données» recèle un potentiel énorme. Elle permet par exemple de combiner des sources de données afin de mieux satisfaire les utilisateurs de données: répondre à des besoins non couverts jusque-là ou répondre mieux à des besoins couverts jusqu'alors de manière basique seulement. La science des données offre ainsi de nouveaux moyens de générer aussi rapidement que possible des informations pertinentes concernant différentes problématiques pour les mettre à la disposition des responsables politiques.

On peut par ailleurs subdiviser les **types de données** en **trois catégories**:

- les données structurées: données formatées et transformées selon un modèle clairement défini. Les données brutes sont reprises dans des champs prédéfinis, à partir desquels il est ensuite possible de les extraire aisément

---

<sup>1</sup> Les données statistiques sont toutes les données qui sont recueillies ou communiquées à des fins statistiques, en particulier conformément à la loi sur la statistique fédérale, et qui sont dès lors soumises au secret statistique.

<sup>2</sup> Les données en libre accès (données ouvertes) sont des données qui peuvent être utilisées, traitées, analysées et transmises librement, sans restrictions légales, financières ou techniques particulières. Du point de vue légal, ces données doivent pouvoir être utilisées et traitées gratuitement. Du point de vue technique, les données ouvertes doivent pouvoir être traitées par ordinateur. La publication de données à titre de données ouvertes doit garantir le respect des dispositions régissant la protection des données, la protection de l'information et le droit d'auteur ainsi que des secrets commerciaux.



et de les lire, par exemple au moyen d'un langage de requête structuré (SQL, de l'anglais *Structured Query Language*). Des bases de données relationnelles utilisant le langage SQL, faites de tableaux comprenant des lignes et des colonnes, de même que des tableaux classiques comprenant des lignes et des colonnes sont des exemples de données structurées;

- les données semi-structurées (ou partiellement structurées): données qui présentent quelques caractéristiques constantes et univoques, sans se limiter à une structure rigide, à l'instar de celle requise par exemple pour les bases de données relationnelles. Les photos en sont un exemple: lorsqu'une photo est prise par un smartphone par exemple, elle présente quelques attributs structurés, comme une géomarque, l'identité de l'appareil et une estampille temporelle. Une fois la photo enregistrée, il est possible de lui attribuer une étiquette, comme «animal domestique» ou «chat», en vue de pouvoir l'inclure dans une structure;
- les données non structurées: données à l'état entièrement brut. En raison de la complexité de leur ordonnancement et de leur formatage, ces données sont difficiles à traiter. L'absence de structure permet de conserver ensemble des données de différentes formes, par exemple des billets publiés sur les réseaux sociaux, des dialogues en ligne, des photos aériennes et des images par satellite<sup>3</sup>, des données provenant de capteurs IdO, des courriels et des présentations.

---

<sup>3</sup> Les photos aériennes et les images par satellite peuvent se présenter aussi bien sous forme structurée (matrice de mesure) que sous forme non structurée (images non calibrées).



### 3. Terminologie propre à l'«apprentissage à partir de (sources) de données»

Terme	Définition
<b>Science des données</b>	<p>La science des données (<i>data science</i> en anglais) est la <b>science</b> interdisciplinaire <b>de l'apprentissage à partir des données</b> (de la compréhension des données). Elle vise à tirer des enseignements des données pour pouvoir prendre des décisions fondées.</p> <p>La science des données couvre l'entier du processus qui va de la formulation de problèmes à la communication et à la publication des résultats, en passant par la saisie, la sélection, la préparation et l'analyse des données, ainsi que l'évaluation et l'interprétation des résultats.</p> <p>Contrairement à la statistique traditionnelle (et aux méthodes statistiques avancées), la démarche de la science des données est toutefois <b>inductive</b>, c'est-à-dire qu'elle se fait à partir des données ou que les données sont déterminantes (<i>data first</i>).</p> <p>Elle se concentre sur le <b>processus de résolution de problèmes</b> 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 («apprentissage automatique», «intelligence artificielle», etc.), des techniques et des pratiques innovantes qui lui sont propres.</p> <p>On dit parfois que la science des données est «l'<b>enfant de la statistique et de l'informatique</b>». 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 entité distincte. Son centre d'intérêt du moins la distingue de ses géniteurs.</p>

Terme	Définition
<b>Intelligence Artificielle (IA)</b>	<p>L'«intelligence artificielle» (<b>IA</b>, en anglais <b>Artificial Intelligence</b> ou <b>AI</b>), appelée parfois «intelligence machine» (<i>machine intelligence</i>), est définie comme «la création ou la programmation d'un ordinateur de telle sorte qu'il soit capable d'exécuter des fonctions qui reposent normalement sur des facultés humaines ou biologiques («intelligence»), comme la perception visuelle (reconnaissance d'images), la reconnaissance vocale, la traduction, la traduction visuelle et l'aptitude à jouer (en respectant des règles précises).</p> <p>L'IA désigne donc des machines «intelligentes» (<i>smart machines</i>) à même d'exécuter des tâches qui sont normalement accomplies par des êtres humains («machines apprenantes», <i>learning machines</i>). En d'autres termes, elle consiste à rendre les machines «intelligentes».</p>
<b>Système IA</b>	<p>Un système IA (AI system) est un système automatique capable d'inférer, sur la base des «inputs» (entrées) qu'il reçoit et pour des objectifs explicites ou implicites, comment générer des «outputs» (sorties) tels que des prévisions, des contenus, des recommandations ou des décisions, et qui, ce faisant, peut exercer une influence sur des environnements physiques ou virtuels. Les systèmes IA peuvent être dotés d'une autonomie plus ou moins grande.</p>
<b>Décisions basées sur l'IA</b>	<p>Les décisions basées sur l'IA sont les conclusions obtenues par des systèmes IA, qui ont des implications dans le monde réel.</p> <p>Elles dépendent toutefois toujours de décisions humaines, que ce soit au niveau de la conception du système, au niveau stratégique (décisions quant à l'utilisation du système) ou au niveau tactique (mode d'interaction avec la personne qui utilise le système).</p>
<b>Technologie d'IA</b>	<p>Une technologie d'IA désigne des fonctions isolées exécutables dans un ordinateur pour obtenir de l'IA («apprentissage automatique», p. ex.).</p> <p>Un système IA désigne ainsi une combinaison structurée et contextuelle de technologies d'IA en vue d'atteindre l'intelligence artificielle.</p>



Terme	Définition
<b>Traitement automatique du langage naturel (NLP)</b>	Le traitement automatique du langage naturel ( <b>NLP</b> de l'anglais <i>Natural Language Processing</i> ) est un domaine de l'IA qui traite des activités consistant à analyser, à comprendre et générer des mots et des phrases orales ou écrites (langue naturelle). La plupart des techniques et des méthodes de NLP se basent sur l'«apprentissage automatique» pour extraire la signification et les corrélations du langage humain. Les applications de NLP comprennent par exemple la reconnaissance de textes ( <i>text recognition</i> ), la reconnaissance de la parole ( <i>speech recognition</i> ), les robots logiciels, les agents conversationnels et les assistants numériques.
<b>Apprentissage automatique</b>	L'apprentissage automatique ou apprentissage machine ( <b>Machine Learning, ML</b> ) est un sous-domaine de l'IA qui «dote les ordinateurs de la capacité d'apprendre».  L'apprentissage automatique repose sur l'élaboration d'algorithmes qui, moyennant l'utilisation d'ordinateurs, analysent des données et, ce faisant, acquièrent de nouvelles connaissances, évoluent et s'améliorent (sur la base de règles concrètes établies par l'être humain). Le modèle statistique qui en résulte permet par exemple d'établir des prévisions et des classifications de données (non encore vues) pouvant être utilisées comme des aides à la décision.  L'apprentissage automatique est le principal sous-domaine de l'intelligence artificielle qui traite du problème de la pensée inductive. Son approche est donc celle du <b>raisonnement par induction</b> , dans lequel les données sont déterminantes ( <i>data first</i> ).
<b>IA générative</b>	L'«IA générative» est une notion large qui se réfère aux systèmes d'IA entraînés sur la base de grandes quantités de données provenant du monde physique et virtuel afin de générer des données de manière autonome (textes, images, enregistrements sonores, vidéos, simulations, codes, etc.). Ils sont souvent multimodaux, avec par exemple des inputs et/ou des outputs selon une ou plusieurs modalités (texte, image, vidéo, etc.).

Pour simplifier, on distingue trois méthodes d'apprentissage automatique: l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. Cette distinction est utile pour comprendre les différentes composantes



de l'apprentissage automatique. La recherche se concentre toutefois actuellement sur les possibilités de combiner ces différents types d'apprentissage automatique.

Terme	Définition
<p><b>Apprentissage supervisé</b></p>	<p>Dans le cas de l'apprentissage supervisé (<b><i>supervised learning</i></b>), l'algorithme d'apprentissage automatique suit un entraînement basé sur des données spécifiques et ses progrès sont évalués par un observateur humain ou à l'aide de données de référence.</p> <p>Les algorithmes de l'apprentissage automatique sont soumis à un processus itératif utilisant des données d'entraînement qui contiennent un grand nombre d'entrées (<i>inputs</i>) et une variable réponse ou sortie (<i>output</i>, «juste» ou «faux», p. ex.). Ce faisant l'algorithme apprend automatiquement des modèles (complexes) reliant les entrées et les sorties sans recevoir d'instructions explicites (d'un humain). Sur la base des liens appris, il peut ensuite prédire ou classer de nouvelles données (qu'il n'a pas encore vues).</p> <p>L'apprentissage profond (<b><i>deep learning</i></b>) est un sous-domaine de l'apprentissage supervisé qui utilise une forme particulière d'algorithmes appartenant au groupe des «réseaux de neurones artificiels».</p>
<p><b>Apprentissage non supervisé</b></p>	<p>L'apprentissage non supervisé (<b><i>unsupervised learning</i></b>) vise à repérer des corrélations et des dépendances parmi des données et, le cas échéant, d'en déduire des nouvelles entrées.</p> <p>Dans le cas de l'apprentissage non supervisé, les données d'entraînement sont constituées d'entrées (<i>inputs</i>) sans indications quant à la performance que l'on attend du système. C'est-à-dire qu'aucune sortie («juste» ou «faux», p. ex.) n'est prédéfinie.</p> <p>Contrairement à ce qui prévaut dans l'apprentissage supervisé, l'entraînement de l'algorithme repose uniquement sur des entrées. À partir de ces entrées, l'algorithme identifie de manière autonome des modèles dans les données. Cette méthode permet d'extraire des connaissances à partir de structures non visibles dans les données (p. ex. en utilisant le groupement de données ou <i>clustering</i>).</p> <p>Étant fondé sur l'observation, les expériences et les analogies, l'apprentissage non supervisé est similaire à l'apprentissage humain.</p>

Terme	Définition
<p><b>Apprentissage par renforcement</b></p>	<p>L'apprentissage par renforcement (<b><i>reinforcement learning</i></b>) désigne une série de méthodes d'apprentissage automatique qui consistent, pour un «agent», à apprendre une «stratégie» de manière autonome, afin de maximiser des récompenses reçues. Dans ce processus, rien ne permet à l'«agent» de savoir au préalable quelle action est la meilleure dans une situation donnée, mais il reçoit une récompense, positive ou négative, à des moments spécifiques. À l'aide de ces récompenses, l'«agent» détermine par approximation une «fonction d'utilité», qui décrit la valeur d'un certain état ou d'une certaine action.</p> <p>L'apprentissage par renforcement est principalement utilisé pour résoudre des problèmes de planification, comme la navigation (un «agent» doit par exemple atteindre un but et la récompense peut être associée avec le temps requis ou la distance à parcourir) et les jeux vidéo (l'«agent» doit p. ex. jouer et sa récompense est associée au résultat final du jeu).</p>



## 4. Terminologie accessoire

Terme	Définition
<b>Chaîne de blocs / blockchain</b>	Une chaîne de blocs (en anglais <i>blockchain</i> ) permet à un ensemble d'acteurs de gérer un registre commun fiable, même en l'absence d'autorité centrale digne de confiance. La technologie dite de chaîne de blocs est aussi appelée technologie de registre distribué ( <i>Distributed Ledger Technology, DLT</i> ).
<b>Dialogueurs / agents conversationnels</b>	Les agents de dialogue sont des systèmes de dialogue capables d'utiliser le langage naturel (écrit ou oral). Ils sont souvent utilisés en association avec des avatars statiques ou animés, sur des sites Web ou dans des systèmes de messagerie instantanée, où ils guident les personnes intéressées ou les clients vers les prestations ou les produits appropriés, s'enquière de leurs problèmes ou répondent à des questions.
<b>Informatique en nuage / services en nuage ou infonuagiques</b>	L'informatique en nuage désigne la mise à disposition de ressources informatiques (serveurs, capacité mémoire, base de données, éléments de réseaux, logiciels, fonctions d'analyse, fonctions intelligentes, etc.) via Internet. L'utilisateur ne paie généralement que les services infonuagiques ( <i>services</i> ) qu'il utilise vraiment (facturation à l'utilisation ou <i>pay as you go</i> ). Les prestations offertes sous forme de «services en nuage» sont en général très standardisées. Elles sont toutefois disponibles extrêmement rapidement (souvent en l'espace de quelques minutes).
<b>Internet des Objets (IdO)</b>	L'«internet des objets» (IdO, en anglais <b>Internet of Things, IoT</b> ) est une expression qui englobe les technologies d'une infrastructure globale de la société de l'information, qui permet d'interconnecter des objets physiques et virtuels et de les faire collaborer en recourant à des technologies de l'information et de la communication.



## 5. Principes fondamentaux pour une science des données (et une IA) centrée sur l'être humain et digne de confiance<sup>4</sup>

Principe fondamental	Définition et explication
<b>Protection des données et de l'information</b>	<p>L'objectif de la <b>protection des données</b> (<i>privacy</i>) est de protéger la sphère privée de tout individu. Dans ce sens, la protection des données protège contre l'utilisation abusive des données («autodétermination en matière d'information») et spécifie si certaines données peuvent être collectées et traitées. Elle définit le traitement autorisé de ces données. La protection des données sert tout spécialement à protéger les données personnelles, qui sont des informations se rapportant à une personne identifiée ou identifiable. La protection des données est une protection «abstraite» des données personnelles contre une utilisation abusive (voir en comparaison la protection «pratique» sous «sécurité des données»).</p> <p>La <b>protection de l'information</b> définit les niveaux «abstraites» de confidentialité de l'information («non classifié», «interne», «confidentiel», «secret») destinés à préserver les intérêts d'un pays ou d'une organisation.</p>
<b>Sécurité de l'information</b>	<p>La <b>sécurité de l'information</b> consiste à garantir l'intégrité, la disponibilité et la confidentialité des informations et des systèmes qui les traitent, indépendamment de la manière dont les informations sont présentées, stockées et utilisées. Elles sont ainsi protégées contre la perte, la falsification et la divulgation non souhaitée.</p>
<b>Sécurité des données</b>	<p>La <b>sécurité des données</b> concerne la protection des voies d'acheminement et de l'utilisation de données de toute nature, c'est-à-dire pas seulement des données personnelles. L'intégrité, la disponibilité, la confidentialité et l'authenticité des données sont garanties par des solutions techniques ou des mesures organisationnelles, personnelles et physiques locales. Il s'agit d'une protection «pratique» des données de tout type.</p>

<sup>4</sup>«Code de bonnes pratiques de la Confédération pour une science des données centrée sur l'être humain et digne de confiance»: <https://www.bfs.admin.ch/bfs/fr/home/dscc/dscc.assetdetail.29325685.html> & <https://www.bfs.admin.ch/bfs/de/home/dscc/dscc.assetdetail.28405275.html>



Principe fondamental	Définition et explication
<b>Gouvernance des données</b>	Alors que la «gestion des données» met en œuvre la gestion technique des données, la <b>gouvernance des données</b> prescrit les directives et les procédures internes qui régissent la gestion des données au sein d'une organisation. La gouvernance des données définit donc des rôles et des responsabilités et établit des processus pour le traitement des données au sein d'une organisation. Elle a pour objectif de garantir la qualité, l'intégrité et la disponibilité des données dans un système (de traitement de l'information). Elle englobe également l'harmonisation et la standardisation des données, tout en devant garantir que ces données sont utilisées conformément aux lois sur la protection des données et aux normes éthiques en vigueur.
<b>Non-discrimination</b>	La <b>non-discrimination</b> signifie qu'il est interdit de désavantager des personnes ou des groupes de personnes sur la base de certaines caractéristiques. Elle vise en particulier toute discrimination ou préférence fondée sur un handicap, le sexe, la filiation, la langue, l'origine ethnique ou sociale, les caractéristiques génétiques, l'orientation sexuelle, la religion ou la vision du monde, ainsi que les opinions politiques ou autres opinions d'une personne. Il s'agit donc d'une protection contre les inégalités de traitement discriminatoires, qui sont interdites par la loi. Ce principe comprend aussi bien la discrimination directe qu'indirecte (c'est-à-dire indirecte ou dissimulée).
<b>Explicabilité</b>	L'explicabilité se réfère à la capacité d'une approche de résolution de problèmes en science des données à expliquer ses résultats et les connaissances acquises d'une manière compréhensible pour les êtres humains. Les exigences concrètes en matière d'explicabilité varient en fonction du groupe cible et du contexte. Elles désignent par exemple l'obligation de divulguer les algorithmes et les méthodes utilisés ainsi que les sources et les traitements de données sur lesquels ils se basent ( <i>completeness</i> ), de fournir des explications claires et compréhensibles pour les êtres humains ( <i>interpretability</i> ) ou d'améliorer continuellement l'explicabilité (dans la mesure où cela est autorisé par la législation sur la protection des données et d'autres directives). Dans l'ensemble, l'explicabilité vise à garantir que les approches de résolution de problèmes tirées de la science des données sont équitables, dignes de confiance et acceptables et que leur utilisation correspond aux intérêts de la société.

Principe fondamental	Définition et explication
<b>Traçabilité</b>	<p>La <b>traçabilité</b> fait référence à la possibilité de suivre, de vérifier et de corriger les aides à la décision basées sur les données, obtenues grâce aux méthodes, aux techniques et aux pratiques issues de la science des données. Ces aides à la décision (sous forme de recommandations, p. ex.) doivent être documentées de manière claire, traçable et compréhensible et leur impact sur le contexte dans lequel elles sont appliquées doit être compris. Les exigences concrètes en matière de traçabilité varient selon le groupe cible et le contexte. Il s'agit par exemple de l'obligation de documenter les approches de résolution de problèmes basées sur la science des données, de fournir des explications claires et compréhensibles sur les outils d'aide à la décision basés sur les données, ou de fournir un retour d'information et des corrections afin de garantir que les processus de prise de décision sont continuellement améliorés et que les outils d'aide à la décision basés sur les données deviennent de plus en plus précis et fiables. La traçabilité vise donc à garantir que les aides à la décision basées sur des données résultant de l'application de la science des données sont cohérentes et fiables et qu'elles répondent aux besoins et aux exigences du contexte dans lequel elles sont appliquées.</p>
<b>Transparence</b>	<p>La <b>transparence</b> désigne l'obligation d'exposer et de communiquer ouvertement l'utilisation de méthodes, de techniques et de pratiques issues de la science des données, ainsi que les étapes de la résolution de problèmes et les décisions qui y sont liées. Les exigences concrètes en matière de transparence varient en fonction du groupe cible. Par exemple, cela implique l'obligation de divulguer les sources et le traitement des données, ainsi que les méthodes et les algorithmes utilisés, ou l'obligation d'expliquer les décisions et les processus (dans la mesure où la législation sur la protection des données et d'autres dispositions le permettent). La transparence vise donc à donner une image claire de l'ensemble du processus de résolution de problèmes en science des données et de ses différentes étapes. Elle sert ainsi à garantir que les approches de résolution de problèmes en science des données sont équitables et transparentes et que l'application de la science des données au sein de l'administration fédérale peut gagner la confiance du public.</p>

Principe fondamental	Définition et explication
<b>Reproductibilité</b>	<p>La <b>reproductibilité</b> se réfère à l'obligation de faire en sorte que les résultats et les connaissances acquises via l'application de méthodes, de techniques et de pratiques issues de la science des données puissent être reproduits par d'autres personnes. Elle inclut l'obligation de documenter et de communiquer clairement les données, les sources de données, les algorithmes et les méthodes utilisés, ainsi que les résultats et les conclusions, afin que d'autres personnes soient en mesure de vérifier et de reproduire les résultats de manière indépendante (dans la mesure où la législation sur la protection des données et d'autres directives le permettent). Elle comprend également l'obligation de fournir des informations détaillées sur la reproduction des résultats afin de permettre des comparaisons entre différentes méthodes et approches. Associée à tous les autres principes fondamentaux énumérés ici, la reproductibilité vise donc à garantir que les approches de résolution de problèmes en science des données reposent sur une base solide et que les résultats sont fiables et pertinents.</p>
<b>Neutralité</b>	<p>En général, la <b>neutralité</b> signifie qu'une personne ou une organisation agit de manière impartiale, sans idées préconçues et sans préjugés ni intérêts personnels susceptibles d'influencer ses décisions. Cela vaut également pour une personne ou une organisation lorsqu'elle utilise la science des données. La divulgation de conflits d'intérêts potentiels est un exemple de promotion de la neutralité.</p>



Principe fondamental	Définition et explication
<b>Objectivité</b>	L' <b>objectivité</b> vise à garantir que les approches de résolution de problèmes reposant sur la science des données sont impartiales et neutres et que les décisions et les résultats sont basés sur des critères objectifs et indépendants. Une personne ou une organisation qui développe, applique et diffuse des méthodes, des techniques et des pratiques de résolution de problèmes basées sur la science des données le fait en préservant son indépendance scientifique et en travaillant de manière objective, professionnelle et transparente.
<b>Traitement éthique des données et des résultats</b>	Le <b>traitement éthique des données et des résultats</b> vise à ce que les données et les résultats soient utilisés de manière éthique, responsable et durable, afin de garantir le respect des droits et de la dignité des personnes et de tenir compte de l'impact de cette utilisation sur la société et l'environnement. Ce principe comprend par exemple l'obligation de garantir la protection et la sécurité des données, de promouvoir la transparence et la franchise ou de veiller à ce que l'utilisation des données et des résultats soit conforme à des principes et à des normes éthiques qui respectent la dignité humaine, l'autonomie et la sphère privée. Son respect inclut la prévention de la discrimination et des préjugés, ainsi que l'obligation de tenir compte de l'impact que les données et les résultats de toute approche de résolution de problèmes basée sur la science des données produisent sur les parties prenantes ( <i>stakeholders</i> ) et sur la société et l'environnement en général.