



Terminologie

Kompetenznetzwerk CNAI

Autor	Geschäftsstelle CNAI
Version	2.1
Datum	21.12.2023

Geschäftsstelle CNAI

Bundesamt für Statistik BFS
Espace de l'Europe 10
CH - 2010 Neuchâtel

[CNAI.swiss](https://cnaai.swiss)

CNAI@BFS.admin.ch

INHALTSVERZEICHNIS

1. ZIEL DIESES DOKUMENTS UND GELTUNGSBEREICH	3
2. TERMINOLOGIE IN BEZUG AUF DATEN(QUELLEN)	4
3. TERMINOLOGIE IN BEZUG AUF «LERNEN AUS DATEN(QUELLEN)»	6
4. ZUSÄTZLICHE TERMINOLOGIE.....	11
5. GRUNDPRINZIPIEN FÜR MENSCHENZENTRIERTE UND VERTRAUENSWÜRDIGE DATENWISSENSCHAFT (UND KI)	12



1. Ziel dieses Dokuments und Geltungsbereich

Mit der Etablierung eines Kompetenznetzwerks für künstliche Intelligenz («*Competence Network for Artificial Intelligence*» – CNAI) soll der Einsatz von künstlicher Intelligenz und anderen neuen Technologien innerhalb der Bundesverwaltung (und darüber hinaus) nachhaltig gefördert werden.

Eines der Ziele des Kompetenznetzwerks CNAI ist die Einführung einer einheitlichen Terminologie zum Erreichen eines gemeinsamen Verständnisses für die im Zusammenhang mit künstlicher Intelligenz und neuen Technologien relevanten Begriffe auf Ebene der Bundesverwaltung.

Eine gemeinsame Sprache erleichtert den aktiven Erfahrungs- und Wissensaustausch innerhalb des Kompetenznetzwerks CNAI und darüber hinaus. Zudem vereinfacht sie die Kommunikation bezüglich Ideen, Vorhaben und Dienstleistungen in diesem Bereich.

Das entsprechende Terminologie-Management seitens der Geschäftsstelle CNAI stellt somit einen zentralen Grundstein für die Funktionen des CNAI im Allgemeinen und seiner Geschäftsstelle im Speziellen dar.

Dieses Terminologie-Management basiert auf den im BFS und dessen Kompetenzzentrum für Datenwissenschaft («*Data Science Competence Center*» – DSCC) bereits vorhandenen Dokumenten und auf den Dokumenten, welche im Rahmen der Bundesratsentschlüsse zum Aufbau des Netzwerks CNAI, inklusive Geschäftsstelle, erarbeitet wurden.

Ergänzungen bereits erklärter Begriffe sollen nur dann erfolgen, wenn sie für die Abgrenzung zu anderen neuen Technologien erforderlich sind. Das Glossar wird in Zusammenarbeit mit den Netzwerkknotenpunkten zielführend weiterentwickelt und soll die zukünftige Begriffswelt des CNAI veranschaulichen.

Besitzerin des Dokuments ist die Geschäftsstelle CNAI.

Das vorliegende Dokument wird periodisch überarbeitet und die aktuelle Version (Versionsnummer und Datum als Referenz) hat Gültigkeit.

Schreibkonvention

Solange ein Begriff nicht explizit definiert wurde, wird der Begriff in Anführungszeichen gesetzt, z. B. «Datenwissenschaft».



2. Terminologie in Bezug auf Daten(quellen)

Im Allgemeinen können **drei Arten von Datenquellen** unterschieden werden:

- verwaltungsinterne (z. B. «statistische Daten»¹, administrative Daten, amtliche Daten, Geodaten des Bundes, *Open Government Data* – OGD²) und
- verwaltungsexterne Datenquellen (z. B. offene Daten) sowie
- cyberphysische Systeme (wie Daten von vernetzten Sensoren im Rahmen des «Internets der Dinge» – IoT), welche sowohl eine verwaltungsinterne als auch eine verwaltungsexterne Datenquelle sein können.

Die integrierte Verwendung der verwaltungsinternen und -externen Datenquellen zusammen auch mit Daten aus cyberphysischen Systemen eröffnet der Daten(r)evolution für politische Entscheidungsträger ein enormes Potential. Gerade mit Blick auf eine Nutzung dieses Datenpotentials unter der Verwendung «datenwissenschaftlicher Methoden» kann ein grosses Anwendungspotenzial für die «Datenwissenschaft» erschlossen werden. So können bspw. Datenquellen zusammengeführt werden, um bisher nicht abgedeckte Nutzerbedürfnisse zu erfüllen oder um grundsätzlich abgedeckte Nutzerbedürfnisse besser zu erfüllen. Daraus ergeben sich neue Möglichkeiten, um möglichst zeitnah zu verschiedenen Problemstellungen relevante Informationen zu generieren und den politischen Entscheidungsträgern zur Verfügung zu stellen.

Zudem können im Allgemeinen **drei Arten von Datentypen** unterschieden werden:

- strukturierte Daten, die formatiert und in ein wohldefiniertes Datenmodell umgewandelt wurden. Die Rohdaten werden in vordefinierte Felder abgebildet, die dann einfach extrahiert und, z. B. über SQL, gelesen werden können. Relationale SQL-Datenbanken, bestehend aus Tabellen mit Zeilen und Spalten, und klassische Tabellen mit Zeilen und Spalten sind Beispiele für strukturierte Daten; und
- semi-strukturierte (oder halbstrukturierte oder teilweise strukturierte) Daten, welche einige konsistente und eindeutige Merkmale aufweisen, ohne sich auf eine starre Struktur zu beschränken, wie sie z. B. für relationale Datenbanken erforderlich ist. Ein Beispiel sind Bilder: wenn ein Bild beispielsweise von einem Smartphone aufgenommen wird, weist es einige

¹ Statistische Daten sind alle Daten, welche zu statistischen Zwecken erhoben oder weitergegeben wurden, insbesondere gestützt auf das Bundesstatistikgesetz und unterliegen damit dem Statistikgeheimnis.

² Offene Daten meint Daten, die frei, ohne wesentliche rechtliche, finanzielle oder technische Einschränkungen, genutzt, verarbeitet, ausgewertet und weitergegeben werden dürfen. Rechtlich muss die kostenfreie Nutzung und Weiterverarbeitung der Daten gewährleistet sein; technische Offenheit betont, dass offene Daten maschinell bearbeitbar sein müssen. Bei der Publikation von Daten als offenen Daten müssen Datenschutz-, Informationsschutz- und Urheberrechtsbestimmungen sowie Geschäftsgeheimnisse gewahrt bleiben.

strukturierte Attribute wie Geotag, Geräte-ID und Zeitstempel auf. Nach dem Speichern können Bildern auch Tags wie «Haustier» oder «Katze» zugewiesen werden, um eine Struktur bereitzustellen; und

- unstrukturierte Daten sind Daten, welche in absoluter Rohform vorliegen. Diese Daten sind aufgrund ihrer komplexen Anordnung und Formatierung schwer zu verarbeiten. Die unstrukturierte Datenhaltung kann Daten aus vielen Formen annehmen, darunter Social-Media-Posts, Chats, Luft- und Satellitenbilder³, IoT-Sensordaten, E-Mails und Präsentationen.

³ Luft- und Satellitenbilder können sowohl in strukturierter Form (Sichtweise einer Messmatrix) wie in unstrukturierter Form (unkalibrierte Bilder) vorliegen.

3. Terminologie in Bezug auf «Lernen aus Daten(quellen)»

Begriff	Definition
Datenwissenschaft	Datenwissenschaft («<i>Data Science</i>») ist die interdisziplinäre Wissenschaft des Lernens aus Daten (des Datenverstehens), mit dem Ziel, Erkenntnisse aus Daten zu gewinnen, auf deren Basis datenbasierte Entscheidungsgrundlagen erstellt werden können. Die Datenwissenschaft umfasst, wie die Statistik, den gesamten Prozess der Problemformulierung, der Erfassung, Auswahl, Vorbereitung und Analyse von Daten sowie der Evaluation, Interpretation, Kommunikation und Bereitstellung der gewonnenen Erkenntnisse. Im Gegensatz zu traditioneller (und erweiterter) Statistik ist die Vorgehensweise der Datenwissenschaft aber das induktive Vorgehen, das von Daten ausgeht, d. h. «<i>data first</i>». Im Zentrum der Datenwissenschaft stehen ein Problemlösungsprozess sowie ein Prozess der kontinuierlichen Verbesserung, der darauf abzielt, komplexe, unstrukturierte und datenreiche Probleme durch die Anwendung innovativer datenwissenschaftlicher Methoden (z. B. Methoden aus dem «maschinellen Lernen» und dem Bereich der «künstlichen Intelligenz»), Techniken und Praktiken zu lösen. Datenwissenschaft wird auch als «das Kind von Statistik und Informatik» bezeichnet. Diese Kind-Metapher schlussfolgert angemessen, dass die Datenwissenschaft (idealerweise das Beste) von beiden Elternteilen erbt, aber schließlich zu einer eigenen Entität heranwächst. Ihr Fokus trennt sie von ihren Eltern.

Begriff	Definition
Künstliche Intelligenz (KI)	Künstliche Intelligenz (KI – «Artificial Intelligence» – AI), heute manchmal als «maschinelle Intelligenz» («Machine Intelligence») bezeichnet, wird definiert als «einen Computer so bauen oder programmieren, um Dinge zu tun, die normalerweise menschliche oder biologische Fähigkeiten («Intelligenz») erfordern», z.B. visuelle Wahrnehmung (Bilderkennung), Spracherkennung, Sprachübersetzung, visuelle Übersetzung und Spiele spielen (mit konkreten Regeln). Bei KI geht es um «intelligente» Maschinen («<i>smart machines</i>»), die Aufgaben ausführen können, die normalerweise von Menschen ausgeführt werden («lernende Maschinen»; «<i>learning machines</i>»), d. h. Maschinen «intelligent» machen.
KI-System	Ein KI-System ist ein maschinenbasiertes System, das für explizite oder implizite Ziele aus den empfangenen Inputs schlussfolgert, wie es Outputs wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erzeugen kann, welche die physische oder virtuelle Umgebung beeinflussen können. KI-Systeme können mit unterschiedlichem Ausmass an Autonomie ausgestattet werden.
KI-Entscheidungen	KI-Entscheidungen sind Schlussfolgerungen von KI-Systemen mit realweltlichen Auswirkungen. Diese sind auf der Ebene des Designs des Systems, der strategischen Ebene (Entscheid über Einsatz des Systems) und der taktischen Ebene (Ausgestaltung der Interaktion mit der Person, die das System nutzt) von menschlichen Entscheidungen abhängig.
KI-Technologie	KI-Technologie bezeichnet einzelne, in Computer implementierbare Funktionen für die Erreichung von KI (z. B. «maschinelles Lernen»). Ein KI-System bezeichnet somit eine strukturierte, kontextgebundene Kombination von KI-Technologien zwecks Erreichens von KI.

Begriff	Definition
«Natural Language Processing» (NLP)	«Natural Language Processing» (NLP) ist ein Teilgebiet der KI, welches sich mit der Analyse, dem Verständnis und der Generierung von geschriebenen und gesprochenen Wörtern und Sätzen (natürlicher Sprache) beschäftigt. Die meisten NLP Techniken und Methoden basieren auf «maschinellern Lernen» und extrahieren so die Bedeutung und Zusammenhänge aus menschlicher Sprache. Anwendungsgebiete sind z. B. Texterkennung («Text Recognition»), Spracherkennung («Speech Recognition»), Bots, «Chatbots» und digitale Assistenten.
Generative KI	«Generative KI» ist ein weit gefasster Begriff, der sich auf KI-Systeme bezieht, die auf grosse Mengen von Daten aus der physischen und virtuellen Welt trainiert werden, um selbst Daten zu generieren (z.B. Texte, Bilder, Tonaufnahmen, Videos, Simulationen, Codes). Sie sind oft multimodal, z.B. mit Inputs und/oder Outputs in einer oder mehreren Modalitäten (z.B. Text, Bild, Video). Verschiedene «Modellarchitekturen», einschliesslich «Diffusionsmodelle» und «Transformer-basierte Modelle», können für generative Aufgaben verwendet werden.
Maschinelles Lernen (ML)	Maschinelles Lernen (ML – «Machine Learning») ist ein weiteres Teilgebiet der KI, welches «den Computern die Lernfähigkeit verleiht». ML untersucht die Konstruktion von Algorithmen, die durch den Einsatz von Computern Daten analysieren und dabei automatisch lernen, sich anpassen und verbessern (anhand von konkreten vom Menschen vorgegebenen Regeln). Das resultierende statistische Modell ermöglicht bspw. Vorhersagen und Klassifizierungen von (noch nicht gesichteten) Daten, welche entscheidungsunterstützend eingesetzt werden können. In der KI ist ML das wichtigste Teilgebiet, das sich mit dem Problem des induktiven Denkens auseinandersetzt. Die Vorgehensweise von ML ist somit das induktive Vorgehen, das von Daten ausgeht, d. h. «data first».

ML kann grob in drei verschiedene Teilbereiche unterteilt werden: überwachtes, unüberwachtes und verstärktes Lernen. Diese Trennung ist nützlich, um die grundlegenden ML Bausteine zu verstehen. Die aktuelle Forschung untersucht allerdings, wie man/frau die verschiedenen ML Arten miteinander kombinieren kann.

Begriff	Definition
Überwachtes Lernen	Beim überwachten Lernen («supervised learning») wird der ML-Algorithmus mit Daten trainiert, und ein menschlicher Beobachter bzw. Benchmark-Daten überwachen den Lernerfolg. Konkret werden ML-Algorithmen mithilfe eines Trainingsdatensatzes, welches eine Vielzahl an «<i>Features</i>» (Inputs) und eine «Zielvariable» (Output, z. B. «richtig» oder «falsch») beinhaltet, in einem iterativen Prozess trainiert. Dabei lernt der Algorithmus automatisch (komplexe) Muster zwischen den «<i>Features</i>» und der Zielvariable ohne explizite (menschliche) Anweisungen. Mithilfe der gelernten Muster können für (noch nicht gesichtete) Daten Vorhersagen oder Klassifizierungen vorgenommen werden. Bspw. ist tiefes Lernen («Deep Learning») eine Teilsparte des überwachten Lernens und verwendet eine spezialisierte Form von ML-Algorithmen aus der Gruppe der «künstlichen neuronalen Netze».
Unüberwachtes Lernen	Das unüberwachte oder unbeaufsichtigte Lernen («unsupervised learning») zielt darauf ab, Zusammenhänge und Abhängigkeiten in Daten ausfindig zu machen und diese gegebenenfalls zu Merkmalen weiter zu verarbeiten. Beim unüberwachten Lernen bestehen die Trainingsdaten aus Eingabedaten (Inputs), ohne dass ein Hinweis auf die erwartete Leistung des Systems vorliegt, d. h. eine Zielvariable (z. B. «richtig» oder «falsch») liegt nicht vor. Im Gegensatz zum überwachten Lernen wird der ML-Algorithmus somit nur mit «<i>Features</i>» trainiert. Auf Basis der «<i>Features</i>» entdeckt der Algorithmus selbstständig Muster in den Daten. Dies ermöglicht es von nicht sichtbaren Strukturen in den Daten zu lernen (bspw. mittels «<i>Clustering</i>»). Unüberwachtes Lernen ähnelt dem menschlichen Lernen; durch Beobachtungen, Erfahrungen und Analogien.

Begriff	Definition
Verstärktes Lernen	Das verstärkte Lernen («reinforcement learning») steht für eine Reihe von ML-Methoden, bei denen ein «Agent» selbstständig eine «Strategie» erlernt, um erhaltene Belohnungen zu maximieren. Dabei wird dem «Agenten» nicht vorgezeigt, welche Aktion in welcher Situation die beste ist, sondern er erhält zu bestimmten Zeitpunkten eine Belohnung, die auch negativ sein kann. Anhand dieser Belohnungen approximiert er eine «Nutzenfunktion», die beschreibt, welchen Wert ein bestimmter Zustand oder Aktion hat. Verstärktes Lernen findet bspw. seine Hauptanwendung in Planungsproblemen wie der Navigation (z. B. der «Agent» muss ein Ziel erreichen, und die Belohnung kann mit Zeit oder Entfernung verknüpft werden) und Videospielen (z. B. der «Agent» muss spielen, und seine Belohnung ist mit dem Endergebnis des Spiels verknüpft).

4. Zusätzliche Terminologie

Begriff	Definition
Blockchain	Eine «Blockchain» erlaubt es einem Kollektiv von Akteuren, auch ohne vertrauenswürdige zentrale Instanz einen vertrauenswürdigen gemeinsamen «Ledger» zu führen. Blockchaintechnologie wird auch als « <i>Distributed Ledger</i> »-Technologie (DLT) bezeichnet.
Chatbots	<i>Chatbots</i> sind Dialogsysteme mit natürlichsprachlichen Fähigkeiten (Text oder Audio). Sie werden, oft in Kombination mit statischen oder animierten Avataren, auf Websites oder in Instant-Messaging-Systemen verwendet, wo sie Interessenten und Kunden zu relevanten Produkten oder Dienstleistungen führen respektive sich um Anliegen der Interessenten und Kunden kümmern oder Fragen beantworten.
Cloud Computing / Cloud Services	<i>Cloud Computing</i> ist die Bereitstellung von Computingressourcen (z.B. Server, Speicher, Datenbanken, Netzwerkkomponenten, Software, Analyse- und intelligente Funktionen) über das Internet. Dabei wird in der Regel nur für die Clouddienste (« <i>Services</i> ») gezahlt (« <i>pay as you go</i> »), die tatsächlich genutzt werden. In der Regel werden hoch standardisierte Leistungen als <i>Cloud-Services</i> angeboten. Im Gegenzug sind diese Leistungen extrem schnell verfügbar (oft innerhalb von Minuten).
Internet der Dinge (IoT)	Das Internet der Dinge (« Internet of Things » – IoT) ist ein Sammelbegriff für Technologien einer globalen Infrastruktur der Informationsgesellschaften, die es ermöglicht, physische und virtuelle Gegenstände miteinander zu vernetzen und sie durch Informations- und Kommunikationstechniken zusammenarbeiten zu lassen.

5. Grundprinzipien für menschenzentrierte und vertrauenswürdige Datenwissenschaft (und KI)⁴

Grundprinzip	Definition und Erläuterung
Daten- und Informationsschutz	Zweck des Datenschutzes («privacy») ist der Schutz der Privatsphäre eines jeden Menschen. Der Datenschutz schützt in diesem Sinne vor der missbräuchlichen Verwendung von Daten («informationelle Selbstbestimmung») und beantwortet die Frage, ob bestimmte Daten erhoben und verarbeitet werden dürfen. Er definiert den erlaubten Umgang mit diesen Daten. Personendaten sind das zentrale Schutzobjekt des Datenschutzrechts. Sie sind Angaben, die sich auf eine bestimmte oder bestimmbare Person beziehen. Beim Datenschutz handelt sich um einen «abstrakten» Schutz der persönlichen Daten vor Missbrauch (siehe im Vergleich dazu den «praktischen» Schutz unter «Datensicherheit»). Der Informationsschutz legt, im Hinblick auf die Wahrung der Interessen eines Landes oder einer Organisation, die «abstrakten» Vertraulichkeitsstufen für Informationen fest («nicht klassifiziert», «intern», «vertraulich», «geheim»).
Informationssicherheit	Informationssicherheit steht für die Sicherstellung der Integrität, Verfügbarkeit und Vertraulichkeit von Informationen und informationsverarbeitenden Systemen, ungeachtet der Art ihrer Darstellung, Speicherung und Verwendung. Informationen werden dadurch vor Verlust, Verfälschung und unerwünschter Offenlegung geschützt.
Datensicherheit	Die Datensicherheit betrifft den Schutz der Wege und Verwendung von Daten aller Art, d.h. nicht nur von personenbezogenen Daten. Die Integrität, Verfügbarkeit, Vertraulichkeit und Authentizität der Daten werden durch technische Lösungen oder organisatorische, personalbezogene sowie lokal-physische Massnahmen gesichert. Es handelt sich um einen «praktischen» Schutz von Daten aller Art.

⁴ «Verhaltenskodex des Bundes für menschenzentrierte und vertrauenswürdige Datenwissenschaft (und KI)»: <https://www.bfs.admin.ch/bfs/de/home/dscc/dscc.assetdetail.29325686.html> & <https://www.bfs.admin.ch/bfs/de/home/dscc/dscc.assetdetail.28405274.html>

Grundprinzip	Definition und Erläuterung
Datengouvernanz	Während «Datenmanagement» die technische Verwaltung von Daten umsetzt, schreibt die Datengouvernanz die internen Richtlinien und Verfahren für die Verwaltung der Daten innerhalb einer Organisation vor. Datengouvernanz definiert somit Rollen wie auch Verantwortlichkeiten und legt Prozesse für den Umgang mit Daten in einer Organisation fest. Das Ziel von Datengouvernanz ist es, Qualität, Integrität und Verfügbarkeit von Daten in einem (informationsverarbeitenden) System zu gewährleisten. Zur Datengouvernanz gehört auch die Harmonisierung und Standardisierung der Daten. Gleichzeitig soll die Datengouvernanz sicherstellen, dass diese Daten in Übereinstimmung mit geltenden Datenschutzgesetzen und ethischen Standards verwendet werden.
Nichtdiskriminierung	Nichtdiskriminierung bedeutet, dass Personen oder Personengruppen nicht aufgrund bestimmter Merkmale benachteiligt werden dürfen. Darunter fallen die Benachteiligung oder Bevorzugung wegen einer Behinderung, des Geschlechts, der Abstammung, der Sprache, der ethnischen oder sozialen Herkunft, der genetischen Merkmale, der sexuellen Ausrichtung, der Religion oder der Weltanschauung, als auch der politischen oder sonstigen Anschauung einer Person. Es geht somit um einen Schutz vor diskriminierenden Ungleichbehandlungen, welche gesetzlich verboten sind. Dies umfasst sowohl direkte Diskriminierung als auch indirekte (d.h. mittelbare oder versteckte) Diskriminierung.
Erklärbarkeit	Erklärbarkeit bezieht sich auf die Fähigkeit eines datenwissenschaftlichen Problemlösungsansatzes seine Resultate und gewonnen Erkenntnisse in einer für Menschen verständlichen Weise zu erklären. Die konkreten Anforderungen an Erklärbarkeit unterscheiden sich je nach Zielgruppe und Kontext. Beispielsweise beinhaltet dies die Verpflichtung zur Offenlegung der verwendeten Algorithmen und Methoden und der Datenquellen und -verarbeitungen, auf denen sie basieren («completeness»), zur Bereitstellung von klaren und verständlichen Erklärungen für Menschen («interpretability»), oder zur kontinuierlichen Verbesserung der Erklärbarkeit (soweit gemäss datenschutzrechtlichen und anderen Vorgaben zulässig). Insgesamt zielt die Erklärbarkeit darauf ab, sicherzustellen, dass datenwissenschaftliche Problemlösungsansätze fair, vertrauenswürdig und akzeptabel sind und dass sie auf eine Weise eingesetzt werden, die den Interessen der Gesellschaft entspricht.

Grundprinzip	Definition und Erläuterung
Nachvollziehbarkeit	Nachvollziehbarkeit (oder «Rückführbarkeit») bezieht sich darauf, dass die datenbasierten Entscheidungsunterstützungen, die auf der Grundlage der Anwendung von datenwissenschaftlichen Methoden, Techniken und Praktiken getroffen werden, nachverfolgt, überprüft und korrigiert werden können. Solche Entscheidungsunterstützungen (z.B. in Form von Empfehlungen) sollen auf eine klare, nachvollziehbare und verständliche Weise dokumentiert und ihre Auswirkungen auf den Kontext, in dem sie angewendet werden, verstanden werden. Die konkreten Anforderungen an Nachvollziehbarkeit unterscheiden sich je nach Zielgruppe und Kontext. Beispielsweise beinhaltet dies die Verpflichtung zur Dokumentation von datenwissenschaftlichen Problemlösungsansätzen, die Bereitstellung von klaren und verständlichen Erklärungen von datenbasierten Entscheidungsunterstützungen, oder die Bereitstellung von Rückmeldungen und Korrekturen, um sicherzustellen, dass die Entscheidungsprozesse kontinuierlich verbessert werden und dass die datenbasierten Entscheidungsunterstützungen immer genauer und zuverlässiger werden. Somit zielt die Nachvollziehbarkeit darauf ab, dass datenbasierte Entscheidungsunterstützungen resultierend aus der Anwendung von Datenwissenschaft konsistent und zuverlässig sind und dass sie die Bedürfnisse und Anforderungen des Kontexts, in dem sie angewendet werden, erfüllen.
Transparenz	Transparenz bezieht sich auf die Verpflichtung, dass die Anwendung von datenwissenschaftlichen Methoden, Techniken und Praktiken sowie die damit verbundenen Schritte der Problemlösungsfindung und Entscheidungen offen ersichtlich dargelegt und kommuniziert werden. Die konkreten Anforderungen an Transparenz unterscheiden sich je nach Zielgruppe. Beispielsweise beinhaltet dies die Verpflichtung zur Offenlegung von Datenquellen und -verarbeitung, Algorithmen und Methoden, oder die Verpflichtung zur Erklärung von Entscheidungen und Prozessen (soweit gemäss datenschutzrechtlichen und anderen Vorgaben zulässig). Somit zielt die Transparenz darauf ab, ein klares Bild vom gesamten Problemlösungsprozess der Datenwissenschaft und auch von dessen einzelnen Schritten zu geben. So wird sichergestellt, dass datenwissenschaftliche Problemlösungsansätze fair und transparent sind und dass die Anwendung von Datenwissenschaft in der Bundesverwaltung das Vertrauen der Öffentlichkeit gewinnen kann.

Grundprinzip	Definition und Erläuterung
Reproduzierbarkeit	Reproduzierbarkeit bezieht sich auf die Verpflichtung, dass die Ergebnisse und Erkenntnisse, die durch die Anwendung von datenwissenschaftlichen Methoden, Techniken und Praktiken gewonnen werden, durch andere Personen reproduzierbar sind. Dies beinhaltet die Verpflichtung, dass die verwendeten Daten, Datenquellen, Algorithmen und Methoden sowie die Ergebnisse und Schlussfolgerungen klar dokumentiert und kommuniziert werden, damit andere Personen in der Lage sind, die Ergebnisse unabhängig zu überprüfen und zu reproduzieren (soweit gemäss datenschutzrechtlichen und anderen Vorgaben zulässig). Weiter umfasst dies die Verpflichtung zur Bereitstellung von detaillierten Informationen zur Reproduktion der Ergebnisse, um Vergleiche zwischen verschiedenen Methoden und Ansätzen zu ermöglichen. Somit zielt die Reproduzierbarkeit in Kombination mit allen anderen hier aufgelisteten Grundprinzipien darauf ab, sicherzustellen, dass datenwissenschaftliche Problemlösungsansätze auf einer soliden Grundlage stehen und dass die Ergebnisse vertrauenswürdig und aussagekräftig sind.
Neutralität	Allgemein bedeutet Neutralität, dass eine Person oder Organisation unparteiisch und unvoreingenommen handelt und keine Vorurteile hat oder persönlichen Interessen verfolgt, die ihre Entscheidungen beeinflussen können. Dies gilt auch für eine Person oder Organisation bei der Anwendung von Datenwissenschaft. Ein Beispiel zur Förderung von Neutralität ist die Offenlegung möglicher Interessenkonflikte.

Grundprinzip	Definition und Erläuterung
Objektivität	Die Objektivität zielt darauf ab, sicherzustellen, dass datenwissenschaftliche Problemlösungsansätze unvoreingenommen und neutral sind und dass Entscheidungen und Ergebnisse auf objektiven und unabhängigen Kriterien basieren. Eine Person oder Organisation, welche datenwissenschaftliche Methoden, Techniken und Praktiken bei der Problemlösungsfindung entwickelt, anwendet und verbreitet, macht dies durch die Wahrung der wissenschaftlichen Unabhängigkeit und eine objektive, professionelle und transparente Arbeitsweise.
Ethischer Umgang mit Daten und Ergebnissen	Der ethische Umgang mit Daten und Ergebnissen strebt an, dass Daten und Ergebnisse in einer ethischen, verantwortungsvollen und nachhaltigen Weise verwendet werden, um sicherzustellen, dass die Rechte und die Würde von Personen respektiert und die Auswirkungen auf die Gesellschaft und die Umwelt berücksichtigt werden. Beispielsweise beinhaltet dies die Verpflichtung, Datenschutz und -sicherheit zu gewährleisten, Transparenz und Offenheit zu fördern oder sicherzustellen, dass die Verwendung von Daten und Ergebnissen ethischen Grundsätzen und Standards entspricht, welche die menschliche Würde, die Autonomie und die Privatsphäre respektieren. Darunter fallen die Vermeidung von Diskriminierung und Vorurteilen sowie die Verpflichtung, die Auswirkungen von Daten und Ergebnissen etwelcher datenwissenschaftlicher Problemlösungsansätze auf Anspruchsgruppen («Stakeholder») und die Gesellschaft und die Umwelt im Allgemeinen zu berücksichtigen.