

Von Big Data zu Smart Data

Themen, Chancen und Risiken

Herausgeber

DFKI Projektbüro Berlin
c/o Deutsches Forschungszentrum für Künstliche Intelligenz GmbH
Alt-Moabit 91c
10559 Berlin

Stand

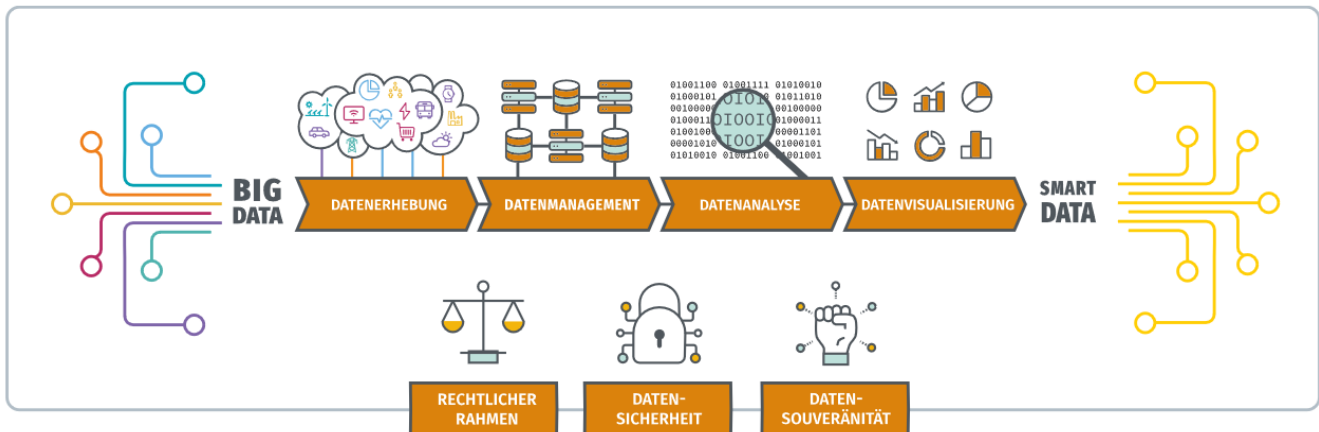
August 2019

Inhalt

1. **Big Data & Smart Data – Was ist das überhaupt?**
2. **Smart Data Technologien**
 - a. **Datenerhebung**
 - b. **Datenmanagement**
 - c. **Datenanalyse**
 - d. **Visualisierung und Benutzerinteraktion**
3. **Smart Data Governance**
 - a. **Datensicherheit**
 - b. **Rechtlicher Rahmen**
 - c. **Datensouveränität**
4. **Smart Data Anwendungen**
 - a. **Gesundheit**
 - b. **Mobilität**
 - c. **Energie**
 - d. **Industrie**

1. Big Data & Smart Data – Was ist das überhaupt?

BIG DATA & SMART DATA – WAS IST DAS ÜBERHAUPT?



Big Data ist das Ergebnis der fortschreitenden Digitalisierung vormals analoger Datenquellen sowie der Entstehung neuer digitaler Datenquellen: riesige Datenmengen, die sich mit herkömmlichen Methoden nicht mehr verarbeiten lassen.

Unter **Smart Data** verstehen wir die wirtschaftliche und gesellschaftliche Wertschöpfung aus den rasant wachsenden digitalen Datenbergen in allen Lebensbereichen. Im Begriff wird bereits die entscheidende Rolle künstlicher Intelligenz bei der Bewältigung der stetig anfallenden Daten und bereits bestehenden Datensilos angedeutet. Neben der Betrachtung der notwendigen Technologien für Datenhaltung und Datenanalyse müssen aber auch die rechtlichen und gesellschaftlichen Rahmenbedingungen berücksichtigt werden.

In dieser Broschüre wollen wir die Kernelemente des Themenfeldes Smart Data vorstellen: die Technologien, gesellschaftlich-politische Steuerungsaufgaben (Governance) und Anwendungen. Dabei konzentrieren wir uns auf knappe Darstellungen der jeweiligen Herausforderungen und Chancen und stellen zusätzlich kostenlose Tools und Repositorien vor.

2. Smart Data Technologien

Zwischen einem ungeordneten Datenpool und einer nutzerfreundlichen Anwendung oder Entscheidungshilfe liegen mehrere Verarbeitungsphasen, welche jeweils geeignete technische Werkzeuge erfordern. Daten müssen zunächst gefunden und dann gefiltert und aufbereitet werden. Nur wenn Daten von hoher Qualität vorbereitet oder die richtigen Datenquellen verknüpft wurden, lassen sich mithilfe verschiedener Analysemethoden Erkenntnisse gewinnen. Wurden Muster und Zusammenhänge in den Datenmassen erkannt, müssen diese noch für verschiedene Nutzergruppen in ganz konkreten Anwendungsszenarien visualisiert und eine weitere Bearbeitung ermöglicht werden.

Datenerhebung

Riesige Datenmengen werden im Sekundentakt von Milliarden von Geräten erzeugt: von Smartphones, Wearables, Social Media, über das Internet der Dinge und cyber-physikalische Systeme in der Industrie 4.0. Dazu kommen vormals analoge Daten, wie beispielsweise Krankenakten oder wissenschaftliche Dokumente, die immer mehr digitalisiert werden. Daten aus so unterschiedlichen Quellen können in völlig verschiedenen Formaten vorliegen und sollen dennoch im Zusammenhang ausgewertet werden. Wer gezielt Erkenntnisse aus Daten gewinnen will, muss zunächst einmal klären, welche Art von Daten benötigt werden, welche zur Verfügung stehen und ob diese mit Problemen behaftet sind. Dafür kommen sowohl Daten aus dem eigenen Projekt oder Unternehmen als auch externe Daten in Frage. Um eine sinnvolle Datenerhebung und -nutzung zu gewährleisten, müssen auch umfangreiche Datenspeichermöglichkeiten eingesetzt werden.

Verschiedene Datenquellen haben verschiedene Formate. Aufgrund der Vielfalt der Datenquellen werden Daten sowohl in einem strukturierten, aber überwiegend in einem unstrukturierten Format erzeugt. Das macht es schwierig, Daten aus mehreren Quellen für den Analyseprozess zusammenzuführen. Aufgrund stark gefallener Speicherkosten ist es nun allerdings in vielen Anwendungsfeldern möglich, zunächst alle Daten in unstrukturierter Form zu sammeln und später einer Analyse zu unterziehen. Eine der Herausforderungen im Zusammenhang mit diesen „Data Lakes“ besteht darin, dass die Daten ohne Indexierung gespeichert werden. Dies macht es schwieriger, die Daten abzurufen, da sie keine definierten Mechanismen zur Katalogisierung haben. Für große Datenmengen ist außerdem die Skalierbarkeit der Datenspeicher ein wichtiges Thema. Eine weitere Herausforderung besteht darin, Daten so zu speichern, dass eine reibungslose Portabilität von Informationen zwischen Rechenzentren oder Cloud-Service-Providern gewährleistet ist. Die fehlende einheitliche Datenstruktur für alle Quellen macht es für Datenanalysten schwierig, die Daten zu aggregieren.

Ausgewählte Open-Data-Repositories

Neben den großen Datensätzen, welche von Forschenden, Behörden und Unternehmen selbst erzeugt werden, können diese auch auf die gewaltigen Mengen an offenen Daten zurückgreifen, die allen zur freien Verfügung stehen. Hier wollen wir ein paar Beispiele für Portale und Repositorien auflisten, die Sie bei Ihrer Datensuche unterstützen können.

global	Europa	Deutschland	Gesundheit	Mobilität	Energie
Re3Data ↗	European Data Portal ↗	GovData ↗	Gesundheitscloud ↗	mCloud ↗	SMARD ↗
World Bank Open Data ↗	EU Open Data Portal ↗	BMBF-Datenportal ↗	RKI-Gesundheitsmonitoring ↗	MDM ↗	Open Power System Data ↗
World Economic Outlook Databases ↗	Zenodo ↗	Offenesdatenportal.de ↗	EudraCT ↗	DB Open-Data-Portal ↗	OpenEI ↗
GDelt ↗		RADAR ↗	OpenTrials ↗	Transforming Transport ↗	
DBpedia ↗			GHDx ↗	opentraffic ↗	
World Data Systems ↗			Federal Health Monitoring System ↗	Uber Movement ↗	
Data One ↗			DRYAD ↗		
			Global Health Observatory ↗		
			cBioPortal ↗		

Datenmanagement

Damit Daten analysiert werden können, bedarf es oftmals einiger Vorbereitung. Verteilte Daten müssen zusammengeführt werden, bevor sie gemeinsam ausgewertet werden können und ihrem Verwendungszweck entsprechende Qualitäten besitzen. Der Aufgabe, verschiedene Daten zu organisieren sowie die nötige Qualität herzustellen und zu erhalten, widmet sich das Datenmanagement. Die Wahl der richtigen Infrastruktur, um Erkenntnisse aus Big Data zu gewinnen, ist der wichtigste Faktor für das Datenmanagement. Die Vielzahl verschiedener Daten (strukturiert, unstrukturiert und halbstrukturiert) erschwert die Identifikation und den Einsatz des am besten geeigneten Tools. Die Variety, Veracity, Veracity und Volume (4 V's) von Big Data erzeugt hohe Anforderungen an die Datenanalyse-Software.

Die derzeitigen Infrastrukturen weisen Mängel bei der Analyse großer Datenmengen auf. Die Infrastruktur ist entweder aufgrund mangelnder Skalierbarkeit nicht vollständig in der Lage, große Data-Mining-Workloads zu bewältigen oder sie kann die schwankenden Datenströme nicht effizient bewältigen. Das Verfügbarmachen zusätzlicher Netzwerk- und Verarbeitungskapazitäten kann dabei helfen das Problem zu lösen, sobald aber neue Funktionen hinzugefügt werden, steigen die Anforderungen an die Systemleistung und werden selbst zu einer neuen Herausforderung. Dabei muss die Infrastruktur variable Arbeitsabläufe bewältigen und hohe Durchsätze bieten.

Open-Source-Instrumente für Verteilte Systeme

Datenverarbeitung durch Systeme wie Apache Flink, Spark, Hadoop u.a. hat die Visualisierung und Detektion von Abweichungen oder Fehlern von großen Datenmengen zum Ziel. Es bestehen zwei Möglichkeiten, um große Datenmengen zu analysieren – in Batch und Stream.

Batch-Verarbeitung bedeutet, dass große Datenmengen in Stapeln oder „Batches“ analysiert werden. Die Informationen werden über einen längeren Zeitraum gesammelt und später zusammen verarbeitet. Das heißt, die Batch-Datenverarbeitung dauert länger (in der Regel Sekunden) als die Datenstromverarbeitung, die normalerweise in Echtzeit erfolgt. Sie kann auch offline verwendet werden. Die Batch-Verarbeitung ist z.B. für die Erstellung von Lohn- und Gehaltsabrechnungen oder Kontoauszügen sinnvoll.

Die Systeme für die **Datenstromverarbeitung** helfen dem Anwender, in Echtzeit nützliche Erkenntnisse aus Daten zu gewinnen. Damit dies funktioniert, müssen die Daten ab dem Zeitpunkt der Generierung direkt in das Analysesystem einfließen. Die Datenstromverarbeitung kann z.B. bei der Betrugserkennung, der Protokollüberwachung oder der Analyse des Kundenverhaltens effizient eingesetzt werden. Sie funktioniert nur online.

Batch-Verarbeitung	Datenstromverarbeitung
Apache Spark ↗	Apache Flink ↗
Apache Flink ↗	Apache Spark Streaming ↗
Apache SystemML ↗	Apache Kafka ↗
Apache Hadoop ↗	Apache Storm ↗
	Apache Beam ↗

Datenanalyse

Die Datenanalyse konzentriert sich auf die Gewinnung wertvoller Erkenntnisse aus heterogenen Datenquellen. Eine nachvollziehbare und umfassende Visualisierung der Ergebnisse gehört ebenfalls zum Analyseprozess dazu, sodass die Informationen aussagekräftig und gebrauchsfertig dargestellt werden können. Dies hilft Unternehmen, intelligentere Entscheidungen zu treffen, Kosten zu senken und neue Produkte und Dienstleistungen von höherer Qualität zu entwickeln.

Die Gewährleistung von Fehlertoleranz bei der Ausführung von komplexen Analysen ist eine schwierige und kostspielige Aufgabe. Ihr Ziel ist es, die erfolgreiche Ausführung von komplexen Analyse-Tasks auf potenziell ausfallender Hardware innerhalb eines vorgegebenen Zeitrahmens sicherzustellen. Eine weitere Herausforderung stellt die Skalierbarkeit dar. Die Hauptschwierigkeit liegt darin, wie verschiedene Aufgaben auszuführen sind. Unstrukturierte Daten verlängern die Liste der Probleme, da die Umformung von allen unstrukturierten zu strukturierten Daten unmöglich ist. Daher kann Datenhomogenität zu falschen statistischen Schlüssen führen, wenn keine besseren adaptiven Verfahren entwickelt werden.

Machine Learning

Wenn wir von Datenanalyse sprechen, müssen wir auch Machine Learning (ML) erwähnen. Dies ist ein Prozess, bei dem Computer aus großen Datenmengen lernen, indem sie Muster erkennen und Informationen extrahieren. Das kann Maschinen helfen, Entscheidungen ohne menschlichen Eingriff (Programmierung) zu treffen. Die Algorithmen müssen mit großen Datenrepositorien trainiert werden. Deep Learning gehört dabei zu den wichtigsten Methoden, um Computer lernen zu lassen und kann erfolgreich für die Informationsbeschaffung, Objekterkennung, Stimmungsanalyse, personalisierte Medizin und mehr eingesetzt werden. Für das Trainieren von Deep-Learning-Modellen existiert eine größere Zahl von Bibliotheken und Frameworks.

Visualisierung und Benutzerinteraktion

Neue Smart-Data-Lösungen erzeugen einen hohen Bedarf an neuen Schnittstellen zwischen Menschen und Computern – Sprachdialogsysteme, Brain-Computer-Interfaces, Co-adaptive Lernsysteme oder verschiedene Arten von Augmented-Reality-Schnittstellen (AR). Mit Geräten wie AR-Brillen oder AR-Helmen für Bauarbeiter zielt Augmented Reality unter anderem darauf ab, die Wahrnehmung der Realität zu erweitern und Umgebungsinformationen interaktiv und digital manipulierbar zu machen.

Die Kommunikation zwischen Mensch und Maschine ist eine ständige Herausforderung für die Industrie. Ohne die Fähigkeit, verbale und vor allem nonverbale Interaktion zu verstehen, gelingt es Maschinen nicht, die Absichten des Menschen ausreichend zu interpretieren, um angemessen zu reagieren zu können. Dasselbe gilt für Menschen, die nicht sehen können, was in der Blackbox eines Datenanalyse-Systems passiert. Um die Mensch-Maschine-Interaktion zu unterstützen, sind höhere Belastbarkeit und Flexibilität erforderlich. Denn da sich das Nutzerverhalten im Laufe der Zeit ändert, müssen sich die Maschinen ohne großen Programmieraufwand schnell an die Bedürfnisse der Anwender anpassen können. Darüber hinaus müssen Frameworks zur Echtzeit-Sensordatenverarbeitung so verbessert werden, dass sie mehr Sicherheit, Zuverlässigkeit und Benutzerfreundlichkeit bieten.

MXNet – Open-Source-Framework, das auf Deep Learning basiert und für das Training und die Ausführung von tiefen neuronalen Netzwerken entwickelt wurde. Durch seine Skalierbarkeit auf mehrere GPUs und Computer kann das System Algorithmen schnell trainieren. Es unterstützt Programmiersprachen wie Python, JavaScript, Go, C++, Scala, Matlab, R und viele mehr.

TensorFlow – Open-Source-Software-Bibliothek, die numerische Berechnungen durchführt. Sie funktioniert auf vielen Plattformen.

3. Smart Data Governance

Der Bedarf an qualitativ hochwertigen Daten erfordert zunehmend den freiwilligen Austausch von Daten durch Privatpersonen, Behörden und Unternehmen. Voraussetzung dafür ist aber, dass sowohl die Datentransaktionsprozesse als auch die Verwendung der Daten sicher und Vertrauenswürdig funktionieren. Um diese Rahmenbedingungen zu schaffen, sind verbindliche Standards und politische Maßnahmen nötig.

Datensicherheit

Digitalisierung schafft viele Chancen für Unternehmen und Verbraucher, aber gleichzeitig auch Herausforderungen. IT-Sicherheitsthemen wie Cyber-Angriffe, Identitäts- und Datendiebstahl stellen bereiten vielen Nutzern Sorgen.

Unternehmen aller Branchen verfügen über wertvolle Daten und Kenntnisse. Gerade wenn sie Sicherheitsmaßnahmen vernachlässigen, werden sie daher zum Ziel von Hackern, die sich diese Informationen beschaffen oder das System eines Unternehmens gefährden wollen. Laufende aktualisierte Vorgaben zu Sicherheitsstandards können Unternehmen und Privatpersonen dabei unterstützen, ein robustes Schutzsystem einzurichten.

Aufgrund der hohen Kosten für den Einsatz von Sicherheitstools haben viele Systeme nur eine einzige Schutzschicht. Das macht sie anfälliger für Angriffe. Veraltete Geräte und Anlagen, die ursprünglich nur für die Verwendung im Intranet vorgesehen waren oder für die keine Updates mehr bereitgestellt werden sowie die willkürliche Verschaltung von Sensorgeräten schaffen Sicherheitslücken in IoT-Netzwerken und führen zu einer wachsenden Anzahl von Cyberattacken. Eine weitere Herausforderung ist die Sicherheit von Cloud-basierten Speichersystemen. Umso mehr Daten geteilt und an Dritte weitergeleitet werden, desto größer wird die Notwendigkeit, an allen Punkten deines Netzwerks umfassende Sicherheitsmaßnahmen zu implementieren. Bei fehlenden Sicherheitsebenen kann sonst der Datenschutz vieler Nutzer gleichzeitig gefährden werden. Grundsätzlich kann fehlende Verschlüsselung bei der Datenübertragung Hackern Datenmanipulationen ermöglichen.

Rechtlicher Rahmen

Die ständige Weiterentwicklung digitaler Technologien wirft auch neue rechtliche und ethische Fragen auf. Um einen einheitlichen Verhaltenskodex für den Umgang mit neuen Technologien festzulegen, bedarf es eines kollektiven Dialogs, der Mitglieder der Zivilgesellschaft, der Wissenschaft, Ethikkommissionen, Juristen und Politiker zusammenbringt. Autonomes Fahren, humanoide Roboter, Drohnen oder Gesichtserkennungssysteme sind nur einige Beispiele für Bereiche, die in den rechtlichen Rahmen jedes Landes eingefügt werden müssen. Obwohl die EU-Datenschutzgrundverordnung (DSGVO) die Datenerhebung und -verarbeitung EU-weit gründlich regelt, müssen Regierungen der tatsächlichen Durchsetzung weiterhin besondere Aufmerksamkeit widmen.

Digitalisierung und die Verbreitung von Big-Data-Anwendungen schaffen vor allem Herausforderungen in den Bereichen Datenschutz und Urheberrecht. Die europäische Datenschutzgrundverordnung (DSGVO) stellt zwar sicher, dass die Vertraulichkeit, Integrität und Transparenz der Daten während der Verarbeitung gewahrt bleibt, die Einhaltung der neuen Vorschriften ist aber zum Teil eine schwierige Aufgabe für die Datenverarbeiter. Auch die Durchsetzung von Ansprüchen aus dem europäischen Recht im Ausland, wo ein Großteil der Daten gespeichert und verarbeitet wird, ist problembehaftet. Eine weitere Schwierigkeit besteht darin, dass Analysen von Massendaten zu Ergebnissen führen können, die in gewisser Weise diskriminierend sein und daher negative soziale oder wirtschaftliche Auswirkungen haben können. Der EU-Rechtsrahmen sieht auch keine Regelung für das Dateneigentum vor – so herrscht Unsicherheit darüber, wer die Daten in einem ganzen Daten-Wertschöpfungskreislauf besitzt. Außerdem müssen die geistigen Eigentumsrechte neuer Technologien und Anwendungen gewahrt, gleichzeitig aber auch der Zugang für die Allgemeinheit sichergestellt werden.

Datensouveränität

Datensouveränität ist als Teilbereich von Digitaler Souveränität ein unscharfer Diskursbegriff und wird verwendet, um verschiedene Aspekte der individuellen oder gesellschaftlichen Befähigung zum selbstbestimmten Umgang mit Big Data zum Ausdruck zu bringen. Grundannahme ist dabei, dass aktuelle rechtliche Regelungen, politische Agenden und individuelle Kompetenzen nicht ausreichend an die neuen Bedingungen der Massendatenverarbeitung angepasst sind. Der Datensouveränitätsbegriff wird von verschiedenen Akteuren unterschiedlich geprägt. Da er anders als „informationelle Selbstbestimmung“ stärker die Ermöglichung von Datenerhebung und Datenverarbeitung betont, wird er von manchen Datenschützern auch als „Lobbybegriff“ bezeichnet. Die in dem Zusammenhang mit Datensouveränität diskutierten Probleme gehen jedoch deutlich über die Verfügbarmachung von Daten hinaus. Der Diskurs ist uneinheitlich und verzweigt, außerdem ergeben sich zahlreiche Überschneidungen mit anderen Themen wie rechtlicher Rahmen und Datensicherheit.

Rechtliche Fragen bilden den kontroversesten Aspekt der Debatten um Datensouveränität. Hier beschreibt Datensouveränität häufig den Ansatz, Daten aufgrund der wertvollen Erkenntnisse, die sich aus ihnen gewinnen lassen, weniger als individuelles Eigentum, sondern als kollektives Gut zu betrachten. Daraus würde auch folgen, die Widersprüche zwischen den Grundsätzen des aktuellen Datenschutzes (Datensparsamkeit, Einwilligungserfordernis und Zweckbindung) und den Anforderungen von Massendatenanalysen (Datenreichtum, automatischer Mustererkennung ohne vorentworfenes Ziel und Rekontextualisierung) zugunsten letzterer aufzulösen. Nutzer von Digitaltechnologien sollten aber auch unter Big-Data-freundlichen Bedingungen in der Lage sein, informiert und selbstbestimmt zu entscheiden, was mit ihren Daten geschieht. Dazu müssten nicht nur Transparenz in den Prozessen geschaffen, sondern auch Einschränkungs- und Wahlmöglichkeiten im Verlauf des Datenanalyseprozesses ermöglicht werden. Datenschützer befürchten jedoch den Verlust der im Datenschutzrecht verankerten informationellen Selbstbestimmung.

Im politischen Diskurs wird der Begriff der Datensouveränität verwendet, um den angestrebten Zustand der Unabhängigkeit von ausländischen Datenverarbeitungsunternehmen zu bezeichnen. Eng damit verbunden ist das Ziel, in Deutschland bei der Entwicklung von Künstlicher Intelligenz und der Stärkung der Datenwirtschaft eine internationale Führungsposition zu erlangen. Dazu strebt die Bundesregierung die Förderung von Forschung und Innovation, beispielsweise im Rahmen einer nationalen KI-Strategie, aber auch um den Ausbau der digitalen Infrastruktur als deren Grundbedingung an. Deutsche Datenverarbeitungsunternehmen und Forschungseinrichtungen hoffen auf den Abbau regulatorischer Hürden für die Sammlung und Auswertung von Daten und benötigen technisch wie auch rechtlich sichere Repositorien sowie Austausch- und Verarbeitungsplattformen für qualitativ hochwertige Datensets.

Zu überblicken, welche Daten während der Interaktion mit Digitaltechnologien übermittelt werden und wie die Prozesse zur Verwertung dieser Daten eigentlich funktionieren, ist nur wenigen Spezialisten möglich. Es gibt nicht genug spezialisierter Fachkräfte, um den Ambitionen von einer deutschen Datenwirtschaft mit Spitzenniveau gerecht zu werden. Zusätzlich müssen auch auf der Anwendungsebene die Nutzer Datenverarbeitungsmechanismen verstehen, um informiert entscheiden zu können, welcher Datennutzung sie zustimmen. Digitale Technologien überfordern Teile der Bevölkerung oder bleiben ihnen aufgrund zu hoher Komplexität und eingeschränkter Bedienbarkeit gänzlich unzugänglich. Um Datensouveränität im Sinne von Kenntnissen und praktischer Befähigung – als Digitalkompetenz – zu erreichen, muss der Vorsprung bei der Big-Data-Entwicklung aufgeholt werden: nicht nur in den Bereichen Bildung, Inklusion und Usability sondern auch im Sinne einer gesellschaftlichen Reflexion.

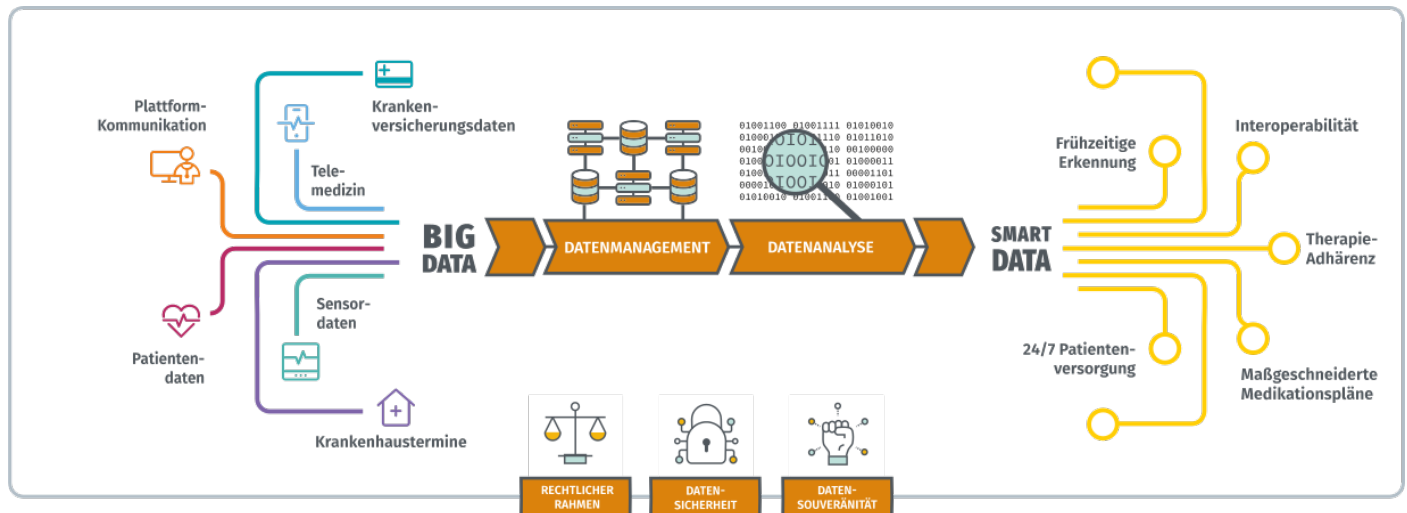
Herausforderungen und Chancen

- Anpassung des rechtlichen Rahmens
- Bereitschaft zum Datenteilen
- Lockerung der Monopolstellung großer Plattformen
- Transparenz der Prozesse
- Digitalkompetenzen der Bevölkerung
- Fachkräfteausbildung
- Privacy-by-Design
- Good Big Data Practice
- Nationale Daten- und Wissensinfrastruktur
- Nutzerzentrierung der Angebote
- Open-Data-Verpflichtung

4. Smart Data Anwendungen

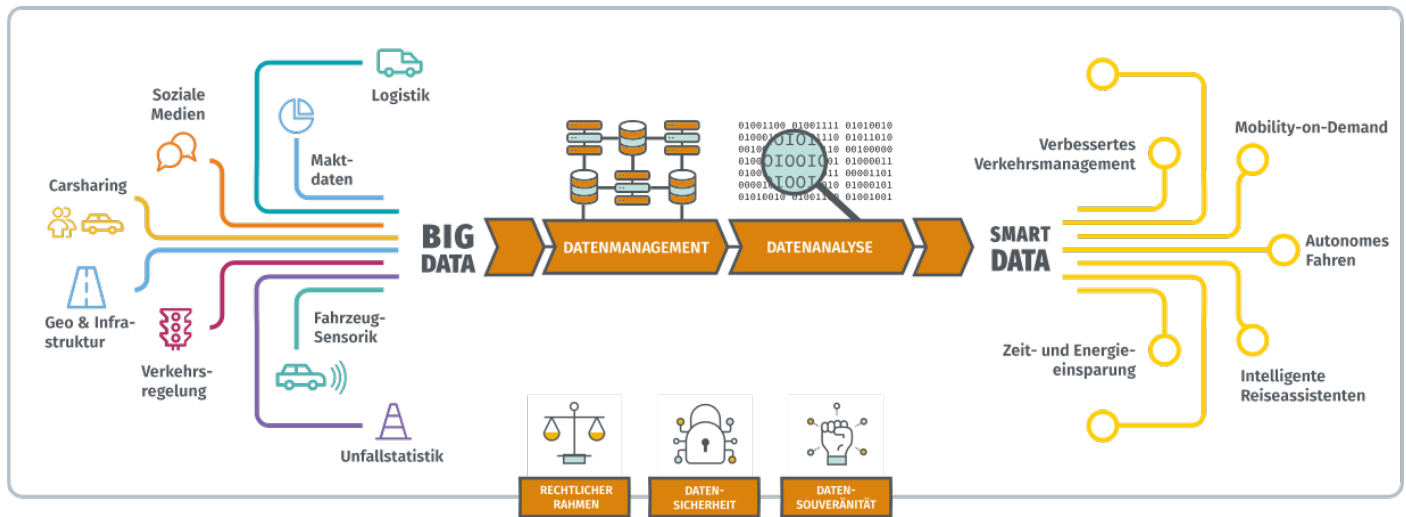
Die intelligente Auswertung von großen Datenmengen birgt das Potential, grundsätzlich in allen Bereichen des Lebens nutzbringend angewandt zu werden. Am umfangreichsten und am weitesten fortgeschritten sind Smart-Data-Anwendungen aber in den Bereichen Gesundheit, Mobilität, Energie und Industrie. Die spezifischen Chancen und Herausforderungen innerhalb dieser Sektoren wollen wir im Folgenden überblicksartig skizzieren.

Gesundheit



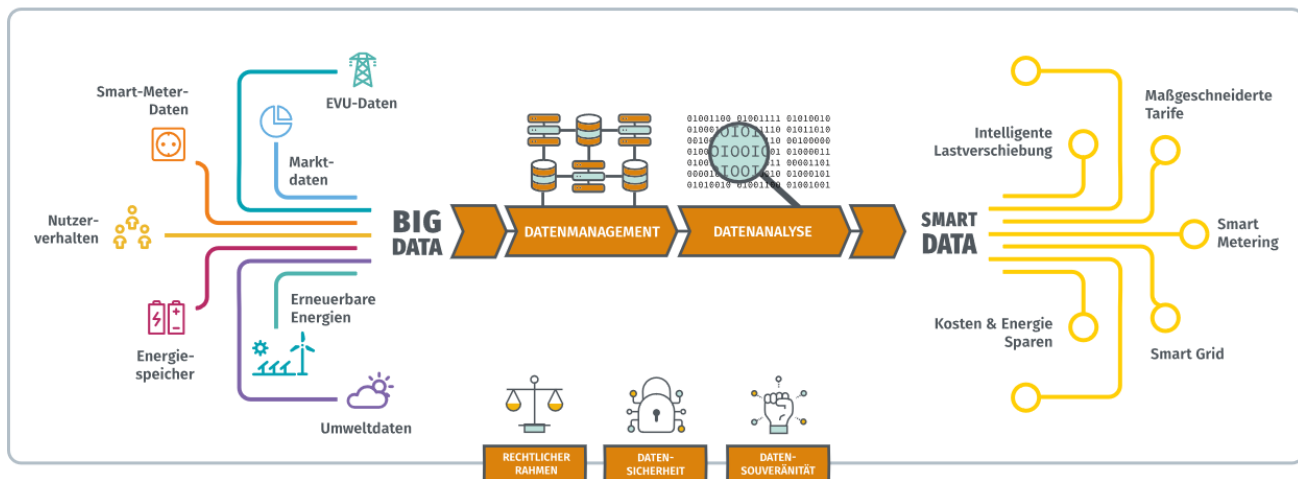
Biosensoren, High-Throughput-Technologien in den Biowissenschaften, Lifelogging, mobile Apps, Online-Patiententagebücher – der Umfang digitaler, patientenbezogener Daten schreitet dank neuer Möglichkeiten massiv voran. Daneben stellen internationale Forschung und partizipative Medizin in Open-Data-Initiativen immer größere Mengen an Rohdaten online zur Verfügung. Diese Datenbestände sollen zur Verbesserung von Therapien und zur Versorgung von Patienten, beispielsweise in der Telemedizin eingesetzt werden, ohne dabei das Vertrauensverhältnis zwischen Arzt und Patienten zu gefährden.

Um die in der Digitalisierung liegenden Chancen der Gesundheitsversorgung voll auszuschöpfen, müssen Patientendaten standardisiert, strukturiert und mit relevanten Daten aus persönlichen Geräten oder mit demographischen Informationen verknüpft werden. Dies ermöglicht die Erstellung präventiver Strategien und prädiktiver Modellen. So kann beispielsweise KI-gestützte Diagnostik Fachärzten dabei helfen, große Mengen an verfügbaren Daten zu vergleichen, mögliche Gesundheitsprobleme zu identifizieren und geeignete und personalisierte Therapieverfahren vorzuschlagen. Um zu einem einheitlichen digitalen Gesundheitssystem zu gelangen, müssen jedoch zunächst Probleme des Datenschutzes und der Interoperabilität gelöst werden. Auf dieser Basis kann eine Infrastruktur für Netzwerke von Kliniken und Fachärzten geschaffen werden, die eine IT-gestützte individuelle Medikamentenplanung und patientenzentrierte Therapien ermöglicht. Weiterhin ist es wichtig, die digitale Gesundheitskompetenz der Bürger zu verbessern, um präventive Behandlungsmöglichkeiten zu finden und die medizinische Versorgung auch aus der Ferne effizient zu nutzen. Die meisten dieser Herausforderungen spiegeln sich im Konzept der P4-Medizin wider: präventiv, prädiktiv, partizipativ und personalisiert.



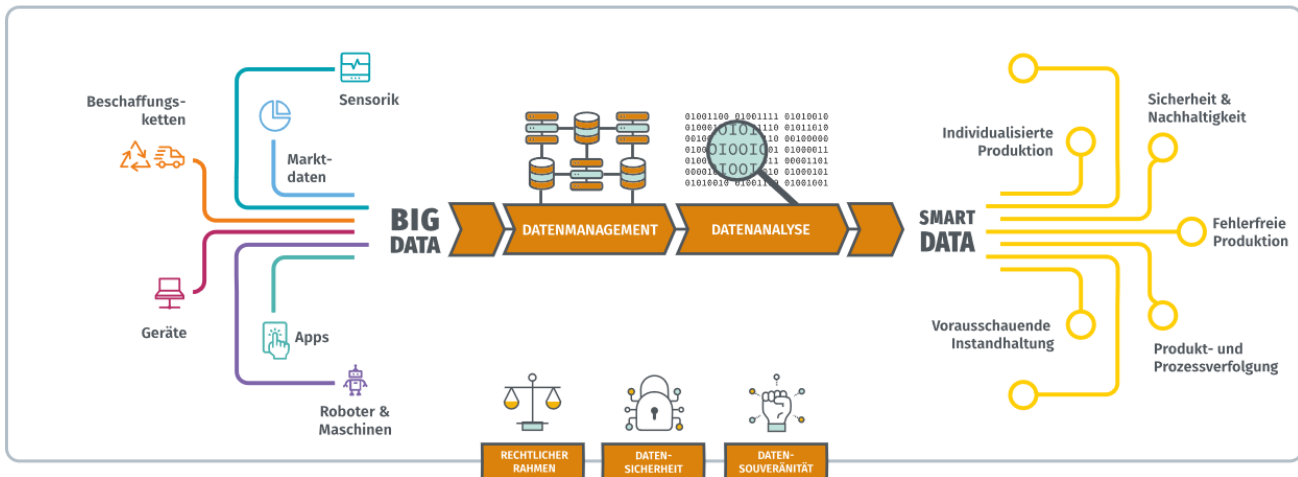
Floating Car Data, Car2X, Fahrbahnsensoren und Signalanlagen, Mobilfunkdaten, Strecken und Fahrpläne, Infrastrukturdaten, Twitter – in kaum einem anderen Bereich stehen so viele Datenquellen für Optimierungen zur Verfügung wie im Verkehrssektor. Aus diesem Datenschatz lassen sich neue Datenwertschöpfungsketten und Lösungen für individuelle Assistenzsysteme oder Methoden in der integrierten Verkehrsplanung und Lenkung schöpfen. Große Chancen liegen dabei im Autonomen Fahren, Stauvermeidung, Mobility-on-Demand, Verkehrssicherheit und in intelligenten Reiseassistenten. Smart Mobility kann das Reisen schneller, billiger, angenehmer und sicherer machen.

Der Mobilitätssektor entwickelt sich rasant zu einem voll digitalisierten System. Trotz wichtiger Fortschritte in diesem Bereich sind noch viele Herausforderungen zu bewältigen. Das autonome Fahren erfordert erhebliche Infrastrukturverbesserungen wie Sensoren, Netzabdeckung, Breitbandzugänge und niedrigere Latenzzeit. Die heutigen Datenübertragungstechnologien stoßen an ihre Grenzen. LTE kann Daten nicht vollständig in Echtzeit übertragen, was jedoch eine Voraussetzung für viele Smart-Mobility-Anwendungen darstellt. Es wird erwartet, dass der kommende 5G-Standard dieses Problem lösen wird. Um den datenschutzrechtlichen Anforderungen und ethischen Fragen einer Welt neuer Geschäftsmodelle gerecht zu werden, muss zudem der rechtliche Rahmen angepasst werden.



Smart Meter, Smart Grids, virtuelle Kraftwerke – die Energieversorgung von morgen reicht von innovativen Smart-Home-Technologien hin zu dezentralen Energienetzwerken. So bieten neue Informations- und Kommunikationstechnologien Potenzial zur effizienteren Auslastung vorhandener Netzkapazitäten, sowie relevante Daten für einen intelligenten Netzausbau und die Integration erneuerbarer Energien. Durch intelligente Lastverschiebungen und Analysen des Verbrauchsverhaltens in Unternehmen und Heimenergieanlagen können Smart-Energy-Technologien auch im privaten Bereich helfen, Energie und Kosten einzusparen.

Die Entwicklung alternativer Energiequellen und die Dezentralisierung der Energieversorgung sind die wichtigsten Herausforderungen der Energiewende in Deutschland. Spannungsschwankungen, die bei der Verwendung erneuerbarer Energiequellen auftreten, müssen ständig ausgeglichen werden können. Dazu sind Energiespeichermöglichkeiten so weiterzuentwickeln, dass auch in Zeiten ohne Wind- und Sonnenenergieerzeugung eine konstante Stromversorgung gewährleistet ist. Darüber hinaus müssen sensorbasierte Echtzeitmessung und adaptive Energiemanagement-Software entwickelt, getestet und implementiert werden. Bestehende Smart-Energy-Technologien müssen zur Einsparung von Energiekosten für gewerbliche Großverbraucher und Privathaushalte angepasst werden.



Das Internet der Dinge erobert zunehmend die Produktionshallen und Versorgungsnetzwerke. Gleichzeitig lässt sich aufgrund der Komplexität der hochautomatisierten Produktion die Qualität, die Effektivität und die Produktivität der Prozesse kaum noch „händisch“ überwachen. Datenstromanalysen in Echtzeit können nicht nur rechtzeitig Verschleißerscheinungen an Maschinen erkennen, sie liefern auch wertvolle Hinweise für zukünftige Produkt- und Prozessoptimierungen, verbesserte Kapazitätsauslastung und Services. Zu den Visionen der Industrie 4.0 gehören die Individualisierte Produktion (Losgröße 1), lückenlose Produkt- und Prozessverfolgung und vorausschauende Instandhaltung.

Angesichts der Vielzahl von Geräten, die Daten erzeugen, analysieren und an andere Maschinen übertragen, ist die Datensicherheit für die angeschlossenen Systeme ein zunehmendes Problem. Sicherheitssysteme müssen weiterentwickelt werden, um unbefugte Eingriffe von außen zu verhindern. Darüber hinaus ist mehr Standardisierung für den cyberphysikalischen Informationsaustausch erforderlich, um volle Systemzuverlässigkeit und Kompatibilität zu gewährleisten. Volle Funktionalität der Hardware und Fehlertoleranz sind weitere Herausforderungen, die den Produktionsprozess beeinflussen. Außerdem wächst die Befürchtung, dass der Ersatz von Menschen durch Maschinen zu massiven Jobverlusten führen könnte.