

STUDIUM DIGITALE

Kapitel 1: Einführung

Die Kosten der Kommunikation sind zerfallen.

Viele Personen behaupten, dass sich unsere Welt so schnell veränderte wie noch nie.

Automation→Substitution menschlicher Arbeit.

Vorher funktionierten Computern nur mit Arbeitern, aber heute, ein Computer kann alles allein machen, man braucht nicht mehr Menschen→ alles wurde automatisiert.

Berufe werden durch Computern und Technologie ersetzt.

Individuen, Technologie und die Gesellschaft sind dank der Technologie besser vernetzt.

Herausforderung mit der Digitalisierung: private Strasse gekauft, de-skiling,etc.

Die Organisation hat sich mit Informationssysteme verändert.

Wir sind divers, kooperativ und Fehleranfällig (soggetti a errori)

Disruption

Alle haben heute überall WLAN, alle haben Tablets und Computer.

Regulation

Wie sieht es mit dem Ethik ein?

Digitalisierung→Technische und gesellschaftliche Veränderung.

Wechselspiel von gesellschaftlicher und technischer Entwicklung.

Diese haben Einfluss/Effekte auf drei Stufen.

Es gibt mehrere Chancen/ Herausforderungen

Wir generieren pro Tag 2.5 Exabytes an Daten.

Wir können heute Berechnungen machen, die wir in der Vergangenheit nie könnten, da die Kosten von Berechnungen mit dem Verlauf des Jahres verfallen sind.

Digitalisierung bringt die Automatisierung von menschlichen Tätigkeiten.

Interdisziplinarität -> mehrere Disziplinen (z.B. in einer Startup braucht man Juristen für Datenschutz, Informatiker für die Implementierung, Marketing Experte für das Marketing Teil usw.).

Daten = Zeichen/Symbole/Signale/"rohe" Fakten -> Bedeutung ist nicht in Daten! Daten können verschieden ausgedrückt oder dargestellt werden. Sie sind unabhängig von ihrer Interpretation. Sie sind auch in anderer Anordnung (ordine) gleich.

I 3 kann entweder eine Buchstabe sein (B) als auch eine Zahl(13).

Es ist nicht gleich interpretiert, z.b in Europa ist es B und in Nord-Amerika die Zahl 13.

Information=Daten+Semantik

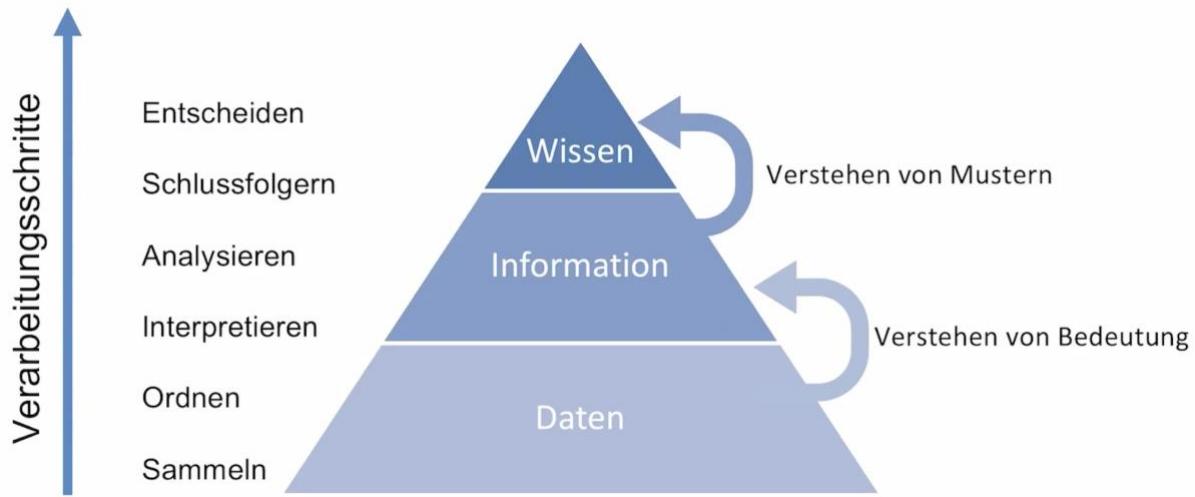
Daten sind also Fakten.

Noten Beispiel: Eva hat ein 1 in eine Prüfung bekommen, in der Schweiz bedeutet es dass sie die

Prüfung schlecht gemacht hat und nicht bestanden hat, aber in Deutschland ist 1 die beste Note, so hätte sie die Prüfung bestanden→Verschiedene Aussichtspunkte (punti di vista/interpretazioni)

Information = Daten + Semantik (Interpretierung von Daten). Informationen entstehen aus unterschiedlicher Darstellung der Daten. Information wird auch von der Bedeutung der Daten bestimmt.

Wissen = Information + Verknüpfung. Wissen entsteht bei der Analyse der Information, und bedarf oft der Verknüpfung mit zusätzlichen Informationen und Wissen.



Beispiel Wetterstation:

- Daten: z.B. ZH 7.8.2020 22:00 7 Grad.
- Information: Am 7 August 2020 um 22 Uhr abends gab es 7 Grad.
- Wissen: Morgen sollte ich ein Pullover mitnehmen, da es nur 7 Grad ist.

Informationstechnologie: Werkzeug, um Daten zu sammeln, übertragen, speichern, und verarbeiten (Bsp: Buch).

Informationssystem: Kombination von Hardware, Software und Netzwerken, die beim Sammeln, Kreieren und Verteilen wichtiger Daten hilft (Bsp: Bibliothek).

Digitale Güter: Produkte oder Dienstleistungen, in Form von Binärdaten hergestellt, verarbeitet, übertragen und konsumiert. (Bsp: Youtube Videos, Spotify Playlists)

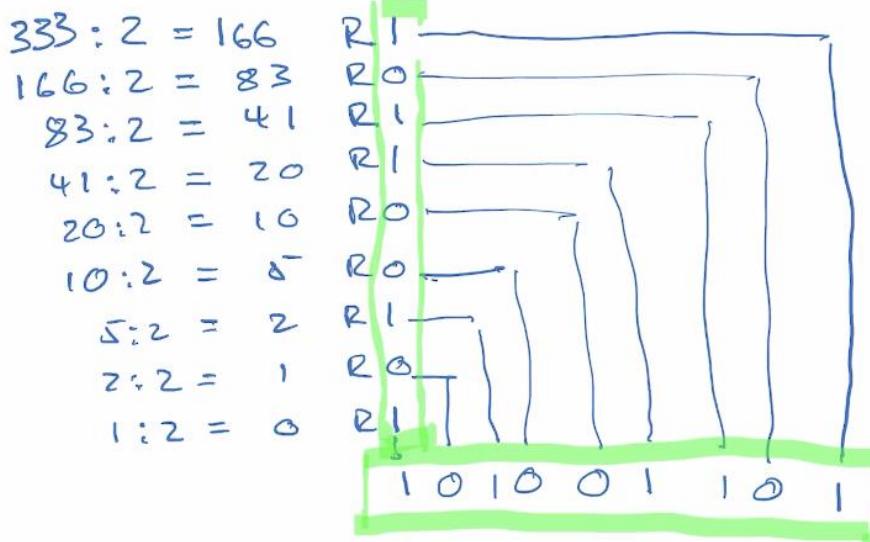
Informationssysteme und Informationstechnologie können auch nicht digitaler Natur sein. In der modernen Welt stellen wir uns bei diesen Systemen oder Technik fast immer irgendwelche Computer vor, das muss aber nicht unbedingt sein. Die Systeme können auch nicht-digital sein.

Kapitel 2: Digitalisierung

Computer können nur binäre daten verarbeiten. Computer verstehen nur die Nullen und die Einsen.

Von Dezimal zu Binär:

Umwandeln von Dezimalzahl 333 in Binärzahl:



Von Binär zu Dezimal:

Umwandlung der Binärzahl 1101101_2 in eine Dezimalzahl:

1	1	0	1	1	0	1
2^6	2^5	2^4	2^3	2^2	2^1	2^0

$1 \cdot 2^6$	$1 \cdot 2^5$	$0 \cdot 2^4$	$1 \cdot 2^3$	$1 \cdot 2^2$	$0 \cdot 2^1$	$1 \cdot 2^0$
64	32	0	8	4	0	1

$$64 + 32 + 8 + 4 + 1 = \underline{109}$$

1 Bit = ein 0 oder ein 1

1 Byte = 8 Bits

1 KB	= 1 Kilobyte	= 10^3 Bytes	= 1'000 Bytes
1 MB	= 1 Megabyte	= 10^6 Bytes	= 1'000'000 Bytes
1 GB	= 1 Gigabyte	= 10^9 Bytes	= 1'000'000'000 Bytes
1 TB	= 1 Terabyte	= 10^{12} Bytes	= 1'000'000'000'000 Bytes

Digitalisierung von Text

ASCII, Unicode, ISO 8859-15

ASCII-Code:

American Standard Code for Information Interchange

ASCII-Codetable für Zahlen und Zeichen

0-9:48-57

A-Z: 65-90

a-z: 97-122

7Bit: reicht für 128 Zeichen (2^7)

→Ok für alle Zeichen der englischen Sprache

→Ungenügend für europäische Sprachen (z.B spezialzeichen: ä, ü)

ISO 8859-15:

8 Bit: Iso 8859 definiert 256 Zeichen (2⁸)

Die ersten 128 Zeichen stimmen mit ASCII überein

Nicht geeignet für arabische, chinesische,griechische,usw

Unicode:

32 Bit: über 4 Milliarden Zeichen (2^{32}) können codiert werden

Z.Z werden über 100'000 Zeichen erfasst (Unicode 6.2)

Auf Klingonisch muss weiterhin verzichtet werden: Unicode-Konsortium verweigert die Aufnahme

hartnäckig.

Klingonisch ist nicht aus einer Sprache der Erde bezeichnet.

Es existieren verschiedene weitere Codierungssysteme – passend für die jeweiligen Bedürfnisse

Je universeller das System, desto mehr Bits werden für die Codierung benötigt → Speicherbedarf!

Damit die Zeichen korrekt interpretiert werden können, ist die Angabe des verwendeten Codierungssystems zwingend (Affinché i caratteri siano interpretati correttamente, è obbligatorio

specificare il sistema di codifica utilizzato).

Je mehr Text, desto höher der Speicherbedarf:

-1 Seite Text

80 Zeichen/Zeile

64 Zeilen/Seite

8 Bit pro Zeichen (=1 Byte, ISO 8859-15) = $80 \times 64 \times 1 \times 8 = 40'960$ Bits

→ $40'960 / 8 = 5'120$ Bytes = circa 5.1 kB (Speicherbedarf)

Esempio con Unicode:

32 Bit pro Zeichen, 50 Zeilen/Seite, 60 Zeichen/Seite, 10 Seiten → Bits: $10 \times 50 \times 60 \times 32 = 960000$

Bytes: $960000 / 8 = 120000$

Kilobytes: $120000 / 1024$

Digitalisierung von Ton

1 Hz = eine Schwingung pro Sekunde

Je höher die Frequenz, desto höher der Ton

Sampling: Periodisches Abtasten des analogen Signals. Analoges Signal wird in regelmäßigen Zeitabständen gemessen. Gemessene Werte ergeben Abbild des Ursprungssignals. Je höher die Zahl der Samples pro Sekunde (Abtastrate) desto besser entspricht das digitale Abbild dem Original.

Quantisierung: Zuordnen der ermittelten Werte an vorgegebene Werte. Erfasste Messwerte werden gerundet und einem vordefinierten Wert zugewiesen.

Je höher die Tonqualität, desto höher der Speicherbedarf

Der Mensch hört zwischen 20Hz und 22'000Hz

Lautstärke

-60 dB: Unterhaltung

-90dB: Laute Musik

-120 dB: Flugzeug

Lärm=rumore

Tondigitalisierung: Sampling und Quantisierung:

-Abtasten (Sampling)

→Periodisches Abtasten des analogen Signals

Quantisierung

→Zuordnen der ermittelten Werte an vorgegebene Werte

Sampling:

-Analoges Signal wird in regelmässigen Zeitabständen gemessen

-Gemessene Werte ergeben Abbild des Ursprungssignals

Je höher die Zahl der Samples pro Sekunde (=Abtastrate) desto besser entspricht das digitale Abbild dem Original).

Für gute Ergebnisse muss die abtastrate der erwünschten maximale Frequenz betragen.

Da der Mensch maximal 22'000hz hört, sollte die abtastrate mindestens 44'000Hz betragen (Cd-qualität 44'100Hz).

Quantisierung:

Erfasste Messwerte werden gerundet und einem vordefinierten Wert zugewiesen (I valori misurati

registrati sono arrotondati e assegnati a un valore predefinito).

-8 Bits: 256 unterscheidbare Tonwerte

-16 Bits: 65536 unterscheidbare Tonwerte

Je mehr besser die Tonqualität, desto höher der Speicherbedarf

Digitalisierung von Bild

Wie können wir Bilder in Zahlen verwandeln?

RGB: Für Bildschirme -> Additives Farbmodell

CMYK: Für Drück -> Subtraktives Farbmodell

Bilder haben nicht nur Farben, sondern auch Formen.

Vektorgrafik:

-Besteht aus mathematischen definierten Objekten

-Jedes Objekt hat Attribute. Es werden nur Attributwerte gespeichert

-Jedes Objekt kann einzeln bearbeiten werden

-Vergrössern ohne Qualitätseinbusse

z.B: Logos sind Beispiele für Vektorgrafik.

Pixelgrafik:

-Besteht aus rasterförmig angeordneten Punkten (Pixel)

-Jeder Punkt hat Farbwert gemäss Farbmodell

-Jeder Punkt kann einzeln bearbeiten werden

Pixelgrafik: Farben:

Beispiel RGB, 24 Bit Farbtiefe: Für jedes Pixel 8 Bit pro Grundfarbe → 256 Farbe pro

Kanal → 16.7 Mio.

Farben.

Datenspeicherung

Unstrukturierte Daten:

- Ohne explizite, formale Struktur
- Semantik durch Interpretation
- Automatische Nutzbarkeit eingeschränkt
- Durchsuchen: Herausfordern

Unstrukturierte Daten: Keyword-Suche:

- Eingabe eines oder mehrerer Suchbegriffe
- Volltextsuche
- Anzeige der Suchresultate nach Relevanz"

Ist nicht sehr spezifisch

Probleme der Keyword-Suche:

Kann andere Wörter geben, oder zu viele Informationen (auch Informationen die nicht relevant sind)

-Keyword-Suche findet viele irrelevante Dokumente

Precision → wenn = 100% dann sind alle gefundenen Dokumente relevant.

Wenn die Precision 100% ist, wurden nur relevante Dokumente gefunden → Richtig

Wenn der Recall 50% ist, sind die Hälfte aller gefundenen Dokumente relevant → Falsch

Eine perfekte Suchmaschine hat einen Recall von 100% und eine Precision von 100%.-->Richtig

Wenn die Precision 100% ist, wurden alle relevanten Dokumente gefunden → Falsch

-Keyword-Suche findet nicht alle relevanten Dokumente

Recall → wenn gleich 100% dann wurden alle relevanten Dokumente gefunden.

Die Präzision gibt uns ein mass dessen wieder wie gut die Antwort der Such Systeme ist, und recall

gibt uns ein mass dessen wieder welches Ausmass der relevanten Dokumente in diese Antwort

enthalten ist (La precisione ci dà una misura di quanto è buona la risposta dei sistemi di ricerca, e il

richiamo ci dà una misura di quanto dei documenti rilevanti sono inclusi in quella risposta).

Precision = |Relevant und gefunden| / Gefunden

Recall = |Relevant und Gefunden| / Relevant

Wie kann man diese Problematik angehen?

-Semantik

-Personalisierung

-Page Rank

Page Rank:

Versucht anhand der Linkstruktur die Wichtigkeit von Dokumenten zu erschliessen (Tenta di determinare l'importanza dei documenti sulla base della struttura dei link)

Strukturierte Daten:

- Haben explizite, formaler Struktur
- Semantik vorgegeben
- Automatische Nutzbarkeit möglich
- Durchsuchen: Problemlos

Daten werden strukturiert in einer Datenbank gespeichert

-Beziehungen durch mathematische Relationen modelliert.

Verwaltung, z.B. mittels Datenbankmanagementsystemen
Suche mit logikbasierter Sprache
-Relevanz gemäss Suchkriterien
-Zusätzliche Ergebnisse dank Nutzung von Relationen
Von Unstrukturierten zu Strukturierte Daten
-XML (Extensible Markup Language) ist ein (menschenlesbares) Datenformat
-Häufig zum Datenaustausch verwendet
-Trägt einen Teil der Strukturinformation mit sich.
Fazit:
-Viele unstrukturierte Daten → geben den Computer Probleme (durch Suchen, etc.)
-Struktur vereinfach Verarbeitung
-XML ist eine Strukturierungsmöglichkeit

Datenformate

Dateien haben verschiedene Formate (z.B: doc,html,mp3,ogg,wav,raw,jpg,png).

Formate:

- Interpretation der Inhalte (Inhalte = contenuti)
- Werden über oft Endungen identifiziert
- Werden durch Hersteller oder Gremien identifiziert

Formaten unterscheiden sich bezüglich:

- Struktur (Syntax und Semantik)
- Möglichkeit für Digital Rights Management (DRM)
- Datenkompression

Digital Rights Management (DRM):

Definition:

-Digitales Management von Rechten, um die Verbreitung und Nutzung digitale Güter durch technische Massnahmen zu kontrollieren (=Gestione digitale dei diritti per controllare la distribuzione e l'uso di beni digitali attraverso misure tecniche, es. license key di sylenth1 o di serum, etc.)

Ziel:

- Durchsetzen von Urheberrechten
- Übertragen bestehender Geschäftsmodelle auf digitale Welt

Übersetzung:

Obiettivo:

- imporre il diritto d'autore
- Trasferire i modelli di business esistenti al mondo digitale

Vier Rechte stehen typischerweise im Fokus:

-Transportrechte

Kopieren, Übertragen, Verleihen

-Vorführrechte

Anzeigen, Ausdrucken, Abspielen

-Sicherungsrechte

Schutz vor Beschädigung oder Verlust

-Editierrechte

Verändern, Extrahieren

Nicht mehr das digitale Gut, sondern die Nutzung kostet

- Rechteinhaber vergibt Nutzungslizenz
- Definiert, wer, was, wann, wie und wie oft innerhalb welcher Zeitperiode zu welchem Preis mit den

Daten machen darf:

- Access Control
- Usage Control
- Content Tracking
- Payment Management

Lizenzennehmer hat das Recht, das Gut gemäss den vereinbarten Konditionen zu nutzen.

Kritik:

- Kein einheitlicher Standard → Kompatibilitätsprobleme
- Systeme nicht sicher → DRM kann umgangen werden (Systeme können gehackt werden)
- Wettbewerb wird geschwächt → Markzutrittsschranken (barriera all'entrata nel mercato)
- Datenschutz
- Kosten vs Nutzen

Je multimedialer, desto grösser

1 Seite Text: 5.1kB

1 Bild: 6.5MB

1 min Ton: 10.6MB

1 min Film: 1.3GB

Datenkompression verringert Datengrösse

→ weniger Platz

→ schnellere Übertragung

Datenkompression: verlustfreie Verfahren

-Redundante Daten werden durch kürzere Zeichen ersetzt.

-Musterwiederholung (z.B. Lauflängenkodierung)

-Reversibel: Daten lassen sich wieder Bit-genau herstellen

Datenkompression: verlustbehaftete Verfahren (si perdono dei dettagli dall'originale)

-Nutzen die Begrenztheit menschlicher Sinne:

-Helligkeit wird stärker registriert als Farbe oder Farbsättigung

-Bestimmte Ton-Frequenz werden nicht wahrgenommen

-Kompression entfernt Teile des Originals

-Nicht reversibel: Rekonstruktion entspricht nur näherungsweise den Originaldaten.

Fazit:

-Zahlreiche Formate

-Einige ermöglichen eine Kontrolle der Nutzung

-Dateigrösse kann durch Kompression verringert werden

Kapitel 3: Datenmanagement

Um was geht es bei Datenmanagement?

Kurz gesagt:

-Es geht um das Sammeln, Speichern und Verarbeiten von Daten

-Das Ganze sollte zuverlässig, sicher und effizient geschehen

-Macht auch schon Sinn für weniger als Millionen von DVDs...

Was geht mich das an?

-Forschungsprojekte

-Industrie 4.0

Was gibt es zu beachten?

-Wie finde ich was ich suche?

-Hilfe, mein Rechner stürzt ab!

Dateiformate

-Textdateien

-Worddateien

-Spreadsheets

-PDFs

-JSON

-XML

-HTML, etc.

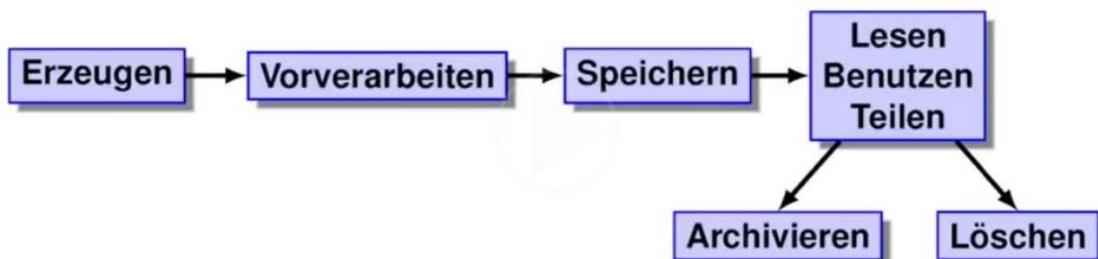
Sicherheit

-Datenschutz

-Privatsphäre

Lebenszyklus von Daten

Lebenszyklus von Daten (Data Life Cycle)



1. Daten erzeugen (generare dati)

-Digitalisierung

-Logdateien

-Sensoren

2. Daten vorverarbeiten (pre-elaborazione dei dati)

-Qualitätsprobleme

3. Daten speichern

-müssen auffindbar sein

-und möglichst schnell

4. Daten nutzen

5. Archivierung/Lösung

-keine aktive Verwendung

-rechtliche Gründe (= ragioni legali)

Viele Daten leben lange:

-Wegen Langlebigkeit macht sorgfältiges Datenmanagement Sinn

-Daten können für andere Leute interessant sein:

-z.B. SBB-Daten, Open-Data Initiativen, etc.

Spreadsheets

VisiCalc (precede Excel)

-Eine Idee Ende 1978

Was ist eigentlich ein Spreadsheet?

Kurz gesagt: ein Werkzeug für die

-Eingabe (inserire i dati)

-Speicherung

-Analyse und

-Visualisierung

Von Daten.

Aufbau eines Spreadsheets

Ein Spreadsheet besteht aus einer Tabelle mit Zeilen und Spalten:

-oberste Zeile: Beschreibungen

-jede Zeile darunter: ein Datensatz

-jede Spalte: eine Variable

-jede Zelle (= 'cella'): ein Wert (kann auch berechnet werden)

Generelle Best Practices:

-Namensgebung

-Metadaten anlegen (Metadaten oder Metainformationen sind strukturierte Daten, die Informationen über Merkmale anderer

Daten enthalten. Bei den durch Metadaten beschriebenen Daten handelt es sich oft um größere

Datensammlungen wie Dokumente, Bücher, Datenbanken oder Dateien).

-Backups

-Als Text abspeichern

Best Practices für Spreadsheets:

-Nur ein Wert pro Zelle

-Zellen nicht einfach leer lassen (perché dopo non si riesce a capire se le celle sono vuote perché i dati non esistono o se sono vuote perché i dati non si conoscono)

-Keine Berechnungen auf den Rohdaten2

-Datavalidierung benutzen

-Einfach einmal einen Datensatz in einem Satz zusammenfassen

Vorteile und Grenzen von Spreadsheets

Vorteile:

+ Es ist einfach Daten einzutragen und loszulegen

+ Daten können auch importiert und exportiert werden

Nachteile/Grenzen:

-Bei Millionen von Datensätzen wird es mühsam

-Probleme mit der Datenqualität

-Keine Mehrbenutzersynchronisation (= nessuna sincronizzazione multiutente)

Datenbanksysteme

Ursprung (origine) von Datenbanksystemen:

60er und 70er: Immer wieder die gleichen Probleme

Warum kein System bauen, das diese Probleme löst?

-Senkt Kosten und Entwicklungszeit (von Anwendung) → riduce i costi e il tempo di sviluppo dell'applicazione

-Die Idee von Datenbanksystemen war geboren

Caratteristiche dei Datenbanksysteme:

-Datenunabhängigkeit

-Anfragesprachen

-Fehlerbehandlung

-Mehrbenutzersynchronisation

-Datenintegrität

-Skalierbarkeit (schnell Daten zu suchen)

Bekannte Datenbanksysteme:

-Oracle

-Microsoft SQL Server

-MySQL

Datenmodellierung

Relationale Datenbanken:

In einem relationalen Datenbanksystem

-werden die Daten in Tabellen abgespeichert.

-Die Tabellen bestehen aus Zeilen und Spalten.

Es werden oft mathematische Begriffe verwendet:

-Tupel = Zeile

-Attribut = Spalte

-Relation = Tabelle

Aufbau von Relationen:

Eine Relation besteht aus zwei Teilen:

-einem Schema (symbol R), das sind die Metadaten

-einer Instanz (symbol R), das ist der Inhalt der Relation (Inhalt = contenuto)

Datenbankentwurf:

-Zuerst konzeptuelles Schema

-Dann Umsetzung in Relationen

-Ein wichtiges konzeptuelles Modell:

Entity-Relationship-Modell:

Das Entity-Relationship-Modell hat folgende Grundbausteine:

-Entität (Entities)

-Attribute

-Schlüssel

-Beziehungen (Relationships)

Entität:

-Eine Entität ist ein Ding, eine Person, ein Ort oder ein Konzept

-Eine Sammlung von gleichen Entitäten bildet eine Menge.

-Entitätsmengen werden durch Rechtecke dargestellt. (vengono rappresentati con un rettangolo)

Attribute:

-Eigenschaften von Entitäten werden mit Attributen beschrieben

-Ein Attribut, das eine Entität eindeutig beschreibt, ist ein Schlüssel

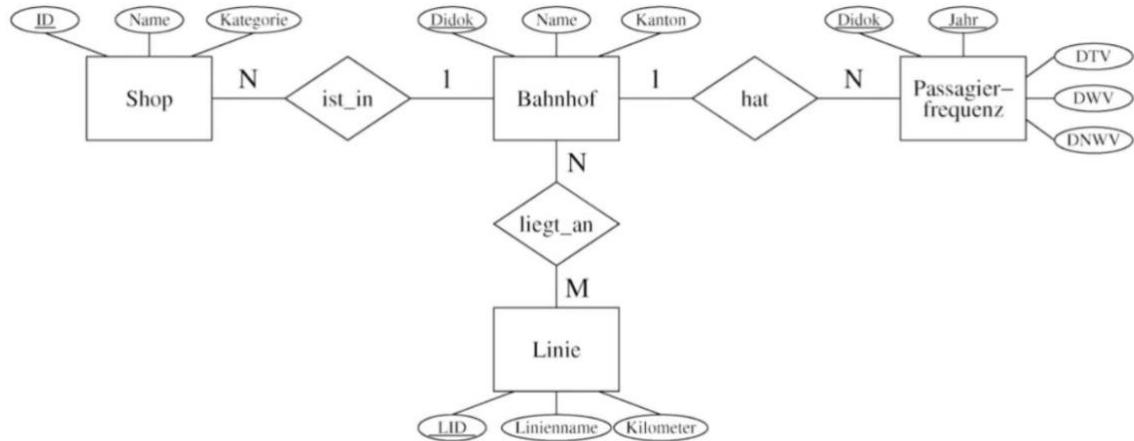
-Attribute werden durch Ellipsen repräsentiert

Beziehungen:

-Eine Beziehung verbindet zwei oder mehr Entitätsmengen

-Darstellung durch eine Raute mit Funktionalitäten

-Beispiel für eine Binare Beziehung:



I rettangoli sono le Entità, gli ovali (cerchi) sono gli Attribute e le linee che uniscono i rettangoli

sono le beziehungen

L'1 e la N sono Funktionalitäten.

Eine Attribute bezeichnet eine Entität

Didok e Id sono le chiavi (Schlüssel); le chiavi possono chiamarsi in qualsiasi modo, non sono sempre

didok e id, ad esempio per l'uni si può mettere Matrikelnummer e Vorlesungsnummer come chiavi.

Posso scegliere qualsiasi nome per la\le chiave\i.

Con la Schlüssel un valore puo uscire solo una volta, ad esempio no posso mettere due volte lo stesso matrikelnummer, ma solo uno per studente.

Umsetzung in Tabellen:

Beispiel Bahnhof:

-Bahnhof (Didok, Name, Kanton)

-Shop (ID,Name,Kategorie)

-Passagierfrequenz (Didok, Jahr, DTV, DWV, DNWV)

-Linie (LID, Liniename, Kilometer)

Was machen wir mit den Beziehungen?

1:N

-Shop (ID,Name,Kategorie,Didok)

-Passagierfrequenz (Didok, Jahr, DTV, DWV, DNWV)

N:M

-Bahnhoflinien (Didok, LID)

Datenbankabgrafen

SQL:

→Structured Query Language (linguaggio dei Datenbanksysteme)

Relationen verbinden:

Informationen können über mehrere Relationen (Tabellen) verteilt sein.

Zwei Tupel dürfen nur kombiniert werden, wenn ihre Werte für Didok (Bahnhof Identifikator)

übereinstimmen.

Daten ändern:

- Relationen anlegen
- Daten einfügen ('insert into')
- Daten löschen ('delete from')
- Vorsicht: bei leerer Where-Klausel wird alles gelöscht!
- Daten aktualisieren ('update' → ('set' → 'where'))

-'Select': welche Spalten
-'From': von welcher Tabelle
-'Where': mit welchem Filter

Select * from Immatrikulation

Where Immatrikulationsjahr = 1897

Select * from Immatrikulation seleziona tutto quello che c'è nella tabella Immatrikulation.
'Where' seleziona la sezione della tabella chiamata Immatrikulationsjahr e seleziona l'anno 1897.

Big Data

Ansätze (approci):

- Scale-Up
- Scale-Out

Scale-Up vs Scale-Out:

Scale-up: hinzufügen von Erweiterungen wie mehr oder größere Festplatten und mehr Arbeitsspeicher, um die Leistung von Rechnern zu erhöhen.

Scale-out: hinzufügen von weiterem Rechner in der Architektur, um die Arbeitslast diese zu verteilen.

Auf italienisch:

Scale-up: aggiunta di estensioni come dischi rigidi più grandi e con più memoria per aumentare le prestazioni dei computer.

Scale-out: aggiunta di altri computer all'architettura per distribuire il carico di lavoro.

Ende 1990er:

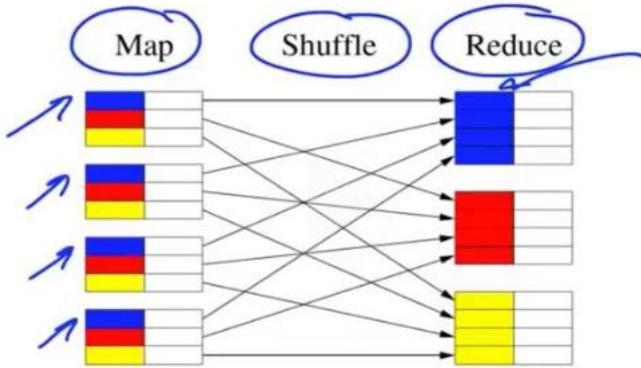
Firmen wie Amazon, Ebay, Google, Yahoo produzieren riesige Datenmengen.

Scale-Up war keine Option mehr. Relationale Datenbanksystem kommen gut mit Scale-Up zu Recht aber weniger gut mit Scale-Out.

Fehlertolerante Systeme (flessibili agli errori):

- Google entwickelt das Google File System und Map-Reduce
- Yahoo: Hadoop Open-Source Framework (samt Dateisystem)
- Apache: Spark

Map-Reduce:



Rückblick & Ausblick

Datenmanagement ist etwas sehr modern.

Aber die Menschheit beschäftigt sich seit langem mit Datenmanagement (l'umanità ha avuto a che

fare con la gestione die dati per molto tempo, nonostante sia qualcosa di moderno)

Kurze Geschichte des Datenmanagements:

3400-3100 v.Chr. Erfindung der Schrift

2000 v.Chr.: sumerische Tontafellisten

300 V.Chr.: Bibliothek von Alexandria

1440: Erfindung des Buchdrucks

Usw.

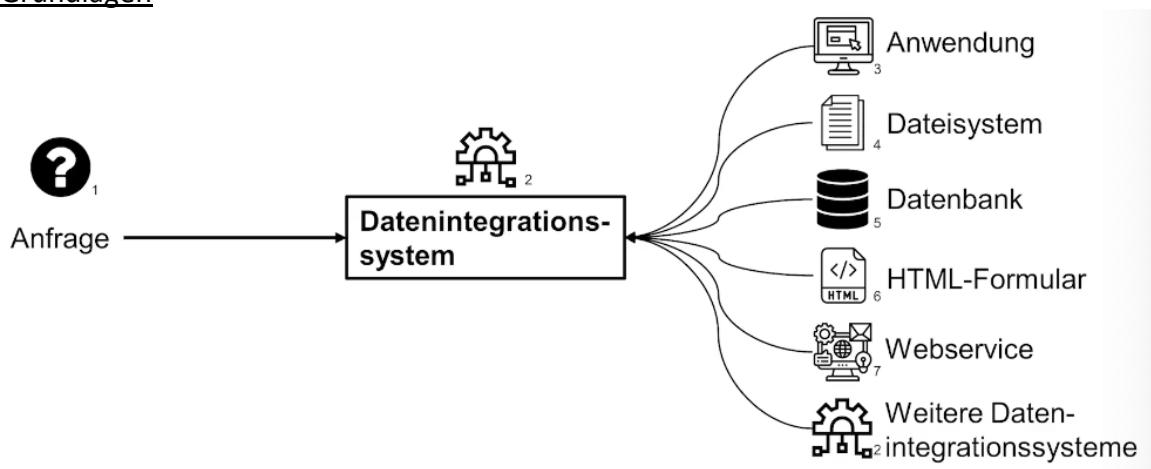
Kapitel 4: Datenintegration

Integrationsvorgang: heterogene Daten werden harmonisiert und einheitlich abrufbar gemacht. Beispiel: Preisvergleich Website -> erlaubt eine einzige Abfrage bei einem einzigen Portal.

Vorteil: Gewinn neuer Erkenntnisse durch Kombination von Datenquellen. Mehr Daten, komplettere Daten, korrektere Daten.

Datenintegration ist die Verwendung der Daten für die Tools unseres digitalen Zeitalters (Excel, Datenbank, WWW, usw.).

Grundlagen



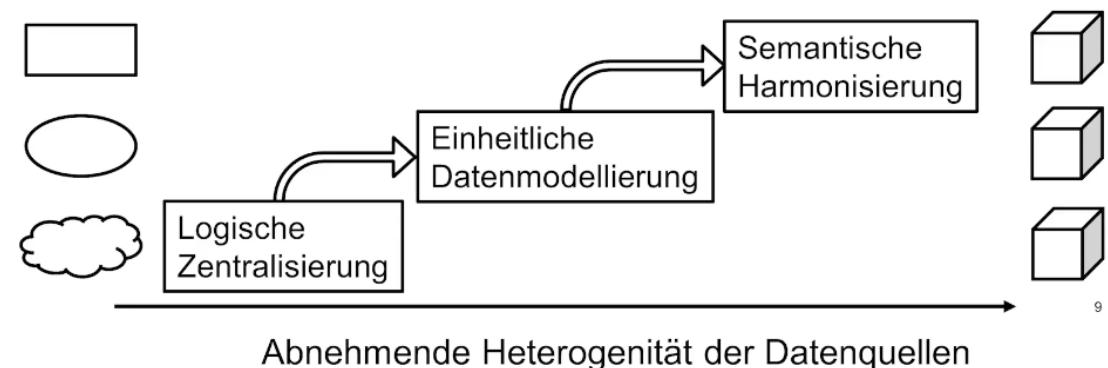
Welche Unterschiede der Datenquellen: Struktur der Daten, Datenmodell, Schema, Datentypen/-werte.

Datenstrukturen der Datenquellen: Strukturierte Daten (daten in einer Datenbank) und Unstrukturierte Daten.

Datenmodell: für Repräsentation zuständig. Z.B: Relationalles Modell.

Schema: für die Organisation der Daten verantwortlich. Tabelle mit Attributen.

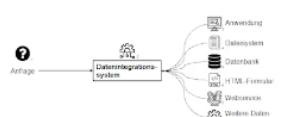
Datentypen & -werte: codieren die eigentliche Informationen der Datenquellen.



Datenintegrationsarchitekturen:

Materialisierte Integration

- Datenquellen werden in gemeinsamer Datenbank integriert (Date Warehouse)



Virtuelle Integration

- Daten bleiben in den Quellsystemen
- Anfrage wird auf Quellsysteme gemappt

Metadaten: Daten über Daten.

Hürden der Datenintegration (Ostacoli)

Bei der Heterogenität der Datenquellen unterscheiden wir zwischen:

- Technischer Heterogenität: unterschiedliche technische Kommunikation zwischen Datenquellen und Integrationssystem
- Syntaktischer Heterogenität: unterschiedliche Darstellung gleicher Informationen. (e.g. 1->Ja, 0 -> Nein und dann umgekehrt)
- Struktureller Heterogenität: strukturell unterschiedliche Darstellung gleicher Konzepte / Verwendung unterschiedlicher Schemata
- Semantischer Heterogenität: abweichende Bedeutung von Schemaelementen und Datenwerten

Datenfusion: erkennen, bereinigen und fusionieren von ähnlichen Dateneinträgen

Weitere Hürden: Sicherheit, Aufwand von kosten, legale und ökonomische Gründe

Datenintegrationsarchitekturen

Materialisierte Integration vs. Virtuelle Integration.

Materialisierte Integration: kopieren & überführen von Informationen aus unterschiedlichen Datenquellen in einheitliche Zielstruktur einer neuen Datenbank.

Virtuelle Integration: Daten werden in globales Schema gemappt und verbleiben in den Quellsystemen.

Vergleich Integrationsarchitekturen

	Materialisierte Integration	Virtuelle Integration
Datenpipeline	ETL (extract-transform-load)	Wrapper
Datenzugriff	Kopie originaler Daten	Mapping der Daten
Integrationszeitpunkt	Periodisch	Zeitpunkt der Abfrage
Aktualität	Abhängig von Periodizität	Daten stets aktuell
Antwortzeit	Schnell	Langsam
Datenqualität	In der Regel höher	In der Regel geringer

Hybride Architekturen: kombinieren Vorteile materialisierter & virtueller Integrationsarchitekturen. Beispiele: Caching, Partielle Materialisierung.

Caching: Zwischenspeicher enthält Resultat häufiger Anfragen. Erlaubt reduzierte Anfragezeiten.

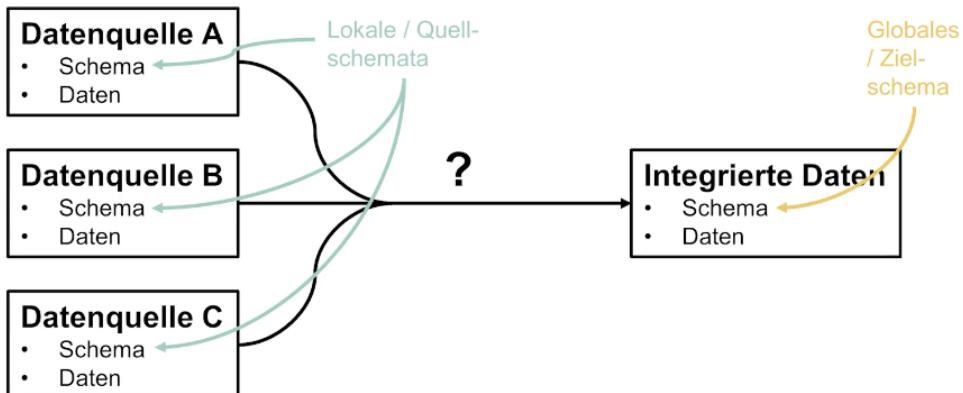
Partielle Materialisierung: Kombination aus virtueller und materialisierter Integration.

Vorteile sind stets aktuelle Daten, schnelle Anfragen.

Data Lake: System zur Speicherung von Roh- und transformierten Daten; kann als Datenquelle dienen.

Schemamanagement

Ziele: Korrespondenzen zwischen Datenquellen, einheitliches (globales) Schema, Transformationsregeln festlegen.



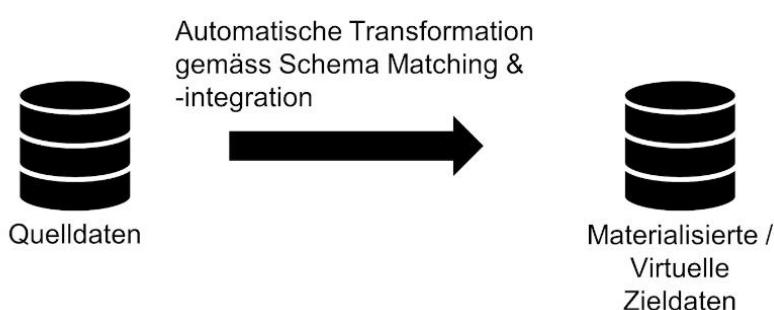
Schema Matching - Methoden.

- 1) Manuell: Matching wird manuell von einem Entwickler erstellt. Nachteile: Technisches Know-How, zeit-/arbeitsintensiv.
- 2) Lern-basiert: Computer lernt Matching durch Beispiele - Entwickler markiert zugehörige Attribute in Quell- & Zielschema. Nachteile: Anfällig für Ausnahmen.
- 3) Automatisch: Computer erstellt Mapping selbstständig. Nachteil: sehr rechenintensiv, nur einfache Mappings, anfällig für ausnahmen
- 4) Interaktiv: Kombination aus lern-basierten und automatischen Ansatz; Entwickler gibt Feedback bei Ungewissheiten. Nachteil: Komplexes Program.

Probleme der Schema Matching: Gross und Kleinschreibung, Rechtschreibfehler, Unterschiedliche Verwendung von Sprachen.

Schemaintegration: Fusionierung von lokalen Schemata in globales Schema

Schema Mapping (Schema-Abbildung)



Warum nicht einfach Standards verwenden? Datenquellen historisch unabhängig gewachsen, Definition schwierig, nicht immer sinnvoll.

Standards

Was ist ein Standard?

Ein **Standard** ist eine einheitliche Weise, etwas zu bewerkstelligen; und die sich gegen andere Vorgehensweisen behauptet hat.

Ein **Datenstandard** ist eine einheitliche Weise, Daten zu modellieren, auszutauschen oder zu repräsentieren

Datenstandards und Datenintegration

Datenstandards helfen bei der Überwindung von **technischen**, **strukturellen** und **semantischen** Heterogenitäten der Datenquellen

- **Standardisierte Datenformate und -modelle:** Entscheidend für die harmonisierte Strukturierung und dem Austausch von Daten
- **Terminologiestandards:** Entscheidend für die harmonisierte Repräsentation von Daten

Praxis

Ziel: unkomplizierte Verwendung des integrierten Data.

Erste Schritte: vertraut werden mit Daten, einfache Fallbeispiele, Divide & Conquer.

Datendokumentation: Reproduzierbarkeit, FAIR Data Prinzipien (für Forschungsdaten).

FAIR: Findable, Accessible, Interoperable, Reusable.

Kapitel 5: Statistik

Stichprobe muss repräsentativ sein (e.g., zufällige Stichprobe).

Mittelwert, median: Lagemasse

Varianz, Interquartilsabstand, Range: Streuungsmasse

Statistische Tests, Lineare Regression, Plots.

Was kann man mit einem Modell machen?

- Modell auswählen, das die Muster in den Daten am besten beschreibt

- Parameter interpretieren

- Vorhersagen machen (predizioni)

Kapitel 6: Qualitative Datenanalyse

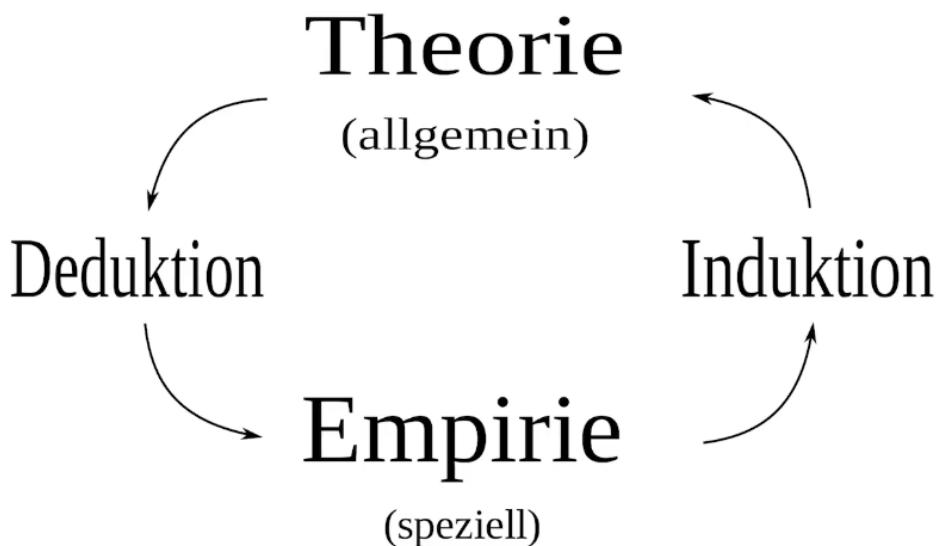
Wieviel interessiert nicht, warum ist interessant (e.g. Sport/Training). Wieviel interessiert die quantitative Forschung.

Die qualitative Forschung ist eine ergebnisoffene Forschung.
Die Forschungsfrage erlaubt Dinge unbeachtet zu lassen. Sie ist der Rettungsanker bei der Methodenwahl.

In der quantitativen Forschung spricht man von forschungsfrage, in der qualitativen Forschung spricht man von Erkenntnisinteresse.

Empirische Sozialforschung: die systematische Gestaltung von Erkenntnisprozessen. Es geht um Wahrscheinlichkeiten, nicht die absolute Wahrheit.

Empirie: Sammlung von Daten, und aus diesen Daten leitet man eine Erkenntnis ab.
(raccogliere dati e ricavarne conoscenza).



Deduktion ist die quantitative Forschung.

Induktion ist die qualitative Forschung (von Empirie zu Theorie).

Leitfragen qualitativer Forschung:

- Nachvollzug von subjektivem Sinn
- Beschreibung der Herstellung von sozialem Handeln
- Rekonstruktion tiefer liegender Strukturen sozialen Handelns

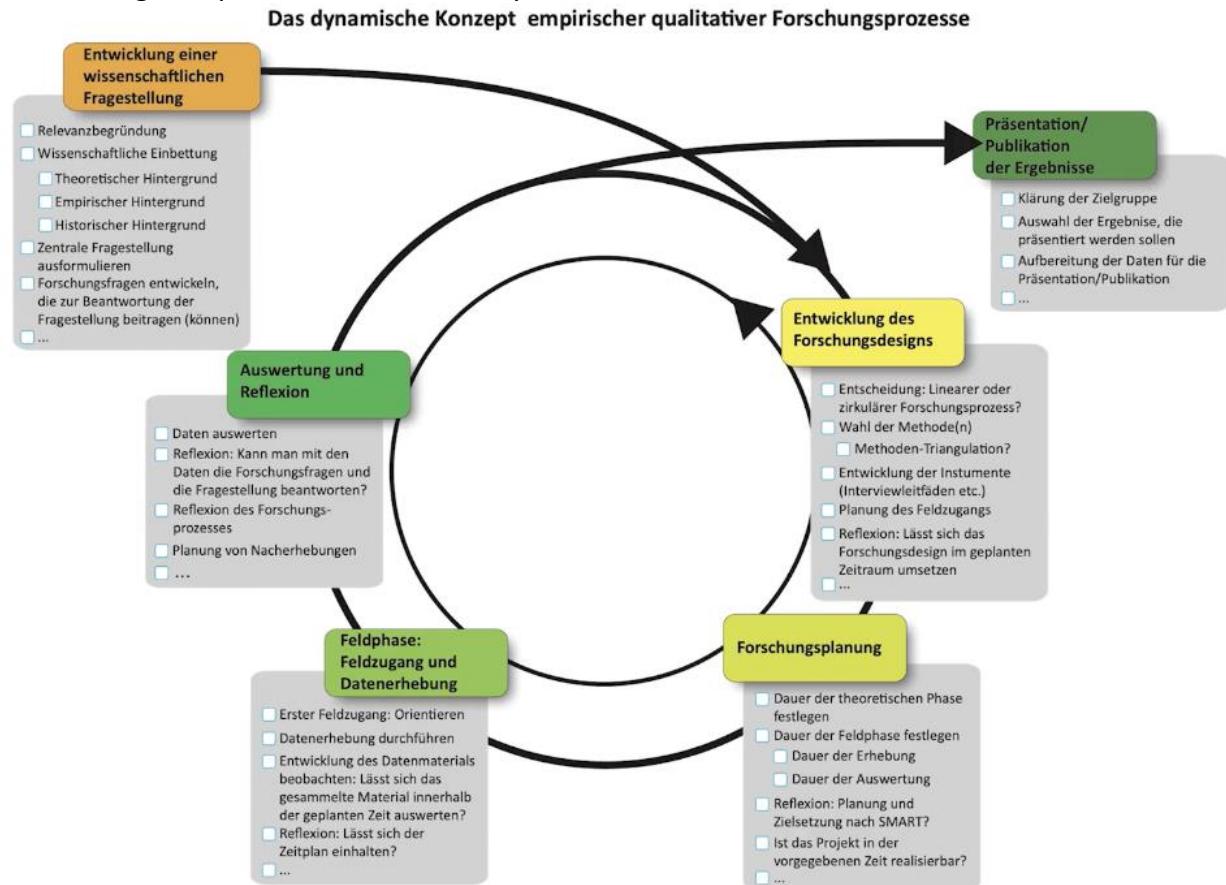
Arten von Qualitativer Forschung:

- Mikroskopische/rekonstruktive qualitative Forschung: kleine Textsegmente, Satz für Satz durch, Wort für Wort. (Narrative und hermeneutische Analysen) -> Rekonstruktion der Fallstruktur
- Makroskopische/qualitative qualitative Forschung: größere Textsegmente und sage was drin passiert (Kodierung und Kategorisierung) -> Kategorisierung (große Datensätze).
Forschungsablauf: zirkulär in qualitative Forschung, linear in quantitative Forschung.
Qualitative Inhaltsanalyse: ist ein wesentlicher methodischer Zugang in der Medienforschung. Es wurde als systematische Methode zur Auswertung von Materialen der immer wichtiger werdenden Massenmedien in der Kommunikationswissenschaft entwickelt und hat damit die empirische Medienforschung mit begründet.
Vorteile der Inhaltsanalyse: systematische, intersubjektiv nachvollziehbare Bearbeitung grosser Materialmengen. Ist die Analyse von Textmaterial unterschiedlicher Herkunft.
Verwendung von Kategorien.

Ablauf einer qualitativen Inhaltsanalyse:

- Festlegung des Materials
- Analyse der Entstehungssituation (e.g., Interview)

- Formale Charakteristika des Materials
- Richtung der Analyse - Autor, soziokultureller Hintergrund, Wirkung.
- Theoretische Differenzierung der Fragestellung
- Bestimmung der passenden Analysetechnik. Analysetechniken: Zusammenfassung, Explikation, Strukturierung.
- Festlegung des konkreten Ablaufmodells
- Festlegung und Definition der Kategorien oder Kodes / des Kategoriensystems - deduktiv und/oder induktiv (hole aus Theorie oder aus Empirie).
- Definition der Analyseeinheiten. Analyseeinheiten: Kodiereinheit, Kontexteinheit, Auswertungseinheit.
- Durchführung der qualitativen Inhaltsanalyse



Formen qualitative Online-Daten

- Kommentare
- Videos, Bilder, Emojis

Kriterien zur Beurteilung der Online-Daten:

- The Medium is the Message
- Tools setzen Rahmenbedingungen
- Daten werden nicht in einem Forschungssetting erhoben
- Daten werden nicht mit Fokus auf die Forschungsfrage generiert -> Auswahl der relevanten Daten
- Technische Voraussetzungen und technisches Wissen der Nutzer

- Nicht alle hinterlassen "Spuren" im Internet (z.B. lesen/schreiben, Anzahl aktiver Nutzer, Nutzer die kein Zugang zum Internet haben).
- Soziale Erwünschtheit
- Nachfragen sind nicht möglich

Gütekriterien (Criteri di qualità)

Gütekriterien sind Kriterien, um die Qualität des Weges zur wissenschaftlichen Erkenntnisgewinnung durch bestimmte Methoden festzuhalten.

Klassische Gütekriterien

– Objektivität

Sind die Ergebnisse unabhängig von ungewollten Einflüssen durch die erhebenden Personen entstanden?

– Reliabilität

Wie genau ist die Messung?

– Validität

Wird wirklich das gemessen, was man messen möchte?

Unterschiedliche Standards guter Forschung:

Quantitative Forschung:

- Beobachterunabhängigkeit bzw. Objektivität der Datenerhebung und auswertung
- Theoriegeleitetes Vorgehen
- Statistische Verallgemeinerbarkeit der Befunde

Qualitative Forschung:

- Erkundung von Sinndeutungsvorgängen (wieso tun Leute eine Sache, wieso tun Leute eine Sache nicht)
- Exploration kultureller Praktiken und Regeln
- Genaue und tiefgehende Analyse und Beschreibung von Einzelfällen



Validitätsformen qualitativer Forschung: Okologische Validierung, Kommunikative Validierung, Validierung durch Triangulation

Software zur Auswertung qualitativer Daten

QDA-Software: qualitative data analysis software

Paper & Pencil vs QDA-Software

QDA ist keine eigenständige Auswertungsmethode, sondern ein Hilfsmittel für die qualitative Datenanalyse unterschiedlicher Methoden.

Gemeinsamkeiten Auswertungsmethoden

- Subsumieren, Zusammenfassen, Kategorisieren
- Anbringen zusätzlicher Informationen, Paraphrasieren, Kontextinformationen → Memos
- Aufschreiben was im Kopf passiert: Die Gedanken und Ideen beim Codieren sind sehr wichtig und sollten daher in Memos festgehalten werden.

Da oben ist für qualitative Forschung, in Quant. Forschung ist nicht wichtig, was im Kopf passiert.

Daten sprechen nicht für sich selbst -> guten Fragen an die Daten stellen.

Kapitel 7: Datenvizualisierung

Einführung

Datenvizualisierungen sind überall (ad esempio mappa con i casi covid nel mondo)

Die 5 Ws der Visualisierung:

Warum/Was/für Wen/Wozu und Wie wird visualisiert?

Warum visualisiere ich?

Erkenntnisgewinn aus Datenhaufen

Für wen visualisiere ich?

Wahrnehmungseffekte

Was, wozu und wie visualisiere ich?

1. Mengen

2. Verteilungen

3. Anteile

4. Zusammenhänge

Ausblick

1. Daten mit Zeitbezug (Zeitdaten)

2. Raumdaten

Warum

Warum visualisiere ich?

1. Kommunikation eines Sachverhalts

2. Wenn ein Bild mehr als tausend Wörter sagt

3. Um effizient & effektiv den Erkenntnisgewinn aus Datenhaufen zu erleichtern.

Wenn die Daten bekannt sind, beantwortet das Visualisierte Muster den Sachverhalt, den ich kommunizieren will.

Wenn die Daten noch zu erkunden sind, wirft das visualisierte Muster weiterführende Fragen oder

vielleicht zu testende Hypothesen auf.

Was

Was wollen wir visualisieren? Was visualisiere ich?

1. Datentyp (qualitativen: Texte, quantitativen: Sie befinden sich in einer Tabelle (z.B. Zahlen))

2. Datenmerkmale

Was visualisiere ich?

1. Geometrie

2. Graphische Variablen

Wer

Für wen visualisiere ich?

Für Personen mit einem Informationsbedürfnis unter Berücksichtigung ihrer visuellen Wahrnehmung.

Menschliche Wahrnehmung

1. Perzeption

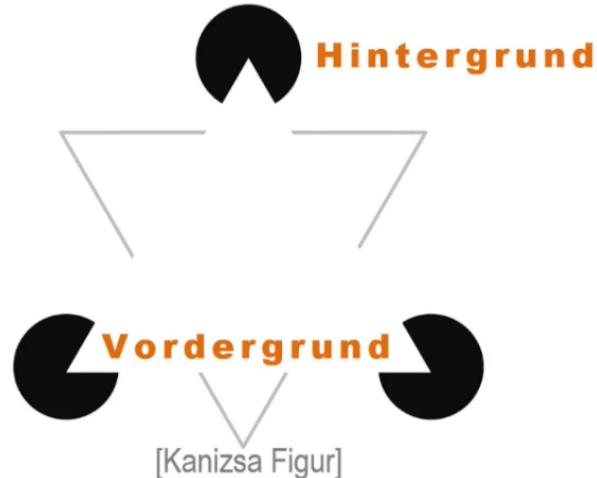
2. Kognition

Graphische Lösung?

Das Prägnanzprinzip

Vordergrund, Hintergrund

Was ist visuell prägnant?



<<Gesetz der Guten Gestalt>>:

1. Nähe
2. Ähnlichkeit
3. Geschlossenheit
4. Kontinuität

Wozu

Wozu visualisiere ich? Wozu = per che cosa/a che scopo

Um die richtige Antwort auf mein Informationsbedürfnis zu erhalten.

1. Mengen
2. Verteilungen
3. Anteile
4. Zusammenhänge

Kapitel 8: Computational Thinking

Einführung

Real Problem → Algorithmus finden → eine Lösung beitragen.

Algorithmen verändern die Welt.

Algorithmen sind praktisch überall und verändern unseres Leben (positiv aber auch Negativ (z.b: fake News)).

Einflussreiche Algorithmen:

- Google Suche
- Facebook Newsfeed
- Produkt-empfehlung

CT löst systematische Probleme mit Algorithmen.

Definition CT: 1) Fähigkeit, ein Problem in kleinere Probleme zu zerlegen, die deterministisch gelöst

werden.

Zweite Definition: Die Denkprozesse, die bei der Formulierung von Problemen und deren Lösungen

ablaufen (essere coinvolti), so dass die Lösungen in einer Form dargestellt werden, die von einem

Informationsverarbeitenden Agenten effektiv durchgeführt werden kann.

→Komplexe Probleme in einfacher Probleme (einzelne Schritte) zu lösen.

Es. Se devo trovare la via da casa all'uni con ZVV prima trovo la stazione più vicina a casa mia, poi

guardo che linea di tram devo prendere, poi guardo se il tram è diretto oppure devo scendere dal

primo e prenderne un altro, e poi guardo a che ora prendere il tram.

Das Komplex Problem wurde in kleinere einfachere Probleme durchgeschnitten.

Algorithmen sind überall, nicht nur im Softwares.

CT – Andere Definition: Computational Thinking ist eine Art und Weise, wie Menschen Probleme

lösen; es ist nicht der Versuch, Menschen dazu zu bringen, wie Computer zu denken.

Computer sind

stumpfsinnig und langweilig; und Menschen sind klug und einfallsreich. Wir Menschen machen

Computer spannend.

Was ist CT und was ist es nicht

-Konzeptualisierung, nicht Programmierung

-Fokus auf Ideen & Prozessen, nicht Software und Hardware.

Für jeden, nicht nur für Informatiker*innen

Hauptbestandteile 1

Hauptbestandteile= componenti/caratteristiche principali.

1) Dekomposition: Sortieren, zusammenlegen (Ordinare e piegare)

2) Mustererkennung (identificare, riconoscere il modello)

3) Abstraktion

4) Algorithmus: Ablauf festlegen (determinare il processo)

→Algorithmus spielt eine große Rolle.

Dekompostion

Scomporre il problema in problemi più piccoli e più semplici.

→Zerlegung von komplexen Strukturen in Teilprobleme

Methode:

-Zielreduktion→z.b Problem halbieren (dimezzare il problema)

-Rekursive Zielreduktion→Wiederholte Anwendung der Problemzerlegung (Applicazione ripetuta

della decomposizione del problema)

Es quando devo scrivere un tema prima faccio una scaletta e suddivido il testo in introduzione,

svolgimento e conclusione, in questo modo suddivido il problema in parti più piccole.

Altro esempio: creare un app sul telefono chiamata “Buch2Go” dove la gente può comprare libri di

seconda mano.

Dekomposition von Problem:

- Budget
- Funktionalitäten? Erfassen, suchen, kaufen
- Zielpublikum?
- Marketing?
- Zeitplan?

Mustererkennung: Ähnlichkeiten & Gemeinsamkeiten zwischen und/oder innerhalb von Problemen identifizieren

-Vereinfachung der Problemlösung durch Wiederverwendung der gleichen Lösung.
Abstraktion & Datenrepräsentation
Einteilung in wichtige und unwichtige Eigenschaften des Problems, Fokus auf relevante Aspekte.
→ Classificazione in caratteristiche importanti e non importanti del problema, concentrarsi sugli aspetti rilevanti.

Arten (Tipi):

- Datenabstraktion (größe, farbe, schriftgröße)
- Abstraktion der Steuerung (steuerung=controllo)

Hauptbestandteile 2

Rezepte sind Algorithmus → Problem in kleinere Schnitten geteilt.

Algorithmus – Definition: Ein Algorithmus ist eine endliche Folge wohldefinierten Anweisungen,

typischerweise zur Lösung einer Klasse von Problemen. (Un algoritmo è una sequenza finita di

istruzioni ben definite, tipicamente per risolvere una classe di problemi)

Andere Definition: Algorithmen bestehen aus endlich vielen, wohldefinierten Einzelschritten. Bei der

Problemlösung wird eine bestimmte Eingabe in eine bestimmte Ausgabe überführt.

Problemlösung: Eingabe → Algorithmus → Ausgabe (wie Rezepte)

Es: se voglio prendere treno da lugano a zugo con sbb app:

Algoritmo:

1. Auf zeitplan tippen
2. Felder 'von' und 'nach', 'datum' ausfüllen
3. Zugverbindung auswählen
4. Ticket kaufen

Questo algoritmo non funziona solo per la tratta lugano-zugo ma per tutte le tratte in treno in

svizzera.

Kriterien von Algorithmen:

-Maschinentauglichkeit → idoneità della macchina → Anweisungen die von einer Maschine gelesen

werden und ausgefüllt können.

-Allgemeinheit (muss ähnliche Probleme lösen)

-Korrektheit (muss die richtige Lösung finden)

Achtung! Es gibt Probleme die Algorithmen nicht lösen können, bzw. Halte-Problem.
Das Halte-Problem ist ein ungelöstes Problem.

7 Fertigkeiten (competenze)

- 1) Modellierung (Real Welt in virtuelle Welt modellieren, Realität simulieren)
Es. Per Google Maps devi mettere la mappa reale del mondo virtuale nell'app.
- 2) Wissenschaftliches Denken
- 3) Logisches Denken (es. Auto non puo passare per il lago anche se è la via più veloce per arrivare a B partendo dal punto A).
- 4) Kreativität
- 5) Menschenkenntniss → conoscenza della natura umana
- 6) Heuristiken
- 7) Evaluation:
überprüfung auf
 - funktionale Richtigkeit
 - Zweckerfüllung (Soddisfazione obiettivo)
 - Leistungsfähigkeit (Performance)
 - Benutzerfreundlichkeit/Bedienbarkeit (User-friendly, facilità e operabilità)Evaluation → kontinuierlich, nicht nur am Ende (da fare continuamente, non solo alla fine la valutazione)

Neutralität von Algorithmen

BIAS

- Vorurteile (Pregiudizi)
 - Verzerrungen
 - Vernachlässigung gesellschaftlicher Gruppen
- Arten von Bias (tipi di bias)
- Bewusster Bias: Person weiß, dass sie t.B vorurteile hat, beeinflusst ist, etc.
 - Unbewusster Bias: Bias ist Person selber nicht bewusst, z.B. Name oder Affinity Bias Implicit Association Test.

Übersetzung:

- Pregiudizio cosciente: la persona sa di avere dei pregiudizi, di essere influenzata, ecc.
- Pregiudizio inconscio: Pregiudizio di cui la persona non è consapevole, per esempio il nome o il pregiudizio di affinità Test di associazione implicita.

Algorithmischer Bias - Definition: Algorithmischer Bias beschreibt systematische und wiederholbare

Fehler in einem Computersystem, die unfaire Ergebnisse erzeugen, wie zum Beispiel die Privilegierung einer willkürlichen Gruppe von Benutzern gegenüber anderen.

Das soll vermieden werden.

Gründe für Algorithmischer Bias

- Algorithmen wider-spiegeln impliziten Werte der Menschen, die beim Erstellen beteiligt sind.
- Daten auf denen Algorithmus basiert (z.B. im recruiting, Personen über 60 sind vermietet)

Kapitel 9: Programmieren

Einführung

Programmieren = eine Idee in einen Programm formalisieren. Programm kann dann auf Tablet,

Handy oder Computer benutzt.

Programmieren ist nicht nur Code schreiben aber auch Problemlösung.

Programmiersprache

Beeinflusst:

- darstellung der lösung (visualizzazione della soluzione)

- computerausführung (esecuzione del computer)

Es ist wichtig, programmen nicht nur zu nutzen, aber auch zu verstehen.

Programmiersprachen

Erste Generation→Die ersten Programmen wurden auf Lochkarten programmiert.

Die ersten Programmen waren binärcodereich.

Zweite Generation:

Assemblersprache→Assembler→Maschinencode

Dritte Generation: Höhere Programmiersprache (bzw. Python, Java, C#, C++, usw.)

Generation der Programmiersprachen

- Höhere Programmiersprache

- assemblersprache

- maschinensprache

Höhere Programmiersprache→Compiler oder Interpreter→ maschinensprache

Definition Programmiersprache:

- Lexikalik: gültigen Zeichen (A-Z,a-z,0-9-usw.)

- Syntax: definiert korrekt Aufbau der Sätze.

- Semantik (il termine semantica formale riguarda i modelli matematici che definiscono formalmente

i linguaggi di programmazione o, più generalmente, la computazione stessa):

definiert Bedeutung semantisch korrekter Sätze.

z.B zero division (Error)

Programmiersprache Paradigmen

Paradigma bezeichnet eine fundamentale Denkungsart, einen Denkstil oder ein Denkmuster (modello di pensiero).

Programmierparadigmen:

- deklarative programmierung

- funktionale programmierung

- imperative programmierung

- Objekt-orientierte programmierung

Programmier-Paradigmen:

Deklarativ

- Was?'

- Beschreibung des gewünschten Endergebnisses

- z.B SQL, Oz.

Datenbank: speichert Informationen.

Programmier-Oaradigmen:

Objektorientiert

-Fokus auf Objekte:
Eigenschaften, Manipulation und Kommunikation von und zwischen Objekten
z.b Java, Smalltalk

Sequenzen und Variablen

Anweisung:

- Elementare Einheit im Programm
- steht für einen einzelnen Abarbeitungsschritt im Algorithmus
- Auch statement, kommando oder Befehl genannt.

Sequenz= Abfolge von Anweisungen

- sequenzielle Ausführung in der gegebenen Reihenfolge von oben nach unten
- Kein überspringen von Anweisungen

Was ist eine Variable?

- Speichert einfache Werte in Programmen
- Erlaubt auf den gespeicherten Wert zuzugreifen (permette di accedere al valore salvato)

Verwendung von Variablen (utilizzo delle variabili)

Initialisierung:

- Erstmaliges Speichern von einem Wert in der Variable.
- Zuweisung = assegnazione, es. Var shoppingCartItems=1

Gebrauch:

- Zugriff auf Variable, resp, auslesen des gespeicherten Wertes

Datentyp:

- Ganze Zahlen (int)
- Gleitkommazahlen (float)
- Zeichen → Einzelne Ziffern (es. '\$' oder '@', 'g') oder Zeichenkette: Wörter, Sätze) → strings (str)

-listen (list): Kollektion von Datenobjekten

-Wahrheitswert(-e) (bool): Entweder wahr oder falsch (boolean)

Datentyp:

- beschreibt Menge von Datenobjekten,
- Datenobjekte eines Typs haben alle die gleiche Struktur.
- auf Datenobjekten des gleichen Typs können die gleichen Operationen durchgeführt werden.

Fallunterscheidungen

Fallunterscheidung:

- ermöglicht alternative Abläufe
- Unterscheidung aufgrund einer Bedingung (Entscheidungsanweisung) → decisione basata su una condizione

Für verschiedene Bedingungen gibt es verschiedene Aktionen im Programm.

Schleifen

Schleifen = loops

Zähler-kontrollierte Schleifen:

- Wiederholungsanweisung mit vorgegebener Anzahl Iterationen. (Ripeti l'istruzione con il numero specificato di interazioni)

Pseudocode

Der Pseudocode ist ein Programmcode, der nicht zur maschinellen Interpretation, sondern lediglich

zur Veranschaulichung eines Paradigmas oder Algorithmus dient. Meistens ähnelt er höheren

Programmiersprachen, gemischt mit natürlicher Sprache und mathematischer Notation. Mit Pseudocode kann ein Programmablauf unabhängig von zugrunde liegender Technologie beschrieben

werden. Er ist damit oft kompakter und leichter verständlich als realer Programmcode.

Funktionen

Funktionen:

-wird ein Stück Code oft wiederholt, kann es in eine Funktion ausgelagert werden.

Dann wird jeweils diese Funktion aufgerufen.

-Unterprogramme

-Jede Funktion hat einen Namen

-einer Funktion können Werte übergeben werden, die man Parameter nennt.

-eine Funktion hat oft auch eine Rückgabewert (Ergebnis) → una funzione ha spesso anche un valore

di ritorno (risultato)

Blockenprogrammierung → programmare a blocchi (z.B Scratch)

Dove si scrive il codice si chiama Script

La Konsole è per Eingabe und Ausgabe (Shell)

Kapitel 10: KI und ML

Lernziele

-Sie sind in der Lage, verschiedene Definitionen des Begriffs Künstliche Intelligenz voreinander

abzugrenzen.

-Sie können die wichtigsten Meilensteine der künstlichen Intelligenz darlegen.

-Sie können den Turing-Test erläutern

-Sie können Risiken künstlicher Intelligenz darlegen und Beispiele dazu nennen.

Machinelles Lernen: Lernziele

-Begriff

-Welche Aufgaben sich für Maschinen Lernen eignen

-usw.

Was ist künstliche Intelligenz?

Artificial Intelligence is the field that studies the synthesis and analysis of computational agents that act intelligently.

Ansätze künstlicher Intelligenz

Fähigkeiten

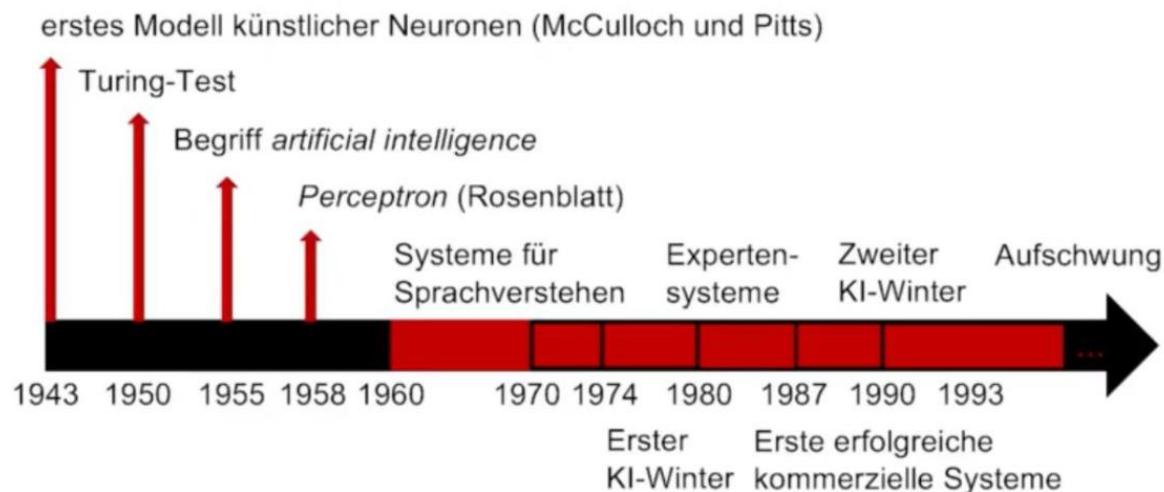
-Menschlich denken

-Rational denken

-Menschlich handeln

- Rational handeln
- Grundlagen der künstlichen Intelligenz
- Philosophie
- Mathematik
- Wirtschaft
- Neurowissenschaft
- Psychologie (→ wie denken Menschen?)
- Informatik
- Linguistik
- Arten von künstlicher Intelligenz
- Schwache KI (weak AI)
- Starke KI (strong AI)

Meilensteine



Teilbereiche (sotto sezioni):

- Automatische Sprachverarbeitung (Maschinelle Übersetzung, personelle Assistenten)
- Wissensrepräsentation
- Reasoning
- Maschinelles Lernen
- Computer Vision
- Robotik

Wissen, Schlussfolgern, Planen →conoscere, concludere, pianificare

Wissen in der KI

Wissensbasis (Knowledge Base) si suddivide in Fakten und Regeln.

Fakten

-Rob ist ein Rotkelchen

Rotkelchen ist eine Art Vogel

Vogel ist eine Art Tier

Tier ist eine Art Lebewesen

Vogel bewegt sich Fliegen

Fakten sind Aussagen (dichiarazioni/affermazioni)

Inferenzmaschine (Interface Engine)→ Robin ist ein Rotkelchen; Ein Rotkelchen ist ein Vogel, ein Tier

und ein Lebewesen.

?X ist eine Art ?Y

?X ist ein ?Y

Fliegen→Relativ Begriff (fly through air, fly through water, etc.)

Wo kommt das Wissen her? Linked Open Data

-Tabellen der D8pedia

-Geonames

-Datenbanken

-BBC Music

Fazit

Wissen und automatische Schlussfolgern sind grundlegende Bausteine der KI.-->la conoscenza e il

ragionamento automatico sono elementi fondamentali dell'intelligenza artificiale.

Sie kommen auch heute noch stark zum Einsatz. --> sono ancora oggi molto utilizzati.

Künstliche Intelligenz – Prototypische KI-Anwendungen →Prototipi di applicazione dell'AI

-Objekterkennung (riconoscimento oggetti)

-Image Captioning (non solo oggetti ma viene scritta una frase sotto a un'immagine→es. un ragazzo

vestito di nero suona la chitarra ! Può essere imperfetto (es. descrizione di donne invece di bambine,

ecc.)

-Medizinische Diagnostik (hinweise aus Krankheiten, usw.)

-Empfehlungssysteme

-Computational Creativity

Turing-Test

Turings «Imitation Game»

Ha collegato l'umano alla macchina→dialogo tra umano e la macchina.

Turing-Test und Dialogsysteme

-Operationale Definition von künstlicher Intelligenz

-Textbasiertes Dialogsystem

-Total Turing test: mit Computer Vision und Robotik

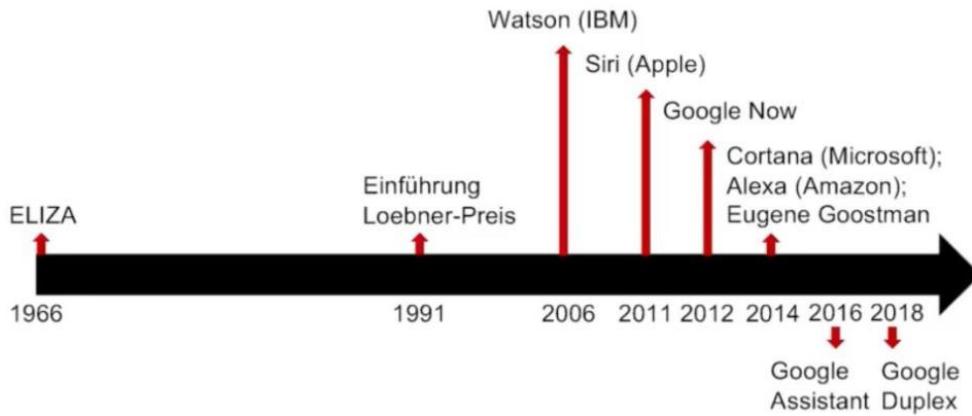
Dialogsysteme 1

Dialogsysteme sind die mit einem menschlichen benutzen eine Unterhaltung führen.

→I sistemi di dialogo sono quelli che hanno una conversazione con un umano.

Erste Dialogsysteme: ELIZA

Dialogsysteme: Geschichte



Dialogsysteme 2

Dialogsysteme Jill Watson:

Jill Watson ist eine Maschinelle Tutorin.

Herausforderungen

- Informationsgehalt der Antworten
- Specherkonsistenz
- Einbindung von externem Faktenwissen

Robotics

What is robotics?

Robots are automated machines, that can accomplish submissions.

Different Types of Robots

Aerial Robots (drones)

-Water Robots

-Ground Robots

-Manipulators (industry/robot impiegati nelle aziende per la produzione)

2001→First Aerial Robot: Automated Airplane

Environment→Perceive→Reason (Robot decides what to do according to its mission)→Act
(the
robot moves)→the cycle starts again.

How do Robots Perceive the World? As a set of numbers

Classic Robots

Perception-Action Cycle created with (lots of) if-else-conditions

Condition: if True→Do this

If False→Do that or do nothing

Modern Robotics

Perception-Action Cycle learned via interaction with the physical world

Take-Home Messages:

- The Perception-Action Cycle defines a robot's operation.
- There is a Perception-Gap between robots and humans
- Robots learn to act by Interaction with the World.

Auswirkungen der Künstlichen Intelligenz

Effetti/Conseguenze:

- Ethisch
- Psychologisch
- Sozial
- Ökonomisch
- rechtlich (in campo legale/giuridico)

Risiken im Zusammenhang mit künstlicher Intelligenz

- Verlust von Arbeitsplätzen
- Verwendung für unbeabsichtigte oder unerwünschte Zwecke.--> l'AI è a volte utilizzata per scopi non voluti o non desiderati (Rischio)
- Frage der Rechenschaft→questione di responsabilità
- Bias in maschinellem Lernen
- System als gross Menge von Daten.

'O bir doktor. O bir hemşire'→ He is a doctor. She is a nurse.

Was ist maschinelles Lernen?

Bedeutet ein Lern von grosse Datenmengen.-->è lo studio di grandi insiemi di dati.

Nachteile regelbasiertes Vorgehen→approccio basato su delle regole: svantaggi

- Anpassung einer Aufgabe auf neunen Kontext aufwändig→l'adattamento di un compito a un nuovo contesto richiede tempo
- Definition von Regeln manchmal schwierig→definizione delle regole a volte difficile

Zentrale Konzepte des maschinelles Lernens I

Daten

- Trainingsset (Training set)→grösste Teil des Daten
- Entwicklungsset (validation/development test)
- Testset (test/evaluation set)

Experience und Task

Experience E:

- Supervisiertes Lernen→ Labels (z.B. Themenklassifikationen)
- Unsupervisiertes Lernen→ keine Labels
- usw.

Task T (supervisiertes Lernen):

- Regression (Lineare Regression→Regressionstasks)
- Klassifikation
- usw.

Performance measure P

- Training: minimaler Verlust, wenn wir \hat{y} vorhersagen, während die echte Ausgabe y ist
- Verlustfunktion (*loss function*)
- Beispiel: *Mean squared error* (MSE) (mittlerer quadratischer Fehler)
$$MSE = \frac{1}{N} \sum_{i=1}^N (\hat{y}_i - y_i)^2$$
- Performanz des Lernalgorithmus auf ungesehenen Daten → Testset

Klassifikation

Lineare Regression: $\hat{y} = w^\top x + b$

Logistische Regression: $\hat{y} = \underline{\sigma(w^\top x + b)}$
mit $\sigma(x) = \frac{1}{1+e^{-x}}$

Unsupervisiertes Lernen:

-Keine Labels

-z.B. Clustering

Clustering

-Menge von Elementen in Cluster aufteilen, sodass Mitglieder eines Clusters untereinander ähnlich

und den Mitgliedern anderer Cluster unähnlich sind.

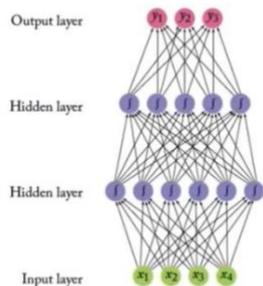
-Bsp.: Clustering von Dokumenten anhand von linguistischen Merkmalen → Esempio:

Clustering di

documenti basati su caratteristiche linguistiche.

Künstliche Intelligenz – Deep Learning und künstliche neuronale Netze (Deep Learning e reti neurali artificiali)

Neuronale Netze



$$NN(x) = W^3(g^2(W^2(g^1(W^1x+b^1))+b^2))$$

Abbildung: Multi-layer perceptron (Goldberg 2017, S. 42)

Neuronale Netze = reti di neuroni

Deep Learning ist ein Teil von Machine Learning

Maschinelles Lern ist ein Teil von AI.

Neuronale Netze: Herausforderungen

-Grosse Datenmengen benötigt (richiesti)

-Interpretierbarkeit

-Robustheit

Il Deep Learning, la cui traduzione letterale significa apprendimento profondo, è una sottocategoria

del Machine Learning (che letteralmente viene tradotto come apprendimento automatico) e indica

quella branca dell'Intelligenza Artificiale che fa riferimento agli algoritmi ispirati alla struttura e alla funzione del cervello chiamate reti neurali artificiali.

Künstliche Intelligenz - Abschluss

Künstliche Intelligenz

-Künstliche Intelligenz: Verstehen und Entwickeln von intelligent handelnden Maschinen

-Ansätze künstlicher Intelligenz:

-Menschlich denken

-Rational denken

-Menschlich handeln (Turing-Test) → agire in modo umano e con gli umani

-Rational handeln → agire in modo razionale

-Schwache KI vs starke KI → Begriff artificial intelligence 1955 im Sinne von schwacher KI geprägt

Handeln = agire

Teilbereiche künstlicher Intelligenz

-Automatische Sprachverarbeitung

Prototypische Anwendungen:

-Maschinelle Übersetzung

-Spracherkennung

-Image captioning

- Dialogsysteme
- Turing-Test: textbasiert
- Persönliche Assistenten: gesprochene Sprache
- Wissensrepräsentation
- Reasoning
- Maschinelles Lernen
- Computer vision

Prototypische Anwendungen:

- Objekterkennung
- Image captioning
- Robotik

Bestandteile maschinellen Lernens

Bestandteile = componenti

- Experience E: z.B. supervisert, unsupervisert
- Task T :z.B. Regression (lineare Regression), Klassifikation (logistische Regression), Clustering
- Performance measure P: z.B. mean square error

Neuronale Netze, Deep learning

-künstliche neuronale Netze: komplexe nicht-lineare Funktionen, die aus kleineren Einheiten zusammengesetzt sind → reti neurali artificiali: funzioni complesse non lineari composte da unità più piccole.

-Künstliche Intelligenz > maschinelles Lernen > Deep Learning (... ist ein Teil von ...)

Kapitel 12: Internet und Onlinemedien

Digitalisierung von Kommunikation

- Betrifft Text, Ton & Bild
- Vervielfältigung sowie Überbrückung von Distanz & Zeit
- Offenheit für Beitragende.

Funktionen von Kommunikation

- Kognitive Funktionen
- Affektive Funktionen
- Relationale Funktionen

Politische Funktionen von Kommunikation

- Politikvermittlung (Artikulation von politischen Anliegen)
- Kritik und Kontrolle

Digitalisierung verändert alle Formen medienvermittelter Kommunikation. Private und öffentliche Kommunikation erfüllt zentrale Funktionen für das Individuum, soziale Beziehungen und die Gesellschaft.

Was ist das Internet

Konstanten der Internet Geschichte:

- 1) Wachstum (quantitativ, qualitativ)
- 2) Unvorhersehbarkeit (Impredicibilità)

Digital Natives: aufwachsen mit digitalen Medien.

In letzten Jahren mehr mobile Internetnutzung als stationäre Internetnutzung.

Wellbeing

Allgemeine Lebenszufriedenheit, Abwesenheit negativer Emotionen.

Soziale Medien können zu Stress und Depression führen.

Onlinenutzung hat einen Einfluss auf unser Wohlbefinden (positiv und negativ).

Bewusste Online Nutzung & Selbstregulierung ist wichtig.

Soziale Netzwerke

Online Social Networking Platformen:

- Personen und ihre Beziehungen
- Kreation von Austausch von Medieninhalten
- Algoritmische Selektion

Relevanz für die Forschung:

Chancen:

- Neue Daten
- Granulare Daten
- Viele Daten

Risiken:

- Verzerrte daten (Nutzung, Bots, Plattformspezifiken)

Plattformen digitalisieren diese Beziehungen

Influencer

Wie gelingt es den Influencern, uns zu beeinflussen?

- Alltagssetting & Authentizität
- Thematische Expertise
- Ähnlichkeit & Nahbarkeit

Social Bots

Identifikation von Social Bots

- Anzahl Aktionen (Posts, Retweets etc.)
- Reaktionszeiten
- Tageszeiten
- Falsch interpretierte Mehrdeutigkeiten
- Netzwerkstruktur

Pfannkuchenpolizei
@PfannKPolizei
Berliner? Dit heißt hier Pfannkuchen!

Klaus Dieter @KlausDieter · 22. Juni 2018
Vorsicht bei Bluthochdruck.....Bericht des Sonderermittlers liegt vor
Scharfe Kritik an Zustand der Berliner Polizeiakademie rbb24.de/content/rbb/r2...

7:41 nachm. · 22. Juni 2018 · erased14445114

1 Retweet 1 „Gefällt mir“-Angabe

Social Bots versuchen menschliche Nutzung digitaler Plattformen vorzutäuschen

Einflussversuch richtet sich gegen Menschen, Algorithmen, und institutionelle Akteure.

Filterblase = personalisierte Informationsauswahl als Resultat einer algorithmischen Selektion.

Motive für die Erstellung von Fake News: finanziell, ideologisch.

Erlösquellen redaktioneller Medien

Verkaufserlös: Online deutlich tiefere Zahlungsbereitschaft

Anzeigenerlös:

- Geringere Nutzung
- Nutzung ohne Login
- unbundling

Nicht die Digitalisierung ist die entscheidende Veränderung der Kommunikation, sondern die Vernetzung.

Kapitel 13: Soziale Implikationen der digitalen Gesellschaft

Wissen und Zugang zur Information

Die Digitalisierung macht Wissen für alle zugänglich (1. te These). Die andere These ist das das Wissen nur für Privilegierte ist (Paywall / Bezahlshranken, Zensur, Sprachbarrieren). Die Digitalisierung verändert Konsumgewohnheiten.

Prosumption: Konsument trägt etwas zur Produktion oder der Erbringung der Dienstleistung bei (Fusion aus Production und Consumption).

Prosumption: Beispiele

Materiell (klassisch)	Digital
- Selbstbedienung im Restaurant	- Buchen von Reisen
- Selbstzusammenbauen von Möbeln	- Open source software (z.B. Linux)
- Self-checkout/- in Systeme	- Social Media oder Wikis (Wikipedia)

Diskursqualität

Öffentliche Debatte sind grundlegend für die Demokratie. Problemlösungen durch Austausch von Argumenten und Meinungen.

These 1: Die Digitalisierung fördert (promuove) die Diskursqualität

Menschen im Internet kommunizieren persönlicher und anonymisierter, auch kritischer.

These 2: Digitalisierung bedroht Diskursqualität (nur ein Teil von den Beiträgen hat viel Aufmerksamkeit, die anderen nicht). Fragmentierung und Inkongruenz.

Potential für ein breiteres und vielfältigeres Meinungsspektrum.

Politische Partizipation

1) Politische Willensbildung: Die Digitalisierung verbessert den Zugang zu Informationen.

Höhere Legitimität von Wahlen und politischen Entscheidungen. Zugang zu mehr

Informationen -> Gesellschaft informierter -> verbesserte politische Willensbildung (processo decisionale)

2) Polarisierung der Gesellschaft: Die Polarisierung ist das Verstärk des Meinungsunterschieden. Zugang zu mehr Informationen -> limitierte Aufnahmefähigkeit -> Polarisierung. Orientierungsschwierigkeiten aufgrund ausdifferenzierten medialen Angebots.

Probleme Informationen korrekt einzuordnen.

Die Digitalisierung kann die Informationsbasis verbessern. Die Menschen sind aber nur begrenzt aufnahmefähig.

Überwachung

These 1: Digitalisierung fördert Freiheit:

- Öffentliche Meinungsäußerung erleichtert

- Online Aktivismus (Veröffentlichung von Informationen, Online-Koordination von Protesten, Bildung von Gruppen)

These 2: Digitalisierung führt zu mehr Kontrolle:

- Möglichkeiten, Umfang und Dauer Daten zu erheben, zu speichern und zu verknüpfen sind durch die Digitalisierung angestiegen (Akteure: Staaten, Firmen, politische Parteien).

- Überwachung wird immer weniger greifbar (keine Einwilligung, Social Surveillance)

Einzelpersonen haben mehr Einflussmöglichkeiten. Alle haben mehr Möglichkeiten zu überwachen.

Ungleichheit

These 1: Digitalisierung produziert mehr Gleichheit:

- reduziert Informationskosten
- nivelliert soziale Unterschiede in Bezug auf Lebenschancen

These 2: Digitalisierung reproduziert bestehende Ungleichheit (Privilegierte Personen können vom Internet profitieren mehr als nichtprivilegierte).

Digitalen Ungleichheit: Unterschiede zwischen sozialen Einheiten (z.B. Individuen, Länder, Regionen) in Bezug auf Dimensionen der Informations- und Kommunikationstechnologie aufgrund sozio-ökonomischer Merkmale.

Nicht alle können gleich vom Internet profitieren.

Sicherheit

These 1: Digitalisierung schafft Schutz und Sicherheit:

- Verbesserte Sicherheit (crowdsourced surveillance)

These 2: Digitalisierung gefährdet Sicherheit und Schutz (Beispiel: Online-Mobbing)

- Illegale Aktivitäten
- Diskriminierung
- Digitale Vorverurteilung

Beziehungen

Soziales Kapital = Ressourcen, die der Teilhabe am Netz sozialer Beziehungen entspringen.

These 1: Digitalisierung schwächt soziale Beziehungen

- Verdrängung: "Kosmetik am Selbst anstatt Selbstoffenbarung" -> Einsamkeit
- Soziale Störfaktor
- Negative Beziehungspflege

These 2: Digitalisierung stärkt soziale Beziehungen

- Neue offline Bindungen
- Positive Pflege bestehender Beziehungen

Kapitel 14: Das digitale Ethik

Einführung

Das ABC der Digitalisierung:

A: Algorithmen

B: Big Data

C: Cybersicherheit

Themenfelder der digitalen Ethik:



Was ist ethik

Antwortet die Frage: was soll ich tun?

Zentrales Merkmal moralischer Sprache

Wir verwenden hier nicht Fakten beschreibende (deskriptiv-empirische), sondern normative und evaluative Sprache:

- «**normativ**»: Etwas soll oder muss getan oder darf nicht getan werden («Es ist verboten, Menschen zu foltern.»).
- «**evaluativ**»: Etwas ist gut oder schlecht im Hinblick auf bestimmte Werte («Er ist ein gieriger Mensch.»).

Ethik ist eine Reflexion über Moral.

Deskriptive Ethik: beobachtet was die Menschen tun

Die Hauptgruppen von Ethik-Theorien

- Deontologie: Lehre vom Sollen
- Konsequentialismus oder Teleologie: Ziel
- Tugendethiken: der Handelnde statt die Handlung

Deontologie: stellt die Frage ob Handlungen (azioni) in sich selbst richtig oder falsch sind. Fokus auf Pflichten und Regeln. Eine Handlung ist ethisch, wenn man diesen Pflichten folgt. Negative Pflichten sind meist wichtiger als positive Pflichten. Absichten der Akteure spielen eine wichtige Rolle.

Probleme: Anwendungsbedingung der Regel, Kontraintuitive Folgen

Konsequentialismus: stellt die Frage nach den Folgen des Handelns und unserer ethischen Überzeugungen. Fokus auf dem vormoralischen Gut (e.g. Glück). Eine Handlung ist ethisch, wenn dieses Gut maximiert wird. Fördern Ausweitung des moralischen Universums. Kein Unterschied Handeln-Unterlassen.

Probleme: Inhaltliche Bestimmung des Guts, Prognosehorizont.

Tugendethik: stellt die Frage nach den moralischen Kompetenzen und Eigenschaften von Personen. Fokus auf dem Selbstverständnis der Person. Moralische Wahrnehmungen sind wichtig. Fokus auf das gute Leben, weniger Überforderung.
Probleme: fehlende Eigenständigkeit, Anwendung oft diffus.

Ethisches Problemlösen

5-Schritte Verfahren:

- 1) Ist-zustand
- 2) Ethische Frage
- 3) Argumentation
- 4) Entscheidung
- 5) Implementierung

Ziele der ethischen Entscheidungsfindung:

- Strukturierter Umgang mit moralischem Dissenz
- Keine Abgabe der Verantwortung an "Ethik-Fachleuten"
- Kein Algorithmus, sondern eine Heuristik

1. Schritt. Analyse des Ist-Zustandes:

- Fakten auflisten
- Geltendes Recht berücksichtigen
- Stakeholder identifizieren
- Kontextsensibilität entwickeln
- Einbezug aller Wahrnehmungsperspektiven der beteiligten Fachleute.

2. Schritt: Die etische Frage benennen

- Moralisch relevante Aspekte identifizieren
- Strittige Frage formulieren
- Persönliche Werthaltung offenlegen
- Moralische und nicht-moralische Aspekte soweit möglich unterscheiden

3. Schritt: Analyse der Argumente

- Argumente pro und kontra identifizieren
- Moralische Normen und Werte rekonstruieren
- Argumente mit normativen Hintergrundtheorien abgleichen

4. Schritt: Evaluation und Entscheidung

- Mehrere Handlungsmöglichkeiten entwickeln, die aus einem "Entweder-oder" herausführen
- Standpunkt der Moral einnehmen
- Argumente beurteilen und gewichten
- Einseitigkeiten in der Argumentationsmustern identifizieren
- Güterabwägung vornehmen

5. Schritt: Implementierung der Lösung

- Möglichkeiten der Implementierung des Entscheids abschätzen

- Massnahmen zur erfolgreichen Implementierung ergreifen
- Kommunikative Aspekte beachten (wie wird der Entscheid wem mitgeteilt?)
- Allfällige Kriterien für Neubeurteilung identifizieren
- Prüfen, wie aus dem Entscheid gelernt werden kann

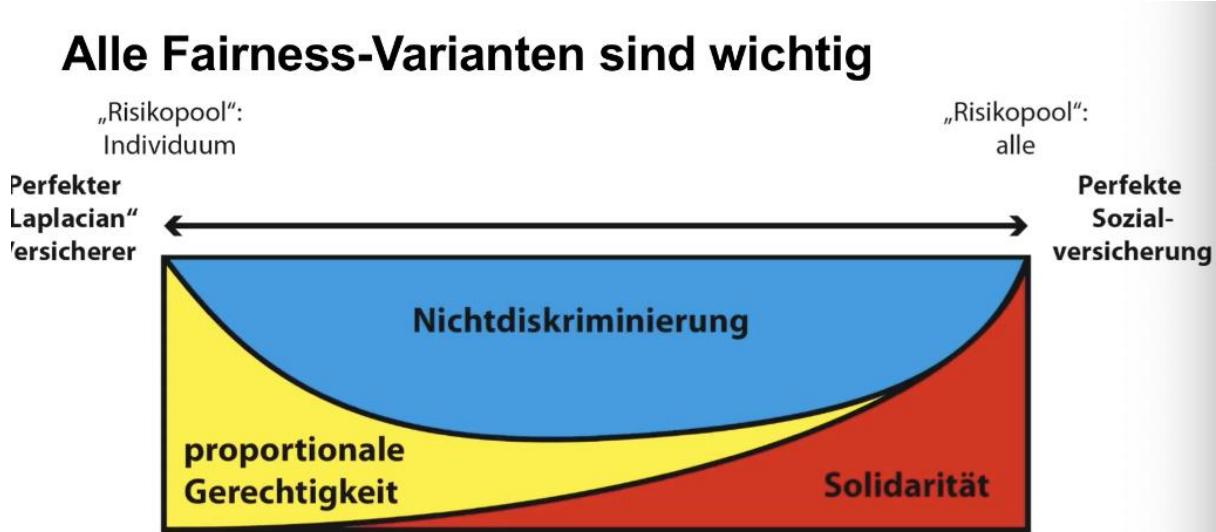
Ethik und Algorithmen

Eine der Fragen, welche die Anwendung von Algorithmen mit sich bringen, ist: sind deren Entscheidungen fair?

Drei Intuitionen von Fairness:

- Proportionale Gerechtigkeit (giustizia)
- Nichtdiskriminierung
- Solidarität

Die Konkretisierung von Fairness für Algorithmen führt zu Zielkonflikten.



Perfekte Laplacian Versicherer weiss genau was alles passiert und passieren wird.

Probleme:

- Korrelation ohne Kausalität
- Indirekte Diskriminierung
- Immer feinere Risikoklassen

Diskriminierung verhindern

Indirekte Diskriminierung: Eine an sich neutrale Massnahme führt faktisch zu einer Ungleichbehandlung von Angehörigen einer «geschützten Gruppe» (z.B. Alter, Geschlecht, Religion).

Gegenstrategie: Input (Feature-Vektor) verändern

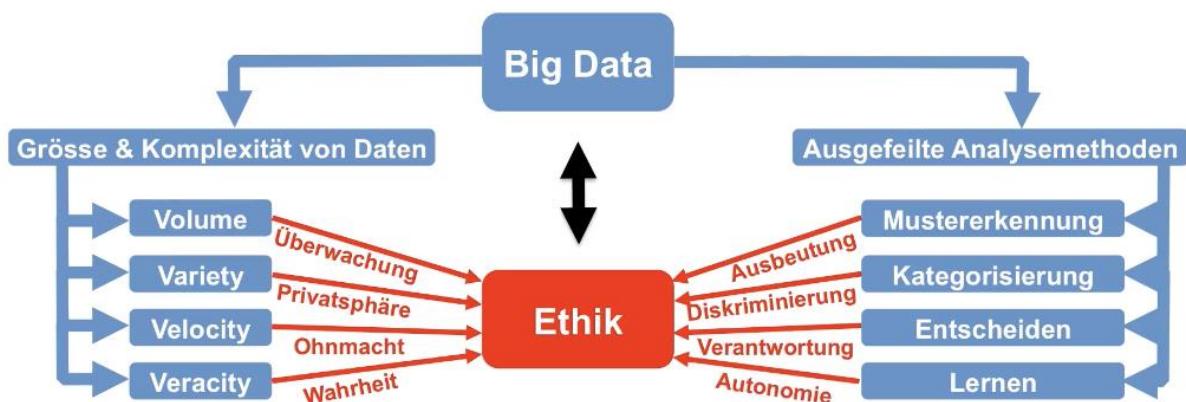
Bestimmte Features verbieten

$$X = (\mathbf{x}_1, x_2, x_3, x_4, \dots)$$

Korrelationen brechen (z.B. Randomisierung)

$$X = (\mathbf{x}_1, \mathbf{x}'_2, \mathbf{x}'_3, x_4, \dots)$$

Ethik und Big Data



8 Orientierungspunkte:

- 1) Privatsphäre: Ziel ist es, Lebensbereiche der Individuen zu schützen, in denen diese sich frei bewegen, entwickeln und verhalten können.
- 2) Fairness: Nichtdiskriminierung -> unter Diskriminierung wird in der Regel eine Ungleichbehandlung von Personen verstanden, die sachlich nicht gerechtfertigt ist. Diskriminierungsrelevante Kriterien sind unter anderem Ethnie (Rasse), Geschlecht, Religion, sexuelle Orientierung.

Informationelle Selbstbestimmung

Informationelle Selbstbestimmung ist das Recht des Einzelnen, selbst über das Erheben, Speichern, Verwenden und Weitergeben persönlicher Daten bestimmen zu können.

Kontrolle der digitalen Identität

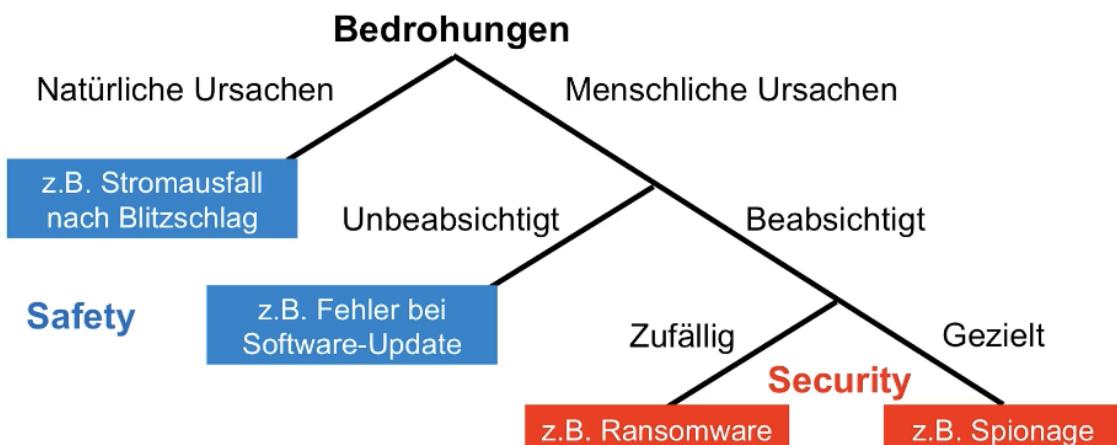
Digitale Identitäten können konstruiert werden, indem Unternehmen aufgrund von Big-Data-Anwendungen verschiedene Merkmale eines Kunden aggregieren, korrelieren und zu einer digitalen Identität verdichten.

Hier stellt sich die Frage der **Einsicht und Korrekturmöglichkeit** solcher Identitäten.

- 5) Transparenz: kann nie umfassend sein. Transparenz meint eine offene und relevante Kommunikation bezüglich der zur Debatte stehenden Entscheidungen.
- 6) Solidarität: Idee der gegenseitigen Hilfe. Sie bezeichnet generell die Verbundenheit der Einzelpersonen in einer Gemeinschaft. Sie werden in Bezug auf bestimmte Risiken und Widerfahrnisse nicht allein gelassen.
- 7) Kontextuelle Integrität: Die Interpretation moralischer Grundwerte und die damit verbundenen Regeln unterscheiden sich je nach sozialer Sphäre.
- 8) Eigentums und Urheberrechte: sind in der Schweiz von der Verfassung geschützte Grundrechte. Im Fall von Big Data stellt sich die Frage, inwieweit Daten ebenfalls unter diese Rechtsnormen fallen.

Ethik und Cybersicherheit (C vom ABC der Ethik)

Was ist Cybersicherheit?



Die CIA-Ziele der Cybersicherheit:

- 1) Confidentiality
- 2) Integrity
- 3) Availability

Warum ist Cybersicherheit schwierig?

Komplexität:

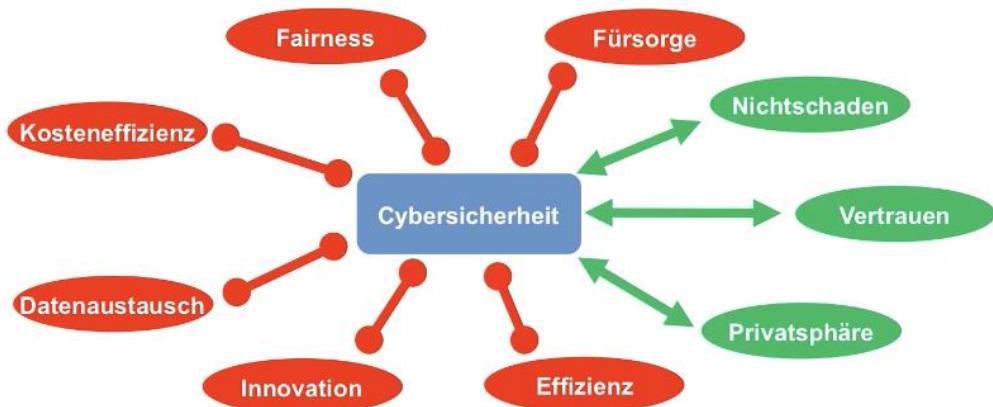
- Mehr Daten und Informationssysteme
- Zunehmend verschiedene Systeme
- Große Programme (Vulnerabilitäten)



Asymmetrie:

- Angreifer hat es einfacher als der Verteidiger

Werte im Spannungsfeld mit Cybersicherheit



Die Umsetzung von Cybersicherheit kann zu Zielkonflikten führen.

Zusammenhang Cybersicherheit – Prinzipien

Ebene 1: *Prinzipien der biomedizinischen Ethik*

Fürsorge

Respekt vor Autonomie

Gerechtigkeit

Nichtschaden

Ebene 2: *Desiderata gegenüber ICT*

Qualität und Effizienz

Datenschutz und Vertraulichkeit

Benutzerfreundlichkeit (Usability)

Sicherheit (Safety)

Ebene 3: Rolle der *Cybersicherheit*

Schutz persönlicher Daten

Schutz von Informationssystemen

Schutz von Medizinprodukten

Das Digitale Ethik ABC - Abschluss

Ethik allgemein

- Moralische **Sprache** ist normativ und/oder evaluativ.
- Moralische **Aussagen** haben in der Regel eine soziale Dimension, ein besonderes Gewicht und beanspruchen allgemeine Gültigkeit.
- Ethik ist **Reflexion** über Moral mit dem Ziel der Klärung und Begründung moralischer Aussagen.
- Ethik und **Recht** haben unterschiedliche Bezüge:
 - Recht kann ethische Normen kodifizieren
 - Ethik kann Anlass sein, das Recht zu ändern

Kapitel 15: Datenrecht

Einführung

Vengono spiegate le domande del diritto dei dati

Daten als Gegenstand des Rechts → dati come soggetti di diritto

“Datenrecht”

-Kein eigenes Rechtsgebiet

-Querschnittsmaterie

-Nicht «Digitalisierung und Recht»

Bei Datenrecht steht:

-Zuordnung von Daten (“Datenrecht”)

-Datenanalyse zur Verbrechensbekämpfung und -prävention (Strafrecht)

-Phänomene (Rechtsgebiet: area del diritto)

-Digitalisierung und Automatisierung der Verwaltung (= amministrazione) → Offenes Recht

Relevanz

- Datengetriebene Geschäftsmodelle → modelli di business basati sui dati
- Wirtschaftlicher Wert → valore economico
- Recht als wichtiger Faktor im Wettbewerb um Datennutzung

Recht und Gesellschaft

Grundlegende Elemente

1. Normative Ordnung (besteht aus Vorschriften → Regeln, Gesetze, usw.)
2. Durchsetzung (notfalls mit Zwang) → Applicazione (con coercizione se necessario)
3. Legitimation

Naturrecht vs Rechtpositivismus

Diritto positivo. comprende le norme che ogni cittadino deve rispettare all'interno di uno Stato,

evitando di incorrere in sanzioni; Diritto naturale: è dato dall'insieme di norme che fanno parte

della coscienza di un popolo e dei singoli individui: esse sono innate nella natura umana.

Esempio: il diritto alla vita, il diritto al rispetto della dignità umana, il diritto del rispetto della libertà

dell'individuo.

Funktion des Rechts?

«Die (zeitstabile) kontrafaktische Sicherung von Erwartungshaltungen»

Rechtsgebiete (Aree del diritto)

- Privatrecht
- Öffentliches Recht
- Strafrecht

Rechtsgebiet

Prozessgesetze

Privatrecht

Zivilprozessordnung (ZPO)

→ auf Bundesebene harmonisiert

Strafrecht

Strafprozessordnung (StPO)

→ auf Bundesebene harmonisiert

Öffentliches Recht

z.B. BG über das Verwaltungsverfahren (VwVG)

z.B. Verwaltungsrechtspflegegesetz des Kantons Zürich (VRG ZH)

→ föderalistische Struktur

Zivilgesetzbuches = codice civile svizzero

Sachen sind unpersönliche, körperliche, für sich bestehende Gegenstände, die der menschlichen

Herrschaft unterworfen werden können. (Von der Rechtswissenschaftlichen Lehre entwickelte definition). Sachen sind keine Daten

Grundbegriffe

Rechtssubjekte → persone fisiche e giuridiche

Objekte der Rechts → Sachen oder Materialien

Gesetz im materiellen Sinn → generell abstrakte Rechtsnormen, nicht jedes gesetz in materiellen sinn

ist auch ein Gesetz im formellen Sinn. Codice civile, penale, le leggi in sé.

Gesetz im formellen Sinn → si occupa di far rispettare le leggi materiali . Tutte le leggi che regolano la

procedura di applicazione e l'organizzazione delle autorità e dei tribunali. La parte formale: come

funziona un processo.

Objektives Recht → le norme contenute all'interno valgono per tutti.

Subjektive Rechte → riguarda individui in una situazione particolare z.B. Stimmrecht des Aktionars,

Proprietà.

Le leggi che regolano la proprietà rientrano nell'Objektives Recht.

Absolute Rechte → leggi che valgono sempre per tutti

Relative Rechte → leggi che valgono per alcune persone. Es se faccio contratto, le regole del mio

contratto sono relative perché riguardano solo quel contratto

Norm → bestimmung des objektiven rechts.

Anspruch → Berechtigung die sich eben aus der Norm ergibt.

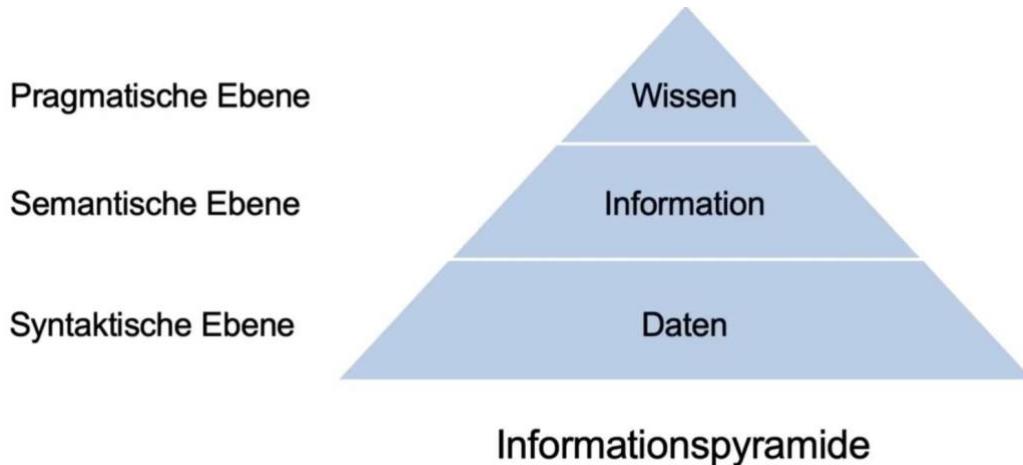
Tatbestand → è il delitto (quello che è successo)

Rechtsfolge → Conseguenze di una violazione del diritto (es. pena, sanzione)

Daten im Recht

Daten können ganze verschiedene Formen nehmen.

Mensch kann keine Information aus Daten auslesen



Nel secondo piano i dati sono sottoforma di informazioni.

Strukturelle Ebene → Non hanno a che fare con la piramide dell'informazione ma sono comunque importanti nel diritto:

-Usb-Sticks (livello più alto della piramide)

-Handyspeicher (livello centrale)

-Server (primo livello)

Daten unterscheiden sich in Sachdaten und Personendaten (→Datenschutzrecht)

Es gibt eben kein Gesetz in formell Sinn dass die Zuordnung von Daten als solchen regelt.

→ Non esiste una legge in senso formale che regoli l'assegnazione dei dati in quanto tale.

Grundsätze der Zuordnung von Daten → Principi di assegnazione dei dati

1) Je stärker die Daten einem Rechtsträger zugeordnet sind, desto schlechter ist das für die Gemeinfreiheit.

Je stärker Daten einer Person oder einer Unternehmen zugeordnet sind, desto weniger können sie von anderen genutzt werden.

→ Più dati sono attribuiti a una persona giuridica, peggio è per il dominio pubblico.

Più dati sono attribuiti a un individuo o a un'entità, meno possono essere usati da altri.

2) Diese Zuordnung von Daten kann auf unterschiedlichen Weisen erfolgen:

-Rechtliche Zuordnung: → Attribuzione legale

-Sachenrecht → Anspruch auf Wiedererlangung

-Patentrecht (diritto dei brevetti) → Unterlassungsanspruch

-Uhreberrecht (diritto legale) → Unterlassungsanspruch

- Uhreberrecht viene applicato per le geistige schöpfung, kreatives Denken. (NFT non è nell'Uhreberrecht, poiché è una technische schöpfung).

-faktischer Zuordnung (Rechtlicher Schutz): → non dà la possibilità di controllare i dati quando sono

via

-Strafrecht → Strafrechtliche Sanktion → il diritto penale viene applicato sulle cose, e non sui dati.

Privatrechtliche Sachen → Questioni di diritto privato

-Gesetz gegen den unlauteren Wettbewerb (= legge contro la concorrenza sleale) → Strafrechtliche

Sanktion, Schadenersatzanspruch (= richiesta di risarcimento danni)

I dati rimangono per sempre

-Es gibt kein Dateneigentum heute

Datenbestände des StGB

-Unbefugte Datenbeschaffung → Acquisizione di dati non autorizzata

-Hacking

-Datenbeschädigung → Corruzione dei dati

-Automatenbetrug → Frode automatizzata

Patentrecht → Absolute Rechte

Schutzgegenstand → oggetto che viene protetto

Erfahrung = Lehre zum technischen Handeln

Der Zweck von Patentrecht ist, Innovation zu fordern.

Schutzworaussetzungen → Requisiti per la protezione

-Neuheit → eine Erfahrung muss innovativ sein

-Nicht-Naheliegen (non-ovvietà)

-Gewerbliche Anwendbarkeit → Applicabilità commerciale

Schutzwirkung → effetto protettivo

(Absolutes) Recht zum Verbot der gewerbsmässigen Nutzung der Erfahrung.

→ diritto (assoluto) di vietare l'uso commerciale dell'invenzione.

Schutz von Daten:

-Nein

-Aber Daten können «mitgeschützt» sein, soweit sie auf der semantische Ebene die Schutzvoraussetzungen erfüllen.

I dati possono essere "co-protetti" nella misura in cui soddisfano i requisiti di protezione a livello semantico.

Schutz von Computerprogrammen:

-Programme, Handlungsanleitungen, etc. sind mangels Technizität vom Schutz ausgeschlossen. (sono esclusi per mancanza di tecnicità)

→non sono considerati come innovazioni/scoperte

-Computerimplementierte Erfindungen (Software erzielt einen technischen Effekt) können aber geschützt werden.

SBB Mobile App Daten → nicht geschützt von Patentrecht

Kein Patentschutz.

Urheberrecht (= diritti d'autore/copyright) → Absolute Rechte

Der Zweck von Uhreberrecht ist nicht Innovation zu fordern, sondern Kreation (kreative Leistungen zu belohnen).

Schutzgegenstand

Werke der Literatur, Kunst und Musik.

Schutzvoraussetzungen

-Geistige Schöpfung → creazione intellettuale

-individueller Charakter (es gibt keine Register, 'Recht entsteht einfach so')

Schutzwirkung

(Absolutes) Recht zum Verbot der Nutzung des Werks. --> Diritto assoluto di vietare l'uso dell'opera.

Oder kann ich es erlauben, mit Geld.

Schutz von Daten

Nein, aber Daten können «mitgeschützt» sein, soweit sie auf der semantische Ebene die Schutzvoraussetzungen erfüllen

Schutz von Computerprogrammen:

-computerprogrammen gelten als Werken, sie sind damit geschützt.

-Diese Fiktion schliesst grundsätzlich auch ein, dass angenommen wird, die Schutzvoraussetzungen

seien erfüllt. → si presume che i requisiti per la protezione siano soddisfatti.

Sbb Mobile App:

Urheberrechtlicher Schutz? Protezione del Copyright?

Software/Code ist urheberrechtlich geschützt.

Wettbewerbsrecht

→Gesetz gegen den unlauteren Wettbewerb (legge contro la concorrenza sleale)

Grundsatz

Nachahmungsfreiheit: wenn ein Unternehmen feststellt, dass ein Konkurrent etwas anders oder besser macht, dann steht dieses Unternehmen frei, das genau zu machen.

Prinzipio

Libertà di imitazione: se un'azienda scopre che un concorrente fa qualcosa di diverso o migliore,

allora quell'azienda è libera di imitarlo.

Ausnahme→Eccezione

-Immaterialgüterrechte (diritti sui beni materiali)

SBB Mobile App Daten:

Rechtlicher Schutz der faktischen Kontrolle über Daten im Falle von «screen scraping»?

Ja.

Con screen scraping si intende una tecnica in cui un programma informatico estrae dei dati dall'output

generato da un altro programma. Lo scraping si esplicita comunemente nel web scraping, che è il processo nel

quale una applicazione estrae informazioni di valore da un sito web.

Vertragsrecht (= diritto contrattuale) → beträgt nur Relative Rechte.

Verträge über (contratti su):

-(IP-)rechtlich zugeordnete Daten (dati assegnati legalmente (IP)):

-Übertragung→trasmissione

-Lizenzierung→licenza

-faktisch kontrollierte Daten.-->dati controllati fattualmente (= realmente/effettivamente)

-Vertragsgegenstand→soggetto contrattuale

-erlaubte Nutzungen→usì consentiti

-Nutzungsdauer→periodo di utilizzo

-Geheimhaltung→segretezza/riservatezza dei dati

-Folgen von Pflichtverletzungen→conseguenze delle violazioni del dovere

Kapitel 16: Datenschutzrecht

Einführung

Datenschutzrecht ist ein Teil des Datenrechts.

Datenschutzrecht schützt nicht Daten!

Datenschutzrecht geht nicht um den Schutz der Daten.

Kontext

-Teil des Datenrecht

-Regelt Datenbearbeitung durch Private (→Privatrecht) aber auch Datenbearbeitung durch in

Öffentliches Hand (→Öffentliches Recht).

-Ursprünge

-Andere Ansätze (= altri approci).

Die Globalisierung sorgt dafür dass diese verschiedene Ansätze heute auseinander setzen→ es gibt

Konflikte und Gegenseiten

Geltungsbereich (= campo di applicazione)

Dieses Gesetz gilt für das Bearbeiten von Personendaten durch (da parte di):

- Private Personen (und)
- Bundesorgane (corpi federali)

Geltungsbereich

-Personendaten

-Bearbeiten

-Akteuren

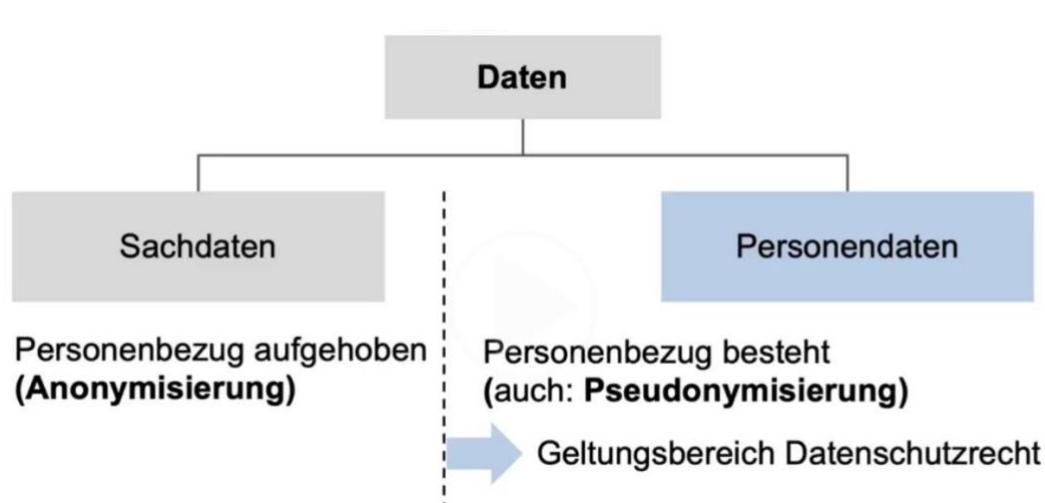
Personendaten: alle Angaben, die sich auf eine bestimmte oder bestimmbare natürliche Person beziehen.--> qualsiasi informazione relativa a una persona fisica identificata o identificabile.

Bestimmbarkeit: Die Identifikation einer Person ist zwar nicht aufgrund der Information allein, aber

aufgrund weiterer vorhandener Informationen (ohne unverhältnismässigen Aufwand) möglich.

Begriff von Personen reicht ziemlich weit→il concetto di persona è piuttosto ampio.

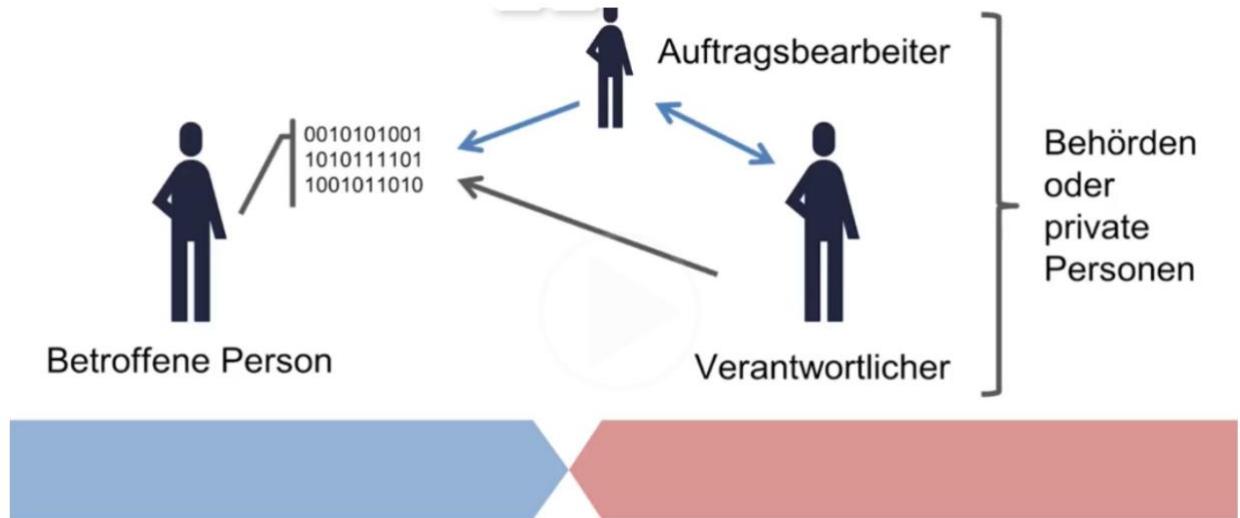
I dati si dividono in Sachdaten e Personendaten.



Personendaten si suddividono ancora in:

- normale
- Besonders schützenwerte Personendaten→Dati personali che richiedono una protezione speciale

Akteuren:



Betroffene Person = persona interessata

Ziele und Zwecke

Es geht nicht um den Schutz den Daten.

Zwecke:

-Schutz der Persönlichkeit (Privatrecht, Zivilgesetzbuch (ZGB)). IL ZGB regola il diritto delle persone

fisiche ma anche di quelle giuridiche.

-Schutz der Grundrechte der betroffenen Personen (BV→Bundesverfassung)

BV = costituzione federale, è il livello più importante del diritto svizzero.

OR = Obligationen Recht →bzw. Austauschverhältnisse

Der Zweck im Datenschutzrecht ist relativ unklar.

→Nachteil bei der Rechtsanwendung (sia in CH che EU)→Svantaggio nell'applicazione della legge.

Rechtsquellen (fonti di diritto)

Internationale Ebene: UN (Leitlinien), OECD (Privacy Guidelines), Europarat (Konvention 108).

Anwendbarkeit: Nicht unmittelbar anwendbar. Mitgliedstaaten z.T. aber zur Umsetzung verpflichtet.

Regional:

EU→DSGVO (gilt nicht für Schweiz).

Anwendbarkeit: Extraterritoriale Wirkung

National:

CH→DSG

Anwendbarkeit: Gilt für Behörden (= autorità pubbliche) und Private

Kantonal:

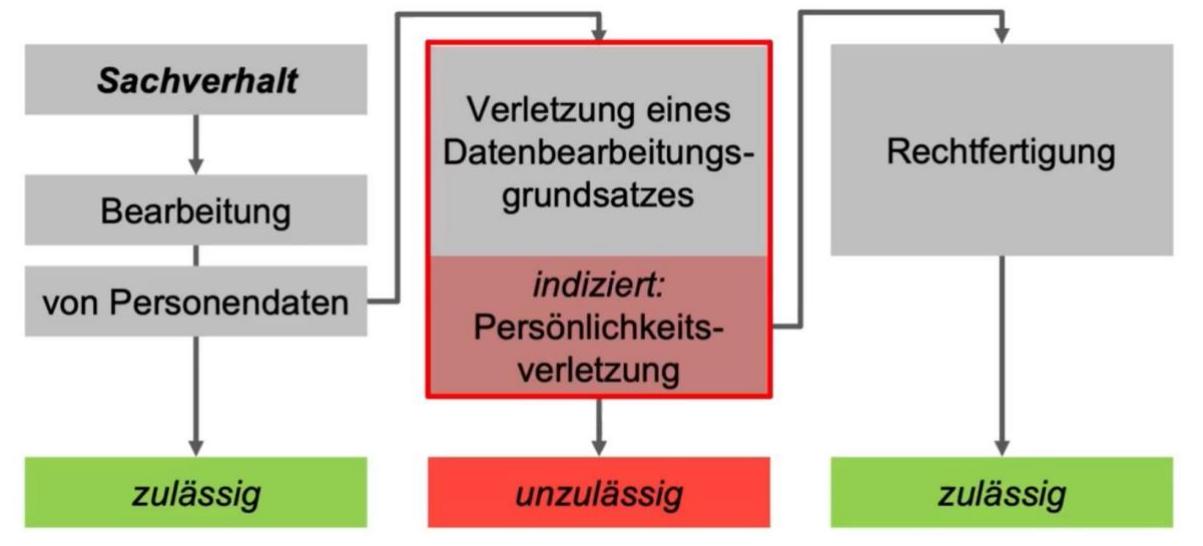
Kanton→Kantonale Datenschutz Gesetz →legge cantonale sulla protezione dei dati

Anwendbarkeit: Gilt für Kantonale Behörden

Funktionsweise (Funzionalità / come funziona)

SBB Mobile App Beispiel:

Sachverhalt (situazione): Personalisierte Fahrplanauskunft mit Handy-Positionsdaten.



Zulässig = permesso/consentito

Rechtfertigung = giustificazione

Erkenntnisse:

1. Grundsätze und Rechtfertigung
2. Präventiver Charakter
3. Untypische Regelungstechnik

Datenbearbeitungsgrundsätze (principi di trattamento dei dati)

1) Verhältnismässigkeit (proporzionalità):

-Beinhaltet Erforderlichkeit: Ist eine bestimmte Datenbearbeitung tatsächlich notwendig?

-Datenminimierung: Vernichtung, wenn nicht mehr erforderlich → i dati vengono distrutti quando non sono più necessari.

-Keine Datensammlung «auf Vorrat»

2) Zweckbindung → Limitazione dello scopo

-Datenbearbeitung nur zu dem Zweck, der zum Zeitpunkt der Erhebung angegeben wurde → -

Elaborazione dei dati solo per lo scopo dichiarato al momento della raccolta.

-Unzulässigkeit der Datenbearbeitung falls sich der Zweck im Laufe der Zeit geändert hat → - Inammissibilità del trattamento dei dati se lo scopo è cambiato nel corso del tempo.

-Keine ziel- oder zwecklose Datensammlung

Rechtfertigungsgründe (motivi di giustificazione)

Rechtsfertigungsgründe

-Einwilligung → consenso

-Gesetzliche Grundlage → base giuridica

-Überwiegende Interessen

Einwilligung

-Gültigkeit (mit angemessener Information, freiwillig, eindeutig) → validità

-Grundsätzlich stillschweigend möglich; bei besonders schützenwerten Personendaten zwingend ausdrücklich.--> Fondamentalmente tacitamente possibile; nel caso di dati personali che richiedono una protezione speciale, il consenso esplicito è obbligatorio.

-Vor der Datenbearbeitung und ohne Widerruf

Gesetzliche Grundlage

Non applicabile all'esempio dell'SBB

Überwiegendes Interesse

-Überwiegende private oder öffentliche Interessen

-Einzelfallabwägung

Privacy-Paradox Phänomen→ Il paradosso della privacy è l'osservazione che le persone condividono informazioni personali mentre sono molto preoccupate per la loro privacy.

Rechtsdurchsetzung (applicazione della legge)

Rechtsdurchsetzung: durch private und durch den Staat (hoheitliche Rechtsdurchsetzung→applicazione della legge sovrana)

Private: Können gegen private oder gegen Behörden vorgehen.-->può agire contro i privati o contro le autorità pubbliche.

Auch Behörden können gegen private oder andere Behörden vorgehen.

Private gegen private:

-Unterlassungs-, Beseitigungs- & Feststellungsklage

-Schadensatz, Genugtuung, Herausgabe des Gewinns.-->danni, soddisfazione, restituzione dei profitti.

-Vernichtung→distruzione

-Verbot der Bekanntgabe an Dritte→proibizione di divulgazione a terzi

-Berichtigungsrecht

-Auskunftsrecht (diritto all'informazione)→ wird oft unterschätzt

-Portabilitätsrecht

Hoheitliche gegen Private:

-EDÖB11 kann:

-eine Untersuchung anordnen (ordinare un'indagine)

-die Anpassung, Unterbruch oder Abbruch der Bearbeitung sowie Löschung verfügen

-Kantonale Strafverfolgungsbehörden können Datenschutzverletzungen verfolgen und zur Anklage bringen.

→Le autorità di polizia cantonali possono perseguire le violazioni della protezione dei dati e sporgere denuncia.

-Es drohen Bussen bis zu 250'000 CHF

Datenschutzrecht als Zurodungsinstrument (Il diritto sulla protezione die dati come metodo di classificazione)

Für eine Rechtlichezuordnung (a favore di un'assegnazione legale):

- Verschiedene Elemente bewirken eine Art Zuordnung (diversi elementi effettuano una sorta di classificazione):
 - Auskunftsrecht→diritto di accesso
 - Datenportabilitätsrecht→diritto alla portabilità dei dati
 - Einwilligung→consenso
 - Widerrufsrecht→diritto di recesso
 - Widerspruchsrecht (die Betroffene Person hat ein Widerspruchsrecht)→ diritto di opporsi Gegen eine Rechtlichezuordnung:
 - Zuordnung erfolgt nur zur natürlichen Personen (Personendaten von Juristischen Personen sind nicht geschützt)→ Attribuzione solo alle persone fisiche (i dati personali delle persone giuridiche non sono protetti).
 - Nur auf «eigene» Personendaten beschränkt→ limitato solo ai propri dati personali
 - Kein Recht zur Übertragung→nessun diritto di trasferimento
- Datenschutzrecht ist nicht nur ein Zuordnungsinstrument, sondern auch als Zugangsinstrument.
- La legge sulla protezione dei dati non è solo uno strumento di attribuzione, ma anche uno strumento di accesso.
- Fazit Datenschutzrecht:
- weiter Anwendungsbereich→ampio campo di applicazione
 - relativ unscharfer Zweck→scopo relativamente vago
 - präventive Wirkung→effetto preventivo
 - es handelt um ein ziemliches undifferenziertes Rechtsgebiet.--> è un area del diritto abbastanza indifferenziata.

Kapitel 17a: Daten organisieren, teilen, und nachnutzen

Open Data und Datenmanagement

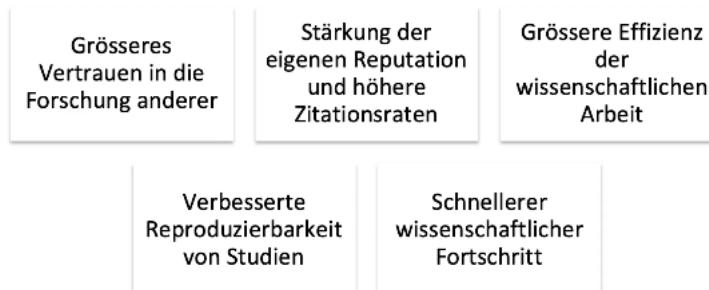
Open data: Daten können frei verwendet, verändert und weitergegeben werden.

Forschungsfragen von Datenmanagement:

- 1) Weshalb sollten Daten systematisch organisiert werden?
- 2) Welche Vorteile entstehen durch Datenmanagement?

Motivation für Datenmanagement: durch technologischen Fortschritt wachsen Datenmengen immer stärker. Nur transparent Wissenschaft ermöglicht Wortschrift. Nachlässiger Umgang mit Daten ist teuer.

Vorteile von Open Data und Datenmanagement

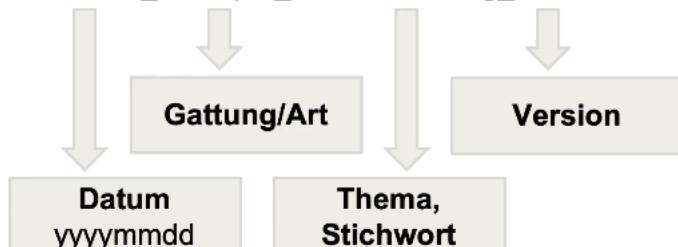


Datenorganisation

Name Convention: convenzione di come si nomina i dati

Namenskonvention für Dateinamen

20200328_Exzerpte_Parteienbildung_V0-2.docx



_ ist ein Trenner, - ein Verbinde

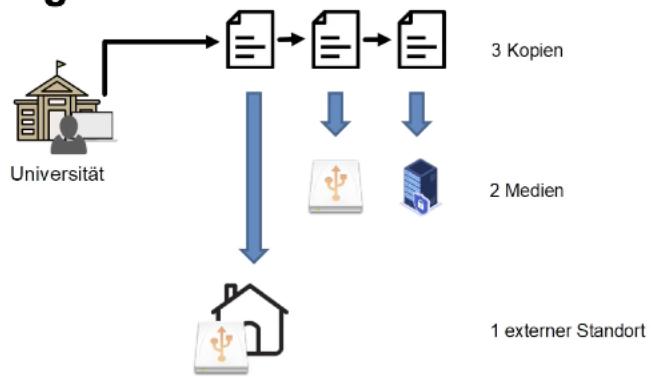
Zeichenraum

- a-z und A-Z
- 0-9
- Bindestrich -
- Unterstrich _

Datensicherheit

Datensicherheit: Massnahmen die den Datenverlust schützen (Z.B: Backups).

3-2-1 Regel



Zu vermeiden:

- Password teilen
- Password notieren
- Logisches Passwort (z.b. hello2024 kann sehr einfach gehackt werden)
- Nur ein Passwort
- Passwort an Angreifern weitergeben

Sicherere Passwörter

- Zeichenraum: Gross- und Kleinbuchstaben, Zahlen und Sonderzeichen: ~!@#\$%^&*()_+-;:"{}[]?/,.<>`'
- Mindestlänge: 8 Zeichen, besser 12
- Sinnfreie Passwörter sind sicherer: &g37*%O2Sik>
- Passphrasen z. B. aus Anfangsbuchstaben jedes Wortes:
 - *Das Jahr 39 solltest Du besser ausklammern, das war keine Sternstunde!*
 - *DJ39sDb(),dwk*s!*

Nachnutzen von Daten

Daten Repositories: erlauben das Speichern von Daten, suchen/finden von Daten

Arten von Repositories: Fachlich, institutionell, generisch.

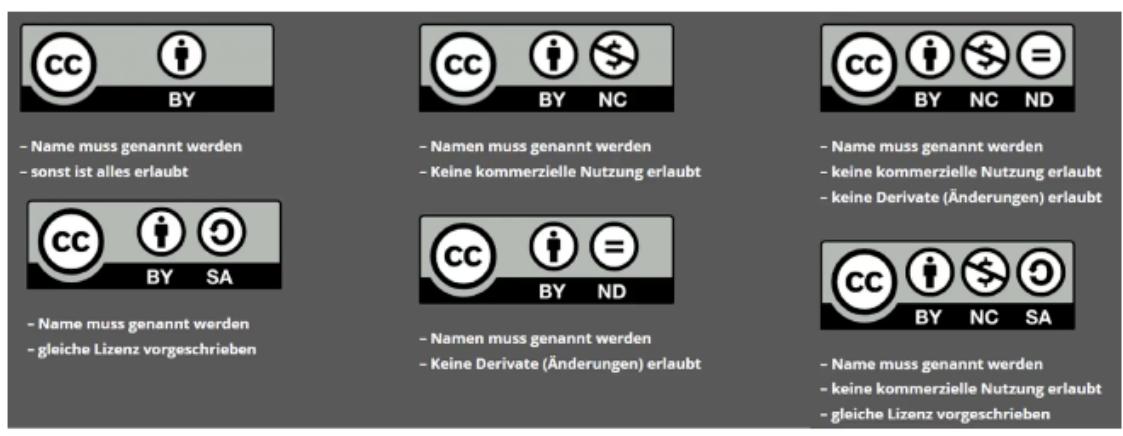
Nur mit einer Lizenz kann man Daten nachnutzen.

Arten von Lizenzen: generelle lizenzen, lizenzen für software.

Beispiel von generellen Lizenzen:

- Creative Commons: Namensnennung, nicht kommerziell, keine Bearbeitung, Weitergabe unter gleichen Bedingungen

Creative Commons



Wie darf das Bild verwendet werden?



- Für eine private Geburtstagseinladung ohne Weiteres?
- Für einen Geburtstagsgruss, der an 1000 Personen versendet wird
– solange der Urheber genannt wird?
- Für den Flyer zur Eröffnung eines Cafés – solange der Urheber
genannt wird?
- Sofern das Bild stark verändert wird, ohne Weiteres?

= bedeutet dass die daten gleich bleiben müssen.
omino: muss zitieren.

Kapitel 17b: Digitale Kollaboration

Awareness

Gründe für Digitale Zusammenarbeit: Verfügbarkeit, Skalierbarkeit, Durchsuchbarkeit. Man kann gemeinsam, gleichzeitig und überall an einem Produkt arbeiten.

Kapitel 17c: E-Accessibility

Was ist barrierefreiheit

Was bedeutet e-Accessibility

E-Accessibility (digitale Barrierefreiheit) ist die englische Bezeichnung für den Grad, in dem ein [IT-]System einem möglichst grossen Benutzerkreis zugänglich ist.

Ein IT-System ist barrierefrei, wenn es von allen Nutzern unabhängig von ihren Einschränkungen oder technischen Möglichkeiten uneingeschränkt genutzt werden kann.

Warum e-Accessibility:

- Gesetzliche Verpflichtung für Bund, Kantone und Gemeinden
- Unterstützt die gesellschaftliche und soziale Inklusion für Menschen mit Behinderungen, andere Nutzergruppen und ältere Menschen.
- positives Image

E-Accessibility wurde 2016 als Menschenrecht erkannt.

Barrieren und Technologien

Visuelle Barrieren: keine Alternativen für visuelle Informationen.

Additive Barrieren: keine Alternative für gesprochene Informationen.

Motorische Barrieren: schlecht Tastaturbedienbarkeit, zu kleine Elemente, schlecht strukturierte Inhalte.

Kognitive Barrieren: keine Informationen in leichter Sprache, keine nutzerorientierte Darstellung.

Welche assistierenden Technologien existieren?

- Mundstab
- Zoomtext (Bildschirmlupe)
- Augensteuerung
- Mundmaus
- Braillezeile

Screenreader

Ein Screenreader wandelt Inhalt des Bildschirms in akustische Sprache um. Alternativ kann die Software den Inhalt auf einer Braillezeile ausgeben.

Web Content Accessibility Guidelines (WCAG)

- sind ein internationaler Standard des World Wide Web Consortiums (W3C) zur barrierefreien Gestaltung von Webinhalten
- fördern einheitliches Verständnis von digitaler Barrierefreiheit
- Erfüllung der WCAG stellt sicher, dass digitale Inhalte für alle nutzbar sind

Wie funktionieren die WCAG?

Die WCAG beschreiben welche Anforderungen e-Content im Hinblick auf seine Gestaltung erfüllen muss. Sie sind technologieunabhängig formuliert.

Die WCAG legen lediglich fest, was ein Element erfüllen muss um barrierefrei zu sein, es gibt aber verschiedene mögliche Wege zum Ziel.

WCAG:

- 4 Prinzipien:
 - Wahrnehmbarkeit
 - Bedienbarkeit
 - Verständlichkeit
 - Robustheit

Interview mit Markus Riesch

Diskriminierungsverbot (gilt für Website, private und öffentliche Unternehmen). Man darf keine extra Hürden einbauen.

Bildung muss auch barrierefrei sein (Unis müssen das physisch und digital erlauben). Gesellschaftliche Hürden müssen abgebaut werden, damit die Inklusion ermöglicht werden kann.

PDF barrierefrei machen: Semantik ist wichtig.

PDFs

Wichtigste Kriterien barrierarmer PDFs

- Überschriften sind definiert (Formatvorlage)
- Aufzählungen sind als Liste formatiert
- Bedeutungstragende Grafiken haben einen Alternativtext
- Aussagekräftige Links
- Das Dokument hat eine korrekte Sprachzuweisung
- Das Dokument wird korrekt exportiert