CONSULTANT DATA SCIENCE

DATENSCHUTZ...

Eva Schabedoth 12.8.-22.8.

KLEINE HISTORIE

Bundesdatenschutzgesetz (BDSG 1977) — diverse Novellierungen bis 2009

Europäische Datenschutzrichtlinie (1995) — Festlegung von Mindeststandards

Europäische Datenschutzgrundverordnung (2016) — nach langen Kontroversen als verbindlich für alle verabschiedet

Europäisches Datenschutz- und Anpassungsgesetz (2017)

Neufassung des BDGS (2017)

Europäische Datenschutzgrundverordnung (DSGVO 2018) — Gültigkeit für alle Mitglieder; höherrangig als nationale Verordnungen

→ **Massive Interventionen speziell der USA**

→ **Kritik in Deutschland**

geht angesichts von Facebook und Co nicht weit genug

vs.

schränkt moderne, innovative digitale Geschäftsfelder zu sehr ein

ZUSTÄNDIGKEITEN

Recht auf informationelle Selbstbestimmung

Grundrecht laut BVerfG

Bundesdatenschutzgesetz (BDSG)

Landesdatenschutzgesetze

Bundesbeauftragter für den Datenschutz

↓

Behörden, Post- und Telekommunikation

Landesdatenschutzbeauftragter

↓

Unternehmen, privater Sektor

Neu im DSGV0:

bei international operierenden/außerhalb der EU ansässigen Unternehmen grundsätzlich das EU-Land, über dessen Bürger Daten gesammelt/verarbeitet werden; konkrete behördliche Zuständigkeiten allerdings unklar

erhebliche Bußgelder bei Verstößen (bis zu 4% des Vorjahresumsatzes)

GRUNDLEGENDE PRINZIPIEN

Datenschutz:

Erhebung, Verarbeitung und Nutzung personenbezogener Daten

auch:
Beschaffen

auch:
Speichern/Löschen

Verbot mit Erlaubnisvorbehalt → **keine Verarbeitung ohne Rechtsgrundlage**

Datensparsamkeit, Datenvermeidung

Erforderlichkeit

Zweckbindung

technisch-organisatorische Maßnahmen zur Sicherstellung erforderlich

formuliert im

↓

Standarddatenschutzmodell (SDM)

GRUNDLEGENDE KONFLIKTE

Datenschutz und Kriminalitätsbekämpfung

Sicherheit vs. Freiheits- und Bürgerrechte

Datenschutz und Informationsfreiheit

Berechtigtes Interesse des Bürgers z.B. an Transparenz des behördlichen/politischen Apparats

Datenschutz und Ökonomie/Unternehmen

„Freier Markt“, Risikoabschätzung, Kundengewinnung etc.

Datenschutz und Wissenschaft

Forschungsfreiheit → **Sonderregelungen**

Datenschutz und Gesundheitswesen

Vertraulichkeit vs. Optimierung

HERAUSFORDERUNG INTERNET

Globale Vernetzung

Standorte/Unternehmen ggf. außerhalb des Geltungsbereichs nationaler/europäischer Datenschutzbestimmungen

Digitale Ökonomie: Wandel -> neue Geschäftsfelder

Wandel der Begriffe „Öffentlichkeit“ und „Privatsphäre“

Verwandlung / Verwischen der Rollen „Produzent“ und „Konsument“

Ideologie der vollständigen Transparenz -> (demokratische) Freiheit

naive Vorstellung der Selbstregulierung

DATENSCHUTZGRUNDVERORDNUNG

Ziel:

Der Bürger soll die Hoheit über seine Daten so weit wie möglich behalten/zurückerhalten.

„*Personenbezogene Daten* [sind] alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person (...) beziehen; als identifizierbar wird eine natürliche Person angesehen, die direkt oder indirekt, insbesondere mittels Zuordnung zu einer Kennung wie einem Namen, zu einer Kennnummer, zu Standortdaten, zu einer Online-Kennung oder zu einem oder mehreren besonderen Merkmalen, die Ausdruck der physischen, physiologischen, genetischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität dieser natürlichen Person sind, identifiziert werden kann.“ (DSGVO, Art.4)

Personenbezogene Daten sind nicht nur Name, Anschrift oder Bestelldaten aus Shops. Google Analytics, Kontaktformulare, Newsletter-Daten, IP-Adressen aus Server-Statistiken, Plugins, Facebook-Like-Button usw. - überall geht es um personenbezogene Daten.

Zentrale Regelungen

Rechtmäßigkeit, Verarbeitung nach Treu und Glauben, Transparenz

Nachvollziehbarkeit — der Prozess der Verarbeitung
der Zusammenhang (und damit auch der Grund der Verarbeitung sowie Zeitpunkt und Grund der Übermittlung an Dritte)

Zweckbindung — das Warum und wofür der Erhebung/Verarbeitung

Marktortprinzip

Einwilligungspflicht

Recht auf Auskunft

Recht auf Löschung / Pflicht zur Löschung

Pflicht Datenschutzbeauftragter

außer bei Kleinunternehmen (bzw. < 10 Personen mit Datentätigkeit)

Die wichtigsten Regelungen auf einen Blick

Pflicht zur Führung eines Verzeichnisses aller Daten-/Auftragsverarbeitungstätigkeiten **Vorlagepflicht** →

Neue Vorgaben für Einwilligungserklärungen online und offline →

Erweitere Vorgaben für Datenschutzerklärungen auf Webseiten

Pflicht zur Datenportabilität

„Recht auf Vergessenwerden“

Pflicht zur Meldung von Datenpannen

Privacy by Design / Default

Anforderung an Konzept / technische Planung, datenschutzfreundliche Architektur

Gratwanderung zwischen Nutzerschutz und Nutzerrecht, seine Daten freiwillig anderen zu überlassen

Regelung für die Auftragverarbeitung

"Erhebung, Verarbeitung oder Nutzung personenbezogener Daten durch einen Auftragnehmer (natürliche oder juristische Person, Behörde, Einrichtung oder andere Stelle), der die Daten im Auftrag des Verantwortlichen verarbeitet)."

z.B.

- **Einsatz eines externen Kundencenters (z.B. Callcenter)**
- **externer Newsletter-Anbieter**
- **Cloud Computing**
- **Einsatz externer Unternehmen beim Marketing**
- **Externes Rechenzentrum**

Nicht nur der Auftraggeber, sondern auch der Auftragnehmer muss Tätigkeiten dokumentieren und ist verantwortlich für technische/organisatorische Maßnahmen zum Datenschutz!

Wer muss ein Verfahrensverzeichnis führen?

Unternehmen, die

- **mehr als 250 Mitarbeiter beschäftigen,**
- **besonders sensible Daten verarbeiten,**
- **bei denen die Verarbeitung ein Risiko für die Rechte und Freiheiten der betroffenen Personen birgt (Videoüberwachung und Ähnliches),**
- **personenbezogene Daten über strafrechtliche Verurteilungen und Straftaten verarbeiten oder**
- ***bei denen die Datenverarbeitung nicht nur gelegentlich erfolgt***

Mindestinhalt eines Verfahrensverzeichnisses

- Name und Kontaktdaten des Unternehmens
- der Zwecke der Datenverarbeitung
- die Kategorien betroffener Personen
- die Kategorien personenbezogener Daten
- die Kategorien von Empfängern der Daten
- die Übermittlungen personenbezogener Daten in ein Drittland
- Fristen für die Löschung der verschiedenen Datenkategorien

Vorgaben für Einwilligungserklärungen online und offline

OPT-IN statt OPT-OUT

Gebot der Freiwilligkeit / Kopplungsverbot

Erklärung darf prinzipiell nicht die Voraussetzung etwa für Vertragerfüllung (Ausnahme erfordert Begründung) oder Downloads sein.

Zweckgebundenheit

Generaleinwilligungen nicht statthaft

Nachweispflicht

dass die Einwilligung vorliegt

Widerrufsrecht

Widerruf so einfach wie die Einwilligung

KRITIK

nicht weitreichend genug in bezug auf große globale Datensammelunternehmen

der kommerzielle Datenmissbrauch geht ungehindert weiter

immer noch fehlende e-privacy-Verordnung

Transparenzgebot immer noch zu viele Ausnahmen

→ speziell auch bei Behörden, z.B. Verfassungsschutz

Aufsichtsbehörden personell, organisatorisch und finanziell zu schwach

nationale Gestaltungsspielräume in der EU schwächen die Verordnung

ohne Zugriff auf Datenbanken, Server, Codes etc. sind die Geschäftsmodelle entsprechender Unternehmen nicht nachvollziehbar

Geplante Erweiterung der DSGVO

Datenschutz in der Privatsphäre und der elektronischen Kommunikation

Schutz der Endnutzer

Die Verordnung bezieht sich auf den *Weg* personenbezogener Daten; die DSGVO setzt erst an, wenn personenbezogene Daten vorliegen.

Mögliche konkrete Bereiche / Folgen, z.B.:

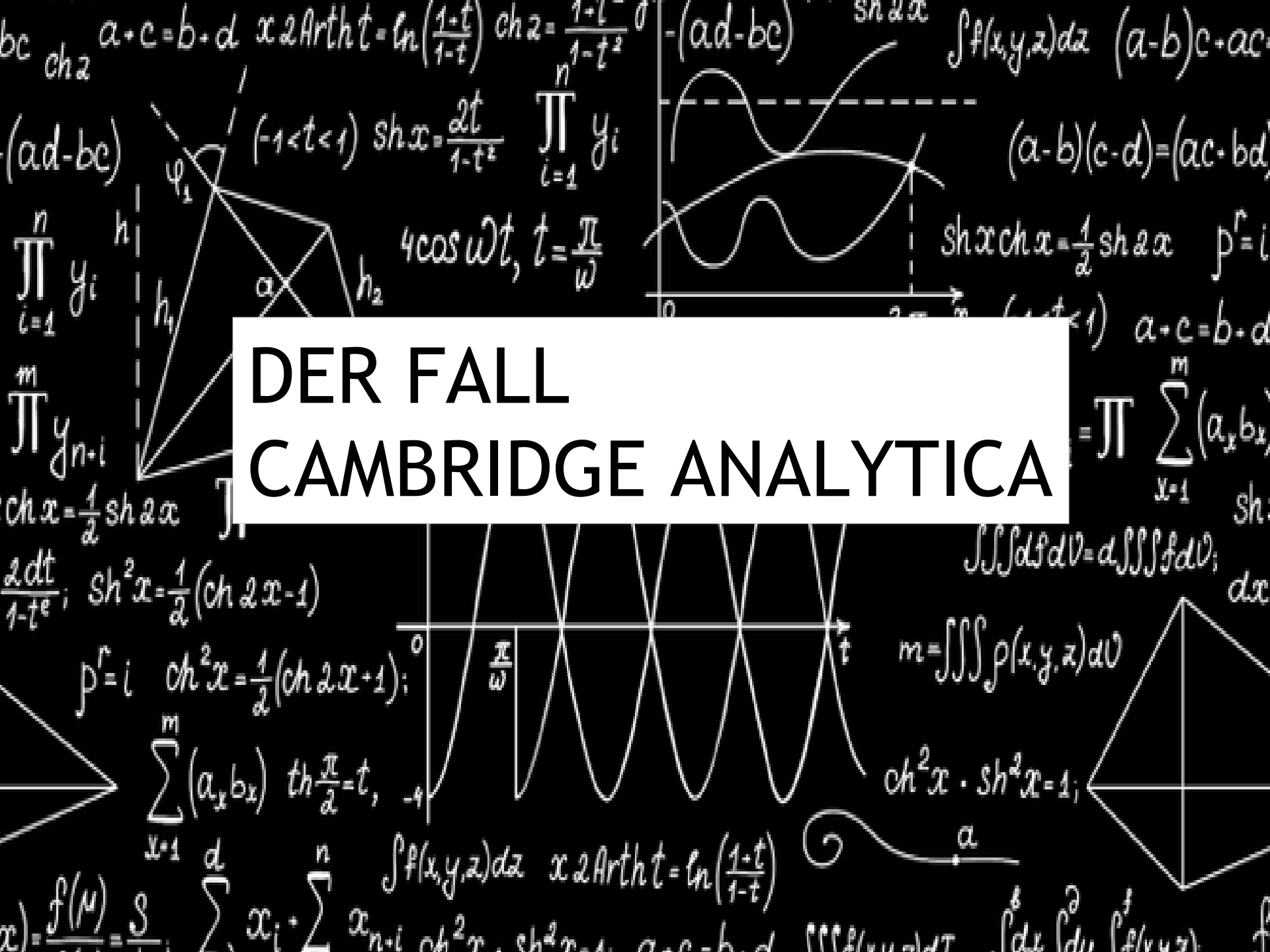
Die Nutzung von Verarbeitungs- und Speicherfunktionen wie Google Analytics wird unzulässig, sofern der Nutzer nicht ausdrücklich darin einwilligt.

→ **explizite Zustimmung zu Cookies, Trackern etc. erforderlich**

Betreiber von Webseiten sollen zukünftig keine Informationen mehr darüber sammeln dürfen, welche Geräte ihre Nutzer verwenden

Direktwerbung = „unerbetene Kommunikation“

Verbesserung/Vereinfachung der Privatsphäre-Einstellungen z.B. bei Browsern. Unbefugte Zugriffe (auch in eine Cloud) müssen technisch unmöglich werden.


DER FALL
CAMBRIDGE ANALYTICA

DIE HAUPTSCHURKEN

Marc Zuckerberg **Alexander Nix** **Steve Bannon**

Quelle: netzpolitik.org

Zeitraum: Wahl zum US-Präsidenten / Wahlkampf Trump

Unrechtmäßiges Sammeln von personenbezogenen Daten

App „thisisyourdigitallife“ (eine Art von Persönlichkeitstest; vom Drittanbieter Global Science Research, extra für diesen Zweck entwickelt)

sammelte Daten nicht nur der ca. 250.000 offiziellen User, sondern ungefragt auch über deren sämtliche Kontakte

→ **geschätzt deutlich über 80 Mio in den USA**

erhielt die Daten über eine regulär vorgesehene Facebook-Schnittstelle und gab diese anschließend an die Analysefirma Cambridge Analytica weiter

Facebook wusste von dieser illegalen Abschöpfung/Weitergabe, hat aber weder die nachfolgende Nutzung und Auswertung unterbunden, noch seine User informiert.

„Schwarzmarkt“ für Facebook-Daten wurde toleriert, da Facebook Gebühren für solche Apps fordert.

Daten gehen anschließend an

→ **CAMBRIDGE ANALYTICA**

Personenprofile auf Basis direkter Information der User

und auf Basis *von Facebook geschlossener* Verhaltensdaten

Klassischer Anwendungsfall von BIG DATA-Modellierung

an einem Datenpool Strukturen, Einstellungs- und Verhaltensprofile analysieren

im großen Stile auf weitere, neue Datensätze (= Personen) anwenden, um diese zu klassifizieren.

„politisches Microtargeting“

Identifikation kleiner und kleinster Personengruppen

lancieren extra zugeschnittener Informationen und Botschaften der Trump-Kampagne; lancieren von Negativinformationen / fake news etc.

Aufsetzen eines ganzen Informationsökosystems aus Webseiten, Blogs o.ä., deren Parteilichkeit nicht erkennbar war

ÜBRIGENS...

Auch Facebook bietet diese Möglichkeiten offiziell an.

Tools:

CUSTOM AUDIENCES

Identifizieren etwa von bestimmten Konsumenten innerhalb von Facebook (z.B. anhand des Abgleichs von email/Telefon)

→ deshalb erbittet Facebook hartnäckig diese Infos

LOOKALIKE AUDIENCES

Aufspüren ähnlicher Personen

AUDIENCE NETWORK

Adressierungsmöglichkeit auf anderen Plattformen und via andere digitale Kanäle von Dritten

Der Miterfinder der berüchtigten Sammel-App ist übrigens mittlerweile als Psycholge bei Facebook tätig...

CAMBRIDGE ANALYTICA

2013 gegründet

Tochterunternehmen der britischen SCL Groups (Strategic Communication Laboratories); Ausgründung speziell für den US-amerikanischen Markt

erhebliche Investitionen durch den ultra-konservativen Milliardär Robert Mercer (der auch einschlägige Politiker und Institutionen finanziert)

Vizepräsident Steve Bannon, vormals Breitbart-CEO, anschließend Chefstrategie von Trump

bei der Datenanalyse Mithilfe eines Mitarbeiters der Überwachungsfirma PALANTIR; deren Gründer sitzt wiederum im Aufsichtsrat von Facebook und unterstützt ebenfalls Trump

2015 erste Enthüllungen durch den britischen GUARDIAN und OBSERVER, im Anschluss kappte Facebook die Verbindung

mit ziemlicher Sicherheit sind die Profilerstellungsmodelle dennoch weiter bei Facebook im Einsatz

zahlreiche Ermittlungen und Anhörung, 2018 Insolvenz

NA UND?

Unterlaufen der demokratischen Öffentlichkeit

Fragmentierung der Öffentlichkeit und der Gesellschaft als Ganzes

Bewusster Versuch der Einschränkung von Information, Schaffung von Räumen mit vorselektierter Info oder *fake news*-Welten

nicht demokratisch legitimierte und zu kontrollierende (oder auch nur immer zu erkennende) Handlungen und Einflussnahmen von privaten Unternehmen auf Politik und Gesellschaft

Thesen

Der Digitale Gläserne Mensch ist ein vielfach und simpel zu manipulierender Mensch.

Der Digitale Mensch ist eine Ware. Sie schafft für Dritte erheblichen ökonomischen Mehrwert, ohne selbst davon angemessen zu profitieren.

Der Digitale Mensch hält den Mechanismus seiner „Abschöpfung“ für seine Belohnung, für die Erfüllung eines Bedürfnisses.

Wir haben gelernt

Dass es Kennwerte gibt, die die in Daten/Merkmalen enthaltende Information verdichten

Dass wir es sehr häufig mit Daten aus Stichproben zu tun haben, die nicht „für sich selbst“ sprechen, sondern als Schätzer der wahren Verhältnisse in der jeweiligen Grundgesamtheit dienen

Dass es verschiedene mathematische Funktionen gibt, in diesem Zusammenhang auch Wahrscheinlichkeitsverteilungen genannt, die wir modellhaft benutzen können, um unsere Stichprobendaten zu beschreiben und zu analysieren

Dass wir insbesondere ihrer Basis die Wahrscheinlichkeit benennen können, mit der die in unserer Stichprobe auftretenden generellen Verhältnisse, Beziehungen zwischen Merkmalen sowie Unterschiede zwischen Fallgruppen auch in der jeweiligen Grundgesamtheit auftreten, d.h. SIGNIFIKANT, sind.

„Das glaube ich nicht!“ lässt sich deshalb in der Regel einfach lösen: Fallzahl, Art der Stichprobe, Verteilung der Merkmale, Test auf Signifikanz. Geht mindestens bei Prozentanteilen auch ganz ohne Computer…