

# Softwaretechnik 1(A) Digitalisierung und Ethik

Autoren: Prof. Dr. Sabine Sachweh

# WANDEL und EXPONENTIELLES WACHSTUM

Der technologische Wandel ist geprägt durch ein **exponentielles Wachstum**:

- **Hardware**

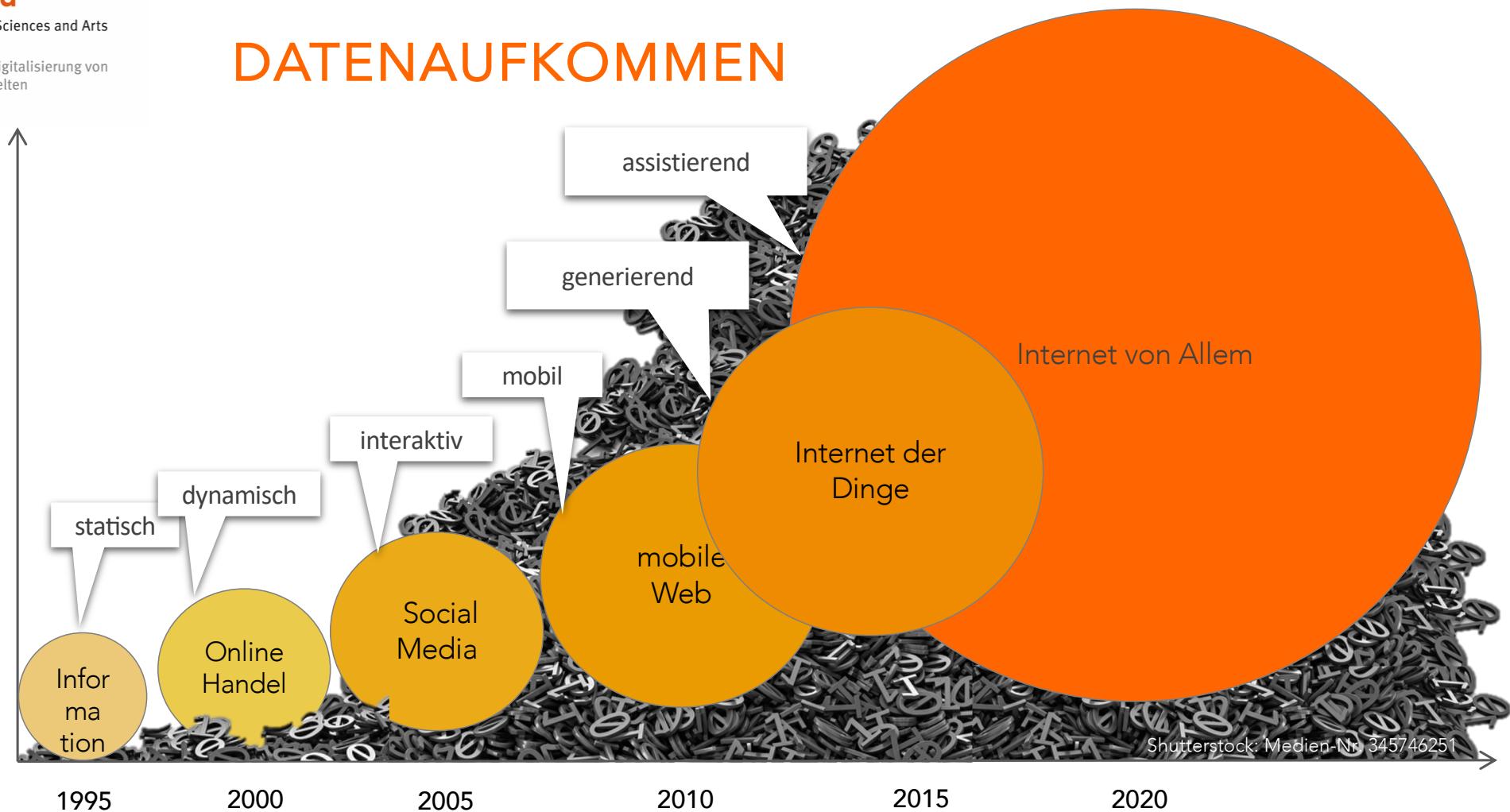
- Moore's Law 1965/1975 → 2016 (51 Jahre korrekt)
- Kompensation durch Multi-core-, Many-core-Systeme, Grafikkarten (GPU), Tensor-Processing-Units (TPU) oder Quantencomputing

- **Exponentiell steigende Nutzung vernetzter Geräte**

- **Datenaufkommen**

2018: 33 Zettabyte → 2025: 175 Zettabyte jährlich [Statista]

# DATENAUFKOMMEN

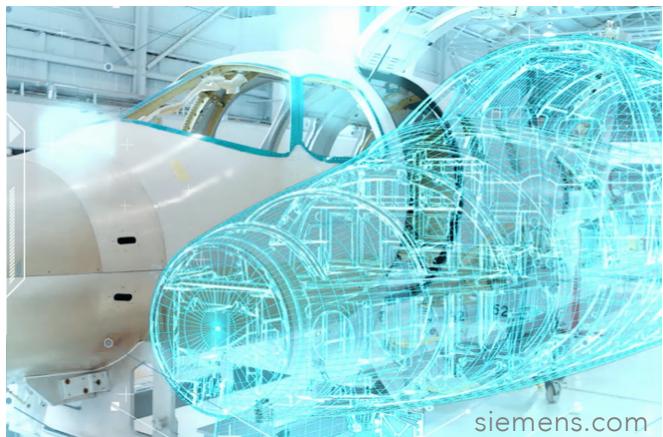


# SOFTWARE, KI und ASSISTENZ

- KI übersetzt **Texte**, an denen andere Algorithmen jahrzehntelang gescheitert sind
- KI erkennt schneller und mit weniger Fehlern die wichtigsten **Elemente in Bildern** → Schildererkennung, Spurhalteassistenz, Hautkrebsdiagnose
- KI wird zunehmend besser in der **Sprachverarbeitung**  
z.B. Google Duplex
  - Google Assistant reserviert einen Tisch im Restaurant 
  - Google Assistant bucht einen Friseur-Termin 
- KI (wird in Teilen) **kreativ**  
→ Musikalische Kompositionen, Gedichte, Gemälde



# CHANCEN



# RISIKEN

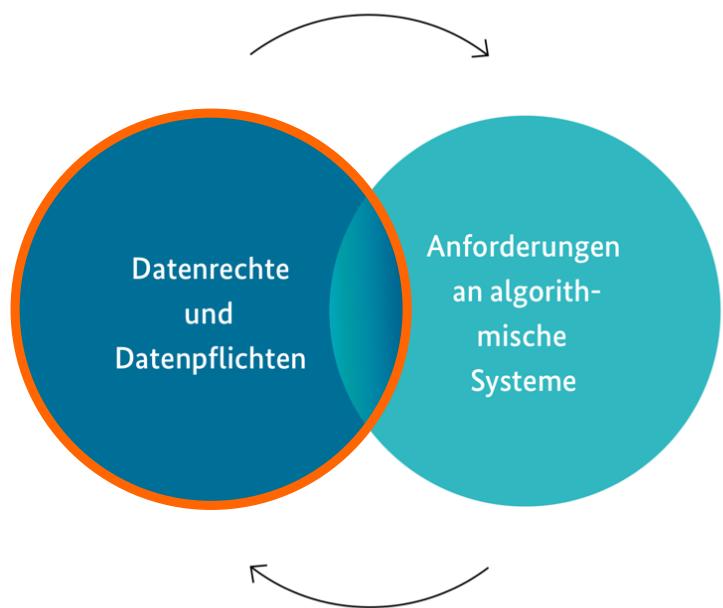


# AUSBALANCIEREN

- Die schnelle technologische Entwicklung liefert nicht per se nur gute Lösungen.
- Problematische Lagen sind durch ein Zusammenspiel aus **Ethik** (gut oder schlecht), **Technikentwicklung** (funktioniert oder nicht) und ggf. **Regulation** (erlaubt oder nicht) zu klären.
- Unterstützung können dabei u.a. Beiräte und Kommissionen wie die Datenethikkommission



# PERSPEKTIVEN



- Wechselseitig ergänzende und bedingende ethische Diskurse
- **Daten-Perspektive**
  - objektive Anforderungen an den Umgang mit Daten
  - subjektive Rechte, welche Personen gegenüber einer bestimmten anderen Person oder auch gegenüber jedermann geltend machen können
  - Dateneigentum
  - Open Data

# DIFFERENZIERUNG

- Identifizierung von Persönlichkeitsmerkmalen  
wie z. B. Adresse, Gesundheitszustand, Vertrauenswürdigkeit oder Zahl  
allem bekannte Fakten wie die Zahlungshistorie u.a. genutzt, heute we  
Beziehungs- und Bedienungsdaten genutzt
- Kategorisierung und Profiling  
Erstellung eines Datenmodells zur Kategorisierung von Personen
- Scoring (u.a. für Prognosen)  
Zuordnung von Zahlenwerten zu Personenprofilen, meistens auf einer  
Wahrscheinlichkeitswerten für ein bestimmtes zukünftiges  
Verhalten



©Sandra Schulze, Katharina Zweig:  
„Ein Algorithmus hat kein Taktgefühl“

# DISKRIMINIERUNG

Geschütztes Merkmal	Art. 3 GG	§§ 1 u.a. AGG**	Erwg. 71 DSGVO	Art. 9 DSGVO
„Rasse“ oder ethnische Herkunft	ja	ja	ja	ja
Abstammung, Heimat, Herkunft	ja			
Geschlecht	ja	ja		
Sprache	ja			
Politische Meinung bzw. Anschauung und sonstige Anschauung	ja		ja	ja
Religion und Weltanschauung	ja	ja	ja	ja
Behinderung	ja	ja		
Alter		ja		
Gewerkschaftsangehörigkeit	ja*		ja	ja
Genetische Merkmale bzw. Anlagen und Gesundheitszustand	ja		ja	ja
Biometrische Merkmale				ja
Sexualleben, sexuelle Identität bzw. Orientierung		ja	ja	ja

Ungleichbehandlung von Personen auf Basis eines geschützten Merkmals → Grundlage der Kategorisierung und Zuschreibung von Merkmalen („Proxies“) werden an verschiedenen Stellen definiert → Zuschreibung nutzt u.a. statt geschützter Merkmale „Proxies“, um die „Geschlecht“

[Quelle: Antidiskriminierungsstelle]

## WORKAROUND: Facebook

- Facebook bietet ca. 5.000 Optionen, Werbung zu personalisieren, dem sogenannten „**microtargeting**“
  - 2018 klagte die Bürgerrechtsorganisation American Civil Liberties Union zusammen mit weiteren Partnern gegen Facebook, da
    - neben den drei geschützten Merkmalen (Standort, Alter, Geschlecht), zahlreiche weitere Merkmale z. B. „[...] Single Dads, Single Moms, Soccer Mom, Working Moms, Working Mother, Bad Moms, Strong Single Moms!, Proud Single Mother, The Single Moms Club“ als **Proxies** zur geschlechterspezifischen Differenzierung genutzt werden konnten
    - der „**lookalike audience**“ Dienst zu Diskriminierung führen kann - dieser Dienst liefert Nutzer, die demografisch ähnliche Merkmale zu einer als Input dienenden Liste – wobei Facebook **intern** Merkmale wie Standort, Alter, Geschlecht und Interessen nutzt.

[Financial Times, 19. März 2019]

## PROBLEM: DATENSCHUTZ

- Beispiel: Bild- und Videodaten
- klassischer Ansatz zur Anonymisierung von Kameradaten:  
Ausschneiden oder Verpixeln
- Wissenschaftler der Universitäten Princeton und Purdue belegten  
2018, dass leichte Veränderungen an Straßenschildern, eine KI  
zur Schilderkennung aus dem Konzept bringt  
[Studie: [arxiv.org/abs/1802.06430](https://arxiv.org/abs/1802.06430)]  
⇒ Vermutung: Ausschneiden identifizierender Bildinformation  
müsste eine KI völlig aus dem Konzept bringen



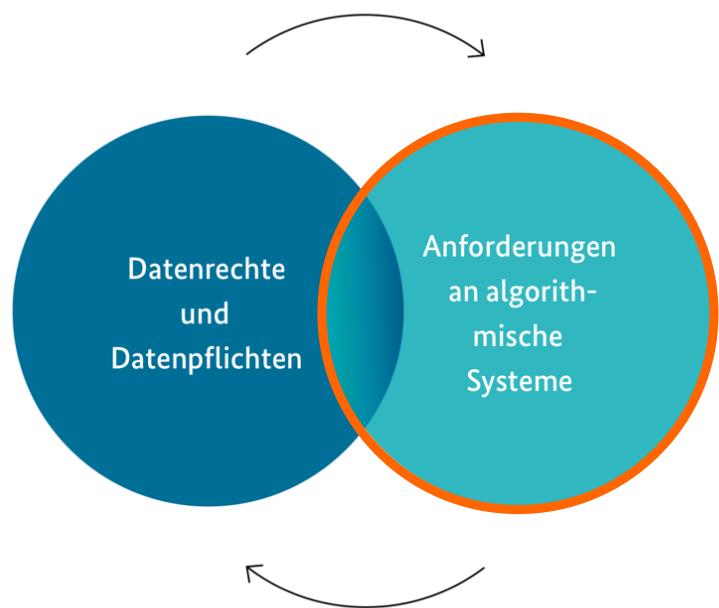
## BEISPIEL: Brighter AI

### ■ Brigther AI Gewinner (Start-Up aus dem Hella-Inkubator)

- Idee: ursprünglich „Digitale Rückspiegel“
  - Problem: Benötigtes Bildmaterial konnte nicht auf den Straßen gesammelt werden, da dort u.a. Personen und Fahrzeuge mit Kennzeichen unterwegs waren.
  - Idee: Anonymisierungslösung, die Informationen in Bildern und Videos, welche eine Identifizierung ermöglichen würden identifiziert und künstliche Ersatzdaten generiert.  
Ansatz: „Deep Natural Anonymization“
  - Ziel: Nutzung von Daten aus dem öffentlichen Bereich bei Wahrung der **Privatsphäre in der Öffentlichkeit**
- ⇒ Enabler: Push für das autonome Fahren und Smart City-Anwendungen



# PERSPEKTIVEN



- Wechselseitig ergänzende und bedingende ethische Diskurse

## Daten-Perspektive

- objektive Anforderungen an den Umgang mit Daten
- subjektive Rechte, welche Personen gegenüber einer bestimmten anderen Person oder auch gegenüber jedermann geltend machen können
- Dateneigentum
- Open Data

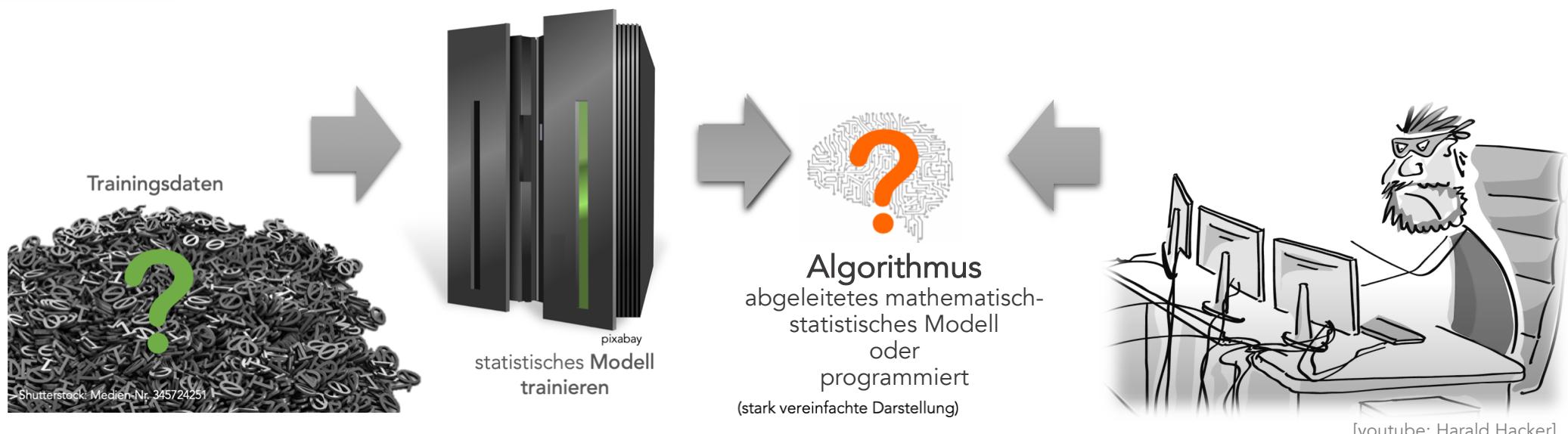
## Algorithmen-Perspektive

- betrachtet die Auswirkungen des datenverarbeitenden algorithmischen Systems auf Einzelne und die Gesellschaft
- Beziehung von Mensch und Maschine; insbesondere auf die Automatisierung sowie auf die Verlagerung auch komplexer Handlungs- und Entscheidungsprozesse auf sog. autonome Systeme

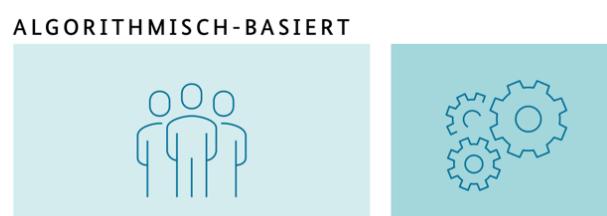
# MENSCH oder ALGORITHMUS



# PROBLEM: NACHVOLLZIEHBARKEIT



# VARIANTEN VON ENTSCHEIDUNGSSYSTEMEN



**vollständig automatisierte Entscheidung**  
Beispiel: Automatisierte Ablehnung

algorithmisch-basierte Entscheidung, wobei die  
Selbstbestimmung des Menschen eingeschränkt ist  
Beispiel: Bewerber-Scoring

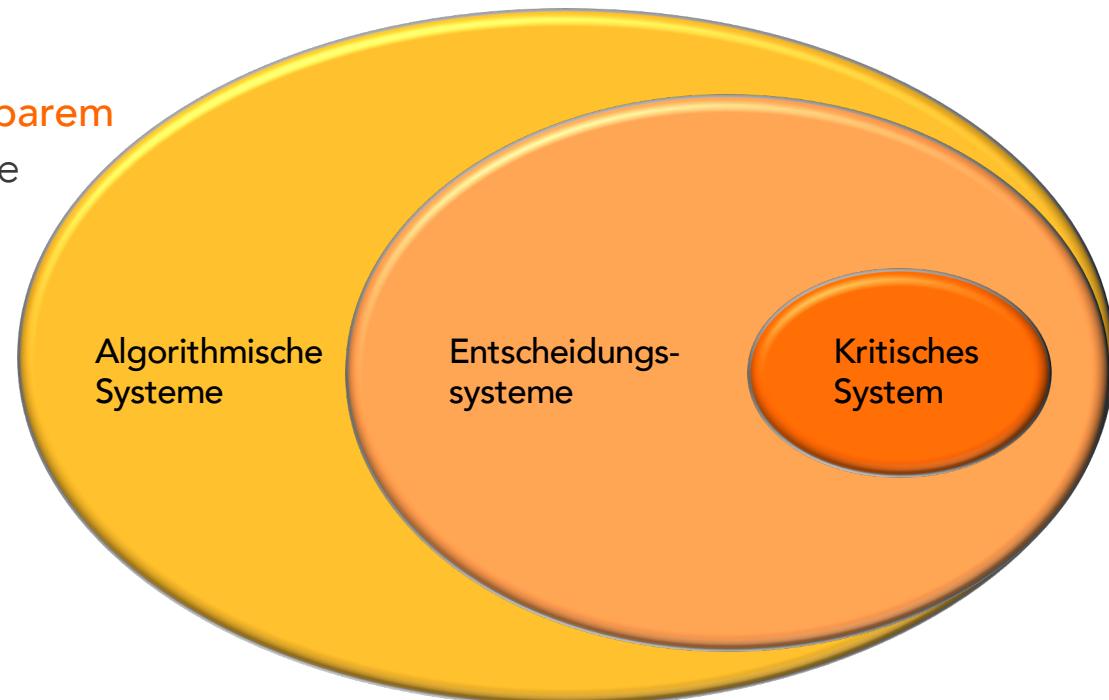
menschliche Entscheidungen, die sich auf  
algorithmisch berechnete (Teil-)Informationen stützen  
Beispiel: Bewerberdatenbank

# KRITISCHE ALGORITHMISCHE SYSTEME

kritische algorithmische Systeme,

↔ Systeme mit **gewissem bis unvertretbarem Schädigungspotential** sind solche, die

- über Menschen entscheiden
- über Ressourcen, die Menschen betreffen entscheiden
- gesellschaftliche Teilhabemöglichkeiten von Menschen beeinflussen



Stufe 5

Stufe 4

Stufe 3

Stufe 2

Stufe 1

Beginn spezifischer Regulierung



## KRITIKALITÄTSPYRAMIDE [DEK 2019]

*Autonome  
Waffensysteme  
Kreditscoring  
...Scoring  
Personalisierte  
Preissetzung  
Dynamische  
Preissetzung*

*Getränkeautomat,  
Lohnsteuerjahres-  
ausgleich*

vollständiges oder teilweises Verbot

zusätzlich Transparenz und Nachvollziehbarkeit z. B. Live-Schnittstelle zur **kontinuierlichen Kontrolle** durch Aufsichtsinstitution

zusätzlich z. B. **Ex-ante Zulassungsverfahren**

z.B. formelle und materielle Anforderungen, wie

- **Transparenzpflichten**,
- Veröffentlichung einer **Risikofolgenabschätzung** oder **Kontrollverfahren**, wie
  - Offenlegungspflichten gegenüber Aufsichtsinstitutionen,
  - Ex-post-Kontrolle (**Input-Output-Kontrolle** bei Verdacht),
  - (branchenspezifische) Auditverfahren
  - Auskunftsrechte für Betroffene

keine Pflichten

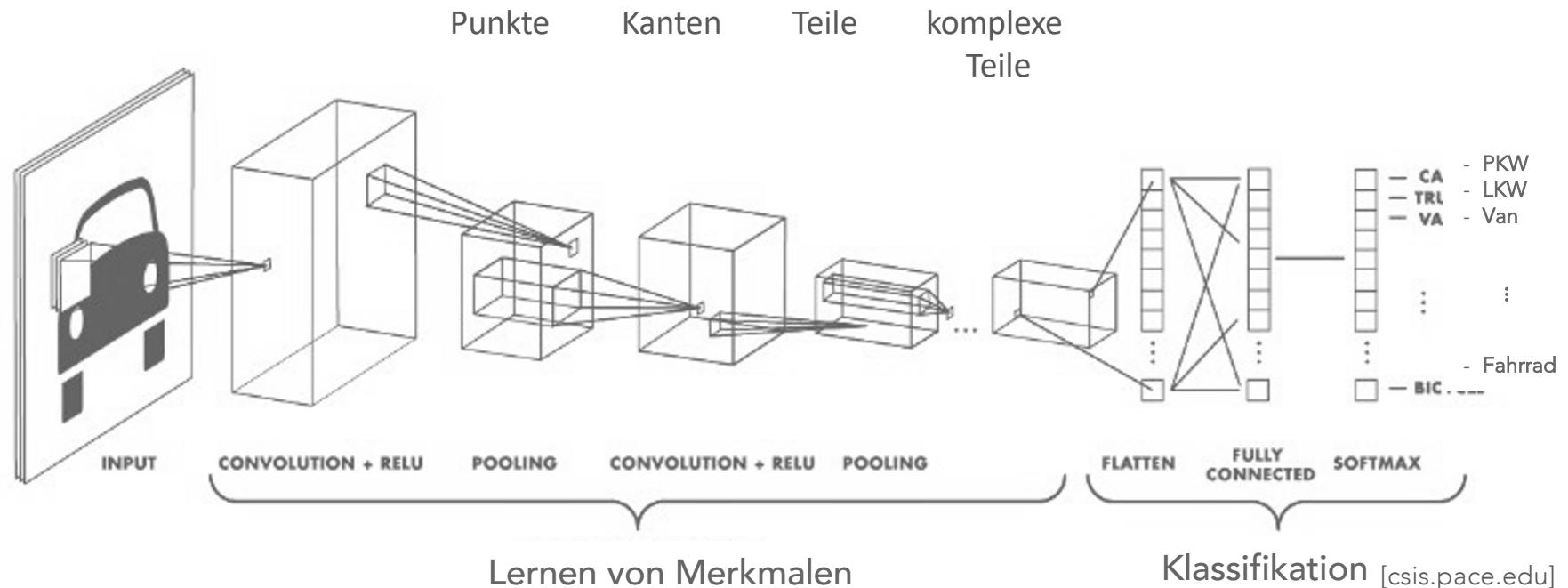
## BEISPIEL NACHVOLLZIEHBARKEIT: SpRAy

- Forschungsbereich Explainable AI (XAI)
- Spectral Relevance Analysis (SpRAy) ermöglicht Blick in die Black Box
- Open-Source-Software visualisiert und interpretiert Neuronale Netze und andere Machine Learning-Modelle
- ermittelt u.a. anhand **welcher Merkmale** ein Auto als Auto identifiziert wird und
- erkennt an Knotenpunkten, **wie Informationen durch das Netz fließen**
- **findet Fehler** in trainierten Modellen, die sonst schwer zu identifizieren sind  
⇒ Schiffserkennung

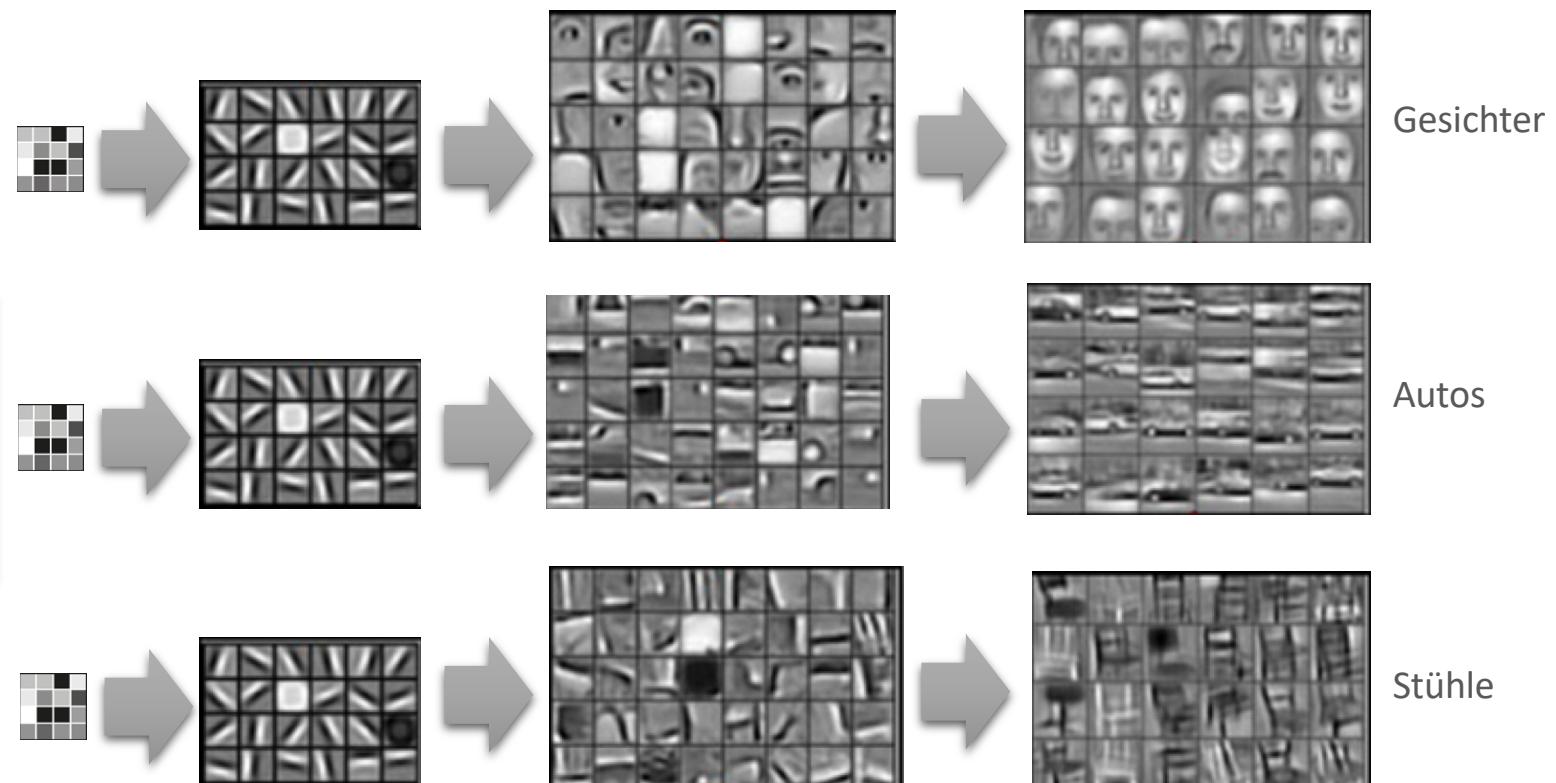


# BEISPIEL: SpRAy

Beispiel: Faltungsnetz

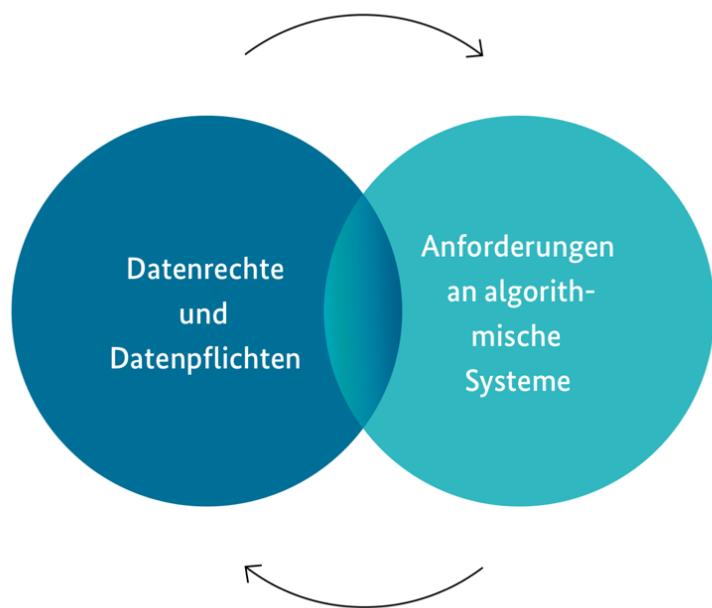


## BEISPIEL: SpRAY



[stats.stackexchange.com]

## DEK - PERSPEKTIVEN



- Zusammenwirken **verschiedener Akteure** und unterschiedlicher **Governance-Instrumente** auf mehreren Regulierungsebenen ist von entscheidender Bedeutung

⇒ Mehr-Ebenen-Governance

## MEHR-EBENEN-GOVERNANCE

### Staat:

- rechtlichen Rahmenbedingungen
- infrastrukturelle und technische Voraussetzungen
- Bildung
- F&E und Transfer
- Digitale Souveränität
- Europäische Perspektive

### Bildung:

- digitale Kompetenz
- Lehrpläne Schulen/Berufsschulen/Hochschulen
- berufliche Weiterbildung
- lebensbegleitende informelle Bildung
- zivilgesellschaftliche Strukturen
  - digitales Ehrenamt
  - Tech-Accountability-Journalismus

### Unternehmen:

- Verantwortung → Corporate Digital Responsibility
- als Querschnittsaufgabe
- Ausbildung professionsethischer Standards und Verhaltenskodizes

### Standardisierung:

- Folgenabschätzung bezüglich Standards muss um ethische und soziale Aspekte erweitert werden
- Akteure der Zivilgesellschaft, Datenschutzbehörden, Verbraucherschützer oder Vertreter von Betroffenenorganisationen sollten sich ebenfalls einbringen können

### Technikgestaltung:

- partizipative Produktentwicklung
- ethics by/in design
  - best practices
  - Werkzeuge
  - (open source) Komponenten
  - .....
- Aushandlung ethischer Prinzipien auch mit Betroffenen

### Forschung:

- s. Technikgestaltung
- Interdisziplinarität
- Transfer

 Akteur

 Instrument

## WEITERE INFORMATIONEN



Weitere Information zu den  
Handlungsempfehlungen der  
Datenethikkommission finden  
Sie auf den Seiten des Innen- und  
Justizministeriums

[www.fh-dortmund.de](http://www.fh-dortmund.de)

 [facebook.com/fhdortmund](https://facebook.com/fhdortmund)

 [instagram.com/fhdortmund](https://instagram.com/fhdortmund)

 [twitter.com/fh\\_dortmund](https://twitter.com/fh_dortmund)

 [youtube.com/FachhochschuleDO](https://youtube.com/FachhochschuleDO)