



KI und Datenqualität – Perspektiven aus Data Science, Ethik, Normung und Recht

# Künstliche Intelligenz

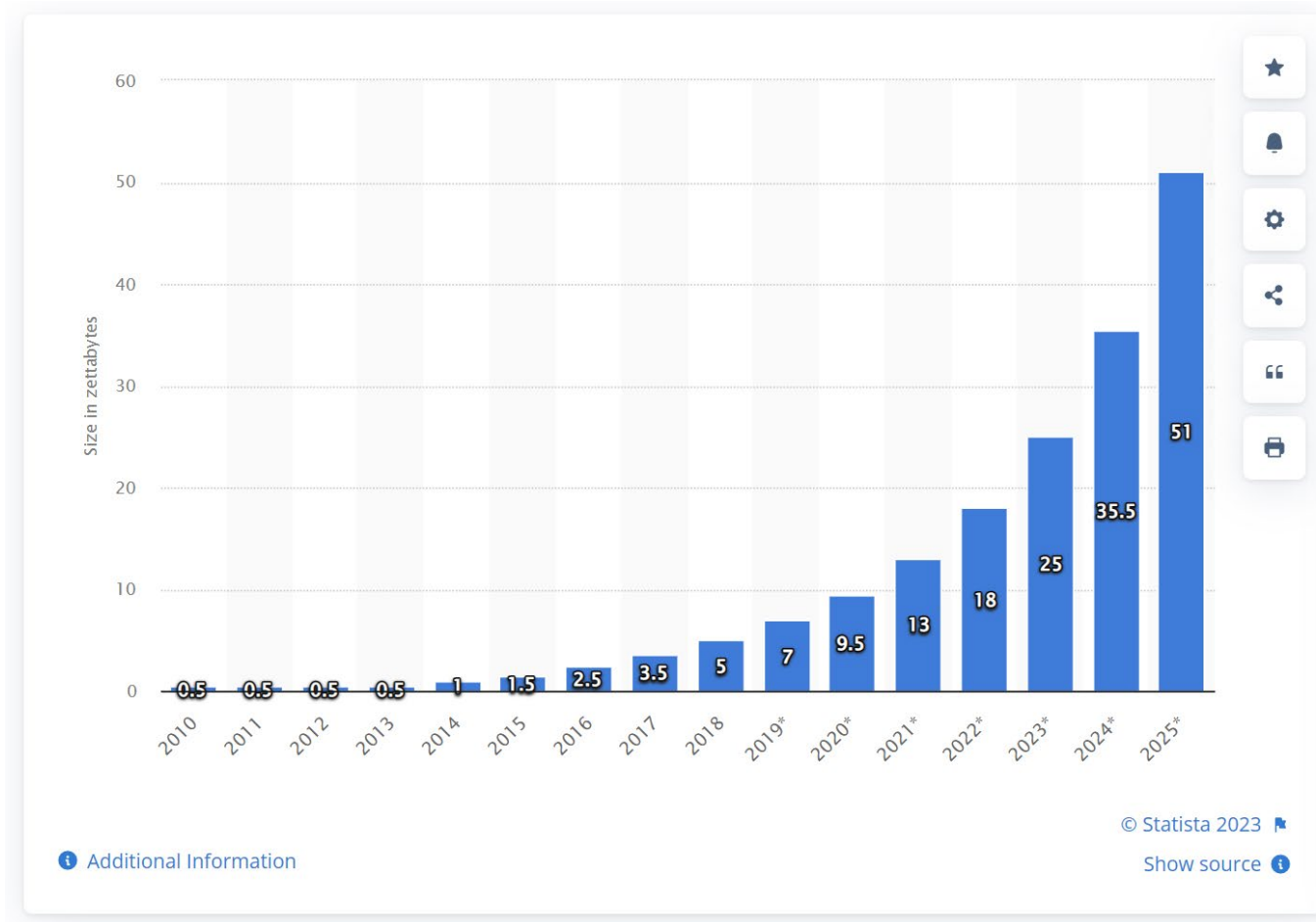
Prof. Felix Naumann

# Big Data

- Datensammlung seit 3000 Jahren
- Digitalisierung seit 100 Jahren
- Datenbanken seit 50 Jahren
  - Unternehmensdaten
  - Transaktionen
  - Messungen
  - Texte, Bilder, Videos
  - Webdaten
- Big Data seit 20 Jahren



# Volume – Der Umfang der Daten



<https://www.statista.com/statistics/949144/worldwide-global-datasphere-real-time-data-annual-size/>

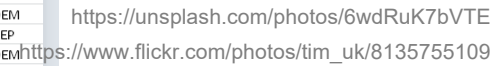
# Velocity – die Geschwindigkeit der Daten



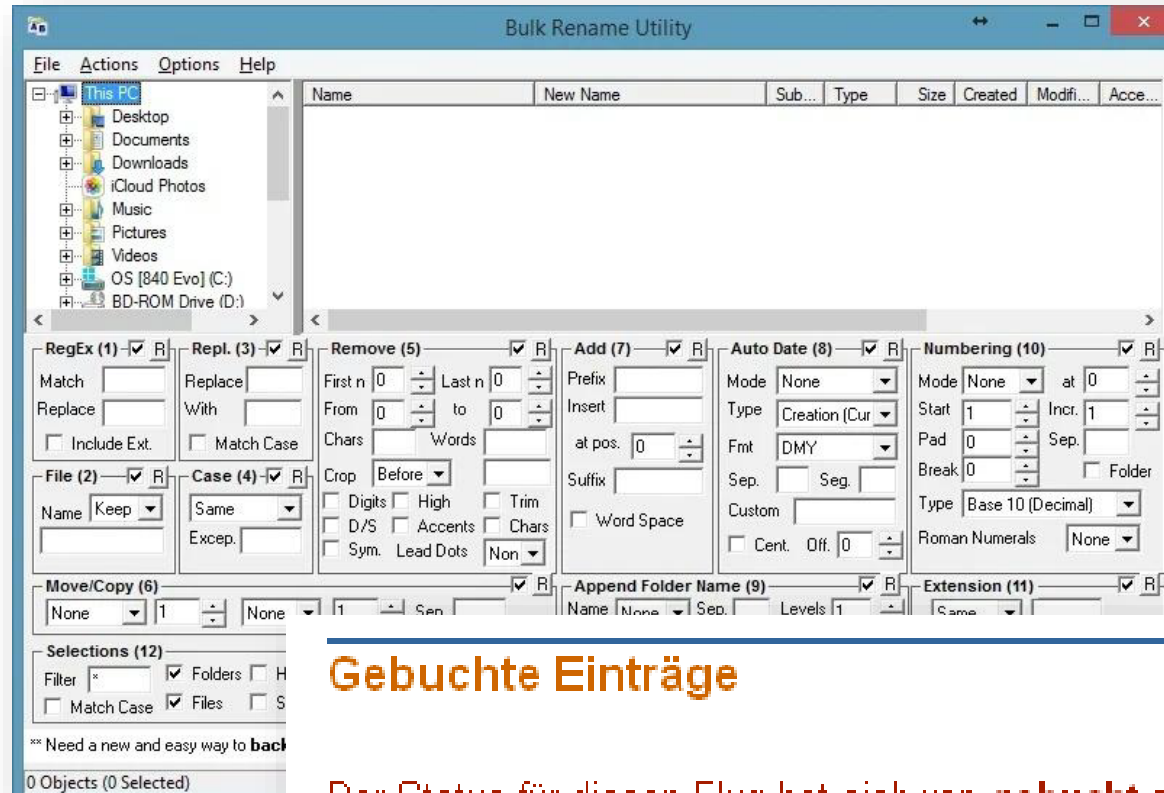
<https://www.flickr.com/photos/78629042@N00/479370088>



Felix Naumann



# Veracity – die Wahrhaftigkeit der Daten



## Gebuchte Einträge

Der Status für diesen Flug hat sich von **gebucht** auf **Gebucht** geändert.  
Bitte überprüfen Sie die nachstehenden Änderungen.

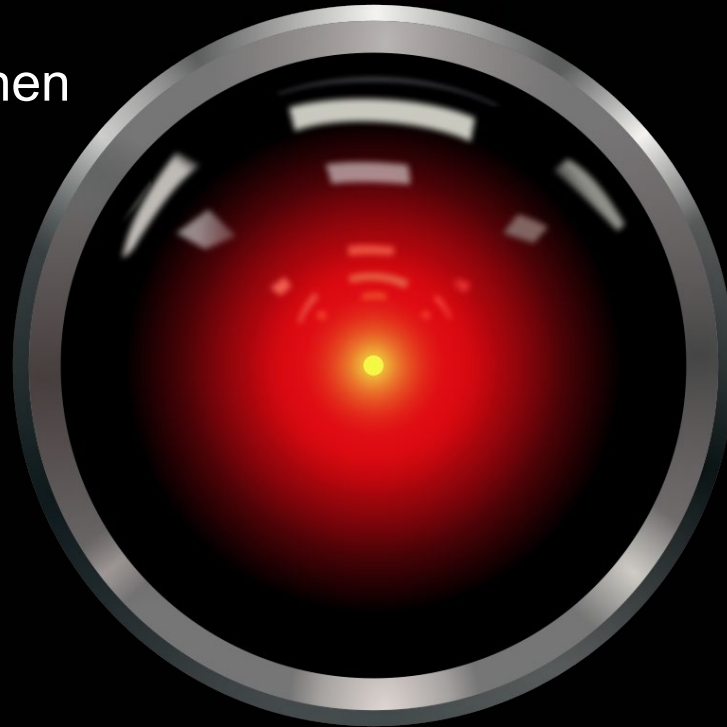
**Flug:** Berlin nach Nürnberg

[zurück](#) ↑

**E-Ticket - Der Kauf wurde von der Fluglinie bestätigt.** Es werden keine Tickets per Post zugestellt. Ihr Ticket, welches das herkömmliche Panieticket ersetzt, erhalten Sie am Check-in.

# Künstliche Intelligenz

Automatisierung intelligenten Verhaltens  
Eigenständig Probleme lösen  
Eigenständig lernen  
Mit Unsicherheit umgehen



Robotik  
Fragen beantworten  
Schlüsse ziehen  
Musteranalyse  
Mustererkennung  
Gesichtserkennung  
Sprachliche Intelligenz

# KI Stufen

- Schwache Intelligenz
  - Konkrete Anwendungsprobleme
  - Menschliches Denken und technische Anwendungen in Einzelbereichen unterstützen
  - Siri, Bilderkennung, Robotik, Spiele, ...
- Allgemeine bzw. starke Intelligenz
  - Auf Augenhöhe mit Menschen arbeiten
  - Verschiedenartige Probleme lösen
  - Bei schwierigen Aufgaben unterstützen
  - „So intelligent wie ein Mensch“

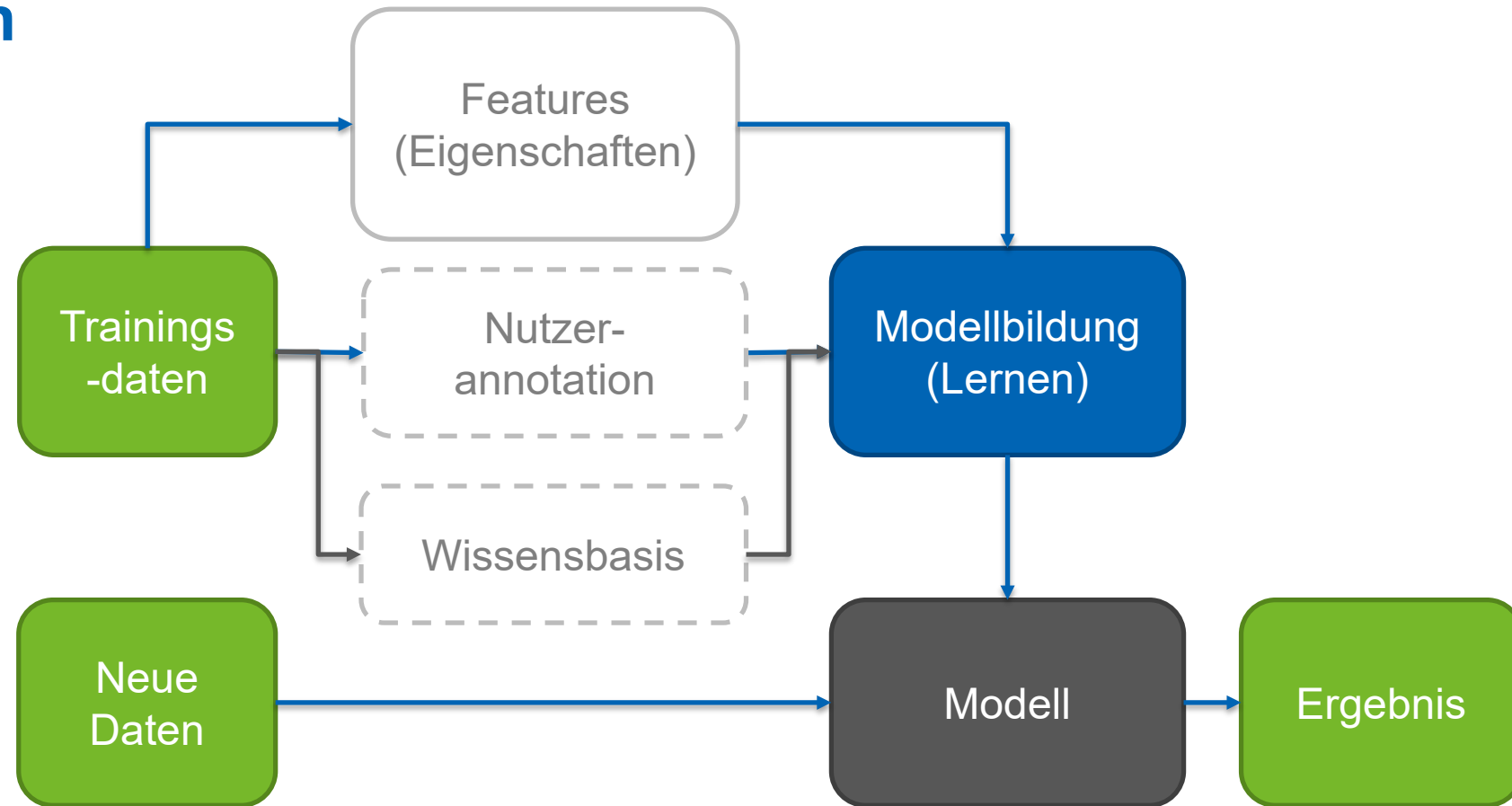


[Rossum's Universal Robots](#)



# Maschinelles Lernen

- Unüberwacht
  - Clustering
  - Assoziationsregeln
  - Word Embedding
- Überwacht
  - Klassifikation
  - Regression
- Weitere Varianten
  - Teilüberwachtes Lernen
  - Aktives Lernen
  - Distanzlernen: Bezug auf Wissensbasis

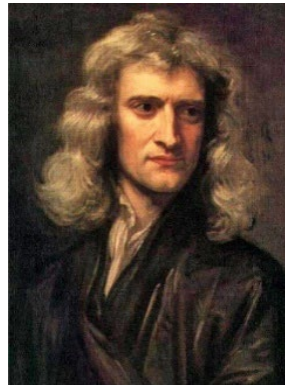


# Big Data = Wissenschaft?

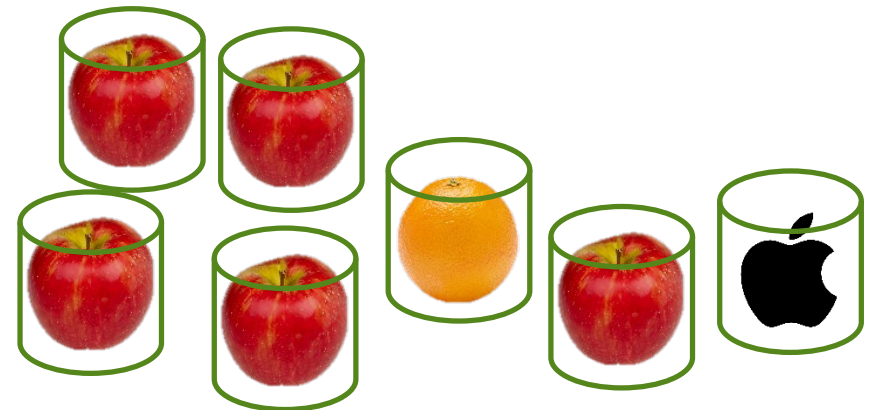
## Vom Ende der klassischen Wissenschaft

- Das Ende der Theorie: Die Datenflut macht die wissenschaftliche Methode obsolet
  - “*All models are wrong, but some are useful.*” (George Box)
  - “*All models are wrong, and increasingly you can succeed without them.*” (Peter Norvig)
- Vor Big Data: Korrelation ist nicht Kausalität!
- Mit Big Data: Na und?
  - Mit Petabytes können wir sagen: “**Korrelation genügt.**”

$$F = G \frac{m_1 m_2}{r^2}$$



vs.



[https://en.wikipedia.org/wiki/Isaac\\_Newton#/media/File:GodfreyKneller-IsaacNewton-1689.jpg](https://en.wikipedia.org/wiki/Isaac_Newton#/media/File:GodfreyKneller-IsaacNewton-1689.jpg)

<https://www.flickr.com/photos/seattlecitycouncil/39074799225/>

Nächstes Video:  
**Wertorientierte Gestaltung von KI**

EBERHARD KARLS  
UNIVERSITÄT  
TÜBINGEN



INTERNATIONAL CENTER FOR  
ETHICS IN THE SCIENCES AND  
HUMANITIES (IZEW)



**VDE**