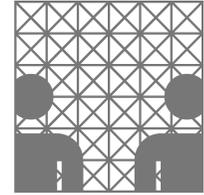




Universität Hamburg

Department  
Informatik  
Arbeitsbereich  
Bildverarbeitung (BV)



## Impulsreferat:

# *Ortslinien und Maschinelles Lernen*

**H. Siegfried Stiehl, Prof. Dr.-Ing.**  
**(„Seniorprofessor“ 😊)**

**Symposium „Ortslinien“ • UE Hamburg • 26. April 2019**

# Interactive Exploration of Digitized Manuscripts (iXMan\_Lab/Sandbox)

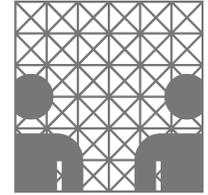


(Source: DAS2018, GI-Jahrestagung 2018 Workshop “CS&DH”)



Universität Hamburg

Department  
Informatik  
Arbeitsbereich  
Bildverarbeitung (BV)



# **B**ig Data .UND. **A**rtificial Intelligence?

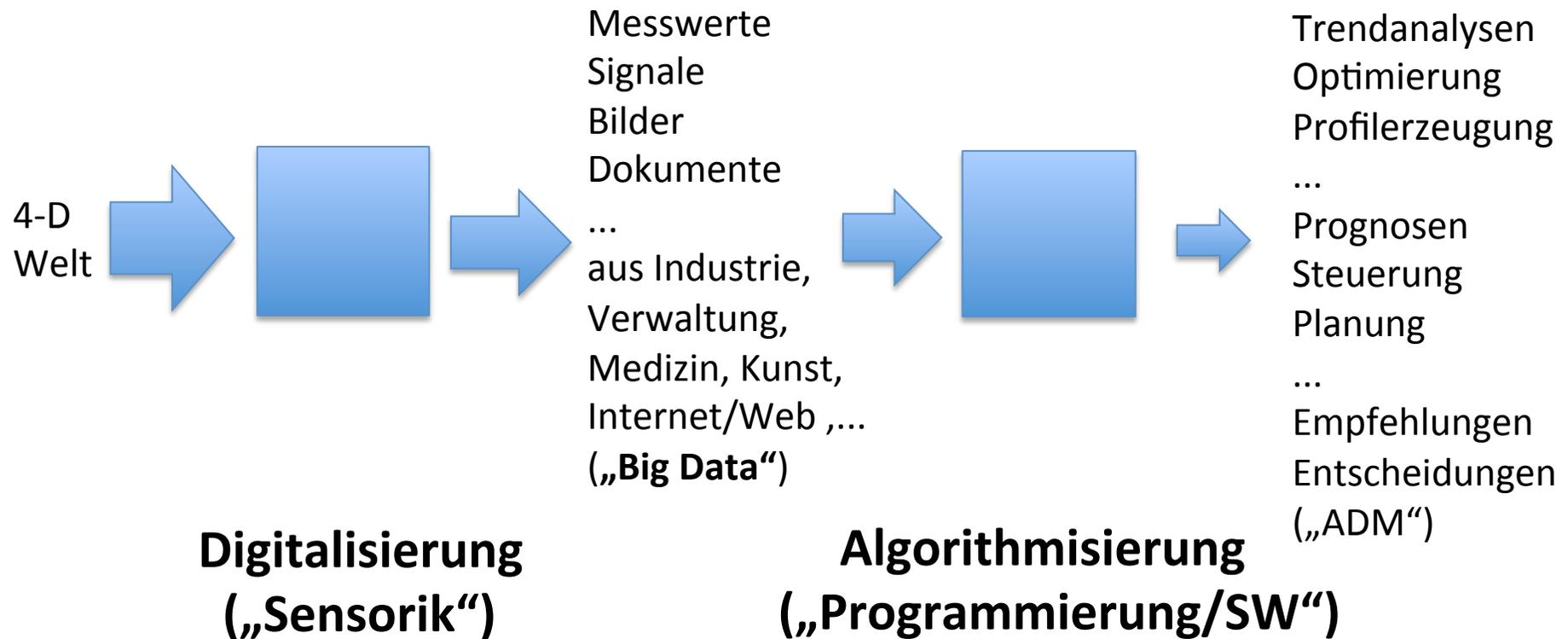
*Oder:*

Wer **B** sagt, muss auch **A** sagen!

H. Siegfried Stiehl, Prof. Dr.-Ing.  
(„Seniorprofessor“ 😊)

Finissage „Freiheit 2.0“ • StadtPalais • Sturgert • 24. Juni 2018

# Elementares (in aller Kürze)



„Big Data Analytics“: „data exploration via **A**(rtificial) **I**(ntelligence)“  
oder „**M**(achine) **L**(earning)“

# Big Data: Definition (**5 Vs**)

- **Volume** (... Assyrer ... Fugger ... Hanse ... Klöster ... Verwaltung ... Industrie x.0 ... **neu**: Datenökonomie)
- **Velocity** (Brieftauben ... Boten ... Reiter ... Kabel ... Glasfaser ... **neu**: Internet bzw. Web und „cloud“)
- **Variety** (Handschriften ... Druckstücke ... Formulare ... Tabellen ... Grafiken ... Bilder ... Sensoren ... Filme ... eMail ... (a)soziale Medien ... **neu**: Konvergenz/Ubiquität)
- **Veracity** := ((Auf-)Richtigkeit ... Realitätstreue ... Qualität der Daten (Geheimtinte ... Standards ... Kryptographie))
- **Value** (Mehrwert ... Ressource ... Währung ...)

# Big Data: Definition (plus 2 Vs)

- **V**olume (... Assyrer ... Fugger ... Hanse ... Klöster ... Verwaltung ... Industrie x.0 ... **neu**: Datenökonomie)
- **V**elocity (Brieftauben ... Reiter ... Boten ... Kabel ... Glasfaser ... **neu**: Internet bzw. Web und „cloud“)
- **V**ariety (Handschriften ... Druckstücke ... Formulare ... Tabellen ... Grafiken ... Bilder ... Sensoren ... Filme ... eMail ... (a)soziale Medien ... **neu**: Konvergenz/Ubiquität)
- **V**eracity ((Auf-)Richtigkeit ... Realitätstreue ... Qualität der Daten)
- **V**alue (Mehrwert ... Resource ... Währung ...)
- **VertrauensV**erlust (NSA, ..., Facebook, ..., DSGVO)

# Big Data Analytics: Definition

- i.w.S. sog. „intelligente“ (*semi-*)*automatische* und/oder *interaktive* computergestützte Analyse/Erschließung (und Visualisierung) von un-/strukturierten Daten
  - aus unterschiedlichen Quellen,
  - in unterschiedlichen Formaten,
  - mit unterschiedlichen Metadaten,
  - mit unterschiedlichem Vorwissen über Domänen und
  - mit vielfältiger Kontextualisierung
- Multimedialität (Text, Grafik/Bild,..., Audio/Video,..., HTML,...)
- Multidimensionalität (1-D, 2-D, 3-D, Raum-Zeit-Diskretuum)
- Multimodalität (Augen, Finger, Hand, Ohren, ..., Körper)

# KI in der Nusschale

- **McCarthy** (1955, Dartmouth Summer Research Project on Artificial Intelligence; McCarthy, Minsky, Rochester und Shannon)

„Die Herstellung einer **Maschine**, die sich auf eine Art und Weise verhält, die wir **intelligent** nennen würden, wenn ein **Mensch** sich so verhielte.“

- **Maschine**: (mathematische) Turingmaschine als Theoriebasis bzw. von-Neumann-Computer <-> **Algorithmus**
- **Künstliche Intelligenz**: *Algorithmisierung* der menschlichen Fähigkeit zum Problemlösen (im weitesten Sinn)

# Zitat (1)

„Wissenschaftlich interessant ist **nicht** die Frage: „Was ist eine intelligente Maschine?“

**Sondern:** „Wer kann eine bestimmte Aufgabe in einem bestimmten Maßstab lösen?“

**Dann** können wir vergleichen, wer diese Aufgabe am besten löst: ein Mensch, ein Tier oder ein Computer.

**Wenn** wir das so sehen, dann ist menschliche Intelligenz nur ein Satz an Fähigkeiten, den wir so benannt haben.

Computer und Tiere verfügen über andere Fähigkeiten.

**Intelligenz ist nur ein Wort – nichts Reales.“**

*(Prof. Z. Ghahramani, U Cambridge, Direktor Leverhulme Centre for the Future of Intelligence, in: TU intern, Nr. 11/2017)*

# NB: Wie ist Intelligenz definiert?

- Definition: Fehlanzeige (keine integrale Theorie)
- Varianten: kognitive, soziale, emotionale, rationale, ethologische Intelligenz ... Bauchgefühl ... Kreativität
- Fähigkeiten, Kompetenzen, ..., Inselbegabungen
- Kontext der Evolution: adäquate Verhaltensorganisation in einer ökologischer Nische auf Basis des Sensoriums
- Wahrnehmen, (Mit-)Fühlen, Entscheiden, Sprechen, Handeln, ... in zyklischer Wechselwirkung mit Umgebung
- Rätsel „Bewusstsein“: Forschungsgegenstand

# Wieso ist all die KI Mathematik ...?

- Strukturwissenschaft und Kulturleistung
- Algebra, ..., Gruppentheorie, ..., Logik, ..., Lineare Algebra, ... ,Analysis, ..., Dynamische Systeme, ... -> **KI**
- „computational theory“ := Algorithmus
- Definition (Kurzform): „Endliche Zahl elementarer wohldefinierter Schritte einer Berechnung, die bei zulässiger Eingabe zu einem eindeutigen und korrekten Ergebnis kommt.“
- Konsequenz: elementare/repetitive Tätigkeiten bzw. Handlungen sind im Prinzip algorithmisierbar

# ... und was ist nun ein Algorithmus?

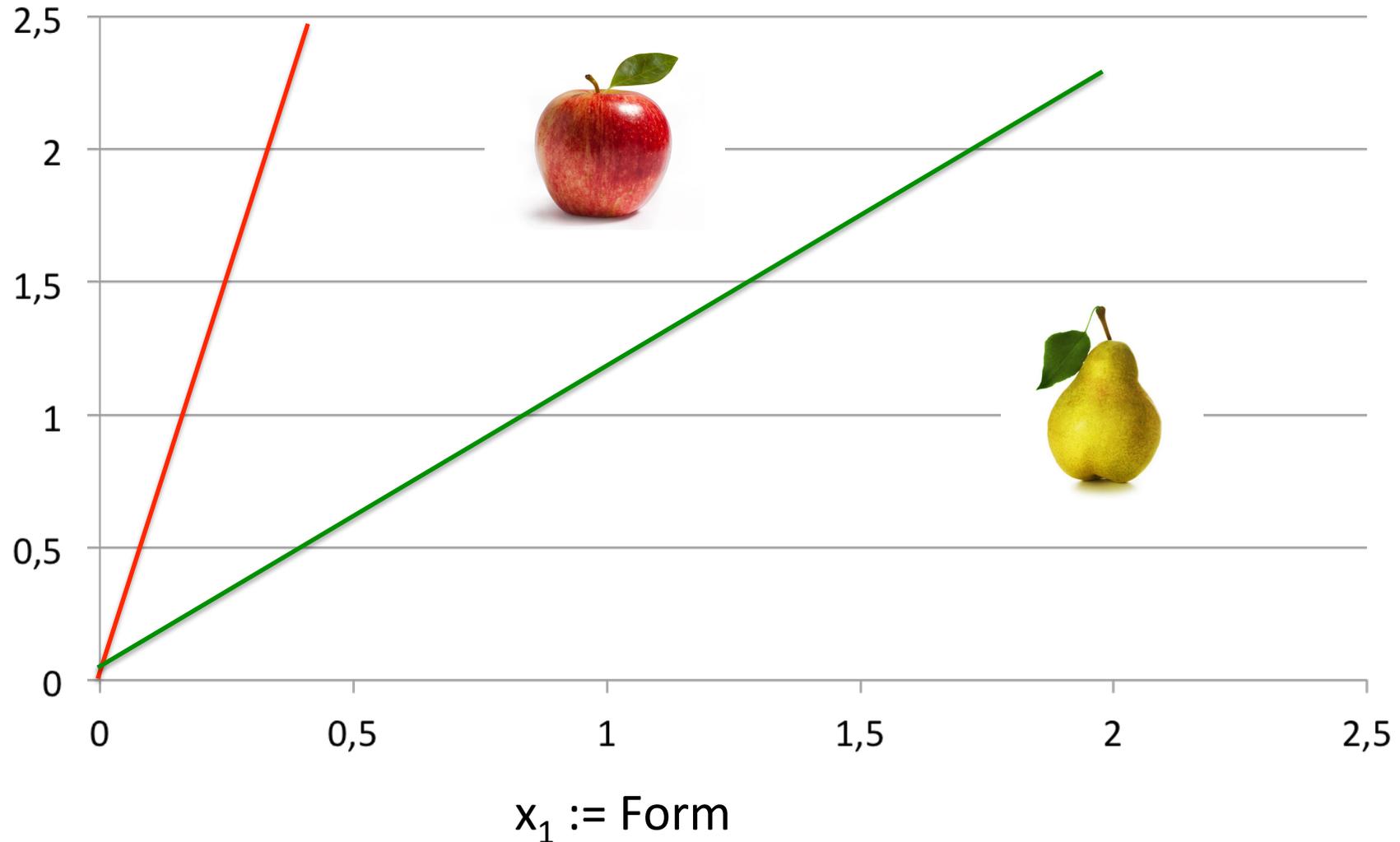
- Kulturleistung und Strukturwissenschaft
- Logik, Boolesche Algebra, Gruppentheorie, Lineare Algebra, Analysis, Dynamische Systeme, ...
- „computational theory“ := Algorithmus
- **Definition** (Kurzform): „Endliche Zahl elementarer und wohldefinierter Schritte einer Berechnung, die bei zulässiger Eingabe zu einem eindeutigen und korrekten Ergebnis kommt.“ (vgl. Turingmaschine)
- **Konsequenz:** elementare/repetitive Tätigkeiten bzw. Handlungen sind im Prinzip algorithmisierbar

# ... und „Maschinelles Lernen“?

- **Maschinelles = algorithmisches Lernen** (oder: Erkennen)
  - aus Daten (mit bekannter Bedeutung/Semantik = Vorwissen)
  - von Ähnlichkeiten, Typizität, Verallgemeinerung, ..., Mustern, Objekten, Zusammenhängen, ..., (Ir-)Regularität, etc.
- **„Hype“**: Neuronale Netze zur Sprach-/Text-/Bildererkennung
  - Daten als „Vektoren“ -> Klassifikation(sverfahren)
  - Statistik der Daten und Komplexität(sbeherrschung)
- **„Non-hype“**: Logische Kalküle
  - Daten als „Ontologie“ -> Schlussfolgerung(sregeln)  
(vgl. semantisches Netz und Web 2.0)
  - Modellierung des Vorwissens und Komplexität(...)

# ML: Äpfel & Birnen

$x_2 := \text{Farbe}$       Daten als Punkte  $\mathbf{x}_i = (x_1, x_2)^T$  der Ebene (Vektoren)



# Konklusio (auf die Schnelle)

- **„DIE künstliche Intelligenz als ELWMS gibt es nicht!“**
  - große Vielfalt an Modellen und Methoden
  - nicht immer verstandene Grenzen der Theorie
  - hohe Anforderungen an Datenqualität
  - nicht immer gewährleistete Machbarkeit (Echtzeitfähigkeit, Sicherheit, etc.)
  - notwendige Beherrschbarkeit in der Anwendung/Nutzung
- Ergo: **„Dem Inscheniör ist nichts zu schwör!“** (Erika Fuchs)
  - klare Zielsetzungen und Anforderungen
  - begründete Auswahl von Modellen/Methoden/Werkzeugen
  - gesicherte Qualität der Daten
  - standardisierte Testverfahren und Qualitätsprüfung („TÜV“)
  - effektiver/effizienter Einsatz in Anwendungen mit Mehrwert

## Zitat (2)

„**Textdeutungen** auf Knopfdruck können Computer **bislang nicht** liefern. ... Dem Computer sind Sinn und Bedeutung gleichgültig.

Sinnerfüllte und historisch informierte Textdeutungen lassen sich **ebenso wenig** maschinell erzeugen wie zwischenmenschliche Erfahrungen und ausgewogene Urteile.

Algorithmen für Einordungsvermögen, ästhetischen Sinn und textsensible Deutungen sind **nicht** in Aussicht. ...

**Den Maschinen sollten wir die Chance geben, uns dabei zu unterstützen – mehr jedoch nicht.“**

*(Prof. Dr. Sandra Richter, seit Januar 2019 Direktorin des Deutschen Literaturarchivs Marbach; in: SZ, 20. April 2019, S. 6)*

## Famous Last Words



- Keep-the-Scholar-in-the-Loop-and-in-Control (UI/UX)
- Visual Programming for Joint Experimentation
- Agility and Rapid Prototyping for Workflow Design/Support
- Platform Independence and Interoperability
- Web-based Platform with OpenX (X=: Data, Methods, Tools, Services)

