



Wenn Computer Meinung machen

LMS Mediengespräche

23.3.2017

Prof. Dr. Katharina A. Zweig

Als das Rennen noch offen war....



Vor der Wahl: „Google manipuliert für Hillary“



- SourceFed behauptete, negative Suchvervollständigungen würden von Google bei Ms Clinton unterdrückt¹.
- Psychologe Prof. Dr. Epstein sah massive Manipulation bei Google und betonte wiederholt, dass bis zu 20% der unentschlossenen Wähler durch Suchmaschinen manipuliert werden können².

¹ Das Originalvideo ist nicht mehr erhältlich (!, 23.3.2017), hier ist eine gekürzte Fassung: https://scontent-frt3-1.xx.fbcdn.net/v/t42.1790-4/13418283_277976322591285_62218208_n.mp4?efg=eyJ2ZW5jb2RlX3RhZyI6InN2ZV9zZCJ9&oh=417e220b5ef8a001cd2a38641aaa783b&oe=58D3C44E

² Zum Beispiel hier, zu demselben Thema: <https://sputniknews.com/us/201609121045214398-google-clinton-manipulation-election/>

Nach der Wahl:



- Ach nee.
- Doch nicht.
- Dann war's Facebook:
- Filter bubble und Echo chambers werden algorithmisch erzeugt, so dass wir in einer postfaktischen Welt leben¹.



Können uns Algorithmen
manipulieren?



Das kleine ABC der Informatik

Gefährden

Algorithmen,

Big Data und

Computerintelligenz

unsere Medienvielfalt und Meinungsfreiheit?



A wie Algorithmus

Ein Algorithmus ist ein Problemlöser

Problem



INPUT

**Der OUTPUT
der uns sagt,
wie Input
mit Output
zusammenhängt.**



OUTPUT

Input: By User:Bluemoose - Own work, [CC BY-SA 3.0](#)

Putput: By Yann (talk) - Own work, GFDL

Output: [CC BY-SA 3.0](#)

Ein Algorithmus ist...



...eine für jede **erfahrene Programmiererin** und jeden erfahrenen Programmierer **ausreichend detaillierte Lösungsvorschrift**, so dass bei **korrekter Implementierung** der Computer **für jede korrekte Inputmenge den korrekten Output** berechnet – in endlicher Zeit.



Beispiel

Problem: Sortieren



Sortieren 1: „Sortieren durch Einfügen“

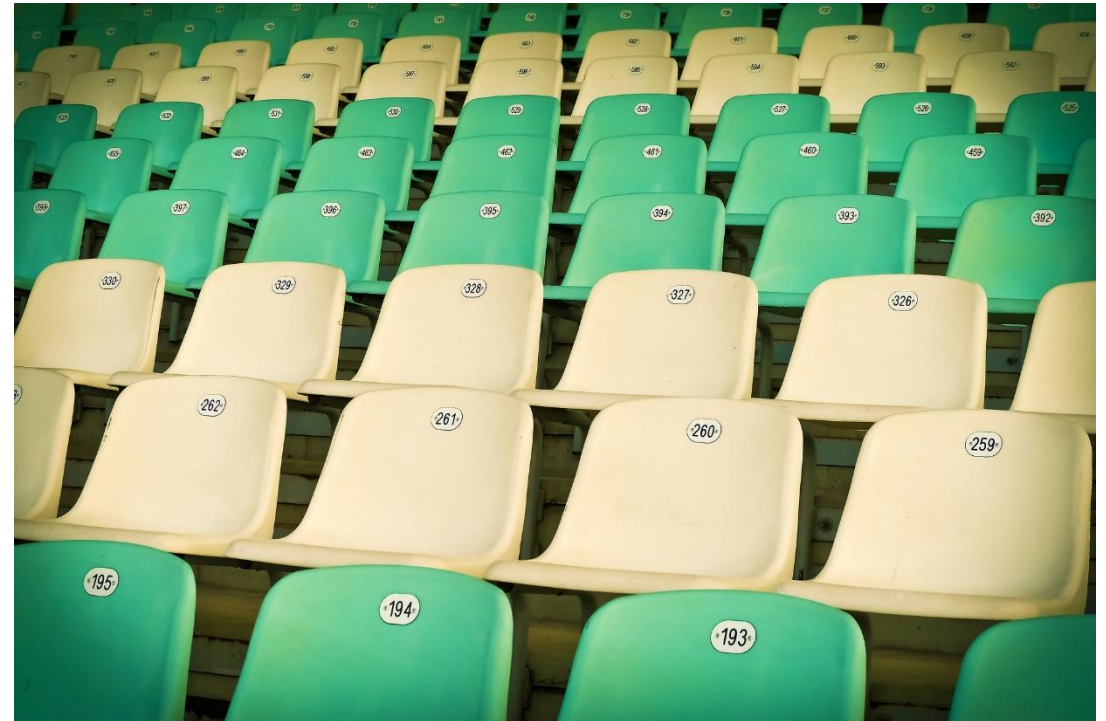


- Hier: Karten mit Zahlen darauf.
- Möglicher Algorithmus, um Zahlen und Besitzer(in) pro Reihe zu sortieren:
 - Alle stehen auf.
 - Eine oder Einer setzt sich hin.
 - Solange noch jemand steht,
 - Geht eine davon an der Reihe entlang und identifiziert die Lücke. Falls rechts davon jemand sitzt, müssen diese Personen alle eins nach rechts weiterrücken.
- Korrektheit: Alle Personen, die schon sitzen, sind bezüglich der ihnen zugeordneten Zahl in der richtigen, relativen Reihenfolge.
- Daher: wenn alle sitzen, sind die Zahlen vollständig sortiert.

Sortieren 2: Aufsteigendes Sortieren



- Alle bleiben sitzen.
- Alle überprüfen immer wieder, ob der linke Nachbar eine kleinere Zahl oder der rechte Nachbar eine größere Zahl hat.
 - Wenn ja, dann bitte tauschen.
 - Nur die Zahlen tauschen, nicht die Sitzplätze!
- Wenn kein Tausch mehr nötig ist, sind alle Karten sortiert.



Manipulationsmöglichkeiten?



Sortierungen können explizit manipuliert werden durch die Designer des Algorithmus.

Klare Verantwortlichkeit:
Designerteam





Komplexe Algorithmen

Beispiel: Suchmaschinenalgorithmen

Suchmaschinen 101



1. Filtern aus allen ihnen bekannten Webseiten diejenigen, deren Text mit den angegebenen Suchbegriffen zusammenhängen.
2. Sortieren diese anhand:
 - Der Vernetzungsstruktur der Seiten untereinander
 - Dem Clickverhalten anderer Nutzer und Nutzerinnen bezüglich derselben Suche
 - Bei Personalisierung: auch nach dem eigenen, bisherigen Suchverhalten

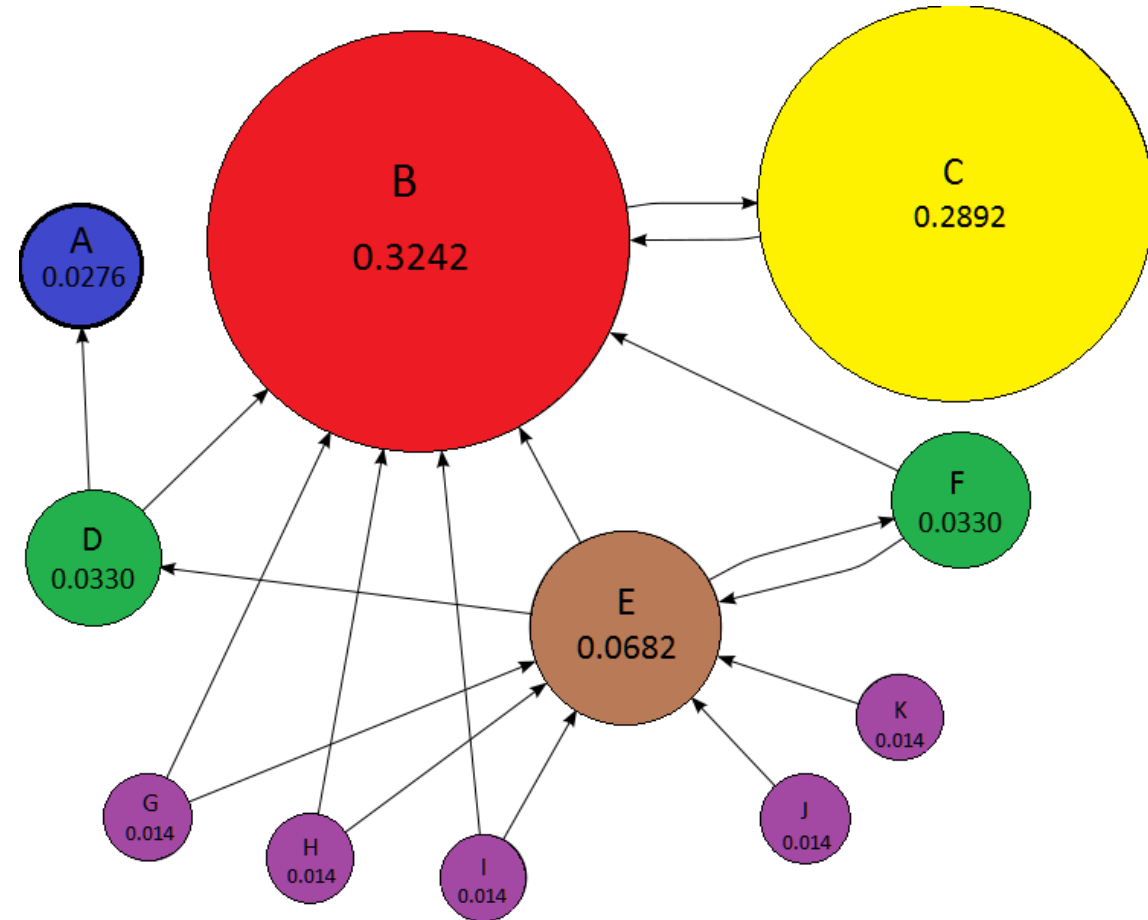
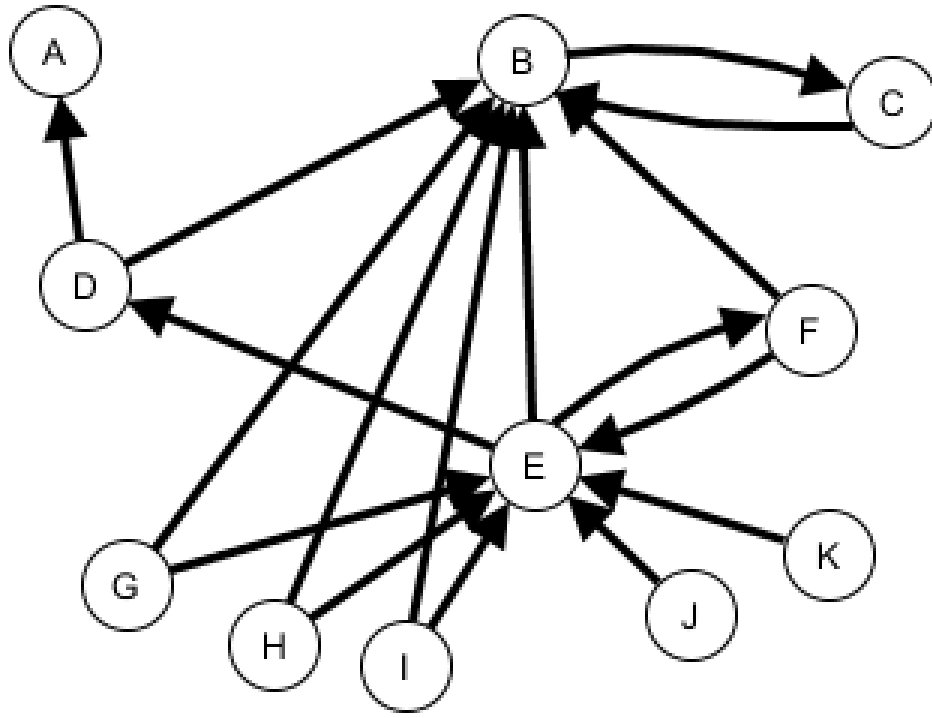
Idee hinter dem Algorithmus



Ein Modell menschlichen Verhaltens: der Random Surfer

- Ein Surfer klickt auf eine Webseite
- Folgt einem der Links auf der Webseite zufällig
- Von Zeit zu Zeit springt er auf eine völlig neue Webseite
 - Modelliert externes Wissen (z.B. Werbung, bekannte Seiten)

PageRank



Manipulationsmöglichkeiten?



- Manipulationsmöglichkeiten:
 - Gezielte Bevorzugung von Webseiten
 - Z.B. eigene Produkte
 - Auch in Abhängigkeit von Inhalt **prinzipiell** möglich
 - +10 für positive Texte über Clinton und negative Texte über Trump
 - -10 für negative Texte über Clinton und positive Texte über Trump
- Wie in jedem guten Krimi:
 - Wo wäre das Motiv?
 - Wie hoch sind die Kosten bei einer Entdeckung, z.B. durch einen Leak?



Externe Manipulationsmöglichkeiten?



- Gibt nur dann relevante Ergebnisse, wenn Webseiten
 - Links auf ähnliche Seiten wie ihre eigene setzen,
 - Links auf relevante, meinungsangebende Seiten setzen, und
 - ihre Links **unabhängig** voneinander setzen.
- Unter dieser Bedingung ist der Algorithmus neutral und gibt das kollektive Wissen der Welt nutzbringend weiter.
- Die Veröffentlichung des Algorithmus führte prompt zu Manipulationen seitens der Webseitenbetreiber.
 - Das ehemals **authentische** Signal war keines mehr.



Das nächste authentische Signal...



Könnte nicht menschliche Aufmerksamkeit wieder ein authentisches Signal bieten für die Relevanz von Webseiten?

Wie kann man Relevanz messen?

Wie kann man sie vorhersagen?





B wie Big Data

Digitales Verhalten

Big Data



- Wie kann Relevanz von Webseiten „quantifiziert“ werden?
- Big Data Methoden nutzen, z.B.:
 - Sprache der Anfrage, Niveau der Anfrage, Wörter, Wortkombination
 - Tageszeit und geographische Informationen, Gerätetyp
 - Ihre bisherigen Suchanfragen und Ihr persönliches Klickverhalten
 - Welche Seiten wurden angeklickt, wie lange betrachtet, kam die Nutzerin wieder zurück zu den Ergebnissen?
 - Metadaten der Nachrichten/Medien: wann erstellt, durch wen, wo publiziert, Verschlagwortung, Wahl der Wörter
 - Verhalten anderer Nutzer, „likern“ auf sozialen Netzwerken, Interaktion mit Beiträgen



C wie Computerintelligenz

Lernende Algorithmen



Big Data + Frage



- Die Algorithmen designerinnen und -designer müssen nun entscheiden, welche der Daten vermutlich mit „Relevanz“ korrelieren.
- Dies sollte am besten in einer einzigen Zahl pro Medium/Nachricht/Webseite münden, so dass man direkt sortieren kann.
- Beispiel Formel:
 - 3 * bisherige Zugriffe
 - Anzahl Tage seit Publikation
 - + Beliebtheitsquotient des Publikationsortes
 - + Beliebtheitsquotient des Verfassers + ...

Allgemein



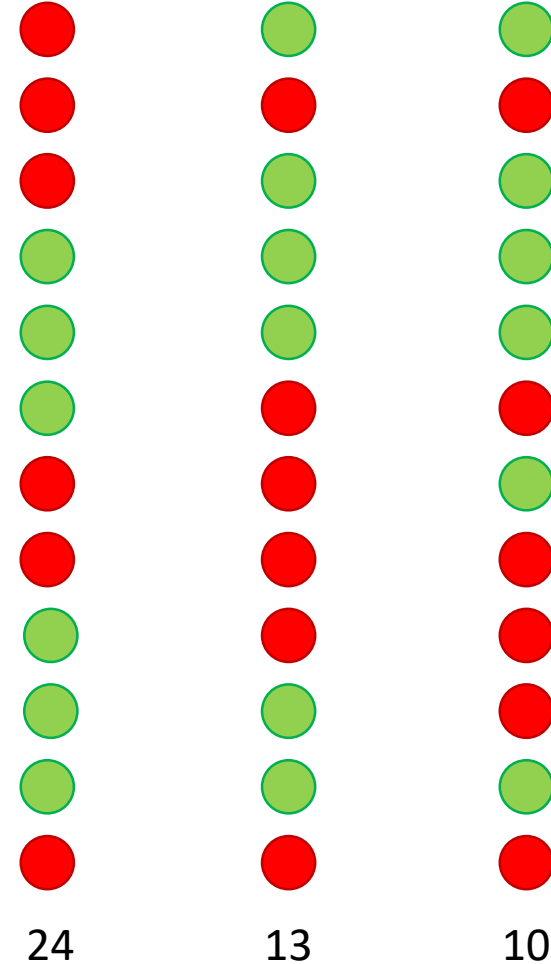
$$\begin{aligned} & w_1 * \text{bisherige Zugriffe} \\ + & w_2 * \text{Anzahl Tage seit Publikation} \\ + & w_3 * \text{Beliebtheitsquotient des Publikationsortes} \\ + & w_4 * \text{Beliebtheitsquotient des Verfassers} + \dots \end{aligned}$$

- Wer bestimmt diese Gewichte, so dass insgesamt die „relevantesten“ (also die, die im Nachhinein am öftesten angeklickt werden) am weitesten nach oben sortiert werden?
- Dazu bedarf es Algorithmen der künstlichen Intelligenz.

Lernen von Gewichten



- Algorithmus probiert Gewichte
- Bewertet jeweils, wieviele angeklickte Inhalte (in rot) möglichst weit oben stehen.
- Die Gewichtung, die das maximiert, wird für weitere Daten genommen.
- Kann im Wesentlichen für alles verwendet werden:
 - News Feed bei Facebook
 - Suchmaschinen
 - Produktempfehlung

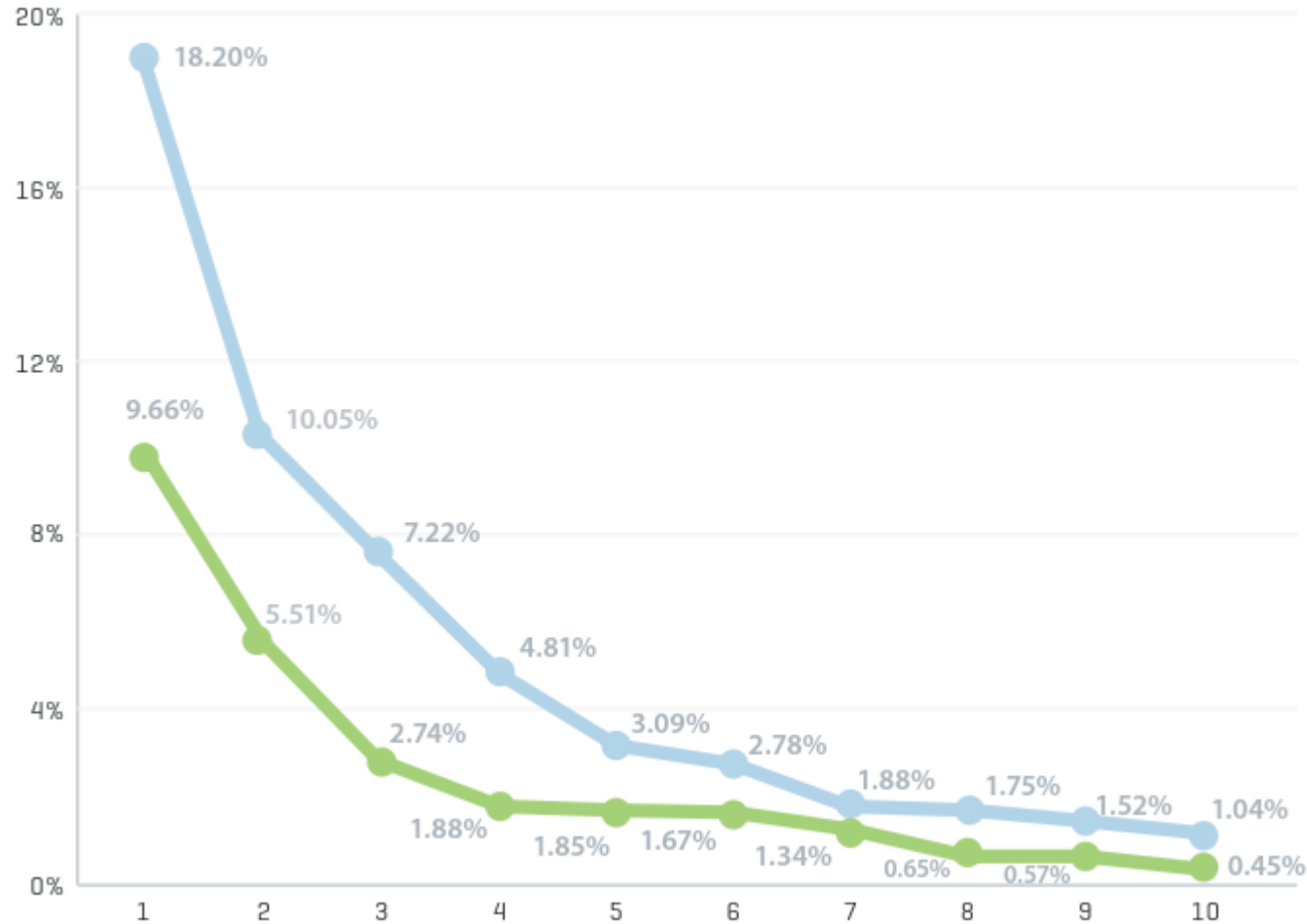




Relevanz von Algorithmen

Vom Wert, der Erste zu sein

GOOGLE VS BING CLICK-THROUGH RATE



Paul Davison at Digital Relevance™: „A Tale of Two Studies: Establishing Google & Bing Click-Through Rates“, Study by Digital Relevance™ using client data from Jan-June 2011, available from http://connect.relevance.com/a-tale-of-two-studies-establishing-google-bing-click_through-rates or research@relevance.com; published 2013.

Manipulationsmöglichkeiten



Intern

- Zusätzlicher Input, der die Gewichtungen verändert
 - Z.B. explizite Verschiebung im Ranking für gewisse Webseiten
- Wieder:
 - Gibt es ein Motiv, das die möglichen Kosten eines Leaks rechtfertigt?

Extern

- Social bots, die sich wie Menschen im Netz verhalten und Interesse für Seiten simulieren.
- Möglich durch Bot Nets (gekaperte Computer durch Viren, etc.).
- Motive:
 - Geld
 - Destabilisierung von Demokratien

Oder waren es ...



- BuzzFeed schreibt, dass ein deutlicher Anteil der Nachrichten auf hyperkonservativen Webseiten falsch oder eine Mischung aus falschen und korrekten Nachrichten ist.
- Ein Teil dieser „Hyperpartisan“-Webseiten stammt aus....
- ...Mazedonien!

Motivation



„Die meisten Texte auf diesen Seiten sind zusammengeschrieben oder einfach abgeschrieben von rechtslastigen Webseiten der USA. Die Mazedonier sahen irgendwo eine Story, versetzten sie mit einer aufgeheizten Überschrift, und veröffentlichten es dann schnell auf ihrer Webseite. Dann teilen sie die Nachricht auf Facebook und veranlassen damit Leute, auf ihre Seite zu kommen. Je mehr Leute sich verführen lassen, desto mehr Geld verdienen sie durch Werbung auf ihrer Webseite.“

Craig Silverman und Lawrence Alexander: „**How Teens In The Balkans Are Duping Trump Supporters With Fake News**“, BuzzFeed Nov. 4th, downloaded on the 27th of November, 2016
https://www.buzzfeed.com/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo?utm_term=.wvzZe7D5#.he3BEIYV

Austrocknen des Sumpfes



- Während Werbetreibende vorher passende Journale gesucht haben, um Werbung zu platzieren (Demographie, Inhalt, ...),
- ist jetzt nur noch die Demographie der Nutzerinnen und Nutzer ausschlaggebend (z.B. jung, weiblich, Single).
- Damit war zeitweise die Umgebung, in der die Werbung lief, unerheblich (und sehr schwer nachzuvollziehen für Werbende)
- Es werden Seiten finanziert, die viele Klicks haben, sich aber nicht durch Qualität auszeichnen.
- Hier muss ein neuer Anreiz gesetzt werden.
- Z.B. Twitter-Kampagne „Kein Geld für Rechts“: Bewusstmachung, wer durch Werbung finanziert wird.

<http://www.spiegel.de/netzwelt/web/breitbart-kritik-an-deutscher-werbung-auf-hetz-seite-a-1124526.html>

Manipulation durch Algorithmen?



- Interne Manipulation durch Software-Entwickler
 - Nicht auszuschließen
 - Als Teil einer Firmenpolitik aber mit hohen Risiken behaftet
- Externe Manipulation
 - Hochwahrscheinlich
 - Sowohl Motiv als auch Möglichkeit und Gelegenheit dazu da
 - Niedriges Risiko der Entdeckung

Müssen wir uns Sorgen machen?



- Pro: Studie von Hagen, In der Au und Wieland zeigt, dass Polarisierung bei Personen ohne Abitur mit Intensität der Nutzung sozialer Medien steigt¹.
- Contra: Neue Studie des Hans-Bredow-Instituts (Schmidt et al., 2017) zeigt, dass Mediennutzung in Deutschland sehr divers – Einfluss der sozialen Medien auf Meinungsbildung eher gering².
- Google Suchvervollständigung
 - Im Wesentlichen gibt es keinen substantziellen Verdacht, dass Google für Clinton manipuliert hat³.

¹ Lutz Hagen, Anne-Marie In der Au und Mareike Wieland: „Polarisierung im Social Web und der intervenierende Effekt von Bildung“, erscheint im März 2017 in der Sonderausgabe „Algorithmen, Kommunikation und Gesellschaft“ von kommunikation@gesellschaft (www.kommunikation-gesellschaft.de)

² Jan-Hinrik Schmidt et al.: „Relevanz einzelner Medienangebote und digitaler Dienste für die Meinungsbildung“, Arbeitspapier des Hans-Bredow-Instituts Nr. 40, 2017 <http://www.hans-bredow-institut.de/de/node/6330>

³ Zum Beispiel Rhea Drysdale auf ihrem Blog: <https://medium.com/@rhea/hillary-clintons-search-results-manipulated-by-sourcefed-not-google-3dd9a5c68ca1#.v4c9tidnj>

Fazit



- Momentan ist die Lage in Deutschland vermutlich nicht besorgniserregend
 - Weitere Grundlagenforschung ist wichtig
- Es ist aber damit zu rechnen, dass der Anteil der Nutzer, die Nachrichten über soziale Netzwerke beziehen, weiter steigt.
- Die aufgebaute Infrastruktur und die Monopolstellung der wichtigen Intermediäre ermöglicht interne und externe Manipulation der Algorithmen und damit Einflussnahme auf Nutzer.
- Vielleicht gibt es ein Recht einer Gesellschaft auf Intransparenz von Algorithmen?
 - Hier beißt sich natürlich die Katze in den Schwanz.
- Auf jeden Fall braucht die Gesellschaft Kontrolle über einige Algorithmen und algorithmische Entscheidungssysteme.



Algorithmen in einer demokratischen Gesellschaft

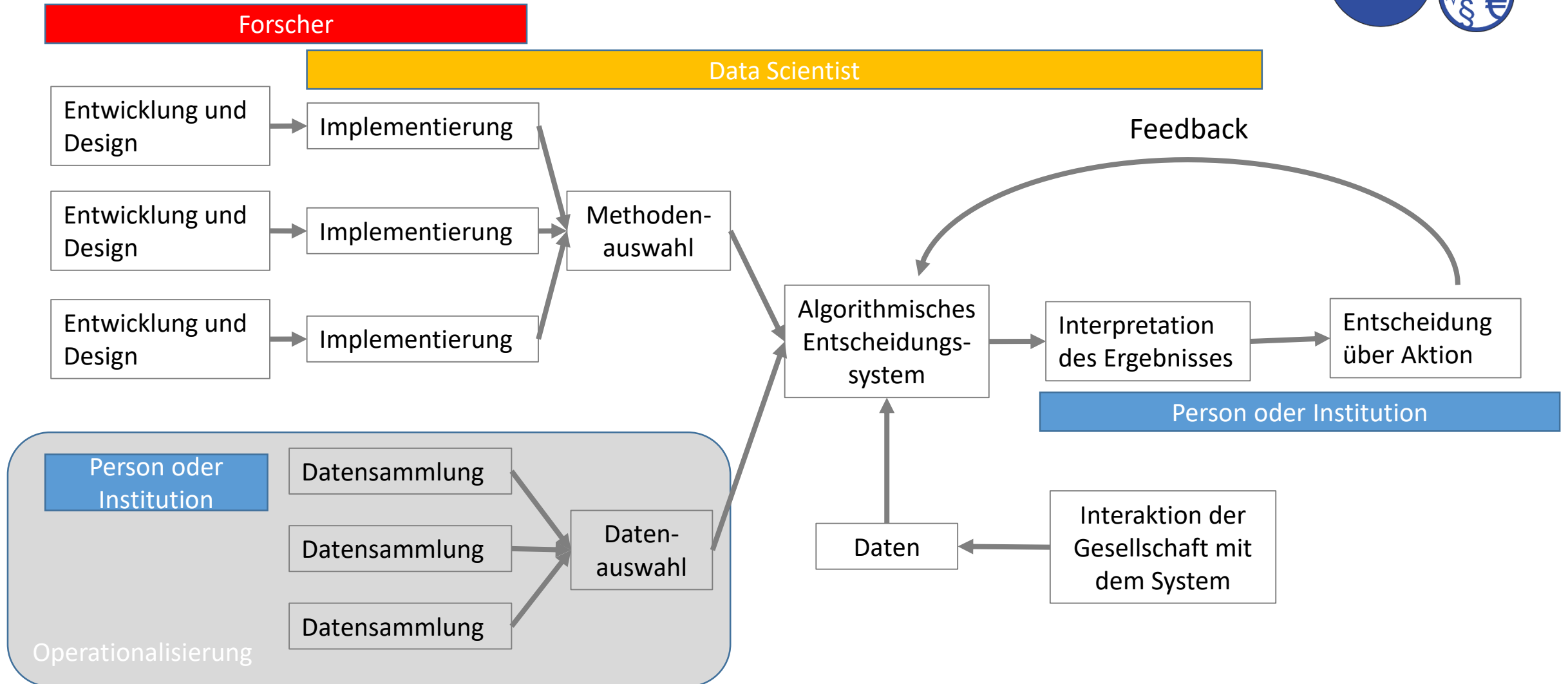
Algorithmische Entscheidungssysteme



- Empfehlungssysteme (wie der NewsFeed) oder Suchmaschinenalgorithmen gehören zu den **algorithmischen Entscheidungssystemen**.
- Weitere Algorithmen gefährden gesellschaftliche Teilhabe:
 - Kreditwürdigkeit, autom. Leistungsbewertung, Bewerbungsortierer, Predictive Policing, Rückfälligkeitsvorhersagealgorithmen, Terroristenidentifikation



Entwicklungsprozess von Entscheidungssystemen



Gründung von „Algorithm Watch“



ALGORITHM
WATCH



Lorena Jaume-Palasi, Mitarbeiterin im iRights.Lab



Lorenz Matzat, Datenjournalist der 1. Stunde, Gründer von lokaler.de, Grimme-Preis-Träger



Matthias Spielkamp, Gründer von iRights.info, ebenfalls Grimme-Preis-Träger, Vorstandsmitglied von Reporter ohne Grenzen.



Prof. Dr. K.A. Zweig, Junior Fellow der Gesellschaft für Informatik, Digitaler Kopf 2014, TU Kaiserslautern

Beipackzettel für Algorithmen



Welches Problem „kuriert“ der Algorithmus?

Was ist das Einsatzgebiet des Algorithmus, was seine Modellannahmen?

Welche „Nebenwirkungen“ hat der Algorithmus?

Schlussformel



... zu Risiken und Nebenwirkungen der Digitalisierung befragen Sie bitte Ihren nächstgelegenen Data Scientist oder den deutschen Algorithmen TÜV.