



Algorithmen und Teilhabe

Lunch Talk in der Bertelsmann-Stiftung
22.11.2016

Prof. Dr. Katharina A. Zweig



Das kleine ABC der Informatik

Wann gefährden

Algorithmen,

Big Data und

Cünstliche Intelligenz

unsere Demokratie?

Wo ermöglichen sie, wo verhindern sie Teilhabe?



A wie Algorithmus

Ein Algorithmus ist ein Problemlöser

Mathematisches Problem



INPUT

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



OUTPUT

1. Beispiel: Navigation



Navigation

Gegeben das Kartenmaterial und weitere Daten, berechne die kürzeste Route zwischen Start und Ziel

Das **Problem** sagt nicht, wie man die Lösung **findet**.



Input: Straßen, Länge, Staus, ...
Start und Ziel



Output: optimale Route

Ein Algorithmus ist...



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



2. Beispiel: Sortieren

Problem: Sortieren



Sortieren 1: „Sortieren durch Einfügen“



- Fange mit einem Buch an, stelle es ins Regal.
- Solange es noch Bücher gibt,
 - nimm das nächste,
 - geh am Regal entlang und sortiere es an der passenden Stelle ein.
- Alle Bücher, die schon im Regal stehen, sind in der richtigen, relativen Reihenfolge.
- Daher: wenn alle im Regal stehen, sind sie vollständig sortiert.

Sortieren 2: Aufsteigendes Sortieren



- Stelle alle Bücher irgendwie ins Regal.
- Gehe das Regal entlang – wenn dabei zwei Bücher in der falschen Reihenfolge nebeneinander stehen, vertausche sie. Tue dies bis zum Ende des Regals und gehe wieder zum Anfang.
- Laufe solange immer wieder am Regal entlang, bis im letzten Durchgang kein Tausch mehr nötig war.
- Wenn kein Tausch mehr nötig war, sind alle Bücher sortiert.

Alle Sortierprobleme auf einen Schlag



- Gegeben eine Menge von Objekten oder Subjekten...
- ... und ein Sortierkriterium, das für je zwei von diesen besagt, welches nach links, welches nach rechts sortiert werden muss,...
- ... kann jeder beliebige Sortieralgorithmus die korrekte Lösung berechnen.

- Eine Interpretation der Ergebnisse (dies sind die relevantesten Nachrichten, die wichtigsten Freunde, die kaufenswertesten Produkte) liefert er **nicht**.



Zusammenfassung Algorithmen



Algorithmen...

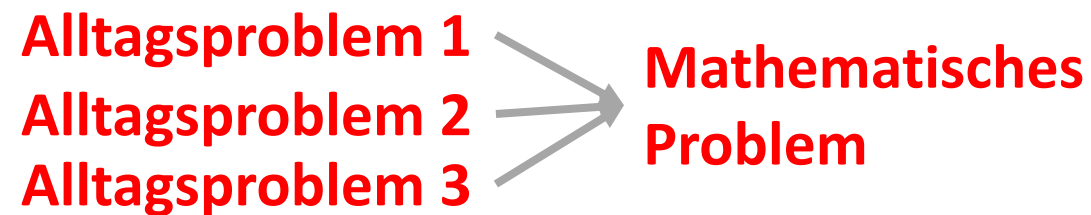
- ...können aus nur einer Formel bestehen,
- aber meistens sind sie eine komplexe Mischung aus Handlungsanweisungen und Berechnungen,
- die genau eine Lösung für das Problem basierend auf den Eingabedaten berechnen.
- Es sind eingefrorene Handlungsanweisungen, von Menschen erdacht, um mathematische Probleme durch Computer zu lösen.
- Das Ergebnis **muss interpretiert** werden.

```
var simple = type.slice(
  forward = type.slice(
    ofType = what === "of-type";
  );
  return first === 1 && last === 0 ?
  // Shortcut for :nth-*(n)
  function( elem ) {
    return !elem.parentNode;
  } :
  function( elem, context, xml ) {
    var cache, outerCache, node, diff, nodeIndex, start,
        dir = simple !== forward ? "nextSibling" :
            "previousSibling",
        parent = elem.parentNode,
        name = ofType && elem.nodeName.toLowerCase(),
        useCache = !xml && !ofType;
    if ( parent ) {
      ...
    }
  }
}
```





Die Zuordnung einer Frage zu einem mathematischen Problem bezeichnet man als Modellierung.



„Harmlose“ Algorithmen



- Die meisten Algorithmen sind harmloser Natur:
 - Suchen, Speichern, Filtern, Sortieren
 - Logistik, Scheduling
 - Bilderkennung in Industrierobotern
 - Verwaltungssoftware
 - Bürosoftware
 - ...
- Hier ist der **Abstraktionsgrad der Modellierung** gering, Fehler werden schnell entdeckt.
- Algorithmen mit (noch) geringer Reichweite, z.B. von Startups
- Die Algorithmen, die Grundrechte von Bürgern und die Grundpfeiler der Demokratie gefährden, haben besondere Eigenschaften.



B wie Big Data

Wie sagt man die Rückfallrate eines Verbrechens voraus?



Predictive Policing



Wir haben schon
auf Sie gewartet!



Vorhersagen,
wann und wo
Straftaten
wahrscheinlich
sind.

Predictive Policing



Ein **Algorithmus**
hat mir geflüstert,
dass Du **fast** ein Krimineller bist.
Dann komm mal mit!

Aber auch: Vorhersagen,
ob ein Individuum
straffällig werden könnte!

Beispiel USA:

- 1) Oregon
- 2) Andere Bundesstaaten



Sozio-
matik

Big Data



- Big Data Methoden nutzen, z.B.:
 - Alter der ersten Verhaftung
 - Alter des Delinquenten (der Delinquentin!)
 - Finanzielle Lage
 - Kriminelle Verwandte
 - Geschlecht
 - Art und Anzahl der Vorstrafen
 - Zeitpunkt der letzten kriminellen Akte
 -
 - Aber nicht: die (in den USA eindeutig zugeordnete) ‚race‘.

Algorithmus



- Die Algorithmen designerinnen und -designer müssen nun entscheiden, welche der Daten vermutlich mit „Rückfallwahrscheinlichkeit“ korrelieren.
- Dies sollte am besten in einer einzigen Zahl münden, so dass man direkt sortieren kann.
- Beispiel Formel:

$$\begin{aligned} & 3 * \text{bisherige Verhaftungen} \\ & - 2 * \text{Anzahl Tage seit letzter Verhaftung} \\ & + 3 * (\text{Wenn Mann, dann 1, sonst 0}) \\ & + 2,5 * (\text{Wenn Raubüberfall, dann 1, sonst 0}) + \dots \end{aligned}$$

Allgemein



$$\begin{aligned} & w_1 * \text{bisherige Verhaftungen} \\ - & w_2 * \text{Anzahl Tage seit letzter Verhaftung} \\ + & w_3 * (\text{Wenn Mann, dann 1, sonst 0}) \\ + & w_4 * (\text{Wenn Raubüberfall, dann 1, sonst 0}) + \dots \end{aligned}$$

- Wer bestimmt die Gewichte so, dass möglichst die einen hohen Wert bekommen, die rückfällig geworden sind?
- Dazu bedarf es Algorithmen der künstlichen Intelligenz.



C wie Künstliche Intelligenz

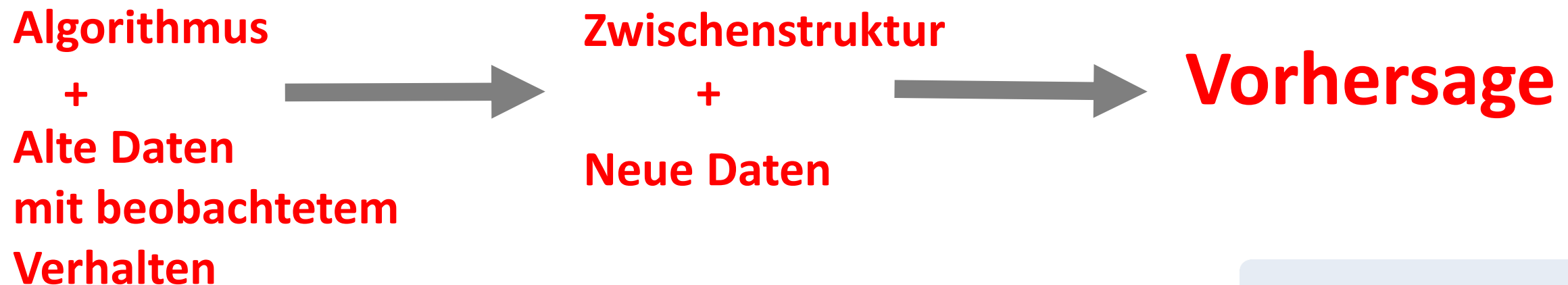
Lernende Algorithmen





Künstliche Intelligenz

- **Problem:** gegeben eine Menge von bekannten Daten, finde Muster, die auf neuen Daten vorhersagen, wie sich etwas oder jemand verhalten wird.
- Algorithmus baut – basierend auf bekannten Daten – eine Zwischenstruktur auf, die dann Vorhersagen für neue Daten generiert.
- Der Algorithmus wird „auf den Daten trainiert“.



Oregon Recidivism Rate Algorithm

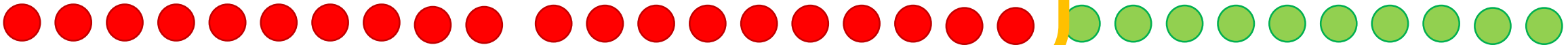
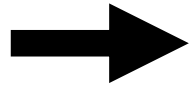


- Das oben genannte Qualitätsmaß dieses Algorithmus: 72 von 100 Paaren werden korrekt sortiert.
- Der in Oregon benutzte Algorithmus hat also, gegeben einen „Rückfall“ und einen „Nichtrückfall“, eine Chance von ca. 1:3 den Rückfall höher zu gewichten als den Nichtrückfall.
- Nur 25% aller so gemachten Prognosen sind falsch!
 - Das klingt doch ganz gut, oder?
- So werden aber keine Urteile gefällt!
- Problem: die Klassen sind ungleich verteilt!
 - 1000 Delinquenten
 - Ca. 2000 werden rückfällig

Optimale Sortierung



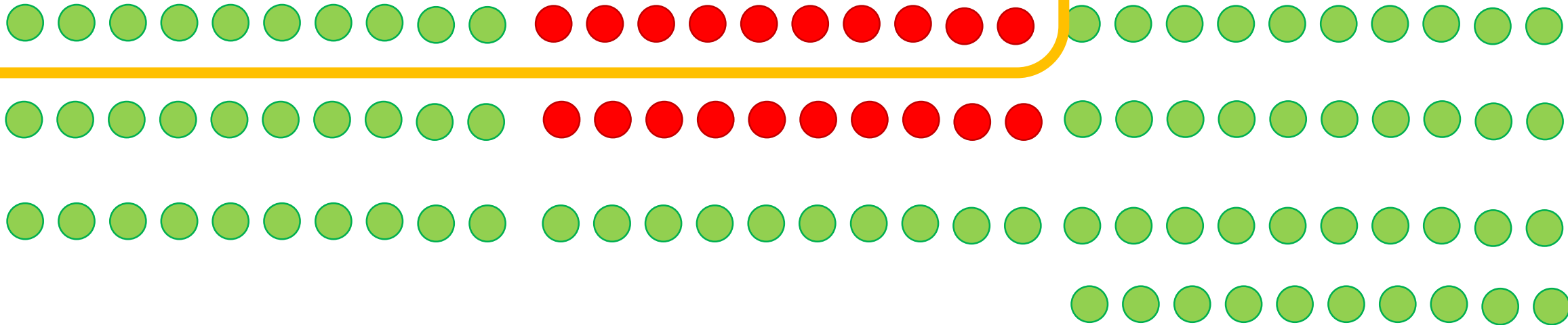
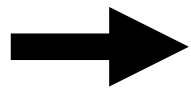
Erwartete 20% „Rückfällige“



Mögliche Sortierung eines Algorithmus mit dieser „Güte“ (75/100 Paaren)



Erwartete 20% „Rückfällige“



Problem: Unbalancierte Klassen



- Bei optimaler Sortierung: die ersten 200 rot – keine Fehlentscheidung.
- Jetzt: nur die Hälfte!
- Damit **50% Fehlentscheidungen**



Rückfallvorhersagealgorithmus ist rassistisch (Propublica)



- In einer Studie von Propublica (anderer Algorithmus) war die Quote noch schlechter:
 - Nur 20% der (vorhergesagten) Gewalttäter begingen eine Straftat
 - Bei allen möglichen Straftaten war die Vorhersage etwas besser als ein Münzwurf.
 - Bei schwarzen Mitbürgern war die Vorhersage immer zu pessimistisch;
 - Bei weißen zu optimistisch.
- Northpoint Software ist eine Firma, der Algorithmus ist unbekannt.
- Rasse ist an sich keine Variable des Algorithmus...

Zweig'sche Regel



Algorithmen der künstlichen Intelligenz werden da eingesetzt, wo es **keine einfachen Regeln** gibt.

Sie suchen **Muster** in hoch-verrauschten Datensätzen.

Die Muster sind daher grundsätzlich **statistischer Natur**.

Versuchen fast immer, eine **kleine Gruppe** von Menschen zu identifizieren (Problem der **Unbalanciertheit**)

Wenn es **einfache Regeln zur Entscheidungsfindung** gäbe, **kennten wir sie schon**.

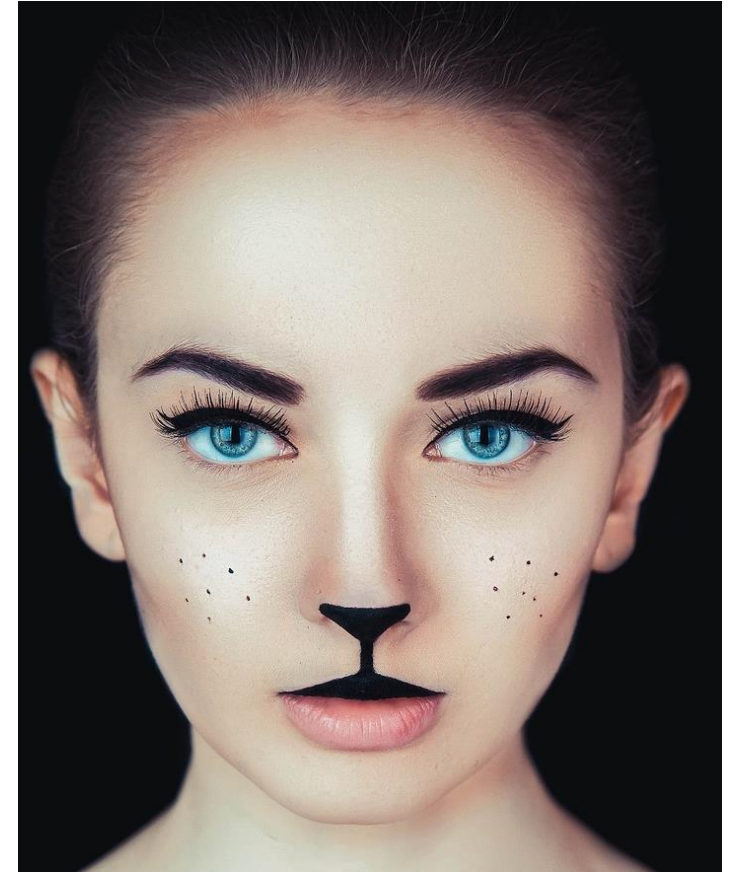


Statistische Vorhersagen über Menschen

Was bedeutet das eigentlich?

Zu 70% ein Krimineller....

- Wenn dieser Mensch eine Katze wäre und 7 Leben hätte, würde er in 5 davon wieder rückfällig werden...
- Nein!
- **Algorithmische Sippenhaftung**
 - Von 100 Personen, die „genau so sind wie dieser Mensch“, werden 70 wieder rückfällig;
 - Mitgefangen, mitgehungen;
 - In einer dem Delinquenten (der Delinquentin) völlig unbekanntem, algorithmisch bestimmten „Sippe“.



Probleme



- Aufmerksamkeitsökonomie der Richter und Richterinnen.
- „Best practice“ erfordert Nutzung der Software.
- Eine Nichtbeachtung der Empfehlung und gleichzeitige Fehleinschätzung wirkt viel schwerer als eine Beachtung der Empfehlung.
- Grundlegende Modellierung und Datenqualität kann schlecht sein.
- Der ins Gefängnis geschickte Delinquent **kann die Vorhersage prinzipiell nicht entkräften!**
 - Dies gilt auch für: Kreditvergaben, Bildungsangebote, Jobs, Personen, die von Drohnen erschossen werden oder als Terrorist eingesperrt werden, ...

Terroristenidentifikation SKYNET



TOP SECRET//COMINT//REL TO USA, FVEY

We've been experimenting with several error metrics on both small and large test sets

Training Data	Classifier	Features	100k Test Selectors		55M Test Selectors	
			False Alarm Rate at 50% Miss Rate	Mean Reciprocal Rank	Tasked Selectors in Top 500	Tasked Selectors in Top 100
None	Random	None	50%	1/23k (simulated)	0.64 (active/Pak)	0.13 (active/Pak)
Known Couriers	Centroid	All	20%	1/18k		
		Outgoing	43%	1/27k		
+ Anchory Selectors	Random Forest		0.18%	1/9.9	5	1
		0.008%	1/14	21	6	

Random Forest trained on Known Couriers + Anchory Selectors:

- 0.008% false alarm rate at 50% miss rate
- 46x improvement over random performance when evaluating its tasked precision at 100

Windows
Wechseln
aktivieren

TOP SECRET//COMINT//REL TO USA, FVEY

<https://theintercept.com/document/2015/05/08/skynet-courier/>

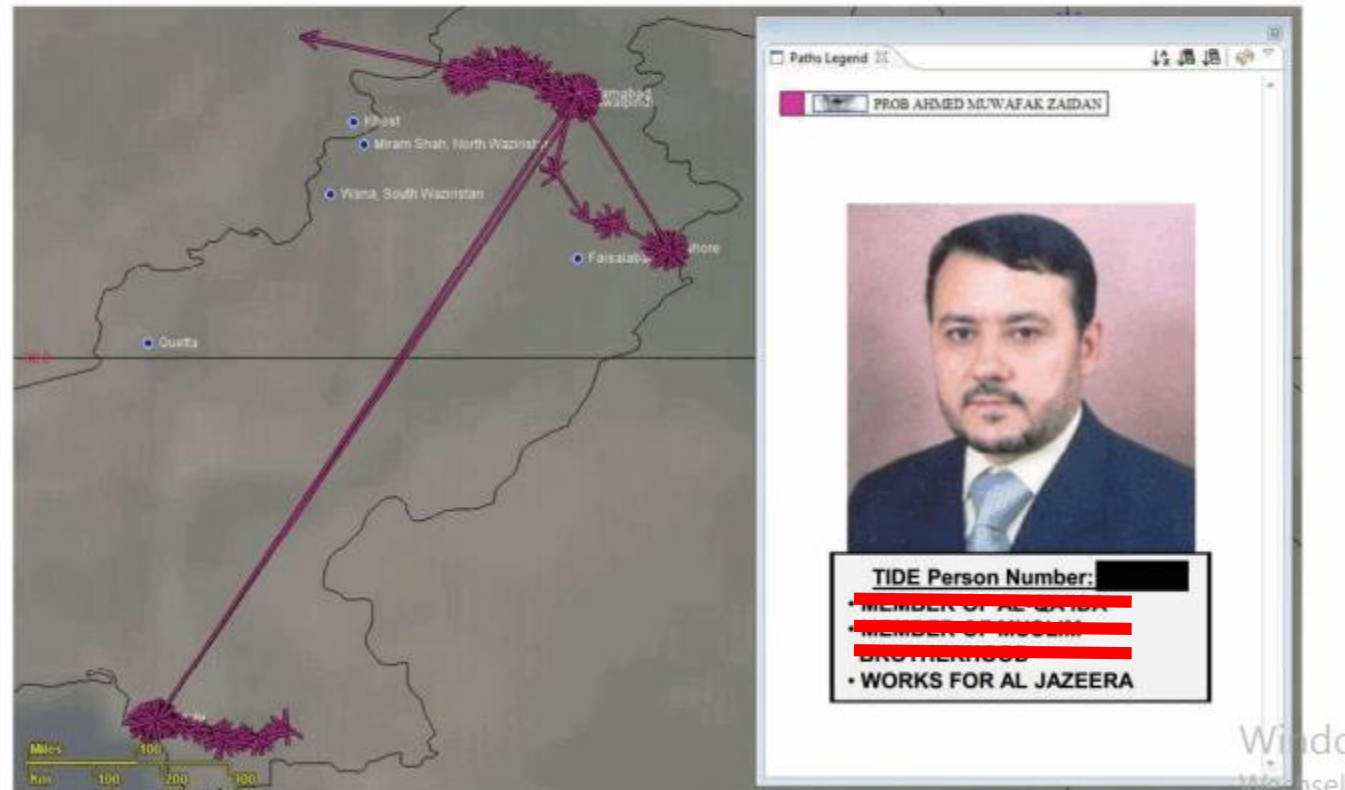
<https://theintercept.com/2015/05/08/u-s-government-designated-prominent-al-jazeera-journalist-al-qaeda-member-put-watch-list/>

Top-“Kurier“ der Terroristen laut Algorithmus ist...



TOP SECRET//COMINT//REL TO USA, FVEY

The highest scoring selector that traveled to Peshawar and Lahore is PROB AHMED Z Aidan





Relevanz von Algorithmen

Gibt es neutrale Empfehlungsalgorithmen in Suchmaschinen, Newsfeed Aggregatoren und sozialen Netzwerken?



- Viele Betreiber behaupten, sie würden Nachrichten **nur** nach Relevanz sortieren.
- Aber: Empfehlungsalgorithmen filtern, lernen und sortieren.
- Sie machen Modellierungsannahmen, wählen Variablen aus und lernen nur aus einer Teilmenge von Daten.
- All diese Schritte können mehr oder weniger gut gelingen, keiner davon ist neutral im Sinne von „objektiv“.

Spielkampsche Regel



**Alle Algorithmen sind objektiv
Bis auf die von Menschen gemachten!**



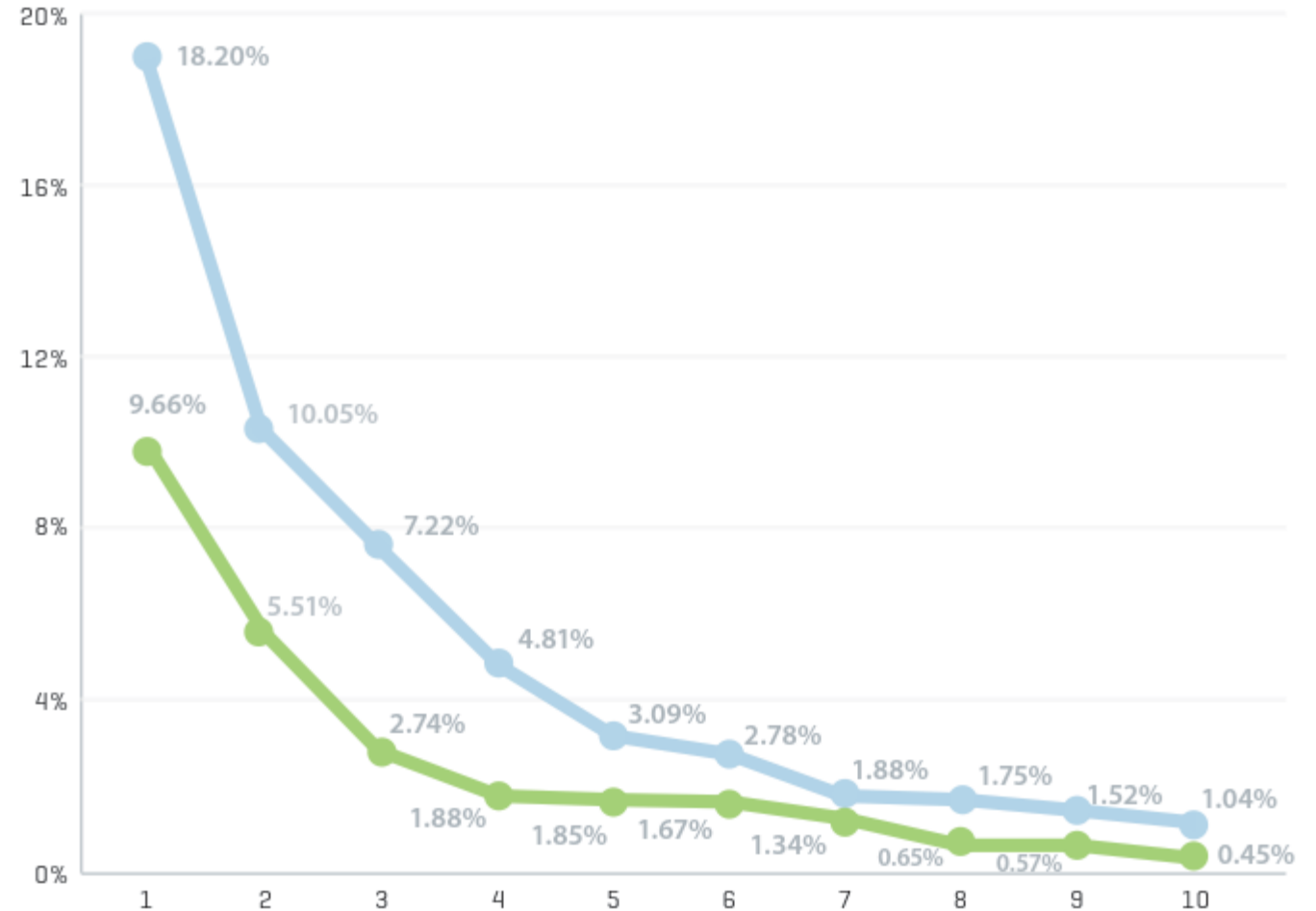
Menschliche Wahrnehmung

Am Beispiel von Suchmaschinen, sozialen Netzwerken und News-Aggregatoren

Vom Wert, der Erste zu sein

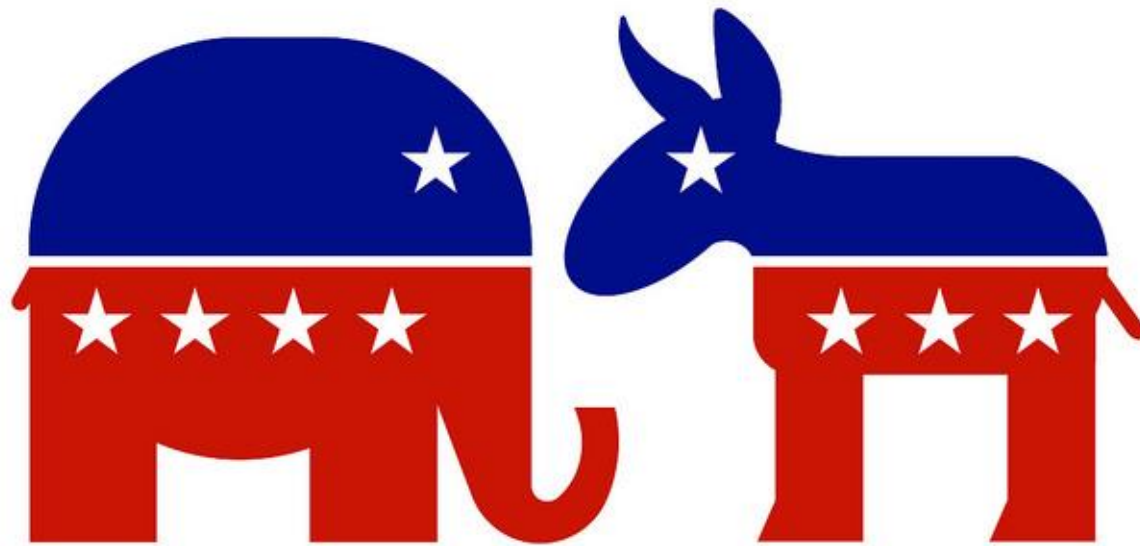
- Der erste bekommt ca. 18,2% (9,7%) aller User, der zweite nur noch 10,1% bzw. 5.51.
- 1. Platz fast doppelt so oft wie der 2., 2,5 mal so oft wie der dritte, fast fünfmal so oft wie der 4.

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.

Bevorzugt Google Demokraten?



DiederHunt

Studie von Trielli, Mussenden und Diakopolous¹:

Unter 16 Präsidentschaftskandidaten (USA) gab es bei Demokraten unter den ersten 10 Suchergebnissen 7 positive Berichte, bei Republikanern nur 5,9.

1 <http://algorithmwatch.org/warum-die-google-suchergebnisse-in-den-usa-die-demokraten-bevorzugen/>

Sind wir beeinflussbar über Algorithmen?



- Suchergebnisreihenfolgen:
 - Manipulierte Suchreihenfolgen werden vom Nutzer nicht bemerkt und können die Tendenz eines unentschlossenen Wähler beeinflussen (Epstein & Robertson, 2015)
- Facebooks „Vote“ bzw. „Ich habe gewählt“-Button
 - Studie von Bond et al. über den Effekt auf das Wahlverhalten.
 - Effekt war klein, aber hochgerechnet ca. 60.000 mehr Wahlstimmen.

Epstein, R. & Robertson, R. E.: "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections", Proceedings of the National Academy of Science, 2015, E4512-E4521

Bond, R. M.; Fariss, C. J.; Jones, J. J.; Kramer, A. D. I.; Marlow, C.; Settle, J. E. & Fowler, J. H.: "A 61-million-person experiment in social influence and political mobilization", Nature, 2012, 489, 295-298

Zuccherosconi



Wollte Zuckerberg der jüngste
Präsident der USA werden –
wer wäre rein technisch noch in
der Lage, als Gegenkandidat über
politische Themen zu punkten?





Algorithmen in einer demokratischen Gesellschaft

Algorithmen und Teilhabe



- Personalisierte, lernende Software als Assistenz erlaubt z.B. mehr Teilhabe für Ältere (Ambient Assisted Living), für Ausländer im deutschen Bildungssystem (z.B. Einblenden von Begriffen in Muttersprache) und Analphabeten.
- Chance für „seltene“ Gruppen nach menschlichen Maßstäben: z.B. Lernbehinderungen, seltene Krankheiten, Hochbegabung.
- Könnte deutlich objektiver sein als menschliche Entscheiderinnen und Entscheider (wenn sie gut gemacht sind, ausreichend unverzerrte Daten vorliegen und sie konstant überwacht werden).

Algorithmen und Teilhabe



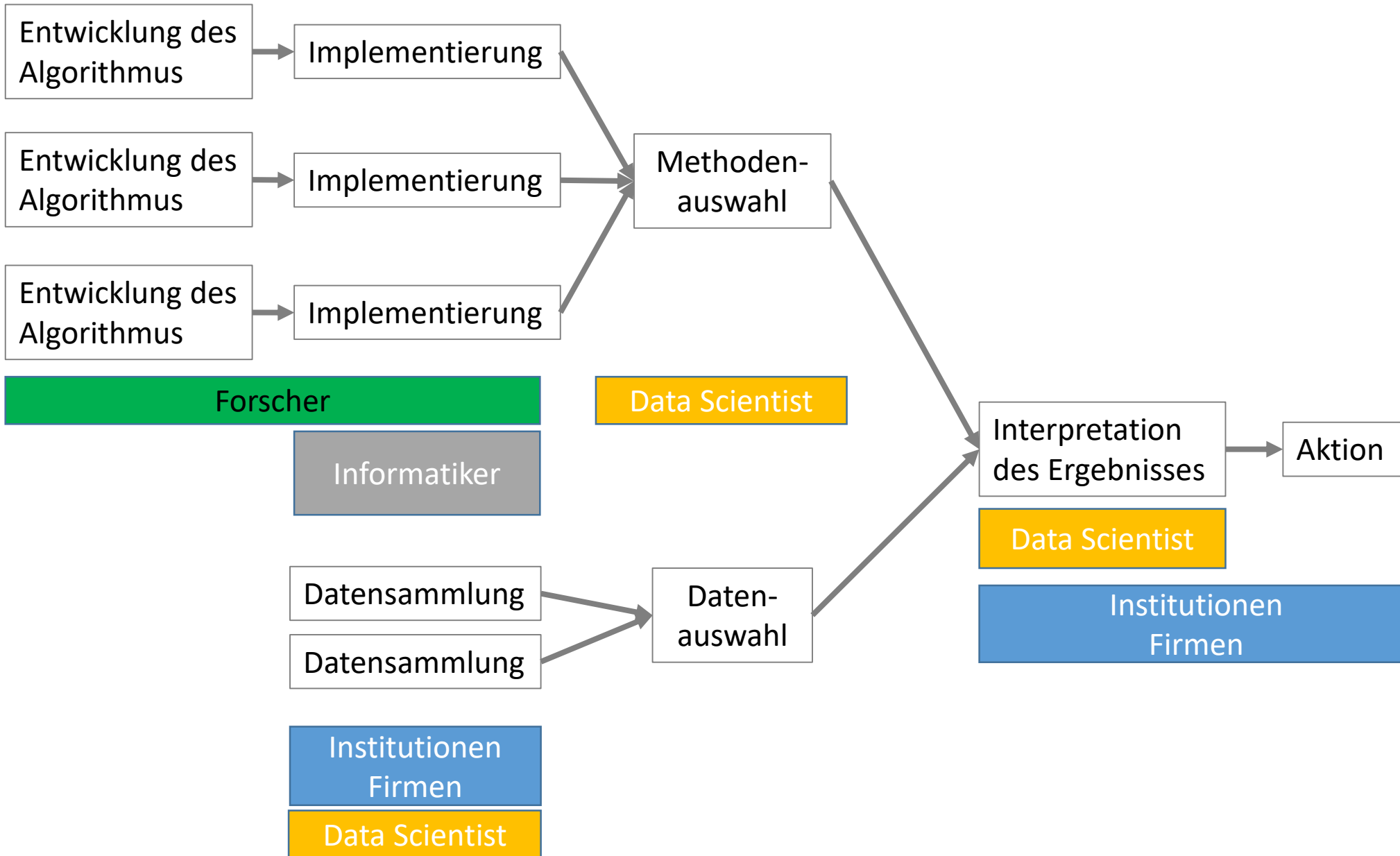
- Klassifiziert uns nach uns unbekanntem Leitlinien trainiert mit unbekanntem Daten, mit nicht ausgehandelten Qualitätsansprüchen (keine gleichberechtigte Teilhabe)
- Könnte uns alle überfordern im Bereich lebenslanges Lernen: wann welche Weiterbildung machen? Trifft besonders weniger gut Ausgebildete.
- Hinterlässt besonders reichhaltige Datenspuren zur Persönlichkeit.

Algorithmen und Teilhabe



- Nicht alle werden den Algorithmen gleichmäßig unterworfen sein. Je individueller (je reicher, gebildeter, eingebetteter), desto weniger, denn Vorhersage ist nur da möglich, wo es viele gibt, und nur da nötig, wo jemand nicht schon bekannt ist.
- Gefahr der Granularisierung der Gesellschaft in Kleinstgruppen (ins. Versicherungen, aber auch Transportkosten?)
- Einmal Ausgeschlossene können sich nicht mehr beweisen: Kredit, Gefängnis, Jobinterview- und vergabe, Studienplatzvergabe, ...

Verkettete Verantwortlichkeiten



Wer überwacht die Auswirkungen auf die Gesellschaft?

Medien?
Gesellschaft?
Politik?
Institutionen?
Firmen?
Recht?



Quis custodiet ipsos algorithmos

Der „Automated Decision Making“-TÜV vulgo: „Algorithmen TÜV“

Gründung von „Algorithm Watch“



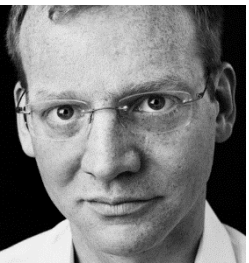
ALGORITHM
WATCH



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



Notwendige Eigenschaften

- Unabhängige Prüfstelle mit Siegelvergabe
- Möglichst auch mit Forschungsauftrag
- Identifikation der **kleinstmöglichen Menge** an zu überprüfenden Algorithmen
 - Die meisten Algorithmen sind harmlos;
 - Produkthaftung ermöglicht, dass andere, z.B. Versicherungen, Interesse an korrekten Algorithmen haben;
 - Wettbewerb ermöglicht, dass andere ‚neutralere‘ Algorithmen anbieten.
 - **Kein weiteres Innovationshemmnis!**
- **Non-Profit**

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.

Kontakt Daten

Prof. Dr. Katharina A. Zweig

TU Kaiserslautern

Gottlieb-Daimler-Str. 48

67663 Kaiserslautern

zweig@cs.uni-kl.de

Algorithmwatch.org

