



**“No more people problems”**

**Kann KI auch HR?**

**Prof. Dr. Katharina Zweig**

**@netwerkerin**

**TU Kaiserslautern**

**Leiterin des Algorithm Accountability Labs**

**Mitgründerin und wissenschaftliche Beraterin von AlgorithmWatch**

# Diskriminierung bei Bewerbungen

- Lebensläufe mit „deutschen“ Namen bekommen 14% mehr Vorstellungangebote als solche mit „türkischen“ Namen<sup>1</sup>.
- US-amerik. Studie: Frauen mit Kopftuch erhalten weniger Jobangebote als solche ohne<sup>2</sup>.



<sup>1</sup> Kaas, L. & Manger, C.: "Ethnic Discrimination in Germany's Labour Market: A Field Experiment", German Economic Review, 2011 , 13 , 1-20

<sup>2</sup> Ghuman, S. & Ryan, A. M.: "Not welcome here: Discrimination towards women who wear the Muslim headscarf , human relations, 2013 , 66(5) , 671-698



# Richter

- Richter müssen vorzeitige Haftentlassungsanträge begutachten.
- Studie: je weiter von der letzten Pause weg, desto weniger risikoreiche Entscheidungen<sup>1</sup>.
- Eine Vielzahl solcher Studien scheint zu beweisen:

<sup>1</sup> Danziger, S.; Levav, J. & Avnaim-Pesso, L.: "Extraneous factors in judicial decisions", Proceedings of the National Academy of the Sciences, 2011, 108, 6889-6892



Menschen – so irrational!

- Richter müssen vorzeitige Haftentlassungsanträge begutachten
- Studie: In den letzten 10 Jahren wurden weniger Haftentlassungsanträge bewilligt. Entscheide
- Eine Vielzahl solcher Studien scheint zu beweisen:

**Menschen sind irrational und vorurteilsbeladen.**

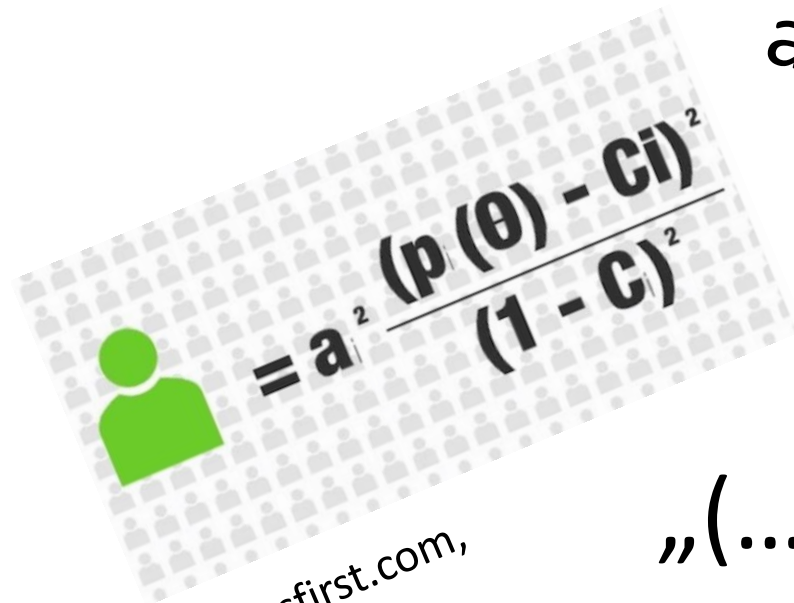
# Könnten Computer das besser?

- Die ersten Firmen testen *algorithmische Entscheidungssysteme*<sup>1</sup>.
- Eigenschaften, nach denen nicht diskriminiert werden darf, können vor ihnen besser verborgen werden.
- Sie entscheiden konsistent.

<sup>1</sup> Claire Miller: "Can an Algorithm hire Better than a Human?", The New York Times, June 25, 2015, <https://www.nytimes.com/2015/06/26/upshot/can-an-algorithm-hire-better-than-a-human.html>



„Employment assessment software“


$$= a^2 \frac{(p(\theta) - ci)^2}{(1 - c)^2}$$

Assessfirst.com,  
16.11.2017

„(...) with the availability of good data, the predictive possibilities are virtually unlimited (...)“

<https://www.inostix.com/predict-hiring-success/>  
16.11.2017

Let's take the emotion out of the process and replace it with a data-driven approach...“

iNostix (by Deloitte),  
16.11.2017

@netwerkerin  
Prof. KA Zweig  
TU Kaiserslautern



A close-up photograph of a person's hands gripping vertical metal bars, likely in a prison cell. The lighting is dramatic, with strong highlights and deep shadows, emphasizing the texture of the skin and the metallic surface of the bars. The background is dark, making the hands and bars stand out.

Forschung

—

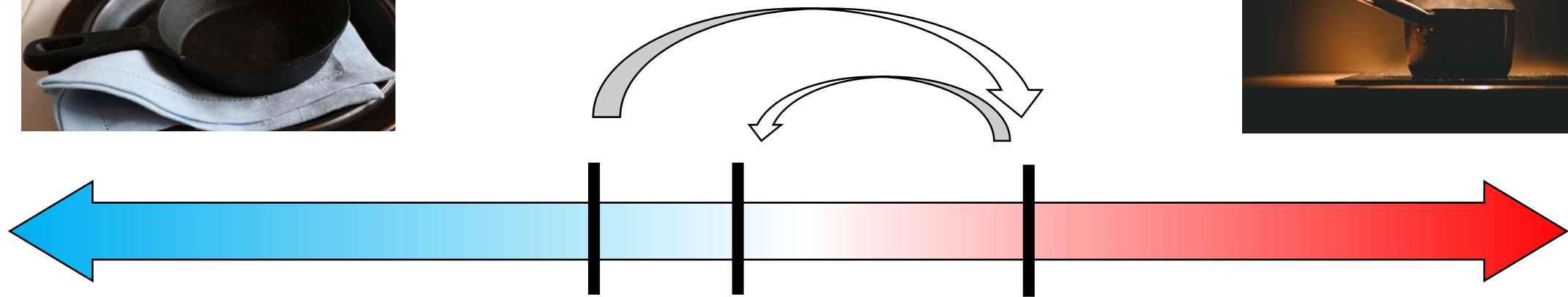
Vorhersage des  
Rückfallrisiko  
von Kriminellen



Wie lernen Computer?



# Sebastian lernt „heiss“ und „warm“



Juli

März

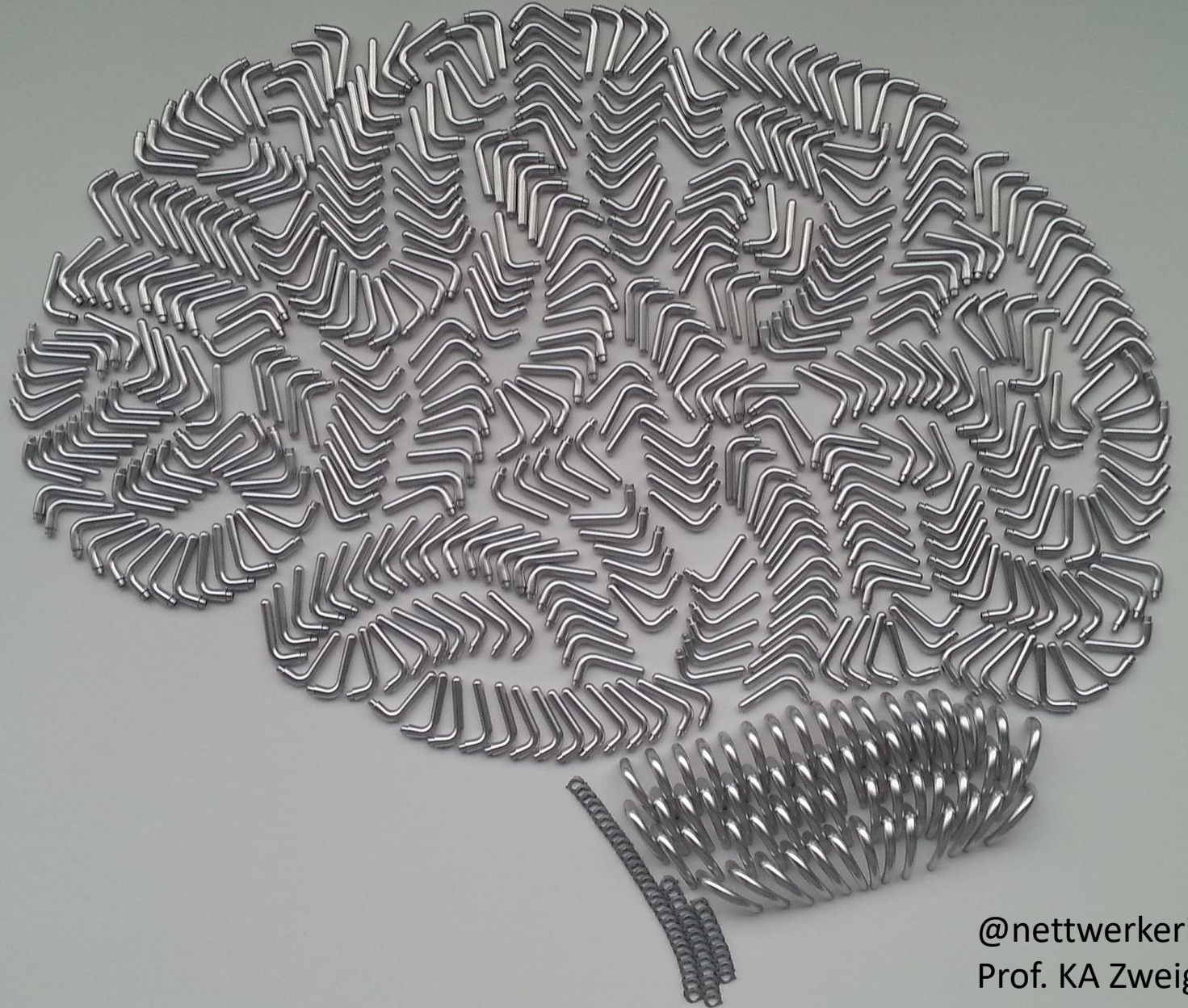
September

**Zu vorsichtig: Darf nicht dampfen Zu mutig geworden**

**Alles muss kalt sein**

## Sebastian lernt...

- Durch **Rückkopplung:** unerwartet heiß, unerwartet kalt
- Durch **Speicherung in einer Struktur:** in Neuronen und deren Verknüpfung.
- Durch **Generalisierung des Gelernten.**

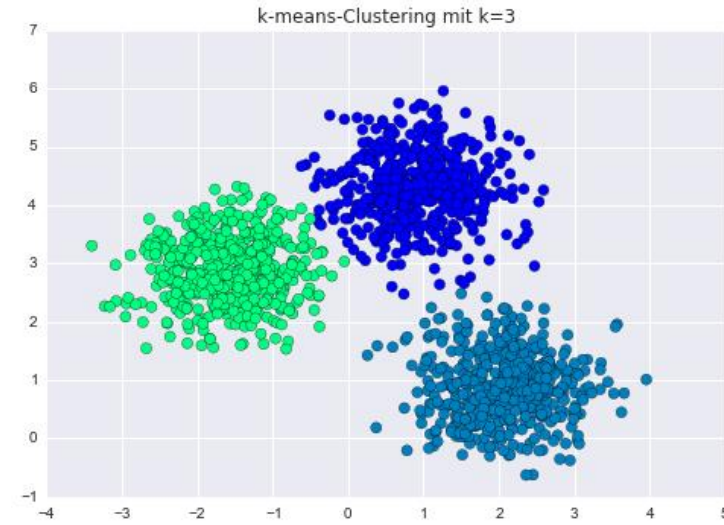
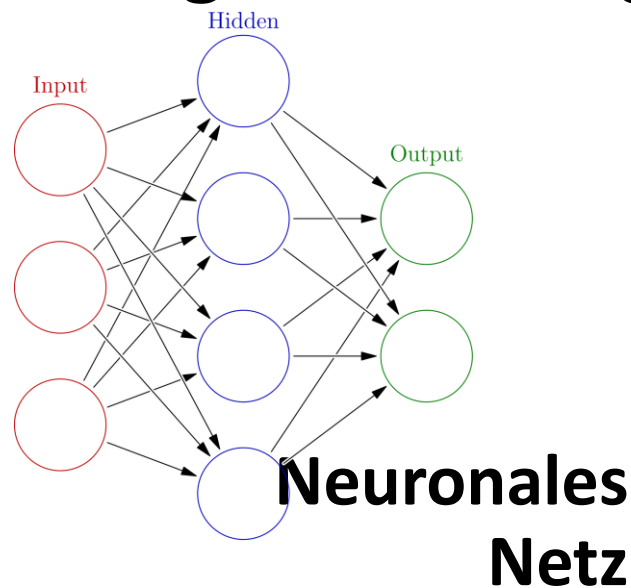


# Computer lernen

Damit ein Computer lernen kann, benötigt er ebenfalls eine **Struktur**, um Gelerntes abzuspeichern.

Optimal auch **Rückkopplung**.

Er lernt **generelle Regeln**.

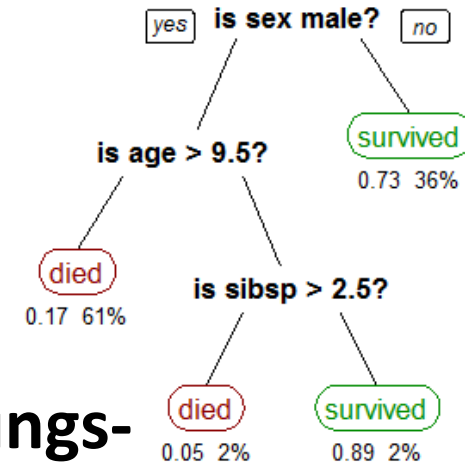


**Clustering**

**Formel**

$$w_1 * \#V_h - w_2 * \#day_i V_h + w_3 * I[g = mc]$$

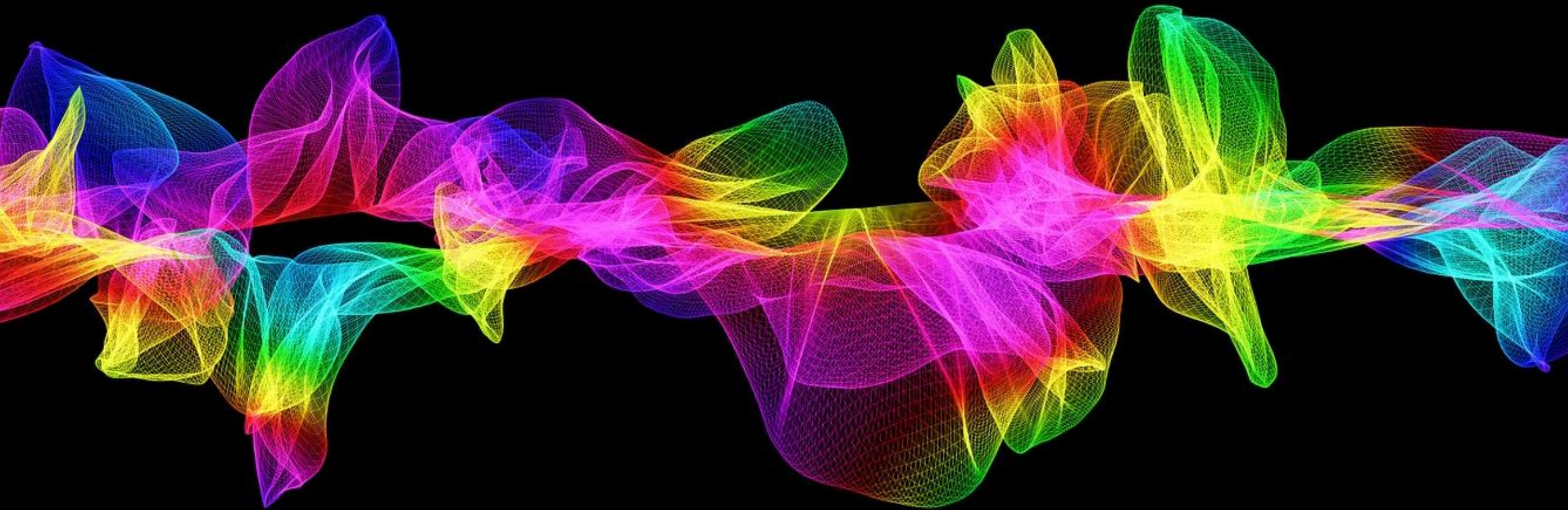
**Entscheidungs-  
bäume**



@netwerkerin  
Prof. KA Zweig  
TU Kaiserslautern

"CART tree titanic survivors" by Stephen Milborrow - Own work. Licensed under CC BY-SA 3.0 via Wikimedia Commons - [https://commons.wikimedia.org/wiki/File:CART\\_tree\\_titanic\\_survivors.png#/media/File:CART\\_tree\\_titanic\\_survivors.png](https://commons.wikimedia.org/wiki/File:CART_tree_titanic_survivors.png#/media/File:CART_tree_titanic_survivors.png)  
By Glosser.ca - Own work, Derivative of File:Artificial neural network.svg, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=24913461>



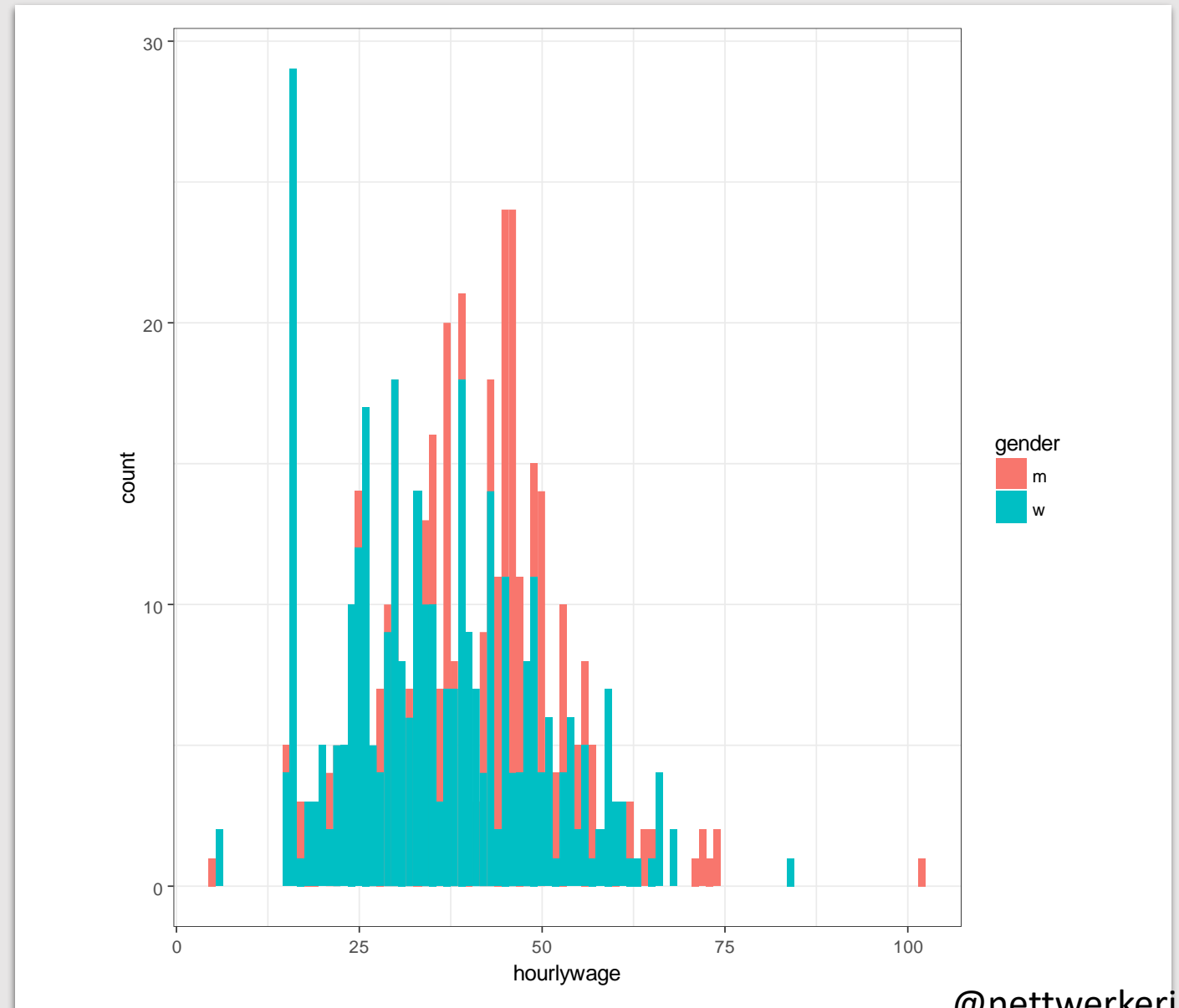


“Lernen” mit Korrelationen

# Gehälter in Seattle

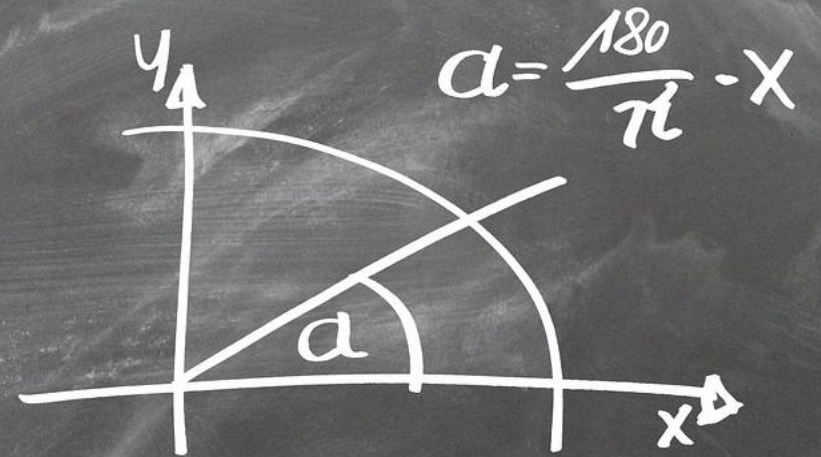
Sie bekommen Daten von einer Person – diese verdient weniger als \$25 pro Stunde.

Basierend auf den Daten, ist die Person weiblich oder männlich?





$$X_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$



$$X^2 + px + q = 0$$

$$x = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$$



$$X_{1/2} = -\frac{p}{2} \pm \sqrt{\left(\frac{p}{2}\right)^2 - q}$$

Lernen mit Formeln



## Datengrundlagen

- Data Mining Methoden nutzen verschiedene Informationen
- Am wichtigsten:
  - **War Einstellung erfolgreich?**

Ausbildung

Leerzeiten

Arbeitgeber  
-wechsel

Alter

Bewerbungs-  
schreiben

Rechtschreibung

Wortvielfalt

Ton

Social Media?

# Regressionsansätze

- Die Algorithmen-designer entscheiden, welche Daten vermutlich mit „erfolgreicher Einstellung“ korrelieren.
- Die Software sollte eine einzige Zahl ausgeben.
- Je höher die Zahl, desto höher die Erfolgswahrscheinlichkeit.
- Beispiel Formel:

$$\begin{aligned} & 3 * \text{Jahre im Job} \\ - & 2 * \text{Anzahl Arbeitgeber} \\ + & 3 * (\text{Wenn Auslandserfahrung,} \\ & \text{dann 1, sonst 0}) \\ + & 2,5 * (\text{Wenn Fortbildung,} \\ & \text{dann 1, sonst 0}) + \dots \end{aligned}$$

## Allgemein

**Der Computer bestimmt die Gewichte und bekommt ein Feedback (Rückkopplung), inwieweit die resultierende Bewertung mit dem (beobachteten) Verhalten übereinstimmt.**

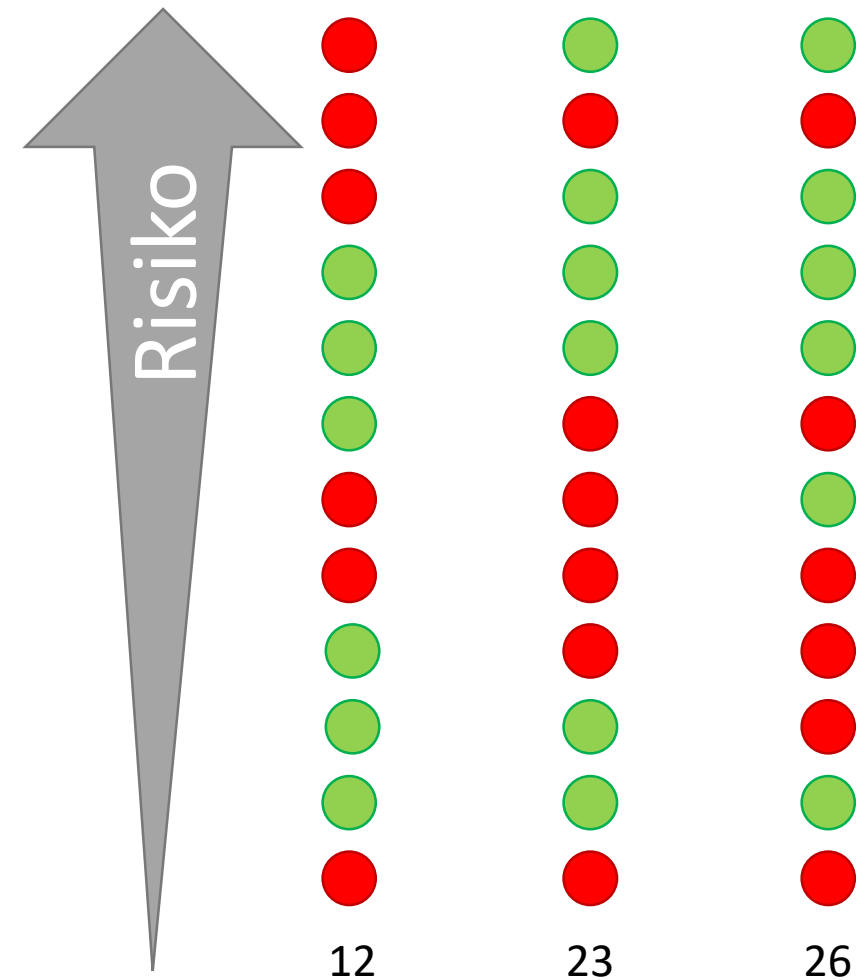
$$\begin{aligned} & w_1 * \text{Jahre im Job} \\ + & w_2 * \text{Anzahl Arbeitgeber} \\ + & w_3 * (\text{Wenn Auslandserfahrung,} \\ & \text{dann 1, sonst 0)} \\ + & w_4 * (\text{Wenn Fortbildung,} \\ & \text{dann 1, sonst 0)} + \dots \end{aligned}$$





Qualität eines Algorithmus

- Grüne Kugeln symbolisieren erfolgreiche, rote nicht erfolgreiche BewerberInnen.
- Optimale Sortierung: Alle grünen oben, alle roten darunter.
- Qualitätsmaß: Paare von rot und grün, bei denen die grüne Kugel über der roten einsortiert ist. (**ROC AUC**)



Ist das Qualitätsmaß sinnvoll?

- Wenn die Stelle **sofort** besetzt werden muss, und nur 5 Bewerber da sind: **ja**
- Wenn es langfristig um die Identifikation der besten Talente geht: **nein**
- Hier müssen andere Qualitätsmaße benutzt werden.





Das ist wie...

„Kaufen Sie diesen wunderbaren Wagen. TÜV? Brauchen Sie nicht! Und sehen Sie nur, die unglaublich gut erhaltenen Sommerreifen. Das ist noch Qualität!“



@nettwwerkerin  
Prof. KA Zweig  
TU Kaiserslautern

# Probleme von algorithmischen Entscheidungssystemen (ADM Systemen) im People Assessment

- 1. Wer entscheidet, wann ein  
ADM System „gut“ ist?**







Wahrscheinlichkeit & Wahrheit



# 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**)

## Algorithmen...

- ... basieren auf Korrelationen von Eigenschaften mit gewünschtem Verhalten.
- **Quasi algorithmisch legitimierte Vorurteile:**
  - Zu 70% erfolgreich heißt:
  - Von 100 Personen, die „genau so sind wie dieser Mensch“, sind 70 nachher erfolgreich.

```
is},a(window).on( load...
e strict";function b(b){return this.each(function(){var
ction(b){this.element=a(b)};c.VERSION="3.3.7",c.TRANSITION_D
.data("target");if(d||(d=b.attr("href"),d=d&&d.replace(/.*(?
ide.bs.tab",{relatedTarget:b[0]}),g=a.Event("show.bs.tab",{r
ar h=a(d);this.activate(b.closest("li"),c),this.activate(h,h
.bs.tab",relatedTarget:e[0]}))}}},c.prototype.activate=fun
Class("active").end().find('[data-toggle="tab"]').attr("ar
[b[0].offsetWidth,b.addClass("in"):b.removeClass("fade"),l
="tab"]').attr("aria-expanded",!0),e&&e()}var g=d.find(">
e").length);g.length&&h?g.one("bsTransitionEnd",f).emula
tab=b,a.fn.tab.Constructor=c,a.fn.tab.noConflict=functionio
"click.bs.tab.data-api",[data-toggle="tab"],e).on("
return this.each(function(){var d=a(this),e=d.data(
function(b,d){this.options=a.extend({},c.DEFAULTS,c
,this)).on("click.bs.affix.data-api",a.proxy(thi
is.checkPosition());c.VERSION="3.3.7",c.RESET=
his.$target.scrollTop(),f=this.$element.offse
l=c?!(e+this.unpin<=f.top)&&"bottom":!(e+
bottom"},c.prototype.getPinnedOffset=fu
get.scrollTop(),b=this.$element.of
this.checkPosition,this) 1))
```

# Probleme von algorithmischen Entscheidungssystemen (ADM Systemen) im People Assessment

1. Wer entscheidet, wann ein ADM System „gut“ ist?
2. **ADM Systeme ergeben nur Wahrscheinlichkeiten, keine Wahrheiten.**







Können Algorithmen diskriminieren?







Und das, wenn ich auf Pixabay nach „Chef“ suche...



## Diskriminierung

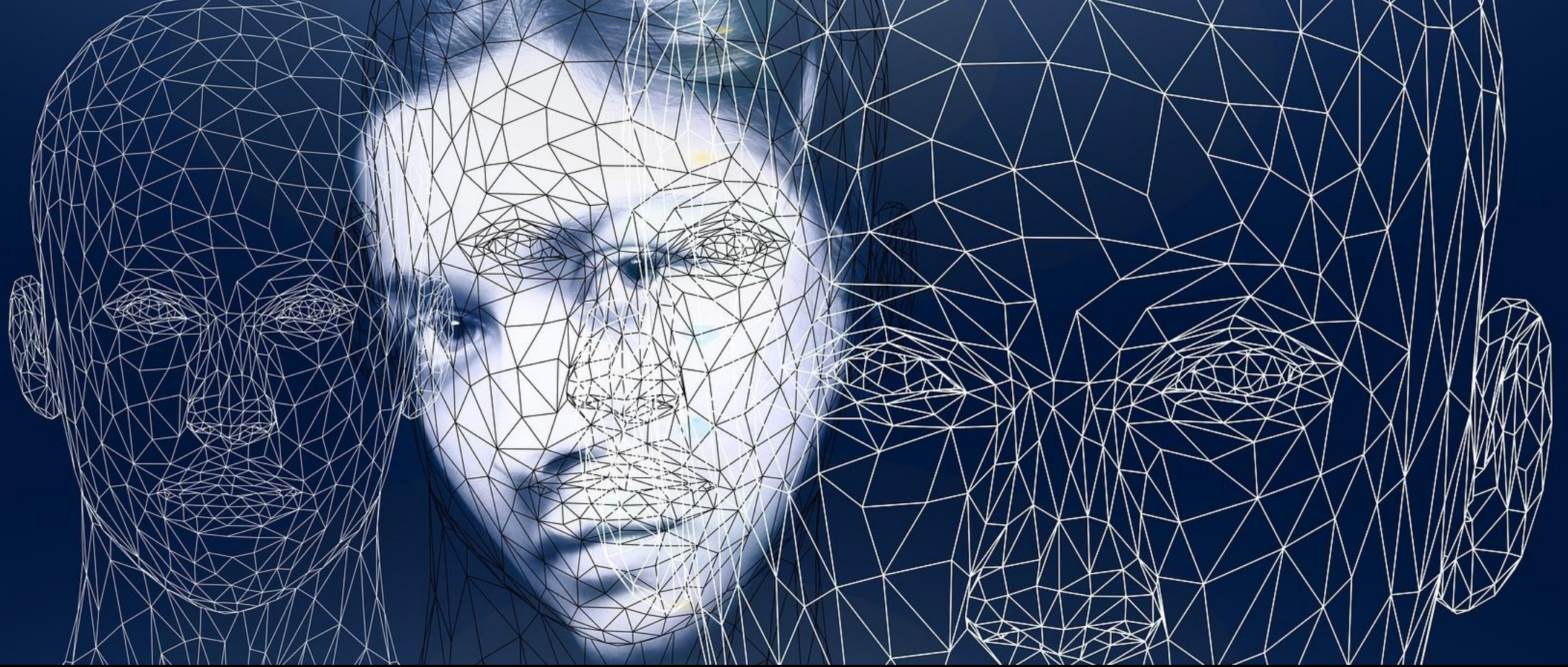
- Google zeigt weiblichen Surfern schlechtere Jobs an.
- Rückfälligkeitsvorhersagealgorithmen sind rassistisch.
- Diskriminierungen in Trainingsdaten werden „mitgelernt“.
- Wenn Trainingsdaten zu wenig Daten über Minderheiten enthalten, werden deren Eigenschaften nicht „mitgelernt“.

# Probleme von algorithmischen Entscheidungssystemen (ADM Systemen) im People Assessment

1. Wer entscheidet, wann ein ADM System „gut“ ist?
2. ADM Systeme ergeben nur Wahrscheinlichkeiten, keine Wahrheiten.
3. **ADM Systeme können diskriminieren.**







# Sozio-informatische Gesamtbetrachtung



## Probleme der Einbettung der ADM in den sozialen Prozess

- **Aufmerksamkeitsökonomie** von Entscheiderinnen und Entscheidern.
- „**Best practice**“ erfordert Nutzung der Software.
- **Delegation von Verantwortung!**
- Manchmal kann ein falsch Beurteiler **die Vorhersage prinzipiell nicht entkräften!**
  - Z.B. abgelehnte Bewerberin

# Probleme von algorithmischen Entscheidungssystemen (ADM Systemen) im People Assessment

1. Wer entscheidet, wann ein ADM System „gut“ ist?
2. ADM Systeme ergeben nur Wahrscheinlichkeiten, keine Wahrheiten.
3. ADM Systeme können diskriminieren.
4. **ADM Systeme können soziale Prozesse verändern.**



# Probleme von algorithmischen Entscheidungssystemen (ADM Systemen) im People Assessment

- 1. Wer entscheidet, wann ein ADM System „gut“ ist?**
- 2. ADM Systeme ergeben nur Wahrscheinlichkeiten, keine Wahrheiten.**
- 3. ADM Systeme können diskriminieren.**
- 4. ADM Systeme können soziale Prozesse verändern.**





## Einschätzung

- People Assessment Systeme könnten dabei helfen, bessere Entscheidungen zu treffen.
- Allerdings ist es schwierig, sie transparent, fair und nachvollziehbar zu gestalten.

Sie sind zuverlässig

Können Entscheidungswege transparenter machen

Könnten Diskriminierung vermeiden





Ein neues Qualitätsprodukt  
aus Deutschland?



Made in Germany



Entwurf eines  
qualitätsgesicherten  
Entwicklungsprozess für

People Assessment Systeme  
und Regeln für ihren  
ethischen (Nicht-)Einsatz.

140

2025



People Assessment Systems  
Made in Germany



Fair  
Nachvollziehbar  
Transparent

140

2025



## Wo Maschinen irren können

Verantwortlichkeiten und Fehlerquellen in  
Prozessen algorithmischer Entscheidungsfindung

## Weitere Literatur

Studie für die Bertelsmann-Stiftung  
(2018)

**Katharina Zweig** mit Sarah Fischer  
und Konrad Lischka

[https://www.bertelsmann-  
stiftung.de/de/publikationen/publikation/did/wo-  
maschinen-irren-koennen/](https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/wo-maschinen-irren-koennen/)

**Katharina Zweig (2018)**

**Auch Algorithmen können  
diskriminieren**

[https://merton-magazin.de/auch-algorithmen-  
koennen-diskriminieren?tags=Personalmanagement](https://merton-magazin.de/auch-algorithmen-koennen-diskriminieren?tags=Personalmanagement)