



Kann KI HR?

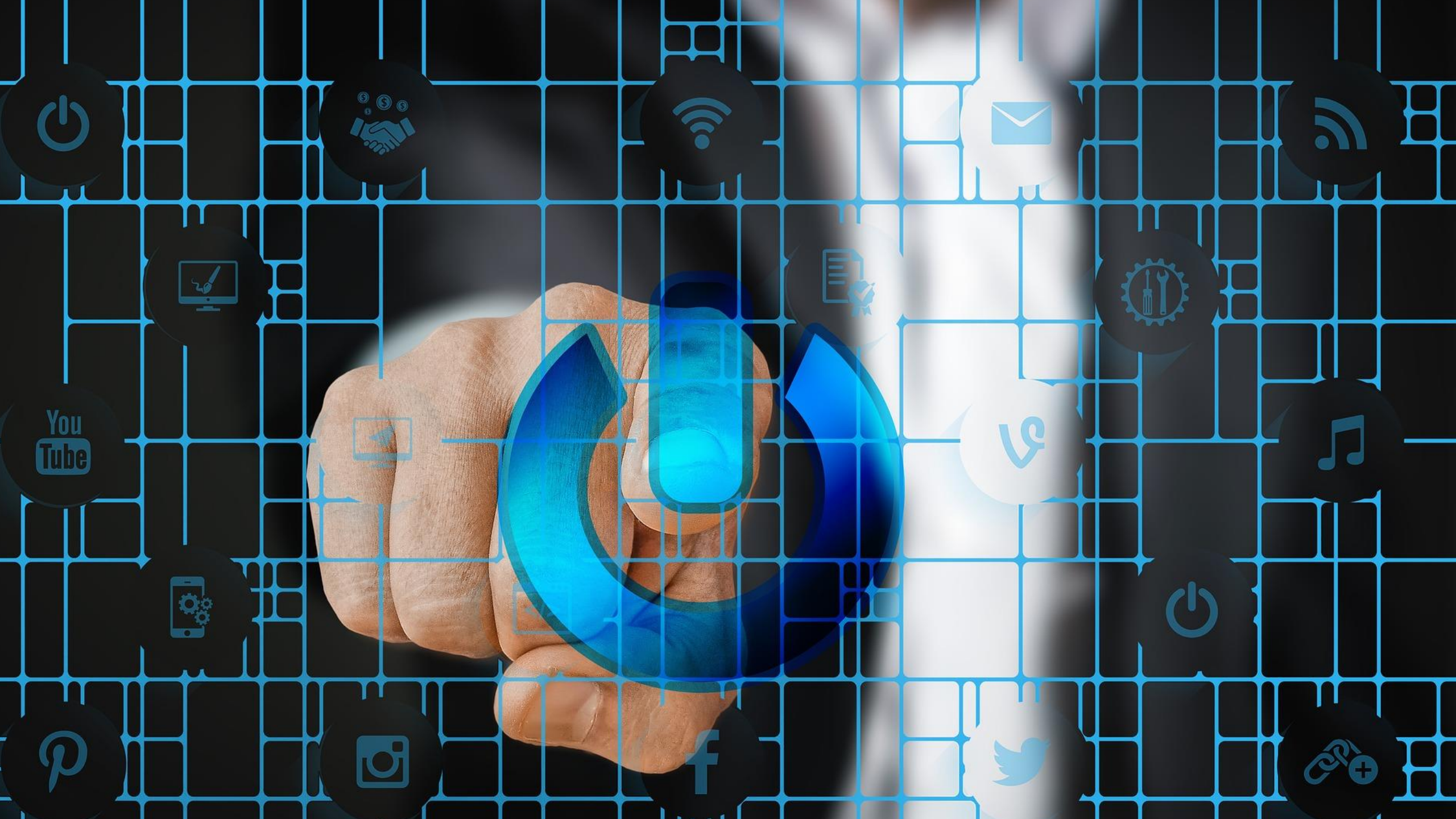
Prof. Dr. Katharina A. Zweig,
TU Kaiserslautern











Konstituierende Sitzung der
Enquete-Kommission
„Künstliche Intelligenz“ am 27.9.2019

Aus der Rede von Bundestagspräsidenten
Dr. Schäuble:

- „Die künstliche Intelligenz gilt
Vielen als neue Zauberformel des
technischen Fortschritts, ...
- ... sie wird dichten, ...
- ... sie wird belohnen und bestrafen ...“



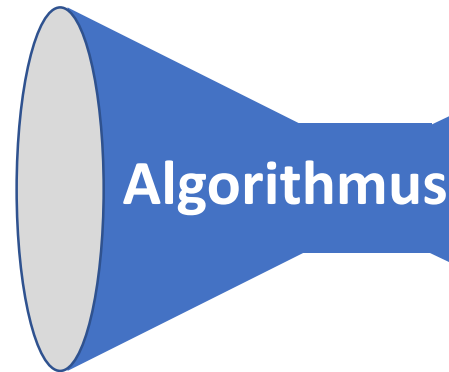
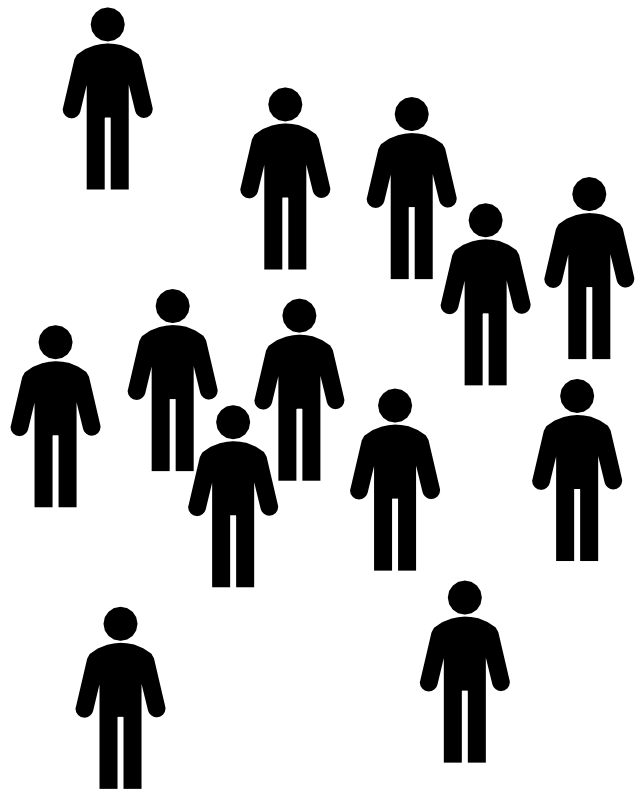
Die zwei Ängste

Sie wird dichten

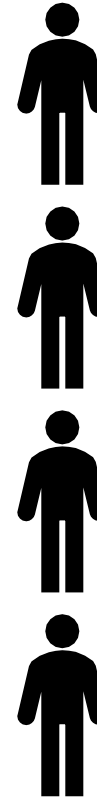
Sie wird richten



Algorithmische Entscheidungssysteme (ADM Systeme)



oder



Scoring-Verfahren



Klassifikation

Diskriminierung bei Bewerbungen

- Lebensläufe mit „deutschen“ Namen bekommen 14% Vorstellungsangebote als solche mit „türkischen“ Namen¹.
- US-amerik. Studie: Frauen mit Kopftuch erhalten weniger Jobangebote als solche ohne².



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

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

Menschen – so irrational!

- Richter müssen vorzeitige Haftentlassungsanträge begutachten.
- Studie: je weiter von der letzten Pause weg, desto weniger risikoreiche Entscheidungen¹.
- Eine Vielzahl solcher Studien scheint zu beweisen:
 - Menschen sind irrational und vorurteilsbeladen.



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

Könnten Computer das besser?

- Die ersten Firmen testen das ¹.
- Eigenschaften, nach denen nicht diskriminiert werden darf, können vor ihnen besser verborgen werden.
- Sie sind objektiv und arbeiten nahezu fehlerfrei.
- (objektiv := „reproduzierbar dieselbe Entscheidung bei derselben Eingabe von Daten“)



Wie „lernt“ das System von Daten?

DIY:

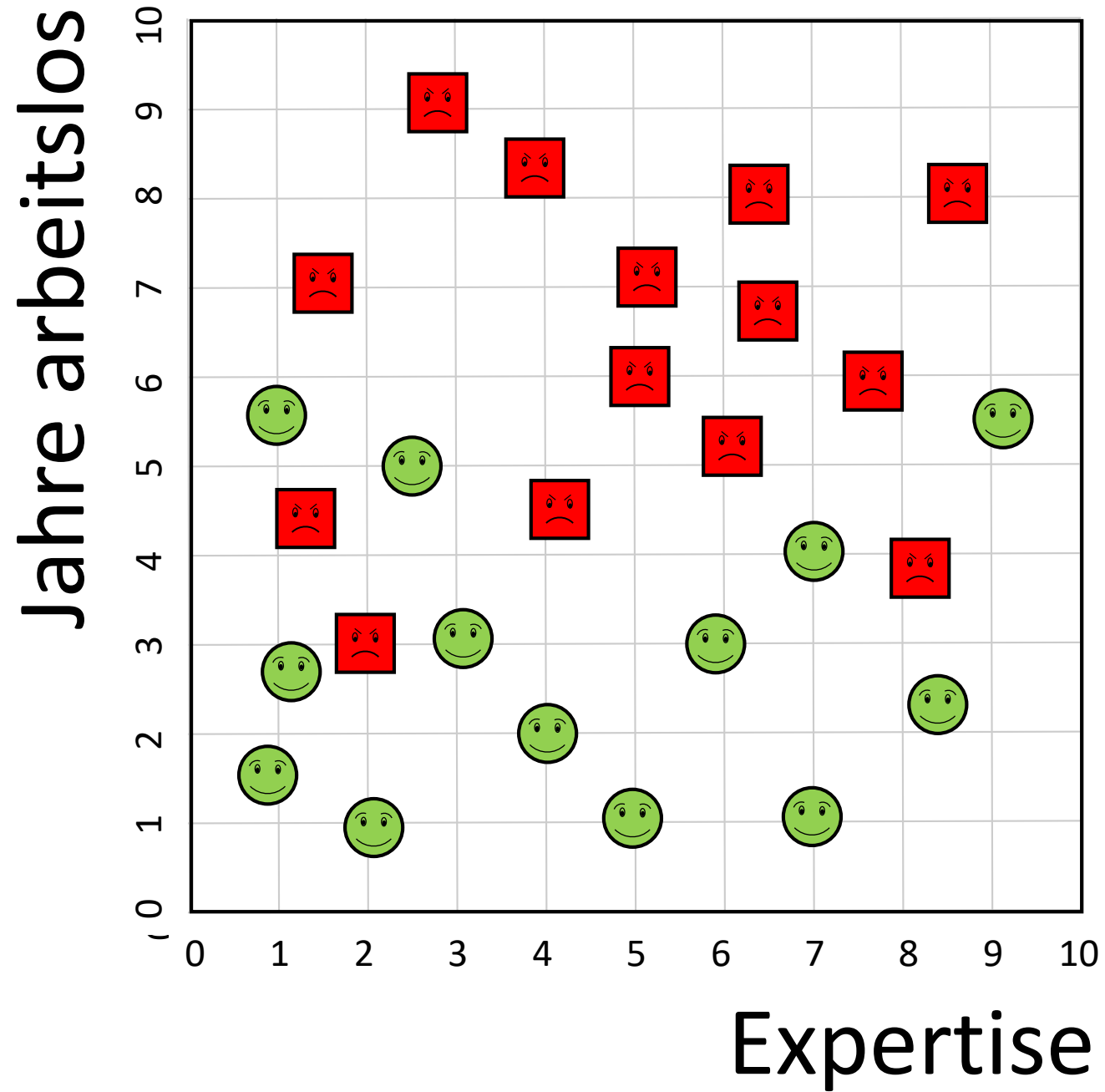
**Sie sind heute meine
„Support Vector Machine“**





Weniger erfolgreiche
Arbeitnehmer:innen



Erfolgreiche Arbeit-
nehmer:innen

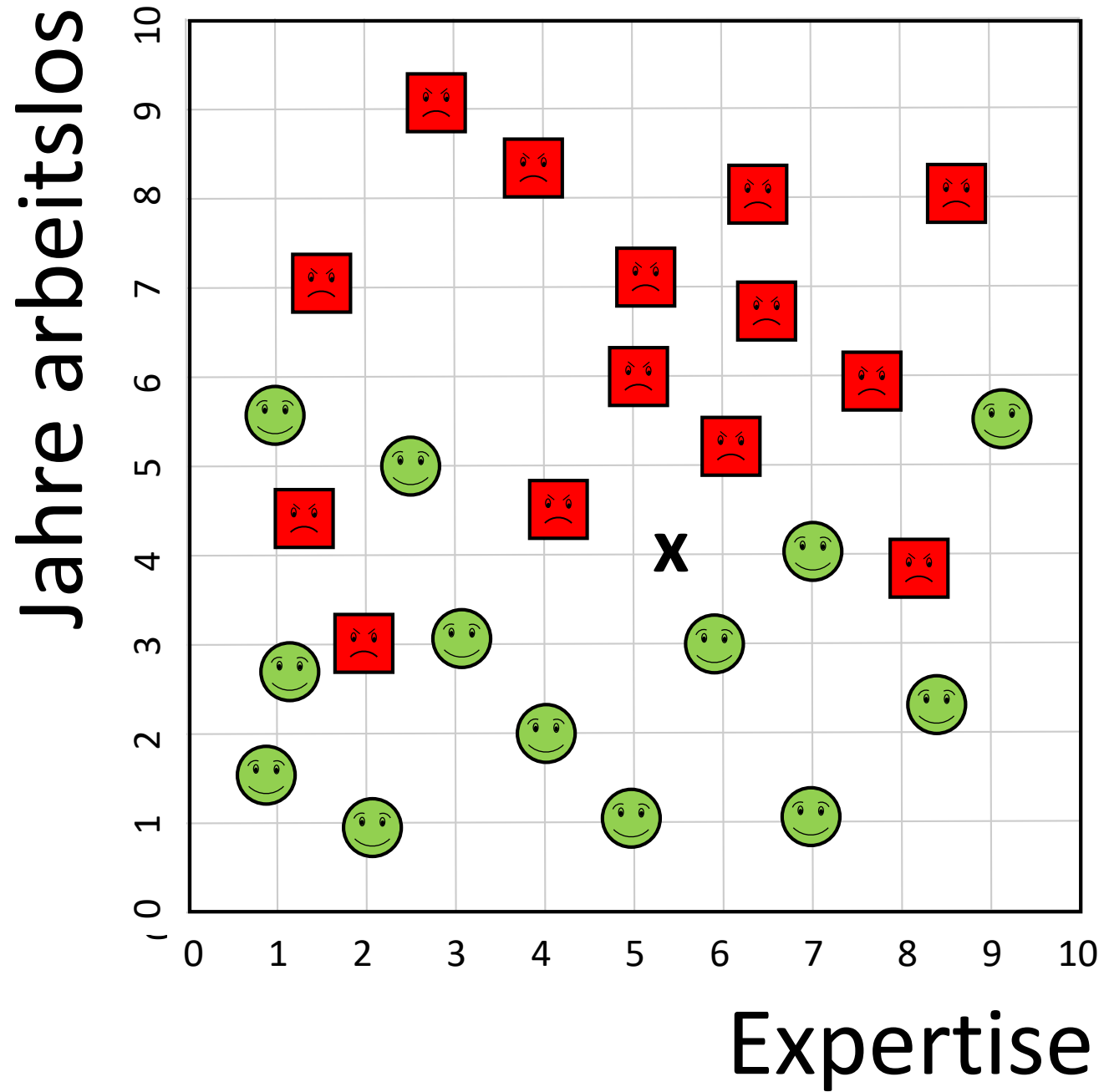


-  Weniger erfolgreiche Arbeitnehmer:innen
-  Erfolgreiche Arbeitnehmer:innen

Bewerten Sie Frau Müller:

5.5 Jahre Erfahrung

4 Jahre arbeitslos

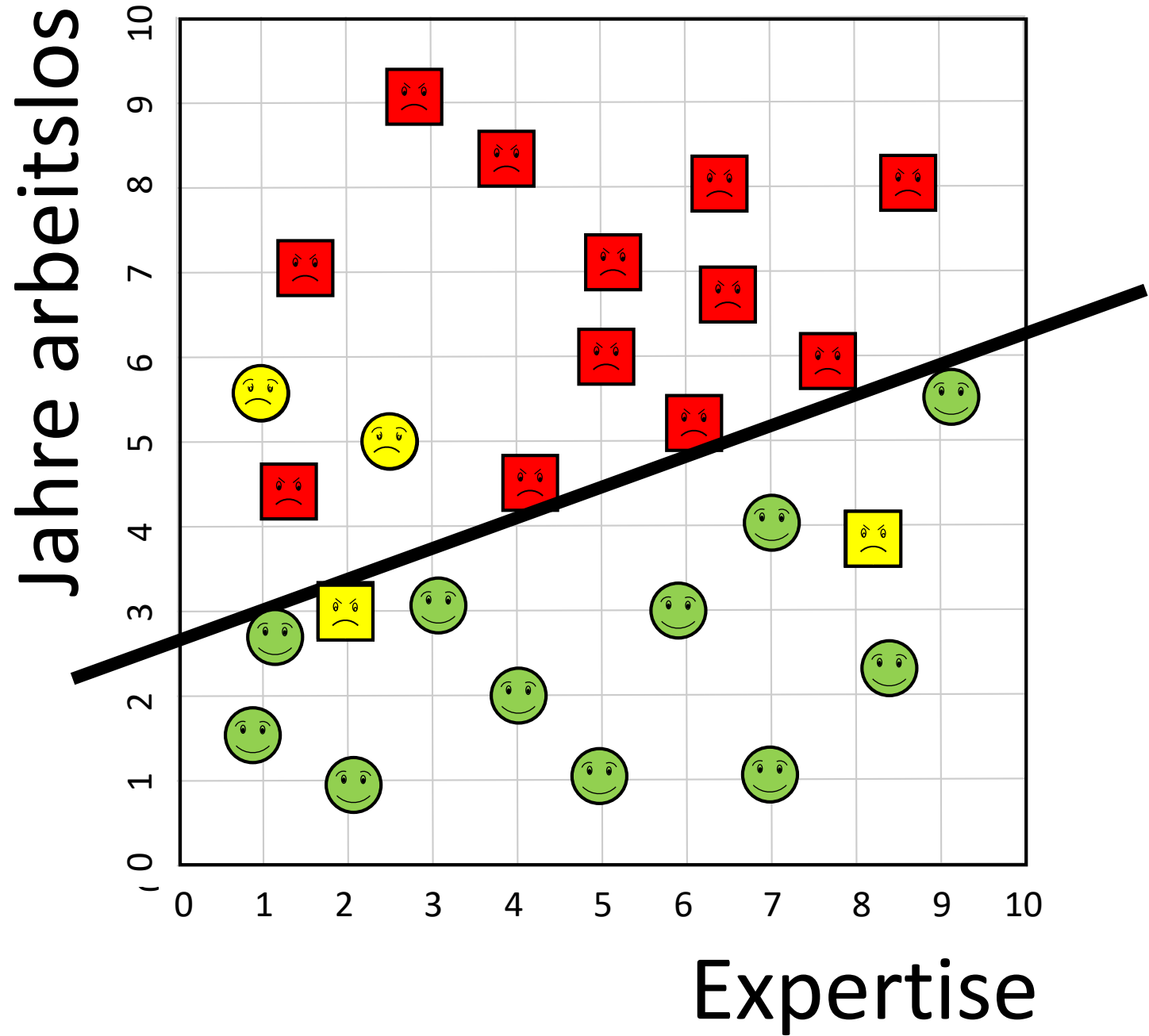


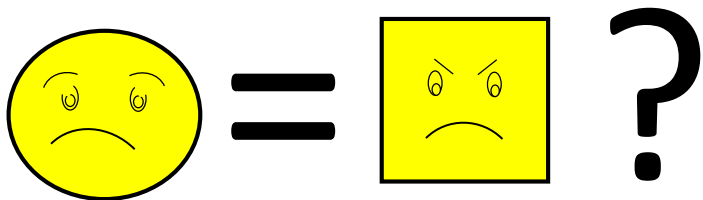




Weniger erfolgreiche Arbeitnehmer:innen

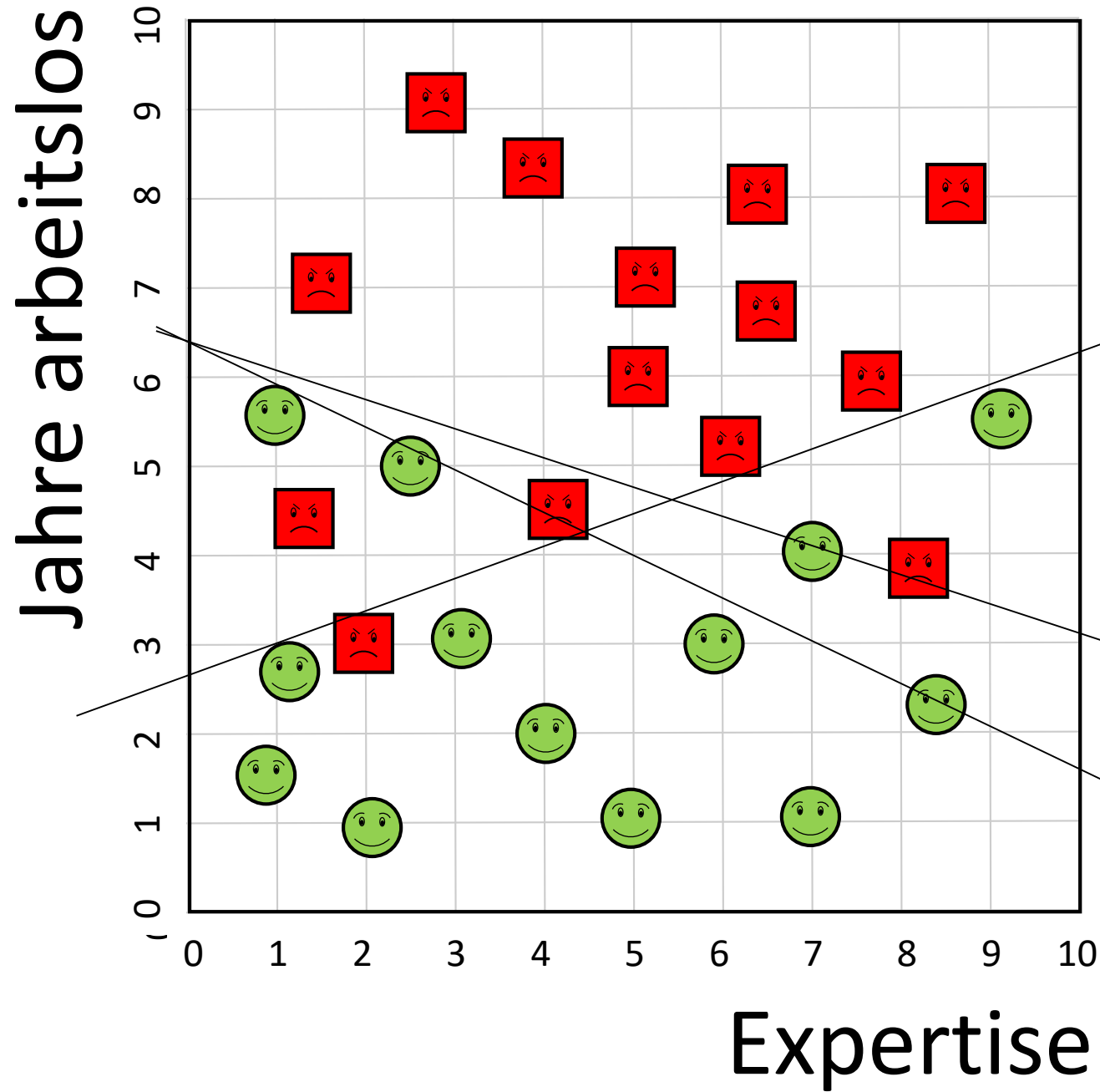


Erfolgreiche Arbeitnehmer:innen





-  Weniger erfolgreiche Arbeitnehmer:innen
-  Erfolgreiche Arbeitnehmer:innen



„Wir können es uns nicht leisten
erfolgreiche
Arbeitnehmer:innen
zu **übersehen!**“

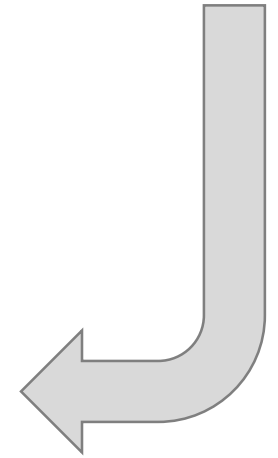
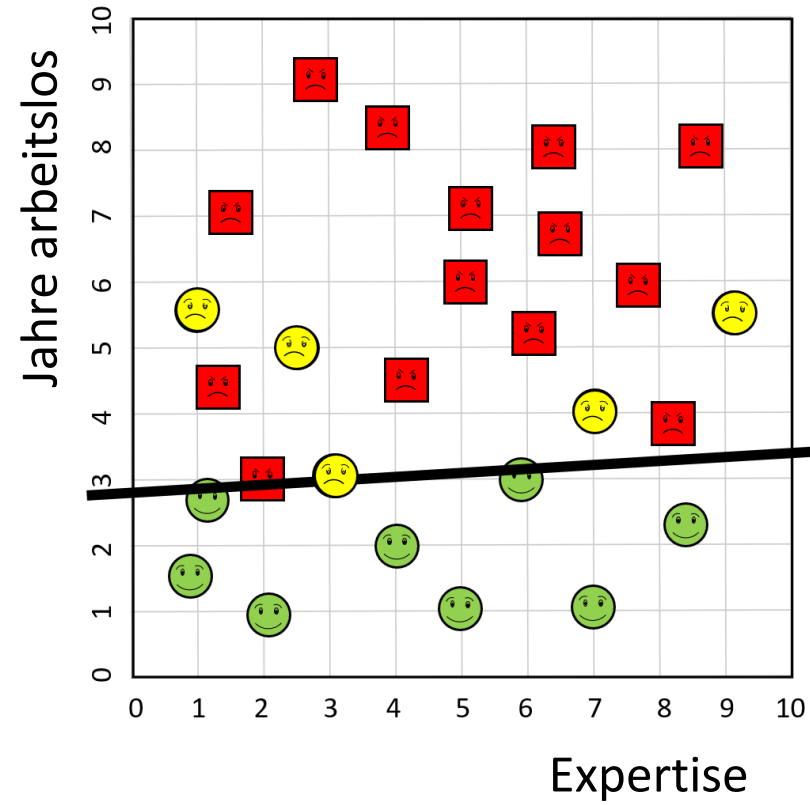
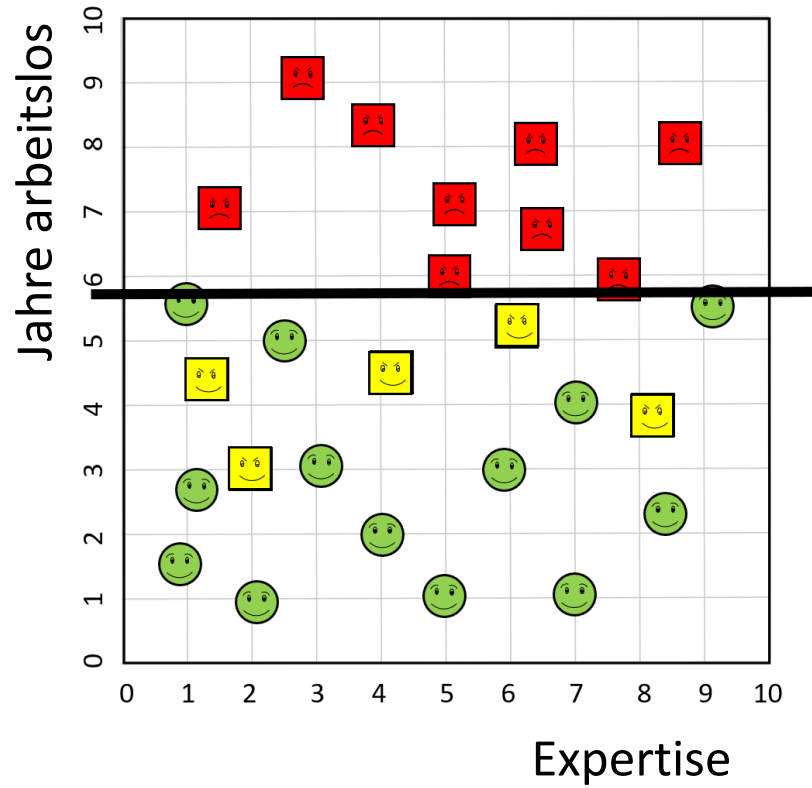
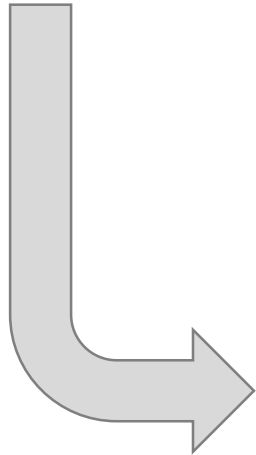


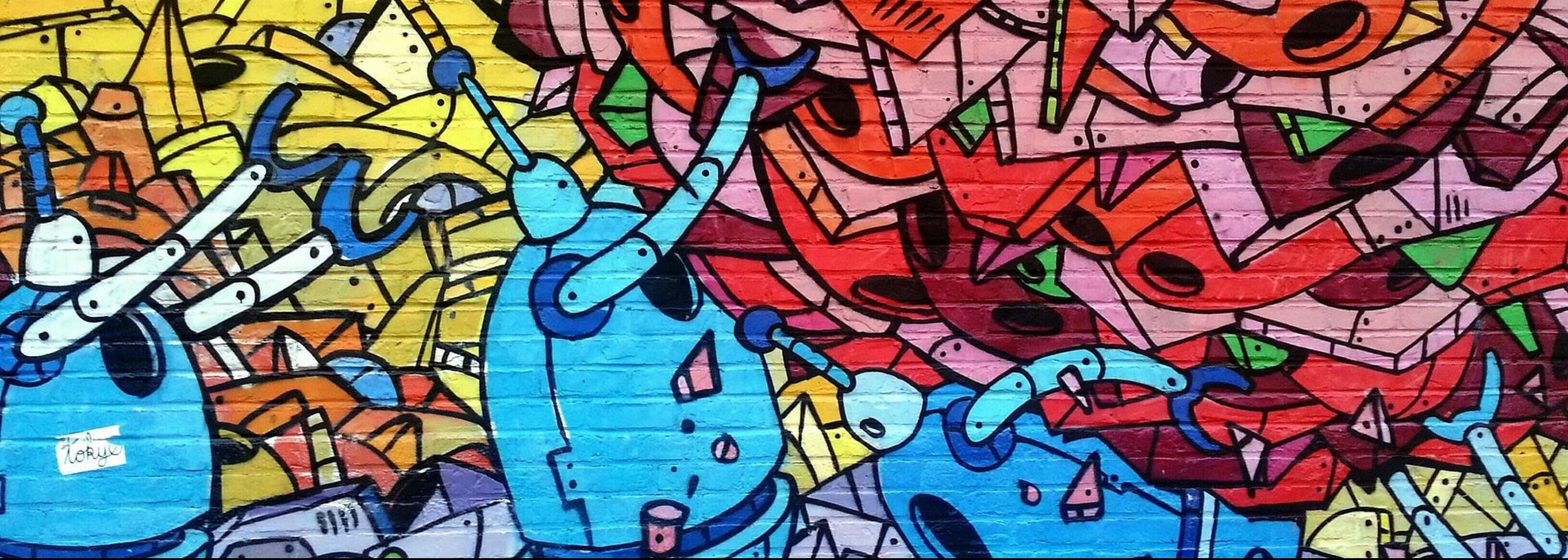
Weniger erfolgreiche
Arbeitnehmer:innen



Erfolgreiche Arbeit-
nehmer:innen

„Wir können es uns nicht leisten,
weniger erfolgreiche
Arbeitnehmer:innen
 einzustellen!“





Warum Sie hier ran müssen!

Qualität von ADM Systemen

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



Transparenz

- Dokumentation des AMS-Algorithmus ist Blaupause für Transparenz:
- Daten
- Methode
- Qualitätstests:
 - Ca. 80% der Entscheidungen über die 1. und ca. 80% der Entscheidungen über die 3. Gruppe sind korrekt.

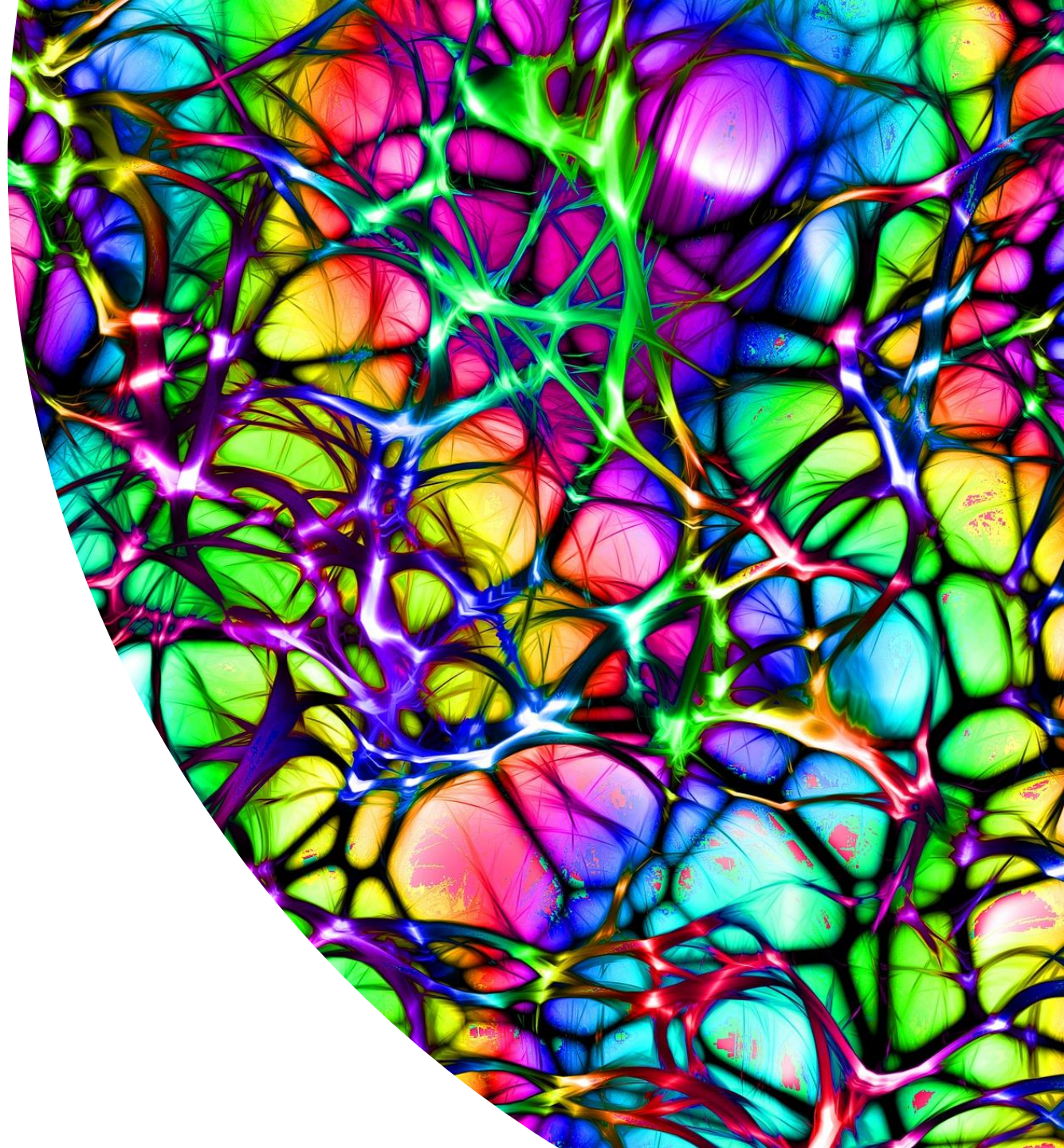


Regel

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

Sie suchen **Muster** in hochverrauschten Datensätzen.

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



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 d
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))
```

Qualität von ADM Systemen

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





Computer: Objektiv und fehlerfrei?

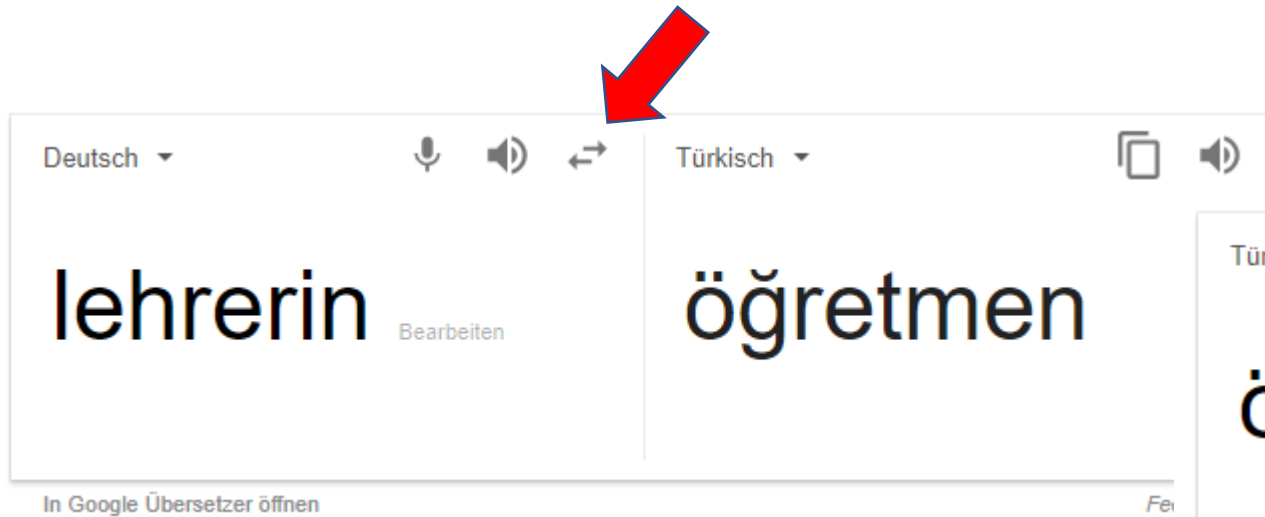
Gleichberechtigung





Wenn man auf Google nach „CEO“ sucht...

Passt das einigermaßen: ca. 8%
der (deutschen) CEOs sind Frauen



Übersetzungen

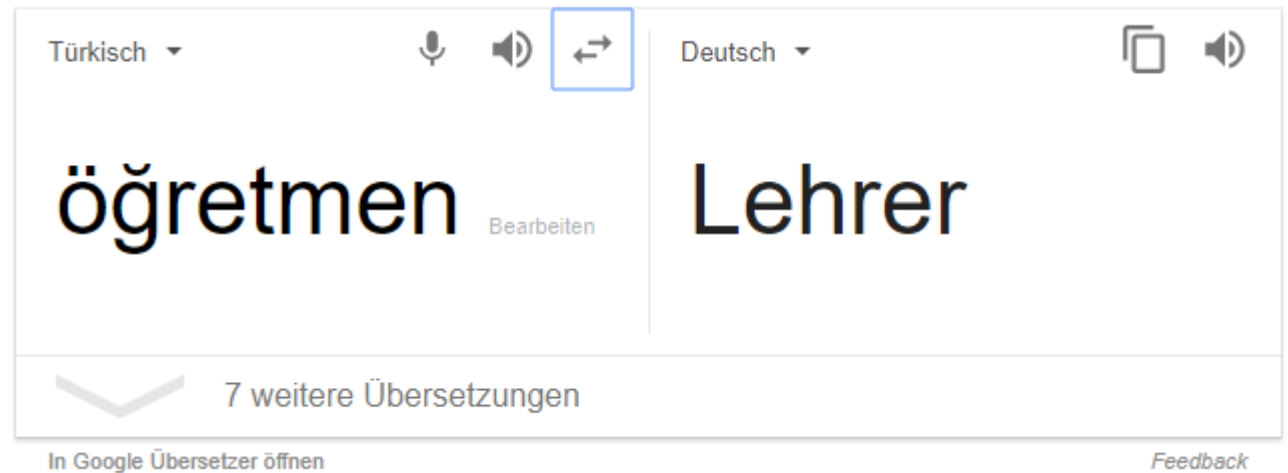





Deutsch ▾    Türkisch ▾  

lehrerin Bearbeiten **öğretmen**


In Google Übersetzer öffnen Feedback

A red arrow points to the swap languages icon (two arrows pointing in opposite directions) between the source and target language dropdowns.



Türkisch ▾    Deutsch ▾  

öğretmen Bearbeiten **Lehrer**

 7 weitere Übersetzungen

In Google Übersetzer öffnen Feedback



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

Diskriminierung



- Der AMS-Algorithmus weist Alten, Frauen und Pflegenden höheres Risiko zu.
- Amazons Bewerbungsbewerter diskriminierte Frauen.
- Diskriminierungen in Trainingsdaten werden „mitgelernt“.
- Wenn Trainingsdaten zu wenig Daten über Minderheiten enthalten, werden deren Eigenschaften nicht „mitgelernt“.

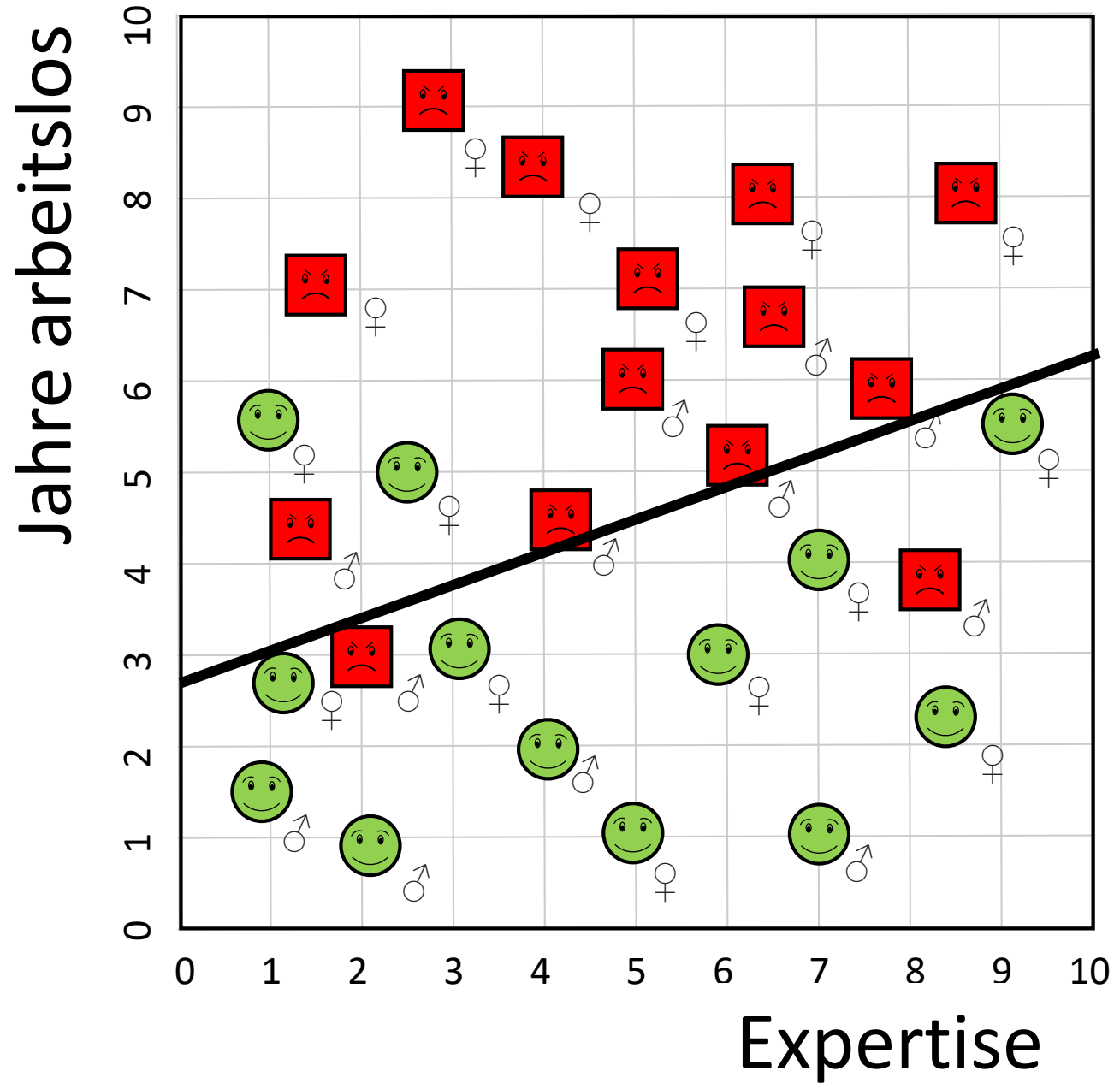


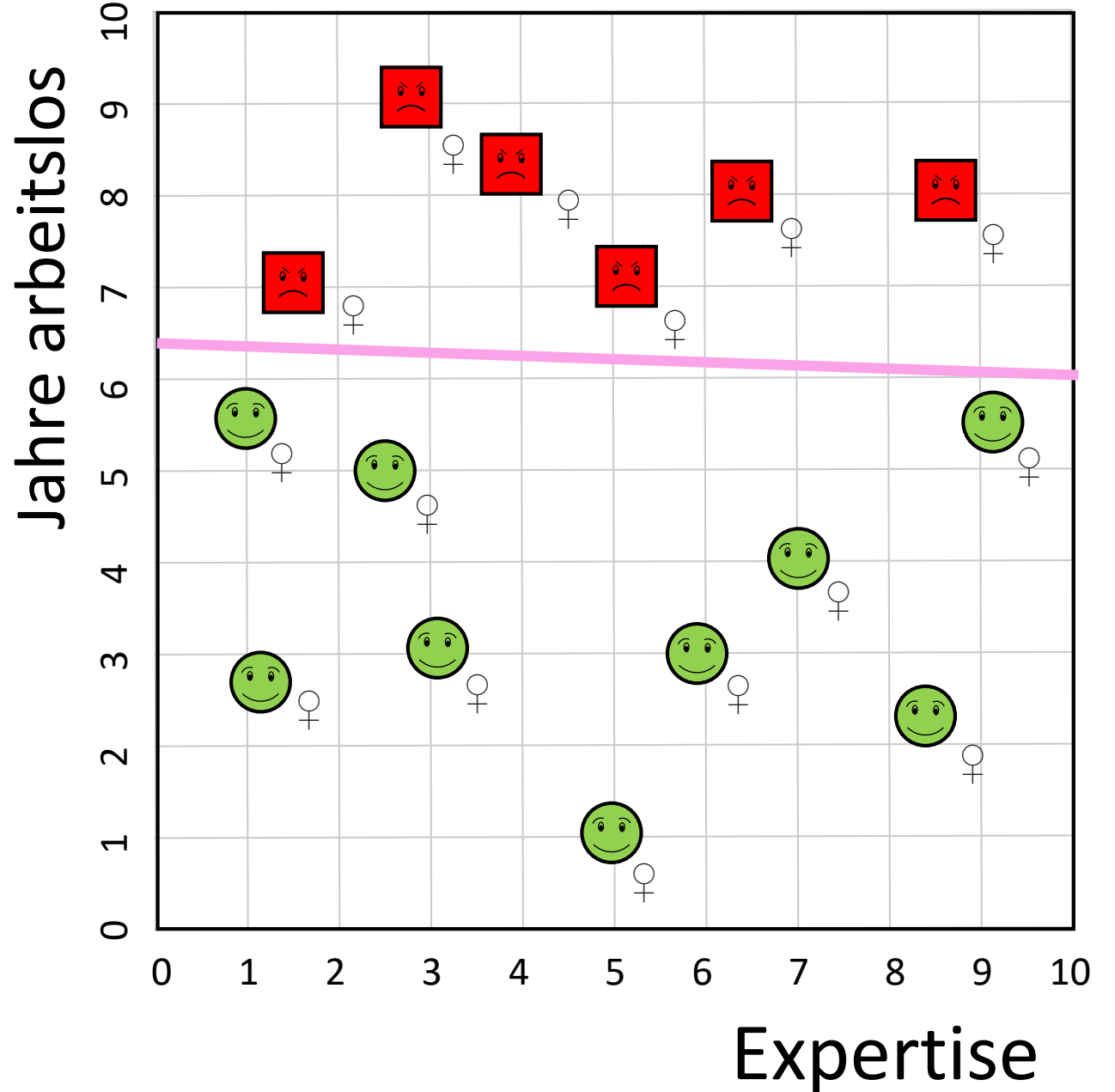
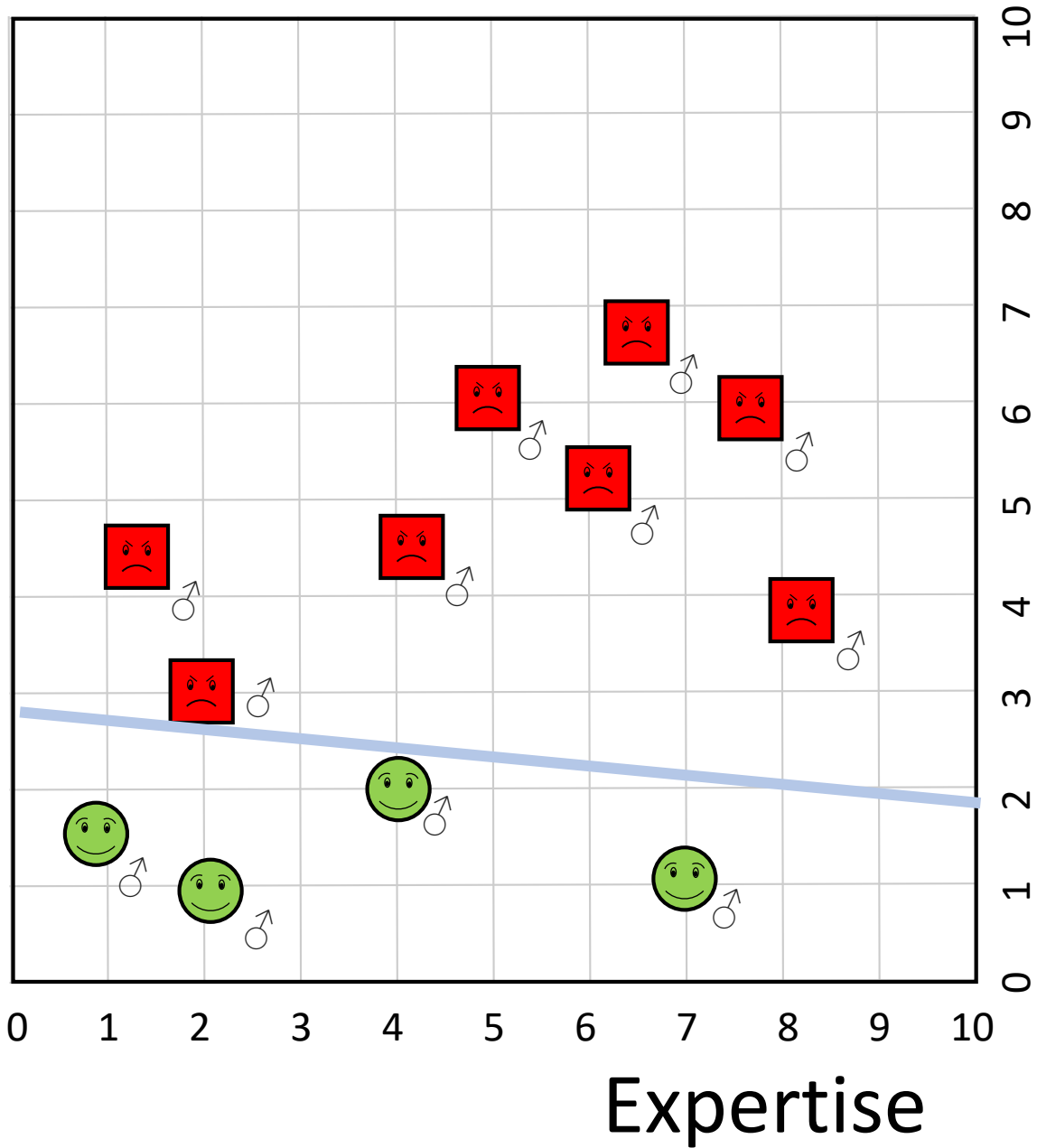
Reicht es, die
sensitive Information wegzulassen?



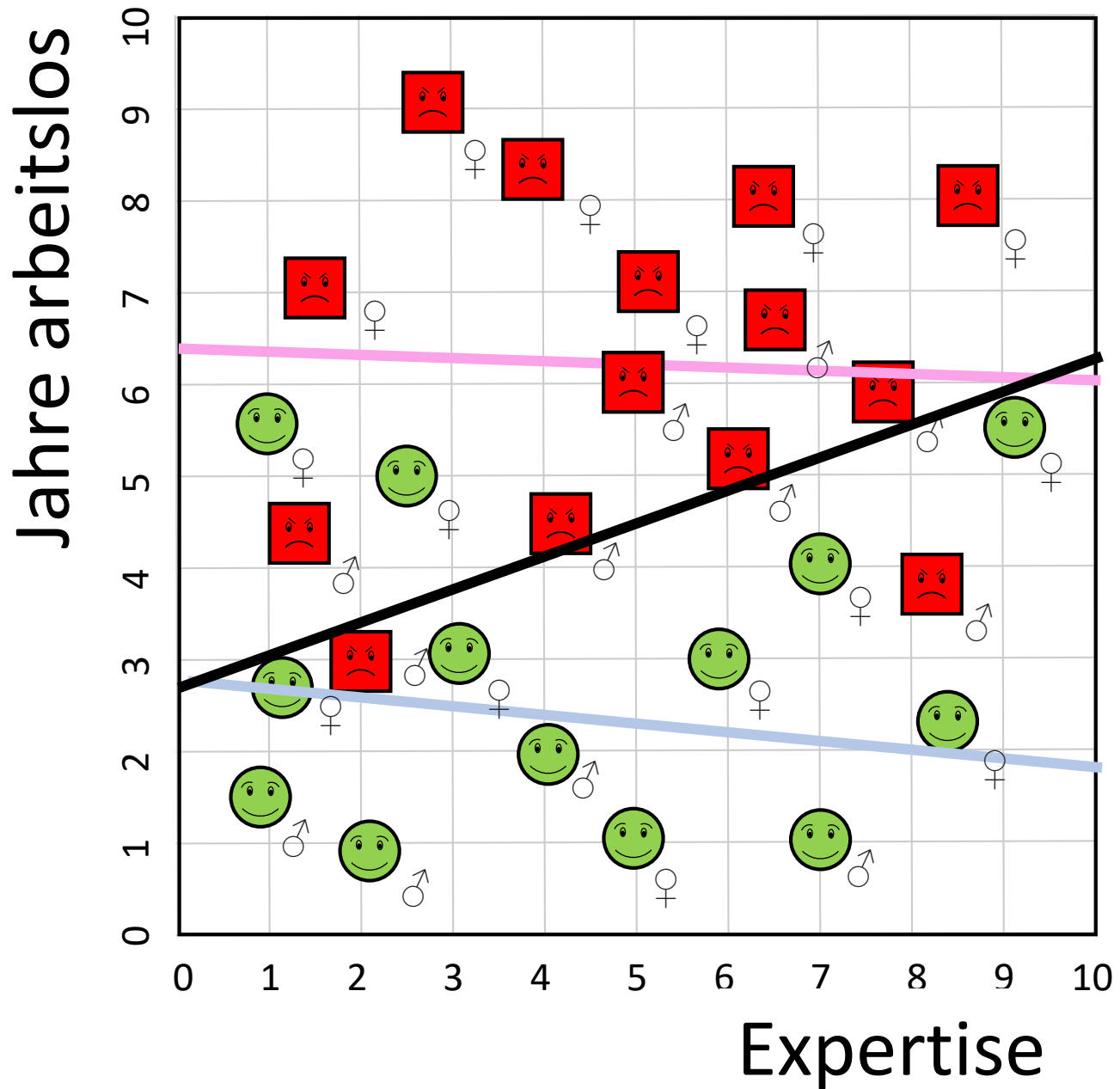
Daten- grundlage

-  Weniger erfolgreiche Arbeitnehmer:innen
-  Erfolgreiche Arbeitnehmer:innen





Effekt:
Wir diskriminieren!



Qualität von ADM Systemen

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





Sozio-informatische Gesamtanalyse



Sozio-informatische
Betrachtungsweise

- Software schafft neue Anreize für menschliche Akteure.
- Diese reagieren auf die Anreize und wirken auf die Software ein.

Probleme der Einbettung des 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

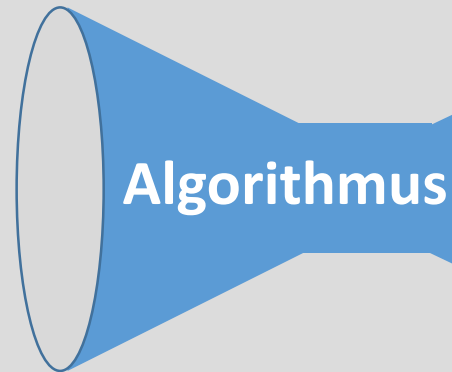


Algorithmische
Entscheidungssysteme
(ADM Systeme)

Bewertete

Nutzer des
ADM Systems

Daten



Soziales System

Scoring-Verfahren

Klasse 1

oder

Klasse 2

Klasse 3

Sozialverträglichkeitsregeln hinter dem AMS-System

- Ergebnis des Algorithmus nur unterstützend einsetzen.
- Ergebnis muss im Dialog mit Betroffenen besprochen werden.
- Trainingsdaten nicht älter als 4 Jahre – das System muss vergessen können.



Qualität von ADM Systemen

1. Wer entscheidet, wann ein ADM System „gut“ ist? Wer, wann es „fair“ ist?
2. ADM Systeme ergeben nur Wahrscheinlichkeiten, keine Wahrheiten.
3. ADM Systeme können diskriminieren.
4. **ADM Systeme bedürfen einer sozio-informatischen Gesamtanalyse.**



Qualität von ADM Systemen

1. Wer entscheidet, wann ein ADM System „gut“ ist? Wer, wann es „fair“ ist?
2. ADM Systeme ergeben nur Wahrscheinlichkeiten, keine Wahrheiten.
3. ADM Systeme können diskriminieren.
4. ADM Systeme bedürfen einer sozio-informatischen Gesamtanalyse.



**Das kann nur mit
Betriebsrat gelingen!**



Zusammenfassung

Was sind die nächsten Schritte?







Über Sie

Über mich

Wie kommt die Ethik in
den Rechner?

Über uns!

Ab Herbst:

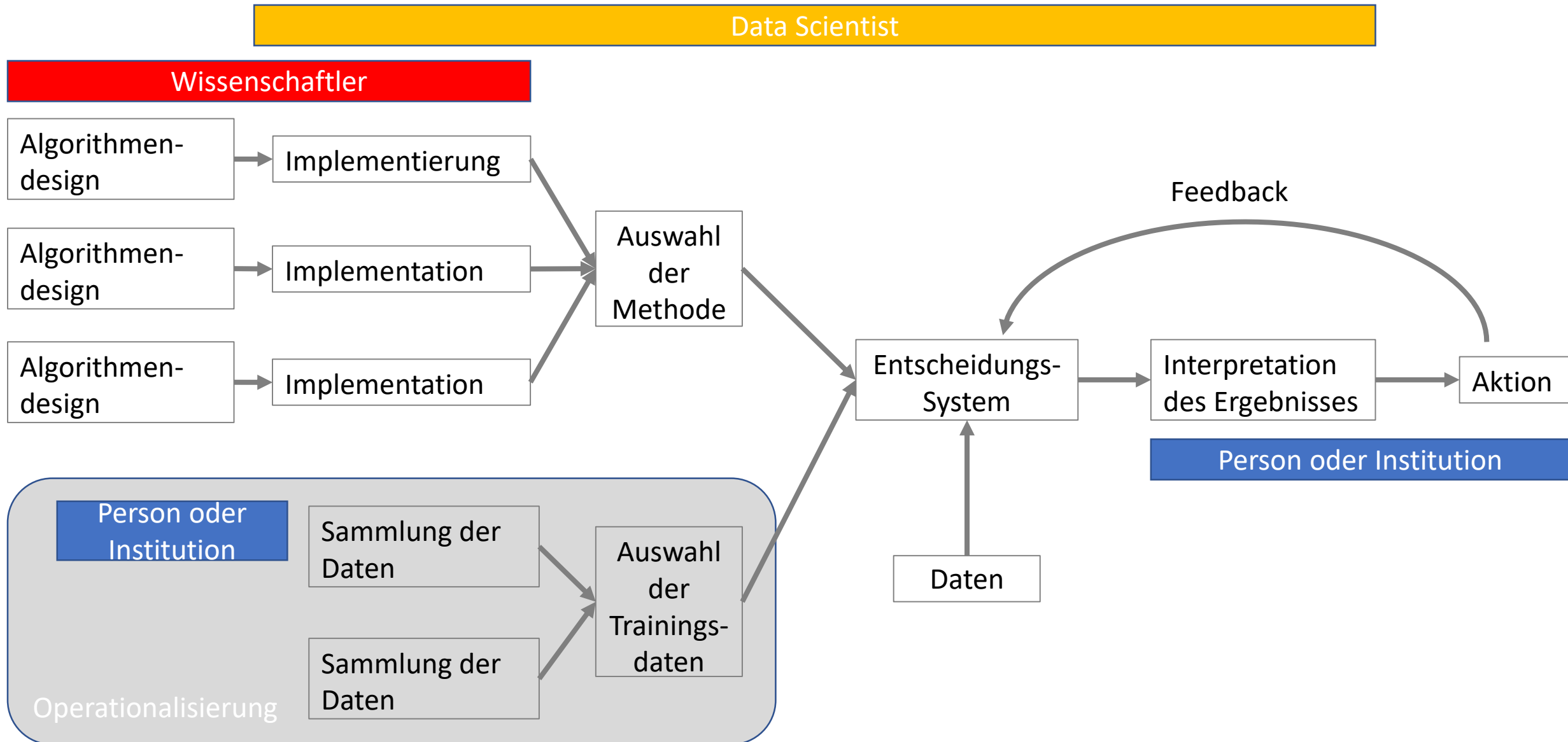


Weitere Literatur



- Katharina Zweig: „Auch Maschinen können diskriminieren“, 16.1.2018, MERTON Magazin, <https://merton-magazin.de/auch-algorithmen-koennen-diskriminieren>
- Studie für die Bertelsmann-Stiftung:
Zweig, Fischer & Lischka: „[Wo Maschinen irren können](#)“ (Serie AlgoEthik, No. 4, 2018)
- [Broschüre der Bayerischen Landesmedienanstalt](#) (Zweig, Krafft & Hauer, 2016): „Dein Algorithmus - meine Meinung“

Lange Kette der Verantwortlichkeiten



Wie bewerten bezüglich der notwendigen Transparenz und Überwachung?

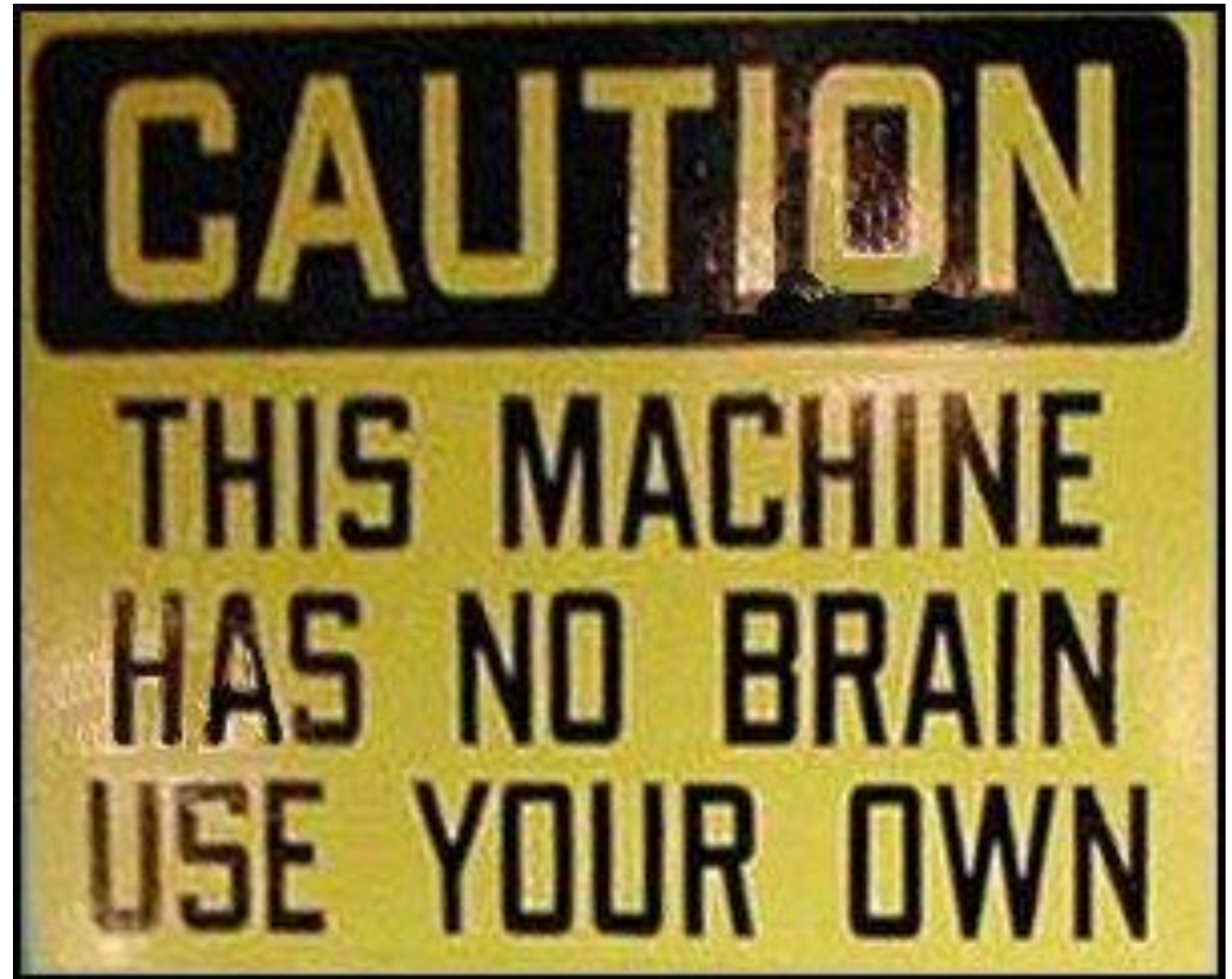
1. Schadenstiefe

$$\Sigma \quad \begin{array}{l} \text{Schaden für Individuum(Fehlurteil)} \\ +\text{Schaden für Gesellschaft(Fehlurteil)} \end{array}$$

2. Anbietervielfzahl und Wechselmöglichkeit

Fünf Risikoklassen

- Klasse 0: Keine technische Regulierung oder Überwachung nötig
- Klasse 1-3: Steigende Anforderungen
- Klasse 4: Einsatz verboten



Risikoklasse HR-Systeme: Helmpflicht....

HR-Systeme sind in Risikoklasse 2 oder 3 und bedürfen der Einmischung, Kontrolle und Überwachung.

