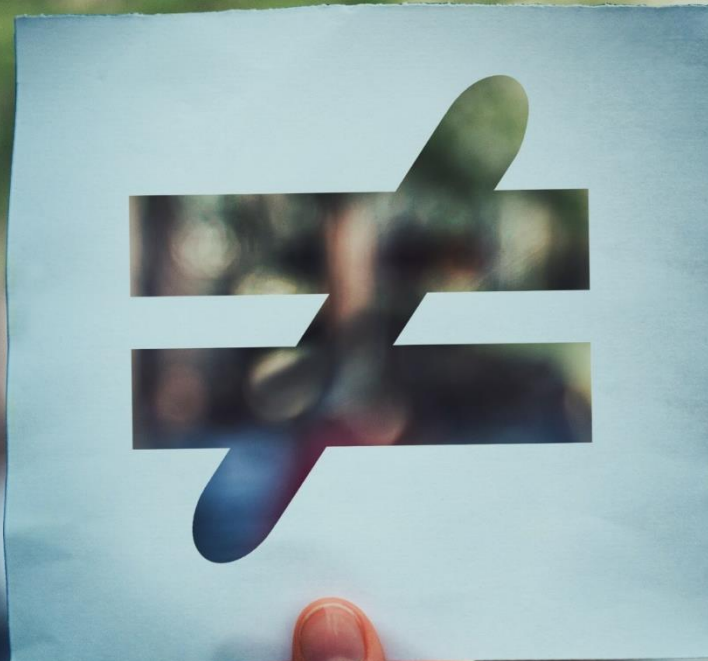




AWK Group

Enabling digital performance.



AI Fairness – Erfahrungen aus der Praxis

Belinda Müller, Consultant

Sophia Ding, Managing Consultant

25.02.2022

Winterkongress, Digitale Gesellschaft

Aufgrund einer externen Untersuchung können keine Beweise für die Diskriminierung von Frauen durch die Apple Card gefunden werden

DHH @dhh · Nov 7, 2019
The @AppleCard is such a fucking sexist program. My wife and I filed joint tax returns, live in a community-property state, and have been married for a long time. Yet Apple's black box algorithm thinks I deserve 20x the credit limit she does. No appeals work.

Steve Wozniak @stevewoz

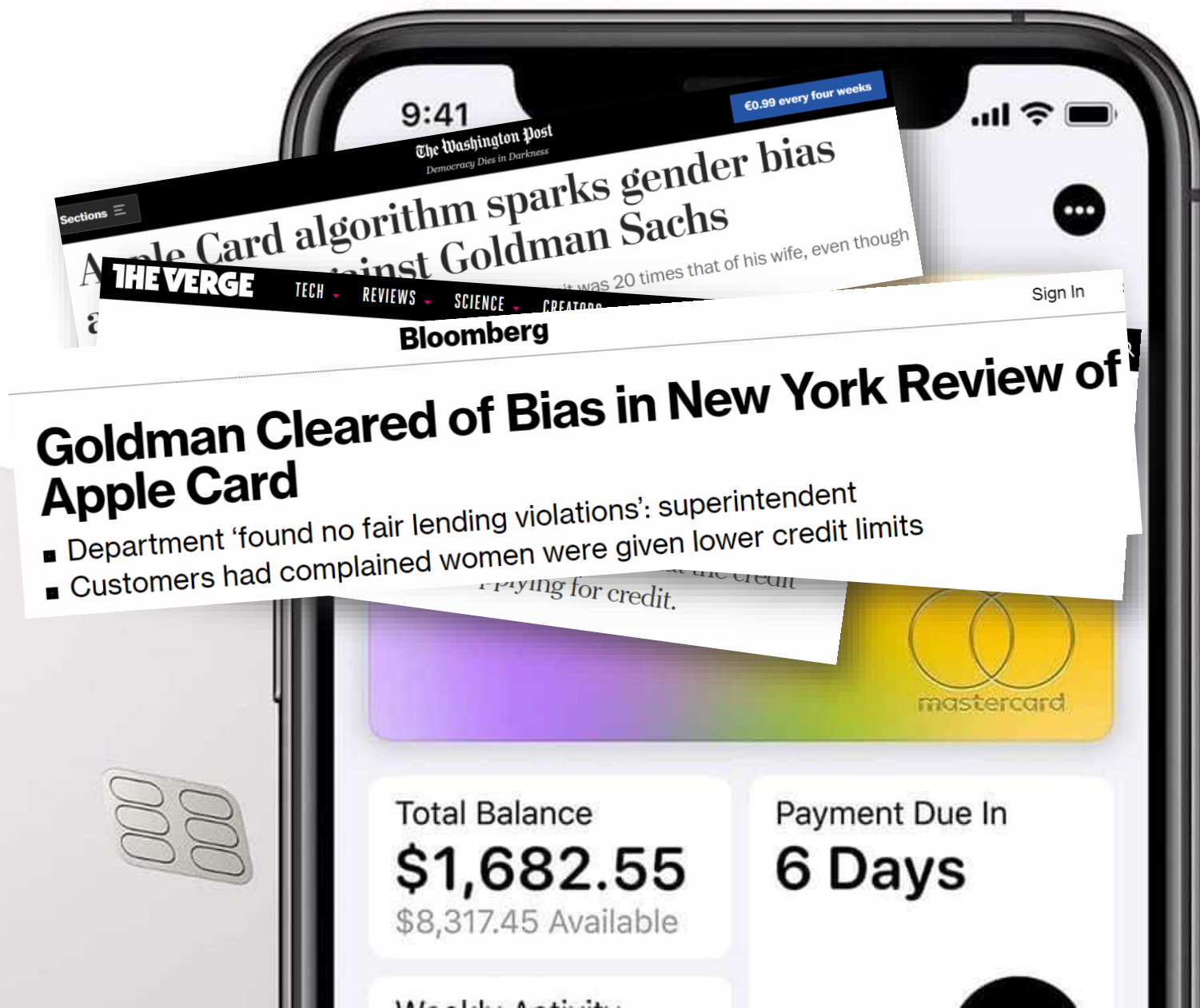
The same thing happened to us. I got 10x the credit limit. We have no separate bank or credit card accounts or any separate assets. Hard to get to a human for a correction though. It's big tech in 2019.

1:51 AM · Nov 10, 2019

3.9K Reply Share this Tweet

Read 115 replies

Marisa Robertson



Diverse Meinungen beim Frühstück...

«Männer sollten im Schnitt keinen Kreditrahmen bekommen als /pft

emand sollte Apple verklagen.»

The same thing happened to us. I g
limit. We have no separate bank or
accounts or any separate assets. He
human for a corre

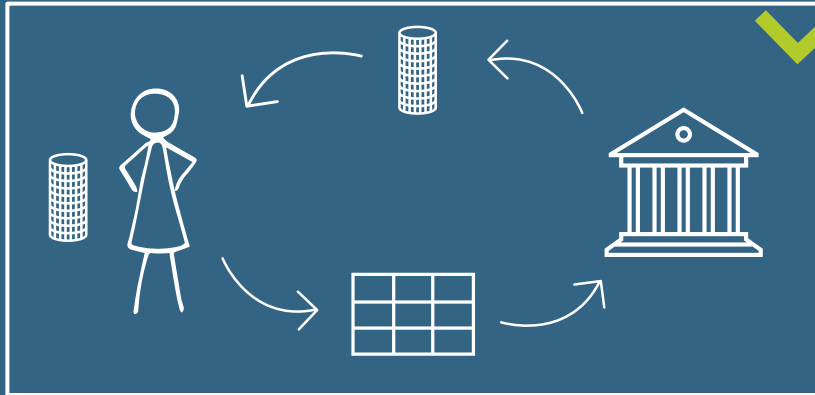
1:51 AM · Nov 10, 2019

«Man mus.

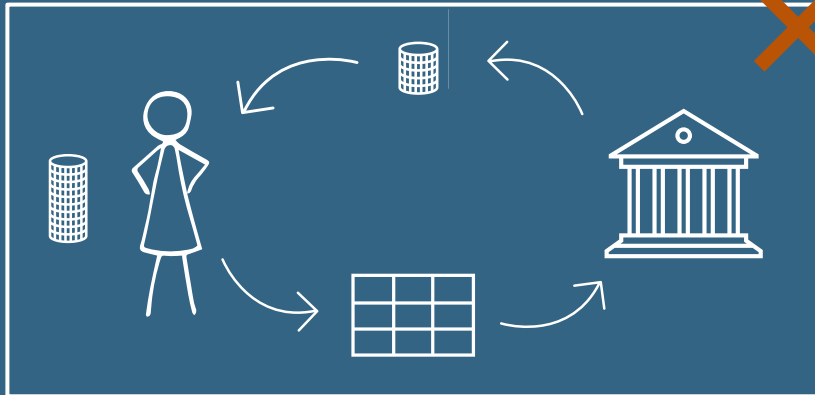
System nicht 100% fair ist,
wir es nicht einsetzen.»



Ethische Perspektive: Was bedeutet Fairness überhaupt?



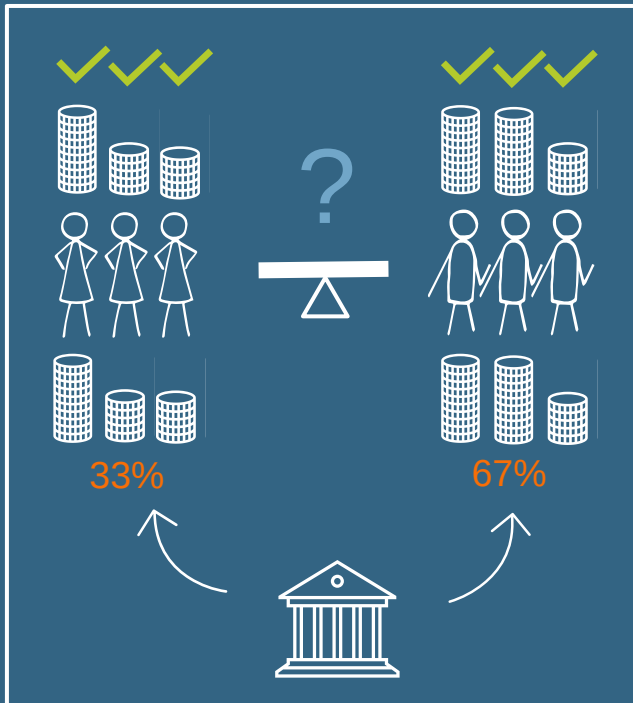
Korrekte Entscheidung:
angemessener Kreditrahmen



Inkorrekte Entscheidung: zu
niedriger Kreditrahmen



Ethische Perspektive: Was bedeutet Fairness überhaupt?



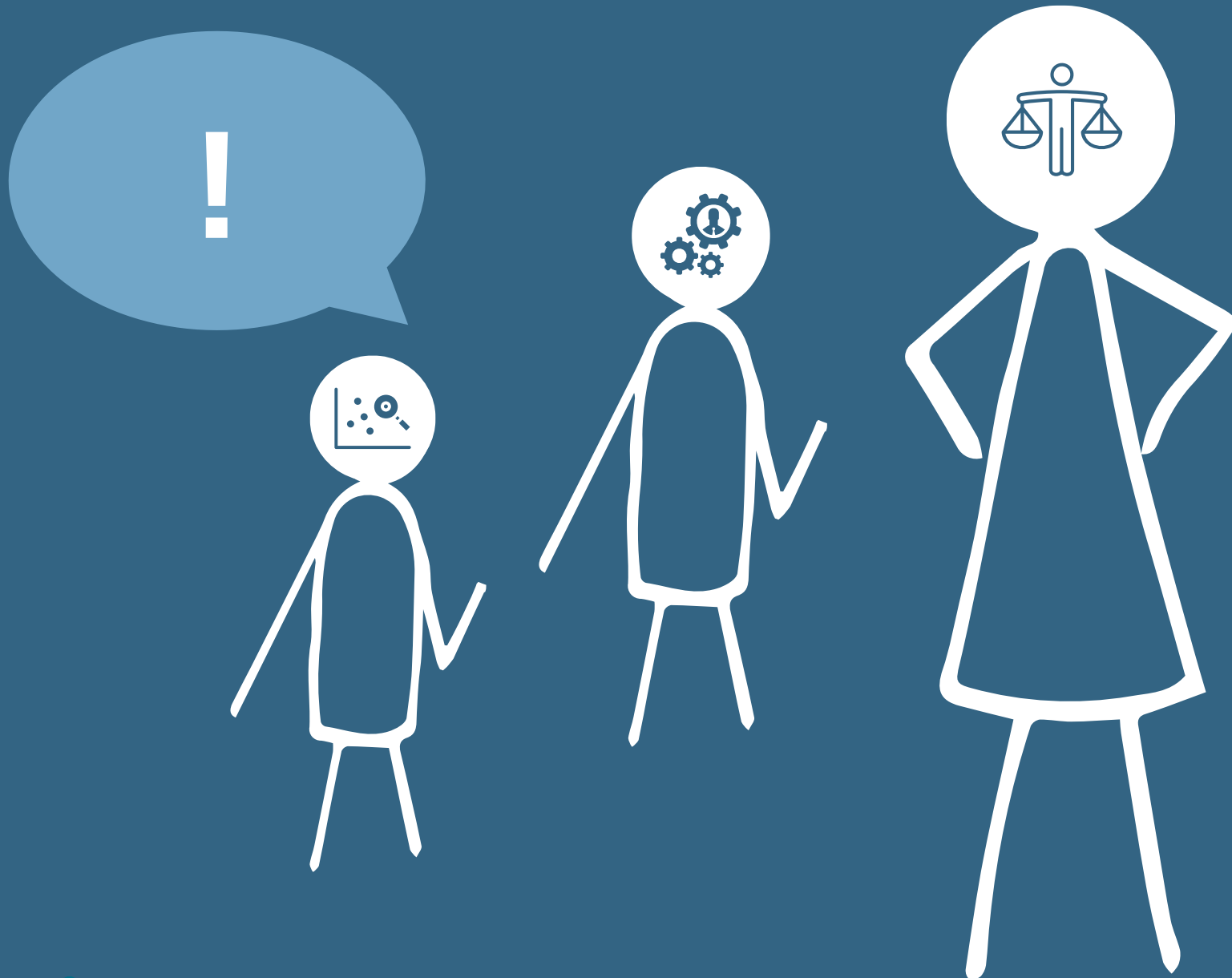
A Gleicher Anteil an hohen/niedrigen Kreditrahmen bei Frauen und Männern

B Gleiche Behandlung bei gleicher Kreditwürdigkeit

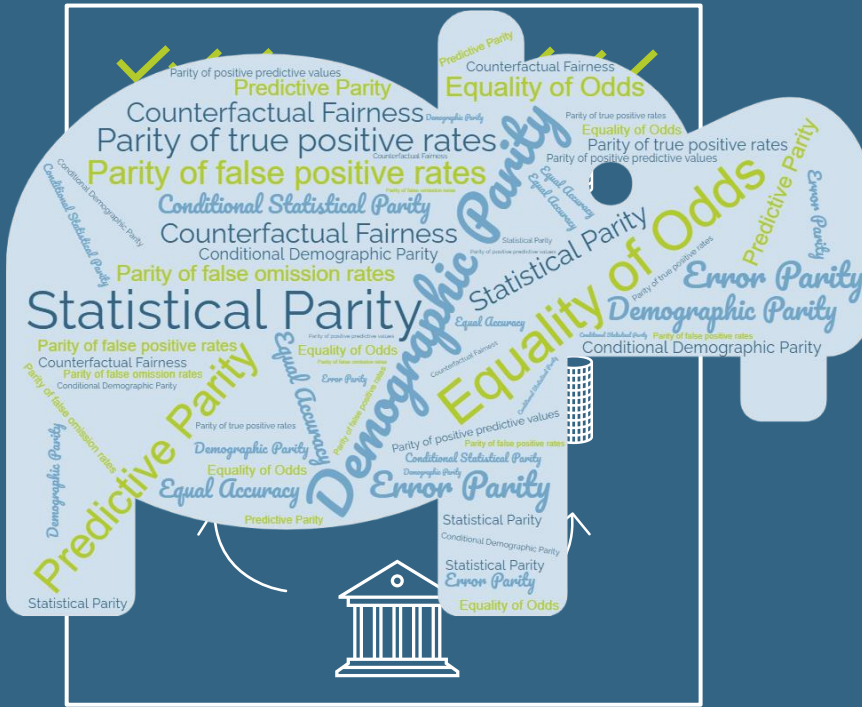
«Männer sollten im Schnitt keinen höheren Kreditrahmen bekommen als Frauen.»

«Frauen und Männer sollten den gleichen Kreditrahmen bekommen, wenn sie die gleichen finanziellen Grundlagen haben.»

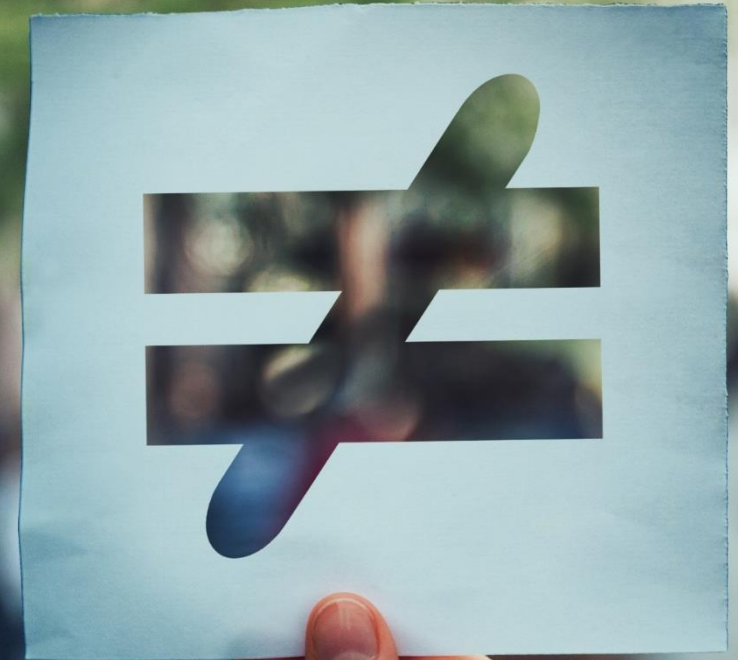
Ethische Perspektive: Wessen Expertise ist gefragt?



Ethische Perspektive vs. quantitative Perspektive



Wie übersetzen wir
den Fairness- Begriff
in Kennzahlen?



A Gleicher Anteil an
hohen/niedrigen Kreditrahmen bei
Frauen und Männern → Statistische Parität

B Gleiche Behandlung bei
gleicher Kreditwürdigkeit → Parität der Fehlerrate (z.B. Falsch-
Negativ-Rate)

Quantitative Perspektive

Welche Gruppen
vergleichen wir?

«Man muss sich einfach
die Kennzahlen
anschauen.»

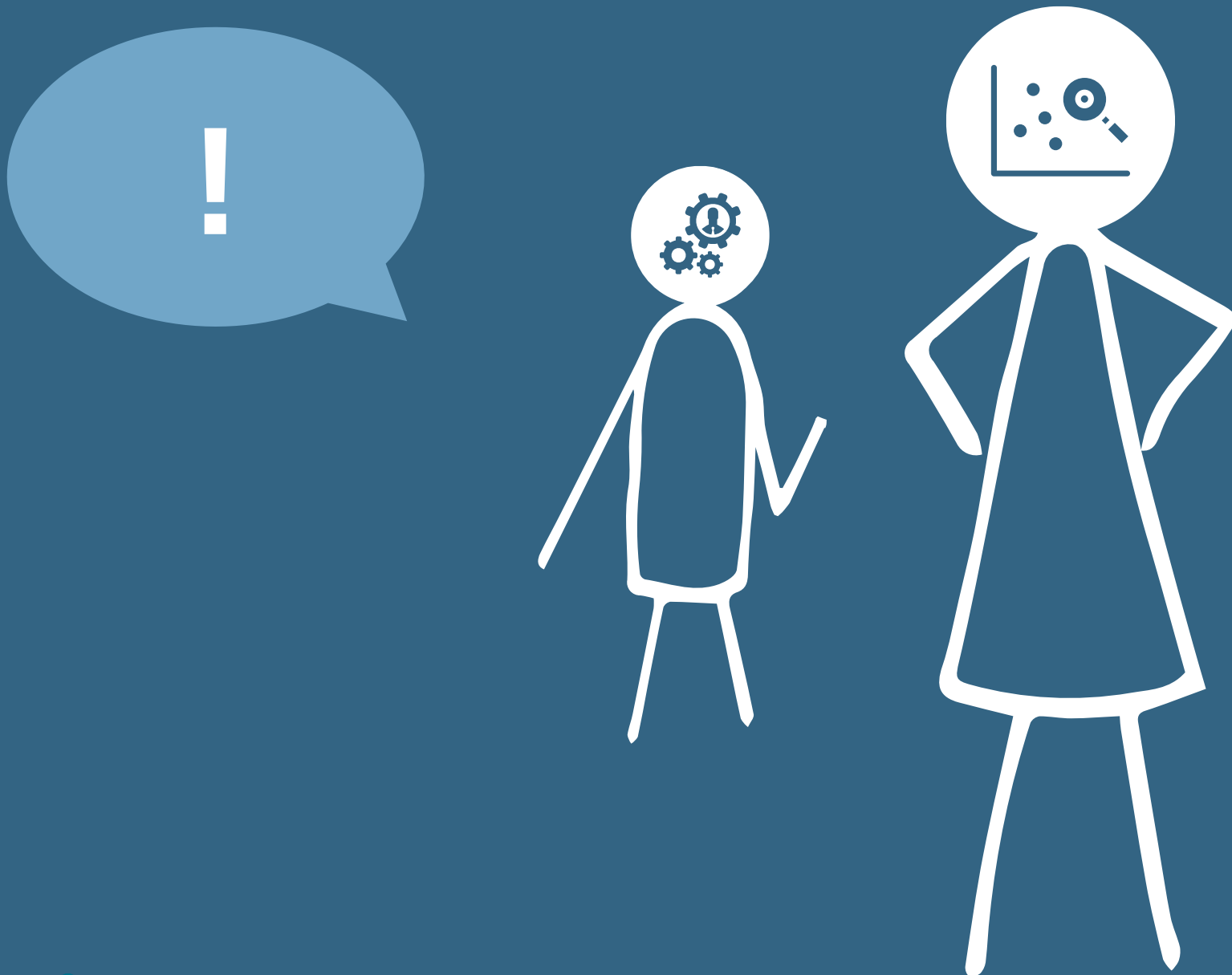


Was machen wir
mit kleinen
Gruppen?

Ab wann ist ein
Unterschied
signifikant und
relevant?

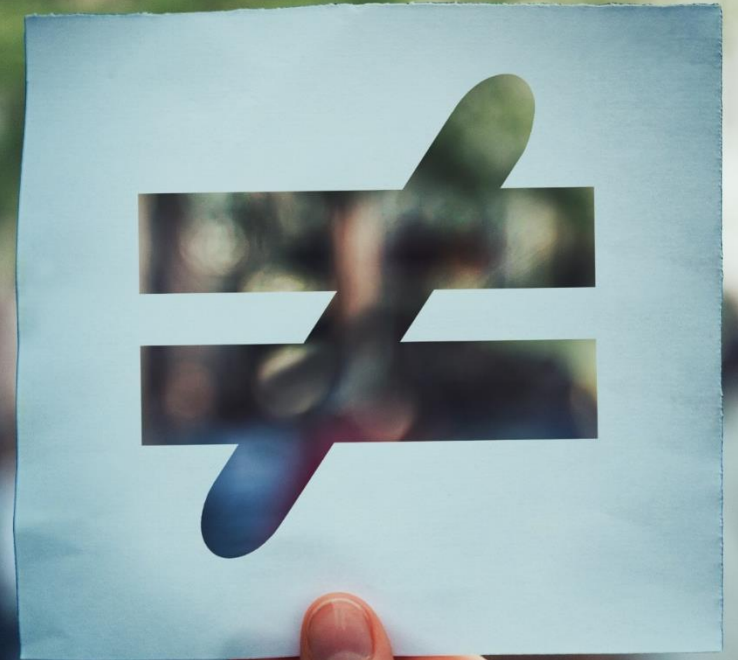


Quantitative Perspektive: Wessen Expertise ist gefragt?



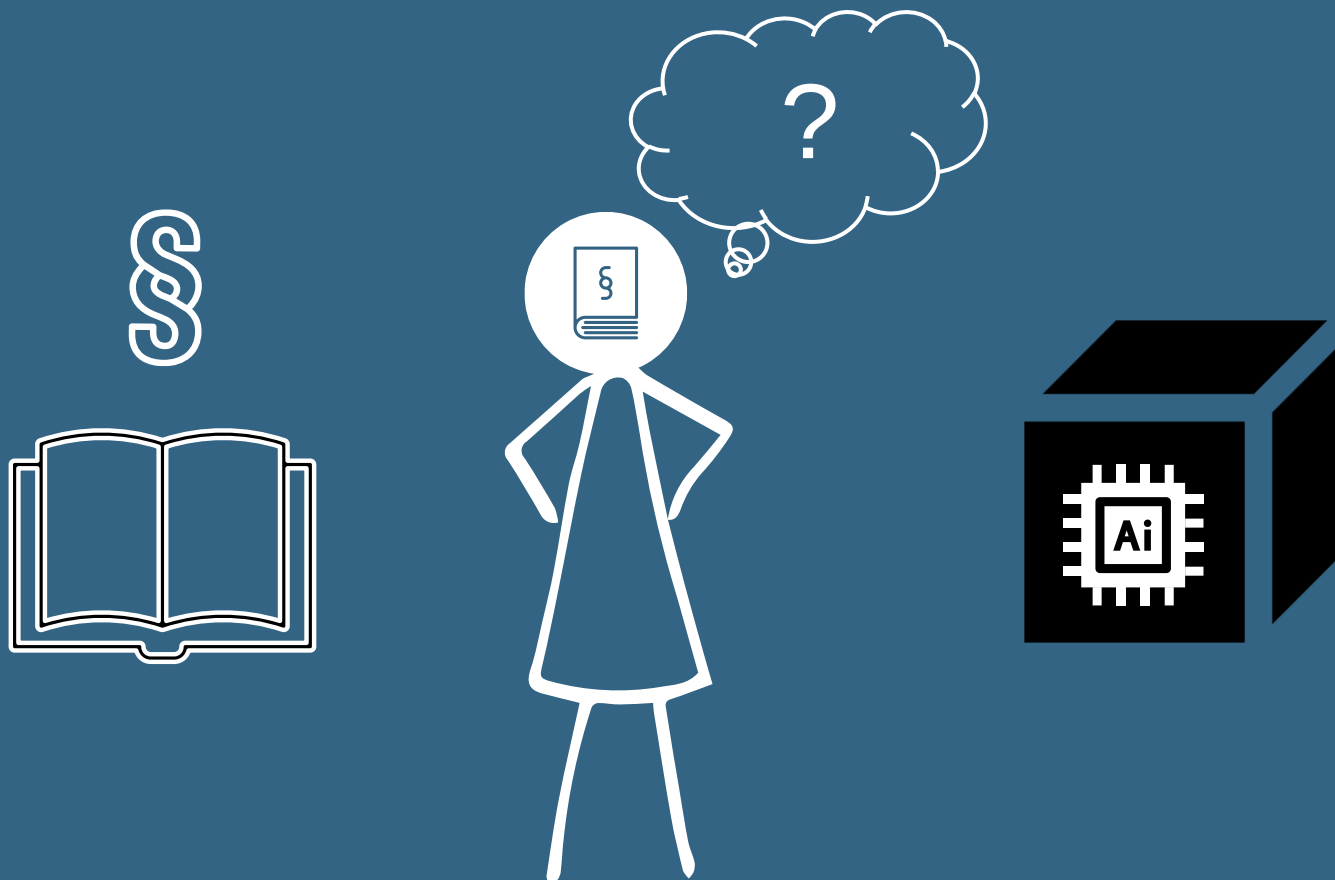
Ethische Perspektive vs. juristische Perspektive

«Jemand sollte Apple verklagen.»



Juristische Perspektive: Gibt es einen Gesetzesverstoss?

«Jemand sollte Apple verklagen.»

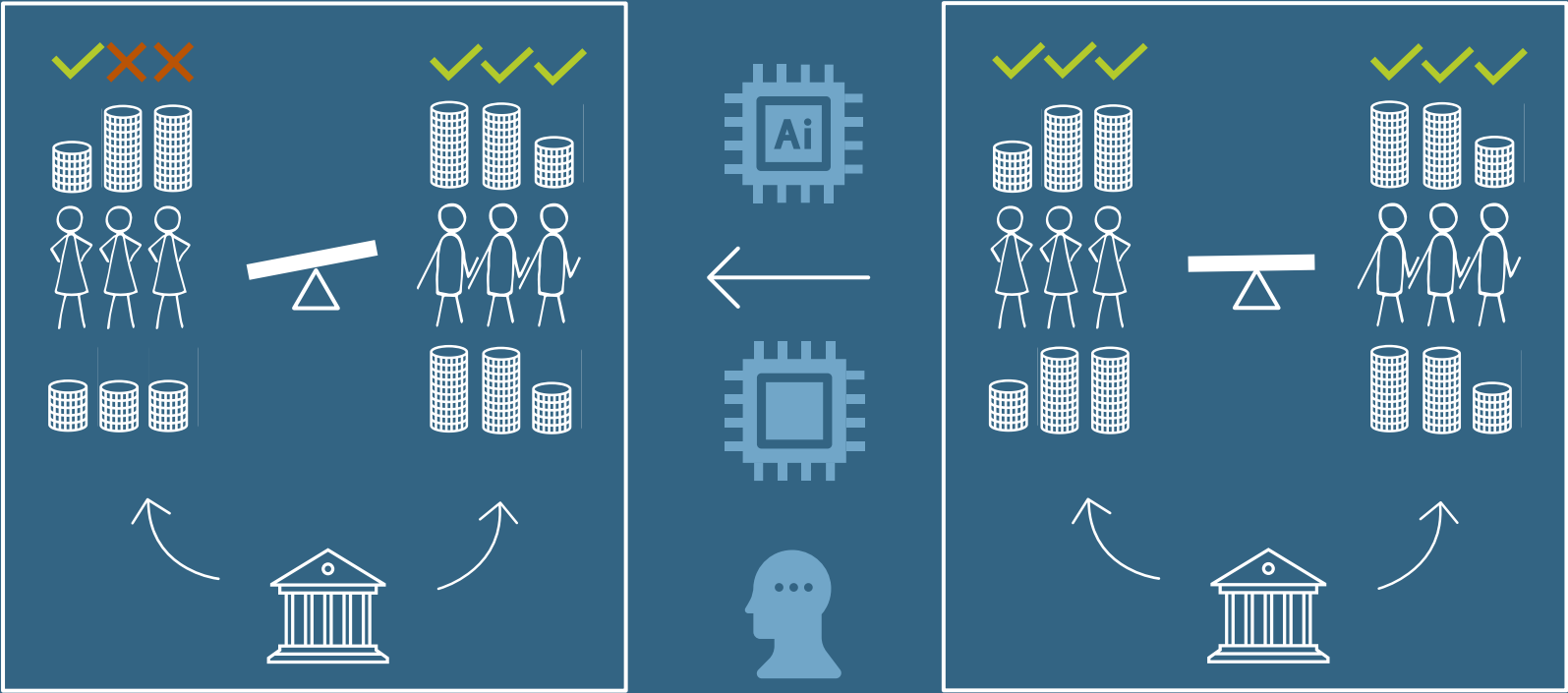


Juristische Perspektive: Wessen Expertise ist gefragt?



Gesellschaftliche Perspektive: Wie fair muss ein KI System sein?

«Wenn ein KI-System nicht 100% fair ist, sollten wir es nicht einsetzen.»



Diverse Meinungen beim Frühstück...

«Männer sollten im Schnitt keinen höheren Kreditrahmen bekommen als Frauen.»

The same thing happened to us. I got 10x the credit limit. We have no separate bank or credit card accounts or any separate assets. Hard to get to a human for a correction.

1:51 AM · Nov 10, 2019

3.9K Reply

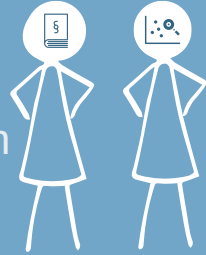
«Man muss sich einfach die Kennzahlen anschauen.»

«Jemand sollte Apple verklagen.»

«Wenn ein KI-System nicht 100% fair ist, sollten wir es nicht einsetzen.»

Diverse Meinungen beim Frühstück...

Die relevante Definition von Fairness sollte auf einer Diskussion mit möglichst vielen Stakeholdern abgestützt sein.



The same thing happened to us. I got 10x the credit limit. We have no separate bank or credit card accounts or any separate assets. Hard to get to a human for a correction.

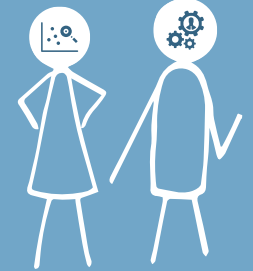
1:51 AM · Nov 10, 2019

3.9K Reply

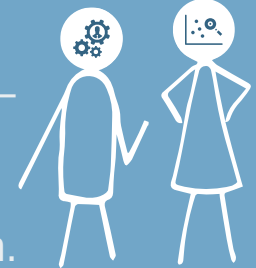
Eine Analyse der Kennzahlen alleine ist nicht aussagekräftig, sondern muss mit Hilfe von technischem, fachlichem und ethischem Expert:innenwissen interpretiert werden.



Eine messbare Ungleichheit bedeutet nicht automatisch eine Diskriminierung im juristischen Sinne.



Der Versuch soziale Probleme technisch zu lösen kann nur scheitern – allerdings hilft uns die Technik ggf. soziale Probleme zu identifizieren.



Apple Card revisited...

Bei der Beurteilung des Apple Cases wurden zwei Definitionen von Fairness herangezogen, die im Gesetz über Kreditvergaben verankert sind. Unklar ist, inwiefern Stakeholder miteinbezogen wurden.



The same thing happened to us. I got 10x the credit limit. We have no separate bank or credit card accounts or any separate assets. Hard to get to a human for a correction.

1:51 AM · Nov 10, 2019

3.9K Reply

Es ist strittig, ob die Art und Weise der Evaluation der Fairness-Kriterien heutigen Standards entspricht.



Für diesen Use Case existiert eine gesetzliche Grundlage, basierend auf der evaluiert werden kann, ob eine Diskriminierung im rechtlichen Sinne vorliegt.



Für alle betrachteten Einzelfälle konnten Gründe gefunden werden, warum die Kreditwürdigkeit der Ehepartner:innen nicht identisch war.



Kontaktiert uns!



**Sophia Ding,
Managing Consultant**

Sophia.Ding@awk.ch

Sophia Ding



- Spezialisiert sich auf die ethischen Aspekte neuer Technologien mit besonderem Augenmerk auf vertrauenswürdige künstliche Intelligenz (AI Trustworthiness) und die praktische Umsetzung ethischer Grundsätze einschliesslich der Anwendbarkeit technischer Hilfsmittel zu diesem Zweck.
- Hintergrund in den Bereichen Datenanalyse, Analytics Translation, empirische Wirtschaftsforschung mit Schwerpunkt auf statistischen und ökonometrischen Methoden sowie der Schnittstelle zwischen Philosophie und Wirtschaft



**Belinda Müller,
Consultant**

Belinda.Mueller@awk.ch

Belinda Müller



- Leidenschaft für den Einsatz von KI zur Erzielung sozialer Wirkung bei gleichzeitiger Abmilderung der damit verbundenen Risiken
- Ausgebildet in Statistik, Kognitionswissenschaft und Linguistik mit starkem Interesse an interdisziplinärer Arbeit
- Berufserfahrung als Data Scientist mit Schwerpunkt Natural Language Processing

