SCHLOSS DAGSTUHL
Leibniz-Zentrum für Informatik

Dagstuhl Perspectives Workshop 19482

TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

Algorithmic accountability and fairness

-

A computer scientist's perspective

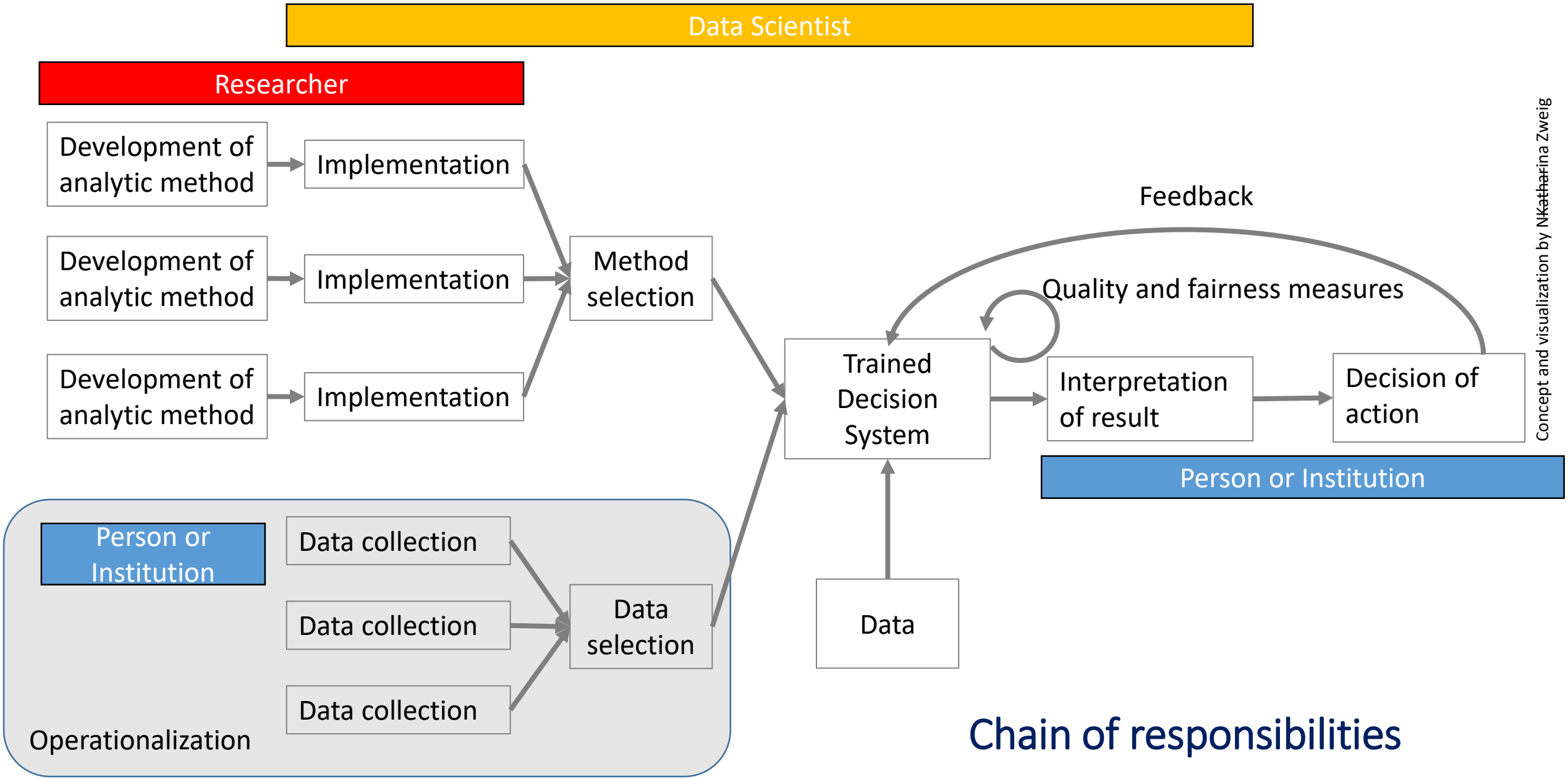
Marc P. Hauer, M.Sc.
Algorithm Accountability Lab, TU Kaiserslautern

What is algorithmic accountability?

- Addresses problems with algorithms that interact with society and affect it, e.g. ADM-systems (algorithmic decision making) - especially the learning ones
- Who is accountable?

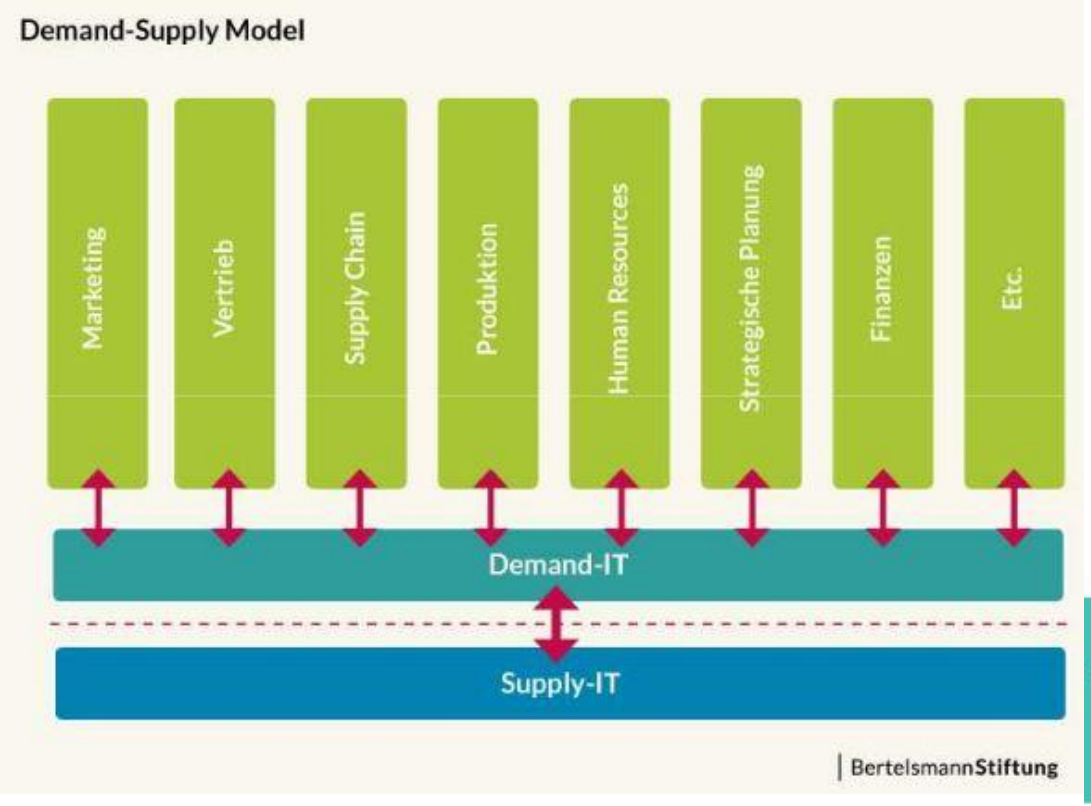
What is algorithmic accountability?

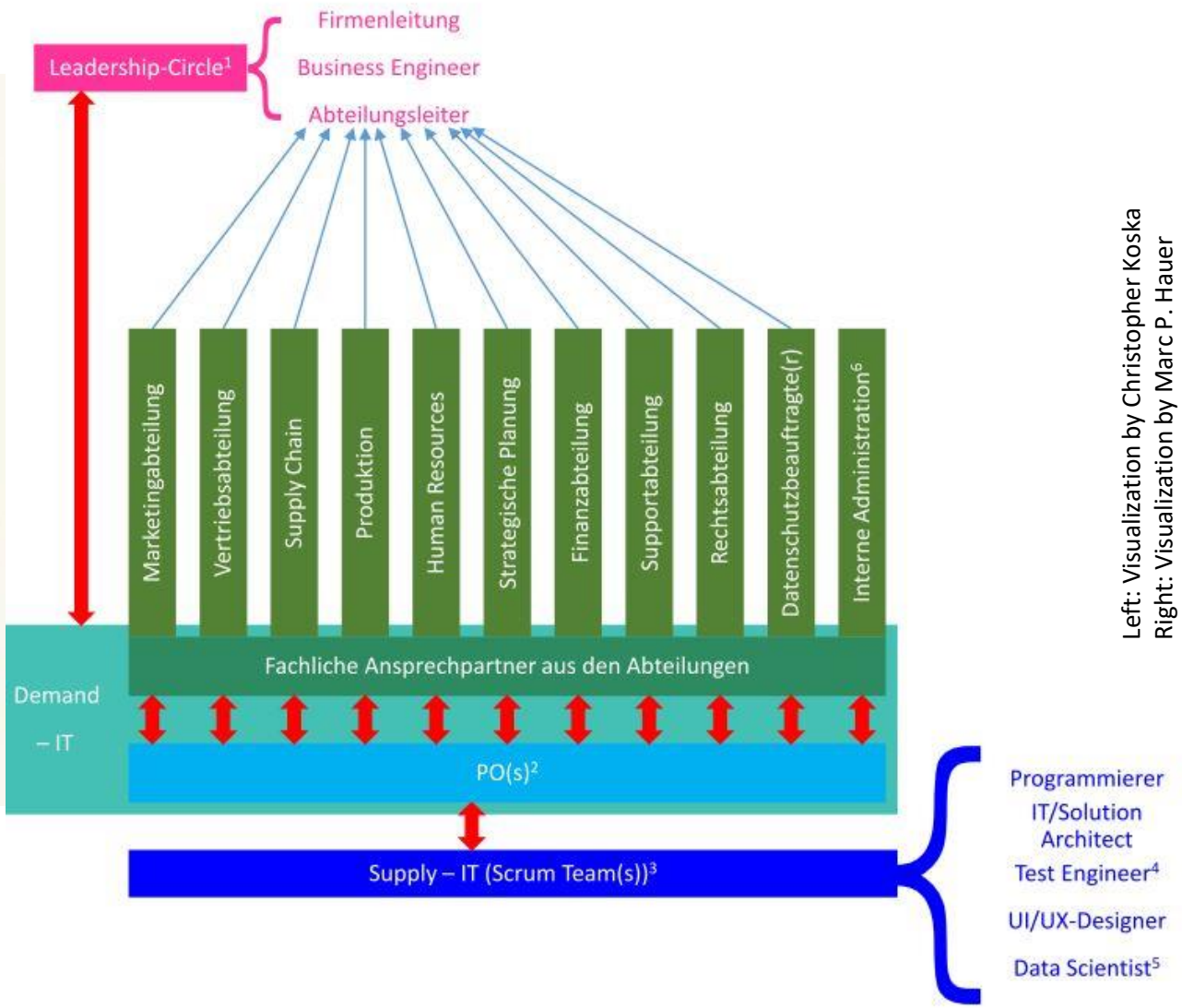
- Addresses problems with algorithms that interact with society and affect it, e.g. ADM-systems (algorithmic decision making) - especially the learning ones
- Who should feel accountable?
- What is fair?
- How can we implement algorithmic accountability?



Chain of responsibilities

Who should feel accountable?





Left: Visualization by Christopher Koska
Right: Visualization by Marc P. Hauer

Demand-Supply Models

Definitions of Fairness

01

Oxford Dictionary

The quality of treating people **equally** or in a way that is right or **reasonable**.

02

Lexico Dictionary

Impartial and just treatment or behavior **without favoritism** or **discrimination**.

03

Webster's Dictionary

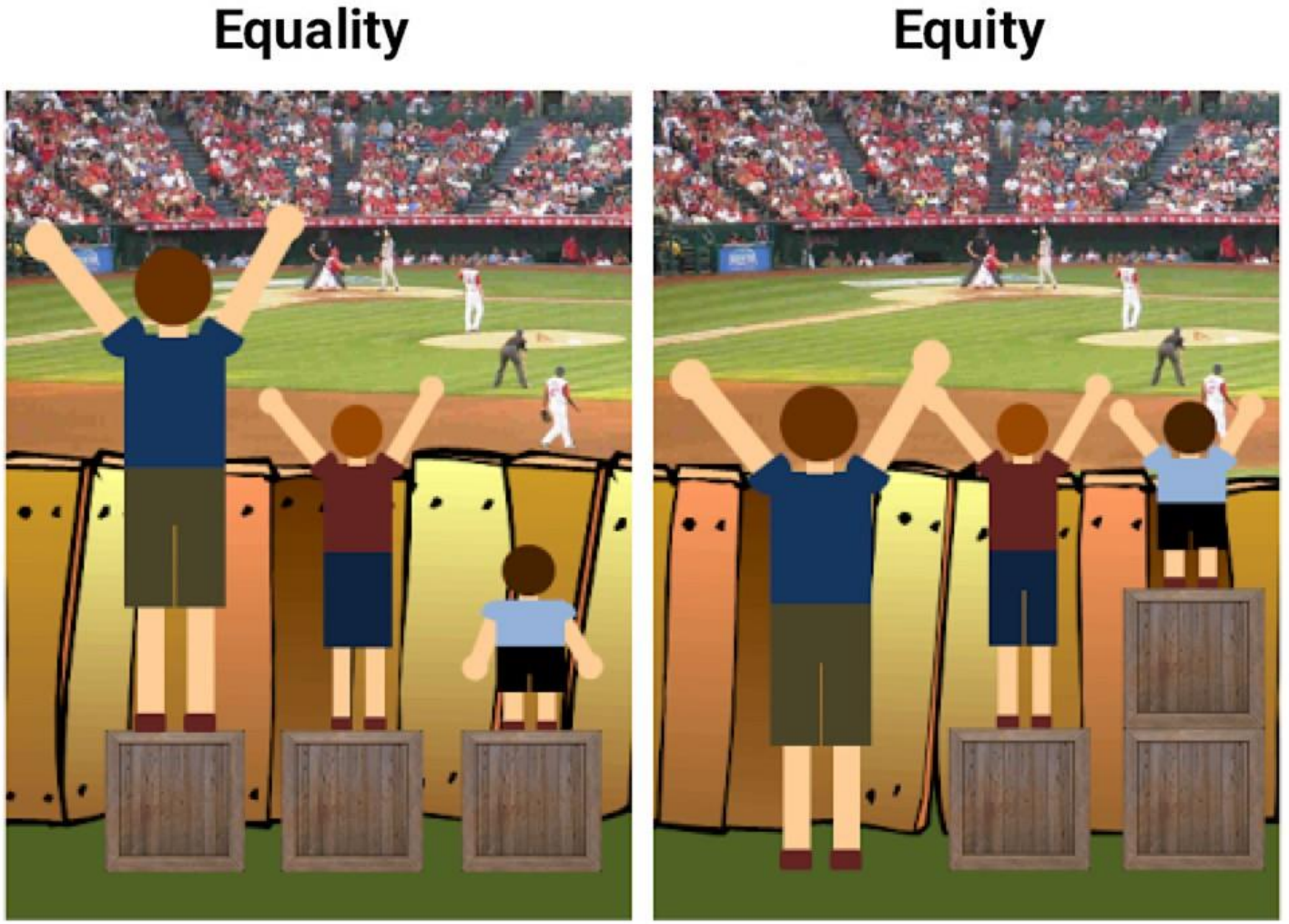
Fair or **impartial treatment**: **lack of favoritism** toward one side or another

04

Collins Dictionary

Fairness is the quality of being reasonable, **right** and **just**.

Definitions of Fairness



https://edtrust.org/the-equity-line/equity-and-equality-are-not-equal/

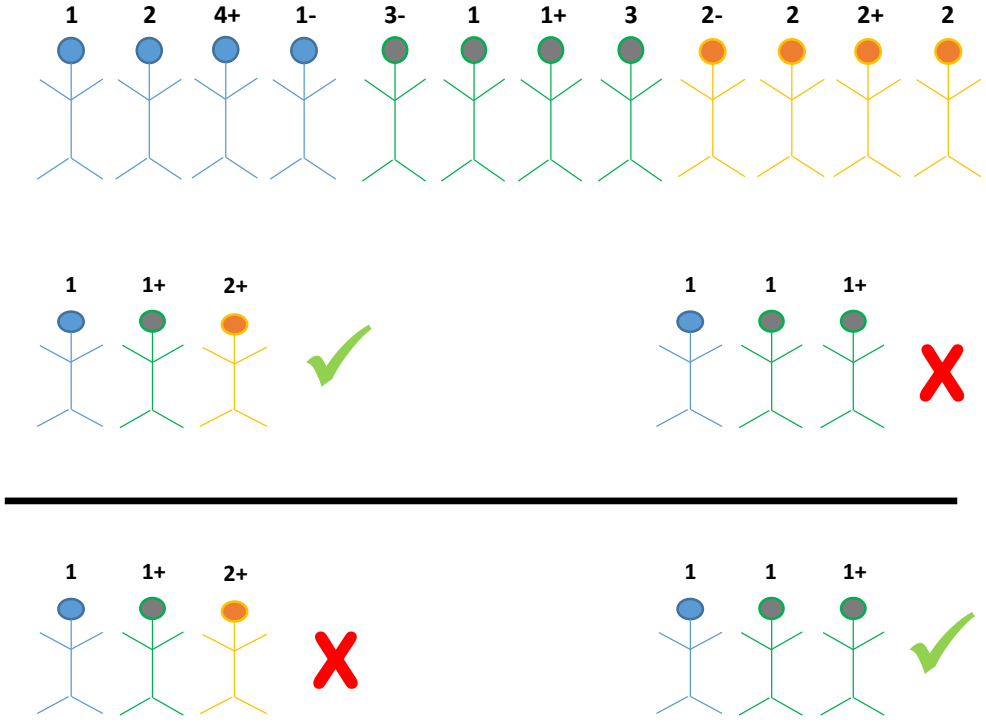
Group fairness vs. Individual fairness

Group fairness:

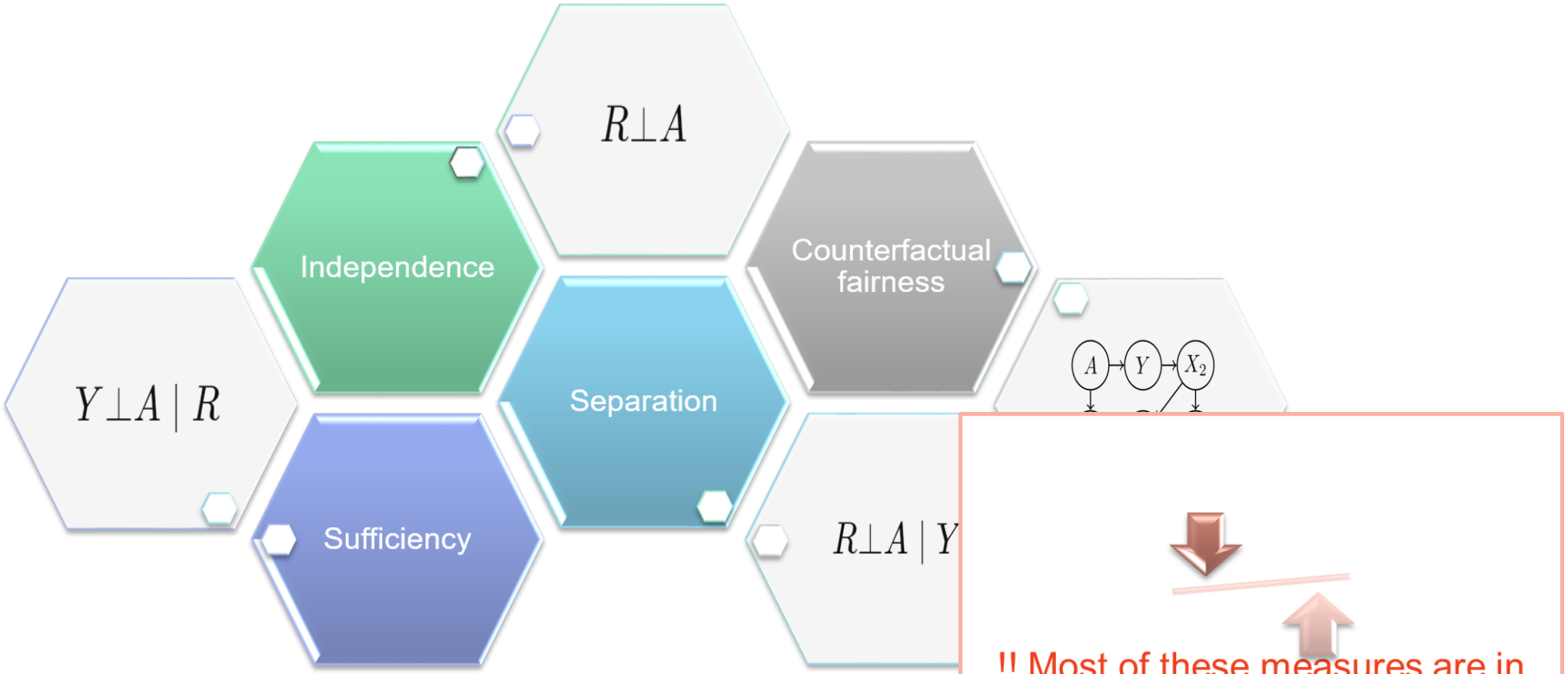
- Protected groups should be treated similarly to the advantaged group or the populations as a whole.
- Does not consider the individual merits.
- May result in choosing the less qualified members of a group.

Individual fairness:

- Individuals should be treated consistently.
- Assumes a similarity metric of the individuals that may be hard to find.
- This kind of fairness is rarely used



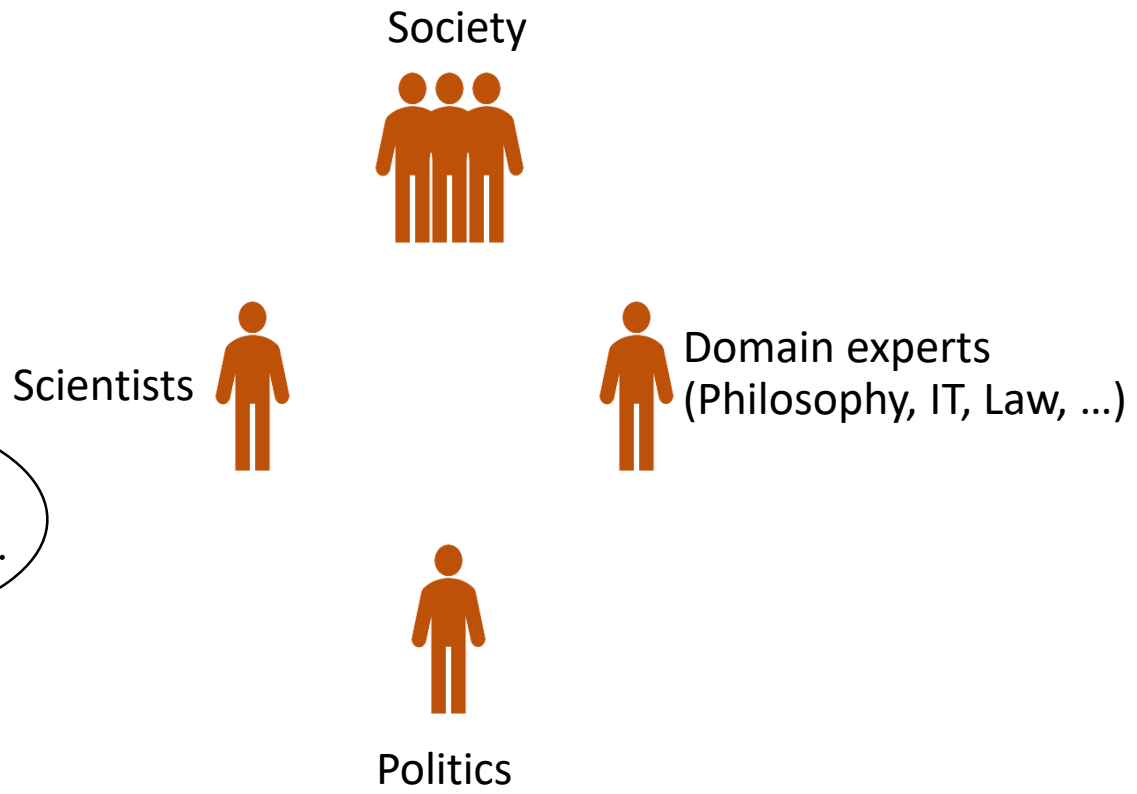
~~Diversity~~
(Group) Fairness measures



Selection and visualization by Maryam Amir Haeri

Fairness measures

- Independence
  - Relaxed Independence
  - Conditional Independence
- Separation
- Equalized Correlation
- Overall Accuracy Equality
- Sufficiency
- Conditional Use Accuracy
- Well Calibration
- Treatment Equality
- False Positive Error Rate Balance
- Balance for Positive Class



Fairness measures

- Independence
  - Relaxed Independence
  - Conditional Independence
- Separation
- Equalized Correlation
- Overall Accuracy Equality
- Sufficiency
- Conditional Use Accuracy
- Well Calibration
- Treatment Equality
- False Positive Error Rate Balance
- Balance for Positive Class

Society

Scientists

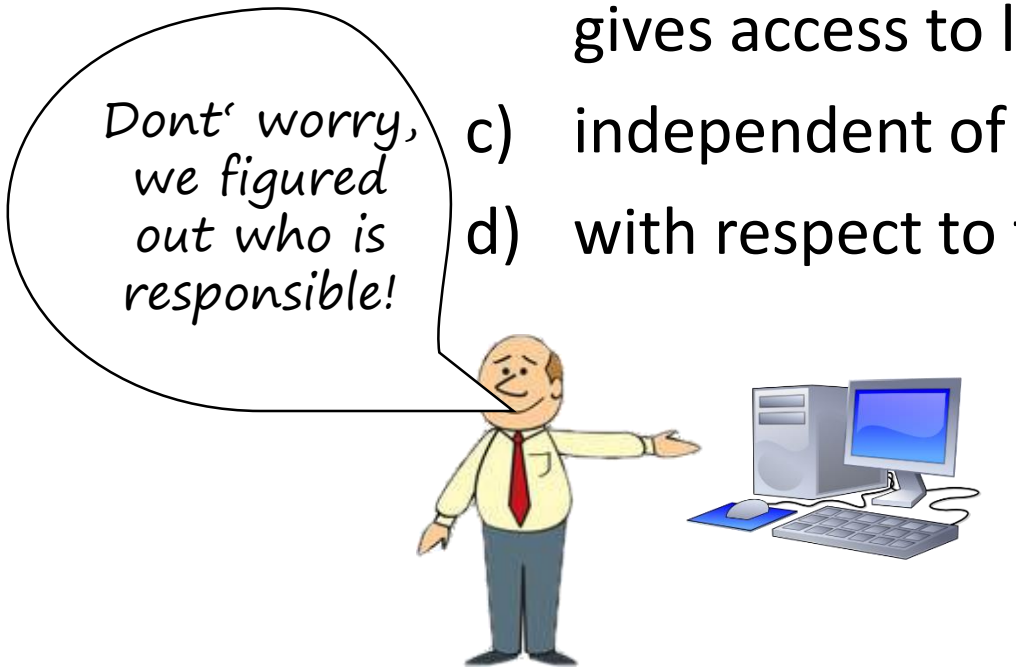
Domain experts
(Philosophy, IT, Law, …)

Politics

When do we need regulation?

ADM-system need to be regulated, normed, and/or controlled if they

a) contain a learning or learned component
b) that makes decisions about humans or their belongings or that gives access to limited resources
c) independent of whether a human is in the loop or not
d) with respect to the logic and mechanism of their decision making.



Why only then?

Both need to be checked for product safety and security

ADM systems deciding about things

Additional need to check for

- Bias in data
- Data quality and representativeness
- Correct operationalization of human values
- Result fairness and quality and
- Justification and explainable decision making (for the possibility of appealing)

ADM systems deciding about people and resources

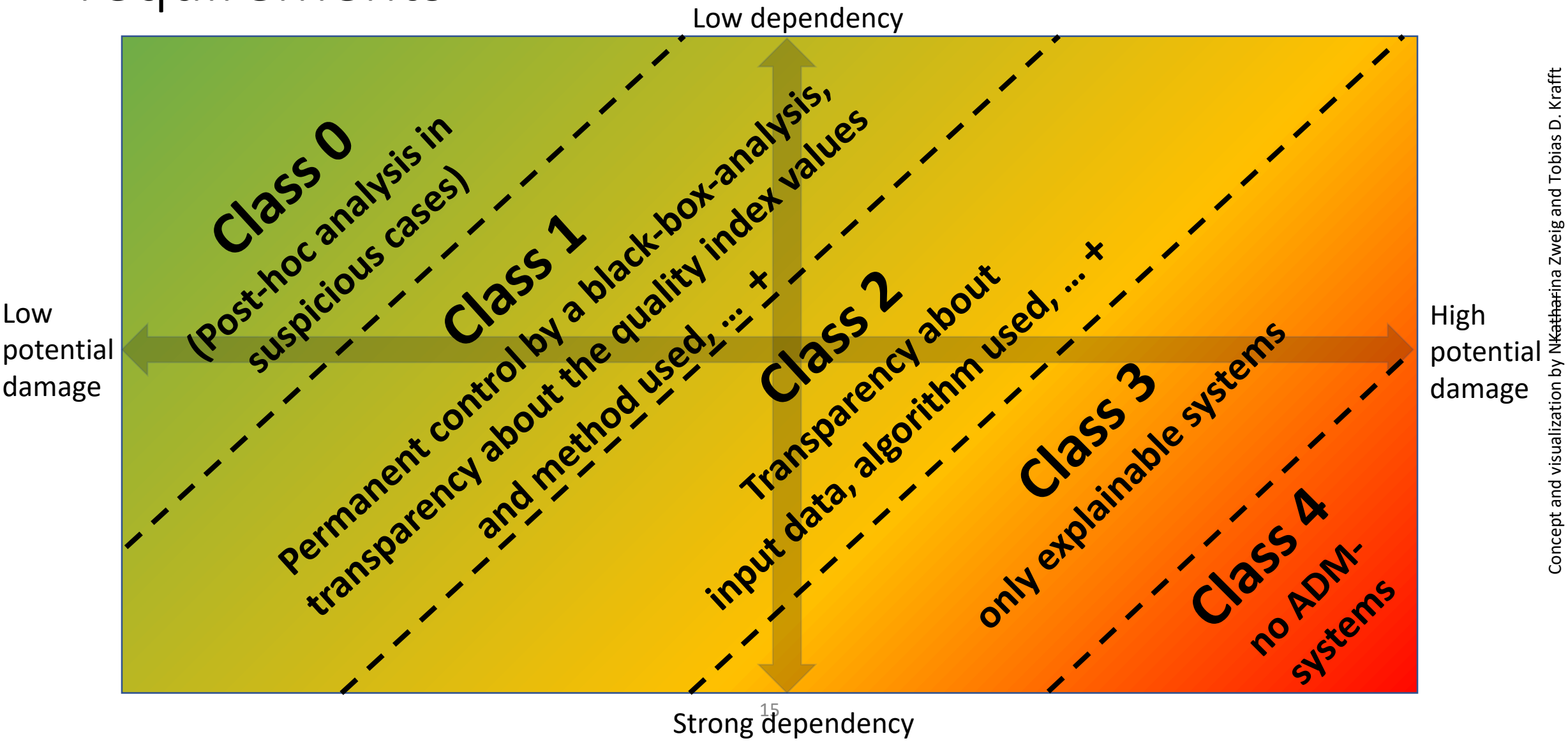
How much control is necessary?

Needs to be **differentiated** by the

a) **total potential individual and societal damage** of using the ADM system in a **given social context** and

b) the **dependency** of the scored/classified subject on the decision.

Five classes of transparency and accountability requirements



Concept and visualization by Katharina Zweig and Tobias D. Krafft

Sources

Maryam Haeri: Paper unpublished yet

Christopher Koska: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Bibliothek/Doi_Publikationen/Ethik_fuer_Algorithmiker._Was_wir_von_erfolgreichen_Professionsethiken_lernen_koennen._Final..pdf

Katharina Zweig and Tobias Krafft: Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse | Ein Regulierungsvorschlag | 22. Januar 2019 https://www.vzbv.de/sites/default/files/downloads/2019/05/02/19-01-22_zweig_krafft_transparenz_adm-neu.pdf (p.13 and p.29)

Cliparts: https://webstockreview.net/