

5th Digital Trust Day

In AI we trust. Do we?

Aspects of Digital Trust in the Realm of AI
and its Impact

Janine Jäger & Jona Karg

In AI we trust. Do we?



Janine Jäger



Jona Karg

In AI we trust. Why is trust relevant?

- AI is increasingly employed for tasks in business, public and private contexts.
- AI decision making processes and outputs are rarely understood by non-experts.
- AI has the potential to help people, but also to hurt them.
- There is still little research on the questions:
 - What makes AI trustworthy?
 - Which factors make people trust or distrust AI?



Picture Credits: ChatGPT4



Picture Credits: ChatGPT4

In AI we trust. Do technology experts trust it?

The New York Times

Harvesting Data for A.I. | A.I. Data Race, Explained | What Is 'Synthetic Data'? | Key Figures in

«Powerful AI systems should be developed only once we are confident that their effects will be positive and their risks will be manageable.»

Elon Musk and Others Call for Pause on A.I., Citing 'Profound Risks to Society'

More than 1,000 tech leaders, researchers and others signed an open letter urging a moratorium on the development of the most powerful artificial intelligence systems.

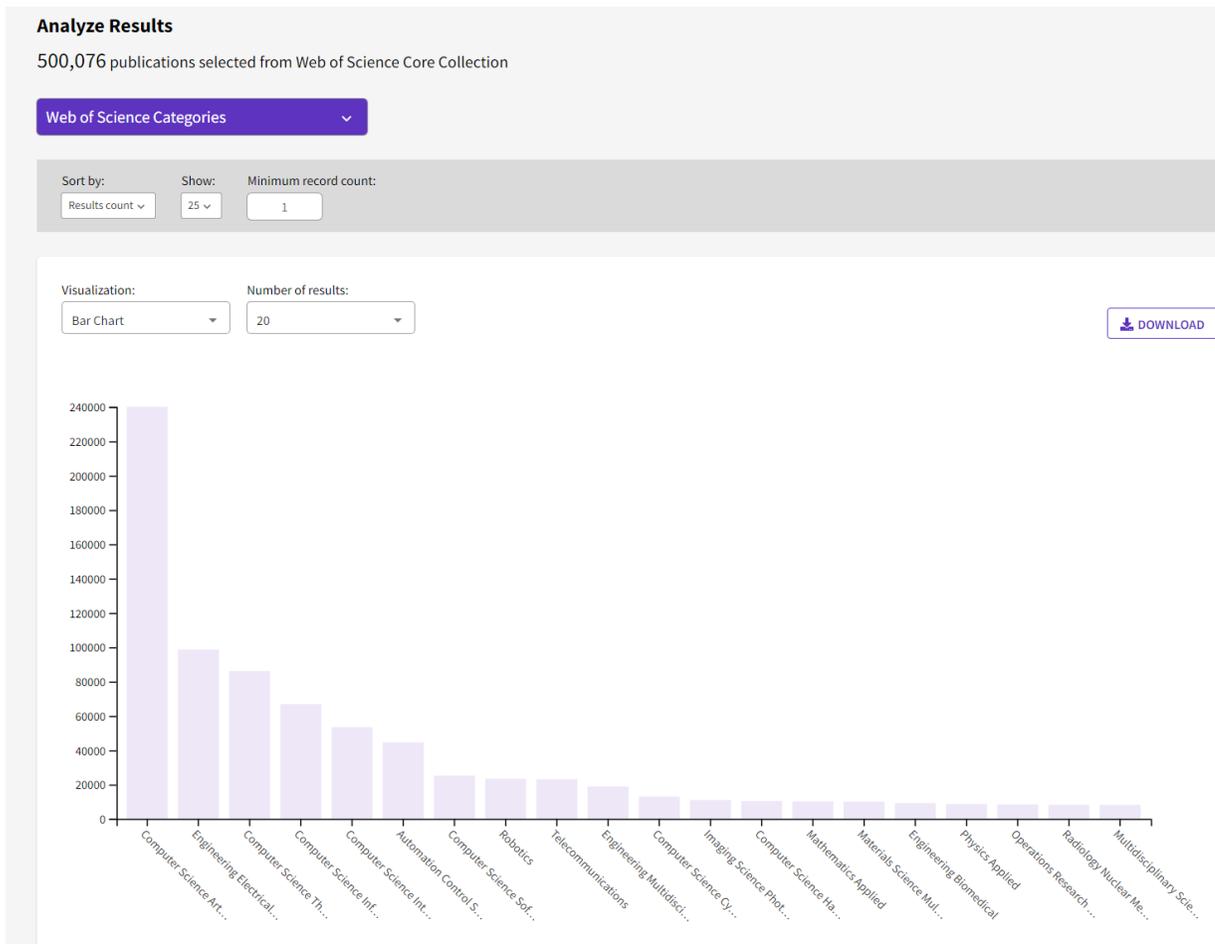
Source: The New York Times (2023) <https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html>

In AI we trust. Does regulation help?

EU Artificial Intelligence Act – Art. 1

The purpose of this Regulation is to improve the functioning of the internal market and **promote the uptake of human-centric and trustworthy artificial intelligence (AI)**, while ensuring a high level of protection of health, safety, fundamental rights enshrined in the Charter, including democracy, the rule of law and environmental protection, against the harmful effects of AI systems in the Union and supporting innovation.

In AI we trust. Does research help?



Most research on AI (naturally) comes from technical disciplines.

Some of that research is interdisciplinary and includes ethical and social aspects.

Research stemming from social sciences is increasing, but still rare.

Research on AI from the fields of (**All times**):

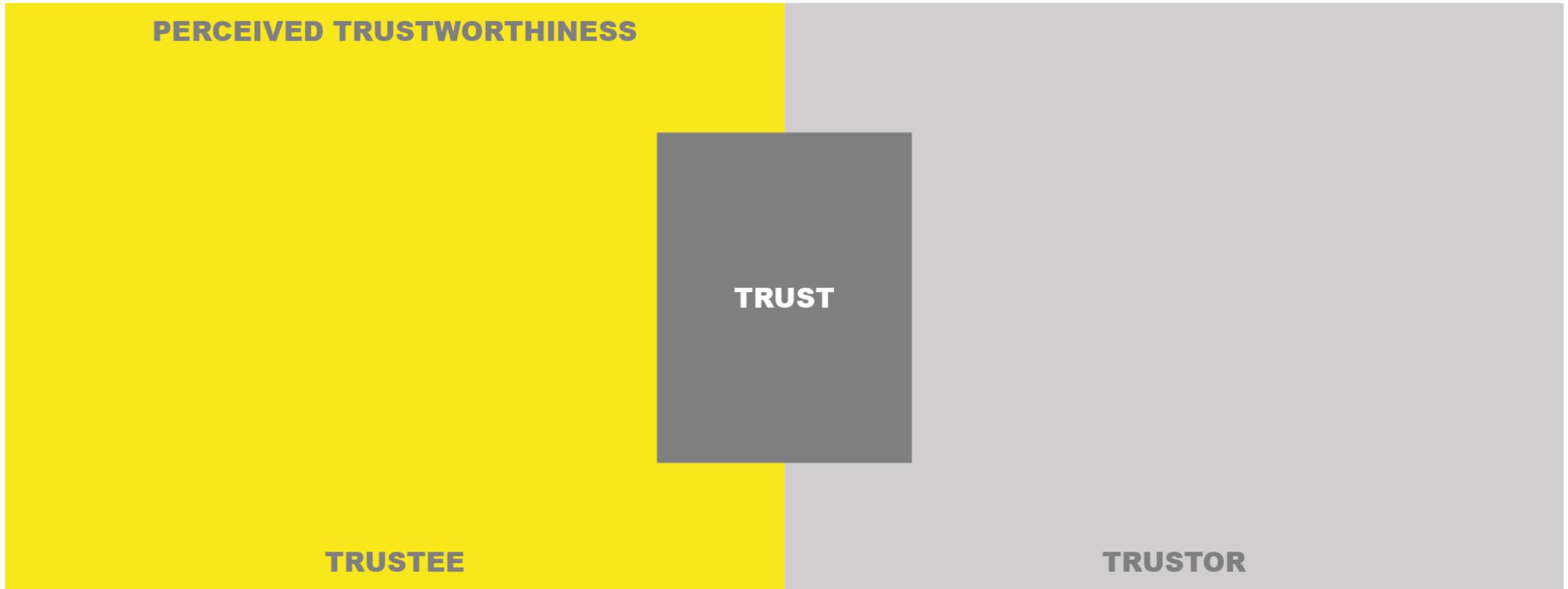
- Ethics = 0.17%
- Social Sciences = 0.33%

A Definition of Trust

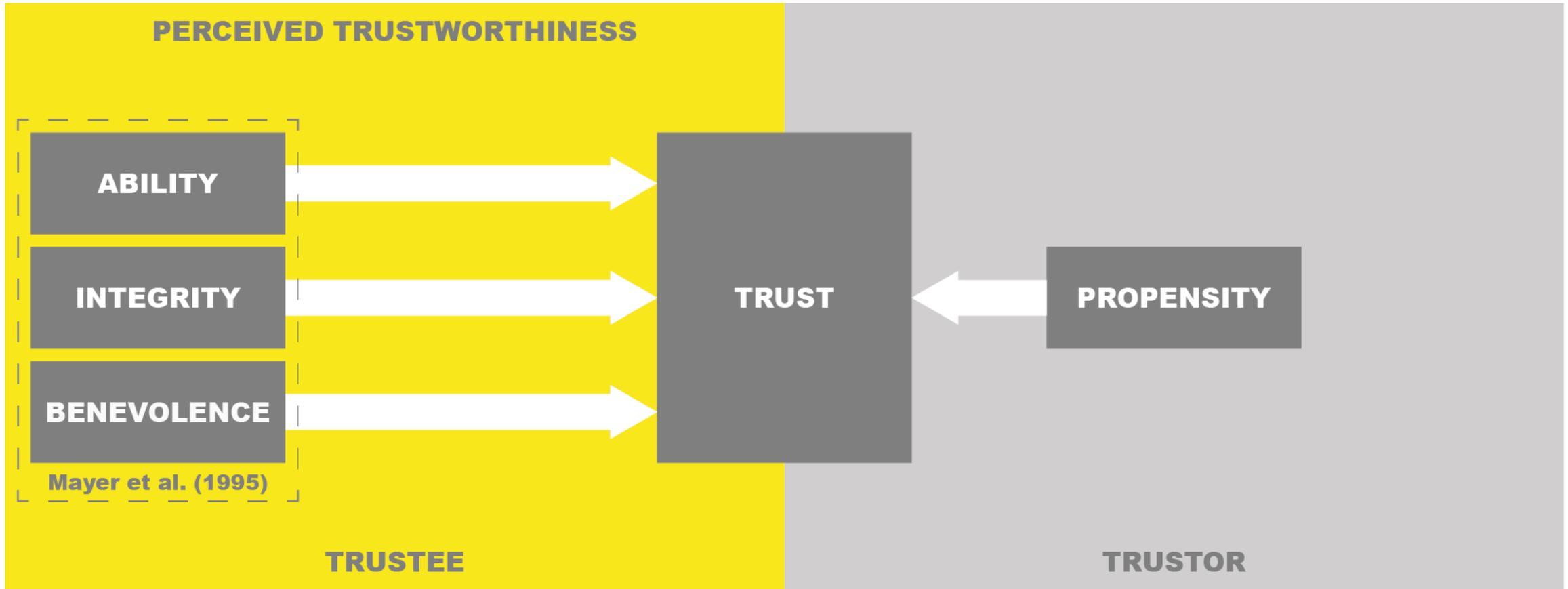
«Trust [...] is the **willingness** of a party **to be vulnerable** to the actions of another party based on the **expectation** that the other will perform a particular **action** important to the trustor, **irrespective of the ability to monitor or control** that other party.»

Mayer et al. (1995, P. 712)

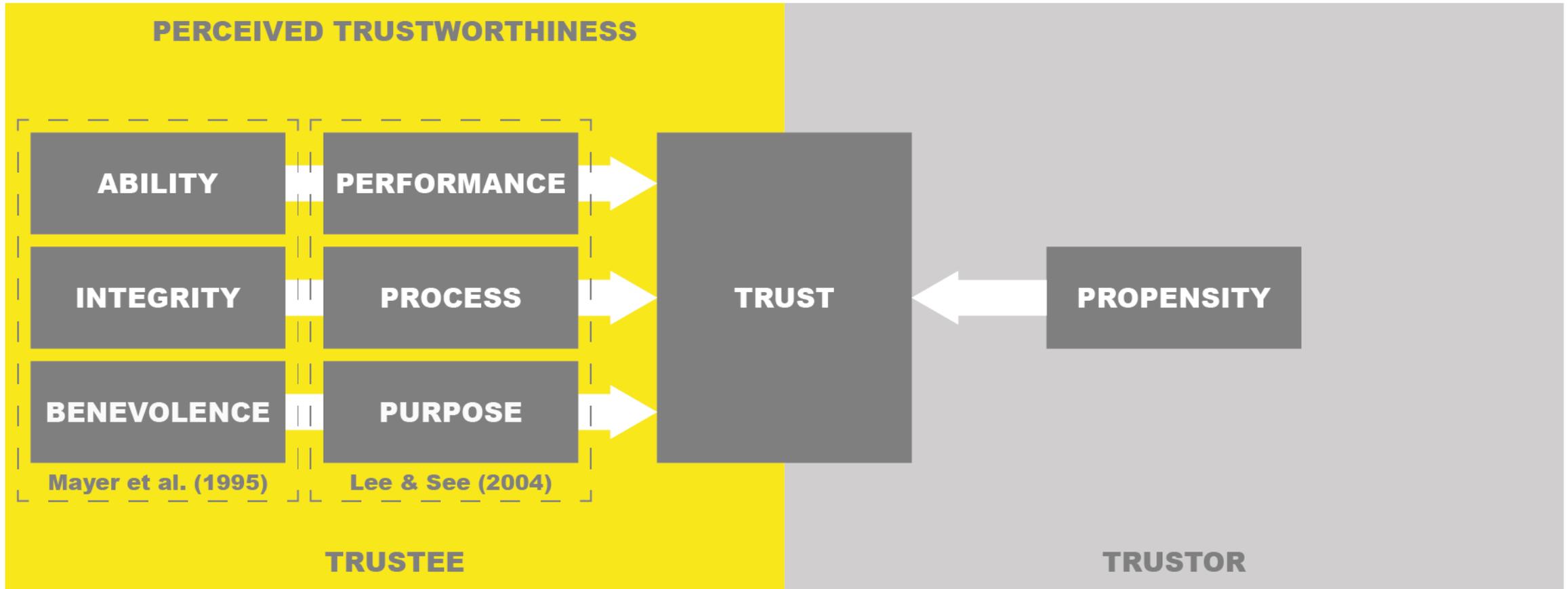
What is Trust?



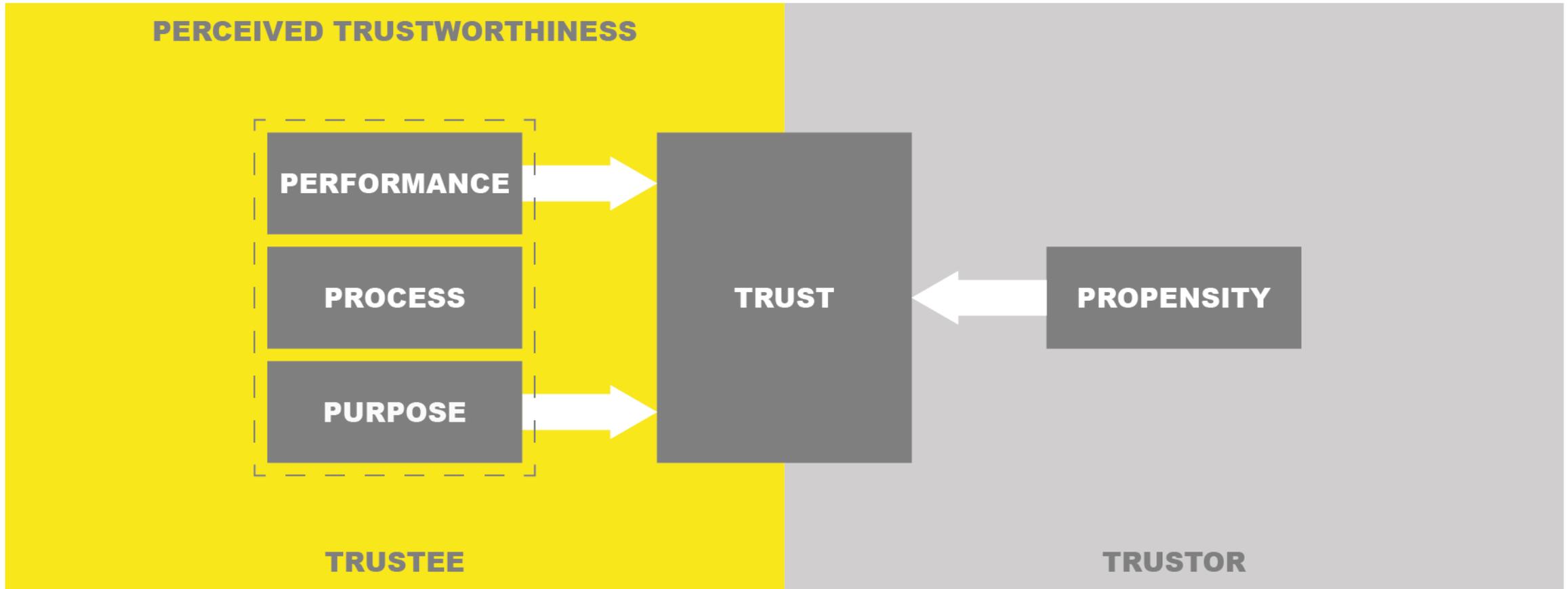
Trust in Organizations



Trust in Automation



Trust in ChatGPT



Opposite of trust – Distrust!?

The screenshot shows a Google search for "opposite of trust". The search results page includes a search bar with the query "opposite of trust", navigation tabs for "Alle", "Bilder", "Videos", "Shopping", "News", and "Mehr", and a "Suchfilter" button. Below the search bar, it indicates "Ungefähr 712'000'000 Ergebnisse (0,36 Sekunden)". The main content area shows the word "trust" with tabs for "Übersicht", "Anwendungsbeispiele", "Aussprache", and "Ähnliche und entgegengesetzte Wörter". Under the "Ähnliche und entgegengesetzte Wörter" section, it lists "Gegensätzlich" (opposite) with the noun "distrust", "mistrust", and "scepticism", and the verb "distrust", "mistrust", and "doubt". Below this, there is a snippet from Merriam-Webster titled "TRUST Synonyms: 166 Similar and Opposite Words" with the text: "noun - misgiving - inconstancy - dubiety - disillusion - disillusionment - disenchantment - nonconfidence - distrustfulness - dubiousness - mistrustfulness." There are also sections for "Weitere Fragen" (Further questions) and "Synonyms of mistrust - Merriam-Webster Thesaurus".

In AI we trust. Do we?

Please pick **one board** with one AI tool and discuss the **two questions** with your peers:

- 1) What makes the AI tool trustworthy for you?
- 2) What else is needed to develop adequate trust for the AI tool?