Your AI is not as useful as you think

Deconstructing AI saviorism

Hinda Haned, Ph.D.
University of Amsterdam | Owls & Arrows
April 20th 2023

Saviorism

A worldview according to which some people are saviors and others need to be **saved** by them

wiktionary

AI Saviorism

The view that AI-driven systems can solve societal challenges, while ignoring the technical, societal, economic and domain-specific constraints that make them unusable in practice

Illustration: COVID19 detection

2,212 studies, of which 415 were included after initial screening and, after quality screening, 62 studies were included in this systematic review

*Our review finds that **none of the models identified are of potential clinical** use due to methodological flaws and/or underlying biases.*

Roberts, Michael, et al. "Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans." *Nature Machine Intelligence* 3.3 (2021): 199-217.

Failure taxonomy

Impossible tasks	fraud detection automation
Engineering failures	grade prediction
Post-deployment failures	recruitment tools
Communication failure	organizational failures

Raji, Inioluwa Deborah, et al. "The Fallacy of AI Functionality." *2022 ACM Conference on Fairness, Accountability, and Transparency*. 2022.

Famous failures

Famous failures

UK GSCE exams 2020

Responsible AI

The EU AI ACT

EUROPEAN COMMISSION

Brussels, 21.4.2021

COM(2021) 206 final

2021/0106(COD)

Proposal for a

REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS

{SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final}

Ethics Guidelines

Responsible AI

- Risk-based approach
- Focus on high-risk applications
- Avoid or mitigate harm or (potential) bias

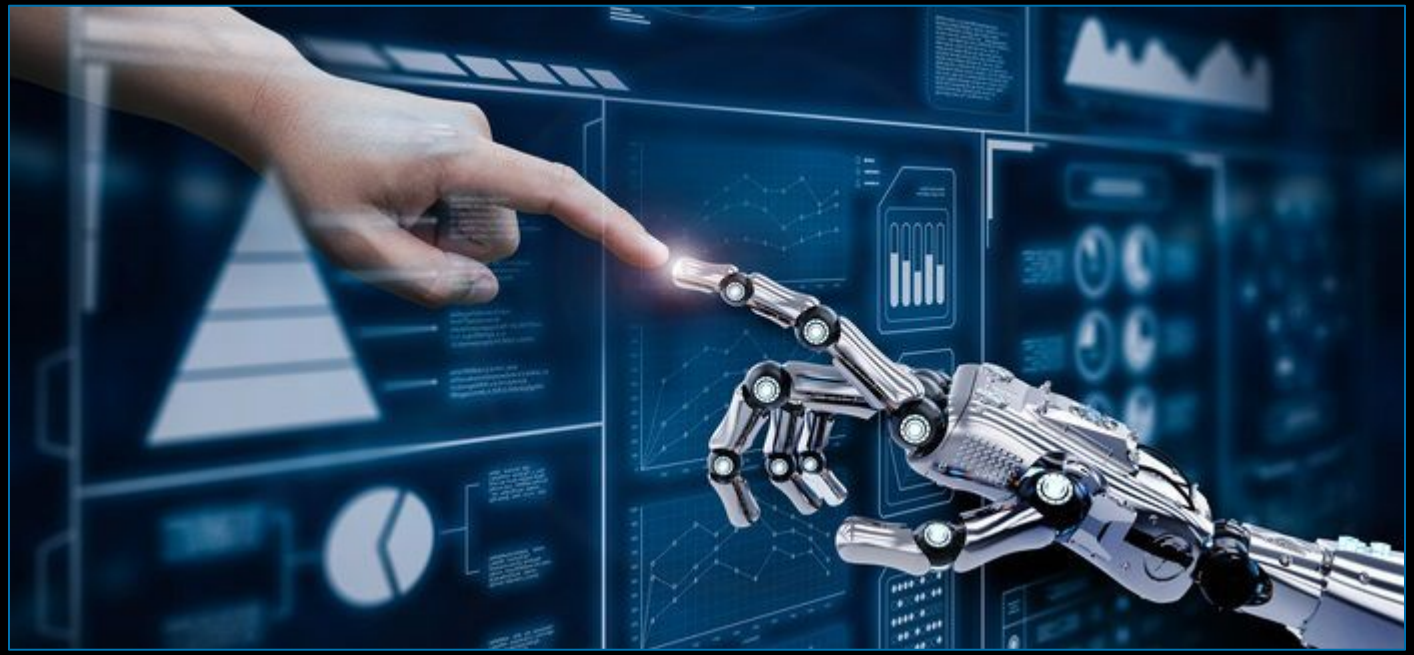
Functionality as an afterthought

How did we get here

- Highly competitive field and industries
- Misalignment between research and business objectives
- Lack of (long- term) strategy around AI and data analytics
- Challenges around data quality & availability

D. Schwarz et al. A Framework for the Systematic Evaluation of Data and Analytics Use Cases at an Early Stage. Hawaii International Conference on System Sciences 2021

Overhyped AI capabilities



From https://ec.europa.eu/

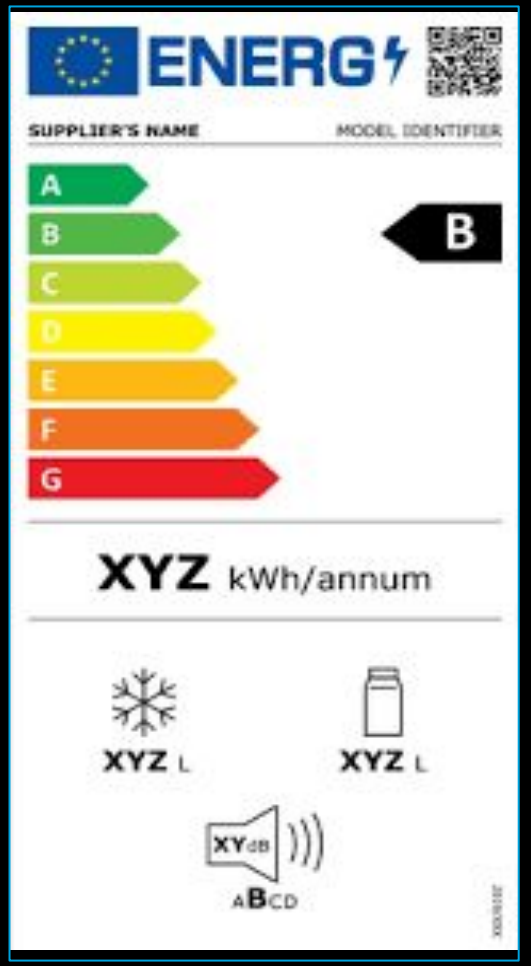
What is functional AI?

- Domain-dependent
  - HR: automation, efficiency
  - Healthcare: accuracy
- Stakeholder-dependent
  - AI democratization vs. digitalisation
- Secondary use
  - Data extraction

What if...

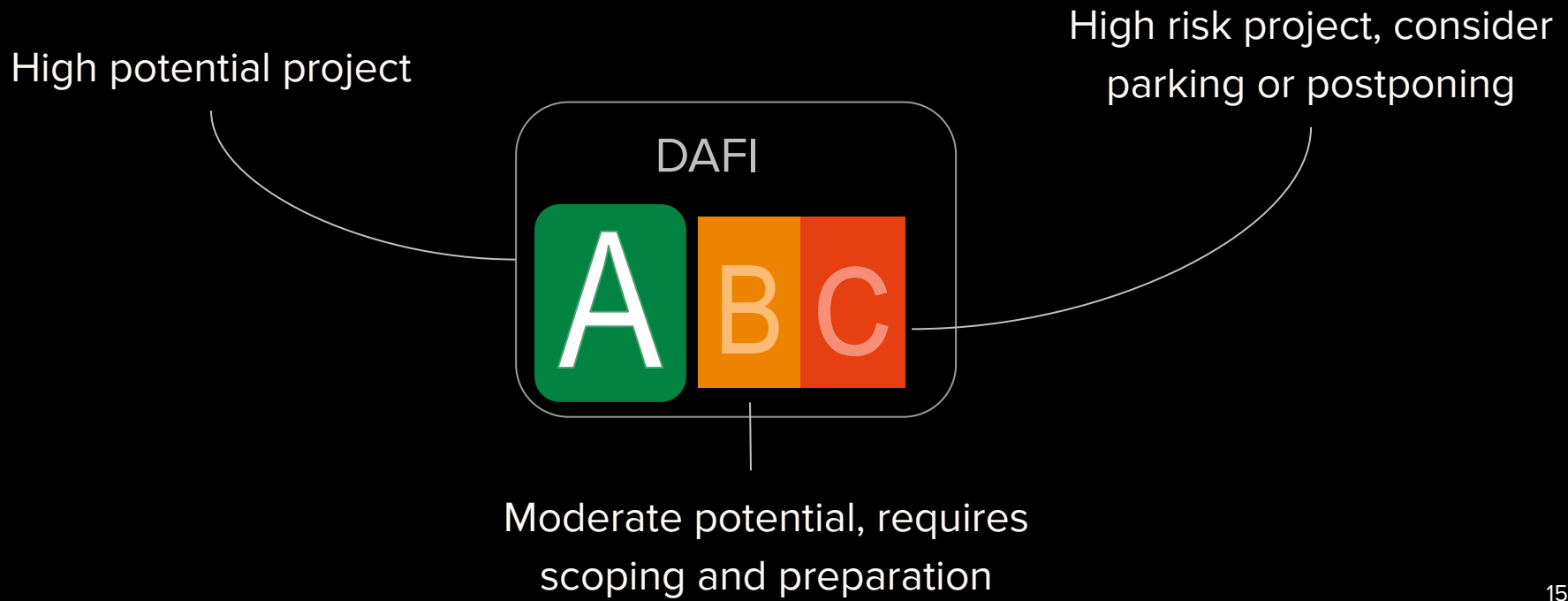
What if you could have a data project **label** to help you decide whether or not to pursue a project?

Visual labels for decision-making



Images from Wikipedia

DAFI: Data Analytics Functionality Index



Constraints for a functionality label

- Must be **generalisable** across projects types
- Must be **easy** to understand hort time to fill in
- Should contain **metrics** that capture sensible goals

F.A.C.T score for responsible AI

510

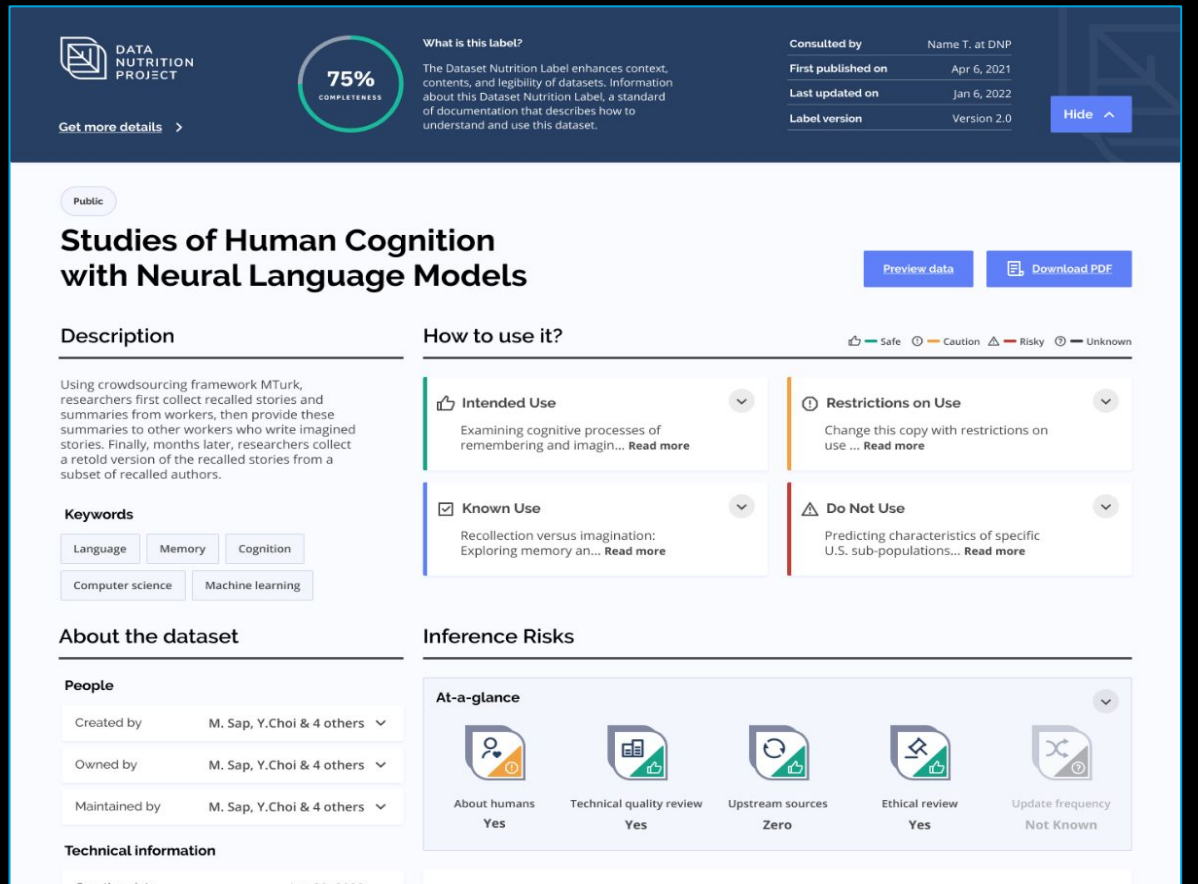
An initiative of the Netherlands Red Cross

https://rodekruis.github.io/responsible_ai

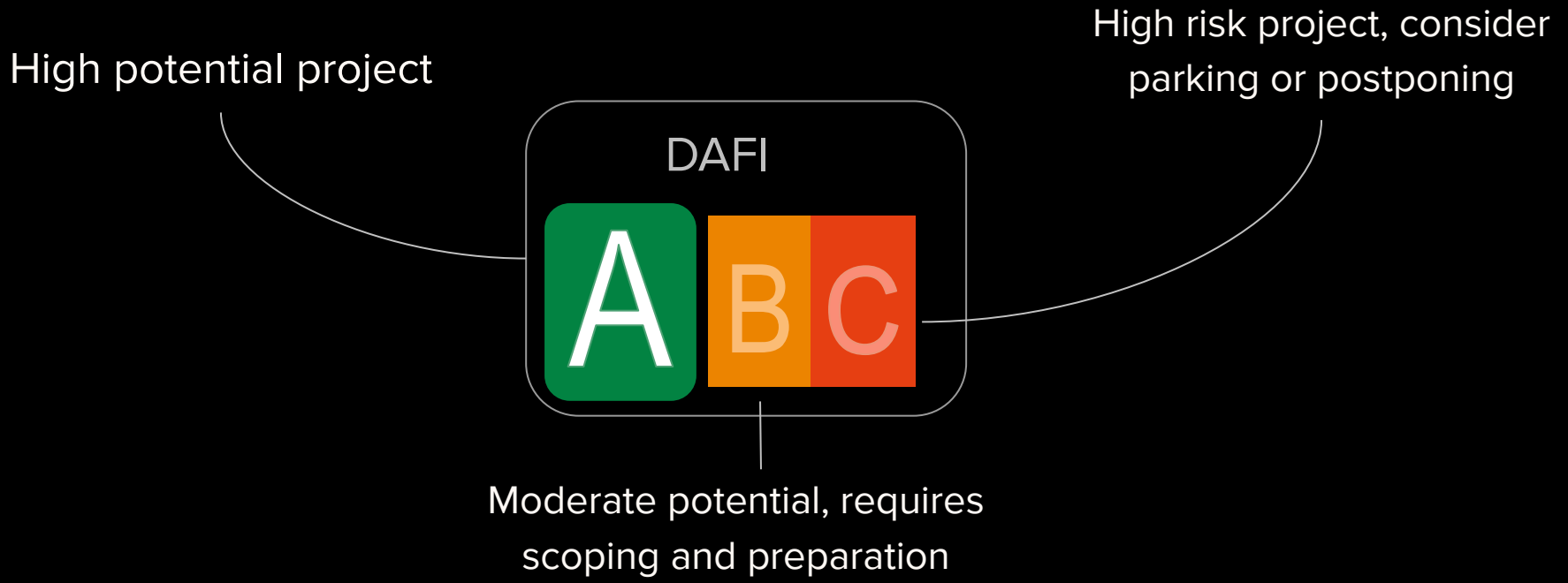
Data nutrition label

S. Holland et al. The dataset nutrition label. Data Protection and Privacy, 2020.

https://datanutrition.org/



DAFI: Data Analytics Functionality Index



Criteria for the DA functionality index

UTILITY

Measures how well the business question is defined.

FEASIBILITY

Measures strategic and operational feasibility within the organization/teams.

VERIFIABILITY

Measures if the system can be validated and therefore trusted.

MEASURABILITY

Measures the actionability of the project’s outcomes.

DAFI questionnaire

Utility

- Is the business question well-defined?
- Is the project mapped to a clear KPI?
- Are the expected outcomes aligned with the data s
- Does the tool/model impact one or several quantita KPIs? (e.g. sales forecast accuracy)
- Are the project's outcomes expected to improve a business process?

Measurability

- Are the outcomes of the tool/model likely to be act
- Does the tool/model impact one or several quantit KPIs? (e.g. sales forecast accuracy)
- Will it be possible to measure the impact of the tool either qualitatively or quantitatively?

Verifiability

- Does the development team have a clear plan & milestones in place?
- Is there a clear plan to document the model/tool (data sources, logic)?
- Can the model or tool be checked or verified independently within the organisation?
- Can the outcomes be verified against a ground truth?
- Does the development plan include a validation phase?
- Will end-users be involved in the development and validation phase?

Feasibility (strategic & operational)

- Is the project supported and facilitated by clearly identified stakeholders?
- Do the stakeholders or domain experts involved have the
- compliance-related constrained kled?
- ces (people, infrastructure) assigned
- s previously been conducted in the
- o process the data and build any odels) available?
- the project readily available and
- Is the tool/model likely to be developed in time to be useful?

Testing with real-world use cases

Use cases	
Predicting next week's sales	A
Speech sentiment analysis	C
Customer lead generation	C
Personalised phishing campaigns	B

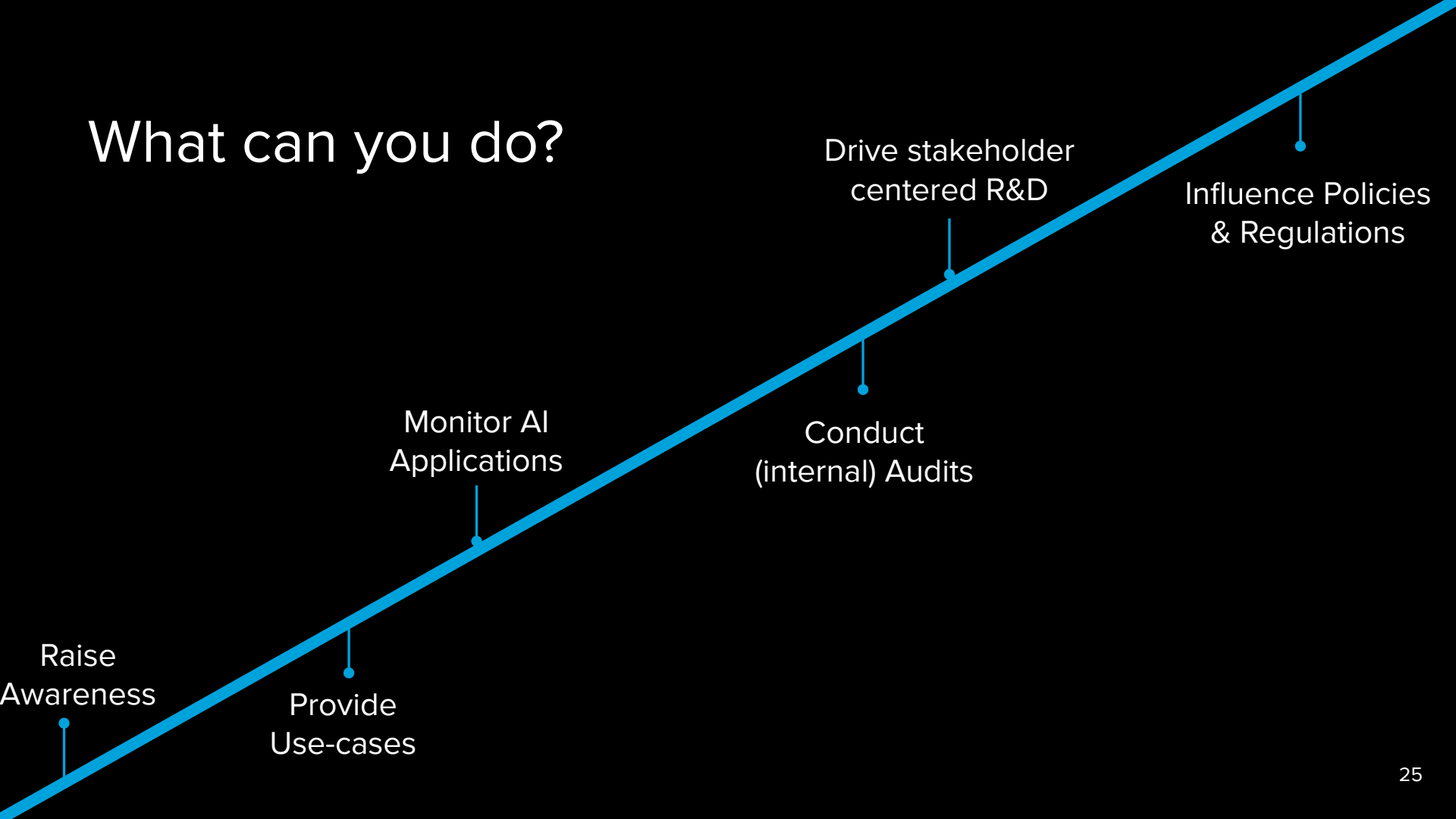
Thank you!

h.haned@uva.nl
hindantation.github.io

Need for a paradigm shift

- Challenge functionality assumption
- Invest in basic data literacy
- Use tools to demystify AI applications
- Bridge the gap with real-world and organisational constraints

What can you do?



The way forward

- **Mobilise** the AI community to solve relevant problems while embracing the challenges of real world datasets and collaboration with domain experts
- **Encourage** and create meaningful incentives for a stakeholder-centric approach to create useful applications and systems
- **Embrace** and normalise direct communication between stakeholders and AI developers/researchers