

# Towards responsible GenAI in education

Dragan Gašević

Oct 19th, 2024  
Global Harwell Online Forum

# Artificial intelligence

Artificial intelligence  
will not go away



Source: <https://bit.ly/47NbA7D>

# Challenge

How do we responsibly  
use GenAI in education and  
prepare for the age of GenAI?

# Responsible

*Having an obligation to do something, or having control over or care for someone, as part of one's job or role.*

Yuval Noah  
Harari



# Nexus

A Brief History of Information Networks  
from the Stone Age to AI

## Key takeaway 1

Responsible GenAI in education  
includes but goes beyond technology

## Key takeaway 2

Responsible GenAI requires to  
rethink skills we develop and assess

# **TECHNOLOGY DESIGN AND DEVELOPMENT**





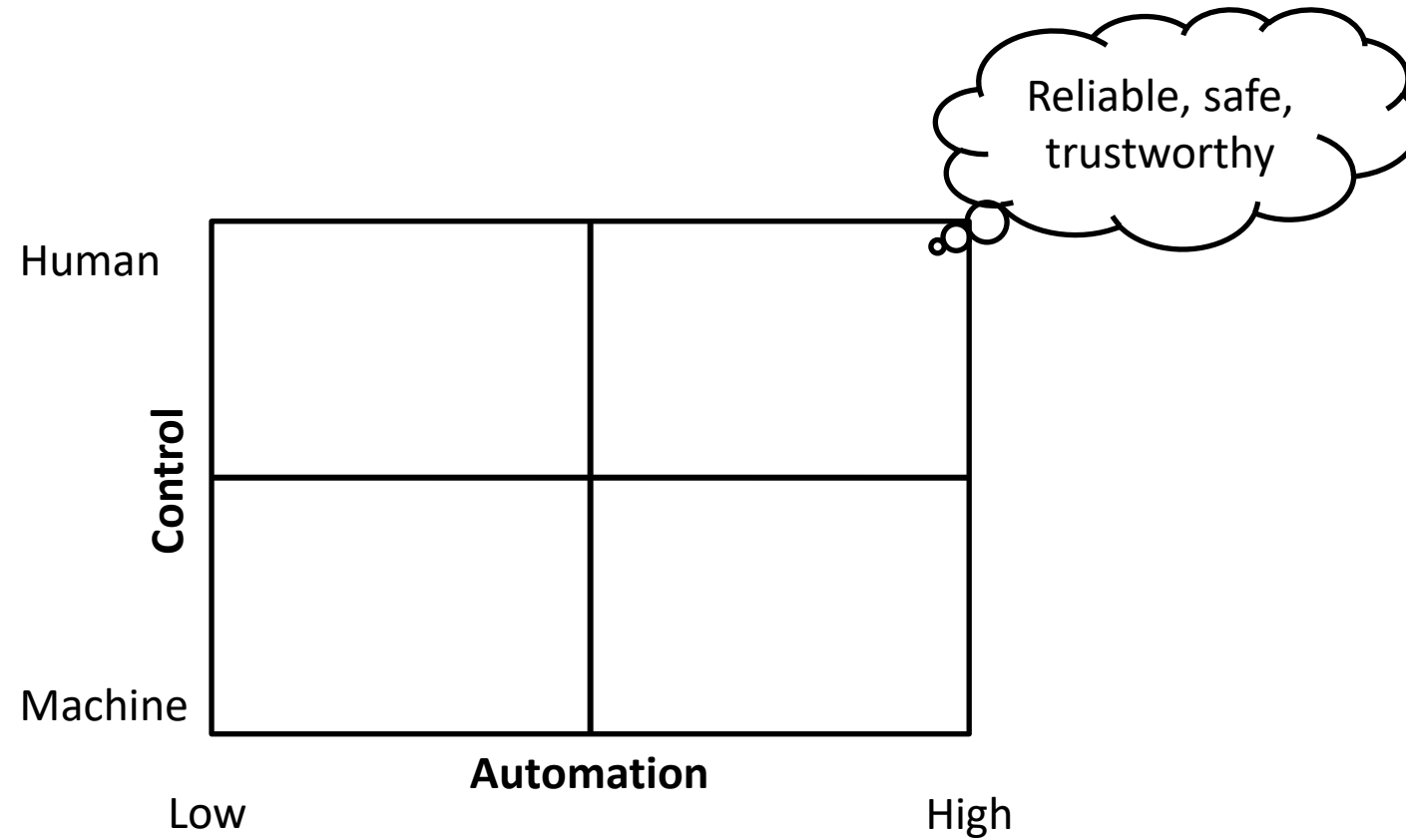
# Consent

How to design GenAI technology in education responsibly?

Giving educators and students control

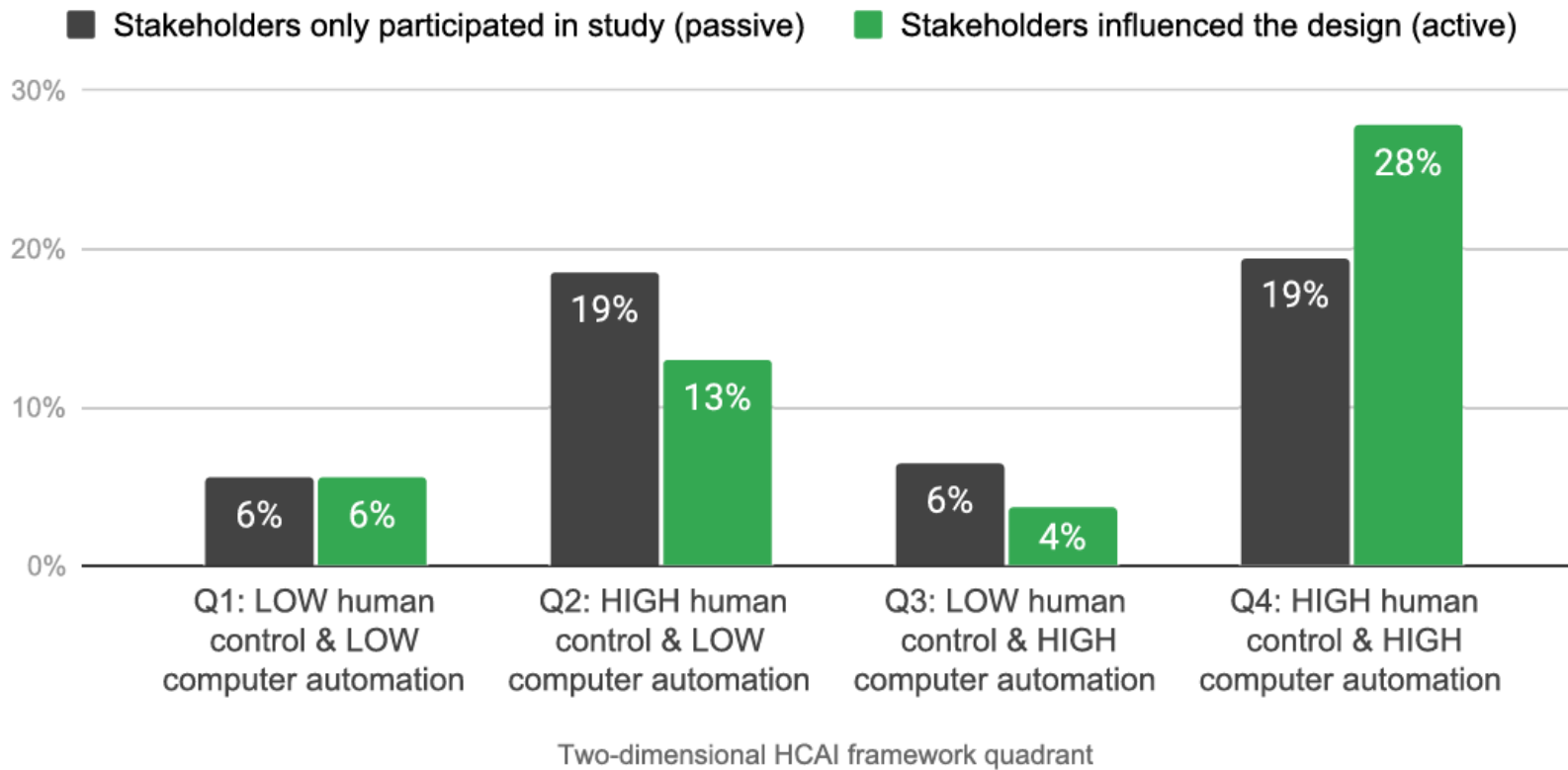


# Direction

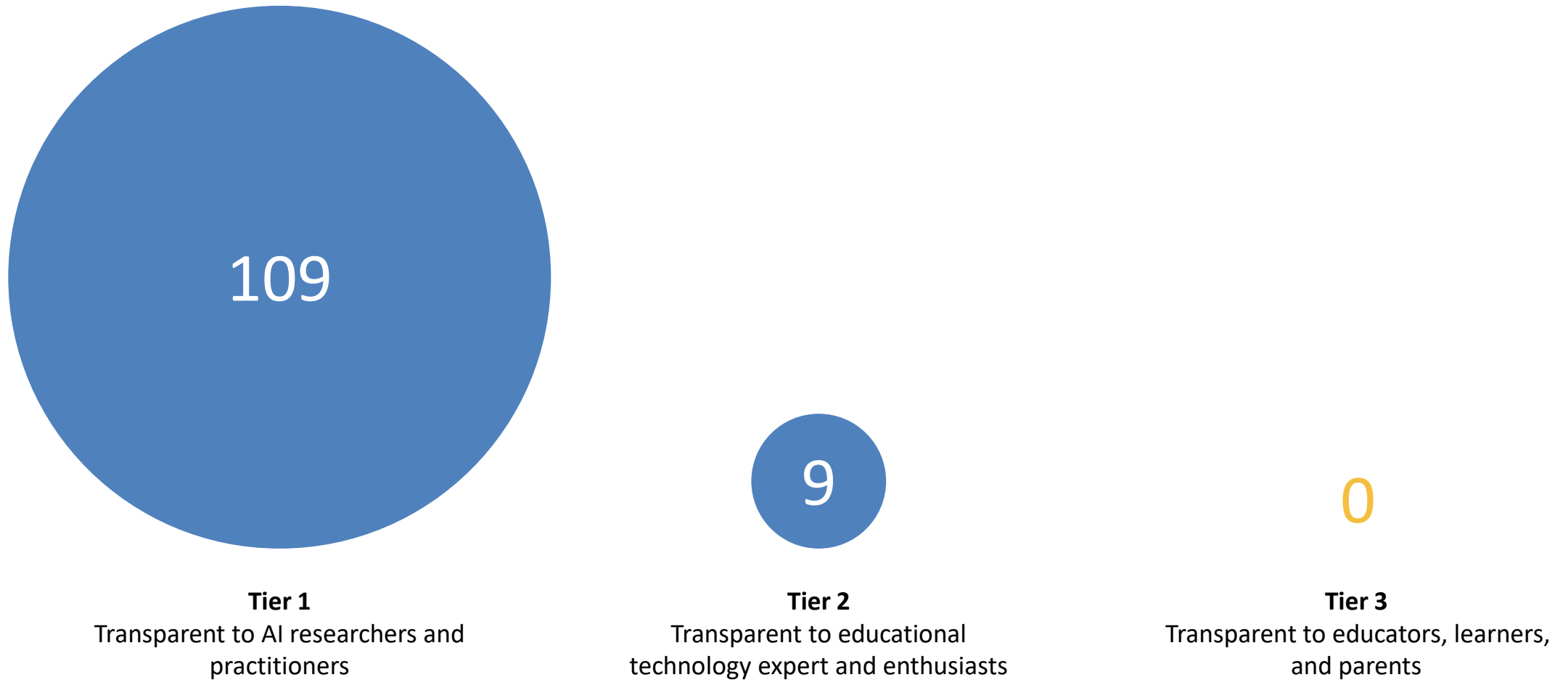


# Human-centred AI in education

## Stakeholder Involvement in HCAI Framework



# LLMs-based system transparency



# Consent

No good mechanism to track where  
and how data are used

Giving educators and students control

# Consent

**THE CONVERSATION**  
Academic rigour, journalistic flair


Search analysis, research, academics...


Arts + Culture Books + Ideas Business + Economy Education Environment + Energy Health Politics + Society **Science + Tech**


## ChatGPT: lessons learned from Italy's temporary ban of the AI chatbot


Published: April 20, 2023 9:54pm AEST


The Italian data protection authority used an emergency procedure to impose the temporary ban. Shutterstock / Diego Thomazini

 Email

 Twitter 31

 Facebook 97


 LinkedIn


 Print

In March 2023, Italy became the first western country to block the advanced chatbot known as ChatGPT.

The Italian data protection authority, Garante, cited concerns over the protection of personal data when making this decision. It has given OpenAI, the California-based company that created ChatGPT, until the end of April to comply with its demands.

**Authors**

**Oreste Pollicino**  
Professor of Constitutional Law, Bocconi University

**Giovanni De Gregorio**  
PLMJ chair in law and technology at Católica Global School of Law and Católica Lisbon School of Law and academic fellow, Bocconi University

# Consent

Hard to *unlearn* something once private data are used for training LLMs

## Responsibility challenge

How do we prepare learners to exercise their privacy agency with GenAI?

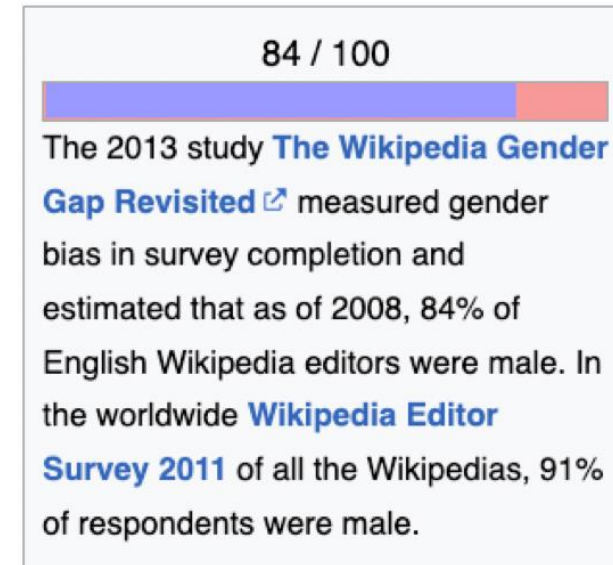


# Fairness in GenAI

## Nationality



## Gender



# Fariness in GenAI

**Midjourney v4**

Prompt:  
patriotic, dog, superhero.



## Fairness in downstream tasks

# Need to mitigate bias in modern LLMs on relevant downstream tasks

Bias in prompting is particularly concerning

# Fairness

Who takes responsibility for  
unintended consequences of bias GenAI  
may have on education?

# **HUMAN KNOWLEDGE AND SKILLS**



# Promise of GenAI

GenAI can enhance  
human performance

# Need for AI literacy

The New York Times

## *Here's What Happens When Your Lawyer Uses ChatGPT*

A lawyer representing a man who sued an airline relied on artificial intelligence to help prepare a court filing. It did not go well.

 Give this article    1.1K



By **Benjamin Weiser**

May 27, 2023

The lawsuit began like so many others: A man named Roberto Mata sued the airline Avianca, saying he was injured when a metal serving cart struck his knee during a flight to Kennedy International Airport in New York.

Weiser, B. (2023). Here's What Happens When Your Lawyer Uses ChatGP. <https://www.nytimes.com/2023/05/27/nyregion/avianca-airline-lawsuit-chatgpt.html>

## **This US lawyer used ChatGPT to research a legal brief with embarrassing results. We could all learn from his error**

ABC RN / By Damien Carrick and Sophie Kesteven for the Law Report, with additional reporting from Reuters.

Posted Sat 24 Jun 2023 at 11:15am



Carrick, D., & Kesteven, S. (2023, June 24). 'Use with caution': How ChatGPT landed this US lawyer and his firm in hot water. *ABC News*. <https://www.abc.net.au/news/2023-06-24/us-lawyer-uses-chatgpt-to-research-case-with-embarrassing-result/102490068>



# Impact of the use of GenAI on learning

## AI Meets the Classroom: When Does ChatGPT Harm Learning?

Matthias Lehmann,<sup>1</sup> Philipp B. Cornelius,<sup>2</sup> Fabian J. Sting<sup>1,2</sup>

<sup>1</sup>University of Cologne, <sup>2</sup>Rotterdam School of Management, Erasmus University

<https://arxiv.org/abs/2409.09047>

## Generative AI Can Harm Learning

Hamsa Bastani,<sup>1\*</sup> Osbert Bastani,<sup>2\*</sup> Alp Sungu,<sup>1\*†</sup>  
Haosen Ge,<sup>3</sup> Özge Kabakcı,<sup>4</sup> Rei Mariman

<sup>1</sup>Operations, Information and Decisions, University of Pennsylvania

<sup>2</sup>Computer and Information Science, University of Pennsylvania

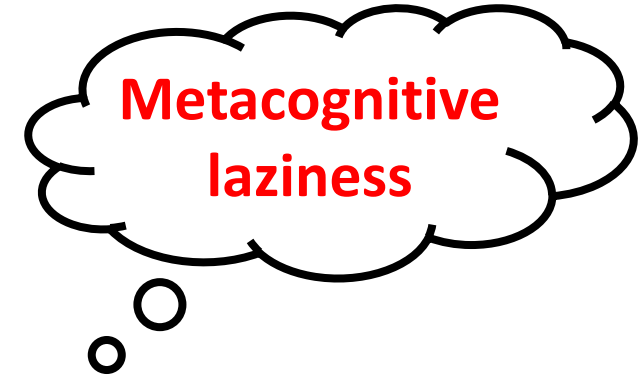
<sup>3</sup>Wharton AI & Analytics, University of Pennsylvania

<sup>4</sup>Budapest British International School

<https://dx.doi.org/10.2139/ssrn.4895486>



# Metacognition in information problem solving



Limited evidence of metacognition when  
using generative AI for help seeking

# Inadvertent deception of ChatGPT

## Inaccurate and verbose, but users prefer ChatGPT responses

Processing fluency – illusion of truth effect

Unclear long term consequences

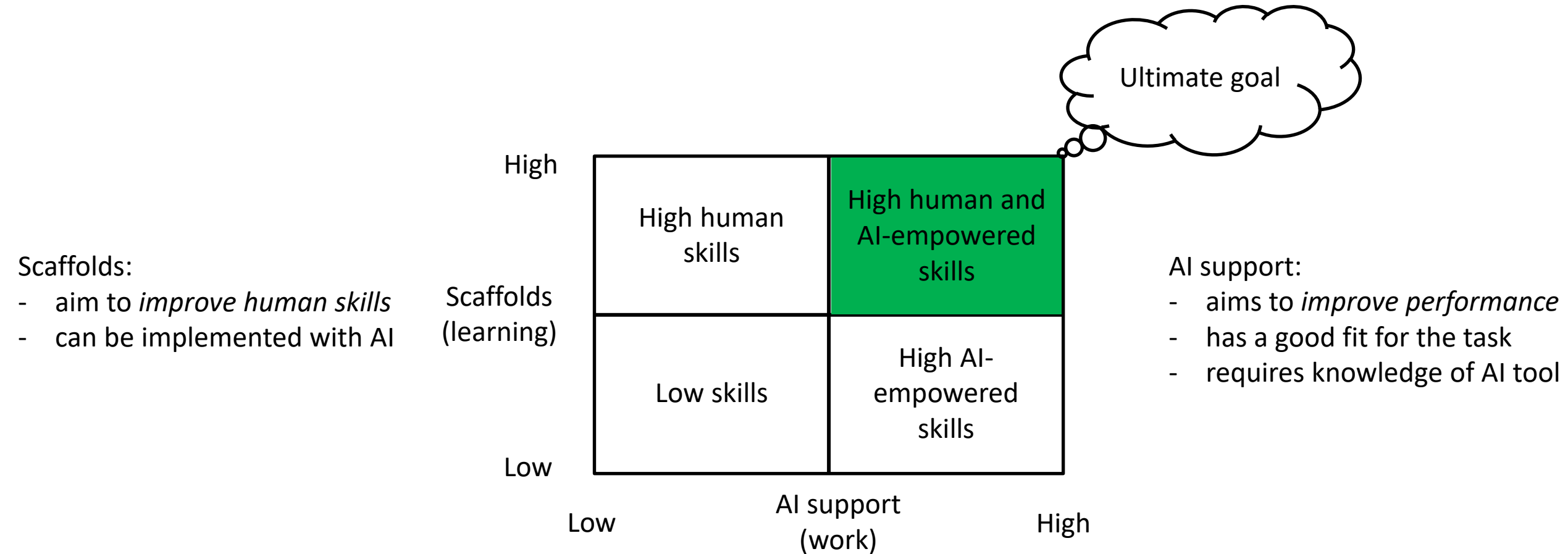
Writing is an important learning tool

What happens when much of that tool is offloaded to  
generative AI?

# Human-AI relationship

What skills and knowledge do we  
need to develop and assess?

# Towards high human and AI-empowered skills



Inspired by, but different conceptualization from

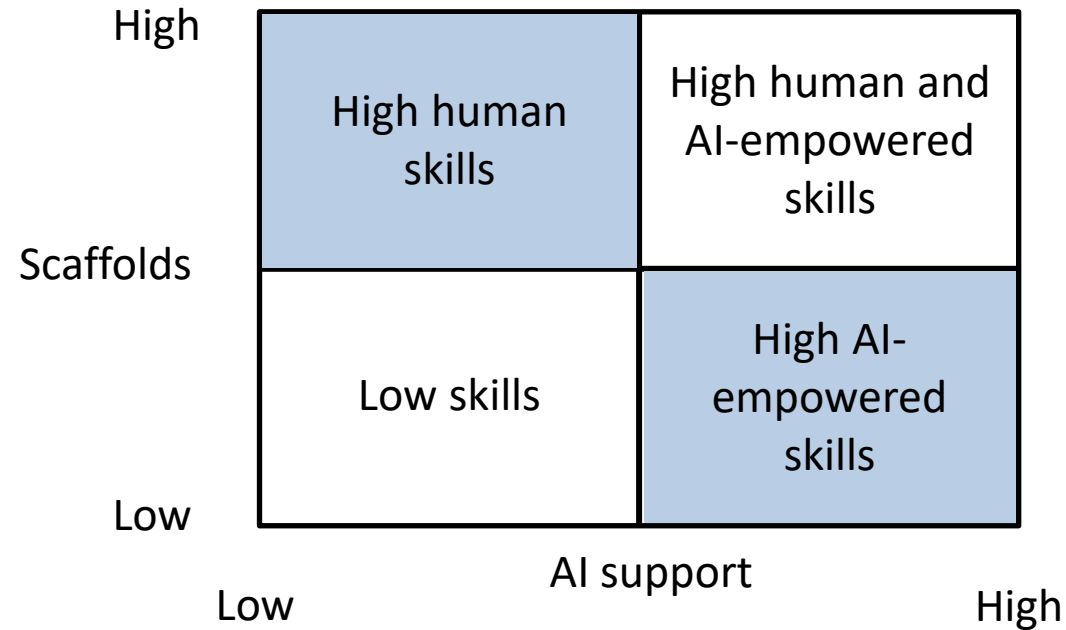
Cukurova, M. (2024). The interplay of learning, analytics, and artificial intelligence in education. *arXiv preprint* arXiv:2403.16081.

Shneiderman, B. (2022). *Human-centered artificial intelligence*. Oxford University Press

## Key challenge

Do we confuse AI-empowered skills for  
high human skills?

# Towards high human and AI-empowered skills



Limited evidence of ***combined*** human and AI-empowered skills

## Call for action

Reimagining our practice of  
learning, teaching, and assessment



# **BEYOND PRODUCTIVITY**



# Aspiration

AI will increase productivity to focus  
on things that matter the most

# Generative AI at Work

**Erik Brynjolfsson, Danielle Li & Lindsey R. Raymond**

WORKING PAPER 31161

DOI 10.3386/w31161

ISSUE DATE April 2023

# Implications

Productivity  $\neq$  Workload

## TECHNOLOGY

ChatGPT Is About to Dump More  
Work on Everyone

# Productivity != Workload

How do we deal with  
increased complexity of working and learning with AI?



“It helped, but sometimes its errors caused more work than doing the task manually.”

# Challenge

## Productivity $\neq$ Workload

Automatic feedback is not associated with decreased teaching load

# Challenge

Workload also includes dealing with  
changed landscape of education

Reinventing assessment, growing AI literacy, and  
dealing with bureaucratic instincts of education systems



# Challenge

How do we deal with issues of  
workload implications responsibly?

PROMISE – CONCERN – NEED – **FINAL REMARKS**

Responsible GenAI goes beyond  
technical notions of fairness and bias

Responsible GenAI concerns  
what skills we develop and assess

# Towards responsible GenAI in education

Dragan Gašević

Oct 19th, 2024  
Global Harwell Online Forum