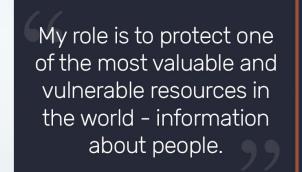
ANNELIES MOENS

Privacy expert



Science &

Fechnology



SUPERSTAR

OF ST

It Takes a Spark STEM conference, The Southport School, Gold Coast 16 June 2023 school







The careers you will have in 10 - 15 years' from now don't exist today.





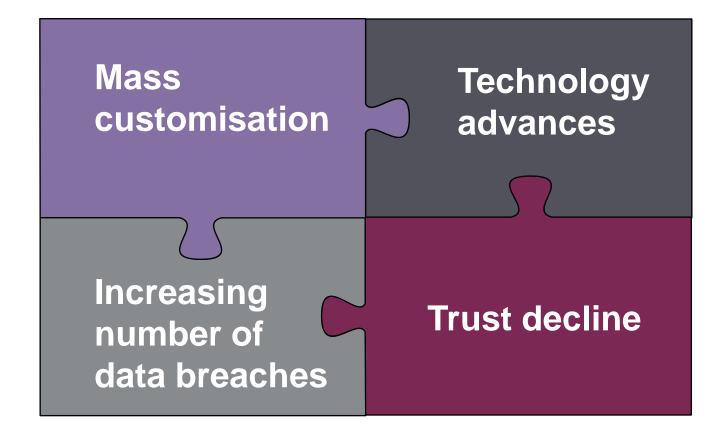


The careers you will have in 10 - 15 years' from now don't exist today.



Why is privacy such a hot topic?

Privcere



 $\mathbf{O} \bullet \bullet \bullet$

Generative Al

Privcere



Privcere imited by your imagination

TECHNOLOGY EXECUTIVE COUNCIL

How ChatGPT is changing the job hiring process, from the HR department to coders

PUBLISHED SAT, APR 8 2023-10:00 AM EDT | UPDATED MON, APR 17 2023-1:23 AM EDT

Kevin Travers

HEALTH TEC

Physician Satisfaction

University of Kansas Health System taps Abridge to roll out Al-based medical transcription for thousands of docs

Artificial Intelligence

By Anastassia Gliadkovskaya • Mar 3, 2023 12:03pm

Workforce

A&O announces exclusive launch partnership with Harvey

SHARE

NEWS 15 February 2023 (9 2 min read

Related people



Press contacts

For media enquiries relating to this topic, please contact

Allen & Overy (A&O), the leading international law firm, has broken new ground by integrating Harvey, the innovative artificial intelligence platform built on a version of Open AI's latest models enhanced for legal work, into its global practice. Harvey will empower more than 3,500 of A&O's lawyers across 43 offices operating in multiple languages with the ability to generate and access legal content with unmatched efficiency, quality and intelligence.

$\mathbf{O} \bullet \bullet \bullet$

Slides by Sasha Luccion

Unintended consequences?

"AGI"

Costs of Generative AI

Chat about any topic

Generate realistic images

False and misleading information Propaganda and deception

Biases and hallucinations

Homogeneity and misrepresentation of language/culture Harmful and violent content

Private information

Copyright infringement

Answer all your burning questions

Do your homework for you

Gather your data to improve models

Exploitation of underpaid workers

Erosion of rich human practises

Raising the barrier to entry in Al Tonnes of carbon emissions

Huge quantities of energy/water

Rare metals for manufacturing hardware

Crowdsourced via Twitter

Figure 1: Credit: Sasha Luccioni



Open Al Limitations: "GPT–4 still has many known limitations that we are working to address, such as social biases, hallucinations, and adversarial prompts. We encourage and facilitate transparency, user education, and wider AI literacy as society adopts these models. We also aim to expand the avenues of input people have in shaping our models."

"We spent 6 months making GPT-4 safer and more aligned. GPT-4 is 82% less likely to respond to requests for disallowed content and 40% more likely to produce factual responses than GPT-3.5 on our internal evaluations"



So, critical thinking skills are essential when deploying and using generative AI.

From a privacy perspective, I see three key categories of risks relating to:

- Data inputs into generative AI
- Data outputs from generative AI
- Governance



Input risks

 Data used to train generative AI - scraped from the Internet

Uses massive amounts of data

- Lack of transparency
- User prompt data used to further train generative AI?
- Data quality



Output risks

- Garbage in garbage out?
- Misinformation disinformation?
- Sources?
- Is it factual, fake, plausible but not real?
- Incorporates and promulgates biases?
- People tend to believe machines over people

Privcore

• Who is responsible if something bad happens?

• How do we allocate responsibility?

• How do we stop the proliferation of bad use cases?

• How do people get a say about what happens to their data?





Limited controls in generative AI – much like the original cars



Use safe environments to conduct testing – not the mass population

Take aways

- Build in security and privacy by design (build the brakes!)
- Respect laws

- Don't believe everything you see or hear apply your critical thinking skills
- Be open to the career that doesn't yet exist today







 <u>The AI Dilemma</u>, presentation created by the Center for Humane Technology

• <u>Women in Security magazine – my career journey</u>

 Parliamentary Inquiry to probe AI use in schools and higher education - a framework to guide schools in harnessing AI tools to support teaching and learning is being drafted

Disclaimer: Generative AI was not used in this presentation as far as Privcore is aware.