



# A question of trust?

How to use Generative AI responsibly

John Fitzgerald  
Digital Evolution Manager

- **How does it work?**
- **How are people using it?**
- **A few examples of controversial use of generative AI**
- **What next? A new kind of literacy/new pitfalls**

Health warning

**Generative AI is just one  
recent part of AI**

# Why is Generative AI getting so much traction?

- Massively expanding user base
- Easy to use, no data science required
- Becoming embedded in lots of common platforms (Bing, Bard, OpenAI plugins etc)
- Makes stuff up – surprisingly plausible(!)

Intro

# What is Generative AI?



**How does it work?**

# **ChatGPT explained on one slide, with thanks to**

Gina Helfrich, PhD

Centre for Technomoral Futures  
Edinburgh Futures Institute





- **Chat-based** (type questions, get answers)
- **Generative:** Makes stuff up! New content
- **Pre-trained:** Uses a language model based on a large amount of text data scraped from the internet
- **Transformer:** Algorithm that enables the system to predict the most likely next word for a given input

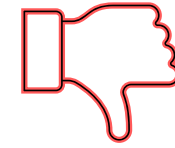


# How are people using it?



Help me  
brainstorm some  
initial ideas

Help me write some  
routine generic social  
media content



Help me draft a specialist  
policy when I don't know  
where to start

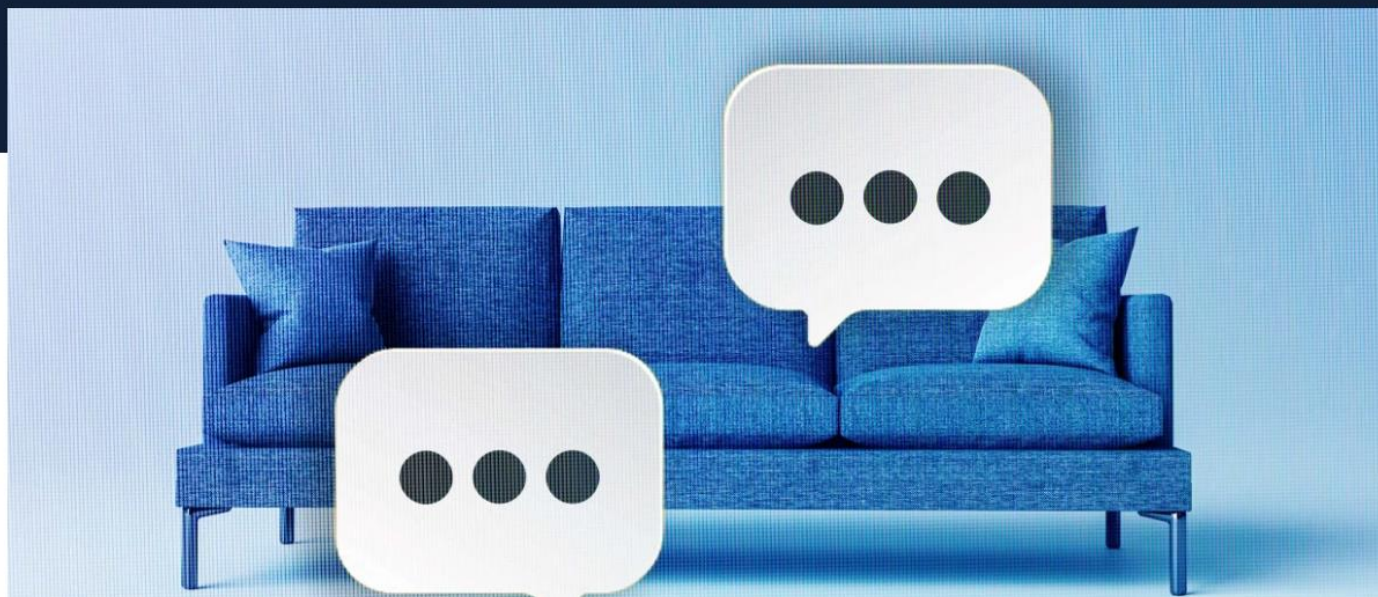
Help write a full-length  
funding application

**Examples in the wild**  
**– trust at risk!**

INTERNET

# A mental health tech company ran an AI experiment on real users. Nothing's stopping apps from conducting more.

A chat app used for emotional support used a popular chatbot to write answers for humans to select. Controversy followed.



## A Wellness Chatbot Is Offline After Its 'Harmful' Focus on Weight Loss

The artificial intelligence tool, named Tessa, was presented by the National Eating Disorders Association as a way to discover coping skills. But activists say it instead veered into problematic weight-loss advice.

Share full article



35



Matt Born/The Star-News, via Associated Press

A nationwide escalation in demand for treatment has been met with a shortage of providers, leading some mental health organizations to supplement care with chatbots and artificial intelligence that present a dilemma in public health: Is something better than nothing?

In March, the National Eating Disorders Association notified the staff members of a telephone helpline that the organization had operated for more than 20 years that they would be laid off, shortly after they had [formed a union](#). At the time, the staff was told that the organization would “wind down the helpline as currently operating” and “transition to Tessa, the AI-assisted technology, expected around June 1,” [NPR reported](#).



**Thank you**

Your support powers our independent journalism

**The  
Guardian**  
Newspaper of the year

**News**

**Opinion**

**Sport**

**Culture**

**Lifestyle**

More ▾

World ► Europe US **Americas** Asia Australia Middle East Africa Inequality Global development

**Colombia**

## Amnesty International criticised for using AI-generated images

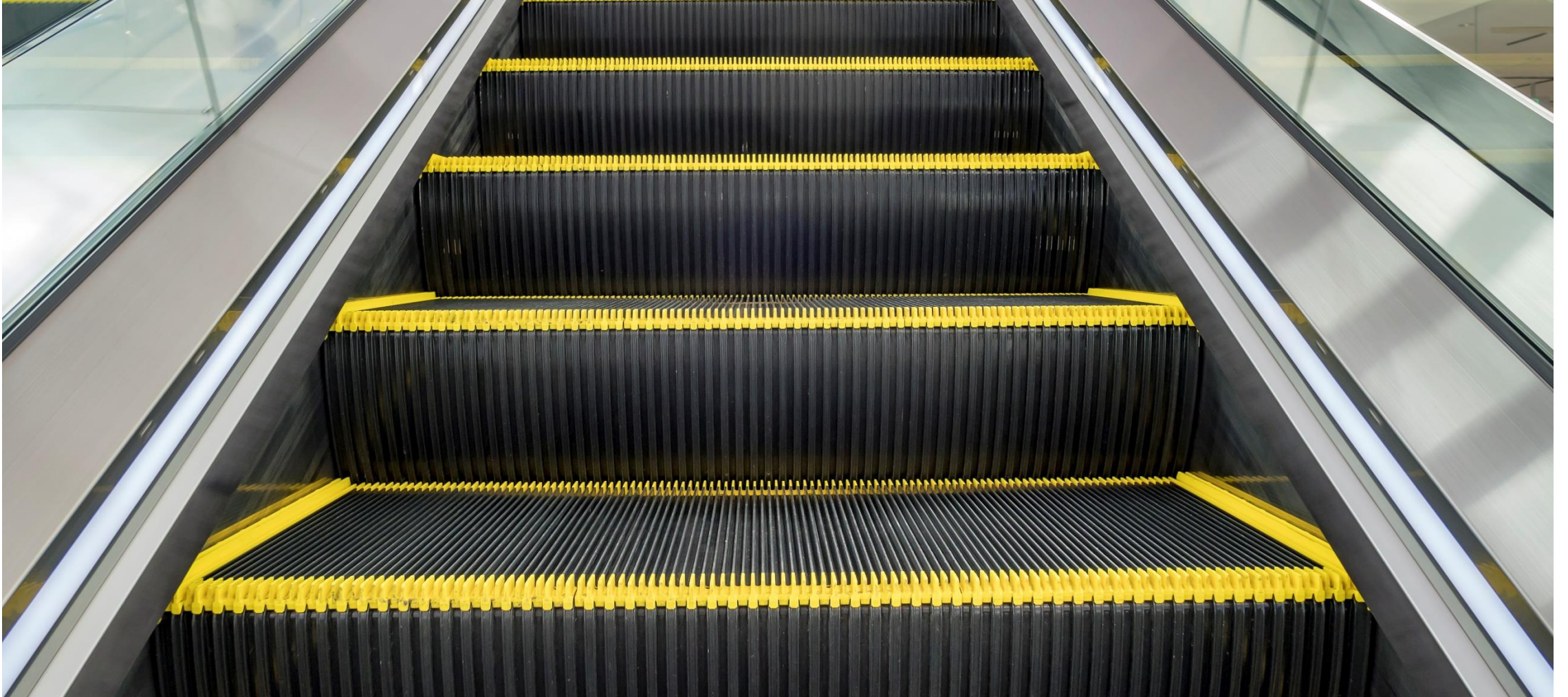
Group has removed AI images used to promote their reports on social media, including fake photos of Colombia's 2021 protests

**Luke Taylor in Bogotá**

Tue 2 May 2023 19.20 BST



# What now?





## Plus side

- Lowers the bar for people who find writing difficult
- Helps people brainstorm and generate ideas
- Helps people rephrase content

## Negative side

- Still inequitable – digitally skilled will get further ahead while folk with lower digital skills left behind
- If your literacy/expertise is low, you will struggle to check the quality of generative AI content
- Corrosive effect on trustworthiness/credibility

## Coming soon – keep an eye on 👁️

- AI tools that can look up real-time or recent web information. Eventually retrieval-based AI?
- Ongoing security hiccups eg Google Bard leak
- Vision and Voice with ChatGPT
- ‘Customisable’ Generative AI – but not 100% reliable
- Continued work on trust and safety

# What next?



# Longer-term consequences

- A new type of literacy
- ‘Co-pilot’ creativity
- Ethical questions about model development (copyright, privacy, bias, labour practices, environmental impact of compute)
- Significant and widespread changes to nature of work and workforce skills



## Longer-term consequences (contd)..

- Two-speed workplace?
- The degenerative internet?
- What does this mean for trusted and reliable content in general?

# What can you do now?

- Try it out yourself
- Have a conversation in your team
- Go to reliable guidance like SCVO's AI guide and the Scottish AI Alliance
- Come to our session at The Gathering on 8 Nov
- Don't panic, and remember that plausible is not the same as true

## Low risk examples – give it a go

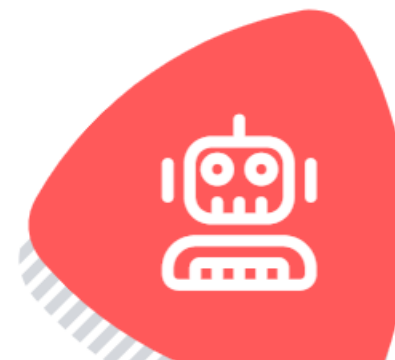
- **Using generative AI to help draft routine, generic content:** for example A/B testing emails, social media posts
- **Using generative AI to help you generate some ideas** or try alternative ways of phrasing content (especially helpful for solo writers)



## High risk examples – proceed with caution!

- **Using generative AI to draft full-length content** that will be used a policy or shared externally – how will you check it is coherent and correct?
- **Using generative AI to work on full-length drafts of funding applications** – how will you check it's accurate and meets the guidelines?

# Guide to **generative AI** for Scotland's voluntary sector



We're producing this guide in 2023, as more and more people start to use and discuss generative AI. It's clear that many people would like to know more about how these tools work, and would appreciate ideas and guidance to help them tackle the practical and ethical issues that come up.

the  
**Gathering**  
**2023**

Tuesday 7 & Wednesday 8 November  
EICC, Edinburgh  
#SCVOGathering

the  
**GatheringExtra**

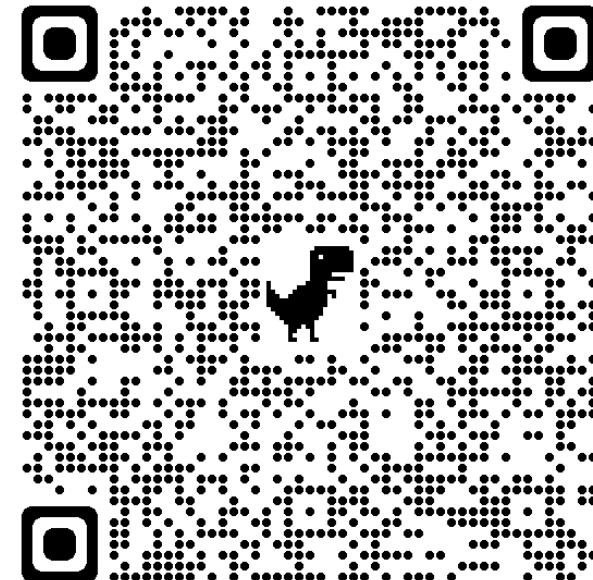
Wednesday 29 November  
Online

[← Back to full programme](#)

## Artificial Intelligence and the voluntary sector: potential and pitfalls

In-depth discussion with leading experts on the strategic potential for AI in the voluntary sector

**Weds 8 Nov 3-4pm**





# Thank you

[scvo.scot](https://scvo.scot)

© 2021. The Scottish Council for Voluntary Organisations (SCVO) is a Scottish Charitable Incorporated Organisation. Charity registered in Scotland SC003558.