

*Yikes, this is changing fast!*



# Contemporary Generative AI: Some headlines on developments and regulation

# Quick Orientation on Timelines & Concepts

AKA: The fastest diffusion of mediating technology in human history

- Excluding decades of previous conceptual development of AI models, including Geoffrey Hinton's 1986 publication (on backpropagation) that enabled the massive advancements in machine learning models, **we can begin the explosive current AI era with OpenAI's 2018 GPT model.**
  - **GPT:** Generative Pre-Trained Transformer - A Large Language Model (LLM)
  - **Transformer models** can self-direct their learning and may be let loose on massive datasets to make sense of them
- OpenAI was established (as a non-profit) in 2015, and transitioned to for-profit in 2019. The company is responsible for **ChatGPT** (Highly flexible and wildly popular chatbot), **DALL-E** (image generation tool), and **Whisper** (speech translation and transcription tool).
- The commercial and public impact of ChatGPT since launch (*30 November 2022!*) has prompted releases of similar products, genesis of new companies, and attention on new and pre-existing GenAI tools.
  - Important here: **Microsoft partnership with OpenAI** in January 2023 (AI-enabled **Bing**); Google's continued development of tools (LaMDA, PaLM, **Bard**); **Anthropic** launch and development of the **Claude** AI, **MidJourney**'s use (and misuse) of art online, and **StableDiffusion**'s legal trouble with *Getty Images*.

# Recent Important Developments

AKA: It's been a crazy last ten days

- OpenAI have announced (and opened beta access for) **vision and speech on ChatGPT**. The model can receive images and analyse them, including toward abstract goals; the model can also listen to and understand voice, as well as respond with speech.
  - Note the likely impact of this on our current perceptions of *prompt engineering*
- OpenAI have partnered with **Spotify** for translation of podcasts, using the vocal models of their hosts.
- Amazon have partnered with **Anthropic**, resulting in tool development for Amazon and access to data and AWS (and \$4B!) for Anthropic.
- **Meta** have announced **AI Chatbots**, leveraging US celebrities, which emphasise personality-based interaction. (US beta launch, likely to extend quickly; focus on keeping users on-platform).
- **HeyGen** received related attention for their tools for **AI avatar creation**, with speech and naturalistic movement

# Regulatory Landscape

AKA: How are we going to manage the most powerful tech?

- The US context for tech regulation has been a fraught space, defined by reluctance of political intervention based on philosophical views on innovation, personal liberty, etc.
  - 10 US states have state-level laws in force or drafted which impact AI use. California and NY are notable in mirroring aspects of EU regulation and linking to data privacy.
- The EU context is part of the wider *Europe's Digital Future* strategy for tech transformation, and builds on data privacy, tech regulation, and reporting efforts linked to **GDPR** and **DSA**.
- The draft **European Union Artificial Intelligence Act (EU AI Act)** is an attempt at **horizontal legislation** that will broadly define the regulatory prohibitions for tech in this space, the use of personal data involved, and the levels of risk associated with AI.
  - It will identify unacceptable risks, and classify others.
  - It will require transparency obligations from AI tool producers that are directed at informing the user about the process of interaction.
  - It will include structural guidelines on data governance and link to existing data regulation

# Important parts of the EU AI Act

AKA: What the court battles will be fought over

- **Article 5** of the draft *EU AI Act* will specifically ban certain tools or technology:
  - Subliminal harmful effects of AI on humans: "techniques beyond a person's consciousness" that materially affect behaviour that (§5.1.a)
  - Tools exploiting vulnerable groups and tools that trustworthiness or related social scoring
  - Real-time remote biometric identification tools - except for certain carve-outs related to missing children, terrorism, etc (§5.2.a/b)
- **Article 6** will classify high-risk systems and tools with unacceptable risk, and is still subject to revision and development

# Regulatory Imperatives

AKA: What should I be thinking about now, as I use GenAI?

- ✓ Stay up to date on developing tools and consider how these might be relevant either to your own communications or to the *shifting expectations and norms of interaction* of your audiences.
- ✓ Stay up to date on partnerships, which can provide an indication of the likely diffusion of innovations and path to increase market share and *interaction type* by platforms and online services
- ✓ Be attentive to the developing EU AI Act, and note the likely requirement to address transparency in your usage of AI, especially in decision-making, profiling, targeted communications, etc.
- ✓ Consider developing (and regularly updating) a policy on AI usage, and communicate this to employees and co-workers.
- ✓ Be strategic about the testing of new tools in a space of rapid change
- ✓ Subscribe to some tech podcasts!