

Nordic Region

Building a Human-Centered and Responsible AI

PECB REGIONAL WEBINARS

Olena Sarkar



Co-founder, CEO at APE AI

Simplifying implementation of AI with a no-code platform and ready-to-use AI agents

Background

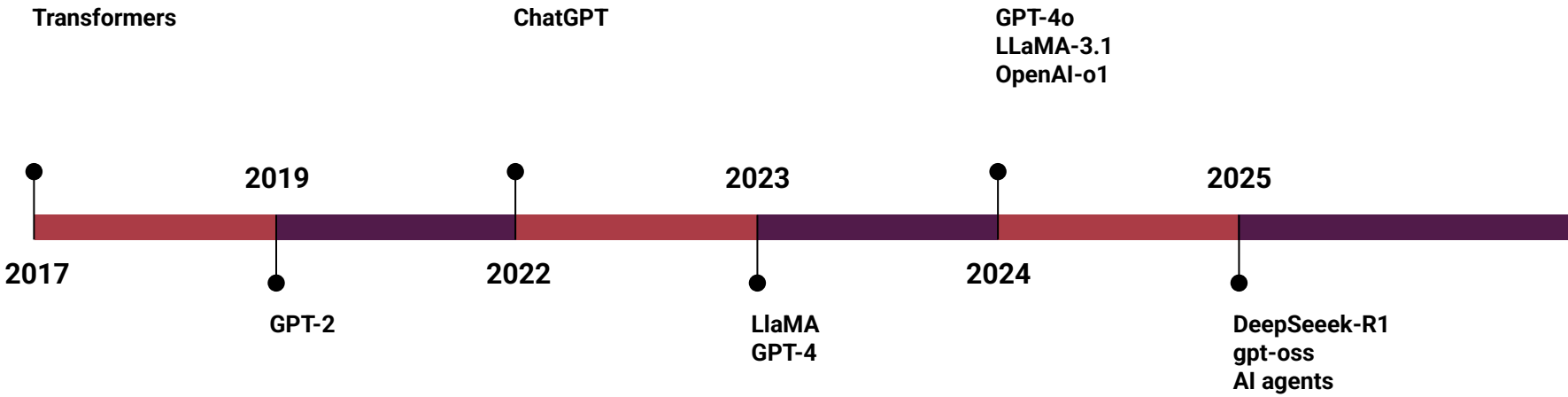
MBA from Copenhagen Business School
Consulting public and private sector on large scale digital transformation
Infrastructure Ontario, Capgemini, Orsted, etc.

https://ape-ai.dev

Agenda

1. What is HCAI and why we need it.
2. Key principles of HCAI.
3. Responsible AI and its key principles.
4. The level of autonomy.
5. Human-computer interaction.
6. Other key considerations for designing AI systems.
7. Q&A (15 mins)

AI and LLMs have been developing rapidly



Building effective AI systems for humans

'Traditional' AI	Human-centered AI
• Technical performance is highest priority	• User experience, seamless human-AI interaction
• Automation above all	• Augmenting rather than replacing human abilities
• Focus on productivity	• Foster creativity and innovation
• ***Example***: *AI that automates grading process*	• ***Example***: *AI platform for teachers to improve content*

Effective AI systems are not about blind automation. They don't view AI as a standalone tool. **Cooperation between humans and machines is a fundamental belief** when it comes to developing such systems.

Why we need HCAI?
-> more inclusive
-> trustworthy
-> beneficial AI applications

Principles of HCAI

1. Balance between automation and human control
2. Transparency and explainability
3. Empathy and understanding the user (user centered design)
4. Inclusive design

Responsible AI

A practice and framework that involves the consideration of a broader societal impact of AI systems and the measures required to align these technologies with stakeholder **values, legal standards and ethical principles.**

Principles of Responsible AI

1. Fairness and inclusiveness
2. Reliability and safety
3. Privacy and security
4. Transparency
5. Accountability

4. The level of autonomy.

Why would we not want fully autonomous AI?

For most tasks the roles will look like this:
AI = generation, human = verification

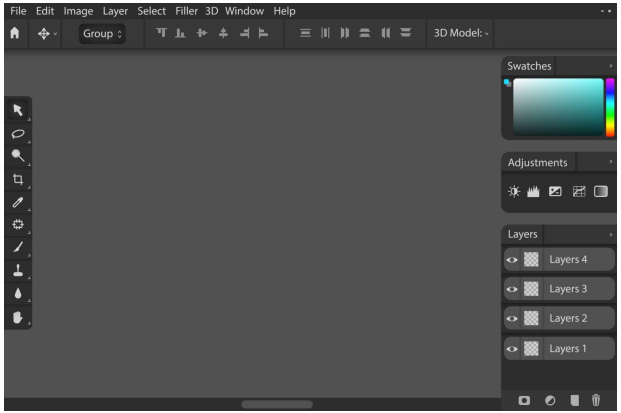
Full autonomy	Partial autonomy
• Risk of errors from unexpected scenarios	• Gradual approach, fast implementation and ROI
• Difficulty to interpret outputs due to volume	• Small chunks of output lead to catching errors faster
• Slow verification loop	• AI with narrow focus will yield less errors, scalable
• ***Example***: *AI writing code for entire software*	• ***Example***: *AI for working with spreadsheets*

It is best to decide the amount of autonomy based on specific task at hand.
E.g. deploying airbag vs preparing presentations.

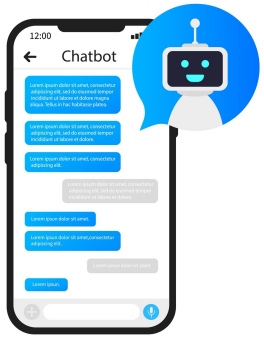
5. Human-computer interaction.

How do we align HCI with HCAI core principles?

HCI before AI and its interface



Conversational Interface



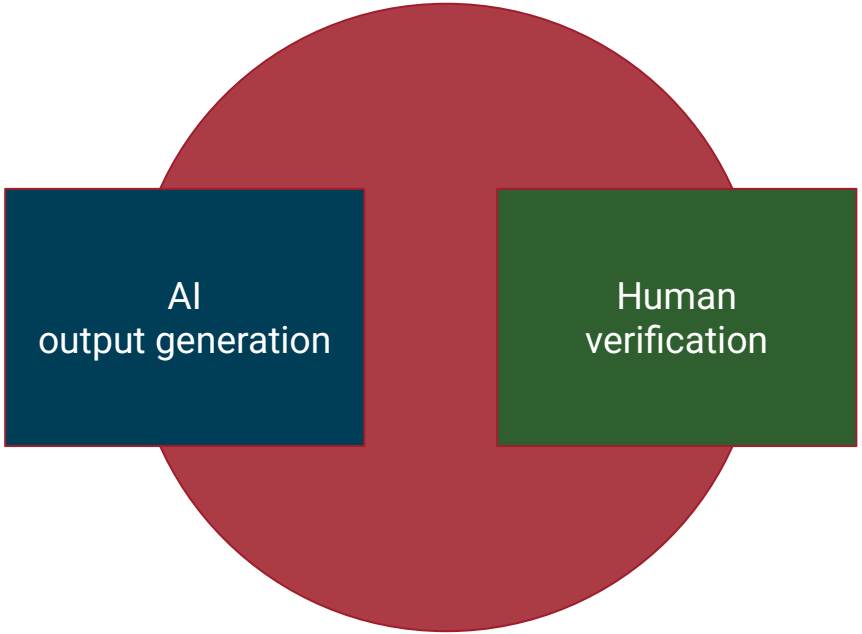
Custom GUI fit for task



Principles of HCAI

1. Balance between automation and human control
2. Transparency and explainability
3. Empathy and understanding the user (user centered design)
4. Inclusive design

Who needs to be in the loop, when and why.



Who:
- domain expert / user

When:
- design
- development
- testing
- using AI products and services

Why:
- governance and monitoring
- verifying outputs for any errors
- making ultimate decisions

Privacy and security in AI

Some measures:

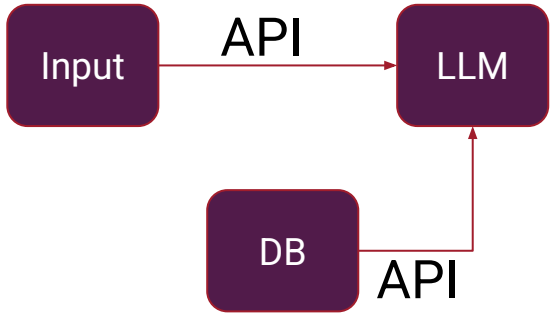
- Encryption
- Secure storage
- Strict access controls

Best approach:

- Security by design

Minimizing data exposure:

Instead of relying on APIs,



Leverage open source LLMs and self hosting

LLaMA by Meta

gpt-oss

Techniques to improve accuracy of AI outputs

1. Fine tune AI system to real world use cases and workflows
2. Leverage multi-agent systems
3. Apply relevant context
4. Use model fine-tuning
5. Select suitable LLM
6. Work on prompts and instructions

In conclusion

Humans are the users of AI products and services
Augment human capabilities, balance autonomy and control
AI applications need GUI that is fit for task at hand
Human in the loop approach increases reliability
Security by design, open source LLMs, self hosting for privacy

Q&A Session

