HCI4AI:

The EU AI Act and the Crucial Role of Human Computer Interaction in Developing Human-Centric AI

Dympna O'Sullivan
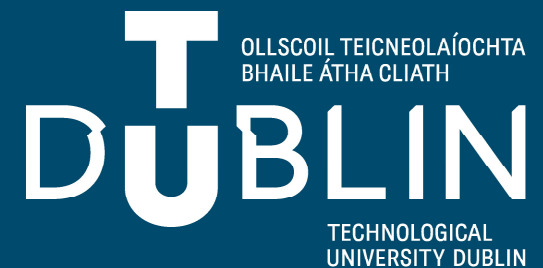
About Me

ACADEMIC LEAD
DIGITAL FUTURES RESEARCH HUB, TU DUBLIN
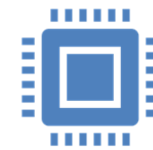
MY WORKS INTERSECTS
AI AND HCI

Outline

Why do we need regulation for AI?

Brief introduction to new AI regulation – the EU AI Act

What part can HCI play in achieving the aims of the EU AI Act?

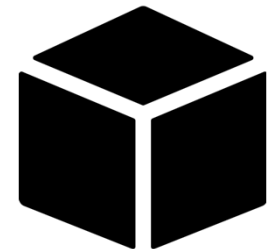# Why do we regulation for AI?

Ubiquitous Nature of AI in Society

Medicine | Recruitment | Entertainment | Judicial System

Known AI Issues

- **Fairness – Bias and Discrimination**
- **Accountability**
- **Lack of Transparency**
- **Reliability**
- **Privacy**
- **Copyright**
- **Anthropomorphism**
- **Trust Gap**
- **Failure of regulators to get a grip on social media**

The EU AI Act

- A legal framework governing the sale and use of artificial intelligence in the EU

- Set to become the world's first comprehensive legal framework for AI

Key Objectives of the Act

- **Focus on building trust, ensuring transparency, and fostering innovation**

- **Ensuring the creation of safe and accountable AI systems**

- **Establishing a competitive European AI ecosystem**

- **Enhancing public trust and acceptance of AI technologies**

Core Principles of the Act

- **Human-Centric AI**
- **Transparency**
- **Accountability**

Risk-Based Classification of AI Systems

04 Unacceptable risks: prohibited
Social scoring, social manipulation...

03 High risk: compliance
Education, justice, law, labour...

02 Limited risks: transparency
Chat bots, deep fakes...

01 Minimal risk: General security
Spam filters, video games...

Risk-Based Classification of AI Systems

AI DEVELOPERS MUST UNDERGO A CONFORMITY ASSESSMENT TO ENSURE COMPLIANCE WITH THE RISK-BASED APPROACH

NEED FOR CONTINUOUS MONITORING AND REPORTING OF AI SYSTEMS TO DETECT AND ADDRESS EMERGING RISKS PROMPTLY

REGULATORY MEASURES ARE PROPORTIONATE TO THE IDENTIFIED RISKS

Compliance and Penalties

Conformity assessment bodies for assessing and certifying AI systems' compliance with the EU AI Act

The heftiest fines are imposed for violating the prohibition of specific AI systems, up to 40M EUR or 7% turnover

Lowest penalties for providing incorrect, incomplete or misleading information, up to 5M EUR or 1% of annual worldwide turnover

Impact and Implications

Impact on businesses

Cost of compliance

New models for responsible innovation

Impact on Consumers:

Benefit from increased transparency and clear information including potential risks

International Influence

AI regulation may influence international standards

Where is the EU AI now?

- Entering the Trialogue process
  - Final text expected to be adopted in January 2024
  - Transition phase
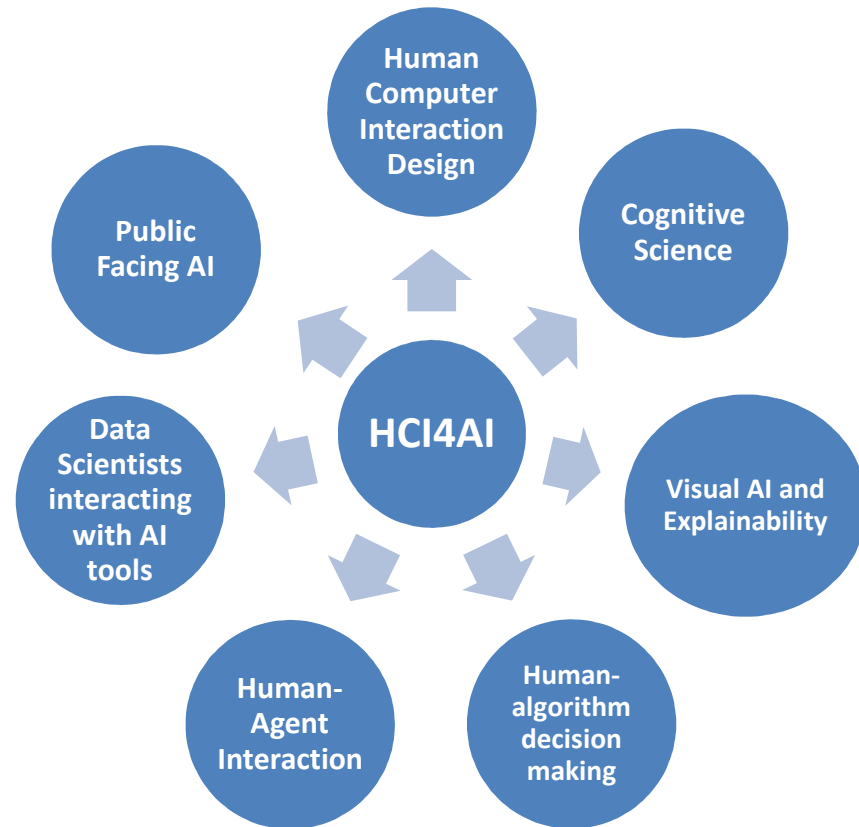 - Become law in 2025

Key Requirements of Trustworthy AI

From: https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai

The EU AI Act and the role of HCI

- The human experience of AI
  - It is imperative to understand AI systems from a human perspective
  - Insights into Human AI Interaction will help us drive better design of AI systems
  - It will also allow us to better understand their implications for individuals and society

The Human Experience of AI

Human Computer Interaction Design

Cognitive Science

Public Facing AI

HCI4AI

Visual AI and Explainability

Data Scientists interacting with AI tools

Human-Agent Interaction

Human-algorithm decision making

HCI as a methodology to drive human-centric, interpretable and trustworthy AI

Human Computer Interaction Design for AI

Sense and Sensemaking

Interfaces for Human AI collaboration

UX for AI

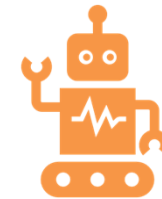Co-design and co-creation

Human in/on the loop

Cognitive Science

To better understand and replicate human intelligence

Understanding mental models for human-AI interactions

Impact of AI on individual users' sense of agency

Visual AI and Explainability

- What are the most effective ways for an AI to translate its actions into terms the user understands?
  - What are the most effective ways for an AI to promote transparency?
  - What should different explanations for different end users look like?
  - How can we measure the effectiveness of AI transparency and explainability?

Human-Algorithm Decision Making

Collaborative decision support systems

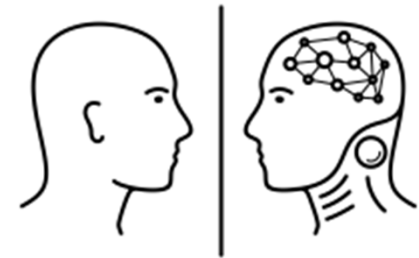Interactive visualizations

Context-aware interfaces

Customizable interfaces

Ethical considerations for interface design

Human-Agent Interactions

- Interfaces that support natural language interaction to enable users to communicate with agents using everyday language

- Social cues for agent interfaces to make interactions more relatable and engaging

- Interfaces that allow agents to adapt to individual user preferences over time

Data Scientists interacting with AI tools

- Interactive visualizations to help explore and interpret complex datasets and AI model outputs

- Prioritize the explainability of AI models and provide transparent insights into their decision-making processes

- HCI designs that highlight and address ethical considerations
  - fairness, accountability, and bias mitigation contribute to responsible and ethical data science practices

Public Facing AI

- Community workshops to introduce AI concepts, applications, and their societal impact

- Creating AI-driven installations that engage the public and showcase the intersection of AI and creativity

- Collaborating with museums to create interactive exhibits that explain the principles of AI to provide an accessible way for the public to learn about AI in a cultural context

- Public demonstrations focused on how AI can be used for social good and solving real-world challenges
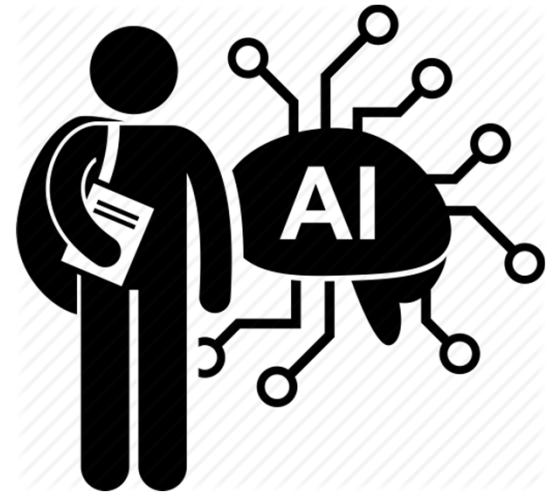
Conclusions – HCI4AI

- User understanding results in interactions that are intuitive and align with the way humans naturally interact with technology
- By focusing on transparency and explainability, HCI enhances user trust in AI systems
- Working with end-users ensure AI accessibility and inclusion
- Contributes to user education and empowerment about AI systems
- Ensuring that AI systems are developed and deployed in an ethically responsible manner
- Well-designed AI systems results in positive user experiences, making users more likely to embrace and trust AI technologies