SHARON OVIATT
Human-Centred AI
Monash University

WHAT'S HUMAN-CENTERED A.I. (HCAI)?

Beyond the basic tools, algorithms, and technical foundations of A.I.,
human-centered A.I. will enable a deeper understanding of how
A.I. technology can be designed to support ***human goals, activities, and values.***

For several decades, the fields of ***A.I. and HCI/Human-centered systems have been fusing***,
with the shift most evident in the transition of HCI from its origin in discrete GUIs
to ***recognition-based natural multimodal interfaces.***

HCI COMPETENCIES REQUIRED TO DESIGN A.A. SYSTEMS

1. What interaction paradigms & modalities best support interaction with users?

Deeply human-centered ones, based on natural multimodal behavior & communication modes.

2. How do A.I. systems fail due to poor knowledge of HCI theories, principles and methods?

By adopting an engineering/M.L. approach without multidisciplinary input to model people and domains (e.g., social and medical sciences)

3. What HCI design and user testing practices could be adopted by A.I.?

To succeed in supporting human goals, activities and values, A.I. systems need to embrace proactive user design, participatory design, and iterative system evaluation.

4. What theories & methods should be used to create A.I. systems that best empower users?

A.I. system design needs to be based on broad multidisciplinary teams (e.g., ethicists, social scientists, political scientists, engineers) to ensure prosocial human-centered systems. This includes, but goes well beyond, simply teaching "HCI."

FOUR EXAMPLES OF CORE AREAS OF HUMAN-CENTERED A.I.

Design of synergistic human-A.I. systems that augment human intelligence, maximize net system accuracy and minimize net bias in different domains

Development of explainable A.I. systems, including (1) explanation at a level that successfully enables human understanding needed to complete tasks in given domains; (2) at the level of theory

Design of deeply human-centered A.I. systems that can adapt to contextual information about users' cognitive, emotional, health and mental health status (i.e., not simply concrete factors like location)

Comprehensive evaluation of the impact of A.I. systems on human lives and societies to ensure desirable outcomes for diverse constituencies (e.g., not just governments or corporations)

HUMAN-CENTERED A.I. COURSE: TOPICS

(planned for A.I. Masters program)

1. Introduction, Foundations, Models & Theory for Human-centered A.I. Systems
(e.g., how AI differs from HCAI; multidisciplinary foundations of HCAI; models of HCAI; goals & strategies of designing HCAI systems; concepts of transparent & fair A.I. systems)

2. History & Design of Deeply Human-centered A.I. Systems
(e.g., systems that understand natural human multimodal communication, and that conduct predictive analytics on human behavior, health, learning, etc.)

3. Platforms & Tools for Developing Human-Centered A.I. Systems
(e.g., datasets for exploring HCAI; tools for analyzing human-centered signal patterns; models & platforms for developing HCAI systems) **[Student Projects]**

HUMAN-CENTERED A.I. COURSE: TOPICS (cont.)

4. Major Trends in Impact of A.I. on Users and Society

(e.g., summary of positive impacts, such as medical innovation & cases where AI exceeds human accuracy; cases of negative impacts, such as biased systems, systems that degrade human autonomy, learning or mental health, or that disseminate propaganda and fuel militaristic or totalitarian regimes)

5. Legal & Regulatory Dilemmas related to A.I. Ethics

(e.g., legal accountability for wrongful death & other human loss; international privacy laws safeguarding human welfare & dignity; public policy & regulatory challenges)

Discussion: What should an undergraduate course or major in HCAI require?

(core courses in HCI methods, ML techniques, ethics & society, multidisciplinary domain perspectives & "challenge" cases)