

H9HCAI: Human Centered Artificial Intelligence

Module Code:	H9HCAI
Long Title	Human Centered Artificial Intelligence APPROVED
Title	Human Centered Artificial Intelligence
Module Level:	LEVEL 9
EQF Level:	7
EHEA Level:	Second Cycle
Credits:	10
Module Coordinator:	Paul Stynes
Module Author:	Shauni Hegarty
Departments:	School of Computing
Specifications of the qualifications and experience required of staff	
Learning Outcomes	
<i>On successful completion of this module the learner will be able to:</i>	
#	Learning Outcome Description
LO1	Demonstrate expert knowledge of the theory and concepts underpinning human centered AI.
LO2	Determine the design requirements for human centered AI systems.
LO3	Critically analyse the capabilities and limitations of AI systems based on the governance structures of the human centered AI.
LO4	Investigate and critically assess the impacts of reliability, trustworthiness, fairness, accountability, and transparency in AI.
LO5	Evaluate and present the adherence of AI systems to the human centered AI guidelines
Dependencies	
Module Recommendations	
No recommendations listed	
Co-requisite Modules	
No Co-requisite modules listed	
Entry requirements	

H9HCAI: Human Centered Artificial Intelligence

Module Content & Assessment			
Indicative Content			
Introduction to Human Centered Artificial Intelligence • How do rationalism or empiricism provide sound foundations? • Are people and computers in the same category? • Will automation, AI, and robots lead to widespread unemployment? • Harnessing the benefits of emulating humans and empowering people • Trade-offs between emulating humans and empowering people			
Rising above the levels of automation • How to design to safely increase human performance? • Understand the situations to apply full human or full computer control • Balance between human and computer control			
Two dimensional HCAI framework • Introduction to two-dimensional HCAI framework • Categorisation of systems using the framework			
Design guidelines and examples • Introduction to the HCAI guidelines • Application of the guidelines to design HCAI systems • Example of systems developed using the HCAI guidelines			
Defining reliable, safe & trustworthy systems • What means a reliable, safe, and trustworthy system? • What determines the reliability, safeness, and trustworthiness of a system? • How to measure reliability, safeness, and trustworthiness of a system?			
Governance Structures for HCAI How to bridge the gap from ethicsto practice • Introduction to the three-layer governance structure for HCAI systems • Application of the governance structure			
Reliable AI systems Audit Trails and Analysis Tools • Verification and Validation Testing			
Trustworthy certification by independent oversight • Introduction to oversight methods • Government Interventions and Regulations • Accounting Firms Conduct External Audits • Insurance Companies Compensate for AI Failures • NGO, Professional Organisations, and Research Institutes			
Fairness in AI • The meaning of fairness with respect to AI • Perceptions of algorithmic bias and unfairness • Legal, social, and philosophical models of fairness • Methods, tools, and standards for ensuring that algorithms comply with fairness policies (e.g., IEEE P7003 TM) • Mitigating biases in systems, or discouraging biased behaviour from users			
Accountability in AI The meaning of accountability with respect to algorithmic systems • Strategies for developing accountable systems • Methods and principles for accountable algorithms (e.g., FAT/ML Principles for Accountable Algorithms, Social Impact Statement for Algorithms)			
Transparency in AI • The meaning of transparency with respect to algorithmic systems • Tools, models, and principles for AI explainability and transparency (e.g., ACM Principles for Algorithmic Transparency and Accountability, NIST Principles of Explainable AI) • Trade-offs between privacy and transparency • Tools and Frameworks for conducting ethical and legal algorithm audits			
Human Centered Approach to AI Ethics • Problems in AI and robot ethicsthat arise due to cognitive states • Define what is welfare and responsibility • Review of HCAI approach to resolve some of ethics problems			
Assessment Breakdown			%
Coursework			30.00%
End of Module Assessment			70.00%
Assessments			
Full Time			
Coursework			
Assessment Type:	Formative Assessment	% of total:	Non-Marked
Assessment Date:	n/a	Outcome addressed:	1,2,3,4
Non-Marked:	Yes		
Assessment Description: Formative assessment will be provided on the in-class individual or group activities. Feedback will be provided in written or oral format, or on-line through Moodle. In addition, in class discussions will be undertaken as part of the practical approach to learning.			
Assessment Type:	Continuous Assessment	% of total:	30
Assessment Date:	n/a	Outcome addressed:	1
Non-Marked:	No		
Assessment Description: Discuss the challenges an organisation faces for adopting AI due to the differences between humans and computers. How can HCAI improve and enhance this experience?			
End of Module Assessment			
Assessment Type:	Terminal Exam	% of total:	70
Assessment Date:	End-of-Semester	Outcome addressed:	1,2,3,4,5
Non-Marked:	No		
Assessment Description: The examination will be of two hours duration and may include a mix of: theoretical, applied and interpretation questions			
No Workplace Assessment			
Reassessment Requirement			
Coursework Only <i>This module is reassessed solely on the basis of re-submitted coursework. There is no repeat written examination.</i>			

H9HCAI: Human Centered Artificial Intelligence

Module Workload				
Module Target Workload Hours 0 Hours				
Workload: Full Time				
Workload Type	Workload Description	Hours	Frequency	Average Weekly Learner Workload
Lecture	Lectures	24	Per Semester	2.00
Independent Learning	Independent Learning	202	Per Semester	16.83
Tutorial	Practical/Tutorials	24	Per Semester	2.00
Total Weekly Contact Hours				4.00

Module Resources	
<i>Recommended Book Resources</i>	
Shneiderman, B. (2022). Human-Centered AI. Oxford University Press. Dubber, M. D., Pasquale, F., & Das, S. (Eds). (2020). The Oxford Handbook of Ethics of AI. Oxford University Press. [ISBN 978-0190067397].. O'Keefe, K. & O'Brien, D. (2018). Ethical Data and Information Management. Kogan Page. [ISBN: 978- 0749482046]..	
<i>Recommended Article/Paper Resources</i>	
Chrisley, R. (2020). A human-centered approach to AI ethics: A perspective from cognitive science. The Oxford Handbook of Ethics in AI. Oxford University Press. DOI: 10.1093/oxfordhb/9780190067397.013.29.. Shneiderman, B. (2020a). Human-centered artificial intelligence: Three fresh ideas. AIS Transactions on Human-Computer Interaction, 12(3), 109-124. DOI: 10.17705/1thci.0013.. Shneiderman, B. (2020b). Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human-Computer Interaction, 36(6), 495-504. Shneiderman, B. (2020c). Design lessons from AI's two grand goals: Human emulation and useful applications. IEEE Transactions on Technology and Society, 1(2), 73-82. Shneiderman, B. (2020d). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. ACM Transactions on Interactive Intelligent Systems. 10(4), Article 26. DOI: 10.1145/3419764. Hassani, H., Silva, E. S., Unger, S., TajMazinani, M. and Mac Feely, S. (2020). Artificial Intelligence (AI) or Intelligence Augmentation (IA): What is the future? AI, 1(2), 143-155. DOI: 10.3390/ai1020008.	
<i>This module does not have any other resources</i>	
Discussion Note:	