

Basic details

UID		Cohorts covered	Earliest cohort 2026-27
Long title	Responsible AI ethics at work		
New code	ELEC70141	New short title	
Brief description of module (approx. 600 chars.)	This module supports you in a discursive, case study centred, exploration of moral tensions and extended responsibilities that arise during processes of AI design and development. It guides you through debates about the systemic sociotechnical dynamics with which AI are increasingly entangled. You will learn about managing complex stakeholder relationships over time and become better informed about the regulatory regimes that apply to different forms of AI around the world. The module will focus on the specificities of AI and the process of making trades offs between opportunities and risks so that we don't systematically exclude the vulnerable or exacerbate prejudice.		
Available as a standalone module/ short course?	N		

Statutory details

	ECTS	CATS	Non-credit	
Credit value	5	10		HECOS codes
FHEQ level	Level 7			

Allocation of study hours

	Hours	
Lectures	15	
Group teaching	0	<i>Incl. seminars, tutorials, problem classes.</i>
Lab/ practical	0	
Other scheduled	5	<i>Incl. project supervision, fieldwork, external visits.</i>
Independent study	105	<i>Incl. wider reading/ practice, follow-up work, completion of assess</i>
Placement	0	<i>Incl. work-based learning and study that occurs overseas.</i>
Total hours	125	
ECTS ratio	25.00	

Project/placement activity

Is placement activity allowed?	No
--------------------------------	---------------------------------

Module delivery

Delivery mode	Taught/ Campus	Other	
Start term	Autumn	Other	

Ownership

Primary department	Electrical Engineering / I-X
Additional teaching departments	ICBS
Delivery campus	White City

Collaborative delivery

	Collaborative delivery?	N
External institution	N/A	
External department	N/A	
External campus	N/A	

Associated staff

Role	CID	Given name	Surname
Module Leader		Susan	Scott

Learning and teaching

Module description

Learning outcomes	At the end of the module you should be able to: (1) Explain why the call for responsible AI has emerged and identify the approaches adopted by stakeholders within different organizational contexts. (2) Critically discuss theories and frameworks relating to the processes and practices of responsible AI. (3) Deepen understanding of the contemporary emerging technologies covered on the course with particular emphasis on the dynamic relationship between applications of AI and their ethical implications. (4) Critically evaluate the discourses that characterise inclusive and reflective approaches to the management of contemporary emerging technologies in workplaces. (5) Assess practices of accountability, responsibility, fairness, and sustainability in relation to digital innovation projects. (6) Formulate your own considered views on particular aspects of responsible AI examined in the course. (7) Critically appraise case study data gathered as part of the course assessment with a view to both making and taking an informed research perspective on responsible digital innovation both in writing and in class discussion. (8) Enhance your research skills by consulting relevant sources of information for the purpose of presenting and discussing case studies both in writing and in class discussion. (9) Develop your understanding of the social, political, and economic context in which stakeholders are attempting to achieve responsible AI and manage its consequences.
-------------------	--

Key topics covered session-by-session (20 hr module organised as 5 x 4 hr sessions):

1i. Understanding the responsibility gap: Examines how a tech ethics crisis emerged, illustrated by high profiles cases. Relates Christensen's notion of modularization and specialisation to the emergence of gaps. Discusses how emerging contemporary technologies overflow existing regulatory frameworks. Identifies responsibility gaps. Asks students to share their stories of hope and concern surrounding AI technologies.

1ii. How do we become responsible? Introduces multiple, different theories of responsibility attribution and weighs up their strengths and limitations. Argues that a conventional notion of responsibility is premised upon attributing duties, obligations and/or blame to an individual human actor which becomes challenging considering the distributed agencies of contemporary emerging technologies. Introduces the notion of extended agencies and a range of sociotechnical approaches. Using a stakeholder approach, tensions and issues relating to the attribution of responsibility are explored in different settings.

2i. Understanding the accountability gap: Notes that responsibility mechanisms are more effective when they are situated in a conducive institutional context. Provides a brief overview of how approaches to CSR have evolved over time and how the notion of 'stakeholding' has changed (shared value, sustainability etc). CSR is critiqued to highlight multiple dominant discourses and cultural reference points to achieve a more global perspective. Considers whether the recent adoption of ethics by big tech might constitute 'ethics washing'. Examines where, when, and how accountability mechanisms are designed into systems of work.

2ii. Session on case studies and essay preparation

3i. How do we work responsibly with AI? Explores the challenges and opportunities of AI collaboration using case studies from different settings. Theories of collaboration are weighed up using examples in different settings. We explore how much we understand human-human collaboration and then consider what it means to have an AI co-pilot. How do AI collaborations reconfigure the attribution of responsibility? Do responsibility gaps emerge and what kind of accountability mechanisms might be used to counter the ethical tensions and issues in different settings?

3ii. Understanding distributed/shared responsibility: Introduces approaches to responsible innovation, exploring their strengths and limitations. Reviews historical work design, how different configurations of work have impacted wellbeing. Considers the different frameworks, guidelines, principles, tools, and practices that might lead to sustainable responsible innovation. Discusses the notions of transparency, explainability, and how they relate to responsibility and accountability.

4i. How to enact fairness in big data and AI tools? Breaks down the Neutrality Fallacy and investigates how, when, and why bias enters data practices. Examining forms of human bias that emerge in the workplace and how it relates (or not) to machine bias. Debating fairness and understanding it as a historically sensitive, situated notion. Discussing other ways of developing Information Ethics. Exploring Floridi's notion of Artificial Evil and Environmental Macroethics.

4ii. Learning about corporate ethics in action (guest talk). For example, discussing the difference between compliance and ethics with a corporate legal counsel and company ethicist. Hearing about real cases and the trade-offs that were needed to achieve more positive outcomes. How might we achieve change from within large corporations? Are born digital firms better placed to make a difference?

5i. Reimagining the responsible firm: Reviewing the various approaches (in theory and practice) to better integrate ethics at work. Examining the motivations and conditions of possibility for Purposeful Organizations; 360 degree corporations; Regenerative Organizations; Senge's Fifth Stage of Sustainability. Discussing different examples of ethical business e.g. social innovation, ethics by design, sustainability start-ups. Considering the extent to which redesigning platforms and infrastructures might enable ethical outcomes e.g. privacy enhancing AI technologies.

5ii. Managing the balance between hope and concern - Explores the implications of opening up an ecology of responsibility relations. If we have to arrive at trade-offs, where, when, why and how do we make them? Considers the notions of institutional entrepreneurship and liminal innovation. Encourages students to reflect on how their sense of hope and concern has changed during the course. Discussion of the case studies that they have engaged with on the course, both on the syllabus and through their own assignments. Maps out careers that would lead to working as corporate

Learning and Teaching Approach	Teaching will be practice-centred using case studies as the basis for discussions and debates. Case studies are an appropriate method for teaching and learning about challenges in the workplace surrounded by uncertainty where there is no 'right' and 'wrong' answer. This teaching method enables discussion and debate, applying theories and ideas to a concrete situation leading to better understanding of different management approaches. The learning and teaching approach is reflective, using a case study approach to examine how the uncertainties and trade-offs surrounding ethics at work relate to theoretical debates. The ethical challenges surrounding contemporary emerging technologies are the core focus with specific case studies varying from year to year. Each session will be comprised of short lectures (10-15mins), interspersed with group discussion of set readings (selected journal articles, book chapters, and policy documents), and interactive exercises. The case study teaching will involve a mix of role play, small group exercises, debates, and group discussion. Where appropriate, guest speakers will be invited to share their experience of how ethics at work is enacted in practice to ground and illustrate key themes on the course. Students are encouraged to reflect on the similarities and differences between theory and practice in different contexts.
Assessment Strategy	Individual - mini case study report - 40% - 2000 words Individual - Essay - 50% - 2500 words Class participation - 10% - Assessed by leader/GTA Description and logic of assessment A formative assignment will be submitted at the end of week 3. This is comprised of a short stakeholder analysis identifying interest and power dynamics over time in a given setting. It prepares them for the mini case study report in which the different perspectives need to be examined and put in context. The original mini case study (2000 words) submitted approx. half way through the module is worth 40% of the total mark for the module. This provides an opportunity for module leaders to discuss the strengths and limitations of different case study options with the students. It enables the module leader to initiate discussions about key points of knowledge and development of skills that you will need to be active learners in this topic area. For example, our discussion of the different case studies being produced by class members will give us an opportunity to practice how to deconstruct the data they gather. This will involve stakeholder analyses (building on formative) to help you foreground the dynamic relationship between digital innovation, practices, and their ethical implications. In acknowledgement that management practice is under pressure to catch up with the pace of digital innovation the case study assignment is structured around processes of problematisation: How did we get here? What is this a case of? These discussions ensure that the students thoroughly understand the kind of ethical issue or challenge that best fits the course criteria. It also motivates students to engage deeper with the

that best fits the course criteria. It also motivates students to engage deeper with the foundational principles of the course and appropriate literature. This learning cycle helps you to pace the development of your work with skills that are fresh in your mind.

The summative assessment builds on the formative assignment (stakeholder analysis) and the case study exercise. Using examples from different case study research on digital innovation studied on the module and researched for your essay, you will illustrate an ethical issue or management challenge relating to contemporary technologies in a specific workplace. You will be advised to check your essay titles with the module leader in advance.

The relationship between the formative and summative assignments is designed to reflect the research process of grounded theorists, a systematic methodology used by qualitative researchers. An empirical study is undertaken to establish a corpus of systematically data gathered from multiple, carefully cross-referenced sources with an emphasis on original data (rather than secondary). Where grounded theory researchers in the field compose theoretical and fieldwork memos during their longitudinal studies, you will write up and reflect upon your formative mini-case study. Having practiced this approach to handling and analysing case study data, you then weigh up and assess similar (but different) data as you illustrate your line of argument for the summative assignment.

The focus of the research essay assignment will be a critical discussion of leadership, corporate governance and business ethics relating to the management of contemporary emerging technologies.

The research essay (2500 words) is submitted at the end of the module is worth 50% of total mark for the module. The period of time between the formative and summative assignments is intended to mirror the grounded theory process, i.e. it provides time for reflection and distillation so that insights can be produced from systematic handling of data. You will spend time studying case studies on the module and searching for further cases of relevance to your chosen essay topic. The class discussions and formative

You will be expected to show mastery of the concepts, marshal the literature, and produce a well-structured line of argument. You must show that you have learned the difference between opinion and the use of data to support a line of argument i.e., how to be data-informed. We are also hoping to see capacity for critical thinking about how to manage ethical issues surrounding AI in your future roles as business leaders.

Feedback

As per the assessment strategy above, formative and summative feedback will be provided in a timely way. Final grades will be given within two weeks of submission. Tutorial meetings to guide you through the assignments that will be available throughout the module- you must book at least two 1:1 tutorials with Prof Scott during the preparation of your assignments.

Reading list

Indicative reading list: Ananny, M. and Crawford, K. 2018. Seeing without

knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973-989.

Coeckelbergh, M. 2020. Artificial intelligence, Responsibility attribution, and a relational justification of explainability. *Science and Engineering Ethics*, 26: 2051-2068.

Coeckelbergh, M. 2022. The Ubuntu robot: Towards a relational conceptual framework for intercultural robotics. *Science and Engineering Ethics*. 28:16.

Cybulski, J.L. and Scheepers, R. (2021). Data science in organizations: Conceptualizing its breakthroughs and blind spots. *Journal of Information Technology*, 36(2), 154-175.

Floridi, L. 2023. *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.

Mitchell, M. 2019. *Artificial Intelligence: A Guide for Thinking Humans*, Farrar.

Mittelstadt, B. Principles alone cannot guarantee ethical AI. *Nat. Mach. Intell.* 1, 501 (2019).

Noble, S.U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

Rai, A., Constantinides, P. & Sarker, S. Next-generation digital platforms: Toward human–AI hybrids. *MIS Q.* 43, iii–x (2019).

Stahl, B. C. 2022. Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence. *Int. J. Inf. Manage.* 62, 102441. Stahl, B. C., Rodrigues, R., Santiago, N. & Macnish, K. 2022. A European agency for AI: Protecting fundamental rights and ethical values. *Comput. Law Secur. Rev.* 45.

Stilgoe, J., Owen, R., and P. Macnaghten (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568–1580.

Vesa, M. and Tienari, J. 2022. Artificial Intelligence and rationalized unaccountability: Ideology of the elites? *Organization*, 29(6), 113-1145.

Quality assurance

Date of first approval	
Date of last revision	June 2026
Date of this approval	

Notes/ comments

Assessment details

Grading method	Numeric
----------------	---------

Pass mark

50%

Assessments

[illegible]

100%