IE HUMANITIES

IE University
Professor: **THEODORE LECHTERMAN**

E-mail: tlechterman@faculty.ie.edu

Academic year: 22-23
Degree course: FIRST
Semester: 2º
Category: COMPULSORY
Number of credits: 6.0
Language: English

PREREQUISITES

None.

SUBJECT DESCRIPTION

ETHICS OF EMERGING TECHNOLOGIES

Artificial intelligence, blockchain, self-driving cars, autonomous weapons, and other emerging technologies offer incredible potential for solving problems and making our lives better. Their many promises include accelerating scientific discovery, curing diseases, improving efficiency, and widening access to opportunities. But these technologies also raise numerous questions about the values they promote and the risks they impose. Emerging technologies can threaten human safety, destroy the environment, and create or reinforce patterns of discrimination. They can upend cherished traditions and ways of living and empower some people and belief systems over others. What are the potential benefits and costs of different innovations and applications? How should value trade-offs be resolved and how should opportunities and impacts be fairly distributed? And who should decide and enforce the answers to these questions? This course will acquaint students with several forms of emerging technology and the ethical debates that surround them. These debates include concerns about (e.g.) privacy, safety, fairness, explainability, wellbeing, sustainability, democracy—and disagreements about how these values should be defined, weighed, and applied. Students will leave the course equipped to defend convincing positions on these questions. And they will leave with general tools of philosophical reasoning that can be applied to ethical questions in other domains.

IE IMPACT

IE IMPACT-HUMANITIES is the first Pillar course in the IE Impact Learning Journey.

IE IMPACT is a multi-bachelor, multi-school mandatory academic program for all IEU students whose mission is to foster the skills, mindsets, and knowledge we at IE University believe our future graduates need to develop into leaders of positive change.

IE IMPACT includes three "pillar courses": the Humanities, Technology and Entrepreneurship. It culminates in a fourth, hands-on course called the IE Challenge, whose design embodies IE's value of Diversity and its commitment to Sustainability.

The IE Impact Learning Journey aims to help IEU students transform into leaders of positive change by developing:

- humanistic approaches to interpersonal relations, decision-making and critical thinking;
- familiarity with the technologies that are applied to solve some of the world's greatest challenges, as well as understanding the implications of applying such technologies;
- entrepreneurial mindsets, know-how, and skills to identify and solve problems worth solving, and to create validated business models that enable solutions to be enduring and scalable;
- a curiosity and an awareness about how sustainability-driven enterprises are sources of positive change; and
- adaptive leadership skills, as well as personal and team-based skills to foster innovation through diversity.

OBJECTIVES AND SKILLS

- Development of analytical, critical, and creative thinking.
- Comprehensive reading of complex texts.
- Writing and argumentation.
- Research and information competence.
- Sensitivity to other cultures and human experiences.
- Familiarity with concepts and approaches in analytic philosophy.
- Detailed understanding of technological phenomena and their social implications.
- Reassessment of one's own predispositions and conventional wisdom.
- Improved skills of group deliberation.
- Ability to articulate and defend normative positions on the design, operation, and governance of technology.

METHODOLOGY

The course will be taught employing IE's Liquid Learning methodology, which combines three essential elements for a complete and dynamic learning experience: synchronous interactions, asynchronous interactions and individual inquiry and discovery. Synchronous Interaction is learning that happens in live, in real-time. Asynchronous interaction and individual inquiry and discovery are learning experiences that happen interactively and asynchronously using collaboration tools and digital platforms.

Students are required to read the assigned materials for each session in advance. Asynchronous sessions will consist of a multimedia viewing (e.g., film), event, and/or activity. In-person sessions will consist of an interactive lecture, to be followed in many cases by a discussion and/or group exercise.

Teaching methodology	Weighting	Estimated time a student should dedicate to prepare for and participate in
Lectures	23.33 %	35 hours
Discussions	20.0 %	30 hours
Exercises	23.33 %	35 hours
Group work	6.67 %	10 hours
Other individual studying	26.67 %	40 hours
TOTAL	100.0 %	150 hours

PROGRAM

SESSIONS 1 - 2 (LIVE IN-PERSON)

Introduction: Technology and ethical conflict

Examples of technology raising ethical questions. Review of the syllabus and expectations. Discussion of students' backgrounds, interests, and prior experience with the topics of the course.

Readings:

Book Chapters: Stephanie Hare, Technology Is Not Neutral, selections (See Bibliography)

SESSIONS 3 - 4 (LIVE IN-PERSON)

What is ethics? A crash course in ethical theory

How do you understand "ethics," "morality" and related concepts? How do different research disciplines treat these concepts? What are the differences between ethics, law, and social norms? What are the different levels of analysis in ethical theory? What are some of the main theoretical approaches?

Readings:

Book Chapters: James Rachels, Elements of Moral Philosophy, selections (See Bibliography)

Multimedia Documentation: Helena de Bres, "The Pink Guide to Philosophy

Multimedia Documentation: Jim Pryor, "Guidelines on Writing a Philosophy Paper,(Optional)

Other / Complementary Documentation: Douglas Portmore, "Tips on Writing a Philosophy Paper(Optional)

SESSION 5 (ASYNCHRONOUS)

Introduction to AI ethics

Video: Viewing of the film Coded Bias (Netflix)

SESSIONS 6 - 7 (LIVE IN-PERSON)

Artificial (narrow) intelligence, bias, and explainability

What is artificial intelligence? What are some of the major current problems in AI ethics?

Readings:

Article: Safiya Umoja Noble, "Missed Connections: What Search Engines Say About Women", (Bitch 54 (2012): 36-41)

Book Chapters: Reid Blackman, Ethical Machines, selections (See Bibliography)

SESSIONS 8 - 9 (LIVE IN-PERSON)

Autonomous vehicles and transportation justice

How should autonomous vehicles navigate moral dilemmas? What are the social impacts of introducing autonomous vehicles, and how should they be managed?

Readings:

- Participate in the Moral Machine Experiment: https://www.moralmachine.net/.
- J.-F. Bonnefon, A. Shariff, and I. Rahwan. "The Social Dilemma of Autonomous Vehicles." *Science* 352, no. 6293 (June 24, 2016): 1573–76. https://doi.org/10.1126/science.aaf2654.
- J. Himmelreich, "Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations." *Ethical Theory & Moral Practice* 21, 669–684 (2018). https://doi.org/10.1007/s10677-018-9896-4

SESSION 10 (ASYNCHRONOUS)

Introduction to AGI

1. View the following clip from 2001: A Space Odyssey (Kubrik, 1968)
2. Essay preparation.

Video: View the following clip from 2001: A Space Odyssey (Kubrik, 1968) (Youtube)

Video: View the whole film: 2001: A Space Odyssey (Kubrik, 1968).(Optional)

SESSIONS 11 - 12 (LIVE IN-PERSON)

Artificial (general) intelligence and value alignment

Is AGI a threat to humanity? What does it mean for advanced forms of AI to be aligned with human values?

Readings:

Book Chapters: Stuart Russell, Human Compatible (London: Allen Lane, 2019), chs. 5-6.

Book Chapters: Wendell Wallach and Shannon Vallor, "Moral machines." In S. Matthiew Liao (ed.),
The Ethics of Artificial Intelligence (Oxford: Oxford University Press, 2020): 383-412. (See
Bibliography)

Article: Iason Gabriel, "Artificial Intelligence, Values, and Alignment," (Minds and Machines 30, no. 3
(2020): 411–37) (ced) (Optional)

SESSIONS 13 - 14 (LIVE IN-PERSON)

Surveillance and privacy

What is privacy and why is it valuable? How should we weigh the value of privacy against security and convenience?

Readings:

Book Chapters: Carissa Véliz, Privacy is Power, selections (See Bibliography)

SESSION 15 (LIVE IN-PERSON)

Midterm essay due.

SESSIONS 16 - 17 (LIVE IN-PERSON)

Social media, truth, and violence

What are the benefits and costs of social media for individuals and society? How should social media be regulated?

Readings:

Book Chapters: Joshua Cohen and Archon Fung, "Democracy and the Digital Public Sphere," In
Digital Technology and Democratic Theory, edited by Lucy Bernholz, Hélène Landemore, and Rob
Reich, 23–61. Chicago: University of Chicago Press, 2021.

SESSIONS 18 - 19 (LIVE IN-PERSON)

Governing technology

What role should businesses and engineers play in the governance of emerging technologies? What role should the public play?

05th October 2022

Readings:

Book Chapters: Rob Reich et al., System Error, selections (See Bibliography)

Article: Annette Zimmermann et al., "Technology Can't Fix Algorithmic Injustice," (Boston Review, 9 Jan. 2020) (Optional)

Article: Theodore Lechterman et al., "The Ethics of Self-Regulation for High-Risk Technology" (TBC)(Optional)

Article: Iyad Rahwan et al., "Society-in-the-Loop: Programming the Algorithmic Social Contract." (Ethics and Information Technology 20, no. 1 (March 2018): 5–14) (ced) (Optional)

SESSION 20 (ASYNCHRONOUS)

Making Technology Safe for Democracy: A Conversation with Rob Reich and Emma Saunders-Hastings

Attend the event with Prof. Reich and Prof. Saunders-Hastings.

SESSIONS 21 - 22 (LIVE IN-PERSON)

Unpiloted and autonomous weapons

Under what conditions, if any, should unpiloted and autonomous weapons replace conventional weapons?

Readings:

Book Chapters: Peter Asaro, "Autonomous Weapons and the Ethics of Artificial Intelligence," in S. Matthew Liao (ed.), The Ethics of Artificial Intelligence (Oxford: Oxford University Press, 2020). (See Bibliography)

Article: Ryan Jenkins, "The Ethics of Killer Robots, (E-International Relations, Jul 23 2014) (Optional)

Article: Bradley J. Strawser, "Moral Predators: The Duty to Employ Uninhabited Aerial Vehicles," (Journal of Military Ethics 9, no. 4 (December 2010): 342–68) (Optional)

Article: Linda Eggert, "The Human Factor" (TBC)(Optional)

SESSIONS 23 - 24 (LIVE IN-PERSON)

Automation and the future of work

Is automation something to be welcome or resisted? What do societies owe to workers whose jobs are lost to robots?

Readings:

Working Paper: Joshua Cohen, "Good Jobs"

Book Chapters: Elizabeth Anderson, Private Government, selections (See Bibliography) (Optional)

Article: Branko Milanovic, "Three Fallacies that Make You Fear a Robot Economy," (Evonomics, Sept 1., 2016) (Optional)

Article: John Danaher, "Will life be worth living in a world without work? Technological unemployment and the meaning of life," (Science and Engineering Ethics 23, no. 1 (2017): 41-64.) (Optional)

SESSION 25 (ASYNCHRONOUS)

Work on final essay

SESSIONS 26 - 27 (LIVE IN-PERSON)

Automating government

How should technology be used to enhance democracy? Should we replace democracy with algocracy?

Readings:

Book Chapters: Jamie Susskind, Future Politics, selections

SESSIONS 28 - 29 (LIVE IN-PERSON)

Blockchain and decentralized power

What are the social benefits and costs of blockchain for various applications? Should we seek to replace governments with blockchains?

Readings:

Article: Marcella Atzori, "Blockchain Technology and Decentralized Governance: Is the State Still Necessary?" (Journal of Governance and Regulation 6, no. 1 (2017): 45–62)
Book Chapters: Chiara Cordelli, The Privatized State, selections (See Bibliography)

SESSION 30 (LIVE IN-PERSON)

Final session

1. Review, conclusion, and celebration of our work together.
2. Final essay due.

BIBLIOGRAPHY

Recommended

- Stephanie Hare. *Technology Is Not Neutral: A Short Guide to Technology Ethics.* ISBN 9781907994999 (Digital)

- James Rachels. *Elements of Moral Philosophy.* ISBN 0078038243 (Digital)

- Reid Blackman. *Ethical Machines.* ISBN 1647822815 (Digital)

- S. Matthew Liao (ed.). *The Ethics of Artificial Intelligence.* ISBN 9780190905071 (Digital)

- Carissa Véliz. *Privacy Is Power: Why and How You Should Take Back Control of Your Data.* ISBN 1473583535 (Digital)

- Elizabeth Anderson. *Private Government: How Employers Rule Our Lives (and Why We Don't Talk about It).* ISBN 9781400887781 (Digital)

- Rob Reich, Jeremy Weinstein, and Mehran Sahami. *System Error: Where Big Tech Went Wrong and How We Can Reboot.* ISBN 1529356741 (Digital)

- Jamie Susskind. *Future Politics: Living Together in a World Transformed by Tech.* ISBN 9780192559494 (Digital)

- Chiara Cordelli. *The Privatized State.* ISBN 0691211736 (Digital)

EVALUATION CRITERIA

Preparation and participation (20%). Students are expected to contribute actively and thoughtfully to the class activities and discussions. This requires careful reading of all required materials in advance of each session. It also requires cultivating habits of responsible deliberation. We will review preparation strategies and participation guidelines in the first session.

Midterm evaluation (20%). Students will write a short essay of 1000 or more words on a topic chosen from a list provided. The essay will not require external research. We will practice essay-writing strategies throughout the course and establish progress milestones to help manage the writing process. The essays will be graded against a rubric of criteria. Penalties for late submissions will be applied unless the student can demonstrate to the professor's satisfaction that the circumstances were unforeseeable and unavoidable.

Final evaluation (30%). Students will write a final essay of 1250 or more words on a topic chosen from a list provided. The essay will not require external research. We will practice essay-writing strategies throughout the course and establish progress milestones to help manage the writing process. The essays will be graded against a rubric of criteria. Penalties for late submissions will be applied unless the student can demonstrate to the professor's satisfaction that the circumstances were unforeseeable and unavoidable.

Continuous assessment (30%). Students will write four short essays of 500-750 or more words each throughout the course. These essays will serve to confirm engagement with the course readings, practice the skills of expository writing, and prepare students for the midterm and final essays. The essays will be graded against a rubric of criteria. Penalties for late submissions will be applied unless the student can demonstrate to the professor's satisfaction that the circumstances were unforeseeable and unavoidable. The short essay with the lowest grade will not count towards the final grade.

Criteria	Percentage	Comments
Preparation and participation	20 %	
Midterm evaluation	20 %	
Final evaluation	30 %	
Continuous assessment	30 %	

PROFESSOR BIO

Professor: **THEODORE LECHTERMAN**

E-mail: tlechterman@faculty.ie.edu

PROF. THEODORE LECHTERMAN
Email: tlechterman@faculty.ie.edu

Dr. Theodore Lechterman is Assistant Professor of Philosophy at IE University, having recently joined IE from the Institute for Ethics in AI at the University of Oxford. Lechterman's research addresses questions in applied ethics through the lens of political philosophy. Recent topics of publication include the concept of accountability in AI ethics and the ethics of combatting disinformation with AI. Current projects include the ethics of self-regulation of weaponizable technology and the ethics of democracy-enhancing robots. Lechterman is one of the organizers of the AAAI/ACM Conference on AI, Ethics, and Society, the world's premier interdisciplinary research conference on AI ethics. In addition to his work on AI, Lechterman has made significant contributions to business ethics and the ethics of philanthropy. His first book, *The Tyranny of Generosity: Why Philanthropy Corrupts Our Politics and How We Can Fix It*, was published in 2021 by Oxford University Press. A connecting theme across his work is how the value of democracy applies when power is exercised by private and artificial agents. Lechterman frequently contributes to public debates and advises organizational leaders on frontier ethical challenges in business, governance, and technology. He holds degrees from Harvard and Princeton and completed postdoctoral fellowships at Stanford, Goethe University, and the Hertie School.

OTHER INFORMATION

a)   ATTENDANCE POLICY

Attendance to this Humanities course is mandatory. Just like any other IEU academic course, all students must follow IEU's Attendance Policy.

As per IE University policy, every student must attend at least 70% of sessions. IEU attendance policy applies to any type of session: live in-person and asynchronous sessions. Students attending less than 70% of sessions will be graded with a FAIL for the course. This FAIL will apply to the ordinary and the extraordinary calls of the current academic year.
?For live in-person sessions, students will be marked absent if:

- they do not attend the session in-person or in any mode;

- they attend an in-person session via Zoom, or any remote format;

- they attend virtually a live in-person session, for which they received an authorized attendance waiver to attend the session online, but they do not have their camera turned on, they are not engaged, and they are not seated without distraction during the entire class session.

For asynchronous sessions, students will be marked absent if:

- they do not engage and complete the work expected of them in the asynchronous sessions.

Students are advised to take into account that the great majority of the Humanities sessions are double sessions; therefore, if a student is absent on a day there are 2 sessions scheduled, the student will be marked absent for the 2 corresponding sessions.

Attendance waivers
Please remember that professors are not allowed to grant exceptions to the attendance policy. You may inform them of the reasons of any planned or past absence out of courtesy, but any waiver must be granted by the IE Impact Program Management.

Under exceptional circumstances (serious health problems, visa delays, and travel restrictions), a student can ask for a temporary attendance waiver and may be allowed to attend online (or not in any format) and not have his/her 70% attendance affected. For this to be possible, a written and documented request must be made in advance to the IE Impact/Humanities Program Management Team. Students must email ieimpacthumanities@ie.edu with the request and any related documentation. Last minute changes in attendance or verbal request will not be considered. If the request is approved, an attendance waiver will be granted over one or more sessions and the allowed 30% will only be applied to other absences (without an authorized waiver).

Rules for online attendance (when authorized)
Any student who receives (or expects to receive) authorization for an attendance waiver to attend a live in-person session online (via Zoom) must meet the following criteria to be marked present: must have his/her camera on at ALL times, must be actively engaged during the entire class session and with any and all team-work, and must be located in an appropriate learning environment (e.g. student must be seated and free of distractions for the entire session).

Use of the 30% absences
No sessions or activities may be "skipped". If a student is under the impression that s/he can skip up to 30% of sessions, this is a mistake. The 30% "rule" is reserved for those circumstances that are out of a student's control. Personal trips, interviews, appointments, mild illnesses, family celebrations or ceremonies, or other personal matters will not be treated as exceptional cases and should be considered part of the 30% of allowed absences.

Attending online but still absent
Students who do not have an authorized attendance waiver are allowed to attend a class session online; however they must understand that they will be marked absent, and their participation grade may be influenced negatively by a high percentage of absences during the semester.

Role of professors in attendance
Professors are responsible for recording attendance; however, they have no authority to grant any student permission to have an excused absence of any kind (or to change the format of how a student will attend a session).

Any inquiries or requests must be made by the student to the IE Impact Program Team via email ieimpacthumanities@ie.edu

Professors are instructed to mark as ABSENT:
- students who do not attend a session in its previously scheduled format
- students who attend a session in a different format (online) than what was previously scheduled (i.e. they attend online instead of in-person).
- students who attend a session they were previously authorized to attend online but do not have their cameras on, engage actively or stay seated in an environment free of distractions for the entire session.
- students who do not complete the work adequately assigned for an asynchronous session.
- If a student receives authorization for an attendance waiver and the student was already marked absent, the professor will be informed and the attendance record will be modified from

absent to whatever is authorized (present, excused, etc).

Authorization for Attendance Waiver from a Student's Bachelor Program

If a student has received authorization for an attendance waiver from his/her bachelor degree's program to be absent (from any format of class) or to attend a live in-person session online (via Zoom), then the student must forward the email authorization to the IE Impact Team (ieimpacthumanities@ie.edu) and it will be verified. Upon verification, the professor will be informed to either not record a student as absent or to change a previous attendance record from absent to present, excused, etc.

b)   GRADING SYSTEM

Grade description and equivalents:

Excellent/Sobresaliente: 9.0-10.0 (A- to A+)
Consistently produces work of the highest quality and craft; exhibits notable progress and development over the course of the semester; meets all course objectives at highest level; attendance is near-perfect, and contributions to course discussions are extremely valuable.

Very Good/Notable: 7.0-8.9 (B- to B+)
Completes all assignments with work of above-average quality and craft; exhibits significant progress and development; meets most course objectives; attendance and participation are very good.

Good/Aprobado: 6.0-7.0 (C- to C+)
Completes all assignments with work of acceptable quality and craft; exhibits some progress and development; meets a majority of course objectives. Attendance and participation are acceptable.

Pass/Aprobado: 5.0-6.0 (D)
Assignments are delivered but are incomplete and/or of low quality and craft; exhibits little progress and development; meets few course objectives. Attendance and participation are poor, but absences do not total more than 30%.

Fail/Suspenso: 0-4.9 (F)
Work is incomplete, missing, or does not meet course objectives. Attendance and participation are poor.

Automatic Failure/Suspenso: 0 (F)
Please note that a student who misses 30% or more of the scheduled sessions receives an automatic 0.0, and loses his or her right to the second chance or call (convocatoria.)

c)   RETAKE POLICY

Students have four opportunities to pass a course distributed in two consecutive academic years. It is mandatory to attend 100% of the classes, but if justified, students can miss up to 30% of the classes. If they miss over 30%, they will have to enroll again in the course the following year.

- Students who do not comply with the 70% attendance rule will lose their 1st and 2nd chance, and go directly to the 3rd one (they will need to enroll again in this course the next academic year).

- Students who fail the subject in the first regular period, will have to retake it in July. The dates and locations will be posted in advance by the university informing both professors and students.

- The maximum grade a student can obtain in the second exam session is 8 out of 10.

Students retaking this course will normally be required to retake the midterm exam and (re)submit any failed or outstanding essays during the retake period. Modifications to this policy will be considered on a case-by-case basis.

d) CODE OF ETHICS/HONOR

Cheating and plagiarism are very serious offenses governed by the IE student code of conduct. Any student found cheating or plagiarizing on any assignment or component of this course will at a minimum receive a "0" on the affected assignment. Moreover, the student will also be referred to the University Judicial System for further action. Additional penalties could include a note on your transcript, failing the class, or expulsion from the university.

What is academic integrity? One component of a definition is when one does the right thing even though no one is watching. The core values of integrity, both academic and otherwise include: honesty, fairness, respect, responsibility, and trust. Academic integrity requires that all students within Instituto de Empresa (IE) act in accordance with these values in the conduct of their academic work, and that they follow the rules and regulations concerning the accepted conduct, practices and procedures of academic research and writing. Academic integrity violations are defined as cheating, plagiarism or other violations of academic ethics.

Cheating includes:

a) An act or attempt to give, receive, share, or utilize unauthorized information or unauthorized assistance at any time for assignments, papers, projects, presentations, tests or examinations. Students are permitted to mentor and/or assist other students with assignments by providing insight and/or advice. However, students must not allow other students to copy their work, nor will students be permitted to copy the work of other students. Students must acknowledge when they have received assistance from others.
b) Failure to follow rules on assignments, papers, projects, presentations, tests or examinations as provided by the course professor and/or as stipulated by IE.
c) co-operation or collaboration.
d) with official documents, including electronic records.
e) The impersonation of a student on presentations, exercises, tests or an examination. This includes logging onto any electronic course management tool or program (e.g. Black Board, etc.) using someone else's login and password.

Plagiarism includes:

a) Using the work of others and attempting to present it as your own. For example, using phrases or passages from books, articles, newspapers, or the internet and not referencing them properly in your document is a form of plagiarism. This includes using information from others without citing it, misrepresentation of cited work, and misuse of quotation marks.
b) Submitting an assignment or paper that is highly similar to what someone else has written (i.e., minimal changes in wording, or where the sentences are similar, but in a different order).
c) Plagiarizing is not committing "word for word" copying. "Thought for thought" is also a form of plagiarism.
Other violations of academic ethics include:
a) Not acknowledging that the students' work or any part thereof has been submitted for credit elsewhere.
b) Misleading or false statements regarding work completed.
c) Knowingly aiding or abetting anyone in committing any form of an academic integrity violation.

Academic Misconduct Procedure for Humanities Courses

1. If a Humanities instructor suspects a student has committed academic misconduct (as defined in the student's Code of Ethics) in a Humanities course, he or she must refer the case to the Humanities program director with all the necessary supporting evidence.

2. The Humanities program director will meet with the student and write a brief summary of the instructor and the student's views and pass them on to the Bachelor's program director. The Humanities program director may also suggest what type of sanction would be appropriate for the student.

3. If there is enough objective evidence to sanction the student, the program director will check if this is the first time the student has committed academic misconduct.

4. Is this is a first breach of conduct, the program director will decide the scope of the sanction on the basis of all the above information. (Bachelor or Ethics Committee).

5. If this is a second offense or if, according to the Humanities director's report, the case is serious enough, the case will be redirected to the university's Ethics Committee.

(Approved by the Vice-Rector of Academic Affairs in June 2020)