

Exercises

1	2
---	---

Surname, First name

5ARC0 Human and ethical aspects of AI
Final exam Q1

1	1	1	1	1	1	1
2	2	2	2	2	2	2
3	3	3	3	3	3	3
4	4	4	4	4	4	4
5	5	5	5	5	5	5
6	6	6	6	6	6	6
7	7	7	7	7	7	7
8	8	8	8	8	8	8
9	9	9	9	9	9	9
0	0	0	0	0	0	0

a	<input checked="" type="checkbox"/>	c	d	e	f	→ b
a	b	<input checked="" type="checkbox"/>	d	e	f	→ c
<input checked="" type="checkbox"/>	b	c	<input checked="" type="checkbox"/>	e	f	→ a

a	<input checked="" type="checkbox"/>	c	<input checked="" type="checkbox"/>	e	f	→ b, d
a	b	<input checked="" type="checkbox"/>	d	e	f	→ c
<input checked="" type="checkbox"/>	b	c	<input checked="" type="checkbox"/>	e	f	→ a, d

Fill in your answer(s) to the multiple-choice questions as shown above (circles = one correct answer, boxes = multiple correct answers possible).

Particular Ans on paper exam instructions

- Write in a black or blue pen.
- You answer open-ended questions by using the text box. Provide your answers on the papers inside the answer box underneath a question. **If you need more space for your answers, use the extra space at the end of the exam, and clearly indicate there which question you continue answering. In the text box of the particular question, clearly state that you proceed with your answer on a different page.**
- Hand in all pages. Do not remove the staple. If you remove it anyhow, check that you hand in all pages.

Dear student,

You're about to take an exam. Write down your name and your student ID at the appropriate places above. Make sure that you enter your student ID by fully coloring the appropriate boxes. On the examination attendance card, you fill in the PDF number. You can find the correct number on the top of the first page of your exam (e.g. 1234.pdf).

Please read the following information carefully:

Date exam: 7 November 2023

Start time 09.00

End time: 12.00 (+30 minutes for time extension students)

Number of questions: 15

Maximum number of points/distribution of points over questions: Max 50 points

Method of determining the final grade: Actual score/ 5

Answering style: formulation, order, foundation of arguments, multiple choice:

Multiple choice + Short written answers

Permitted examination aids

- Scrap paper (fully blank)
- Dictionaries. If yes, please specify: English or English-Other Language

Important:

- You are only permitted to visit the toilets under supervision
- Examination scripts (fully completed examination paper, stating name, student number, etc.) must always be handed in
- The house rules must be observed during the examination
- The instructions of subject experts and invigilators must be followed
- Keep your work place as clean as possible: put pencil case and breadbox away, limit snacks and drinks
- You are not permitted to share examination aids or lend them to each other
- Do not communicate with any other person by any means

During written examinations, the following actions will in any case be deemed to constitute fraud or attempted fraud:

- using another person's proof of identity/campus card (student identity card)
- having a mobile telephone or any other type of media-carrying device on your desk or in your clothes
- using, or attempting to use, unauthorized resources and aids, such as the internet, a mobile telephone, smartwatch, smart glasses etc.
- having any paper at hand other than that provided by TU/e, unless stated otherwise
- copying (in any form)
- visiting the toilet (or going outside) without permission or supervision

First-year bachelor students: The final grade for this exam will be announced no later than fifteen working days after the date of this exam, unless this exam takes place in Q4 or the interim period. For Q4 final exams, grades will be announced within five working days after the end of the Q4 final test period. For interim period final exams, grades will be announced no later than five working days before September 1.

All other students: Generally, the final grade for this exam will be announced no later than fifteen working days after the date of this examination. Specifically for bachelor exams administered in the interim period, exam grades will be announced no later than five working days before September 1.

You can start the exam now, good luck!

Main Part

6p 1a Human rights

Here are two of the human rights mentioned in the Universal Declaration of Human Rights. (1) Please explain how **one** of these rights might be threatened by current, real-world applications of machine intelligence. (2) Give a plausible example (real or imaginary) to support the claim of a threat. (3) Rights correspond to duties, and there are normally people or institutions who *hold* the duty (and are therefore obligated to respect the right). Who is the main duty-holder in your example?

- Article 9: No one shall be subjected to arbitrary arrest, detention or exile.
- Article 11, Pt. 1: Everyone charged with a penal offence has the right to be presumed innocent until proved guilty according to law in a public trial at which he has had all the guarantees necessary for his defence.

5p 1b Four principles

Give a simple, deductively valid three-step argument with an ethical conclusion, using one of the “four principles” derived from the Belmont Report as a premise. You must state the principle as a full sentence. The argument should concern AI or robots, and the premises do not all need to be true.

2p 1c Goodall

Which approach to programming ethics into autonomous vehicles does Goodall (2014) recommend, and why?

- ☐ (a) Consequence-based ethics, because this will minimize the damage caused by crash situations.
- ☐ (b) Rule-based ethics alone, because this guarantees respect for the rules of the road such as speed limits.
- ☐ (c) Machine learning based on observations of human behavior, because this will ensure smooth integration in mixed traffic.
- ☐ (d) Rule-based ethics and machine learning based on human ethical judgments, because this allows gradual progress.

2p 1d HCAI

What examples does Schneiderman give for technology with high levels of human control together with high levels of automation (HCAI)? Fill in the following blanks:

6p 1e Human skills and work

Bankins & Formosa give the example of a "third way" of using artificial intelligence for automation, in which it amplifies human capacities. They say that it is different from the simple replacement of jobs with automation, and different from the situation where somebody's original tasks are replaced with the job of managing AI systems. In your opinion, does such an option help to avoid the ethical worries that Zoller raises about the loss of human skills? (1) Explain one of Zoller's arguments. (2) Explain whether the "third way" shows that Zoller's argument is unsound.

6p 1f Environmental impact of AI

Suppose that the impact of machine learning on the environment is very significant, along the lines of what Lin claims. (1) What are Lin's main policy proposals for dealing with this impact? (2) *Without denying that the impact is significant*, give two ideas for plausible counterarguments against Lin's policy proposals. You don't need to spell out the counterarguments in multiple premises.

4p 1g Self-driving cars

Moral responsibility is said to apply to humans but not to machines. What is a simple, valid argument that an autonomous vehicle cannot ever be morally responsible for an accident that it causes? The argument should be based on one or more of the necessary conditions for moral responsibility.

2p **1h** Explainability

Which of the following are basic sources of opacity in relation to AI and machine learning, according to Burrell? You may choose more than one answer.

- ☐ People who operate data science systems are secretive about how these systems work
- ☐ Stakeholders do not have the required understanding of how complex data-driven decisions are made
- ☐ Laws protecting privacy restrict how much information can be learned about particular data subjects
- ☐ People do not trust the explanations that are given to them by companies and engineers

2p **1i** Explainability

Which of the following best describes Zednik's position about solving the Black Box problem?

- ☐ a Different stakeholders have different needs for knowledge and explanation with regard to algorithmic decision-making.
- ☐ b The techniques developed in Explainable AI are likely to solve the Black Box problem for professional users of AI-based systems.
- ☐ c Education in digital literacy and data science is the best general solution to the Black Box problem in algorithmic decision-making.
- ☐ d The privacy of data subjects is the main problem that needs to be addressed by Explainable AI.

5p 1j Medical AI

Hatherley et al. suggest that there are several advantages to interpretable AI in a medical context. (1) What do they mean by interpretable AI? (2) What are three of the advantages mentioned by them, or in the lecture on this topic?

6p 1k AI in military applications

In your opinion, what is the least unethical strategy for automating the tasks of soldiers: a human-centered approach based on meaningful human control, or a technological autonomy approach? Briefly explain the difference between the approaches and give ethical reasons favoring your preferred approach. You should make correct use of at least one of the principles of just warfare in your answer.

2p

1l Generative AI

Which of the following are criteria on which credit for creative works depends according to Khosrowi et al.'s (2023) collective-centered creation (CCC) model of credit for artistic and literary works? You may select more than one answer.

- ☐ Benefit or utility of crediting a particular party
- ☐ Time and effort invested
- ☐ Copyright and intellectual property
- ☐ Originality

1p

1m Generative AI

True or false: According to the CCC model, a generative AI tool should sometimes receive credit for its role as a creator of an artistic work.

- ☐ (a) False ☐ (b) True

1p

1n Generative AI

True or false: According to the CCC model, credit for users of generative AI tools depends partly on the prompts they use as inputs for these tools.

- ☐ (a) False ☐ (b) True

1o Extra Credit: Singularity

In his article "The Singularity: A Philosophical Analysis" David Chalmers says that the most crucial assumption in his argument for the singularity is what he calls the ***proportionality thesis***. What does the proportionality thesis say? (Note: this question is worth two possible extra credit points.)

2 Please use this space if you need to correct or supplement an answer from the exam.

[illegible]

[illegible]