Additional practice exam questions (KEY)

1. **Question: Which of the following is NOT a source of opacity in machine learning systems according to Burrell (2016)?**

   A. Technical illiteracy
   B. Intentional secrecy
   C. System complexity
   D. Computational efficiency

Correct Answer: D

2. **Question: Zednik (2021) discusses stakeholder-dependent opacity in AI systems. Which of the following best describes this concept?**

   A. Opacity that arises from the technical complexity of AI algorithms.
   B. Opacity that varies based on a person's role, prior knowledge, and abilities.
   C. Opacity that is intentionally created by developers to protect proprietary information.
   D. Opacity that results from the lack of transparency in data collection methods.

Correct Answer: B

3. **Question: The guest lecture by Streppel on opacity in AI emphasized that one of the primary challenges in developing ethical AI systems is ensuring that the decision-making processes are _____ to users and stakeholders.**

Correct Answer: transparent, understandable

Possible Wrong Answers:

- **efficient, cost-effective**
- **secure, private**
- **innovative, scalable**
- **fast, reliable**

4. **Which of the following is a key argument made by Vallor regarding the impact of automation on moral skills?**

   - A) Automation inherently improves moral skills by reducing human error.
   - B) Automation can lead to moral deskilling, which is inherently bad.
   - C) Automation has no significant impact on moral skills.

o   D) Automation always leads to moral upskilling by enhancing human capabilities.

Answer: B)

5. **According to Sambasivan and Veeraraghavan, what is a common perception of field workers by AI model developers?**

   o   A) Field workers are seen as essential partners in data collection.

   o   B) Field workers are viewed as lazy, non-compliant, or corrupt.

   o   C) Field workers are considered a source of paid labor for AI models.

   o   D) Field workers are seen as underqualified for their tasks.

Answer: B)

Fill-in-the-Blank Question

6. **Vallor argues that technologies should be designed to help humans become _____, counteracting the influence of culture and capitalism.**

Answer: better people, virtuous

7. How do contemporary artistic practices use Generative AI as an **object of critical inquiry** to address societal, ethical, and cultural issues? Discuss with reference to specific examples and theoretical frameworks presented in the guest lecture by Kalidindi.

Outline of a Plausible Answer

Contemporary artistic practices critically engage with Generative AI by examining its societal, ethical, and cultural implications. Artists like Trevor Paglen and Kate Crawford use projects like "ImageNet Roulette" to highlight biases in machine learning datasets. Theoretical frameworks such as material agency and epistemic practices underscore how technologies shape artistic creation and knowledge production. By focusing on AI's limitations, biases, and ethical concerns, artists challenge the narratives of efficiency and productivity promoted by BigTech, advocating for more ethical and inclusive technological development. This critical engagement fosters a deeper understanding of AI's impact on society and culture.

8. **What is the Orthogonality Thesis as discussed in the lecture?**

   A. The idea that AI will always align with human values.

   B. The concept that intelligence and final goals are independent of each other.

   C. The theory that AI will inevitably become malevolent.

   D. The belief that AI cannot surpass human intelligence.

Correct Answer B

9. **According to the lecture, what is a key challenge in managing AI risks under deep uncertainty?**

    A. Predicting all possible outcomes accurately.

    B. Ensuring AI systems are always benevolent.

    C. Switching from a predict-then-act paradigm to a different paradigm.

    D. Eliminating all risks associated with AI.

Correct answer C

10. **Matching Question**

Match the following terms with their definitions:

1. **Singularity**
2. **Instrumental Convergence**
3. **Existential Risk**
4. **Precautionary Principle**

Definitions:

    A. The idea that intelligent agents with varied final goals will pursue similar intermediate goals.

    B. A theoretical point where AI surpasses human intelligence and becomes uncontrollable.

    C. Lack of scientific knowledge about consequences doesn't justify inaction.

    D. Threats that could lead to the extinction of humanity or drastic loss of potential for future development.

1B; 2A; 3D; 4C

11. AI in military contexts

Essay question: Within the framework of Meaningful Human Control, how would the principles of jus in bello be applied to the deployment of an AI-powered autonomous drone system in a military context? Discuss the ethical considerations and potential challenges.

Sample answer:

Within the framework of Meaningful Human Control (MHC), the ethical principles of jus in bello—proportionality, discriminacy, and military necessity—are "tracked" by the autonomous drone. Proportionality requires that the harm caused by drone strikes is not excessive relative to the military advantage gained. Discriminacy mandates that drones accurately distinguish between combatants and non-combatants to minimize civilian casualties. Military necessity justifies the use of drones only for achieving legitimate military objectives. The "tracing" condition of MHC implies ensuring human oversight, accountability for autonomous decisions, and addressing potential biases in AI algorithms. Challenges include maintaining control and responsibility in complex, fast-paced combat scenarios.

12. In his article "Logic and Conversation", Grice sets out several "Conversational Maxims" according to which we normally communicate in interpersonal conversations. In your opinion, which of the conversational maxims is most difficult for ChatGPT to comply with? Briefly motivate your opinion.

Sample answer The maxim of **Quantity** is most difficult for ChatGPT to comply with, as it often provides more information than necessary to ensure clarity and completeness, which can sometimes overwhelm or deviate from the user's specific query.

13. According to O'Neill and colleagues, which forms of self-improvement does any individual with nearly any set of moral values have a reason to pursue?

Answer: Avoidance of weakness of will, avoidance of biased decisions or decisions based on poor information, improvement of one's understanding of one's own value commitments.

14. Give a three-step, valid normative argument *against* the use of Artificial Ethics Assistants. The premises do not need to be true.

Sample answer:

Artificial Ethics Assistants, when used as intended, lead to basing one's moral judgments on one's past moral commitments, not one's present moral commitments.

One should always base one's moral judgments on one's present moral commitments.

Therefore, one should not use AEAs as intended.