

## Introduction into ethics

2a How would you define "ethics"?

Ethics is the branch of philosophy that helps us make “good” decisions, taking into consideration our values, principles, beliefs and norms. It is a tool that shapes our world by shaping the basis of our thinking.

2b What would you do in the situation with Lee? Explain your choice taking into consideration your norms, values, principles and beliefs.

In the situation with Lee, I would've taken a third approach. I would've taken matters into my own hands by making sure to invite him to a group or one-on-one outing, even if he most likely wouldn't have come. That way, he doesn't feel alienated, and I wouldn't have to divulge his personal mental health situation.

3a What does it take to earn trust in artificial intelligence?

There are 5 pillars when it comes to earning trust in an AI made by your organisation:

1. Fairness – ensuring that the AI model is fair to everyone.
2. Explainable – such that someone would be able to be told what datasets, methods, etc. used to train the AI model.
3. Robustness – nobody can hack such an AI model.
4. Transparency – does the public have access to information about the model?
5. Data privacy

3b What are the 3 principles when thinking about AI in an organization?

1. The purpose of AI is meant to be to augment human intelligence, not replace it.
2. Data and the insights of the data belong to the creator alone.
3. AI systems should be transparent and explainable.

3c How does the socio-technological challenge must be addressed?

With the general and internal public in mind.

3d What does holistically mean?

There are 3 things to think about: people (organisational culture), process (governance), and tooling (how to ensure the 5 pillars).