Text, logo

Description automatically generatedLogo

Description automatically generated with medium confidence

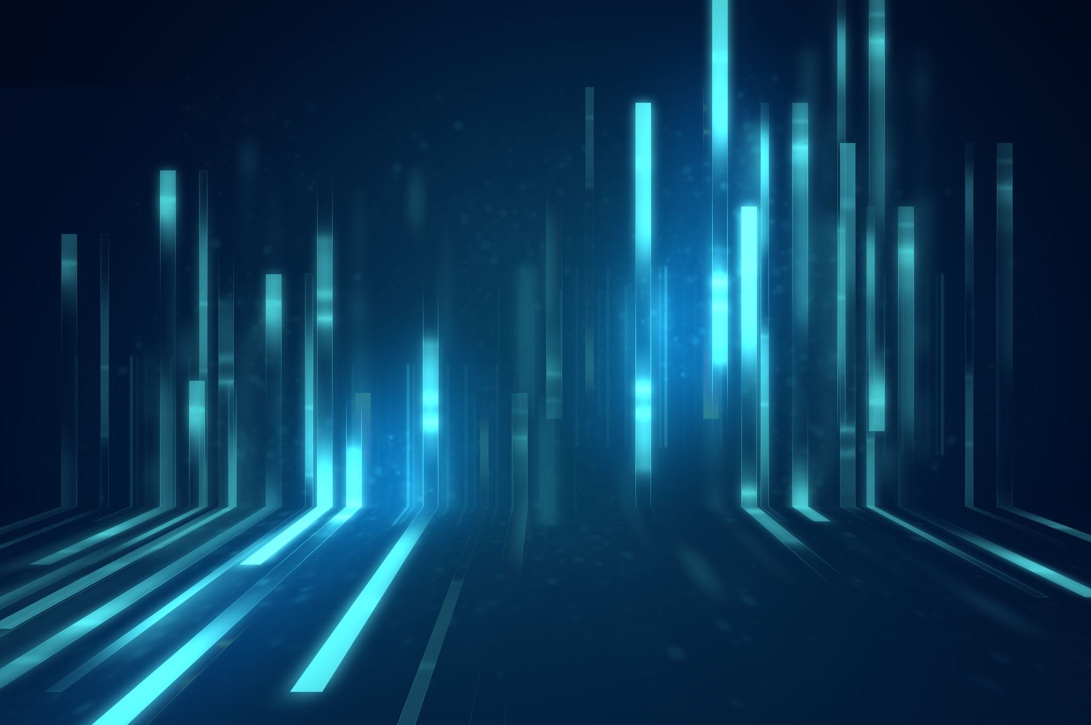
­­­

Fairness

Intermediate

Module 1

Responsible AI



# Introduction

This module will focus on a systematic approach to identify ethical and social impacts. A systematic approach is an important first step to explicate possible ethical concerns around fairness and transparency before attempting to address this concern. The module will explain the ethical impact assessment approach that we use in our company to identify ethical and social concerns (see (Wright, 2011).

­­­

# Reflection Sheet

* At what stages of the technology’s lifecycle can we aim to address ethical concerns?]
* What are the benefits of an EIA?
* Why would an EIA assessment be important when we are designing technology to be transparent and fair?

**Task:**

* Write a three paragraph pitch to persuade a project manager that an ethical impact assessment is important. Show that you understand the benefits of an EIA, how technology can impact ethical values, and what an EIA entails.

* Write another three paragraphs about your own use case:
  + Identify a technology who needs a EIA
  + Explain why they need this assessment
  + Explain if an EIA is still useful
  + Potentially identify some ethical issues

­­­



**Self-assessment pass/fail questions**

1) What was Collingridge contribution to technology assessment?

1. He stated that technology is inherently political;
2. He stated that technology change how we view society;
3. He created a method to explain how values are embedded within the technology;
4. He formulated a dilemma about when to address ethical concerns in emerging technology

2) What is acute technosocial opacity?

1. The inability to identify ethical impacts from new and emerging technology
2. The difficulty new technologies pose for people to shape their lives in meaningful ways
3. The inability to find technological fixes for global societal issues
4. The want of developers and tech companies to hide the effects of technology on society

3) Scenarios to identify and address ethical impacts of technology are predominantly associated with the following methods:

1. Ex-ante methods
2. Intra methods
3. Ex-post

4 ) An ethical impact assessment is which of the following methods?

1. Ex-ante methods
2. Intra methods
3. Ex-post

5) If you want to embed ethical values in the design of the technology, which type of method is most suitable?

1. Ex-ante methods
2. Intra methods
3. Ex-post

6) What is not a step of an ethical impact assessment?

1. Consider whether an ethical impact assessment is relevant
2. Identify stakeholder
3. Publish the report
4. All are relevant

7) What is not mentioned as a challenge of an ethical impact assessment?

1. Publishing a risk assessment
2. Finding experts
3. Identifying operational criteria
4. Participation of stakeholders

**Answers**

Qs 1) D, 2) B, 3) A, 4) B, 5) B, 6) D, 7) A Q Qs

# Pre-reading

Ledford, H. (2019). Millions of black people affected by racial bias in health-care algorithms. *Nature*.

Reijers, W., Wright, D., Brey, P., Weber, K., Rodrigues, R., O’Sullivan, D., & Gordijn, B. (2018). Methods for Practising Ethics in Research and Innovation: A Literature Review, Critical Analysis and Recommendations. *Science and Engineering Ethics*, *24*(5), 1437-1481. <https://doi.org/10.1007/s11948-017-9961-8>

Wright, D., & Friedewald, M. (2013). Integrating privacy and ethical impact assessments. *Science and Public Policy*, *40*(6), 755-766. <https://doi.org/10.1093/scipol/sct083>

Royakkers, L., Timmer, J., Kool, L. et al. Societal and ethical issues of digitization. *Ethics Inf Technol* 20, 127–142 (2018). https://doi.org/10.1007/s10676-018-9452-x

<https://news.harvard.edu/gazette/story/2020/10/ethical-concerns-mount-as-ai-takes-bigger-decision-making-role/>

# References

https://www.bloomberg.com/news/articles/2017-07-09/robert-moses-and-his-racist-parkway-explained

Ledford, H. (2019). Millions of black people affected by racial bias in health-care algorithms. *Nature*. <https://www.nature.com/articles/d41586-019-03228-6#ref-CR1>

Nuffield Council on, B. (2015). The collection, linking and use of data in biomedical research and health care: ethical issues. <https://apo.org.au/node/249651>

Reijers, W., Wright, D., Brey, P., Weber, K., Rodrigues, R., O’Sullivan, D., & Gordijn, B. (2018). Methods for Practising Ethics in Research and Innovation: A Literature

Review, Critical Analysis and Recommendations. *Science and Engineering Ethics*, *24*(5), 1437-1481. <https://doi.org/10.1007/s11948-017-9961-8>

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, *109*(1), 121-136. <http://www.jstor.org/stable/20024652>

Worthington, R. (1982). The Social Control of Technology. By David Collingridge. (New York: St. Martin's Press, 1980. Pp. i + 200. $22.50.). *American Political Science Review*, *76*(1), 134-135. <https://doi.org/10.2307/1960465>

Wright, D., & Friedewald, M. (2013). Integrating privacy and ethical impact assessments. *Science and Public Policy*, *40*(6), 755-766. <https://doi.org/10.1093/scipol/sct083>