THE UNIVERSITY OF MELBOURNE

COMP90087 overview

Tim Miller, Kobi Leins, Simon Coghlan, Marc Cheong

School of Computing and Information Systems
Centre for AI & Digital Ethics
The University of Melbourne

Subject overview

Who are we?

Simon

Tim

Kobi

Marc

Tutors

Kate

Matt

Michael

COMP90087 Learning outcomes

On completion of this subject, students should:

- Understand and describe the social and ethical issues raised by technology in various contexts
- Apply established ethical theories to reason about social and ethical issues related to technology
- Critically examine, analyse, and apply ethical perspectives to design, develop and deploy of digital technologies responsibly
- Communicate perspectives to other technical professionals and the community at large
- Demonstrate a profound respect for that fact that just because you can, does not mean you should create a technology

COMP90087 Assessment

Description	Timing	Percentage
Tutorial participation and contributions to online discussion. Two hours per week. Intended Learning Outcomes (ILO's) 1 to 4 are addressed by tutorial participation and contributions to online discussion. 2% each; best 10 of 12	Throughout the teaching period	20%
Critical evaluation essay. ILO's 1 to 5 are addressed in the essay. • 35-45 hours	Week 7	30%
Research essay. Due first week of examination period. ILO's 1 to 5 are addressed in the essay. • 35-45 hours	During the examination period	30%
Online quiz. ILO's 1 and 2 are addressed in the online quiz. • 1 hours	End of semester	20%

COMP90087 Agenda

One module per week

Each module:

- Readings (1-2 papers)
- Video playlist (30-60 minutes), with notes
- Consultation & discussion hour: 9-10am on Thursdays (on Wednesday in week 2)
- Tutorial
  - Group discussion on relevant topic

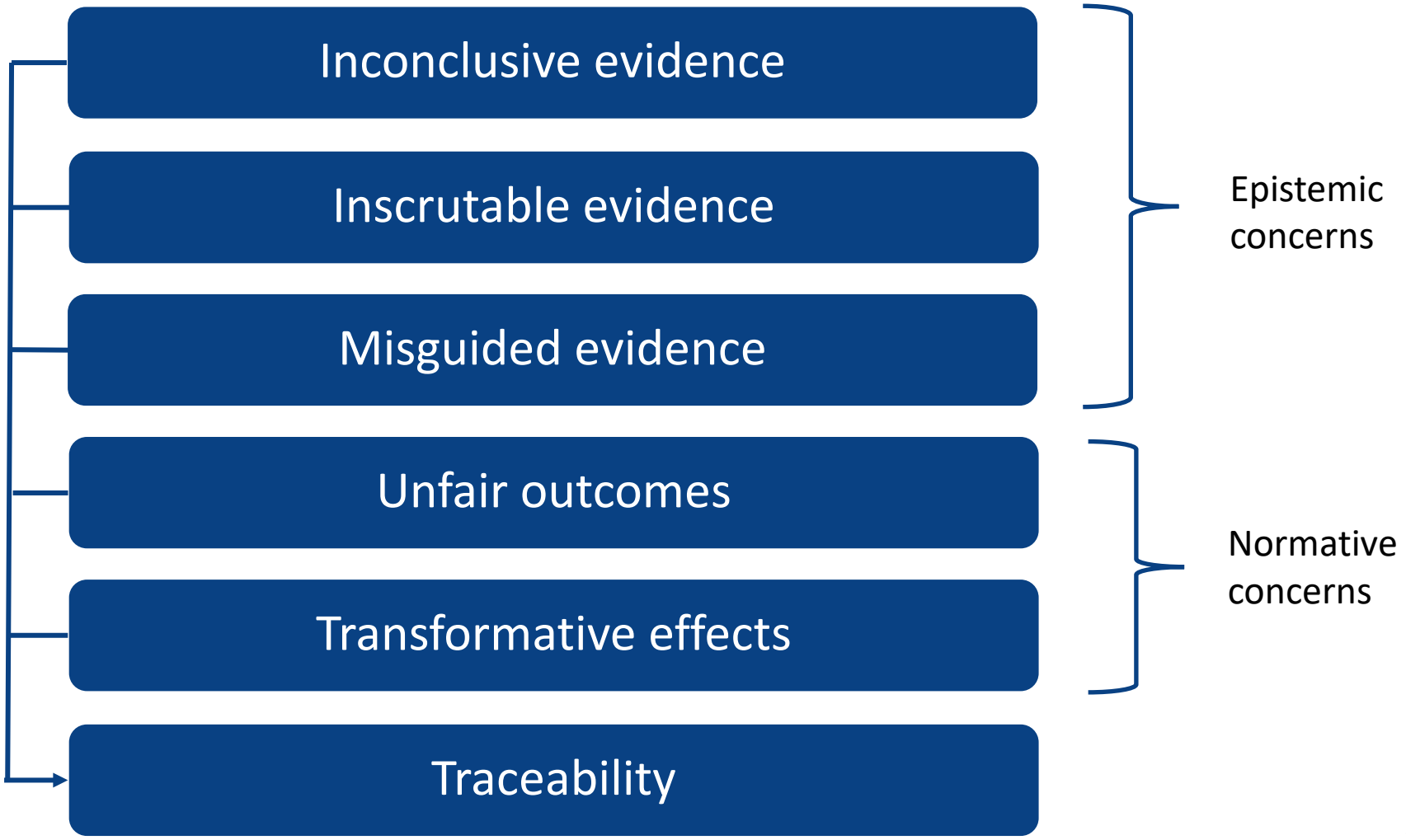
Interaction

When you need advice, input, direction:

- Check Canvas for announcements.
- Check the relevant discussion forum for existing answers.
- Post a question to the relevant discussion forum
  - Feel free to answer others' questions!
- Attend a consultation hour (Thu 9am Melbourne time).
- Ask your tutor during the tutorial.
- For personal matters, send an email:
  - Please do not email more than one staff member independently about the same issue. It duplicates work between us.
- We commit to responding within 1 working day
  - Please do not expect replies over the weekend

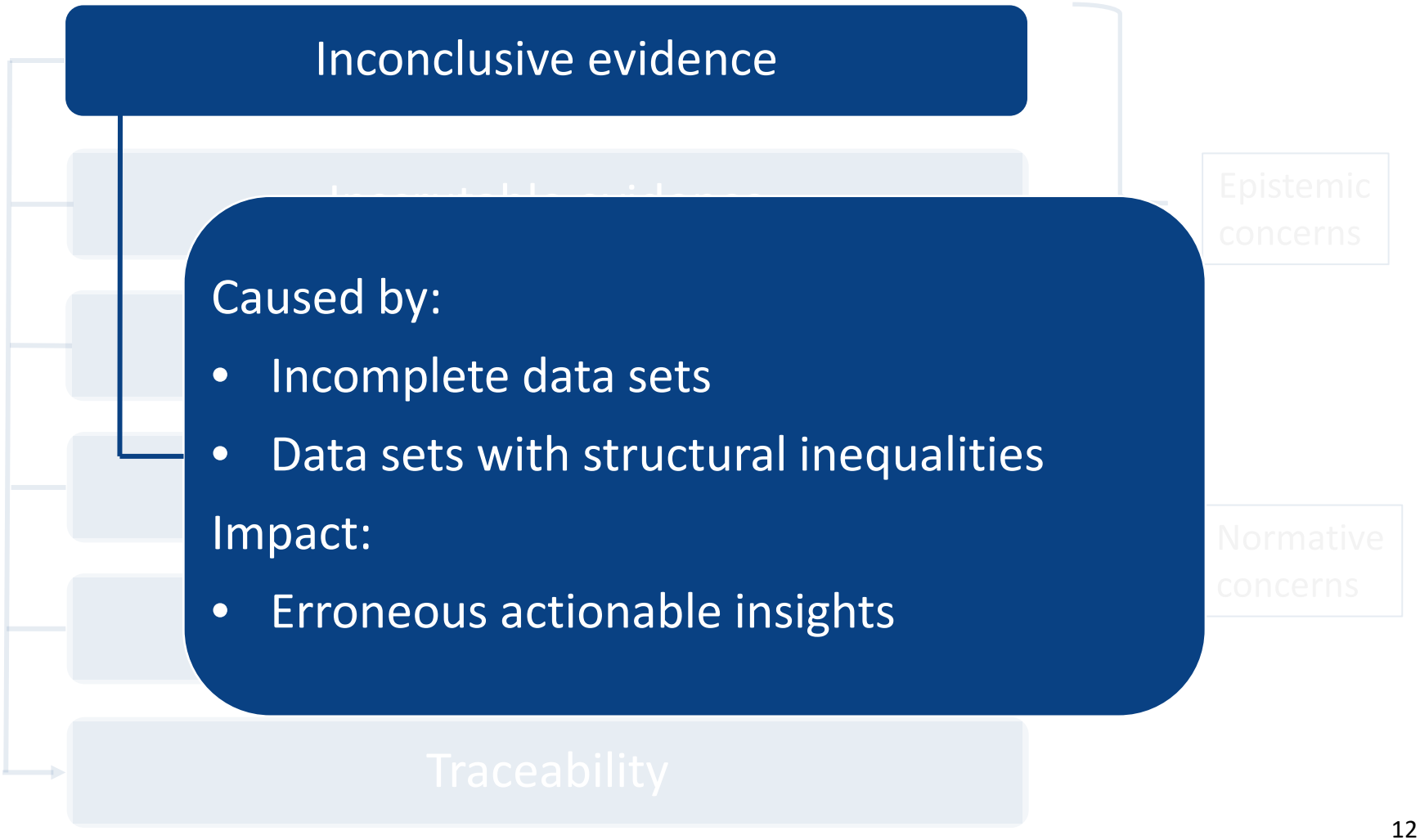
The Ethics of Artificial Intelligence

Six types of ethical concerns



Tsamados, Andreas, et al. "The ethics of algorithms: key problems and solutions." *Available at SSRN 3662302* (2020). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3662302

Six types of ethical concerns



Six types of ethical concerns

Inconclusive evidence

Inscrutable evidence

Caused by:

- Lack of appropriate tools for transparency
- Cognitive impossibility of interpretation

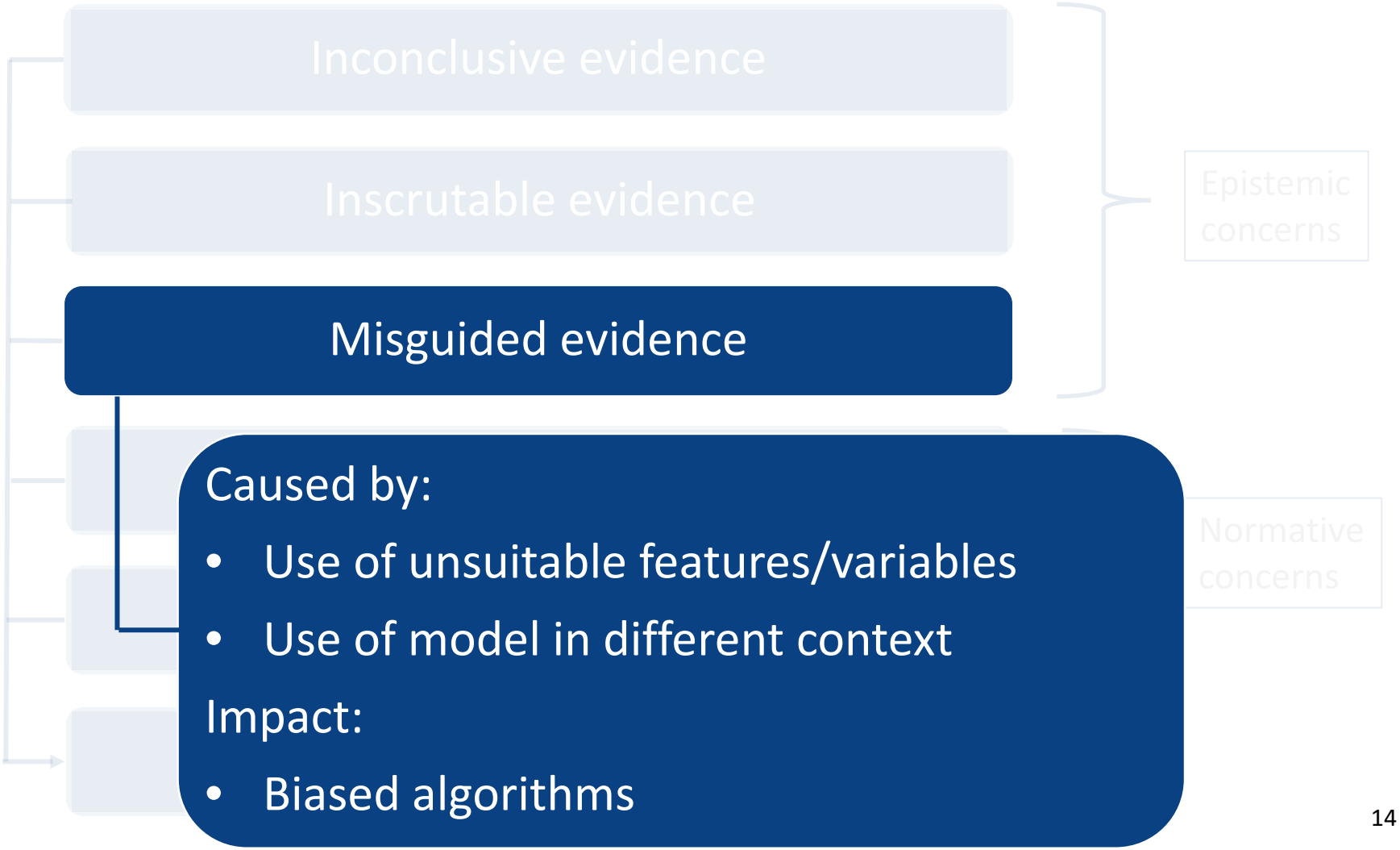
Impact:

- Erroneous actionable insights
- Shirk responsbility

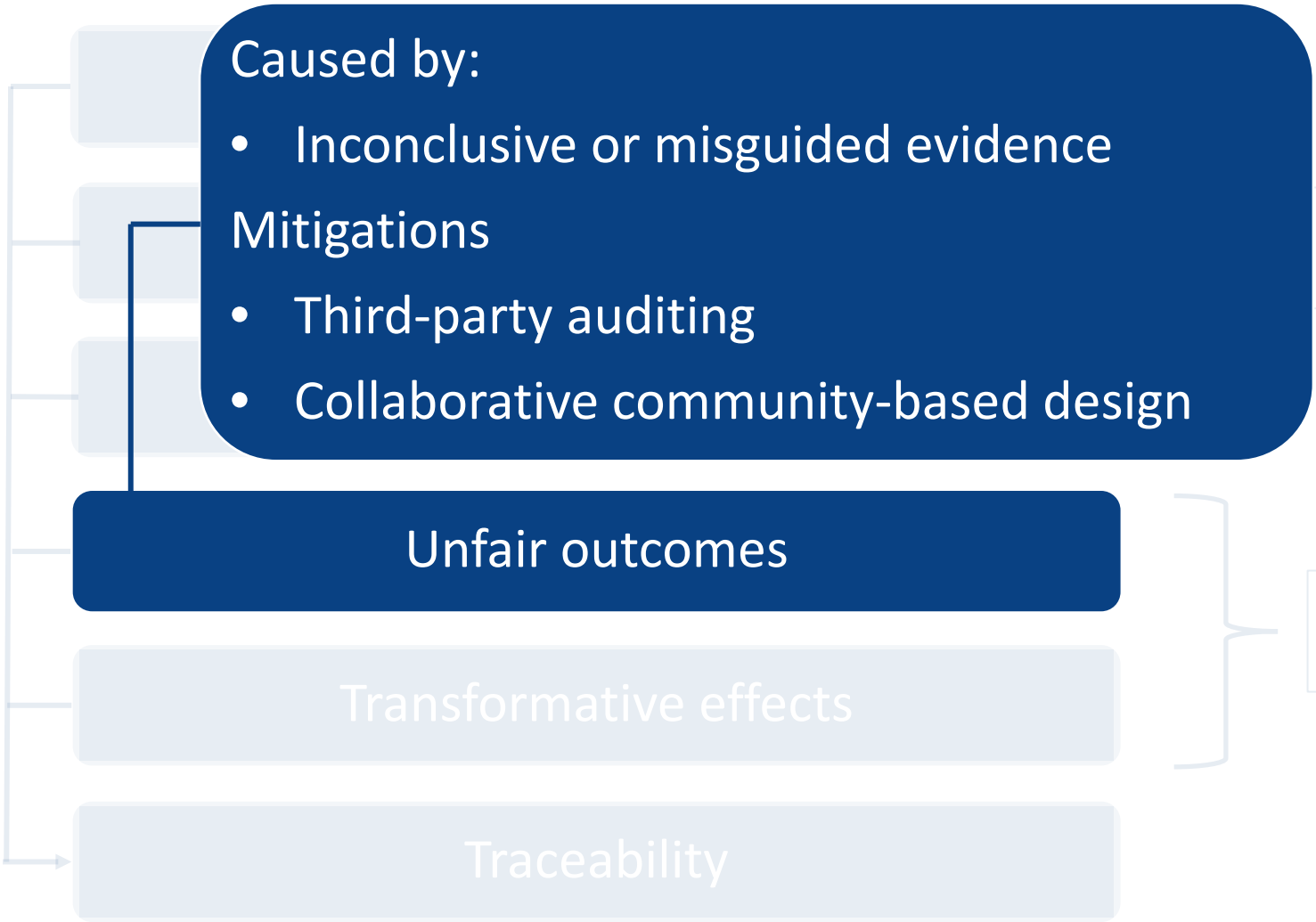
Epistemic concerns

Normative concerns

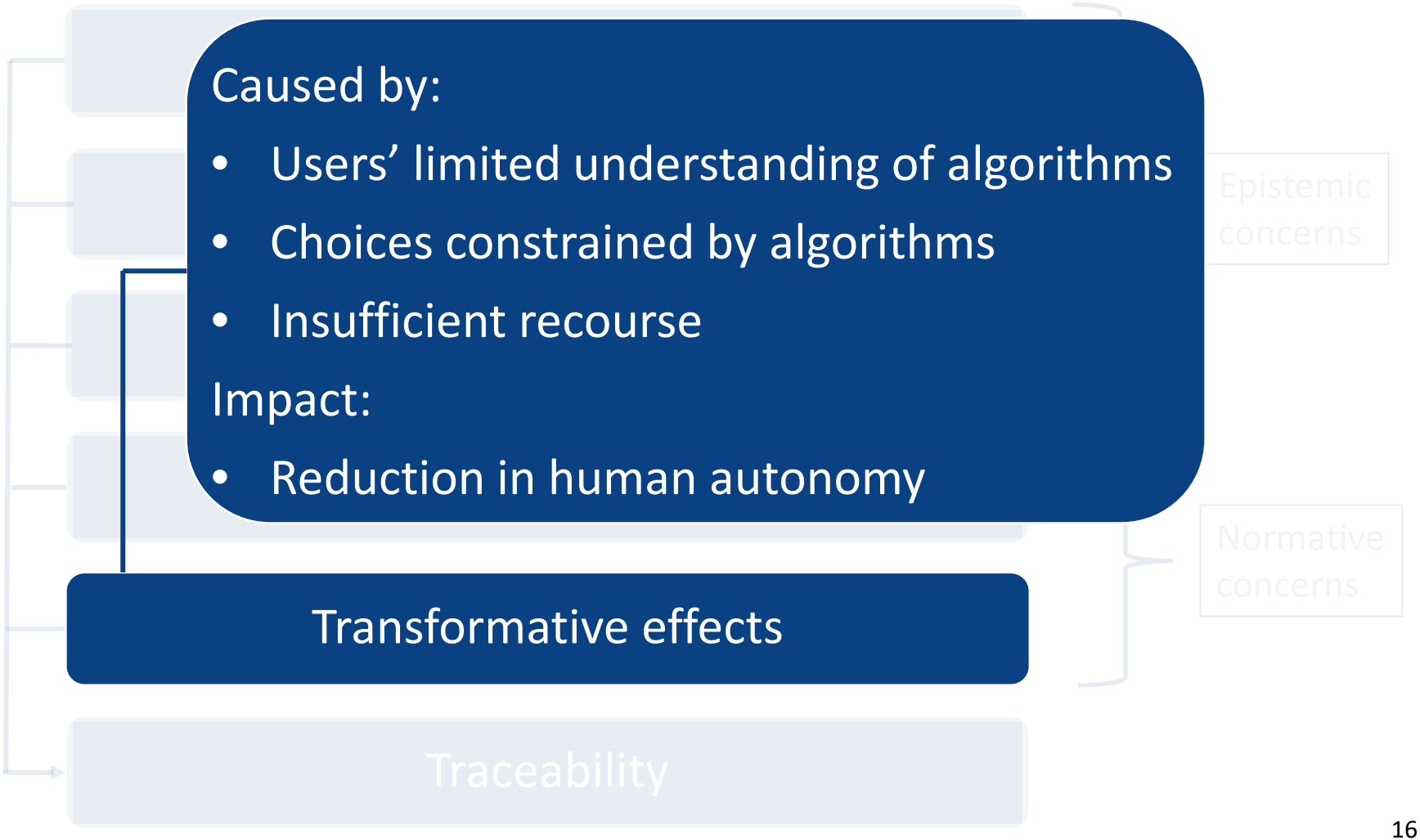
Six types of ethical concerns



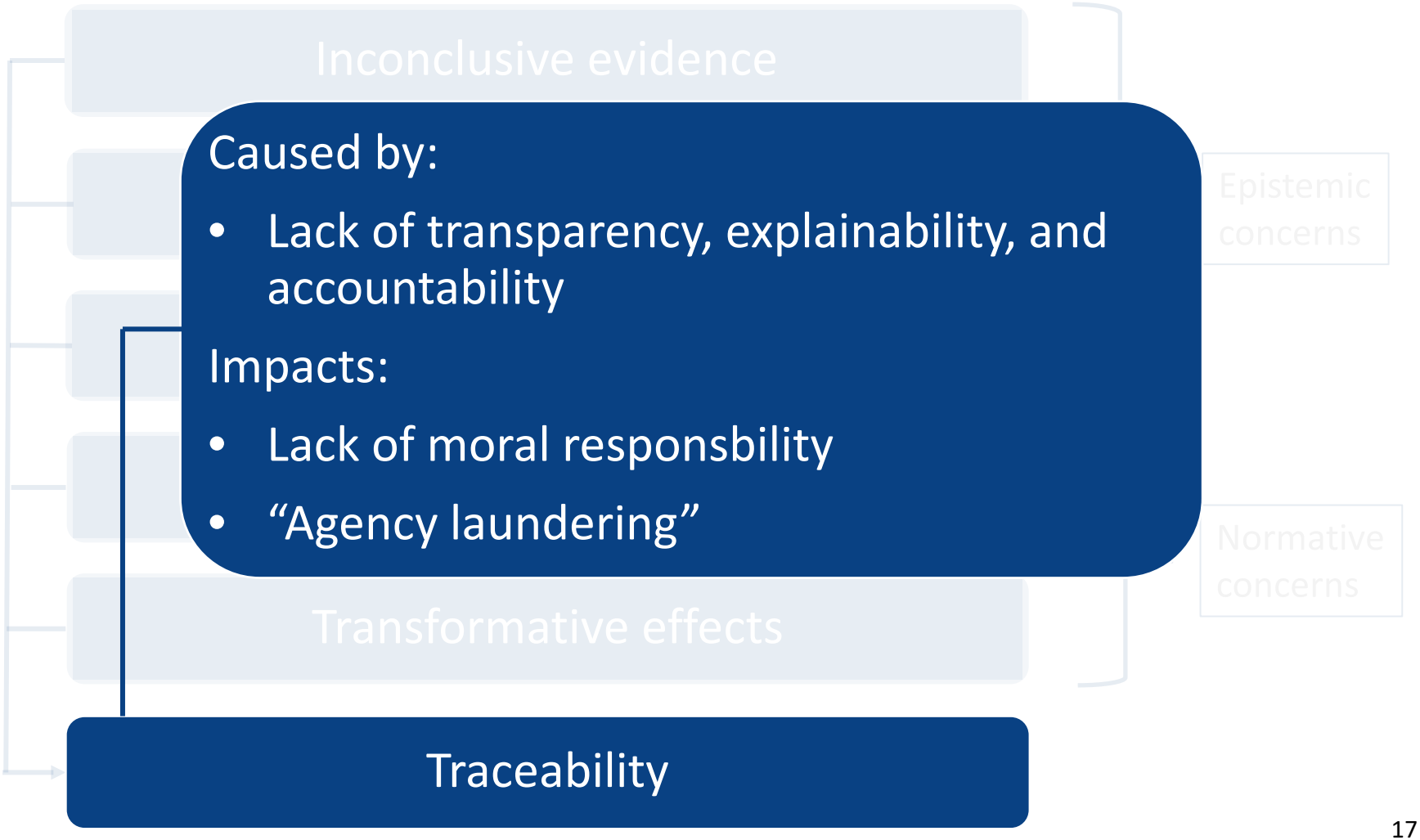
Six types of ethical concerns



Six types of ethical concerns



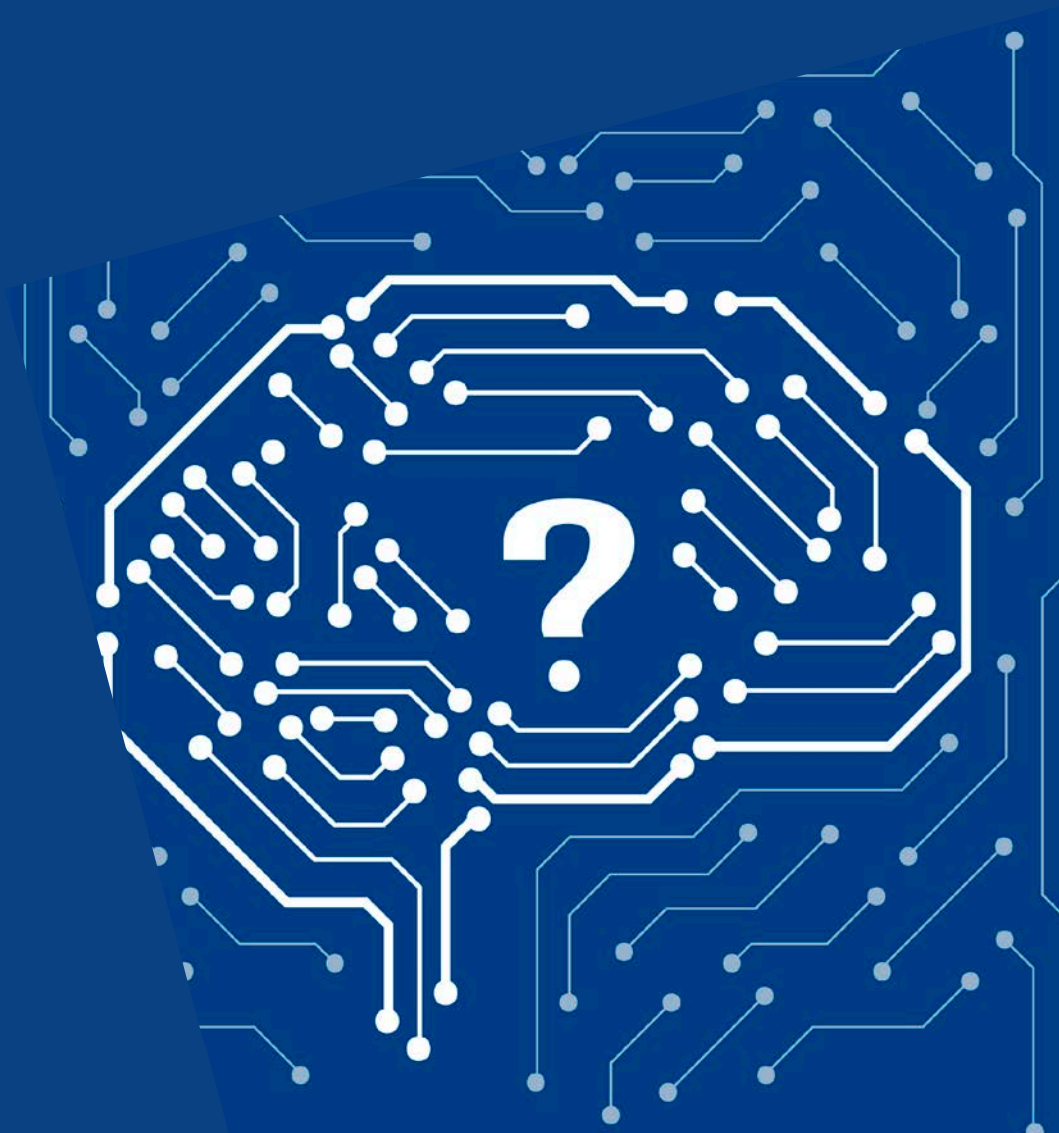
Six types of ethical concerns



Syllabus

COMP90087 Syllabus

Week	Module	Coordinator
1	Trust, machine, and digital ethics	Tim
2	The History of Artificial Intelligence	Tim
3	Philosophy and ethics	Simon
4	Fairness and accountability	Simon
5	Data governance	Marc
6	Accessibility and equity	Marc
7	Transparency: Decisions & processes	Marc
8	Explainability	Tim
9	Politics and policy	Kobi
10	Frameworks and implementation	Kobi
11	Reliability and safety	Kobi
12	Bringing it together	Simon

THE UNIVERSITY OF MELBOURNE

Have fun!