

Applying an ethical lens for more responsible modelling practice

Katrina Szetey^a, **Delphi Ward^{b,c}**, **Sabrina Chakori^d**, **David Douglas^e**, **Elizabeth A Fulton^{b,c}**,
Nicola Grigg^d, **Emma Ligtermoet^f**, **Claudia Munera-Roldan^f**, **Esther Onyango^g**,
Erinne Stirling^h, **Roshni Subramaniam^{i,c}**

a CSIRO Agriculture and Food, Black Mountain, ACT, Australia

b CSIRO Environment, Hobart, TAS, Australia

c Centre for Marine Socioecology, Hobart, TAS, Australia

d CSIRO Environment, Black Mountain, ACT, Australia

e University of New England, Armidale, NSW, Australia

f CSIRO Agriculture and Food, Floreat, WA, Australia

g CSIRO Agriculture and Food, St Lucia, QLD, Australia

h Independent scholar

i Parks Australia, Hobart, TAS, Australia

Corresponding author email: katrina.szetey@csiro.au

Abstract: Modelling is increasingly incorporating human dimensions, not only in systems modelling but also in models which may not explicitly appear to include the ‘social’. In the past, models have sometimes led to unjust or unethical solutions that negatively impacted certain groups. Today it is critically important to consider the ethical implications of model development – something that, in our experience, is rarely done beyond seeking ethics approvals for data collection. However, such approvals are historically grounded in medical research and may not cover the broader societal implications of modelling work, potentially repeating the harms of the past.

Our research took a reflexive and inductive approach to explore these implications, using six case studies of modelling projects in which members of the author team had been involved. We applied the principles of credibility, legitimacy, and salience (Cash et al. 2003) to analyse the case studies, and found them useful for examining the modelling process from an external perspective, being able to append “for whom” to the end of each principle. We further tested these insights through a co-production workshop with modellers at CSIRO, where we used a hypothetical scenario and an interest-influence matrix to highlight issues of power that can exist among stakeholders in modelling and AI projects. After discussion and reflection on their own experiences, participants identified ethical considerations spanning from individual responsibilities to assumptions embedded within the data and models. We conducted a thematic analysis of the workshop outputs and developed a list of 11 ethical dimensions, accompanied by example questions to consider during the modelling process. To support the application of these insights, we provided practical suggestions for developing a responsible modelling practice for modellers, end-users, and institutions.

For this presentation, we will begin by introducing participants to the findings of our research (7 min), followed by a short interactive session with the audience (8 min). This will involve another hypothetical case study – this time, of a generative AI model. We will pre-define a list of stakeholders, and invite the audience to place them on an interest/influence matrix. We will then use an interactive poll (e.g. Slido, Mentimeter, or equivalent) to gather reflections from the audience on how to examine the ethics of this case study guided by the list of 11 ethical dimensions. We will visualise this input in a word cloud.

Keywords: *Ethics, responsible innovation, model development*

REFERENCES

Cash DW, Clark WC, Alcock F, Dickson NM, Eckley N, Guston DH, Jäger J and Mitchell RB (2003)
‘Knowledge systems for sustainable development’, *Proceedings of the National Academy of Sciences*,
100(14):8086–8091, doi:10.1073/pnas.1231332100.