

AI = All In? How can HRECs address the "new research normal"?

Michael A. Martin

Australian National University

AHEC is advising NHMRC about guidelines around AI use in human research

- Credit to colleagues Prof Emma Kowal, Dr Yvonne Ho, Prof Jackie Leach Scully, Prof Erwin Loh
- Guidelines are in draft and being considered by AHEC and should be regarded as provisional at this stage and forthcoming
- Work is also being undertaken by others internationally, including WHO
- Research “involving AI” or “involving the use of AI” means the use of AI in any aspect of the research, whether or not AI is the subject of the research, but is restricted to human research

The problem...

- Researchers are using AI for/within a wide variety of tasks
- Use of AI is now being normalised – even routine web-searching invokes AI behind the scenes, beyond researchers' attention or overt control
- Use of AI has potentially profound ethical impacts on participants, because risks associated with AI use are obscured by “the black box” effect.
- What are the risks?

Risks... what could possibly go wrong?

THEMES / AREAS OF CONCERN

Overarching concern = Risk (individual + collective)

1. Human agency/oversight
2. Transparency/explainability
3. Bias/representativeness
4. Technical robustness/safety
5. Data reliability
6. Data governance/security

How is AI being used in research?

STAGES OF RESEARCH

- Discovery (concept generation and design)
- Literature review
- Planning and development
- Ethics and/or scientific review
- Recruitment
- Intervention (if relevant)
- Data collection
- Data management
- Data analysis
- Publication

The new normal = routine + non-routine use of AI

AI USE SPECTRUM

Intended use	Use of assessment tool required
Routine	No
Non-routine	
Automation	Yes
Generation	Yes
Intervention	Yes

Automation

- Examples of automation in research include
 - selection and randomisation of participants
 - participant reminders
 - creation of and/or conduct of online surveys
 - data cleaning, screening and extraction
 - limited data analysis
 - data visualization
 - research task repetition
- It is likely that much of what is currently considered to be non-routine use of automation in research will rapidly be re-categorised as routine use over time.

Generation

- Examples of generation include but are not limited to
 - using AI to draft text (e.g. for development of participant information sheets, interview questions, policy briefs, or academic publications)
 - for reviewing literature
 - for hypothesis generation
 - for data analysis
- Generative tasks are likely to become increasingly routine as AI tools become embedded in widely available software or in web processes such as search tools. In some cases, the fact that AI is being used may be obscured within software (e.g. search results)

Intervention

- Examples of intervention include but are not limited to the development of AI tools to
 - improve educational or social outcomes
 - to prevent, diagnose or manage disease.
- The development and testing of such tools places a significant onus on researchers to understand and be able to explain what risks participants may face when involved in the research. In one way this is not unlike explaining risks in medical procedures to participants in clinical trials (e.g. when the AI tool involves allocating participants to specific procedures or actions). But in another way, the risks may be harder to identify and describe.

How can researchers inform HRECs?

1. Identify AI use (see Stages of Research)
2. How will the AI tool be used at each relevant stage of the research project?
3. Will participant data be entered into any AI system operated by a third party?
4. What is known about the provenance of the AI tool's training data?
5. What are the characteristics of the AI tool that you intend to use?
6. Describe your plan for oversight of the AI and the outputs of the tools or systems used.
7. Describe the risks associated with your use of the AI tool in your research and your plan for mitigation and management of those risks.
8. What information about the use of AI in the research and the projected impact of any AI-assisted or AI-generated outputs of the research will be made available to participants, and the public (if relevant) and how will it be communicated?

How do these relate to risk?

3. Will participant data be entered into 3rd-party AI systems?

- will the data be retained by the AI system
- are there any restrictions, licensing conditions, or ethical considerations in the tool's creation that may affect participants' rights or data privacy
- whether the data will be used to further train the same model or other models
- how to ensure that the data will not be disclosed, exchanged or sold by the third party without participant consent.

Main issues: oversight of research data, data governance, privacy, consent

How do these relate to risk?

4. Provenance of the AI tool's training data

Do the data sets that are used to train the AI algorithms adequately represent the population/s impacted by or relevant to the research? Is the tool appropriate for the cohort? (Potential issue: Bias)

5. Features of the AI tool to be used

- Has the AI system been validated for its intended use/s?
- Is the AI tool non-adaptive or adaptive (involving continuous learning algorithms while the researcher is using the tool)?
- If adaptive, what are the limits to permitted adaptations?
- reproducibility and/or replicability?
- Will the AI interact with research participants and/or research staff? How?

(Potential issues: Bias, Human oversight, transparency, safety)

How do these relate to risk?

6. Oversight of the AI and related outputs

- Safeguards to ensure human oversight, especially when critical decisions are necessary or unexpected developments arise?
- How to detect biased or unfair outputs? How to mitigate such biases?
- Critical assessment of the outputs of the AI tools/systems and how to intervene to minimise risks to participants
- Monitoring and assessment of the performance of the AI over time

How do these relate to risk?

7. Risk mitigation and management plan

- What are the potential harms that could result from the use of the AI tool and to whom? (**Identify and assess severity**)
- How likely are these harms? (**Likelihood**)
- How will these risks be mitigated and managed, how will this be evaluated?
- Ongoing risk assessment and management process

8. What information about the use of AI in the research and the projected impact of any AI-assisted or AI-generated outputs of the research will be made available to participants, and the public, and how will it be communicated?

Another useful tool - GAIDeT

<https://panbibliotekar.github.io/gaidet-declaration/index.html>

<https://www.tandfonline.com/doi/full/10.1080/08989621.2025.2544331> (paywalled)

<https://futurecampus.com.au/2025/09/01/research-integrity-demands-an-ai-taxonomy/>

GAIDeT Declaration Generator

This is a beta version of the GAIDeT Declaration Generator — an interactive tool for researchers to disclose the delegation of tasks to generative AI (GAI) tools in accordance with the GAIDeT taxonomy.

 Learn more about GAIDeT: Suchikova, Y., Tsybuliak, N., Teixeira da Silva, J. A., & Nazarovets, S. (2025). GAIDeT (Generative AI Delegation Taxonomy): A taxonomy for humans to delegate tasks to generative artificial intelligence in scientific research and publishing. Accountab
<https://doi.org/10.1080/08989621.2025.2544331>

 You can also [request the full text via ResearchGate](#).

To disclose the use of generative AI, fill in the relevant fields, after which the declaration will be generated automatically. Copy its text and insert it into your manuscript.

1. Who delegated the tasks?

e.g., Yevhen Promptovych or 'Collective responsibility'

2. Which LLM (and version) was used?

e.g., ChatGPT-4.5, Claude 3, Gemini 1.5

3. Select Delegated Tasks (from GAIDeT taxonomy)

▼ Conceptualization

- ☐ Idea generation
- ☐ Defining the research objective
- ☐ Formulating research questions and hypotheses
- ☐ Feasibility assessment and risk evaluation
- ☐ Preliminary hypothesis testing

► Literature Review

► Methodology

► Software Development and Automation

► Data Management

The GAIDeT taxonomy - delegated tasks

▼ Conceptualization

- ☐ Idea generation
- ☐ Defining the research objective
- ☐ Formulating research questions and hypotheses
- ☐ Feasibility assessment and risk evaluation
- ☐ Preliminary hypothesis testing

▼ Literature Review

- ☐ Literature search and systematization
- ☐ Writing the literature review
- ☐ Analysis of market trends and/or patent environment
- ☐ Evaluation of the novelty of the research and identification of gaps

▼ Methodology

- ☐ Research design
- ☐ Development of experimental or research protocols
- ☐ Selection of research methods

▼ Software Development and Automation

- ☐ Code generation
- ☐ Code optimization
- ☐ Process automation
- ☐ Creation of algorithms for data analysis

▼ Data Management

- ☐ Data collection
- ☐ Validation
- ☐ Data cleaning
- ☐ Data curation and organization
- ☐ Data analysis
- ☐ Visualization
- ☐ Reproducibility testing

▼ Writing and Editing

- ☐ Text generation
- ☐ Proofreading and editing
- ☐ Summarizing text
- ☐ Formulation of conclusions
- ☐ Adapting and adjusting emotional tone
- ☐ Translation
- ☐ Reformatting
- ☐ Preparation of press releases and outreach materials

▼ Ethics Review

- ☐ Bias analysis and potential discrimination assessment
- ☐ Ethical risk analysis
- ☐ Monitoring compliance with ethical standards
- ☐ Data confidentiality monitoring

▼ Supervision

- ☐ Quality assessment
- ☐ Trend identification
- ☐ Identification of limitations
- ☐ Recommendations
- ☐ Publication support