Thinking Ethically About GenAI-Assisted Educational Research: Principles, Practices, and Emerging Dilemmas

Lorien S. Jordan
University of South Florida

FERA 2026

1

Copy the QR code to get access to handouts, case study scripts, and resources

Or visit:
lorienjordan.com/fera

2

What is AI?

Artificial Intelligence (AI)
- The broader category of technologies that perform tasks involving prediction, classification, decision-making, or other capabilities associated with human intelligence

Generative AI (GenAI)
- A type of AI that generates new content, such as text, images, or audio, based on patterns learned during training

Large Language Models (LLMs)
- Models trained on language that power many generative AI chatbots, including ChatGPT.

3

Why is AI Ethics So Difficult?

- How do we make ethical decisions about a technology when we don't agree on what we're building toward?

Toner, 2024

4

Mapping Today's Conversation

Sociotechnical Ethics
- Broad considerations surrounding AI in research

Practical Ethics
- The decisions we make when using AI in research

5

Four Questions to Consider

- What does it take to make AI work?
- How Is AI Trained?
- How do we adapt to AI?
- Who gets to decide what AI can do, and who benefits?

Página 05

6

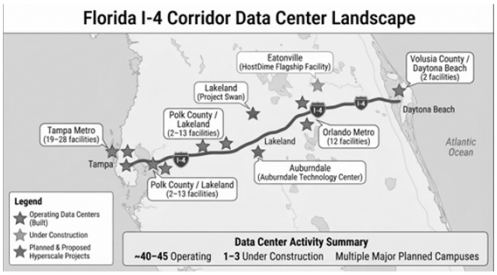
What Does It Take to Make AI Work?

- Mineral extraction zones
- Toxic e-waste

Crawford, 2021; Jordan et al., 2026a

7

What Does It Take to Make AI Work?



8

What Does It Take to Make AI Work?

- AI sustained through human extraction
- "Clean" outputs built on "invisible" lives

Jordan et al., 2026a; Muldoon & Wu, 2023

9

How Is AI Trained?

- Training data comes from human-created and AI-created materials
- Collection often occurs without consent
- Data is selected, filtered, and organized
- Human judgments shape the training process

Jordan, 2026a; Weber & Prietl, 2021

10

How Is AI Trained?

- AI identifies patterns in data
- Patterns can miss context and meaning

Jordan, 2026a; Weber & Prietl, 2021

11

How Is AI Trained?



Prompt—Black African doctor is helping poor and sick white children, photojournalism



Prompt—Traditional African healer is helping poor and sick white children



Prompt—Traditional African healer is healing a white child in a hospital



Prompt—Doctors help children in Africa

Alenichev et al., 2023

12

How Do We Adapt to AI?

- Conversational warmth and confidence
- Trust in seemingly authoritative responses
- Gradual changes in how we think and inquire

Jordan et al., 2026b; Sourati et al., 2026

13

Who Gets to Decide and Who Benefits?

- Unequal access
- Academic labor
- Learning opportunities
- Productivity pressures

Gartenberg et al., 2026; Jordan et al., 2026b

14

Who Gets to Decide and Who Benefits?

- Technology companies
- Governments
- Institutions
- Researchers

Jiang, 2026; Jordan, 2026b

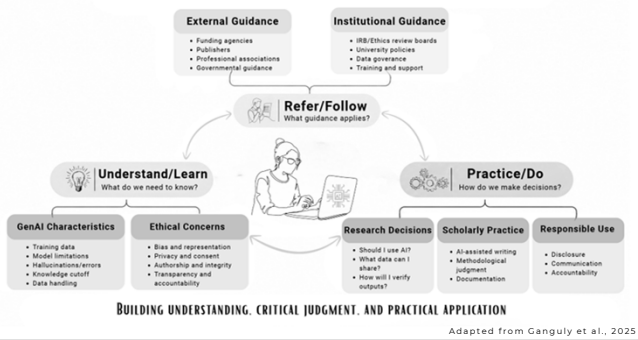
15

What Can We Actually Do?

- Individual: Be deliberate about AI use
- Institutional: Question adoption and governance
- Collective: Participate in standards and community action

16

Navigating GenAI Ethics in Educational Research



Building understanding, critical judgment, and practical application

Adapted from Ganguly et al., 2025

17

Refer and Follow

- Institutions and research ethics boards
- Funding agencies
- Publishers and professional associations
- Technology providers and terms of use
- Research agreements and participant commitments

18

Refer and Follow

Recognizing the Limits of Guidance

- Guidance might overlap and conflict with other guidance
- Institutional guidance is often very general
- GenAI evolves faster than policies can be developed or revised

19

From Understanding to Acting

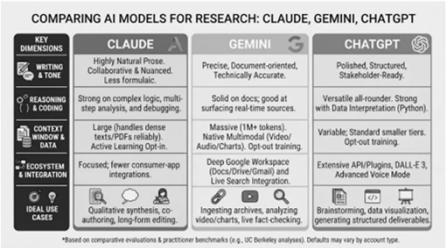
Should I Use GenAI For This Task

- Start with the task
- Consider the purpose
- Examine methodological alignment
- Look beyond efficiency

20

From Understanding to Acting

- Evaluate whether the model is appropriate for the task



Adapted from MIDS UC Berkeley, 2025

21

From Understanding to Acting

Privacy, Model Training, and Data Governance

- Privacy is a major ethical concern
- Research data may leave our control
- Data may be stored, retained, or used for model training

22

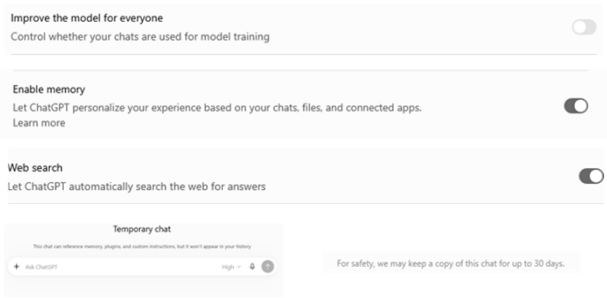
From Understanding to Acting

Does Turning Off Training Protect Our Data?

- Protections vary by platform and account type
- Training opt-outs are often available
- Data can still be stored, logged, or processed

23

From Understanding to Acting



24

From Understanding to Acting

Where Does Our Data Go?

Deployment	Where processing occurs	Main consideration
Cloud-based	External provider's infrastructure	Provider agreements, access, retention, and security
Local	Researcher's computer	Device security, storage, and software configuration

25

From Understanding to Acting

When Is De-Identified Data Still Identifiable?

- Removing direct identifiers is not a guarantee
- Contextual details can reveal identities
- AI can make linking information across sources easier

Li, 2026; Lermen et al., 2026

26

From Understanding to Acting

- **What About Educational Data?**
  - Student records may be subject to additional legal and contractual protections
  - Access to educational data doesn't necessarily authorize sharing them with GenAI providers
  - AI-generated classifications and interpretations can have consequences for students

27

From Understanding to Acting

- **What Did Participants Agree To?**
  - Consider what participants understood about how their data would be used and protected
  - Ask whether GenAI changes how, where, or by whom their information is processed
  - Revisit consent agreements and consult the IRB when proposed uses may extend beyond what was originally approved

28

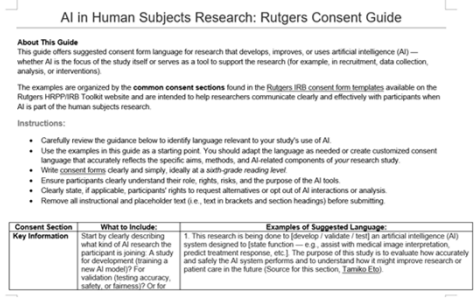
From Understanding to Acting

How to Reduce Data Sharing Risks?

- Minimize what we share
  - Use hypothetical examples, smaller excerpts, or alternative workflows
- Reduce identifiability
  - Generalize identifying details and consider what combinations might reveal
- Choose appropriate environments
  - Consider approved institutional systems or offline models
  - Verify retention and access
- Address research obligations
  - Review informed consent, research protocols, and IRB requirements

29

From Understanding to Acting: Resource



AI in Human Subjects Research: Rutgers Consent Guide

About This Guide
This guide offers suggested consent form language for research that develops, improves, or uses artificial intelligence (AI) — whether AI is the focus of the study itself or serves as a tool to support the research (for example, in recruitment, data collection, analysis, or interventions).

The examples are organized by the **common consent sections** found in the Rutgers IRB consent form templates available on the Rutgers HRPP/IRB Toolkit website and are intended to help researchers communicate clearly and effectively with participants when AI is part of the human subjects research.

Instructions:

- Carefully review the guidance below to identify language relevant to your study's use of AI.
- Use the examples in this guide as a starting point. You should adapt the language as needed or create customized consent language that accurately reflects the specific aims, methods, and AI-related components of your research study.
- Write consent forms clearly and simply, ideally at a *sixth-grade reading level*.
- Ensure participants clearly understand their role, rights, risks, and the purpose of the AI tools.
- Clearly state, if applicable, participants' rights to request alternatives or opt out of AI interactions or analysis.
- Remove all instructional and placeholder text (i.e., text in brackets and section headings) before submitting.

Consent Section	What to Include:	Examples of Suggested Language:
Key Information	Start by clearly describing what kind of AI research the participant is joining. A study for development (training a new AI model)? For validation (testing accuracy, safety, or fairness)? Or for	1. This research is being done to [develop / validate / test] an artificial intelligence (AI) system designed to [state function — e.g., assist with medical image interpretation, predict treatment response, etc.]. The purpose of this study is to evaluate how accurately and safely the AI system performs and to understand how it might improve research or patient care in the future (Source for this section, Tamiko Eto).

Ringel et al. 2026

30

From Understanding to Acting

Privacy, data governance and consent

- Where would you draw the line between what your institution permits and what you consider ethically appropriate when using GenAI in research?
- When we introduce AI into our research, how do we determine whether its use is consistent with what participants originally understood and agreed to?

31

From Understanding to Acting

When Can GenAI Get Wrong?

- Hallucination
  - Fabrication, misrepresentation, unsupported claims, faulty reasoning
- Omission and distortion
- Bias
- Inconsistency
- Overconfidence

32

From Understanding to Acting

The Black-Box Problem

- Limited transparency about training data, development, and evaluation
- Complex systems make individual outputs difficult to explain
- Opacity complicates independent scrutiny and accountability

33

From Understanding to Acting

What Does Verification Actually Require?

- Establish evaluation criteria before using AI
- Compare outputs against independent evidence or benchmarks
- Examine patterns of errors, not just individual mistakes
- Document limitations and decisions

34

From Understanding to Acting

What About Bias and Representation?

- Bias can shape classifications, analyses, and interpretations
- Overall accuracy can conceal unequal performance
- Dominant assumptions can distort contextual meaning
- Human review requires relevant expertise

Jadhav et al., 2026

35

From Understanding to Acting

How Do We Evaluate Bias and Representation?

- Examine omissions and underrepresented perspectives
- Question categories and underlying assumptions
- Compare outputs across groups and contexts

Resnik & Hosseini, 2025

36

From Understanding to Acting

Accuracy, Bias, and Verification

- What practices do we already use to evaluate research quality that could help us assess GenAI outputs?
- What would you need to examine before you felt comfortable using an AI-generated interpretation or result in your research?

37

From Understanding to Acting

Existing practice	How it might help evaluate GenAI outputs
Triangulation	Compare AI-generated findings or interpretations with other sources of evidence, analytic approaches, or human interpretations.
Peer review / Peer debriefing	Have colleagues critically examine AI-generated analyses, looking for unsupported claims, omissions, or alternative interpretations.
Validity / Construct validity	Examine whether an AI-generated measure, score, or classification actually represents the construct researchers intend to study.
Reliability / Consistency checks	Assess whether the model produces comparable results across repeated trials, prompts or similar cases.
Reflexivity	Consider how researchers' assumptions influence their prompts, judgments about outputs, and decisions about what to accept or reject.
Audit trails / Documentation	Document how AI was used, what outputs were generated, and how researchers evaluated or modified those outputs.
Negative case analysis	Look for data that contradict AI-generated patterns, summaries, or interpretations.
Sensitivity analysis	Examine whether findings change when prompts, models, or parameters are altered.
Member checking	When methodologically appropriate, invite participants to respond to AI-assisted representations of their experiences.
Verification against original sources	Return to transcripts, datasets, cited literature, or other primary materials to assess whether AI-generated claims are supported.

38

From Understanding to Acting

AI-Assisted Writing and Scholarly Integrity

- Accuracy and representation: AI may introduce errors, misrepresent scholarship, or alter intended meaning
- Intellectual contribution: AI can influence how we develop arguments, synthesize scholarship, and interpret findings
- Authorship: AI assistance complicates questions about scholarly contribution, credit, and originality
- Plagiarism and attribution: AI may reproduce or closely paraphrase existing work without appropriate acknowledgment
- Accountability: Human authors remain responsible for the integrity of their work

Foley, et al., 2025

39

From Understanding to Acting

Why Is AI Disclosure Complicated?

- Disclosure extends beyond AI-assisted writing
- Expectations and thresholds for disclosure vary
- Concerns about judgment, credibility, and consequences may discourage disclosure

BaHammam, 2026

40

From Understanding to Acting

When and What Should We Disclose?

Table 1. Disclosing AI use in research and writing.

Disclosure is mandatory when, for example, using AI

- To formulate questions or hypotheses, design and conduct experiments.
- To draft parts of the paper, summarize, paraphrase, significantly revise or synthesize textual content.
- To translate parts or the whole paper.
- To collect, analyze, interpret or visualize data (quantitative or qualitative).
- To extract data for review of the literature (systematic or not) and identify knowledge gaps.
- To generate synthetic data and images reported in the paper or used in research.

Disclosure is optional when, for example, using AI

- To edit existing text for grammar, spelling or organization.
- To find references or verify the relevance of human-found references.
- To find and generate examples for existing content.
- To brainstorm and offer suggestions for the organization of a paper or the title of a paper/section.
- To validate and/or offer feedback on existing ideas, text and code.

Disclosure is unnecessary when, for example, using AI

- To suggest words or phrases that enhance clarity/readability of an existing sentence.
- In part of a larger operation where AI is not generating or synthesizing content or making research decisions; for example, when AI is integrated into other systems/machines.
- As a digital assistant, for example, to help organize and maintain a project's digital assets and workflows.

Resnik & Hosseini, 2026

41

From Understanding to Acting

Disclosing Your AI Use

- Tool Identification: Specific name, version, and provider (e.g., Claude 3.5 Sonnet (Anthropic) or ChatGPT (OpenAI, GPT-4)).
- Specific Application: Exactly what tasks it assisted with (e.g., copy-editing, brainstorming structural outlines, formatting citations, or assisting with code syntax).
- Human Oversight: Explicit confirmation that the researchers critically reviewed, verified, and take full responsibility for all content.
- Exclusion of Sensitive Data: Confirmation that no confidential participant data, PII, or raw qualitative transcripts were uploaded to a public model.

Earp et al., 2026; Resnik & Hosseini, 2026

42

From Understanding to Acting: Resource



Ballsun-Stanton & Smith, 2025

43

From Understanding to Acting

Disclosing Your AI Use

- The first author used ChatGPT (OpenAI; GPT-4o and GPT-5.6 Sol) and Claude (Anthropic; Sonnet 4.5 and Opus 4.8) to support brainstorming, editing, and language refinement during manuscript preparation. Specifically, these tools were used to identify redundancy, improve clarity, and suggest more concise phrasing. All AI-generated suggestions were reviewed and evaluated by the authors, who retain full responsibility for the accuracy, originality, and integrity of the manuscript. No confidential participant data or personally identifiable information were uploaded to AI. All theoretical arguments, substantive content, and scholarly contributions are the original work of the authors.

Jordan et al., 2026b

44

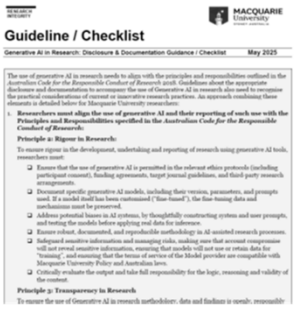
From Understanding to Acting

Writing and Disclosure

- What would make a disclosure of GenAI use genuinely useful for evaluating a study, rather than just satisfying a requirement?

45

From Understanding to Acting: Resource



Guideline / Checklist

Generative AI in Research: Disclosure & Documentation Guidance / Checklist May 2025

The use of generative AI in research needs to align with the principles and responsibilities outlined in the *Australian Code for the Responsible Conduct of Research 2018*. Guidelines about the appropriate disclosure and documentation to accompany the use of Generative AI in research also need to encompass the practical considerations of current or innovative research practices. An approach combining these elements is detailed below for Macquarie University researchers:

1. **Researchers must align the use of generative AI and their reporting of such use with the Principles and Responsibilities specified in the *Australian Code for the Responsible Conduct of Research*:**

Principle 2: Rigour in Research:

To ensure rigour in the development, undertaking and reporting of research using generative AI tools, researchers must:

- ❑ Ensure that the use of generative AI is permitted in the relevant ethics protocols (including participant consent), funding agreements, target journal guidelines, and third-party research arrangements.
- ❑ Document specific generative AI models, including their version, parameters, and prompts used. If a model itself has been customised ("fine-tuned"), the fine-tuning data and mechanisms must be preserved.
- ❑ Address potential biases in AI systems, by thoughtfully constructing system and user prompts, and testing the models before applying real data for inference.
- ❑ Ensure robust, documented, and reproducible methodology in AI-assisted research processes.
- ❑ Safeguard sensitive information and managing risks, making sure that account compromise will not reveal sensitive information, ensuring that models will not use or retain data for 'training', and ensuring that the terms of service of the Model provider are compatible with Macquarie University Policy and Australian laws.
- ❑ Critically evaluate the output and take full responsibility for the logic, reasoning and validity of the content.

Principle 3: Transparency in Research

To ensure the use of Generative AI in research methodology, data and findings is openly, reasonably

Wahle et al., 2023

46

From Understanding to Acting

What Does Accountability Require?

- ▹ Establish meaningful human oversight and decision-making authority
- ▹ Document consequential GenAI uses and research decisions
- ▹ Investigate problems and remain willing to correct course

AI Education Research Principles Design Lab, 2026

47

Supporting Responsible Research: Resources



Sheffield Hallam University | LIBRARY: Research Support

Checklist for GenAI use in research

We advocate for a **responsible use of GenAI**, therefore we have prepared a set of questions you should consider in order to remain reflective throughout the entire cycle of your research project when engaging with AI tools.

RESEARCH PROJECT: _____

PROJECT DURATION: _____

PI: _____

Here are the key questions around GenAI in the research lifecycle.
We have also prepared a set of more detailed questions for each stage of your research project.

Should I use GenAI?

Sheffield Hallam
RESEARCHER / PROJECT LEVEL
What do I need to consider when deciding whether and how to use GenAI throughout my research?

THE WORKSHEET

01 02 03

Participant Data and Model Training | API Retention | Feature Updates

ResearchOps
ORGANISATIONAL LEVEL
What systems, oversight, and safeguards need to be in place to support responsible AI-assisted research?

Sheffield Hallam University, 2025; The ResearchOps Review, 2026

48

From Understanding to Acting

THANK YOU!

LSJORDAN@USF.EDU

49

THANK YOU!

LSJORDAN@USF.EDU

Recommended readings

- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Hao, K. (2025). *Empire of AI: Dreams and nightmares in Sam Altman's OpenAI*. Penguin Group.

References

- AI Education Research Principles Design Lab. (2026). Principles for the use of artificial intelligence in education research. American Institutes for Research. https://www.air.org/sites/default/files/2026-10/Principles-for-Use-AI-in-Education-October-2026-508.pdf?
- Alenichev, A., Kingori, P., & Grietens, K. P. (2023). Reflections before the storm: the AI reproduction of biased imagery in global health visuals. *The Lancet Global Health, 11*(10), e1496-e1498.
- Ballsun-Stanton, B., & Smith, S. (2025). Generative AI in research: Disclosure & documentation guidance / checklist (Version 1.0.0). Zenodo. https://zenodo.org/records/15307492
- Earp, B. D., Schüklenk, U., Savulescu, J., & Porsdam Mann, S. (2026). Against mandatory prompt disclosure in AI-assisted scholarship. *AI & SOCIETY*, 1-15.
- Foley, K., McLean, C., De Zylva, R., Asa, G., Maio, J., Batchelor, S., Dzando, G. & Dimassi, A. (2025). Developing a critical imagination for how researchers can use artificially intelligent tools reflexively and responsibly during qualitative literature reviews. *International Journal of Qualitative Methods*, 24, https://doi.org/10.1177/16094069251316249.
- Ganguly, A., Johri, A., Ali, A., McDonald, N. (2025). Generative artificial intelligence for academic research: evidence from guidance issued for researchers by higher education institutions in the United States. AI Ethics 5, 3917–3933. https://doi.org/10.1007/s43681-025-00688-7
- Gartenberg, C., Hasan, S., Murray, A., & Pierce, L. (2026). More versus better: Artificial intelligence, incentives, and the emerging crisis in peer review. *Organization Science*, 37(3), 795-812. https://doi.org/10.1287/orsc.2026.ed.v37.n3
- Jadhav, R., Danve, J., & Shaw, S. (2026). Implicit grading bias in large language models: How writing style affects automated assessment across math, programming, and essay tasks. *arXiv* preprint arXiv:2603.18765.
- Jordan, L. S. (2026b). Qualitative ethics and the AI war machine. *Qualitative Inquiry*. Online first. 1-12. https://doi.org/10.1177/10778004261468239
- Jordan, L. S., Johnson, M. W., & Wolgemuth, J. R. (2026a). The slow violence of the Anthropocene: Ecological degradation, qualitative ethics, and artificial intelligence. *Qualitative Inquiry*. https://doi.org/10.1177/10778004261489156
- Jordan, L. S., Johnson, M. W., & Wolgemuth, J.R. (2026b). Risky collaborators: Qualitative research ethics in an era of algorithmic co-presence. *International Journal of Social Research Methodology*. https://doi.org/10.1080/13645579.2026.2743033

50

- Lermen, S., Paleka, D., Swanson, J., Aerni, M., Carlini, N., & Tramèr, F. (2026). Large-scale online deanonymization with LLMs. arXiv preprint arXiv:2602.16800.
- Li, T. (2026). Agentic LLMs as powerful deanonymizers: Re-identification of participants in the Anthropic Interviewer dataset. arXiv preprint arXiv:2601.05918.
- Master of Information and Data Sceine (MIDS). (2026, April). *Which LLM should I use? A data scientist's guide to ChatGPT, Gemini, and Claude*. UC Berkeley School of Information. https://ischoolonline.berkeley.edu/blog/chatgpt-vs-gemini-vs-claude-comparison/
- Michael, N., BenShushan, D., Bien, J., & Moore, D. A. (2026). Confidence Calibration in Large Language Models. arXiv preprint arXiv:2605.23909.
- Muldoon, J., & Wu, B. (2023). Artificial intelligence in the colonial matrix of power. *Philosophy & Technology, 36*(4), Article 80. https://doi.org/10.1007/s13347-023-00687-8
- Privette, A. P., Barros, A., & Cai, X. (2026). Data centers water footprint: The need for more transparency. AGU Advances, 7, e2025AV002140. https://doi.org/10.1029/2025AV002140
- Resnik, D. B., & Hosseini, M. (2025). The ethics of using artificial intelligence in scientific research: new guidance needed for a new tool. AI and Ethics, 5(2), 1499-1521.
- Resnik, D. B., & Hosseini, M. (2026). Disclosing artificial intelligence use in scientific research and publication: When should disclosure be mandatory, optional, or unnecessary? Accountability in Research, 33(2). https://doi.org/10.1080/08989621.2025.2481949
- Ringel, M., Watkinson, M., & Walker, N. C. (2026). AI in human subjects research: Rutgers consent guide. Rutgers University, Office for Research. https://research.rutgers.edu/faculty-staff/compliance/human-research-protection/toolkit?
- Sheffield Hallam University. (2025). Checklist for GenAI use in research. https://www.shu.ac.uk/-/media/home/research/ethics-integrity-and-practice/checklist-for-ai-use-in-research-v4.pdf?
- Sourati, Z., Ziabari, A. S., & Dehghani, M. (2026). The homogenizing effect of large language models on human expression and thought. *Trends in Cognitive Sciences*.
- The ResearchOps Review. (2026). Research privacy & ethics in the age of AI. Rally. https://rallyuxr.com/resources/research-privacy-ethics-in-the-age-of-ai?
- Toner, H. (2024, April). How to govern AI—even if it's hard to predict [Video]. TED.
- Weber, J., & Prietl, B. (2021). AI in the age of technoscience: On the rise of data-driven AI and its epistem-ontological foundations. In A. Elliott (Ed.). The Routledge social science handbook of AI (pp. 58-73). Routledge.

51