AI in Research: Policy, SOPs, and Templates

Guidance for Responsible, Reproducible AI Across the Research Lifecycle

[INSTITUTION] — [DEPARTMENT]

2025-09-15

Table of contents

1 Overview & Quick Start

> **Note**
>
> **Purpose.** This packet provides a departmental policy, standard operating procedures (SOPs), stage-by-stage checklists, and ready-to-paste templates for disclosing and documenting AI use in research.

1.1 Who this is for

Researchers, PIs, data stewards, model owners, and editors/reviewers affiliated with **[INSTITUTION]** — **[DEPARTMENT]**.

1.2 Quick-start: Ten Rules

1. **Don't upload confidential material** (unpublished manuscripts, grants, identifiable data, licensed instruments) to public AI tools.
2. **Humans are responsible.** AI is never an author; disclose substantive AI assistance.
3. **Log your AI use** (tool+version, prompts, inputs by type, outputs kept, human checks).
4. **Prefer enterprise or local tools** approved by [INSTITUTION].
5. **Verify claims** and cite original sources, not the model.
6. **No AI for peer review** of confidential manuscripts or proposals.
7. **Protect participants**: IRB approval for AI processing; de-identify first.
8. **Document datasets and models** (Datasheets & Model Cards).
9. **Track risks** (privacy, bias, IP, security, misuse) and mitigations.
10. **Be reproducible**: save prompts, seeds, code, data versions, and environments.

1.3 Roles & accountability

- **PI:** ultimate sign-off on AI use, risk register, and disclosures.
- **Data Steward:** storage, access control, de-identification.
- **Model Owner:** model card, evaluations, updates.

- **Project QA Lead:** verifies logs, prompts, reproducibility bundle.

1.4 How the packet is organized

- **policy.qmd** – Department policy (scope, definitions, roles, permitted/prohibited uses, disclosure, procurement, training, enforcement).
- **sop.qmd** – Stage-by-stage procedures aligned to the research lifecycle.
- **checklists.qmd** – One-page checklists and green/yellow/red lists.
- **templates.qmd** – Disclosure language, IRB snippets, Reviewer attestation, Datasheet & Model Card templates, CSV headers for logs.
- **appendices.qmd** – External references and mappings to national and international guidance.

Tip

Smart defaults for [INSTITUTION]. Enforce a strict ban on public AI use for confidential content, require ICMJE-style disclosure of AI assistance, and adopt NIST AI RMF as the governance spine. Adapt state/funder specifics in *Appendix A*.

2 Department Policy on AI in Research

2.1 3. Principles

1. **Legality & ethics:** Comply with laws, funder rules, publisher policies, and IRB approvals.
2. **Human accountability:** Researchers retain responsibility for all outputs.
3. **Transparency:** Material AI assistance is disclosed.
4. **Privacy & security by design:** De-identify early; use approved systems.
5. **Fairness & quality:** Measure and mitigate bias; validate claims.
6. **Reproducibility:** Preserve artifacts to enable independent verification.

2.2 4. Roles & responsibilities

- **Principal Investigator (PI):** Approves AI use cases; signs risk register and disclosures.
- **Data Steward:** Ensures compliant storage, access control, and de-identification.
- **Model Owner:** Authors and maintains Model Cards; documents evaluation, updates, and limitations.
- **Project QA Lead:** Maintains AI Use Logs, prompt archives, change logs, and reproducibility bundles.
- **Department AI Lead (or designee):** Maintains this policy, reviews exceptions, and coordinates training.

2.3 5. Permitted vs. prohibited uses

2.3.1 5.1 Permitted (with logging)

- Brainstorming, outlining, literature scaffolding on public content.
- Copy-editing nonconfidential text; code linting on toy/synthetic data.
- Summarizing public PDFs with proper citation checks.

2.3.2 5.2 Restricted (require approvals & controls)

- Data labeling/annotation of **de-identified** data.
- Translation of non-sensitive materials.
- Transcription using **enterprise** tools with approved storage.

2.3.3 5.3 Prohibited

- Uploading any **confidential** content to public AI tools.
- Using AI to perform **peer review** of confidential materials.
- Presenting **AI-fabricated data** as empirical observation.
- Generating images or figures that could mislead without explicit labeling.

2.4 6. Disclosure & documentation

All material AI assistance must be disclosed in manuscripts/grants (see templates). Projects must maintain: - **AI Use Log**, **Risk Register**, **Datasheet(s)**, **Model Card(s)**, and a **Reproducibility Bundle** (code, lockfiles, seeds, data access notes, prompt files).

2.5 7. Data governance & privacy

- Apply de-identification at the earliest possible stage.
- Store research data and AI outputs on approved systems.
- Respect licenses and rights (publisher PDFs, test instruments); document TDM legal basis when applicable.

2.6 8. Security & procurement

- Prefer enterprise/private tools approved by [INSTITUTION].
- Vendor vetting is required for any tool touching research data.

2.7 9. Peer review & editorial ethics

- No public AI tools may access confidential manuscripts or grants.
- If a venue permits limited AI assistance, it must be private, logged, and disclosed to the venue.

2.8 10. Training & compliance

- Annual training on AI in research for all researchers and staff.
- Audits may review logs, prompts, risk registers, and artifacts.

2.9 11. Exceptions

Exceptions require written approval from the Department AI Lead and the PI, with documented mitigations and rationale.

2.10 12. Enforcement

Violations may result in corrective actions under [INSTITUTION] policies and sponsor requirements.

3 Standard Operating Procedures (SOPs)

This SOP maps the research lifecycle to concrete steps, artifacts, and gates.

3.0.1 Gate 0 — Project registration (before any AI use)

- File an **AI Use Case** entry: purpose, data types, tools, access, risks, roles.
- Create initial **Risk Register** and **Reproducibility Bundle** skeleton (repo with `env.lock`, `prompts/`, `logs/`).

3.0.2 Stage 1 — Ideas & literature

- Use AI to brainstorm/search; verify against sources.
- Artifact: **AI Use Log** entries; **Source List**.

3.0.3 Stage 2 — Grant/protocol drafting

- Only nonconfidential text may be processed; use enterprise tools.
- Artifact: **Disclosure note** (if AI used for editing), **Access attestations**.

3.0.4 Stage 3 — IRB/ethics & data rights

- Update protocol to reflect AI processing; include consent language.
- Artifact: **IRB-approved language**, **License/TDM memo**.

3.0.5 Stage 4 — Data collection & curation

- De-identify data; produce **Datasheet for Datasets**.
- Artifact: Datasheet v1; **Data License** file; **PII risk assessment**.

3.0.6 Stage 5 — Analysis & modeling

- Use AI for code suggestions/tests; lock seeds and environments.
- Artifact: **Model Card**; **Evaluation report** (accuracy, subgroup fairness, robustness); **Change log**.

3.0.7 Stage 6 — Results verification & reporting

- Independent checks; bias & robustness analyses.
- Artifact: **QA checklist**; **Signed verification** by QA Lead.

3.0.8 Stage 7 — Writing & authorship

- Human-led drafting; disclose AI assistance and verification steps.
- Artifact: **AI Use Statement** in manuscript; prompt archive for major uses.

3.0.9 Stage 8 — Peer review & editorial work

- No public AI use on confidential content.
- Artifact: **Reviewer attestation** (if applicable).

3.0.10 Stage 9 — Publication, sharing & archiving

- Deposit code/data (as permitted) with licenses and metadata.
- Artifact: **Repository DOI**, **README**, **Data/Model Cards**, **AI Use Log** export.

3.0.11 Stage 10 — Deployment & translation

- Define intended use/out-of-scope; user disclosures; monitoring plan.
- Artifact: **Deployment risk assessment**, **User docs**, **Incident log**.

4 Checklists & Allowed Uses

4.1 One-page PI checklist (printable)

☐ AI Use Case registered; roles named.
☐ Tools are enterprise-approved.
☐ IRB reflects AI processing; de-identification complete.
☐ Datasheet(s) and Model Card(s) started.
☐ Risk Register created; mitigations assigned.
☐ Prompts & outputs logged; seeds/environments locked.
☐ Disclosure text prepared.
☐ Bias/robustness tests completed.
☐ Repository prepared for sharing (licenses, README, DOIs).

4.2 Green / Yellow / Red

Green (allowed with logging): brainstorming; literature scaffolding; copy-editing nonconfidential text; code linting on toy/synthetic data; alt-text; captions.

Yellow (approval & controls): summarizing public PDFs; de-identified data labeling; translation of non-sensitive materials; enterprise transcription.

Red (prohibited): confidential manuscript/grant text; identifiable human data to public tools; licensed instruments without permission; automated peer review; AI-fabricated data presented as real; undisclosed AI-generated images.

4.3 Reviewer/editor checklist

☐ I did not use public AI on confidential content.
☐ Any permitted assistance occurred on private, logged systems.
☐ I will not retain manuscript text in external tools.

4.4 QA checklist (analysis & modeling)

- ☐ Evaluation includes subgroup performance.
- ☐ Robustness/shift tests completed.
- ☐ Failure modes documented; limitations section updated.
- ☐ Model Card complete; intended use/out-of-scope defined.

5 Templates

5.0.1 Manuscript AI-use disclosure (short)

We used *[Tool, version]* for *[copy-editing/summarization/code suggestions]* in *[sections]*. Outputs were reviewed and edited by the authors; all accuracy and originality remain the authors' responsibility. No confidential or identifiable data were provided to AI systems.

5.0.2 Grant/IRB language (AI processing of data)

Study data may be processed with machine-learning tools for transcription/annotation/analysis on secure, [INSTITUTION]-approved systems. No public AI services will receive identifiable data. Data will be de-identified prior to any automated processing.

5.0.3 Peer-review attestation (reviewers/editors)

I did not use public AI systems to read, summarize, or draft any part of this review, nor did I disclose manuscript contents to any third-party tool.

5.1 Datasheet for Datasets — template

Dataset name

Version: v0.1
Owners: [Name, email]
Provenance: [Source(s), collection dates]
Licenses/rights: [Link/terms]
Population/coverage: [Who/what/where/when]
Consent & lawful basis: [IRB status, consent language, TDM basis]
Sensitive attributes: [List or N/A]

Known skews/biases: [Describe]
Preprocessing & de-ID: [Methods, date, validator]
Quality checks: [Missingness, noise, audits]
Permitted uses: [Allowed]
Prohibited uses: [Forbidden]
Retention/deletion: [Schedule]

5.2 Model Card — template

Model name

Version: v0.1
Owner: [Name, email]
Intended use: [Scope, users, decisions supported]
Out-of-scope: [Misuse, non-goals]
Training data: [Sources, timeframe, datasheet refs]
Evaluation data: [Datasets, metrics]
Performance: [Overall + subgroup]
Robustness/shift tests: [Methods, results]
Safety mitigations: [Filters, constraints]
Limitations: [Caveats]
Update policy: [Schedule, triggers]
Contact: [CONTACT_EMAIL]

5.3 AI Use Log — CSV header

```
project_id,date,stage,tool,tool_version,prompt_file,input_type,
contains_confidential(boolean),output_kept(desc),human_verification
(desc),reviewed_by
```

5.3.1 Risk Register — CSV header

```
project_id,risk_category,description,likelihood,impact,mitigation,
owner,status,next_review
```

5.3.2 Prompt archive guidance

- Save prompts in `prompts/YYYY-MM-DD_context.txt` .
- For long sessions, export transcripts or maintain a summarized prompt file per analysis step.

6 Appendices & External References

6.0.1 Appendix A — External guidance to align with (curate per your use)

- **National/International:** NIST AI RMF; ICMJE authorship & AI guidelines; discipline-specific reporting (e.g., CONSORT-AI/SPIRIT-AI/TRIPOD-AI in biomed); EU AI Act research exemption vs. deployment obligations; OECD/UNESCO principles.
- **U.S. Federal:** Sponsor and agency rules on AI use for peer review and confidentiality (e.g., NIH); agency public-access plans (article + data).
- **State & Institutional:** State IT AI acceptable-use/procurement; [INSTITUTION] vendor vetting; campus data classification & storage.

Action: Replace this list with citations/links applicable to **[STATE]** and your typical funders (e.g., NIH/IES/NSF). Add any journal-specific policies you frequently encounter.

6.0.2 Appendix B — Mapping table

External rule/guidance	What it says	Our policy hook
[Source]	[Summary]	[Policy section & artifact]

6.0.3 Appendix C — Glossary

Plain-language definitions for AI, GenAI, confidential materials, de-identification, TDM, bias/fairness, robustness, model card, datasheet, etc.