AI in Research: Policy, SOPs, and Templates

Guidance for Responsible, Reproducible AI Across the Research Lifecycle

[INSTITUTION] — [DEPARTMENT]

2025-09-19

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 **Don't upload confidential material** (unpublished manuscripts, grants, identifiable data, licensed instruments) to public AI tools.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.
- **general-doc-analysis** - General analysis of existing documents, sops, and policy on AI use in research.
- **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 AI Research Policy Analysis - General

6.1 Executive summary (what's solid today)

- **Peer review:** U.S. federal funders explicitly **ban** using generative AI to read, analyze, or draft peer-review materials; uploading any application/manuscript content to public AI tools violates confidentiality. Enforced at NIH and mirrored widely. (Grants.gov)
- **U.S. federal governance:** OMB's AI governance memos now run through **M-25-21** (Apr 3, 2025), which **replaced M-24-10**, while **M-25-22** covers AI acquisition. These establish agency-level risk management, inventories, impact controls, and transparency—useful templates for university standards. (The White House)
- **Reporting & authorship:** ICMJE (updated 2024–2025) says **AI tools are not authors**; substantive use must be **acknowledged** and **verifiable**, and journals must instruct reviewers about AI limits. Major publishers (Nature/Science) require disclosure and restrict AI-generated figures/text. (icmje.org)
- **Risk management standards: NIST AI RMF 1.0** (with a 2024 **Generative AI profile**) is the most widely referenced, practical framework for documenting AI risks across the lifecycle; use it as your spine for internal SOPs, audits, and model/data cards. (NIST Publications) ()
- **International law & soft law:** The **EU AI Act** (in force; phased timelines through 2025+) exempts AI **developed and put into service solely for scientific research**, but applies once systems are **placed on the market/put into service**; prohibits certain practices and adds obligations for high-risk uses. UNESCO & OECD principles give high-level guardrails to align local policy with international norms. (EUR-Lex)
- **States (examples):** States are moving on AI governance that can touch public universities: **California EO N-12-23** (GenAI program + procurement guidance), **New York ITS AI acceptable-use policy**, and broader state AI governance (e.g., inventories, restrictions). **Colorado's 2024 AI Act** (developers/deployers) is broader market regulation that may affect research **deployment** to residents. (Governor of California)

6.2 Workflow matrix: what to do & document, and where rules bite

Research stage	Minimum standard (actionable)	What to record/disclose
0) Governance & planning	Adopt **NIST AI RMF 1.0** + GenAI profile as your internal playbook; define roles (PI, data steward, model owner), risk thresholds, and approval gates.	RMF alignment table; risk register; approvals; model/data "owner" of record. (NIST Publications)
1) Idea & literature review	If AI summarizers are used, **retain citations & source passages**; keep a tool-use log.	Tool + version; prompts; output checkpoints; sources verified by human.
2) Grant/protocol drafting	**No AI** may see nonpublic content if sponsor prohibits; keep AI out of confidential sponsor material. Check funder policy.	Statement of AI assistance limited to generic writing aid (if allowed); confirmation none used on confidential text. (Grants.gov)
3) IRB/ethics & data rights	For human data, reflect **GDPR/HIPAA/Common Rule** duties if applicable. If scraping, document legal basis, TDM exceptions, consent language re: AI processing.	Consent language; lawful basis; data minimization; de-ID plan; DUA/DUC terms permitting AI processing. (EUR-Lex)
4) Data collection/curation	Create **Datasheets for Datasets** (who/what/when/how), provenance, and permissible uses; include bias/coverage.	Datasheet; license terms; PII handling; retention/deletion; data quality checks.
5) Model development	Maintain **Model Cards**; record training data, preprocessing, eval sets, metrics, failure modes, and guardrails; run bias/robustness tests; align with RMF "Map/Measure/Manage."	Model Card; risk tests; change log; dependency hashes; reproducible seeds. (NIST Publications)
6) Validation & pre-reg/reporting	For clinical/biomed: follow **CONSORT-AI / SPIRIT-AI / TRIPOD-AI / related** reporting. Outside biomed, use analogous checklists (Model/Data Cards).	Registration; analysis plan; reporting checklist attached to manuscript.
7) Writing & authorship	**Humans write & take responsibility.** If AI assisted drafting, **disclose** which sections and how outputs were verified; **no AI as author.**	"AI use" paragraph in Acknowledgements/Methods; prompts retained; plagiarism checks performed. (icmje.org)

Research stage	Minimum standard (actionable)	What to record/disclose
8) Peer review (as author & reviewer)	**Do not** upload manuscripts or proposals to AI tools. Reviewers must avoid AI assistance unless journal gives a strict, private, compliant workflow.	Reviewer attestation; journal-specific instructions kept on file. (Grants.gov)
9) Publication & sharing	**Public access** (U.S.): follow agency public-access plans (Nelson memo implementation). Share data/code when allowed; include licenses & metadata.	Repository DOIs; metadata; statements on restrictions/sensitive data. (The White House)
10) Deployment/translation	If releasing tools/services to users (esp. EU), check **EU AI Act** scope—research exemption ends when putting into service/on market; follow transparency/high-risk obligations where applicable.	Intended use; user disclosures; conformity route (if applicable); risk mitigations; post-market monitoring plan. (EUR-Lex)

Notes on the evidence base & policies cited above

- NIH peer-review AI ban (binding): reviewers may **not** use LLMs; uploading any application content violates confidentiality. (Grants.gov)
- OMB M-25-21 (2025) now governs agency AI use (M-24-10 **rescinded**); M-25-22 covers acquisition—useful templates for internal governance & vendor vetting. (The White House)
- ICMJE: AI cannot be an author; disclose and take responsibility for any AI assistance; give reviewers AI-use guidance. (icmje.org)
- Nature/Science: disclosure required; restrictions on AI-generated images and some textual uses. (Nature)
- NIST AI RMF + GenAI profile = best-practice backbone for risk & documentation across the lifecycle. (NIST Publications)
- EU AI Act: research exemption (Recital 25) but obligations trigger when systems are **put into service** or **placed on the market**; early prohibitions and penalties phase in during 2025. (EUR-Lex)
- State examples with impact on public universities: CA EO N-12-23 (GenAI program/procurement); NY ITS AI acceptable-use policy. (Governor of California)

6.2.1 "Starter" Standard Operating Procedure (SOP) & documentation kit

Use this immediately; we can refine per your department/university.

A. Governance & registration

- Register each AI use case (project level) with: purpose, data, model(s), access controls, and a named **model owner**. Map to NIST RMF functions (Govern/Map/Measure/Manage). (NIST Publications)
- Maintain an **AI risk register** (privacy, bias, robustness, misuse, IP/copyright), with mitigations and sign-offs at each milestone. (NIST Publications)

B. Dataset documentation

- Create a **Datasheet for Datasets** for every dataset (source, consent, licenses, sensitive attributes, known skews, allowed AI uses).
- Record TDM (text-and-data-mining) legal basis and opt-outs where applicable (esp. EU TDM rules). (EUR-Lex)

C. Model documentation

- Publish **Model Cards** (intended use, data, metrics, failure modes, safety mitigations, update policy).
- Keep full **reproducibility bundle**: code, environment, seeds, data access notes.

D. Authorship & manuscript

- Include an **AI use statement** (template):

  "The authors used <tool & version> to <copy-edit/translate/summarize> sections <X>. Outputs were reviewed and edited by the authors, and all accuracy, originality, and citation responsibilities remain with the authors. No AI systems had access to nonpublic or confidential data/manuscripts."

Align with journal/funder requirements (ICMJE, Nature/Science). (icmje.org)

E. Peer review

- Strictly **no AI** on confidential materials (grant proposals, manuscripts) unless the journal/funder provides a **private, approved** tool and explicit permission. (Grants.gov)

F. Public access & sharing

- Comply with your funder's **public-access plan** (Nelson memo implementation): deposit publications **and** data (or give justified exceptions). Provide machine-readable metadata & PIDs. (The White House)

G. Deployment

- If you release a model/app beyond research, assess EU AI Act/other jurisdictional obligations (transparency, risk management, high-risk conformity). Document "intended use," user notices, and monitoring. (EUR-Lex)

6.2.2 Gaps & where to set your own standard

- **Social/education research outside biomed** lacks CONSORT-AI-style, field-specific reporting. Use **Model/Data Cards + NIST RMF** to fill the gap until discipline-specific extensions emerge. (NIST Publications)
- **State policy patchwork:** campus-level acceptable-use and procurement rules may be stricter than funders' policies—align early with state IT policy (e.g., NY ITS) and your General Counsel. (IT Services)
- **Reviewer guidance** varies by publisher; default to the NIH/ICMJE floor: no public AI tools on confidential content; if any assistance is allowed by a journal, document it and keep it local/private. (Grants.gov)

6.2.3 Suggested deliverables (I can draft next)

1. **Departmental AI in Research Policy** (10 pages): scope, definitions, governance roles, prohibited uses, disclosures, audit.
2. **One-page PI checklist** (by stage).
3. **Templates**: AI use statement; Dataset Datasheet; Model Card; Risk register; Reviewer attestation.

6.2.4 Smart defaults for Auburn/EFLT (today)

- Adopt **NIST AI RMF 1.0** + GenAI profile as the lab/department standard. (NIST Publications)
- Enforce **NIH-style peer-review ban** on AI for all internal and external reviews. (Grants.gov)
- Require **ICMJE-style disclosure** for any AI assistance in manuscripts/grants. (icmje.org)

- For projects touching the EU (data or deployment), track **EU AI Act** timelines and exemptions. (EUR-Lex)
- Align with state-level acceptable-use (e.g., NY ITS example) for university-owned devices/accounts; restrict public AI tools for sensitive data. (IT Services)

6.2.5 Clarifying questions (answer any/all and I'll tailor the next iteration)

1. **Scope of data & populations:** Human subjects only, or also web-scraped/administrative data? Any EU/UK participants (GDPR) or clinical collaborations (HIPAA/FDA)?
2. **Where do you plan to *deploy* models** (beyond publications)—internal dashboards only, or public-facing tools/apps (triggers EU AI Act-like duties)? (EUR-Lex)
3. **Publisher mix:** Which journals/presses do you submit to most (ICMJE-aligned vs. social-science outlets), so I can map exact editorial policies? (icmje.org)
4. **Campus governance:** Do you want the policy written for **your department**, **College of Education**, or as a **university-wide** recommendation that references state IT policy (e.g., CA/NY-style AU)? (IT Services)
5. **Tool posture:** Any enterprise AI tools (private instances) available on campus (e.g., MS Copilot, ChatGPT Teams/Enterprise, Claude for Teams) that we should whitelist for limited uses?

7 Appendices & External References

7.0.1 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.

7.0.2 Appendix B — Mapping table

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

7.0.3 Appendix C — Glossary

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