AI Capability Practice Guide: Research

Practical, Responsible, High-Integrity AI Use for Research & Knowledge Production

---

Who This Guide Is For

This guide is for people engaged in **research and knowledge production** who are already using—or are about to use—AI and want to do so:

- confidently
- responsibly
- with epistemic integrity
- without undermining authorship, methods, or trust
- without exposing themselves, participants, or institutions to avoidable risk

This includes:

- doctoral and postgraduate researchers
- postdoctoral researchers and research fellows
- principal investigators and research leads
- research assistants and analysts
- applied researchers in policy, health, NGOs, and industry

You do **not** need technical AI expertise.
You do **not** need to study AI theory first.
You **do** need to make defensible research decisions under real scrutiny.

This guide assumes **you remain fully responsible** for claims, methods, and interpretations—even when AI is involved.

Who This Guide Is Not For

This guide is **not** designed for:

- automating or outsourcing core research reasoning
- fabricating data, evidence, or citations
- using AI as a ghost author
- bypassing ethics review, supervision, or peer scrutiny
- treating AI output as equivalent to literature, data, or analysis

If you are looking for shortcuts that replace **research judgement**, this guide will feel deliberately uncomfortable.

What You Will Be Able to Do in 30–60 Minutes

By working through this guide, you will be able to:

- decide **when AI is appropriate in research—and when it is not**
- design a clear **human–AI research relationship** for a real task
- recognise **epistemic risk** (hallucination, distortion, false synthesis)
- apply lightweight but credible **research governance**
- document AI involvement in a way you can justify to supervisors, reviewers, or funders
- reflect on AI's effect on your thinking and improve future practice

You will also produce at least **one concrete research artefact** (e.g. an AI-use note or revised workflow) that you can reuse immediately.

FAST START — USE THIS NOW

If you read only one section, read this one.

This Fast Start lets you use AI **responsibly in research** within **10 minutes**, without reading the rest of the guide first.

When to Use This Guide

Use this guide when:

- you feel pressure to use AI because others are doing so
- a task affects **research claims, interpretation, or credibility**
- you are unsure how much to trust an AI-generated synthesis
- the work may be scrutinised by supervisors, examiners, reviewers, or ethics panels
- AI might influence decisions that are hard to reverse
- something “feels off”, but you can’t quite articulate the risk

If none of these apply, formal guidance may not be needed yet.

The 10-Minute Entry Workflow (Research)

Use this sequence **before** opening an AI tool or acting on its output.

Step 1 — Name the Research Task

Write down, in plain language:

> *"I am using AI to help me with: [specific research task]."*

Good examples:

- exploring possible framings of a research question
- identifying themes across papers I have already read
- testing clarity of an argument for a non-expert audience

Avoid vague phrasing like *"help with my research"*.

Step 2 — Decide the Role of AI (Co-Agency Check)

Ask yourself:

- What part of this task can AI **support**?
- What part must remain **human-led**?
- Who is accountable if something here is wrong?

If you cannot answer all three, **stop and refine the task**.

Step 3 — Apply the Research Capability Check

Ask:

- Do I understand how an AI might generate an answer here?
- What would a **confident-sounding but wrong** output look like?
- Where would error cause real harm (credibility, ethics, trust)?

If you cannot identify plausible failure modes, you are over-trusting the output.

Step 4 — Run the Rapid Integrity Screen

Ask:

- Could this distort the literature or misrepresent evidence?
- Could this introduce fabricated or unverified claims?
- Would I be comfortable defending this process in a viva, review, or audit?

If the answer is “no” or “unsure”, slow down and increase verification.

Step 5 — Decide the Action

Choose one:

✅ **Proceed with AI support** (with verification)
🔁 **Revise the task or prompt**
⛔ **Pause and escalate** (consult supervisor, ethics lead, or methods expert)

Document the decision. One sentence is enough.

Worked Example — One Task, Three Outcomes (Research)

Task
Using AI to help shape the literature review section of a paper.

✅ Good Use

- AI suggests possible thematic groupings
- Researcher checks each theme against actual papers
- References are sourced independently
- Final structure is human-designed and defensible

Why this works:
AI supports sense-making without replacing engagement with literature.

⚠ Risky Use

- AI produces a polished synthesis
- Researcher lightly edits language
- Citations are assumed to be correct

Why this is risky:
False coherence and hallucinated references may go unnoticed.

⛔ Unacceptable Use

- AI generates a full literature review section
- References are copied without checking
- Researcher cannot explain why claims are included

Why this fails:
Authorship, evidence, and accountability have been delegated to a system that cannot carry them.

Your First Research Artefact (Create This Now)

Write a short **Research AI Use Note** (3–5 lines):

- **Research task:**
- **Role of AI:**
- **Human responsibility retained:**
- **Key risks considered:**
- **Decision made:**

This single note:

- clarifies your own thinking
- creates an audit trail
- protects you if questions arise later

You have now already **improved your research AI capability**.

How This Guide Works

This guide is designed to be **used**, not read cover-to-cover.

You are expected to:

- dip in when a research situation arises
- apply a tool or checkpoint
- make a judgement
- move on

You are not expected to memorise domains.

The Six Domains as a Living Research Workflow

The AI Capability Framework is built around **six domains**.
In this guide, they operate as a **practical research workflow**, not a theory model.

You will not apply all six every time.
You apply them **as risk and impact increase**.

The research capability flow in practice:

- **AI Awareness & Orientation**
  Understand how AI may mislead in research contexts.
- **Human–AI Co-Agency**
  Decide who does what—and who owns the claim.
- **Applied Practice & Innovation**
  Use AI to explore and test thinking, not replace it.
- **Ethics, Equity & Impact**
  Anticipate effects on people, knowledge, and trust.
- **Decision-Making & Governance**
  Keep research decisions transparent and defensible.

- **Reflection, Learning & Renewal**
  Improve capability over time, not just outputs.

Skipping a domain does not save time.
It usually moves risk downstream, where it is harder to see and harder to correct.

How to Use This Guide Under Time Pressure

Most AI decisions in research are made:

- mid-analysis
- near deadlines
- under publication or funding pressure

This guide is designed for that reality.

If you only have 5–10 minutes

- Use the **Fast Start**
- Clarify co-agency
- Run the integrity screen
- Document the decision

If you have 20–30 minutes

- Identify the relevant situational entry point
- Apply 2–3 domains
- Use one checklist or note
- Capture one reflection insight

If stakes are high

- Work through all six domains
- Focus on ethics and governance
- Prepare documentation for scrutiny

High-impact research requires **slower judgement, not faster automation**.

Stage 2 — How This Guide Works & Situational Entry Points (Research)

HOW THIS GUIDE WORKS

This guide is designed to support **real research decisions under pressure**.

You are **not** expected to move through it linearly.
You are expected to:

- enter when a research situation arises
- apply one or two relevant checks
- make a defensible judgement
- return to the work

This mirrors how research actually happens:
iterative, interrupted, and scrutinised after the fact.

The Six Domains as a Living Research Workflow

The AI Capability Framework consists of six domains.
In this guide, they act as **decision lenses**, not theory labels.

You apply **more domains as risk, impact, or uncertainty increase**.

The Research Capability Flow

- **AI Awareness & Orientation**
  Understand how AI systems can mislead in research contexts.
- **Human–AI Co-Agency**
  Decide who does what—and who owns claims, methods, and interpretation.
- **Applied Practice & Innovation**
  Use AI to explore ideas without shortcutting scholarly reasoning.

- **Ethics, Equity & Impact**
  Anticipate effects on participants, communities, and knowledge itself.

- **Decision-Making & Governance**
  Keep research decisions transparent, auditable, and defensible.

- **Reflection, Learning & Renewal**
  Learn from AI use and refine research practice over time.

Skipping a domain may feel efficient.
It usually **moves risk downstream**, where it is harder to detect and harder to defend.

Why Situational Entry Points Matter in Research

Researchers rarely start with:

> "Which AI capability domain should I apply?"

They start with:

- a looming deadline
- reviewer comments
- uncertainty about evidence
- a supervisor's question
- ethical discomfort
- a result that looks "too clean"

Situational entry points let you start **where the research problem is**, not where the framework begins.

SITUATIONAL ENTRY POINTS — START HERE

Use the entry point that best matches your current situation.
Each one tells you **what to do next** and **which domains to prioritise**.

Entry Point 1 — “I need to move quickly, but I’m uneasy.”

You feel pressure to deliver:

- an analysis
- a draft section
- a response to feedback

AI could help—but something feels risky.

Primary domains to apply

- **Human–AI Co-Agency**
- **Ethics, Equity & Impact**

What to do now

- Clarify which parts of the task must remain human-led
- Identify where errors would undermine credibility
- Increase verification before sharing or submission

Common failure mode

- Letting speed displace epistemic caution
- Treating “draft” outputs as conceptually safe

Entry Point 2 — “I’m not sure AI is appropriate here.”

You could use AI—but may be crossing a boundary:

- authorship
- methods
- interpretation
- ethics approval

Primary domains to apply

- **AI Awareness & Orientation**
- **Decision-Making & Governance**

What to do now

- Ask what the AI can *not* reasonably know
- Identify which judgements must remain human
- Decide whether AI should assist, inform, or be excluded

Common failure mode

- Treating uncertainty as permission to experiment anyway

Entry Point 3 — “The output looks good, but I don’t fully trust it.”

The AI output is fluent, coherent, and persuasive—
but confidence is low.

Primary domains to apply

- **AI Awareness & Orientation**
- **Reflection, Learning & Renewal**

What to do now

- Identify implicit assumptions or missing caveats
- Cross-check key claims manually
- Adjust prompts or workflows for future tasks

Common failure mode

- Mistaking polish for validity

Entry Point 4 — “This affects people beyond me.”

The research output may affect:

- participants
- communities
- policy decisions
- public understanding

Primary domains to apply

- **Ethics, Equity & Impact**
- **Decision-Making & Governance**

What to do now

- Identify who is affected and how
- Examine bias, framing, or representation risks
- Decide what oversight or disclosure is required

Common failure mode

- Treating indirect impact as negligible

Entry Point 5 — “Someone might question how this was produced.”

You anticipate scrutiny from:

- supervisors or examiners
- ethics panels
- reviewers or editors
- funders or auditors

Primary domains to apply

- **Decision-Making & Governance**
- **Human–AI Co-Agency**

What to do now

- Document how AI was used and where humans intervened
- Make accountability explicit
- Avoid reconstructing decisions after submission

Common failure mode

- Trying to justify AI use retroactively

Entry Point 6 — “I want to improve my research practice, not just finish this task.”

You are thinking beyond one project toward:

- doctoral development
- research leadership
- methodological robustness

Primary domains to apply

- **Reflection, Learning & Renewal**
- **Applied Practice & Innovation**

What to do now

- Identify patterns in what AI improves and degrades
- Adjust how you design tasks and prompts
- Capture insights for reuse across projects

Common failure mode

- Repeating convenient workflows without evaluation

How the Rest of This Guide Is Structured

From this point on, the guide moves into the **Core Practice Workflow**.

Each domain section will:

- explain what the domain protects or enables in research
- show how to apply it immediately
- identify common research failure modes
- include a short reflection moment

You can work with:

- one domain
- several domains
- or all six

depending on the research situation you face.

Stage 3 — Core Practice Workflow: Domains 1–3 (Research)

Working Well With AI in Research Contexts

Domains 1–3 help you use AI **without compromising epistemic integrity**, research methods, or authorship.
These domains shape *how* you think with AI—not what AI produces.

Each domain follows the canonical structure:

- **What This Domain Protects / Enables**
- **Apply Now — Key Questions**
- **Tool in Use**
- **Common Failure Modes**
- **Quick Reflection**

DOMAIN 1 — AI Awareness & Orientation (Research Literacy)

What This Domain Protects

This domain protects you from:

- mistaking **plausibility** for **evidence**
- believing AI has **access to knowledge it cannot have**
- accepting **fabricated citations, data, or coherence**
- letting synthetic text distort real literature
- confusing **pattern generation** with **analysis**

Researchers face a unique risk:
AI outputs can appear **academically convincing**, even when wrong.

AI does **not** know—
it predicts.

Awareness ensures you never confuse those two.

Apply Now — Key Questions

Before using or acting on an AI output, ask:

1. **What kind of system am I interacting with?**
   (Predictive text generator, not a reasoning agent.)

2. **What information is it *not* trained on or cannot infer?**
   (Recent publications, proprietary datasets, disciplinary nuance.)

3. **What would a confident-sounding error look like here?**
   (Fabricated citations, overgeneralised claims, missing caveats.)

4. **Which part of this output would cause damage if wrong?**
   (Framings, definitions, relationships, causal suggestions.)

If you cannot identify at least one plausible failure mode, you are over-trusting the model.

Tool in Use — Research Awareness Check

Use this quick template before relying on any AI-generated research text:

AI Awareness Check (Research)

- *What is this AI optimised to do?*
- *What is it not designed to do?*
- *Which parts of my task require evidence the AI cannot provide?*
- *What would a misleading but fluent output look like?*

Write one sentence for each.
This prevents epistemic shortcuts.

Common Failure Modes in Research

- treating AI synthesis as equivalent to reading the literature
- assuming AI “knows” niche or emergent fields
- accepting summaries without tracing claims to sources
- overlooking hallucinated citations that “look right”
- using AI descriptions as conceptual or theoretical definitions

---

Quick Reflection

What did I assume the AI “knew” that it could not reasonably know?
Capture one sentence.
This small habit produces large capability gains over time.

DOMAIN 2 — Human–AI Co-Agency (Research Workflow Design)

What This Domain Protects

This domain protects:

- **authorship**
- **methodological integrity**
- **accountability**
- **interpretive ownership**
- **transparent research practice**

In research, co-agency is **non-negotiable**.
If the boundary between your judgement and AI assistance is unclear, research integrity collapses silently.

AI can assist with:

- idea generation
- exploring multiple framings
- drafting structures
- producing alternative explanations

AI **cannot**:

- judge truth
- determine relevance
- make methodological choices

- interpret data
- justify claims
- carry responsibility

Only the human researcher can.

Apply Now — Key Questions

Ask:

1. **Which parts of this research task can AI support?**
   (Exploration, drafting, counter-arguments.)

2. **Which parts must remain human-led?**
   (Interpretation, evidence evaluation, methodological reasoning.)

3. **Who is accountable if something goes wrong here?**
   (Always you.)

4. **What decisions must be traceable for supervisors, reviewers, or examiners?**
   (How claims were formed, how literature was interpreted.)

Tool in Use — Research Co-Agency Map

A two-minute activity that prevents accidental delegation:

Research Co-Agency Map

- **AI may support by:**
  (e.g. generating alternative framings, proposing structure.)
- **AI may not:**
  (e.g. produce final synthesis, replace reading, justify claims.)
- **I remain responsible for:**
  (method choice, accuracy, interpretation, citation, argument.)

Use this once per meaningful research task.

Common Failure Modes in Research

- accepting AI-generated conceptual categories without justification
- letting AI language shape your argument without noticing
- delegating interpretation of findings (“What does this result mean?”)
- copying structure or coherence that does not reflect actual sources
- mixing literature you read with literature AI fabricated

These failures often remain invisible until challenged by a reviewer.

Quick Reflection

Did I design the research–AI relationship intentionally, or did I let it emerge by default?

One sentence is enough.

DOMAIN 3 — Applied Practice & Innovation (Safe Research Experimentation)

What This Domain Enables

This domain enables **productive, safe experimentation**, allowing AI to:

- broaden conceptual space
- generate alternative framings
- support early drafting
- reveal assumptions
- strengthen reasoning through contrast

This is where creativity lives—
but research innovation must **never** shortcut scholarly rigor.

Innovation is valuable only when **methods, evidence, and interpretation remain intact**.

Apply Now — Key Questions

Ask:

1. **Does AI genuinely improve this research task?**
   (Not every task needs AI.)

2. **What would "good" look like *without* AI?**
   (This prevents dependency.)

3. **Am I using AI to explore, or to avoid thinking?**
   (Be honest.)

4. **Is this experiment low-risk, or could it distort evidence?**
   (Exploration is fine; replacing literature is not.)

Tool in Use — Safe Research Experimentation Prompt

Use this when you want AI support without epistemic shortcuts:

> *"Generate alternative framings or conceptual groupings for this topic.*
> *Do not produce citations or claim accuracy.*
> *I will review for validity, evidence, and alignment with accepted literature."*

This creates **creative divergence** without blurring evidence boundaries.

Common Failure Modes in Research

- using AI to generate citations or summaries instead of reading
- allowing AI's fluency to override uncertainty or nuance
- collapsing conceptual categories based on AI's convenience
- using AI to "decide" what interpretations are strongest
- letting AI produce analysis-like text that seems valid but is not grounded in data

---

Quick Reflection

What improved because I used AI—and what did not?
Capture one insight for reuse.

Domains 1–3 in Practice (Research)

Together, Domains 1–3 establish a research environment where:

- **AI supports thinking** without replacing evidence
- **the researcher remains the epistemic agent**
- **innovation strengthens**, not weakens, research quality

These domains are necessary but not sufficient.
They prepare the foundation for Domains 4–6, which address:

- ethics
- impact
- governance
- long-term capability

These are the guardrails preventing research from drifting into irresponsible use.

Stage 4 — Risk, Responsibility & Renewal: Domains 4–6 (Research)

Governing Impact, Protecting Integrity, and Sustaining Research Capability

Domains 4–6 come into play whenever research outputs:

- affect others,
- enter the public domain,
- influence decisions,
- shape knowledge,
- or require long-term defensibility.

These domains ensure AI use remains **ethical, transparent, equitable, and audit-ready**.

We keep the canonical structure for each domain:

- **What This Domain Protects / Enables**
- **Apply Now — Key Questions**
- **Tool in Use**
- **Common Failure Modes**
- **Quick Reflection**

DOMAIN 4 — Ethics, Equity & Impact (Epistemic Responsibility)

What This Domain Protects

This domain protects:

- fairness in knowledge production
- representation of communities and participants
- epistemic justice
- the integrity of research claims
- avoidance of harm—material, reputational, or interpretive

In research, ethics is **not limited to human subjects**.
AI can distort:

- whose voices are centred or erased
- how phenomena are framed
- which explanations appear “normal”
- how marginalised knowledge is represented

AI amplifies dominant perspectives unless intentionally constrained.

Apply Now — Key Questions

Before sharing, submitting, or embedding AI-generated content, ask:

1. **Who is represented or misrepresented by this framing?**
2. **Does the AI output oversimplify lived experience or sensitive topics?**
3. **Which groups might be disadvantaged or misinterpreted?**
4. **Does this output encode normative assumptions that need unpacking?**
5. **If someone were harmed by this output, how difficult would it be to reverse?**

These questions matter even in desk-based and theoretical research.

Tool in Use — Ethical Impact Scan (Research)

Use this 90-second scan on any AI-assisted research output:

- **Accuracy Risk:** What might be misleading or wrong?
- **Bias Risk:** Whose perspectives are missing, flattened, or treated as universal?
- **Power Risk:** Who has less ability to challenge or contest this framing?
- **Consequence Risk:** What harm would be hardest to reverse?

One sentence per risk is enough.

Common Failure Modes in Research

- trusting AI's "neutral" tone as objective
- summarising sensitive literature without cultural or contextual grounding
- letting AI implicitly define what counts as legitimate knowledge
- using AI explanations for topics involving inequality or lived experience
- assuming non-empirical work has no ethical dimension

Quick Reflection

If I were part of the community represented here, what would concern me most?
Capture one insight.

DOMAIN 5 — Decision-Making & Governance (Research Integrity Infrastructure)

What This Domain Protects

This domain protects:

- transparency in how claims are formed
- defensibility of research decisions
- auditability for supervisors, reviewers, examiners, funders
- traceability of human vs AI contribution
- clarity of authorship and accountability

Research governance is not bureaucracy.
It is the infrastructure that ensures **trustworthiness**.

AI use must be:

- declared
- justified
- reviewable

especially when it shapes interpretation, structure, or claims.

Apply Now — Key Questions

Before finalising a research decision influenced by AI, ask:

1. **Did AI influence this decision directly or indirectly?**
2. **Is this a low-risk support use or a high-stakes judgement?**
3. **What must be documented for future scrutiny?**
4. **Would I stand by this process in a viva, peer review, or audit?**
5. **Does this require supervisor, ethics, or team review?**

If the answer to #5 is unclear, assume it **does require review**.

Tool in Use — Research Decision Transparency Log

This is a lightweight governance artefact.
Use it for any meaningful research decision AI touches:

Decision Transparency Log (Research)

- **Decision being made:**
- **Role of AI:** inform / suggest / draft / structure / generate
- **Human decision-maker:**
- **Verification performed:**
- **Escalation required? (Y/N):**
- **Notes:**

This log can be a notebook entry, a file, or a project page—
but it **must exist**.

Risk Thresholds (Research Escalation Triggers)

Escalate or require additional oversight when AI is used in:

- public-facing research communication
- abstracts, literature reviews, or methods sections
- interpretation of results
- policy-relevant research
- participant-sensitive topics
- safety-critical or regulatory domains
- anything affecting institutional reputation or funding integrity

When in doubt: **escalate**.

Common Failure Modes in Research

- failing to disclose AI-mediated drafting in manuscripts
- no record of how AI shaped the literature review
- mixing human reading and AI hallucinations in the same synthesis
- relying on memory to reconstruct decisions
- assuming “everyone uses AI” means governance is optional

These failures surface during peer review—rarely earlier.

Quick Reflection

Could I explain and justify this decision six months from now?
If not, governance is insufficient.

DOMAIN 6 — Reflection, Learning & Renewal (Sustaining Research Capability)

What This Domain Sustains

This domain sustains:

- long-term research judgement
- epistemic maturity
- adaptive expertise
- resilience against overreliance
- continuous improvement of methods and workflows

AI tools change rapidly.
Without reflection, capability **decays** even as tools "improve."

Reflection prevents research practice from:

- calcifying into routine
- drifting into dependency
- repeating errors
- losing visibility of how AI shapes thinking

Apply Now — Key Questions

After meaningful AI use, ask:

1. **What improved because of AI?**
2. **Where did AI distort or oversimplify?**
3. **What will I change in the workflow next time?**
4. **What new risk did I discover?**
5. **Do I need to update my co-agency boundaries?**

These questions take under five minutes.

Tool in Use — Mini Reflection Cycle (Research)

A simple loop to capture and embed research learning:

Reflect → Adjust → Reapply

- **One insight:** What did I learn?
- **One improvement:** What will I change in my workflow?
- **One boundary reset:** What will I not allow AI to do next time?

This cycle strengthens capability more than hours of reading.

Common Failure Modes in Research

- repeating prompts uncritically
- assuming past success implies future reliability

- ignoring how AI subtly shapes research language
- failing to revisit decisions after feedback or review
- treating reflection as optional instead of capability-building

---

Quick Reflection

What assumption about AI shifted for me during this task?

Capture one sentence—this is your renewal boundary.

Domains 4–6 in Practice (Research)

Together, these domains:

- protect people and communities
- maintain research legitimacy
- ensure auditability
- promote ethical reasoning
- sustain long-term capability
- guard against epistemic drift

They turn good use into **responsible research practice**.

The full capability cycle in research practice is:

Awareness → Co-Agency → Practice → Ethics → Governance → Reflection → Renewal

Skipping steps increases risk.
Revisiting steps deepens capability.

Stage 5 — Capability Self-Check, Worked Research Scenario & Operating Model

This final stage turns the guide into a **usable system**.

You will:

- locate where your research capability is strongest and weakest
- see how all six domains operate together in a real research scenario
- leave with a **personal Research AI Operating Model** you can reuse immediately

Nothing here is evaluative.
Everything here is **practical orientation**.

PART A — RESEARCH AI CAPABILITY SELF-CHECK

This is **not** an assessment and **not** a scorecard.
Its purpose is to answer one question:

"Where do I need to focus next to use AI safely and well in research?"

Complete this in **under five minutes**.

Domain 1 — AI Awareness & Orientation

Ask yourself honestly:

- I understand that AI produces *plausible text*, not verified knowledge
- I routinely question synthesis and summaries rather than trusting fluency
- I can identify where hallucination or false coherence would cause damage

☐ Mostly yes ☐ Mixed ☐ Mostly no

If mostly no:
Pause AI use in literature or framing tasks until awareness improves.

Domain 2 — Human–AI Co-Agency

Consider:

- I explicitly decide what AI supports vs what remains human-led
- I remain comfortable explaining my role if questioned
- I never allow AI to own interpretation, judgement, or claims

☐ Mostly yes ☐ Mixed ☐ Mostly no

If mixed:
Clarify co-agency boundaries before continuing.

Domain 3 — Applied Practice & Innovation

Reflect:

- AI helps me explore ideas rather than shortcut reasoning
- I test multiple framings rather than accept first outputs
- I can imagine completing the task without AI

☐ Mostly yes ☐ Mixed ☐ Mostly no

If mostly no:
Shift from output extraction to safe experimentation.

Domain 4 — Ethics, Equity & Impact

Ask:

- I consider who may be affected by AI-shaped research outputs
- I am alert to bias, misrepresentation, or epistemic harm
- I pause if consequences would be hard to undo

☐ Mostly yes ☐ Mixed ☐ Mostly no

If mostly no:
Ethics must move earlier in your workflow.

Domain 5 — Decision-Making & Governance

Check:

- I can explain how AI influenced key research decisions
- I document AI use when it affects interpretation or outputs
- I know when escalation or review is required

☐ Mostly yes ☐ Mixed ☐ Mostly no

If mixed or no:
Introduce lightweight documentation immediately.

Domain 6 — Reflection, Learning & Renewal

Finally:

- I deliberately reflect on AI use, not just outcomes
- I adjust prompts, boundaries, or workflows over time
- I learn from errors rather than hiding them

☐ Mostly yes ☐ Mixed ☐ Mostly no

If inconsistent:
Adopt a simple reflection habit.

Interpreting Your Self-Check

- Gaps in **Domains 1–2** → slow down and rebuild foundations
- Gaps in **Domains 4–5** → increase oversight before scaling
- Gaps in **Domain 6** → capability will stagnate

Capability grows by **rebalancing**, not maximising.

PART B — WORKED RESEARCH SCENARIO (END-TO-END)

This scenario mirrors **common, high-stakes research reality**.

Scenario

You are writing the **literature review and discussion** for a paper or thesis chapter.
You are under time pressure and considering significant AI support.

Domain 1 — Awareness in Action

You recognise:

- AI can produce convincing synthesis
- AI does not verify literature
- Hallucinated citations are plausible

You therefore:

- avoid asking for “final reviews”
- treat AI outputs as hypotheses, not claims

Domain 2 — Co-Agency in Action

You define boundaries:

AI may support by:

- proposing thematic groupings
- suggesting alternative structures
- rephrasing for clarity

AI may not:

- supply unchecked references
- decide what constitutes a gap
- interpret findings

You remain accountable for every claim.

Domain 3 — Applied Practice in Action

You:

- conduct database searches yourself
- read and annotate papers
- use AI to explore alternative framings
- select structures that reflect *actual* literature

AI expands thinking; it does not replace reading.

Domain 4 — Ethics & Impact in Action

You check:

- which voices or traditions might be flattened
- whether the synthesis privileges dominant narratives
- whether lived experience is oversimplified

You deliberately include complexity and caveats.

Domain 5 — Governance in Action

You create a **Research AI Use Note** stating:

- AI assisted with structuring and language exploration
- all citations were sourced independently
- interpretation remained human-led

If questioned, the trail exists.

Domain 6 — Reflection in Action

After submission or review, you ask:

- Did AI actually improve clarity or rigour?
- Where did it tempt shortcutting?
- What will change next time?

You adjust your workflow accordingly.

What This Scenario Shows

Responsible AI use in research:

- redesigns thinking processes
- preserves epistemic ownership
- makes decisions defensible
- improves future practice

AI capability is not visible in the output alone.
It is visible in **how the work was produced**.

PART C — YOUR RESEARCH AI OPERATING MODEL

This is your **repeatable system** for AI-supported research.

Complete once. Revisit periodically.

---

1 My Valid Reasons to Use AI in Research

Examples:

- exploring alternative framings
- drafting early structures
- stress-testing arguments
- improving clarity of non-substantive text

My reasons:

---

2 My Co-Agency Rules

AI may:

AI may not:

I always retain responsibility for:

---

3 My Ethical Red Lines

I pause or stop AI use when:

- _____
- _____

---

4 My Governance Triggers

I document or escalate AI use when it affects:

- literature interpretation
- analysis or discussion
- public or policy-facing outputs
- sensitive topics or communities

---

5 My Reflection Habit

After meaningful AI use, I ask:

- What helped?
- What distorted or misled?
- What will change next time?

My reflection cadence:
☐ after every task ☐ weekly ☐ per project

---

6 My Renewal Commitment

To keep research capability strong, I will:

- revisit assumptions quarterly
- update boundaries as tools change
- remain alert to new epistemic risks

THE RESEARCH COMMITMENT

I use AI to support—not substitute—scholarship.
I retain ownership of judgement, methods, and claims.
I design AI use intentionally and transparently.
I document decisions so my work can be trusted.
I reflect so my capability grows, not erodes.