A Practical Guide to Using AI in Research Administration: The TaMPER Framework

Large Language Models (LLMs) like ChatGPT are powerful tools that can streamline many research administration tasks. However, to move AI from a personal productivity hack to a trusted, scalable part of your organization's workflows, requires a structured approach. The **TaMPER** framework is a simple mental model to guide you through this process, ensuring your results are reliable, transparent, and secure.

T is for Task (What is the specific job to be done?)

Before you start, clearly define your goal. Don't just think "I'll use AI on this." Instead, specify the exact administrative problem you want to solve. A well-defined task is the foundation for getting a useful result.

Common Tasks for Research Administrators:

- **Summarizing:** Condensing a 50-page funding announcement into a one-page summary for a PI.
- **Information Extraction:** Pulling out all key dates, budget limits, and reporting requirements from a grant agreement.
- **Classification:** Sorting a list of research proposals by scientific discipline or funding agency.
- **Drafting:** Creating a first draft of a routine report, email, or standard operating procedure.

M is for Model (Which AI tool will you use?)

"Model" is the technical term for the specific AI engine you're using (e.g., GPT-4, Claude 3.5, Llama 3). The most important decision here is choosing between a public tool and a private, secure one.

- **Public/Closed-Source Models (e.g., standard ChatGPT):** These are easy to access but often use your data to train their models. **Never input sensitive, confidential, or proprietary information** (like unpublished research data, budgets, or personal information) into these public tools unless it is explicitly approved for these uses by your institution.
- **Private/Open-Source Models:** These can be hosted on secure, local servers. They are the only safe choice for working with confidential data.

Key Question: *Is the information I'm using confidential?* If the answer is yes, a public model is not an option.

P is for Prompt (How will you ask the AI?)

A "prompt" is the instruction you give the AI. The quality of the AI's output depends entirely on the quality of your prompt. Vague questions get vague answers.

A great prompt includes:

1. **Role:** Tell the AI who it should be. *"Act as a university compliance officer."*
2. **Context:** Provide the necessary background. *"...reviewing a new grant award from the National Science Foundation."*
3. **Task:** Give a clear, specific instruction. *"...identify all sub-award requirements and reporting deadlines."*
4. **Format:** Specify how you want the answer. *"...present the information in a table with two columns: 'Requirement' and 'Deadline'."*

E is for Evaluation (How will you check the AI's work?)

This step is critical for building trust and moving AI into official workflows. AI models can make mistakes, invent facts ("hallucination"), or misinterpret instructions. **You are the expert, not the AI.** A rigorous process for verifying the output is non-negotiable for reliable organizational use.

How to Evaluate:

- **Fact-Check:** Always compare the AI's output against the original source document. Verify every date, dollar amount, and policy detail.
- **Assess Quality:** Does the response make sense? Is it well-written and relevant? Did it follow all parts of your prompt?
- **Check for Consistency:** If you run the same prompt multiple times, do you get similar results? Inconsistency is a red flag for a repeatable process.

R is for Reporting (How will you document your process?)

This is how you turn a one-time success into a documented, repeatable, and scalable organizational workflow. Keeping a clear record ensures that you or a colleague can reproduce the results later, builds institutional trust, and makes the process auditable.

Your documentation should note:

- The specific **Task** you performed.
- The **Model** you used (e.g., "ChatGPT-4, accessed on July 10, 2025").
- The exact **Prompt** you used to get the final result.
- A brief summary of your **Evaluation** process (e.g., "All deadlines were manually cross-checked with the original grant agreement.").

By using the **TaMPER** framework, you can move beyond individual experimentation and begin to integrate AI as a strategic, reliable, and auditable tool within your organization, scaling its benefits safely and effectively.

Deep Dive: Defining the "Task"

The first and most critical step in the TaMPER framework is defining the **Task**. How you define the job-to-be-done directly determines the quality, relevance, and accuracy of the AI's output. A poorly defined task leads to generic, unhelpful, or incorrect results, while a well-defined task produces a specific and valuable work product.

Tips for Defining a Clear Task

- **Start with the End in Mind:** Before you write your prompt, ask yourself: "What specific piece of information or document do I need when this is done?" The answer to that question is your true task.
- **Use Action Words:** Frame your task using clear, unambiguous action verbs that describe a specific operation. Good examples include:
    - **Summarize:** Condense a long document.
    - **Extract:** Pull out specific pieces of information (e.g., names, dates, numbers).
    - **List:** Create a bulleted or numbered list of items.
    - **Compare:** Find similarities and differences between two pieces of text.
    - **Classify:** Assign items to predefined categories.
    - **Format:** Re-structure information into a table, JSON, or another format.
    - **Draft:** Create an initial version of a document.
- **Be Specific and Concrete:** Avoid vague or subjective instructions. The more specific your definition, the better the AI can meet your expectations.

Common Errors in Task Definition (and How to Fix Them)

Error 1: The Task is Too General

A vague task forces the AI to guess what you find important, and its guess will likely be wrong. This is the most common reason for disappointing results.

- **The Fix:** Add specific constraints and focus to your task definition. Tell the AI *what* to look for and *why*.

Error 2: The Task is Too Big or Complex

Asking an AI to perform multiple, distinct operations in a single request is a recipe for failure. The model may miss steps, mix up instructions, or produce a low-quality response for all parts of the task.

- **The Fix:** Break down a complex job into a series of smaller, sequential tasks. This is called **task decomposition**. Have the AI perform one operation, check the result, and then use that output as the input for the next step. This gives you more control and leads to a more accurate final product.

Table 1. Example Tasks in Research Administration

Scenario	Improper Task Definition (The Error)	Proper Task Definition (The Solution)
Reviewing a New Funding Announcement	"Read this funding announcement and tell me what's important." *(Error: Too General)*	"Extract the 'Eligibility Criteria' section from the attached funding announcement. Then, create a bulleted list of all specific requirements for the Principal Investigator (e.g., degree, years of experience, institutional affiliation)."
Checking a Proposal for Compliance	"Check this proposal to see if it follows the rules." *(Error: Too General)*	"Compare the 'Budget Justification' section of the attached proposal against the following sponsor requirements: [paste specific requirements]. Identify and list any budget line items in the proposal that are explicitly defined as 'unallowable costs' by the sponsor."
Preparing a Summary for a PI	"Review this grant agreement and write an email to the PI with the details." *(Error: Too Big/Complex)*	**Task 1:** "Extract the following from the attached grant agreement: Project Start/End Dates, Total Award Amount, and all Reporting Deadlines." **Task 2:** "Using the information below, draft a concise email to a Principal Investigator summarizing the key terms of their new award. [Paste output from Task 1]."
Synthesizing Reviewer Comments	"Summarize these reviewer comments." *(Error: Too General)*	"Review the three attached reviewer critiques. First, classify each comment into one of the following categories: 'Strength,' 'Minor Weakness,' or 'Major Weakness.' Then, present the results in a table with the comment and its category."

By investing a few extra moments to properly define and scope your task, you dramatically increase the likelihood of getting a useful, accurate, and reliable result from any AI model.

Deep Dive: Choosing the Right "Model"

Choosing a "Model" is about more than just picking an AI brand name. For a research administrator, it's about making a responsible choice that protects sensitive data and aligns with the practical limitations of the technology. Your primary considerations are **security** and **input readiness**.

The #1 Decision: Protecting Sensitive Data

Before you upload or paste *any* information into an AI tool, you must ask: **"Does this data contain sensitive information?"** In a university setting, this includes, but is not limited to:

- **Student Records (FERPA):** Any personally identifiable information from a student's educational record.
- **Health Information (HIPAA):** Any protected health information (PHI) related to research subjects or employees.
- **Intellectual Property (IP):** Unpublished research findings, novel ideas, patent disclosures, or any proprietary data from the university or a sponsor.
- **Controlled Unclassified Information (CUI):** Unclassified information that requires safeguarding, common in federally funded research (e.g., defense, energy).
- **Confidential Business Information:** Budgets, salaries, or other non-public financial or administrative data.

The Golden Rule of Model Selection:

If you would not be allowed to email the information on a public server or post it on a website, you **cannot** use a public AI model.

- **Public Models (e.g., standard, free versions of ChatGPT, Gemini, etc.):** These services often use your data to train their systems. Submitting sensitive information is a data breach. It violates university policy, federal law, and research sponsor agreements.
- **Private, Secure Models:** These are AI models hosted on secure university or approved third-party servers. They are "sandboxed," meaning your data stays within the secure environment. **This is the only acceptable choice for working with sensitive data.**

Always consult your institution's IT or research security office to understand what secure AI options are available to you.

Common Errors: Practical Input Limitations

Even with the right model, you can run into problems if your source document isn't ready for the AI. Here are two common technical hurdles:

Error 1: The Document is Too Big (Exceeding the "Context Window")

An AI model has a limited "short-term memory," known as its **context window**. This is the maximum amount of text (your instructions + your document) it can consider at one time. If you try to upload a 200-page grant proposal, the model will either reject it or simply ignore the parts that don't fit in its window, leading to incomplete or inaccurate results.

- **The Fix:** Don't try to make the AI "read" the whole document at once. Instead, break the document into smaller, relevant chunks or sections. Copy and paste only the specific section you need for your task (e.g., just the "Budget Justification" or the "Biographical Sketches"). This respects the context window and gives the AI a more focused task.

Error 2: The PDF is Not Machine-Readable (It's an Image, Not Text)

Sometimes a PDF is not a text document; it's just a picture of a document (e.g., an old, scanned file). If the computer can't see the text, neither can many AI models.

- **The Quick Test:** Open your PDF. Can you click and drag your mouse to highlight a specific sentence?
  - **If YES**, the PDF is machine-readable.
  - **If NO**, it's likely an image-based PDF, and an AI model cannot process it.
- **The Fix:** You will need to use **Optical Character Recognition (OCR)** software to convert the image of the text into actual, machine-readable text. Many tools (including Adobe Acrobat Pro) have this feature. Once converted, you can copy and paste the text into your chosen AI model.

Table 2. A Simple Guide to Model Selection

Step 1: Assess Your Data	Step 2: Assess Your Document	Action to Take
Contains SENSITIVE Information (FERPA, HIPAA, IP, CUI, etc.)	*Doesn't matter*	**STOP.** Use only an approved, private, and secure AI model provided by your institution. Do not proceed with a public tool.
Contains NO Sensitive Information	The document is **too long** for the context window.	Pre-process your document. Break it down into smaller, relevant sections before giving it to the AI.
Contains NO Sensitive Information	The PDF is **not machine-readable** (image-based).	Pre-process your document. Use OCR software to convert it to text first.

Contains NO Sensitive Information	The document is a reasonable length and is machine-readable.	You may proceed cautiously with a public AI model but always evaluate the output for accuracy.

Deep Dive: Evaluating the AI's Output

Evaluation is the quality control step that separates professional, reliable AI use from a risky gamble. It is the formal process of verifying the AI's output to ensure it is accurate, appropriate, and trustworthy. For an organization to confidently scale AI into its workflows, every AI-assisted task must be subject to evaluation.

The rigor of your evaluation should directly correspond to the risk associated with a bad answer. A poorly drafted email is an inconvenience; an inaccurate compliance report is a major liability.

Match Your Evaluation to the Risk

Before you evaluate, assess the risk of the task. What are the consequences if the AI gets it wrong?

Risk Level	Example Tasks	Required Evaluation Rigor
Low Risk	• Drafting a routine internal email. • Brainstorming ideas for a presentation. • Summarizing a non-critical news article.	**Spot-Check:** A quick review by the user for clarity and general correctness is sufficient.
Medium Risk	• Creating a first draft of a report. • Summarizing a funding announcement for a PI. • Extracting key dates for an internal calendar.	**Detailed Verification:** The user must perform a careful, line-by-line comparison against the source document to verify all facts, figures, and dates.
High Risk	• Analyzing a grant agreement for compliance risks. • Checking a budget against sponsor policies. • Summarizing patient safety protocols for a clinical trial.	**Expert Review:** A subject matter expert (who may or may not be the user) **must** review and approve the AI's output before it is used. The AI's work is treated as a draft prepared by a junior assistant.

The Pillars of a Quality Answer

Evaluating an AI's output goes beyond a simple "right" or "wrong." For tasks without a single correct answer (like a summary or a risk analysis), you need to assess multiple components of quality.

1. Accuracy (For Factual Tasks with a "Ground Truth")

This applies when there is a single, verifiable correct answer.

- **Question:** Did the AI extract the correct budget number? Is the deadline date correct?
- **How to Evaluate:** Direct comparison with the source document. The answer is either 100% correct or it is wrong.

2. Reproducibility (or Precision)

This is crucial for creating a standardized, repeatable process.

- **Question:** If I run the same prompt on the same document tomorrow, will I get a substantially similar result?
- **How to Evaluate:** Run the prompt 2-3 times. While the exact wording may differ slightly, the core substance of the answer should be the same. Significant variation indicates an unreliable process.

3. Faithfulness (For Interpretive Tasks with No "Ground Truth")

This is the most important metric when evaluating summaries, analyses, or other generated text.

- **Question:** Is the AI's response grounded *only* in the information I provided? Or has it invented facts ("hallucinated") or pulled in outside knowledge?
- **How to Evaluate:** Read the AI's summary or analysis. For every key point it makes, ask, "Can I point to the specific sentence in the source document that supports this?" If you can't, the AI has failed the faithfulness test.

4. Compliance & Quality

This assesses how well the AI followed your instructions.

- **Question:** Did the AI adhere to all the constraints in my prompt? Is the output in the correct format? Is it coherent, logical, and clearly written?
- **How to Evaluate:** Use your prompt as a checklist. Go through each instruction (Role, Context, Instruction, Constraints, Format) and see if the output meets the criteria.

A Practical Evaluation Workflow

1. **Assess the Risk:** Determine if the task is low, medium, or high risk to decide how much scrutiny is needed.
2. **Check for Accuracy:** If the task involves facts and figures, verify every single one against the source.
3. **Check for Faithfulness:** If the task is a summary or analysis, confirm that every statement in the output is supported by the source document.
4. **Check for Compliance:** Review your prompt and ensure the AI followed all instructions and formatting requirements.

5. **Perform Expert Review (for High-Risk Tasks):** The final output must be signed off by a human expert who takes responsibility for its correctness.

Systematic evaluation is what builds trust. It transforms AI from a clever novelty into a dependable tool that your organization can rely on for important administrative tasks. Once common tasks have been evaluated and documented (the next step), they can become trusted parts of your organizational workflows and the frequency and scope of evaluation can be reduced (or even automated).

Deep Dive: Reporting Your Process

Reporting is the final step that elevates an AI-assisted task from a personal shortcut to a transparent, repeatable, and trustworthy organizational workflow. It is the process of documenting your methodology so that a colleague could reproduce your results, a supervisor can trust your process, and your future self can remember exactly what you did.

If you want to scale the benefits of AI across your team or department, a simple, consistent reporting practice is non-negotiable. It creates an auditable "paper trail" that builds institutional confidence.

Why Bother with Reporting?

- **Reproducibility:** A colleague should be able to follow your report like a recipe and achieve the same result. This is the cornerstone of a scalable workflow.
- **Transparency & Trust:** When you document your process, you are showing your work. This builds trust with PIs, department heads, and compliance officers who need to be confident in the data you provide.
- **Training & Onboarding:** A well-documented process is an excellent training tool for new team members, helping them get up to speed on best practices.
- **Process Improvement:** When you have a record of what works (and what doesn't), you can refine and improve your AI-powered workflows over time.

The Four Essential Components of a Report

Your report doesn't need to be long, but it does need to be complete. It should capture the key decisions you made at each step of the TaMPER framework.

1. **Task Definition:**
   - State the specific goal in a single, clear sentence.
   - *Example: "The task was to extract all PI eligibility requirements from the specified funding announcement."*
2. **Model Details:**
   - Name and version of the model used (e.g., GPT-4, Claude 3.5).
   - Date of access (since models are constantly updated).

- o **Crucially:** Specify whether it was a **Public Model** or an approved **Private/Secure Model**.
- o *Example: "The task was performed using the university's secure instance of Llama 3 on July 10, 2025."*
3. **Full Prompt Text:**
   - o Copy and paste the **exact, final prompt** that produced the desired output. Do not summarize or paraphrase it.
   - o *Example: "The following prompt was used: 'Act as a pre-award research administrator...'"*
4. **Evaluation Summary:**
   - o Describe the method you used to verify the output.
   - o State the result of the evaluation.
   - o *Example: "Evaluation was performed via a line-by-line comparison of the AI's output against Section C of the source document. All five eligibility requirements were extracted with 100% accuracy."*

Example: A Simple Workflow Report

Here is an example of a concise report for a common administrative task. This could be saved in a shared folder or attached to the work product itself.

AI-Assisted Workflow Report

Component	Details
Date:	July 10, 2025
Operator:	Jane Doe
Task:	To extract all reporting deadlines from the fully executed "XYZ Corp Sponsored Research Agreement" (Doc ID: #12345) and format them into a table.
Model Used:	University Secure AI Server (GPT-4 Instance). Accessed July 10, 2025.
Prompt Used:	Act as a senior research administrator. You are reviewing the attached, fully executed sponsored research agreement from XYZ Corp. Your task is to identify every reporting requirement mentioned in the document. Extract the name of the report (e.g., "Quarterly Technical Report") and its due date. Provide your final response as a two-column table with the headers 'Report Name' and 'Due Date'. Do not include any other information.
Evaluation:	**Method:** The output table was compared against the original contract (Sections 5.1, 5.2, and Appendix B). **Result:** The AI correctly identified all 4 reporting deadlines and formatted them as requested. The output was 100% accurate.

This simple act of documentation is what makes AI a true organizational capability. It ensures that the powerful work you do is not only effective but also transparent, trusted, and built to last.