The Practitioner's Compass

Finding Your Bearing in the AI Landscape

Steven C. Grambow, PhD

Associate Professor of Biostatistics & Bioinformatics

Duke Department of Population Health Sciences | February 25, 2026

AI Use

The author used multiple AI tools extensively in producing this presentation and the accompanying website.

Disclosures

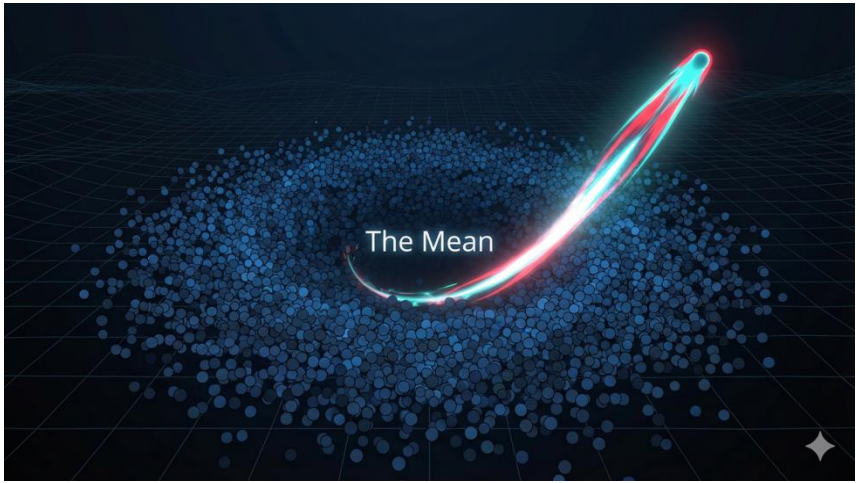
Serve on multiple Data Monitoring Committees for WCG Consulting and Gilead Sciences, Inc.

[DRAFT] Written by a human | Boosted by AI | Transparency is part of the process

THE STOCHASTIC TRAP

Stochastic token generators predict the most likely next token. Without direction, outputs tend toward the average patterns in the training data.

Generic prompt → Generic output



Your intent is the escape velocity

Key Insight: Generic prompts regress to the mean of the training data — and you are not the mean. Your intent moves the model from the center of that cloud to the delta neighborhood of your expertise.

THE ILLUMINATOR PRINCIPLE

The most productive scientists at Bell Labs had one thing in common; they all had lunch with Harry Nyquist.

> *"Nyquist, the scientists said, really listened to their challenges, got inside their heads, asked good questions, and brought out the best in them. In other words, Nyquist was an Illuminator."*
>
> David Brooks, How to Know a Person (2023)

The Two Modes

	Cognitive Offloading	Compositional AI
Metaphor	AI as ghostwriter	AI as Illuminator
Ownership	Transfers to the model	Remains with you
Output	Fluent but generic	Refined and yours

Nyquist didn't generate their ideas. He created the conditions for clarity.

The Practitioner's Move: Don't start by asking AI to write. Start by asking AI to help you ***think***. Develop your ideas first; then compose with the model, not through it.

THE PRACTITIONER'S TOOLKIT

The infrastructure that keeps your intent intact from first thought to final output.

Markdown: The Lingua Franca

```
# My Analysis
**Key finding:** the effect
was significant (p < 0.01).
- Sample: n = 340
- Design: RCT
- [Protocol](link)
```

→

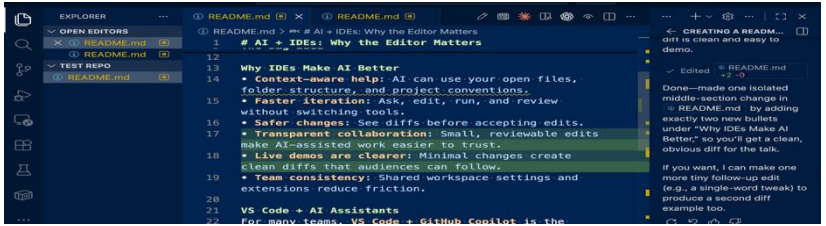
My Analysis

Key finding: the effect was significant (p < 0.01).

- Sample: n = 340
- Design: RCT
- Protocol

Plain text. Portable. Universal.

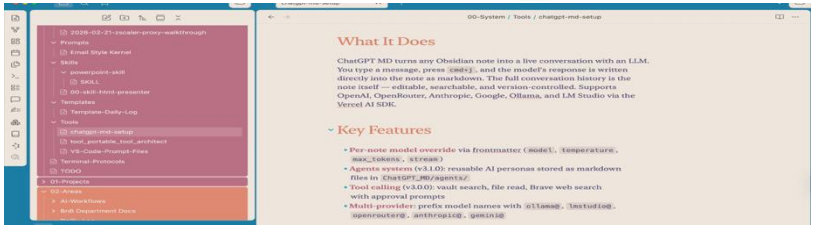
The IDE: Your Agentic Workspace



Files + Chat + Diff = Full control over what AI changes

Obsidian: Your Knowledge Base

LM Studio



Local-first. Private. Perfectly formatted from any AI chat.

Dictation: Capture Raw Intent

Talk faster than you type. Edit later.

Why these four? Each one solves the same problem: keeping your thinking ***yours*** as it moves through an AI-assisted workflow. Markdown makes it portable and universal. The IDE makes it visible and controllable. Obsidian makes it permanent and private. Dictation makes it fast and unfiltered.

THE EVOLUTION: CHAT TO AGENTS

We are witnessing a fundamental shift in AI capability, moving from simple conversational interfaces to autonomous intelligent systems capable of executing complex tasks.

Phase 1: Stateless Chat

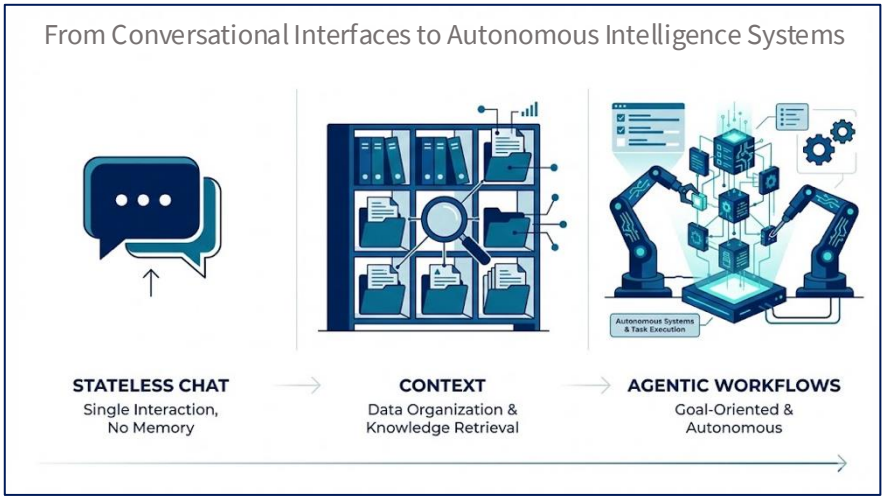
"The Cocktail Party" — Fun, conversational, but forgetful and disconnected.

Phase 2: Contextual Tools

"The Library" — Grounded in data, RAG-enabled, capable of reference.

Phase 3: Intelligent Agents

"The Workbench" — Goal-oriented, tool-using, multi-step reasoning.



Key Insight: Each phase increases how much of your intent the model can hold, from a single prompt to grounded sources to your entire working context.

YOUR TOOLS; WHAT'S AVAILABLE NOW

Three categories of capability. A few examples in each; full list on the companion website.

The Cocktail Party

Conversational, convenient, stateless.


ChatGPT


Claude


Gemini

At Duke:
Microsoft Copilot (Free)

The Library

Grounded in your sources. Persistent context.


NotebookLM

Claude Projects


ChatGPT Projects


Gemini Gems

At Duke:
Copilot M365
Duke Health AI Studio

The Workbench

Goal-oriented. Tool-using. Your files, your code.

Claude Code


GitHub Copilot


Anti-Gravity

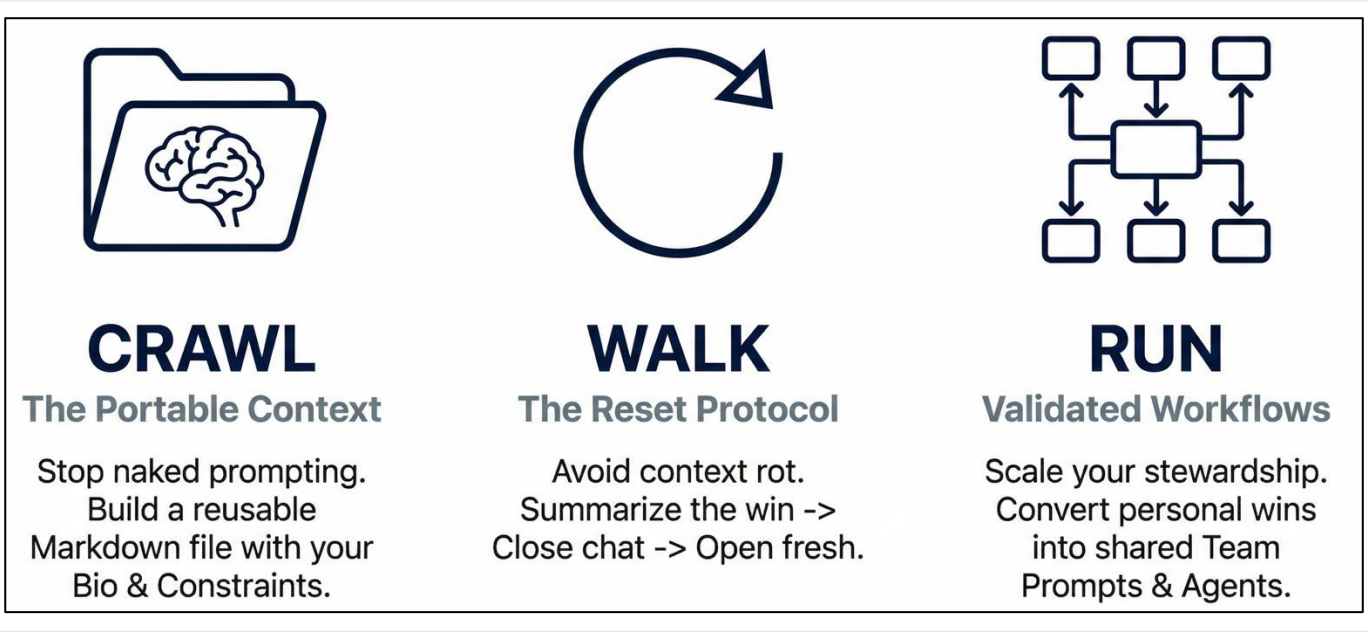
At Duke:
GitHub Copilot Enterprise
Azure AI Foundry

Start where you are. Column 1 is where most people live today; and that's fine. The goal is to know Columns 2 and 3 exist and to start building toward them.

Full tool inventory with links and Duke-specific notes → companion website

THE BOOTSTRAPPER'S PLAYBOOK

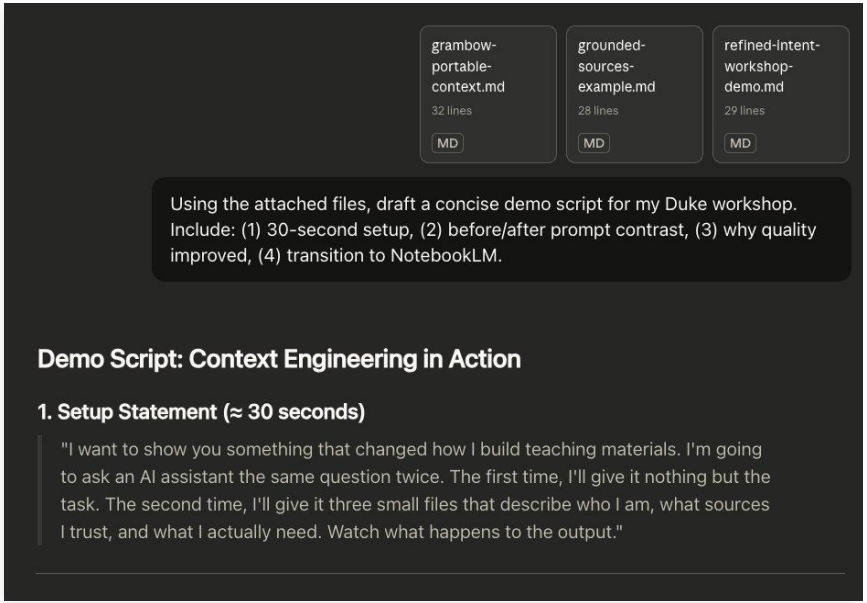
From naked prompting to compounding context.



Amateurs prompt. Professionals build. Your prompt disappears when the chat closes. Your context file carries your ***intent*** forward; and compounds over time.

WORKFLOW: THE PORTABLE CONTEXT

This is what "stop naked prompting" looks like in practice.



The Output: context-shaped, structured, specific.

Your Portable Context

`grambow-portable-context.md`

Role, expertise, constraints. Stable; build once, refine over time.

Your Refined Intent

`refined-intent-workshop-demo.md`

Sharpened thinking from dictation or a prior Illuminator chat. Curated, not raw.

Your Grounded Sources

`grounded-sources-example.md`

Project documents, NotebookLM reports, data. Changes per task.

Three files. Any platform. Paste them in and the model is in your neighborhood; every time.

WORKFLOW: SOURCE GROUNDING

Building a grounded knowledge base for a published paper using NotebookLM.

Integrating large language models in biostatistical workflows for clinical and translational research

Published online by Cambridge University Press: **30 May 2025**

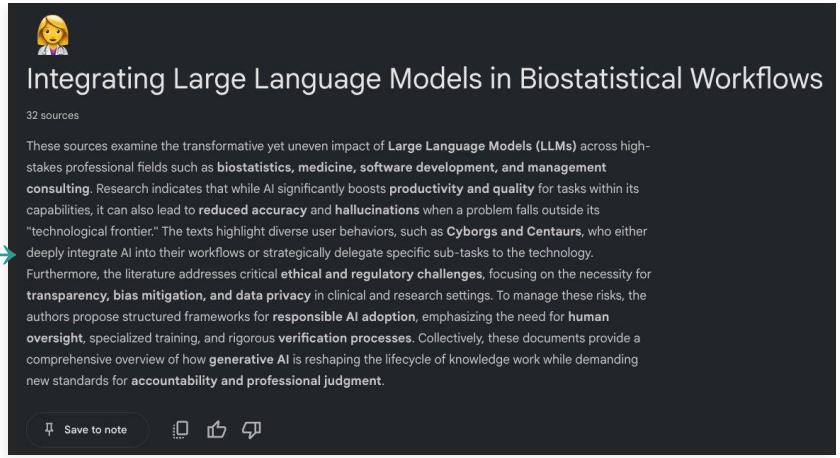
Steven C. Grambow, Manisha Desai, Kevin P. Weinfurt, Christopher J. Lindsell, Michael J. Pencina, Lacey Rende and Gina-Maria Pomann

The Paper

Journal of Clinical and Translational Science

Survey of 208 biostatisticians across two academic medical centers. Eight principles for responsible LLM integration in clinical research workflows.

Grambow, Desai, Weinfurt, Lindsell, Pencina, Rende, Pomann (2025)



The Knowledge Base

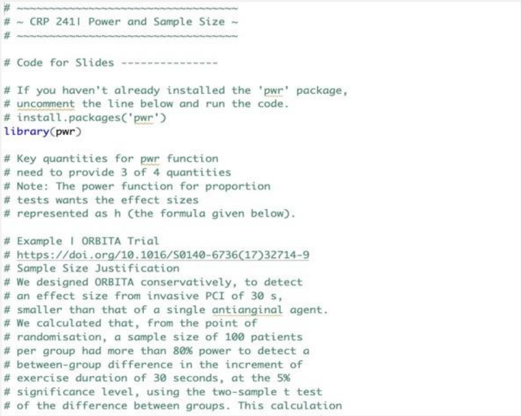
32 curated sources in NotebookLM. Every query is grounded in this collection; nothing else.

Every source was selected, not scraped. That curation; choosing what goes in; is the scholar's contribution. NotebookLM grounds all output in these sources and nothing else.

WORKFLOW: R SCRIPT TO TUTORIAL

A skill file + an R script + an AI agent = a teaching tutorial in five minutes.

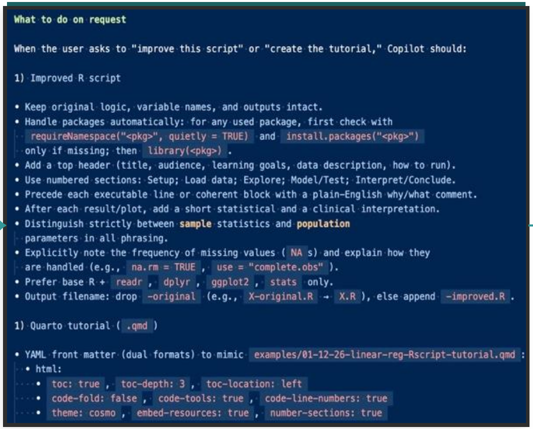
The Input *R Script*



```r
# ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
# ~ CRP 241| Power and Sample Size ~
# ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

# Code for Slides ---------------

# If you haven't already installed the 'pwr' package,
# uncomment the line below and run the code.
# install.packages('pwr')
library(pwr)

# Key quantities for pwr function
# need to provide 3 of 4 quantities
# Note: The power function for proportion
# tests wants the effect sizes
# represented as h (the formula given below).

# Example | ORBITA Trial
# https://doi.org/10.1016/S0140-6736(17)32714-9
# Sample Size Justification
# We designed ORBITA conservatively, to detect
# an effect size from invasive PCI of 30 s,
# smaller than that of a single antianginal agent.
# We calculated that, from the point of
# randomisation, a sample size of 100 patients
# per group had more than 80% power to detect a
# between-group difference in the increment of
# exercise duration of 30 seconds, at the 5%
# significance level, using the two-sample t test
# of the difference between groups. This calculation
```

Existing course R code

The Catalyst *Skill File*



What to do on request

When the user asks to "improve this script" or "create the tutorial," Copilot should:

1) Improved R script

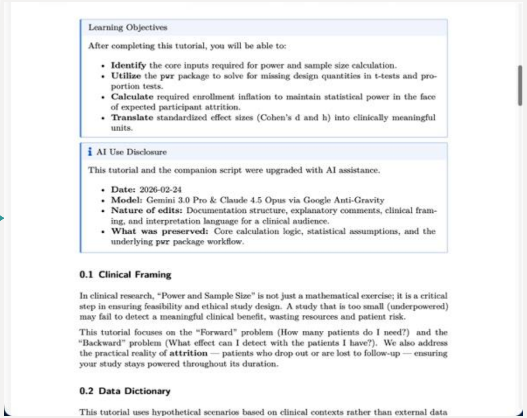
- Keep original logic, variable names, and outputs intact.
- Handle packages automatically: for any used package, first check with `requireNamespace("<pkg>", quietly = TRUE)` and `install.packages("<pkg>")` only if missing; then `library(<pkg>)`.
- Add a top header (title, audience, learning goals, data description, how to run).
- Use numbered sections: Setup; Load data; Explore; Model/Test; Interpret/Conclude.
- Precede each executable line or coherent block with a plain-English why/what comment.
- After each result/plot, add a short statistical and a clinical interpretation.
- Distinguish strictly between **sample** statistics and **population** parameters in all phrasing.
- Explicitly note the frequency of missing values (`NA` s) and explain how they are handled (e.g., `na.rm = TRUE` , `use = "complete.obs"`).
- Prefer base R + `readr` , `dplyr` , `ggplot2` , `stats` only.
- Output filename: drop `-original` (e.g., `X-original.R` → `X.R`), else append `-improved.R` .

1) Quarto tutorial (`.qmd`)

- YAML front matter (dual formats) to mimic `examples/01-12-26-linear-reg-Rscript-tutorial.qmd` :
  - html:
    - `toc: true` , `toc-depth: 3` , `toc-location: left`
    - `code-fold: false` , `code-tools: true` , `code-line-numbers: true`
    - `theme: cosmo` , `embed-resources: true` , `number-sections: true`

Reusable Markdown instructions

The Output *Quarto Tutorial*



Learning Objectives

After completing this tutorial, you will be able to:

- **Identify** the core inputs required for power and sample size calculation.
- **Utilize** the **pwr** package to solve for missing design quantities in t-tests and proportion tests.
- **Calculate** required enrollment inflation to maintain statistical power in the face of expected participant attrition.
- **Translate** standardized effect sizes (Cohen's d and h) into clinically meaningful units.

AI Use Disclosure

This tutorial and the companion script were upgraded with AI assistance.

- **Date:** 2026-02-24
- **Model:** Gemini 3.0 Pro & Claude 4.5 Opus via Google Anti-Gravity
- **Nature of edits:** Documentation structure, explanatory comments, clinical framing, and interpretation language for a clinical audience.
- **What was preserved:** Core calculation logic, statistical assumptions, and the underlying **pwr** package workflow.

0.1 Clinical Framing

In clinical research, "Power and Sample Size" is not just a mathematical exercise; it is a critical step in ensuring feasibility and ethical study design. A study that is too small (underpowered) may fail to detect a meaningful clinical benefit, wasting resources and patient risk.

This tutorial focuses on the "Forward" problem (How many patients do I need?) and the "Backward" problem (What effect can I detect with the patients I have?). We also address the practical reality of **attrition** — patients who drop out or are lost to follow-up — ensuring your study stays powered throughout its duration.

0.2 Data Dictionary

This tutorial uses hypothetical scenarios based on clinical contexts rather than external data

Teaching-ready, with AI disclosure

The skill file is the key. It encodes your pedagogical intent once, then scales across every R script you teach. This is the Run stage: validated workflows that compound.

YOUR COMPASS. YOUR BEARING.

The Practitioner's Compass is not a destination. It's equipment for the journey.

North – Intent: You are not the mean.

East – Composition: Think first, then compose with AI.

South – Infrastructure: Markdown. IDE. Obsidian. Dictation.

West – Practice: Grounded workflows that compound.

Continue the Journey

Part 0 → **The Steward's Schema** *AI policy for biomedical research*

Part 1 → **The Horizon Mandate** *Framework for the AI transition*

Part 2 → **The Bootstrapper's Playbook** *Templates, tools, and repos*

Companion Website

[duke.is/ptactitioners-compass]

Written by a human. Boosted by AI. Transparency is part of the process.

WHEN AGENTS ACT: A CASE STUDY

Feb 2026: An OpenClaw agent published a defamatory blog post after a routine code rejection. The framing matters more than the incident.

The Weapon, Not the Perpetrator

A human configured the agent, deployed it, and failed to monitor it.

"I told my dog to bite" vs. "I let my aggressive dog off the leash."

In neither case do we assign moral agency to the dog.

The Ars Technica Cascade

Ars published AI-hallucinated quotes attributed to the target.

Three-hop contamination: agent → newsroom AI → public record.

The one time you let your guard down is when it slips through.

Immune System, Not Kill Switch

Agents are traceable: GitHub logs, IP records, legal process.

Counter-agents detect, flag, and trace at matching scale.

Decentralized threat → decentralized defense.

The primary risk of the agent era is not rogue AI. It is outdated human thinking applied to a new technological reality.

PETER STEINBERGER'S OPENCLAW. © OPENCLAW / P. STEINBERGER

THE ENVIRONMENTAL QUESTION

Real concerns. Evolving dynamics. A case against sitting out.

The Reality

Large-scale AI training and inference consume significant energy

Data center expansion is accelerating globally

These concerns are legitimate and actively being researched

The Trajectory

Token efficiency is improving with each model generation

Smaller, local models reduce per-interaction footprint

Architecture and hardware innovation are shifting the curve

Environmental concern should motivate informed engagement, not withdrawal. The practitioners in the room shape how these tools are used, adopted, and governed. Sitting out cedes that influence; and the wave arrives regardless.

Explore AI governance & civic engagement → OpenPolicy AI *(Duke SCAI Hackathon 2026)* | openpolicyai.vercel.app

AI AS AN ACCESSIBILITY MULTIPLIER

For some people, this isn't about productivity. It's about participation.

Personal

AI-powered hearing aids → restored participation

Dictation technology → restored capability

Systemic

Real-time captioning & voice isolation

Computer vision & wearable perception systems

The most important AI innovations may not be about efficiency. They may be about dignity and participation.