AI at Your Side

The Student's Guide to Smarter Learning

Forthcoming from Oxford University Press

Sebastian Galiani
Raul A. Sosa

Prosperity Summer Learning Week
The World Bank | July 15, 2026

Book's website

AI Is Already Everywhere

The question is not whether people use it, but how

92%
of UK undergrads use AI tools
Freeman, 2025

85%
of US students use GenAI for coursework
Flaherty, 2025

75%
of knowledge workers already use AI
Microsoft & LinkedIn, 2024

AI can make you *sound* smarter without making you **think** better. Most people use AI for speed. Very few use it for depth.

This book teaches the difference: how to use AI as a **thinking partner**, not a shortcut.

The Dilemma: *Apocalittici* vs. *Integrati*

The Doomsayers

AI will bring mass unemployment, algorithmic control, and the collapse of human agency. Every new technology is a threat.

The Enthusiasts

AI will cure disease, solve climate change, and democratize knowledge. Innovation is unambiguous progress.

Umberto Eco, *Apocalittici e Integrati* (1964): every major cultural innovation splits audiences into pessimists and optimists.

Between these poles lies a simple truth: we do not yet know.

Uncertainty is not an excuse for paralysis. It calls for careful experiments and intellectual honesty.

Two Dystopias

Brave New World

Aldous Huxley, 1932

A **soft** tyranny that pacifies through **pleasure** and numbs with entertainment. AI makes it more personalized and addictive.

1984

George Orwell, 1949

A **hard** tyranny built on **fear**, surveillance, and censorship. Authoritarian regimes already deploy AI for policing and propaganda.

The Paradox: In fleeing Huxley's seductions, we empower Orwellian mechanisms at home. Too much regulation slows AI's development. Too little leaves its risks unmanaged. The task is to promote it within the right limits.

The Third Path: AI on Our Side

- ✔ Lead in AI development. Do not cede the field.
- ✔ Anchor AI in humanist values.
- ✔ Make AI serve people, not the other way around.

We do not need to become post-human. We need to become *better* at being human. That is what putting **AI on our side** really means.

Three Principles That Guide the Book

Augmentation, Not Replacement

The machine extends human capacity. Judgment and creativity remain yours.

Appropriate Automation

Know when to delegate and when to engage. Protect the thinking.

Ethical AI Literacy

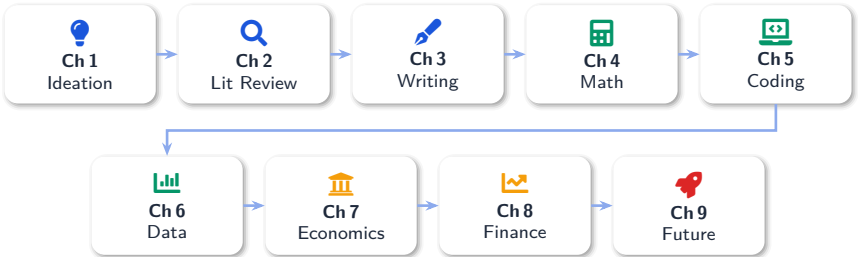
Recognize limitations and biases. Acknowledge your use of AI. Ask yourself: what am I missing?

🎓 AI literacy goes beyond technical skill. It is a **political, educational, and cultural project** that belongs in every classroom and every institution.

Everything the Book Does

A whole discipline, chapter by chapter

The Book at a Glance



+ Appendix: *Understanding AI*, the technical foundations

Each chapter: a **domain-specific harness** (a reusable, rule-bound prompt) + **worked examples** (weak vs. strong prompts) + a **case study**.

Chapter 1: AI for Ideation & Creative Thinking

1

The Ideation Harness

A structured prompt that tells AI your context, defines what counts as a good idea, and requires specific output: research question, constructs, data plan, empirical strategy, and limitations. Keeps AI from generating vague suggestions.

Key techniques

Radiating questions · Perspective-shifting · Dimensional unpacking · Stakeholder simulation

From generic to executable: a vague idea about "public debt" evolves into a fully scoped research plan with testable hypotheses, data sources (IMF, World Bank), and robustness checks.

The point is not the perfect harness. It is building a process where you understand what each line does and how to adapt it.

Chapter 2: AI-Powered Research & Literature Review

The Literature Review Harness

A five-step prompt: explore the field, frame your question, connect it to the live debates, challenge the easy consensus, and focus your contribution, so AI maps the literature without writing it.

Discovery & synthesis

Query engineering · Jargon decoder · Elicit for evidence tables · Synthesis across debates

Three reliability risks

1. Hallucinated citations (fabricated papers)
2. Misrepresented findings (overstated claims)
3. Premature convergence (hidden disagreements)

Remedy: always verify. Ask *where authors genuinely disagree* and what methodological choices drive conflicting conclusions.

Chapter 3: AI for Writing & Academic Productivity

The Writing Harness

A prompt that encodes your thesis, evidence base, and a section-by-section map of which claims go where and what supports them. Forces AI to produce a thesis-driven outline instead of a generic one.

Named frameworks

CREO argument maps · Micro-contracts · Concession-turn · Three-zone integrity

Integrity framework: three zones (supportive / conditional / out-of-bounds), provenance trails, verification-first citations, the "too-close test."

None of these tools wrote the student's paper. They built the discipline that let her write it well.

Chapter 4: AI in Mathematics & Statistics

Core principle: You think first, then bring your thinking to the AI. Math is where AI is most powerful *and* most dangerous.

Coverage

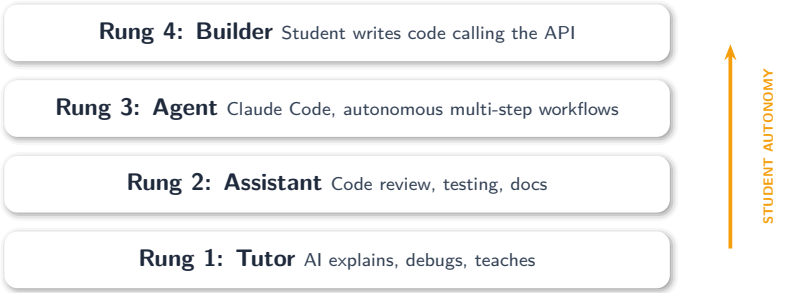
Algebra, calculus, linear algebra · Differential equations · Probability & Bayes · Hypothesis testing & CIs · Proof writing & formal reasoning

Three AI modes

Diagnosis: find where reasoning broke
Model-building: develop intuition before errors occur
Auditing: submit completed work for gap-checking

Protégé effect: student *explains* a method to AI, which flags what is right, wrong, or missing.

Chapter 5: AI in Coding & Computational Thinking



Key insight: the most dangerous AI errors are not crashes but *plausible outputs that answer the wrong question*.

Harnesses: debugging harness, code-review harness (five dimensions), **system prompt as programmatic harness** in API calls.

Chapter 6: AI in Data Analysis & Visualization

Core thesis: The gap between having data and understanding data is *analytical*, not technical. Charts are **arguments**, not neutral windows.

Workflow

Data preparation & merge traps · Exploratory analysis · Visualization refinement · Critical reading of charts

Visual deceptions exposed

Spaghetti charts → the key series lost in the tangle
Dual-axis charts → same data, three stories
Simpson's paradox → aggregate vs. subgroup reversal

Seeing data clearly is itself an analytical skill. The software does not do that for you.

Chapter 7: AI in Economics

7

Formal models are not obstacles to bypass with AI. They are structures to walk through *with* AI.

Three models, one method

1. Public Debt Dynamics (fiscal equations)
2. Becker's Time Allocation (Lagrangian)
3. Spence's Signaling (game theory)

Five ways to work a model

Active correction · Calibration & limits · Interpretation coaching · Stress-testing · Multi-perspective

You derive it yourself. AI checks your steps, translates the intuition, and stress-tests which assumption the result really rests on.

Chapter 8: AI in Finance & Business Strategy

8

The posture: Finance is the rare field where AI cannot predict the answer for you. Treat it as a **reasoning partner** that interrogates your assumptions, never a calculator that hands you numbers.

Worked through AI

Read a 10-K in an afternoon (Costco)
Stress-test a DCF valuation (Ferrari)
Pre-mortem & Monte Carlo on the outcomes
Moat analysis (Greenwald & Kahn)

The Finance Harness

Five moves on any decision: **Decision** · **Assumptions** · **Counter-thesis** · **Sensitivity** · **Blind spots**.

AI never gives live prices or betas. Primary numbers come from primary sources. You verify, then reason.

Chapter 9: AI and the Future of Learning

9

You are the variable. AI has made passing a course nearly free. Because everyone holds the same tool, what you learn now depends on the thinking posture only you bring.

The fork

Every use of AI forks two ways: produce the assignment, or **build the capability**. One leaves a grade, the other leaves you changed.

What you bring

The four capacities AI cannot supply: **knowledge** · the **will** to do the hard part · **judgment** · **metacognition**.

Cognitive debt: the accumulating cost of work done in your name but not in your head. It comes due later, once the tool is gone.

In Practice: AI as a Tool, Not an Objective

In our own classrooms

The classroom philosophy: one-third AI-integrated, two-thirds core material. AI amplifies the ability to learn and to produce research.

What changes

- Students learn to prompt with discipline
- AI use graded on *process*, not just product
- Hallucination checks are mandatory
- An AI Process Log replaces "show your work"

And a six-hour course

We also teach a **six-hour course on the book**. Hands-on, we work through the whole method on real economic problems: prompting, harnesses, agents, and the API.

A Companion Series: Reading Classic Papers in Economics

Rebuilding the classics *with* AI, one paper at a time

Do not ask AI for the summary. **Rebuild the argument with it.** In a public series, we take one classic paper at a time and derive it from the ground up, with the model as a sparring partner that checks every step.

Hayek

The use of knowledge in society

Becker

The allocation of time

Grossman

Health as capital

The series makes the book's method public: you understand a model only when you can **rebuild** it, not when you can quote its last page.

How It Actually Works

Prompting, harnesses, agents, and the API

Prompting: The Skill That Unlocks Everything

Every interaction is only as good as the prompt behind it

The foundational move

Before harnesses, agents, or code, there is the **prompt**. Learning to prompt well is the difference between a vague answer and a rigorous one.

Anatomy of a strong prompt

Context. Who you are, what you know
Role. What the model should act as
Task. The precise question you want answered
Constraints. Scope, length, assumptions
Format. The exact output you want back
Verification. Flag speculation, cite sources

A prompt is not a search query. It is an **instruction set**. The more of your own thinking you put in, the more the model gives back.

Master this, and everything downstream (harnesses, agents, the API) is prompting made **reusable** and **automated**.

From Vague to Precise

The same question, two prompts, two very different answers

A weak prompt

```
"Tell me about public debt in
emerging markets."
```

→ A generic encyclopedia entry. No thesis, no numbers you can use, no way to check it.

A strong prompt

```
"Act as a public-finance economist.
For these six countries, assess
whether external-debt exposure
amplifies fiscal stress.  Use IMF /
WB data.  Define terms.  Cite every
figure.  Flag speculation."
```

→ Structured, sourced, checkable, country-specific.

The strong prompt does three things at once: it sets a **role**, imposes **constraints**, and demands **verification**. Make that reusable, and you have a **harness**.

The Core Innovation: The Harness

A strong prompt, made reusable

Definition

A **harness** is a set of **rules, routines, and constraints** that keep the AI focused, consistent, and honest about what it does not know. Every chapter provides domain-specific harnesses.

A prompt is one-off

Even a strong prompt is written, used once, and rebuilt for the next task. The discipline lives in your head and leaves with you.

A harness is reusable

The same structure, saved and reused. It carries your standards, such as defining terms, citing sources, and flagging speculation, into every task.

The book provides **ready-to-copy harnesses** for ideation, research, writing, math, coding, data work, and domain applications.

Anatomy of a Harness

Five parts turn a one-off prompt into a reusable instrument

What a harness encodes

Context & role. Yours, and the model's
Task. The precise question, reusable
Constraints. Scope, assumptions, limits
Output contract. What shape to return
Verify. Define terms, cite, flag doubt

The public-debt research harness

```
Act as a public-finance economist.
CONTEXT: debt exposure vs. fiscal stress.
DATA: 6 countries, IMF/WB, 2 months.
OUTPUT per idea:  question, hypothesis, plan,
  strategy, limitation + fix, feasibility.
VERIFY: cite every figure, flag speculation.
```

Written once, it holds every project to the same rigor.

A Harness for Every Task

The book ships ready-to-copy harnesses, chapter by chapter

Ideation

Context, a good-idea definition, and a required output: question, data plan, strategy, limits.

Writing (CREO)

Each section mapped to a Claim, its Reasons, the Evidence, and the Objections it must answer.

Finance

On any decision: Decision, Assumptions, Counter-thesis, Sensitivity, and Blind spots.

Literature Review

Five steps: explore, frame, connect to debates, challenge the consensus, focus your contribution.

Math

The protégé effect: you explain a method, the model grades what is right, wrong, or missing.

Model-reading

Your question, the model in your own words, what you understand, and exactly where you are stuck.

Same discipline, tuned to each domain. Copy one, adapt it, make it yours.

A Harness Drafts a Paper

You set the argument. It assembles the draft

A harness turns the model into a disciplined collaborator. You set the thesis, the evidence rules, and the checks. It does the assembly: a first draft that is **structured, sourced, and auditable**, ready for you to refine.

Section by section

A thesis-driven outline: each section bound to its claim, the evidence that backs it, and the objection it answers.

Honest by construction

Every claim carries its source, every gap is marked, and nothing is asserted that the evidence does not support.

At one country, this is a harness. At scale, it becomes DRIL, the same method run on a loop across a whole population of units.

From Chatbots to Agents to API

Chatbot

ChatGPT, Claude, Gemini

You prompt, AI responds.
The harness keeps it honest.

Agent

Claude Code, Cursor

AI plans and executes multi-step work. You supervise.

API

OpenAI, Anthropic, Google

You write code that calls AI.
From user to *builder*.

Agent example: "Inspect this dataset, clean it, run a fixed-effects regression, produce a chart." Claude Code does it all, but you verify every result.

API example: A Python script sends 200 abstracts through a strict system prompt, extracting hypothesis, method, and findings into a CSV.

Agents and the Future of Productivity

If AI raises the productivity of an economy, it will be through **agents**. Not a faster chatbot, but systems that run **entire workflows** in research, in production, and in government.

Research

Datasets, literature, and analyses, end to end.

Production

Whole processes reorganized around agents.

Government

The document-heavy public work that agents are good at.

 **And this is only software.** As the same plan-and-act loop moves into robotics, it reaches the **physical economy**, the next frontier for productivity and development.

Agents at Scale

One instrument, every unit

Deep Research on a Loop (DRIL)

Afonso, Galiani, Gálvez & Sosa (2026)

The idea

Building datasets from primary sources is one of the costliest tasks in empirical economics. **DRIL** applies a **fixed research instrument** across a mapped **unit space** (countries × years), producing structured, evidence-backed records, one unit at a time, on a loop.

Two stages

Design. A human writes the instrument, a codebook of what to measure, its coding rules, and an evidence policy for sources and citations.

Implementation. The agent runs that instrument on every unit, tracking gaps and uncertainty explicitly.

One real run

A 2025 update of the **Global Tax Expenditures Database** for **8 Latin American & Caribbean countries**: **129 sources** and **136 evidence records**, at the cost of **a few hours of research-assistant time**.

Why a Loop Beats a Single Agent

Bespoke answers each time, or one dataset across every unit

One agent, one answer

Ask an agent to research a country and you get a bespoke answer. Ask about the next country and the format drifts, the sources differ, the coding is no longer comparable.

The same instrument, on a loop

The identical codebook runs on every unit. Values are coded the same way, every source is logged, every gap is flagged. The result is a **dataset**, not a pile of memos.

That is what makes agent-produced data **comparable, auditable, and reproducible** across countries, the three properties a development institution cannot do without.

Wherever the information is systematic and public (laws, regulations, budgets, policy documents), a dataset can be built this way.

For the Bank: A Country Report, on Demand

The same loop, pointed at a report instead of a dataset

The loop that built tax-expenditure data across eight countries can draft an **auditable country brief** for **every country the Bank covers**, all to one template. Today an analyst assembles this by hand, over weeks.

What the brief covers

Macro-fiscal position (growth, balance, debt/GDP)
Debt sustainability and external exposure
Tax expenditures and forgone revenue
Institutional indicators (rules, transparency)
Risks and red flags

Each field is its own research problem: find the source, read it, code the value, document the evidence. Multiply by every field, country, and year.

A fixed instrument turns that from weeks of manual work into a loop.

How the Loop Builds It

One instrument, every field, every country

1. The instrument

A codebook of the report's fields, the coding rules, and an evidence policy for sources.

2. The run

The agent fills every field, for every country, from primary sources, coding each to rule.

3. The record

Every figure traced to a source, every gap flagged. It ships with its own audit trail.

This is DRIL, pointed at a report instead of a dataset. A person designs the instrument and audits the result. The loop does everything in between.

What It Looks Like

Illustrative: every figure carries its source, every gap is flagged

Country X Fiscal & Debt Brief · auto-generated draft, 2025

Debt / GDP: 62.4% [IMF WEO 2025, tbl. A1]

Fiscal balance: −3.1% of GDP [Ministry of Finance, 2025 budget]

External-debt share: 48% [World Bank IDS]

Tax expenditures: 34 measures, ≈2.1% of GDP [Global Tax Expenditures DB]

⚠ **Gap flagged:** sub-national debt is not disaggregated in public sources. The figure is omitted, not estimated.

Analyst note: verify Q4 balance against the revised budget before release.

A preview of the *format*, not a real country. The point: **nothing is asserted without a source, and no gap is quietly filled**.

Comparable Across Every Country

The payoff for an institution that works country by country

Why it matters here

- One template for every country
- Every claim sourced and checkable
- Refreshed on demand, at low cost
- Analysts verify, not copy-paste

The same instrument runs on all of them, so a brief for one country is directly comparable to a brief for the next. Coverage stops being a question of how many analyst-weeks you can afford.

A human sets the template and signs off. The loop does the assembly.

Does It Work?

Evidence from a randomized experiment

Does AI Narrow Productivity Gaps?

Cruces, Fernández Meijide, Galiani, Gálvez & Lombardi (2026)

Most evidence studies AI *inside* firms, where hiring already compresses who is in the room. We ran a **randomized online experiment outside firms**, across the general adult population, to ask whether AI narrows productivity gaps by **education**.

The design

1,174 adults, aged 25–45
An incentivized, workplace-style **business problem-solving task**
Randomized: **with** or **without** a generative-AI assistant

The twist: a follow-up

After the main task, everyone completes a second module **without AI**. This reveals whether the gains were real learning or mere **delegation** that vanishes once the tool is gone.

AI Narrows the Gap

Larger gains for lower-education participants

75%

of the education gap in task performance closed, just by putting AI in people's hands

AI increases performance for **everyone**, with substantially larger gains for **lower-education** participants. In task execution, AI levels the field.

The nuance: the follow-up shows the gains are not purely delegation, yet a sizable education gap re-emerges without AI. Human capital still shapes *effective use* and what sticks.

Give in to Huxley and we are entertained into passivity.
Flee toward Orwell and we build cages in the name of protection.
The only real alternative is to put **AI on our side**.

And we have begun to measure what that means. It gives the most to those who start with the least.

AI at Your Side

Sebastian Galiani
Raul A. Sosa

Forthcoming from Oxford University Press

Thank you!