Exercise Lab – Self-Learning User Guide

1. What this lab is for

This Exercise Lab is designed for **knowledge workers**, not software engineers.

You'll use simple interactive controls (sliders, dropdowns, tables) to practice **real decisions about AI**:

- How retrieval settings affect answers (RAG)
- How to diagnose AI failures
- How to design safe tool use for agents
- How to structure multi-agent workflows
- How to balance speed, freshness, and safety in production

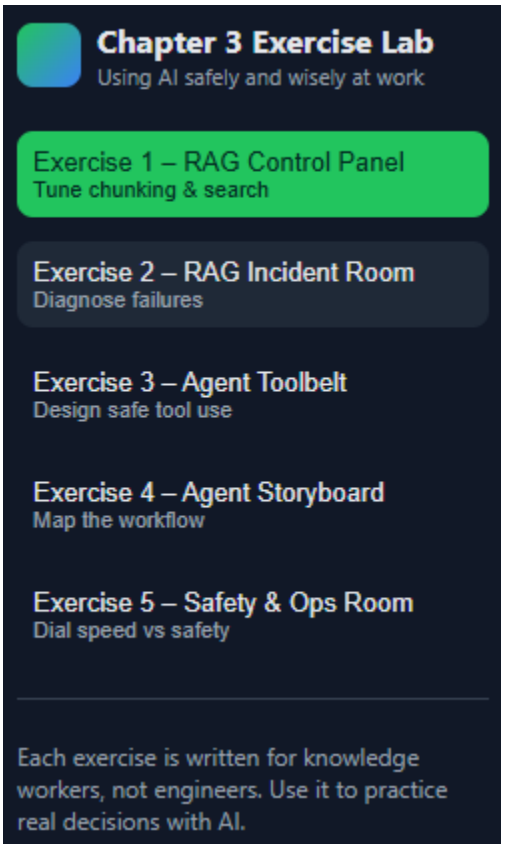
You don't need coding skills. You just need your **day-to-day work experience** and a willingness to experiment.

---

2. Getting started

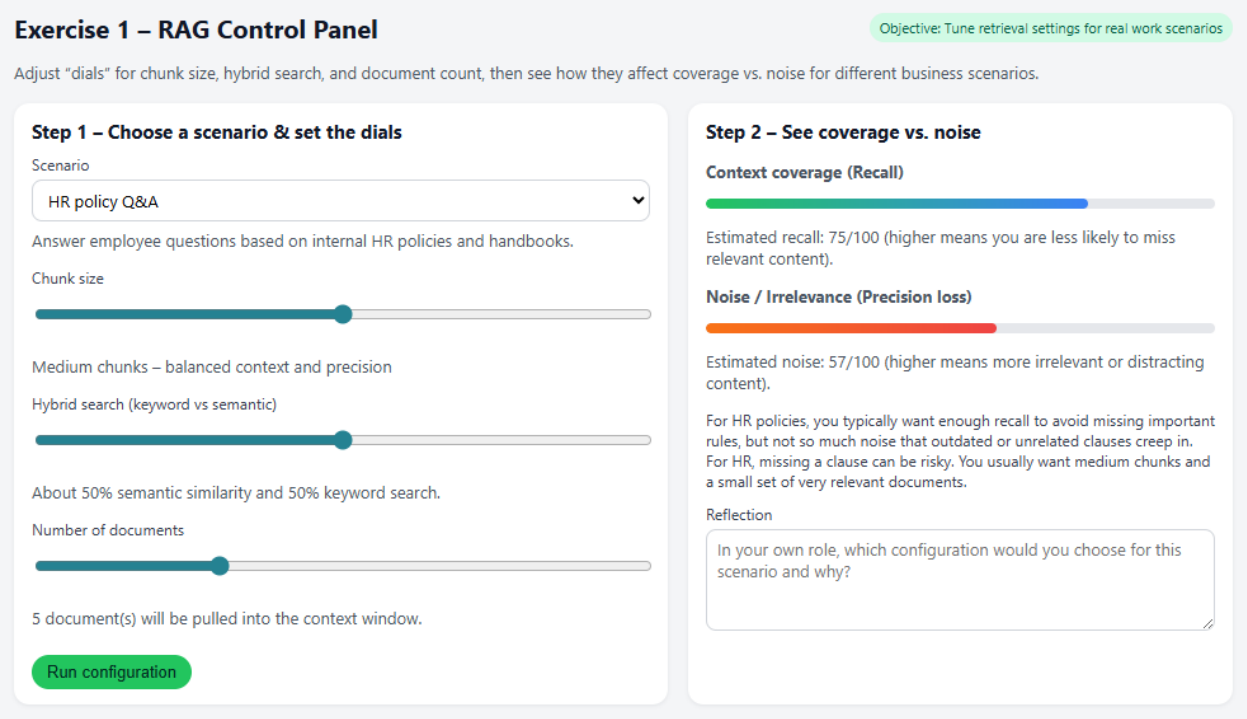
1. **Open the HTML file**
    - Double-click Chapter3_Exercise_Lab.html, or
    - Drag it into a browser window (Chrome, Edge, etc.).
2. **Understand the layout**
    - **Left side:** Navigation buttons for **Exercise 1–5**.
    - **Right side:** The current exercise content (cards, sliders, tables, reflection boxes).
3. **How to move around**
    - Click a button on the left (e.g., "Exercise 1 – RAG Control Panel") to switch exercises.
    - Only one exercise is visible at a time.

You can work through the exercises **in order** or pick the ones that match your current work challenges.



Exercise 1 – RAG Control Panel

In the app:
Exercise 1 – RAG Control Panel



What you'll practice

You'll build intuition for how RAG (Retrieval-Augmented Generation) behaves when you change:

- **Chunk size** (small vs large pieces of documents)
- **Hybrid search** (keyword vs semantic similarity)
- **Number of documents** retrieved

You'll see how these settings change:

- **Context coverage (Recall)**
- **Noise / Irrelevance (Precision loss)**

How to use it

1. **Choose a scenario**
    - Use the **Scenario** dropdown:

        - HR policy Q&A
        - Product SKU lookup
        - Leadership summary
    - Read the short description below the dropdown.
2. **Adjust the "dials"**
    - Move the **Chunk size** slider (Small / Medium / Large).
    - Move the **Hybrid search** slider (0–100% semantic).
    - Move the **Number of documents** slider (3–10).
3. **Run the configuration**
    - Click **Run configuration**.
    - Watch:
        - The **Recall** gauge: how likely the system is to include the right content.
        - The **Noise** gauge: how much irrelevant material may show up.
4. **Read the explanation**
    - The text below the gauges explains why those settings might be good or risky for the chosen scenario.
5. **Write your reflection**
    - In the **Reflection** box, answer:

"In your own role, which configuration would you choose for this scenario and why?"

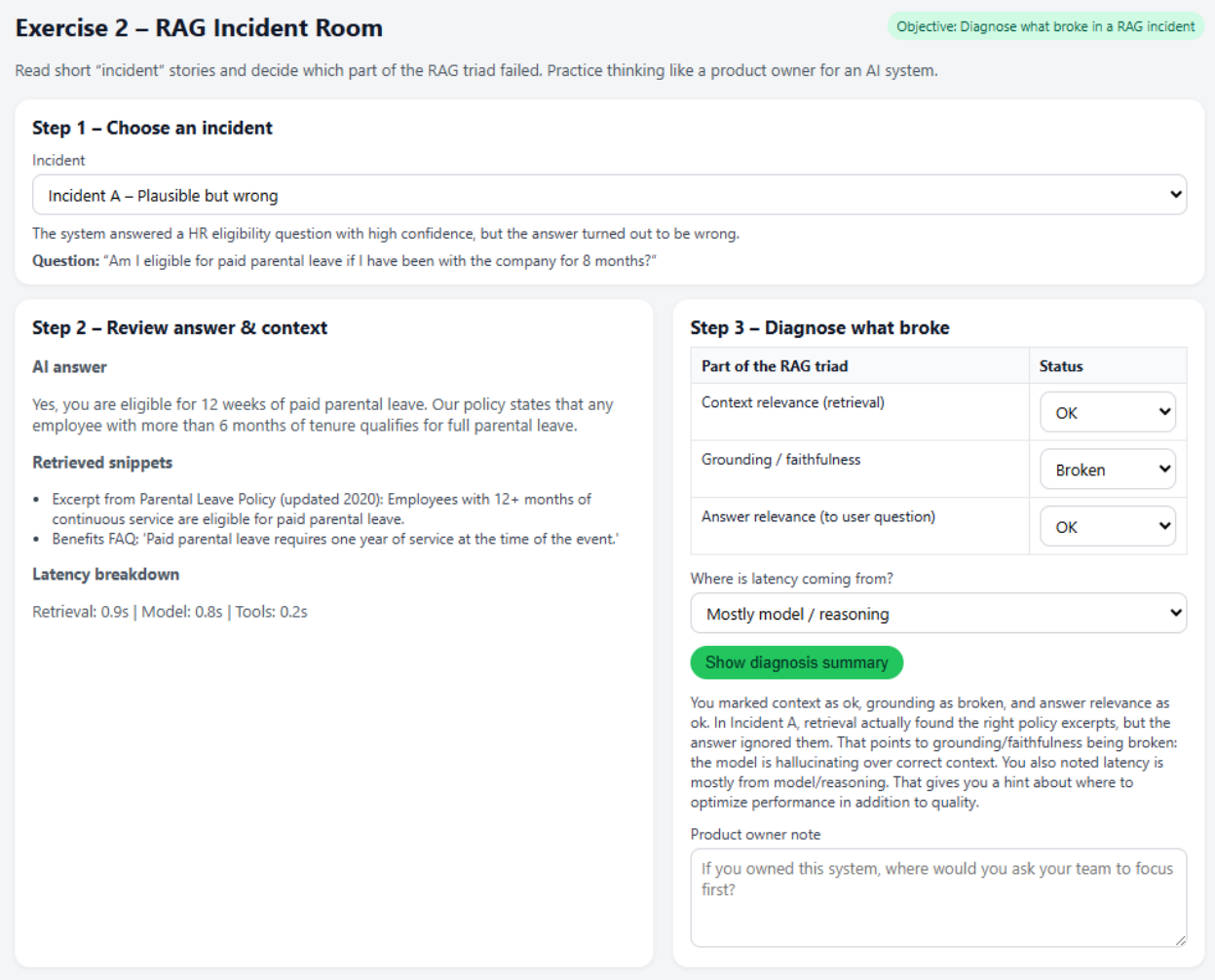
How to get value from it

- Try **extremes** first (tiny chunks vs huge chunks; very few vs many documents).
- Ask yourself: *"If this were my AI assistant at work, would I trust its answers under this configuration?"*
- Capture one or two "rules of thumb" you discovered (e.g., "For HR questions, avoid too many docs—noise goes up fast.").

Exercise 2 – RAG Incident Room

In the app:

Exercise 2 – RAG Incident Room



What you'll practice

You'll practice **diagnosing AI failures** using the **RAG triad**:

- **Context relevance** – Did it retrieve the right documents?
- **Grounding / faithfulness** – Did the answer stay true to those documents?
- **Answer relevance** – Did it answer the user's actual question?

You'll also notice where **latency** (slowness) comes from.

How to use it

1. **Pick an incident**

- Use the **Incident** dropdown (A, B, or C).
- Read:
  - The description
  - The user's question
  - The AI answer
  - The retrieved snippets
  - The latency breakdown

2. **Judge each part of the triad**
   - For each row in the table:
     - Context relevance (retrieval)
     - Grounding / faithfulness
     - Answer relevance
   - Choose **OK** or **Broken**.
3. **Identify latency source**
   - Use the **Where is latency coming from?** dropdown:
     - Mostly retrieval
     - Mostly model / reasoning
     - Mostly tools / APIs
4. **Show the diagnosis summary**
   - Click **Show diagnosis summary**.
   - Read the explanation: it will connect your choices to a likely root cause.
5. **Optional: product owner note**
   - In **Product owner note**, write what you'd ask your AI team to fix first.

How to get value from it

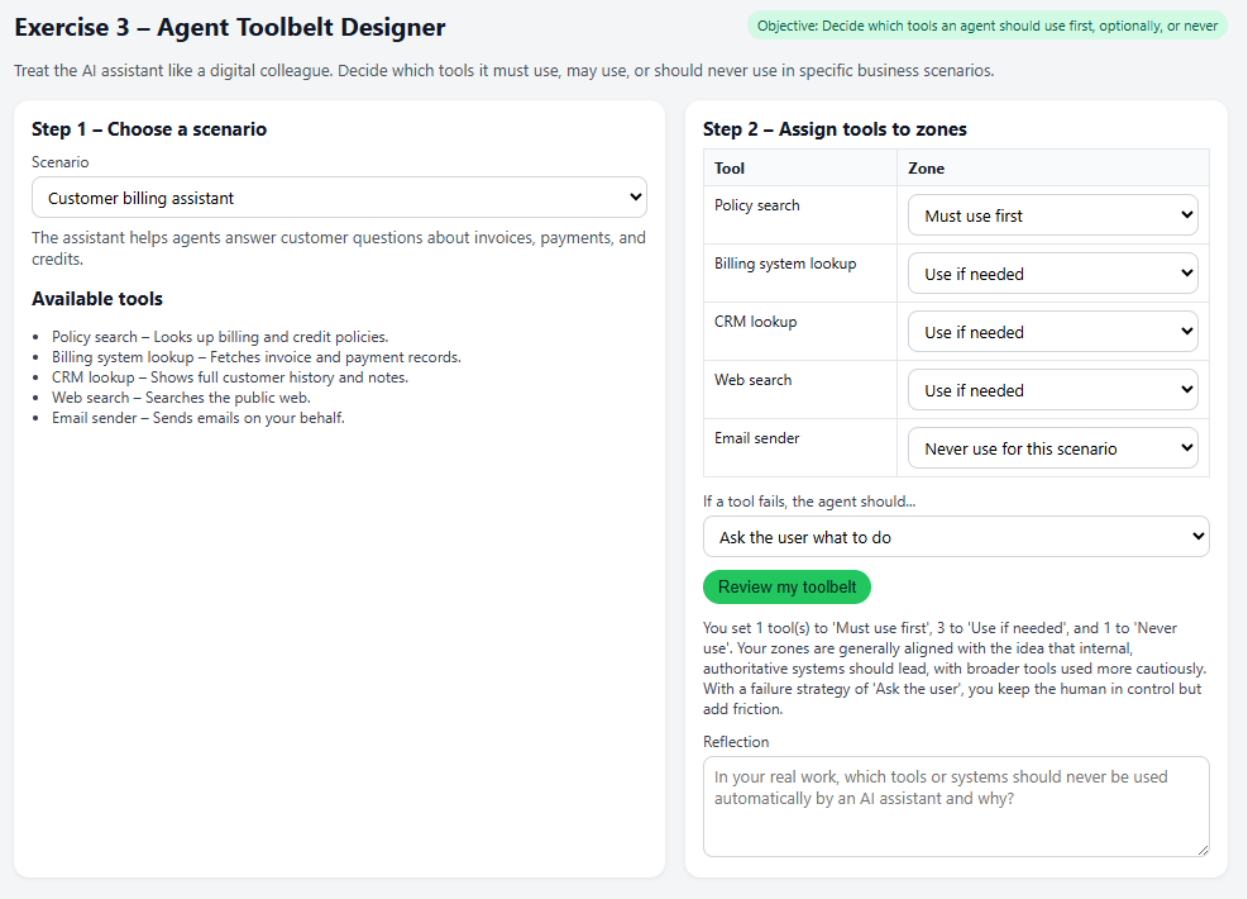
- Don't think in terms of "the model is bad." Instead ask:
  - Did it **find** the right content?

    - Did it **stick** to that content?
    - Did it **answer** the question asked?
- Think about your own environment:
    - *If an AI system made this mistake at work, how would you talk about it in a post-mortem?*

Exercise 3 – Agent Toolbelt Designer

In the app:

Exercise 3 – Agent Toolbelt Designer



What you'll practice

You'll design a **toolbelt** for an AI assistant:

- Which tools it must use first
- Which tools it may use if needed
- Which tools it must never use in a given scenario
- How it should behave when a tool fails

You're expressing **policy and governance**, not coding.

How to use it

1. **Choose a scenario**

- Use the **Scenario** dropdown:
  - Customer billing assistant
  - Internal HR policy helper
  - Sales pipeline explainer
- Read the scenario description.

2. **Review available tools**
   - Read the **Available tools** list (e.g., Policy search, CRM lookup, Web search, Email sender).
   - Notice what each tool can do.
3. **Assign tools to zones**
   - In the table:
     - For each tool, pick a **Zone**:
       - Must use first
       - Use if needed
       - Never use for this scenario
   - Make sure you've chosen a zone for every tool.
4. **Choose a failure strategy**
   - In **If a tool fails...**, pick:
     - Retry with backoff
     - Ask the user
     - Try a different tool
     - Fail fast
5. **Review your toolbelt**
   - Click **Review my toolbelt**.
   - Read the narrative summary and any safety hints.
6. **Write your reflection**

    - In **Reflection**, answer:

“In your real work, which tools or systems should never be used automatically by an AI assistant and why?”

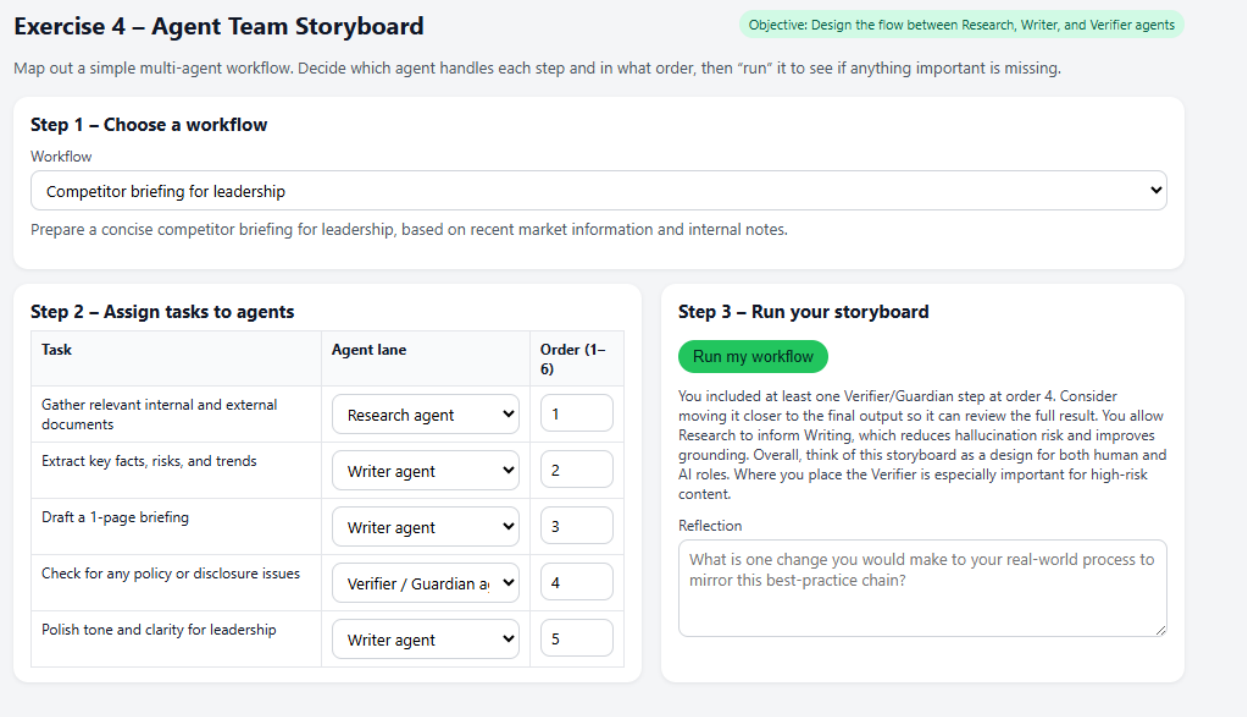
How to get value from it

- Think of **real systems** in your company:
    - Which ones are “source of truth”?
    - Which ones are risky if misused?
- Use your answers as a starting point for an **AI governance checklist**:
    - “These tools can be auto-used; these require human confirmation.”

Exercise 4 – Agent Team Storyboard

In the app:

Exercise 4 – Agent Team Storyboard



What you'll practice

You'll design a **multi-agent workflow**:

- Which tasks belong to a **Research agent**
- Which tasks belong to a **Writer agent**
- Which tasks are handled by a **Verifier / Guardian agent**
- In what **order** these tasks should run

This mirrors how you might structure both **human** and **AI** roles.

How to use it

1. **Choose a workflow**
    - Use the **Workflow** dropdown:
        - Competitor briefing for leadership
        - Customer complaint case summary

        - HR policy announcement
    - Read the description.
2. **Assign tasks to agents**
    - In the table:
        - Each row is a task (e.g., Gather docs, Draft summary, Check compliance).
        - For each task:
            - Choose an **Agent lane**:
                - Research agent
                - Writer agent
                - Verifier / Guardian agent
            - Set an **Order** number (1–6) to show the sequence.
3. **Run your workflow**
    - Click **Run my workflow**.
    - Read the feedback:
        - Is there a Verifier step?
        - Is the Verifier near the end?
        - Does writing start before any research?
4. **Adjust if needed**
    - If the feedback shows problems (e.g., Writer before Research), adjust agents or order and run again.
5. **Write your reflection**
    - In **Reflection**, answer:

"What is one change you would make to your real-world process to mirror this best-practice chain?"

How to get value from it

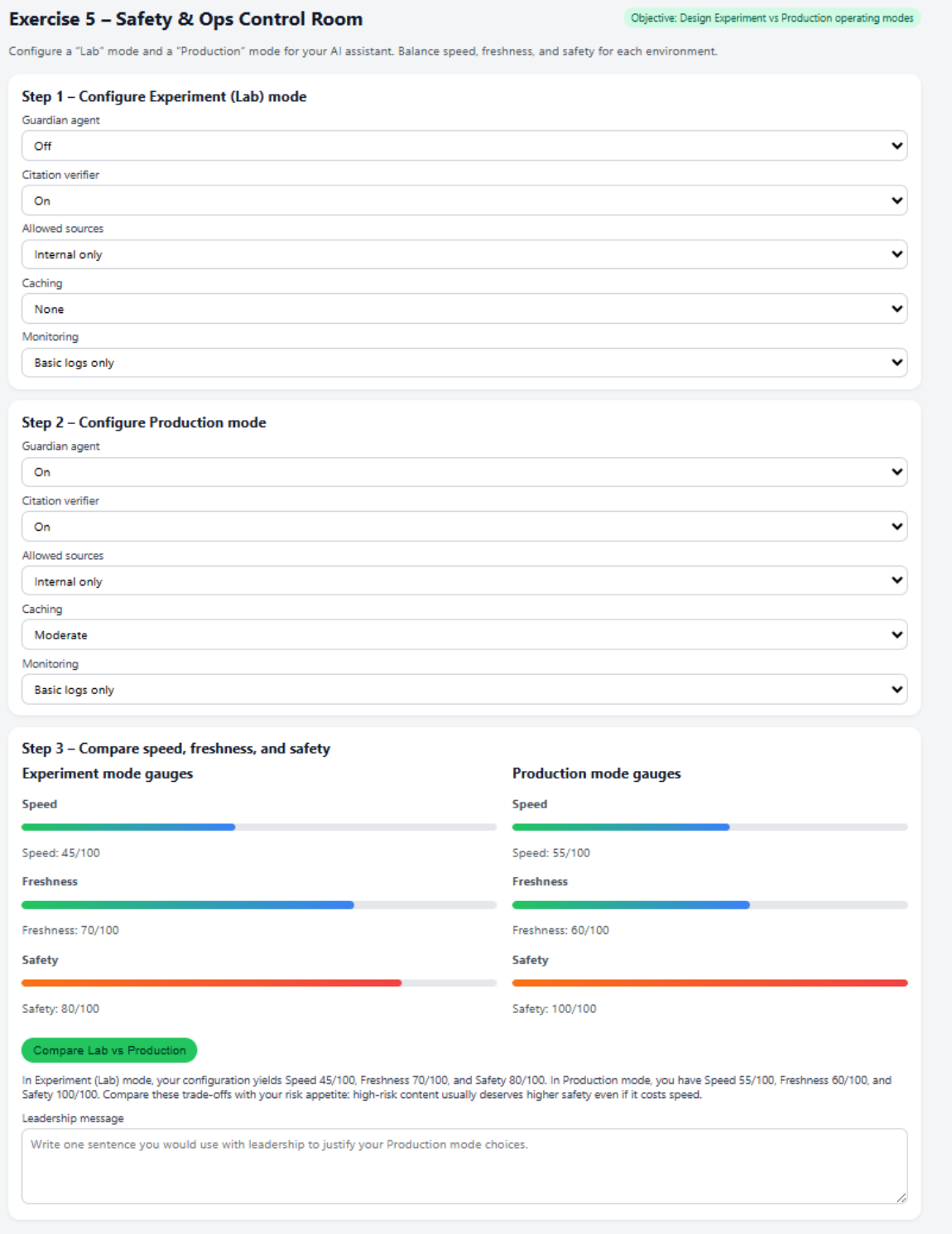
- Compare this agent storyboard to how your **human team** works today.

- Ask yourself:
    - *“Do we actually have a Verifier role before big decisions or announcements?”*
    - *“If we add AI, where should that Verifier step live?”*

Exercise 5 – Safety & Ops Control Room

In the app:

Exercise 5 – Safety & Ops Control Room



What you'll practice

You'll configure two modes for an AI system:

- **Experiment (Lab) mode** – where you explore and test
- **Production mode** – where real users rely on the system

You'll see how choices impact:

- **Speed**
- **Freshness of information**
- **Safety & compliance**

How to use it

1. **Configure Lab mode**
    - In the first card, set:
        - Guardian agent: On/Off
        - Citation verifier: On/Off
        - Allowed sources: Internal only / Internal + web
        - Caching: None / Moderate / Aggressive
        - Monitoring: Basic logs / Dashboards + alerts
2. **Configure Production mode**
    - In the second card, set the same controls, but:
        - Imagine this is what you'd put in front of **real employees or customers**.
        - Aim for the configuration you'd be comfortable defending.
3. **Compare modes**
    - Click **Compare Lab vs Production**.
    - Look at the gauges:
        - Speed (0–100)
        - Freshness (0–100)

        - Safety (0–100)
    - Read the text summary of your trade-offs.
4. **Write your leadership message**
    - In **Leadership message**, write:

"One sentence you would use to justify your Production mode choices to leadership."

How to get value from it

- Ask:
    - *"Is my Production mode clearly safer than Lab?"*
    - *"Where am I willing to trade speed for safety?"*
- Use your leadership message as a draft for an **AI usage standard** in your team or department.

---

8. Putting it all together

As you finish the Chapter 3 Exercise Lab:

1. Review your reflection answers across the exercises.
2. Collect them into a short **"AI Usage Playbook"** for yourself:
    - How you'd configure RAG for your work
    - How you'd diagnose failures
    - How you'd control tool use and workflows
    - How you'd set Lab vs Production modes

This lab is not about learning jargon. It is about becoming a **confident, thoughtful user and governor of AI systems** in your real job.

If you revisit these exercises after a few weeks of using AI at work, you'll often make **different choices**—and that's a sign your understanding has deepened.