

TECH & AI CASE STUDY - MATERIALS SCIENCE - 2026

# A GenAI Research Copilot That Cut Literature Reviews From Six Weeks to Three Days

Based on a delivery engagement with Calderon Advanced Materials

**6 weeks to 3 days**  
LITERATURE REVIEW TIME

**2 this year**  
DUPLICATE TEST CYCLES AVOIDED

**78%**  
WEEKLY ACTIVE SCIENTISTS

**2 wks to 1 day**  
PATENT SEARCH TIME

---

A fillable practice edition: the full story, 8 actions,  
6 reflections and a 90-day plan - work it with your team.

**Free - no email required**  
[emersontechnology.com](https://emersontechnology.com)

## MAKE THIS PRACTICE BOOK YOURS

**Work it, do not just read it**

This practice book turns the Calderon Advanced Materials case study into a working session for your team. Read the full story first - what happened, what was built and what it changed. Then tick the 8 actions you can commit to, write short answers to the 6 reflections, and finish with a 90-day plan. Every field in this PDF is fillable: open it in Adobe Acrobat Reader, Chrome, Edge or macOS Preview, type directly into the boxes and save.

|           |                   |
|-----------|-------------------|
| YOUR NAME | ROLE              |
|           |                   |
| COMPANY   | TEAM / DEPARTMENT |
|           |                   |
| DATE      | EMAIL (OPTIONAL)  |
|           |                   |

  

MY GOAL FOR THIS PRACTICE SESSION

- 01 The story** - What happened at Calderon Advanced Materials - before, during and after the project.
- 02 The results** - The numbers that changed, plus industry benchmarks from public research.
- 03 The 8-action checklist** - Tick boxes as your organization completes each step. Owner and date fields included.
- 04 The 6 reflections** - Fillable questions that pressure-test your own roadmap.
- 05 My 90-day plan** - Three fillable phase cards that turn this case study into your roadmap.

## RUN IT AS A 90-MINUTE TEAM SESSION

|                                                                                                                                    |                                                                                                                                                              |                                                                                                                                                                               |
|------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p><b>25 min</b><br/><b>READ</b></p> <p>Read the story together. Ask everyone to note the one number that surprised them most.</p> | <p><b>40 min</b><br/><b>PRACTICE</b></p> <p>Work the 8-action checklist and the 6 reflections. Assign an owner and a date to every action you commit to.</p> | <p><b>25 min</b><br/><b>COMMIT</b></p> <p>Fill in the 90-day plan. One goal per phase, one number that tells you it worked. Book a review date before you leave the room.</p> |
|------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

## THE STORY

## What happened, and why it worked

### 01 The Company

---

Calderon Advanced Materials develops specialty alloys and ceramic composites for energy and industrial customers, operating 85 scientists across three labs. Its programs run on years-long arcs: a customer need arrives, the state of the art is mapped, compositions are proposed, tested and iterated, and eventually a candidate material moves toward pilot production. In this field the opening move of every program is a literature review — an exhaustive sweep of published papers, patents and, just as importantly, the company's own experimental record — and the quality of that opening move ripples through everything downstream, because a program that starts from an incomplete picture of prior art will pay for the gap later, at pilot scale, where mistakes are expensive. The company's intellectual position made the problem sharper. Its formulations and test results are proprietary and commercially sensitive, and its scientists are exactly the kind of users AI vendors dream of: technically fluent, curious, and already experimenting with public chatbots on their own initiative. Calderon's research leadership needed speed — but speed that came at the cost of pasting confidential compositions into a third-party model was worse than no speed at all.

### 02 Where Things Stood

---

Every new program opened with a literature review averaging six weeks: 40,000 papers and patents, plus twelve years of internal lab notebooks scattered across shared drives and an aging electronic lab notebook. The public literature was at least findable, if slow; the internal record was effectively unfindable. Notebooks were structured well enough for the scientist who wrote them and nobody else, and search across the drives returned filename matches, not knowledge. Scientists spent those six weeks doing work a retrieval system should do — locating, deduplicating, summarizing — before doing the work only they could do.

The cost of unfindable knowledge was not abstract. Duplicate experiments were common: the technical director estimated three redundant test cycles a year, at roughly \$170,000 each, where a team reran an experiment that had already been run — sometimes by a colleague one building away — because nobody could find the first result. Patent-landscape gaps surfaced late, forcing costly redesigns when a program discovered prior art mid-development instead of at proposal. Meanwhile the security team had banned public chatbots outright after finding formulations pasted into consumer tools, a correct decision that solved nothing: it removed the unsafe shortcut without providing a safe alternative, so scientists quietly went back to the shared drives and the six-week slog.

The context outside the company's walls raised the stakes. DeepMind's GNoME model reportedly predicted 2.2 million new crystal structures, including roughly 380,000 rated stable, expanding known stable materials far beyond a century of prior discovery, and Berkeley Lab's autonomous A-Lab reportedly synthesized 41 novel inorganic materials in 17 days. Research loops were accelerating industry-wide, and Calderon's opening move was six weeks of manual search.

### 03 The Challenge

The challenge was to give scientists a research copilot that was simultaneously fast and inside the perimeter — a tool answering questions across the public and internal corpora in seconds, with zero data leaving Calderon's control. The technical shape of that is retrieval-augmented generation over a mixed corpus: published papers and patents with dense, standardized structure, plus twelve years of lab notebooks with dense, non-standard structure — handwritten abbreviations, skipped context, results recorded as the experiments unfolded rather than as conclusions. McKinsey estimates generative AI could add \$2.6-4.4 trillion in annual economic value across use cases, with research and development among the most affected functions; capturing any of that in a materials lab meant solving the internal corpus first, because that is where the duplicated experiments were hiding.

The second constraint was epistemic. In materials science, a confident wrong answer can cost a quarter: a scientist who trusts a hallucinated property value designs a test campaign on fiction. The copilot therefore had to be built so that every answer could be verified — not "trust the model" but "check the source in one click" — and so that its conclusions entered the lab record as leads to verify, never as settled findings. Research leadership was explicit: the tool would be judged not on how impressive its answers sounded but on how quickly a skeptical senior scientist could confirm or refute them.

### 04 Discovery and Diagnosis

Emerson Technology began with a two-month discovery that was itself an exercise in retrieval. The team inventoried the corpora: 40,000 papers and patents, the shared drives, and the aging ELN, quantifying what was digitized, what was scanned-but-unsearchable, and what existed only as structureless PDFs. The audit found the predictable inversion — the most valuable corpus was the worst organized: twelve years of internal notebooks held the negative results, failed compositions and undocumented context that would prevent duplicate experiments, and almost none of it was reachable by any existing tool. Public papers were the easy part; internal knowledge was the highest-value and best-hidden corpus in the company.

The diagnosis process leaned on Calderon's information professionals from the start, and that pairing shaped everything. The librarians knew the taxonomy, the known gaps, and the three most valuable internal documents nobody could currently find — knowledge that saved months of indexing mistakes. With them, Emerson built a librarian-curated evaluation suite of 200 known questions with verified answers, covering prior-art searches, property lookups and internal-record questions, so that every future model update would be gated against expert-written ground truth rather than demo impressions. The interviews also settled the deployment question before it was asked: scientists said, flatly, they would not use a tool they did not fully trust with unpublished formulations, so the security perimeter was a product requirement, not a compliance checkbox.

### 05 The Solution

Emerson Technology designed and built a private retrieval-augmented research platform, deployed inside Calderon's own cloud accounts, that answers research questions with verifiable citations. The architecture has three layers.

#### The private corpus

The 40,000 papers and patents were parsed and OCR'd, and twelve years of lab notebooks were structured — normalized, indexed and linked to programs and scientists — then everything was embedded into a vector index inside Calderon's own AWS virtual private cloud. No data leaves the perimeter: models, indexes and query logs all live inside the company's own accounts, and every query is logged for IP protection. Digitizing and structuring the internal notebooks came before model selection in priority, because an excellent model over an unfindable corpus produces confident answers to the wrong questions. The corpus layer is why duplicate experiments became

preventable: the first result now exists in a system that can find it.

## Citation-backed answers

The copilot answers research questions with hybrid keyword-plus-semantic search and cites every source down to the page and notebook entry, so scientists verify rather than trust. A property claim arrives with the table it came from; a "we already tried this" arrives with the notebook page, the date and the author's initials. The interface is a Next.js application with shared collections, annotations and export to the ELN, so copilot findings flow into the lab record as structured leads — explicitly marked as leads to verify, never as settled findings. This design is what converted the skeptics: the tool never asks a senior scientist to believe it, only to check it, and checking takes seconds instead of weeks.

## Evaluation and rollout

The librarian-curated evaluation suite of 200 known questions gates each model update: any candidate model or configuration change must clear the suite before it reaches scientists, which is why quality regressions never arrived silently. Deployment took four months, and adoption spread lab by lab through two champion scientists — respected senior researchers who tested the tool against questions they already knew the answers to, then taught their labs. Start with review, not discovery: the copilot was aimed at summarizing and searching first, where answers are checkable, and novel hypothesis generation was deferred until trust existed — a sequencing choice the champions later credited for the tool's credibility.

### 06 The First 90 Days

---

Weeks one through six concentrated on the corpus: parsing and OCR of the papers and patents, and the structuring of internal notebooks, with librarians resolving taxonomy conflicts and flagging the collections whose scans were too poor to trust. Weeks seven through ten built and tuned the retrieval stack and ran it against the 200-question evaluation suite, iterating on hybrid search weights until the suite's recall targets held — including the internal-record questions that would prevent duplicate experiments. Weeks eleven and twelve onboarded the two champion scientists and their immediate labmates, instrumented usage, and held weekly feedback sessions where every wrong or unhelpful answer was traced to a corpus gap, a retrieval miss or a model failure and fixed at the root.

The pilot's most famous internal moment came early: a champion asked whether the company had ever tested a particular coating composition, half-expecting another six-week hunt. The copilot returned a completed but unreported internal study from four years earlier in under a minute, notebook page attached. That single retrieval did more for adoption across the labs than any demo, because every scientist who heard the story had personally run an experiment they later learned was a rerun. By the end of the quarter the platform was live to two labs, the evaluation suite was a standing gate, and the third lab was asking when its turn came.

### 07 The Results

---

The opening move of every program changed. Program-opening literature reviews now take three days instead of six weeks, with citation trails attached, and patent searches that took two weeks run in one day — prior art now surfaces at proposal stage instead of mid-development. The copilot has surfaced completed but unreported internal studies that have already prevented two duplicate test cycles this year, and at the technical director's estimated cost of roughly \$170,000 per redundant cycle, that alone covers a meaningful share of the platform's cost. Seventy-eight percent of scientists use the platform weekly, and two alloy-development programs have moved to pilot testing roughly a quarter earlier than planned.

| METRIC                        | BEFORE  | AFTER       |
|-------------------------------|---------|-------------|
| Literature review time        | 6 weeks | 3 days      |
| Duplicate test cycles avoided | -       | 2 this year |
| Weekly active scientists      | -       | 78%         |
| Patent search time            | 2 weeks | 1 day       |

The qualitative changes compounded the quantitative ones. The security team, once the loudest skeptic, now cites the deployment in customer due-diligence answers — a private, logged, citation-backed research platform became evidence of IP discipline rather than a risk to argue about. Program retrospectives now include a line on what the copilot found or missed, which keeps the corpus honest. And the ban on public chatbots stopped being a friction point, because scientists finally had something better: a tool that answers faster than the unsafe alternative ever did, without a single formulation leaving the building.

## 08 What Your Business Should Do

The pattern behind Calderon's result is repeatable in any organization sitting on decades of proprietary knowledge. Digitize lab notebooks before anything: structure and OCR historical notebooks and reports first, because internal knowledge is usually the highest-value and best-hidden corpus in the company — public papers are the easy part. Deploy inside your security perimeter, keeping models, indexes and query logs inside your own cloud accounts or on-premises, because scientists will rarely use tools they do not fully trust with unpublished work. Demand citations for every answer, with source references down to the exact page or notebook entry, so experts verify quickly and catch model errors early. Build an expert evaluation set before launch — 100 to 200 questions with known answers written by senior researchers — and rerun it for every model update, because without it quality regressions arrive silently.

Start with review, not discovery: aim the copilot at summarizing and searching first, where answers are checkable, and let novel hypothesis generation earn trust later. Pair librarians with engineers, because information professionals know the corpus, the taxonomy and where the bodies are buried, and teams that skip them build beautiful tools over incomplete, misleading collections. Measure hours saved per program — literature-review time, duplicate experiments, patent-search turnaround, program by program — because research leaders fund what shows up in program economics, not in impressive demo videos. And treat outputs as hypotheses: write copilot conclusions into lab records as leads to verify, never as settled findings, because in materials science a confident wrong answer can cost a quarter.

Ask the questions Calderon's research leadership now reviews quarterly: how long a literature review takes in your organization, and who counts the cost; how many experiments your organization repeated last year simply because nobody could find the first ones; where your scientists go for answers — your own archive, or tools your security team never approved; what a competitor's four-times-faster R&D loop would do to your patent position; whether your librarians could name the three most valuable internal documents nobody can currently find; and what your research leaders would need to see before trusting an AI tool with program decisions. Work through the eight actions in this practice book with your team, and revisit those reflection questions each quarter to keep the research engine honest.

### REAL-WORLD PROOF & BENCHMARKS

**DeepMind's GNoME model reportedly predicted 2.2 million new crystal structures, including roughly 380,000 rated stable — expanding known stable materials far beyond a century of prior discovery.** - Nature / DeepMind, 2023

**Berkeley Lab's autonomous A-Lab reportedly synthesized 41 novel inorganic materials in 17 days, demonstrating AI-guided experimentation at speeds no manual lab could match.** - Nature, reported 2023

**McKinsey estimates generative AI could add \$2.6-4.4 trillion in annual economic value across use cases, with research and development among the most affected functions.** - McKinsey, 2023

## THE PRACTICE PAGES

## The 8-action checklist

Work through the list with your team. Tick each box as your organization completes the step, and fill in an owner and a target date so every action has a name and a deadline.

☐**1. Digitize lab notebooks before anything**

Structure and OCR historical notebooks and reports first; internal knowledge is usually the highest-value and best-hidden corpus in the company. Public papers are the easy part.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|

☐**2. Deploy inside your security perimeter**

Keep models, indexes and query logs inside your own cloud accounts or on-premises. Scientists will rarely use tools they do not fully trust with unpublished formulations.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|

☐**3. Demand citations for every answer**

Require source references down to the exact page or notebook entry on every response. Citation-backed answers let scientists verify quickly and catch model errors early.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|

☐**4. Build an expert evaluation set**

Have senior researchers write 100-200 questions with known answers before launch, then rerun the suite for every model update. Without this, quality regressions arrive silently.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|

☐**5. Start with review, not discovery**

Aim the copilot at summarizing and searching first; novel hypothesis generation earns trust later. Reviews have checkable answers, which is what skeptics need to see.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|



6. Pair librarians with engineers

Information professionals know the corpus, the taxonomy and where the bodies are buried. Teams that skip librarian input build beautiful tools over often-incomplete, misleading collections.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|



7. Measure hours saved per program

Track literature-review time, duplicate experiments and patent-search turnaround program by program. Research leaders fund what shows up in program economics, not in impressive demo videos.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|



8. Treat outputs as hypotheses

Write copilot conclusions into lab records as leads to verify, never as settled findings. In materials science, a confident wrong answer can cost a quarter.

OWNER

TARGET DATE

STATUS

|  |  |  |
|--|--|--|
|  |  |  |
|--|--|--|

AFTER THE CHECKLIST: OUR TOP 3 PRIORITIES

From the eight actions above, write down the three you will start first - and the single number you expect each one to move. Three priorities with owners beat eight without any.

|                                                                        |                                                                        |                                                                        |
|------------------------------------------------------------------------|------------------------------------------------------------------------|------------------------------------------------------------------------|
| <p>PRIORITY 1</p> <div></div> <p>NUMBER IT SHOULD MOVE</p> <div></div> | <p>PRIORITY 2</p> <div></div> <p>NUMBER IT SHOULD MOVE</p> <div></div> | <p>PRIORITY 3</p> <div></div> <p>NUMBER IT SHOULD MOVE</p> <div></div> |
|------------------------------------------------------------------------|------------------------------------------------------------------------|------------------------------------------------------------------------|

## THE PRACTICE PAGES

## Reflections for your team

Use these questions in your next planning session and type short answers into the boxes. The gaps you find are usually the fastest wins.

**Q1. How long does a literature review take in your organization, and who counts the cost?**

**Q2. How many experiments did your organization repeat last year simply because nobody could find the first ones?**

**Q3. Where do your scientists go for answers — your own archive, or tools your security team never approved?**

**Q4. If a competitor's R&D loop runs four times faster, what does that do to your patent position?**

**Q5. Could your librarians name the three most valuable internal documents nobody can currently find?**

**Q6. What would your research leaders need to see before trusting an AI tool with program decisions?**

## THE PRACTICE PAGES

## My 90-day plan

Finish the session by committing to a plan. Three phases, one goal each, one number that tells you it worked. Bring this page to your next leadership review - or to a free consultation with Emerson Technology, and we will pressure-test it with you.

### Days 1-30 Focus: audit and alignment

**PRIMARY GOAL****KEY ACTIONS****HOW WE MEASURE SUCCESS****OWNER**

### Days 31-60 Focus: build and prove

**PRIMARY GOAL****KEY ACTIONS****HOW WE MEASURE SUCCESS****OWNER**

Days 61-90 Focus: scale and measure

PRIMARY GOAL

KEY ACTIONS

HOW WE MEASURE SUCCESS

OWNER

Tip: save a copy per team so each group keeps its own answers. Revisit the reflections every quarter - the questions stay the same even as the numbers move.

BOOK THE REVIEW BEFORE YOU CLOSE THE BOOK

Plans without a review date drift. Before you close this practice book, agree on when you will come back to these pages, who will run that session, and when you will check the numbers again. If you want an outside pair of eyes, bring this page to a free consultation with Emerson Technology and we will pressure-test the plan with you.

|                                    |                                    |                                      |
|------------------------------------|------------------------------------|--------------------------------------|
| <div>REVIEW DATE</div> <div></div> | <div>REVIEWED BY</div> <div></div> | <div>NEXT CHECK-IN</div> <div></div> |
|------------------------------------|------------------------------------|--------------------------------------|

# Ready for a transformation like Calderon Advanced Materials's?

Emerson Technology designs and builds the data, AI, cloud and automation systems behind results like these. On a free consultation call we will map this exact playbook to your business - your data, your constraints, your numbers.

**Book a free consultation**

**[support@emersontechnology.com](mailto:support@emersontechnology.com)**

or start online: [emersontechnology.com/contact](https://emersontechnology.com/contact)