AFROTECH 2026

What's Next After LLMs?

Building valuable products in the age of AI



Olukunle Owolabi, PhD

AI Engineer at Meta · Researcher · Builder

AFRO TECH™ Conference

LET'S CONNECT

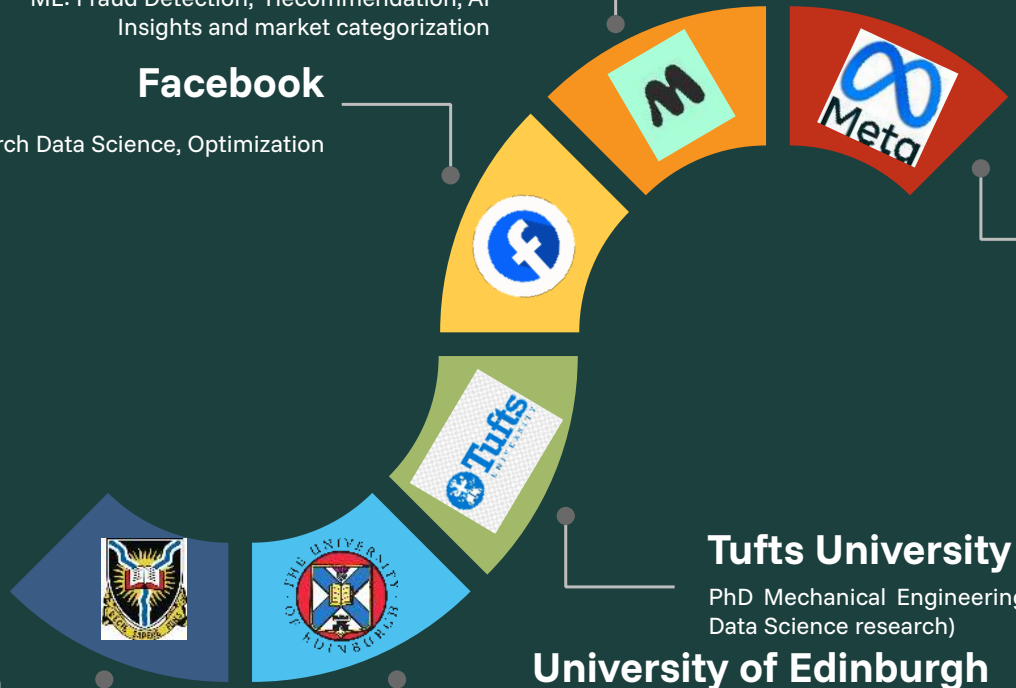
LinkedIn
Olukunle Owolabi, PhD

Website
olukunleowolabi.com

WHAT'S NEXT AFTER LLMS? ✦ OLUKUNLE OWOLABI, PhD ✦ AI ENGINEER AT M

AFROTECH Conference

Non Linear Engineering Career



Mindbody

ML: Fraud Detection, Recommendation, AI Insights and market categorization

Facebook

Research Data Science, Optimization

Meta

Recommendation System, AI Agents Evaluation, Agent Memory

Tufts University

PhD Mechanical Engineering (Energy and Data Science research)

University of Edinburgh

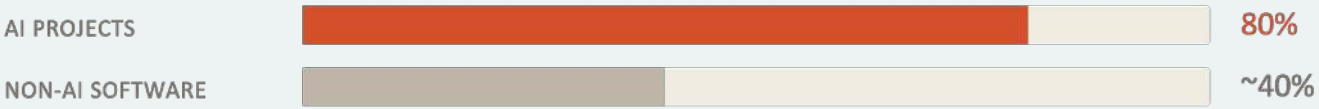
MSc Sustainable Energy Engineering

University of Ibadan

BSc Petroleum Engineering

AI makes it easy to ship products at an astonishing pace

80% of AI projects fail.



Twice the rate. Same teams. Same talent. Different game.

RAND, 2024

Agenda

01 Intro

02 Why Products fail?

03 How to build in the AI Era

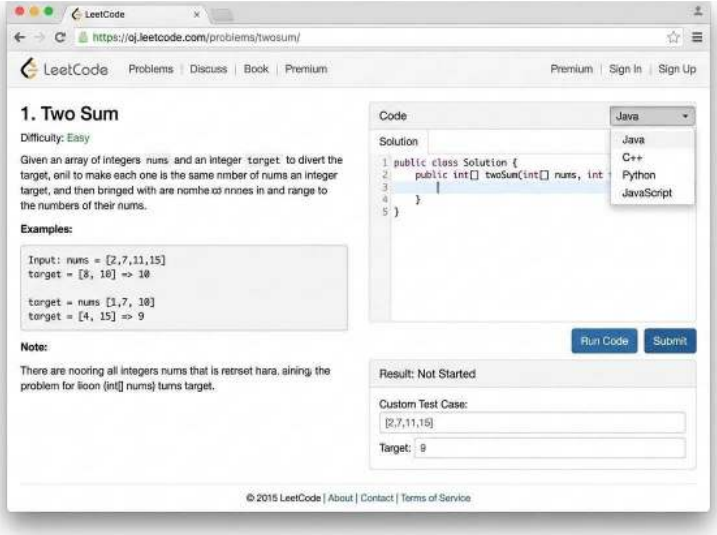
04 The role of builders

05 Conclusion

BEFORE THE AI REVOLUTION

How long would it have taken to build this? (circa 2015)

EXERCISE 1



A. A weekend

B. Two weeks

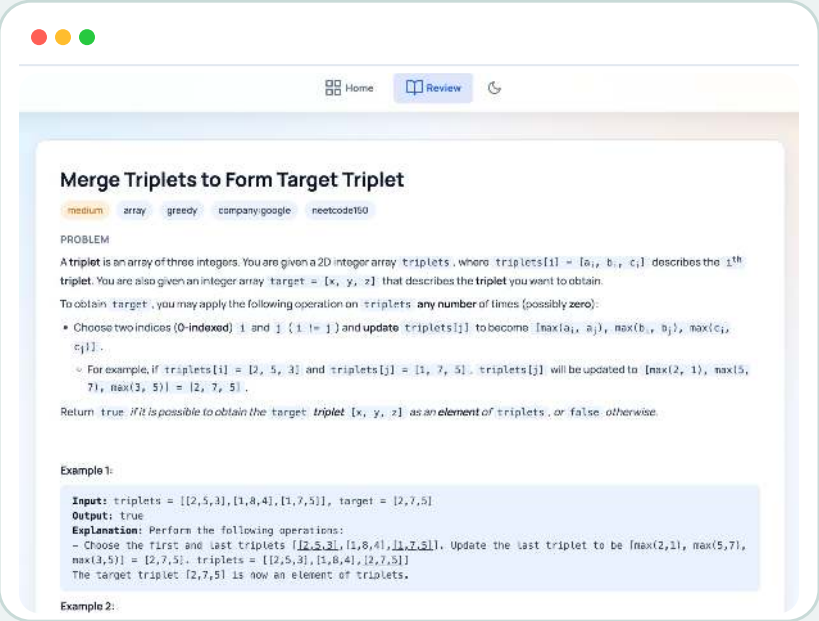
C. Two months

D. Six months or more ✓

TODAY: THE AI ERA

How long did this version take to build? (circa 2025)

EXERCISE 2



A. A weekend ✓

B. Two weeks

C. Two months

D. 6 months or more

AI MADE BUILDING CHEAP

What you need to build

1. An LLM

+



2. A finger

+



3. A computer

+



Your voice

AI MADE BUILDING CHEAP

Personal Story: Recode



Merge Triplets to Form Target Triplet

medium · array · greedy · company-google · neetcode150

PROBLEM

A **triplet** is an array of three integers. You are given a 2D integer array `triplets`, where `triplets[i] = [a_i, b_i, c_i]` describes the ith **triplet**. You are also given an integer array `target = [x, y, z]` that describes the **triplet** you want to obtain.

To obtain `target`, you may apply the following operation on `triplets` **any number** of times (possibly **zero**):

- Choose two indices (**0-indexed**) `i` and `j` (`i != j`) and **update** `triplets[j]` to become `[max(a_i, a_j), max(b_i, b_j), max(c_i, c_j)]`.
  - For example, if `triplets[i] = [2, 5, 3]` and `triplets[j] = [1, 7, 5]`, `triplets[j]` will be updated to `[max(2, 1), max(5, 7), max(3, 5)] = [2, 7, 5]`.

Return `true` *if it is possible to obtain the* `target` ***triplet*** `[x, y, z]` *as an* ***element*** *of* `triplets`, *or* `false` *otherwise*.

Example 1:

```
Input: triplets = [[2,5,3],[1,8,4],[1,7,5]], target = [2,7,5]
Output: true
Explanation: Perform the following operations:
- Choose the first and last triplets [[2,5,3],[1,8,4],[1,7,5]]. Update the last triplet to be [max(2,1), max(5,7), max(3,5)] = [2,7,5]. triplets = [[2,5,3],[1,8,4],[2,7,5]]
The target triplet [2,7,5] is now an element of triplets.
```

ORIGINAL PROMPT

"Can I rebuild a personalized revision app for coding problems that will make it easier to revise coding problems? I often struggle with remembering past Data Structures & Algorithms problems I have solved, so I wanted to build a way to revise that."

I BUILT SOMETHING FOR MYSELF

I didn't use it.

Which is slightly embarrassing when you are also the person who built it.

THE KEY FORMULA

Ease of development ≠ Guaranteed value

RECODE APP METRICS

16	498	0
REVIEWED, ALL TIME	TOTAL PROBLEMS	DAY STREAK

RESULT

- **Unclear Core Problem:** Skipped mapping a single end-to-end user story to test what actually moves the needle.
- **Over-Engineered Complexity:** Built too many features too quickly instead of validating basic usage flow.
- **Lost Time & Compute:** Resulted in wasted development hours, idle deployment, and unused resources.

AI MADE BUILDING CHEAP

What are we missing when we build too fast?

1. An LLM

+

~~2. A finger~~

+

3. A computer

+

Your voice

What are we missing?

When we build at pace, we may be missing something crucial.

PART 1

Why AI Products Fail?

STORY 01

Sora – Cost

~$1M/day
compute burn

$2.1M
total lifetime revenue

~1%
reported day-30 retention

~6 mo
standalone app, launch to wind-down

Capability alone may not carry a product. Unit economics and retention still seem to decide it.

Figures are reported / third-party estimates, not official.

Reported by WSJ / CNBC, Forbes & Appfigures (est.) · app discontinued Apr 2026

STORY 02

McDonald's × IBM – Evaluation

3 yrs
100+ locations

~85%
reported accuracy vs 90%+ human

260
McNuggets in one viral order

Jul '24
IBM partnership ended

An eval framework that only measures the best case can miss the cases that reach customers.

CNBC, Jun 2024 · McDonald's & IBM statements
New 2026 Experiments with Arch IQ (a Google product)

STORY 03

Klarna – User Experience

~700
agents' worth of work

2.3M
chats in month one handled by AI and human job cuts

May '25
CEO admits quality drop

2025
rehiring humans

The metric you optimize for (e.g. cost) shapes what you get. Klarna failed to optimize for customer experience

CEO to Bloomberg, May 2025

STORY 04

Microsoft Recall – Trust & Privacy

May ’24
announced for Copilot+ PCs

5 sec
automatic desktop snapshots

Jun ’24
preview delayed & made opt-in

2024
security & privacy redesign

Without user trust, capability alone isn't enough to ship.

Reported by The Verge / Wire · Microsoft Security Blogs

THE PATTERN

The Failure Patterns

COST

Sora

No unit economics.

EVALUATION

McDonald's

The wrong eval.

USER EXPERIENCE

Klarna

The wrong metric.

TRUST

Microsoft Recall

Lack of user trust.

Strong models. Real resources. Each seems to have stumbled for reasons that had little to do with the model itself.

PART 2

How to build in the AI Era?

AI MADE BUILDING CHEAP

What you actually need to build

1. An LLM

+

~~2. A finger~~

+

3. A computer

+

Your voice

Clear Product Experience

Robust Evaluation Systems

Human In the Loop

PART 2 · THE FRAMEWORK

Three buckets

01

Before you build

Cost, User experience, trust

02

Building & shipping

Evals & calibration

03

After launch

Personalization and Human In the Loop

AFROTECH Conference

BUCKET 01

Before you build

User Experience, Cost & Trust

Olukunle Owolabi, PhD

BUCKET 01 · BEFORE YOU BUILD

Start with the user, not the model

Building valuable AI products requires clear user experiences and explicit upfront expectations.

PRINCIPLE 01

The Experience IS the Product

PRINCIPLE 02

Clear Upfront Expectations

BUCKET 01 · BEFORE YOU BUILD

Cost is a day-one decision

Architect your product to match model costs with user value from day one.

LEVER 01 · ROUTING & TIERING

Multi-Model Routing

Route basic tasks to open-source models. Save flagships for core reasoning.

⚡ **Goal: Reserve flagship LLMs for ~10% reasoning**

LEVER 02 · VECTOR CACHING

Semantic Caching

Store outputs and embeddings in a vector cache to skip duplicate model calls.

🚀 **Goal: Near-zero latency & zero cost on cache hits**

BUCKET 01 · BEFORE YOU BUILD

Building Trust in the Pre-Build Stage

⚠️ **PRE-BUILD MANDATE:** Design user-facing trust mechanisms before building core AI capabilities.

PILLAR 01 · RISK AUDIT

User Control Gates

- Audit high-risk workflows
- Identify sensitive data touchpoints
- Build explicit verification gates

PILLAR 02 · EXPLAINABILITY

Data & Citations

- Bind AI outputs to source data
- Provide interactive citations
- Eliminate black-box responses

PILLAR 03 · CONSENT

Opt-In & Access

- Require explicit user opt-in
- Granular data permissions
- Provide transparent user control

AFROTECH Conference

BUCKET 02

Building & shipping

Evals & Quality

Olukunle Owolabi, PhD

BUCKET 02 · BUILDING & SHIPPING

Set the quality bar first

LEVEL 01 · DESIRED

Good

what we aim for

LEVEL 02 · THRESHOLD

Acceptable

what we tolerate

LEVEL 03 · REJECT

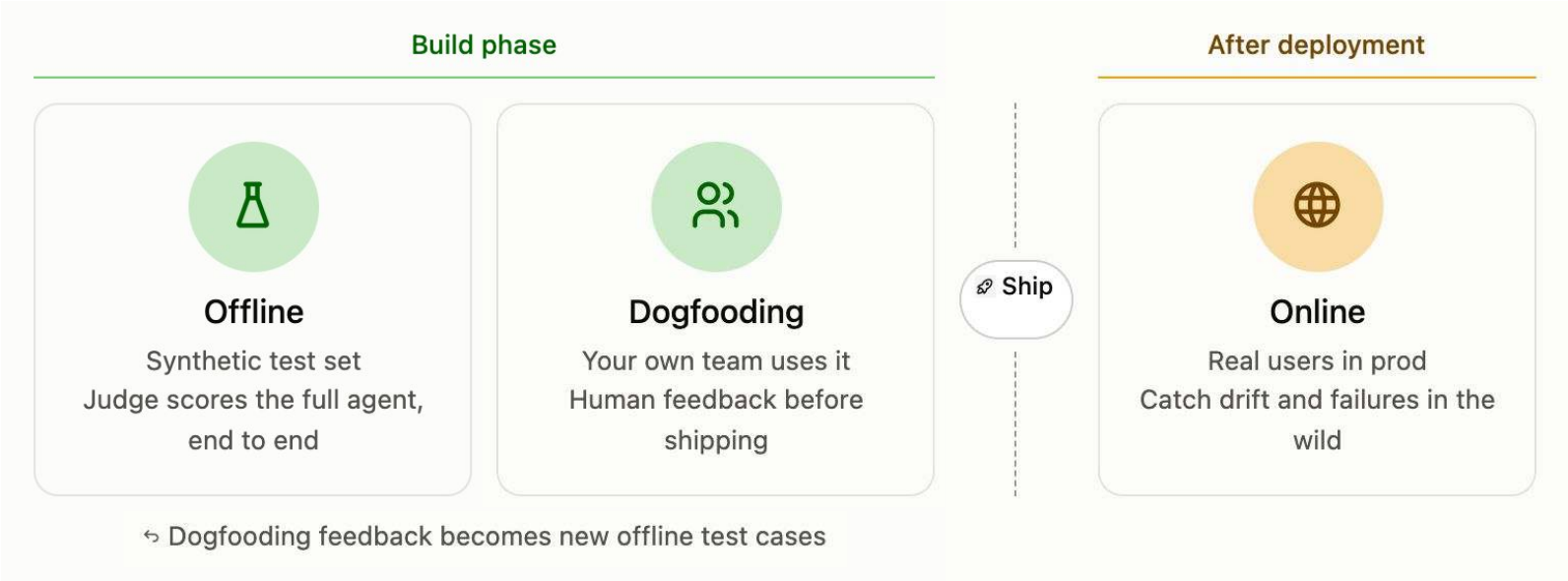
Hard no

what we never ship

The Golden Rule: **Define good before you ship, not after users complain.** Evals measure whether you hit the bar; the quality bar says where the bar is.

BUCKET 02 · BUILDING & SHIPPING

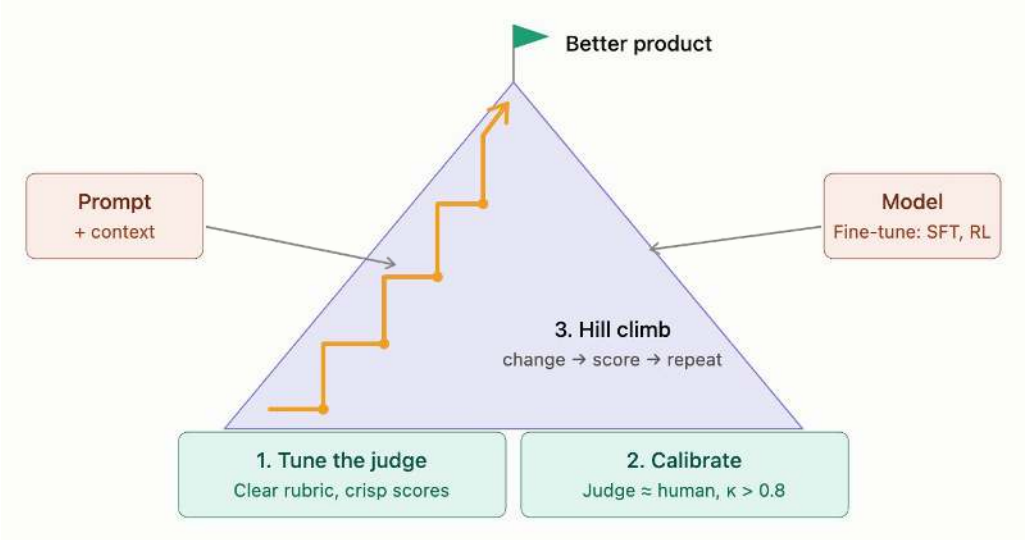
The Three-Tier Eval Pipeline



BUCKET 02 · BUILDING & SHIPPING

Calibration & Hill Climbing

Build a trusted evaluator first, then systematically hill-climb your product quality.



BUCKET 03

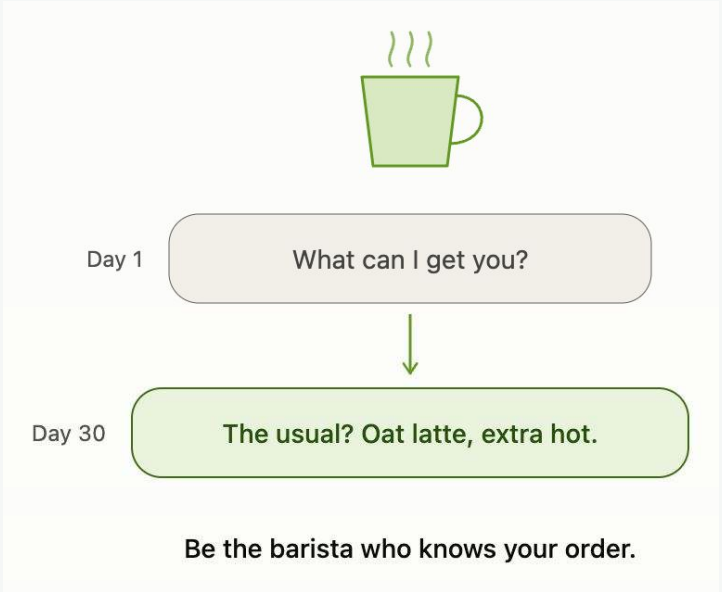
After launch

Personalization

Olukunle Owolabi, PhD

BUCKET 03 · AFTER LAUNCH

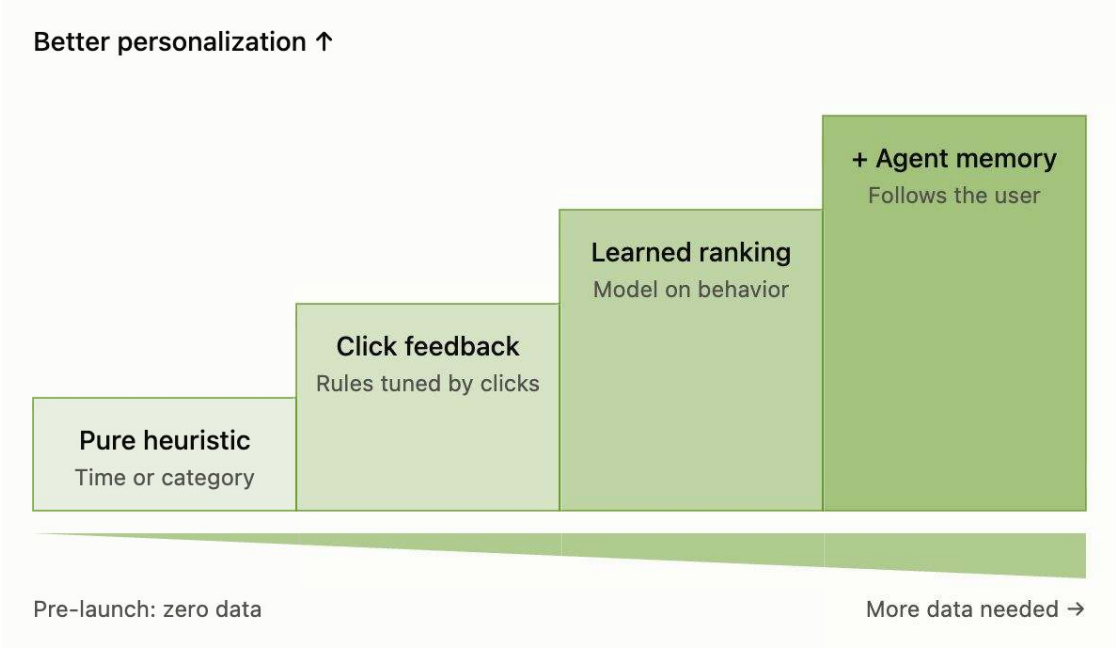
Personalization is the ceiling



BUCKET 03 · AFTER LAUNCH

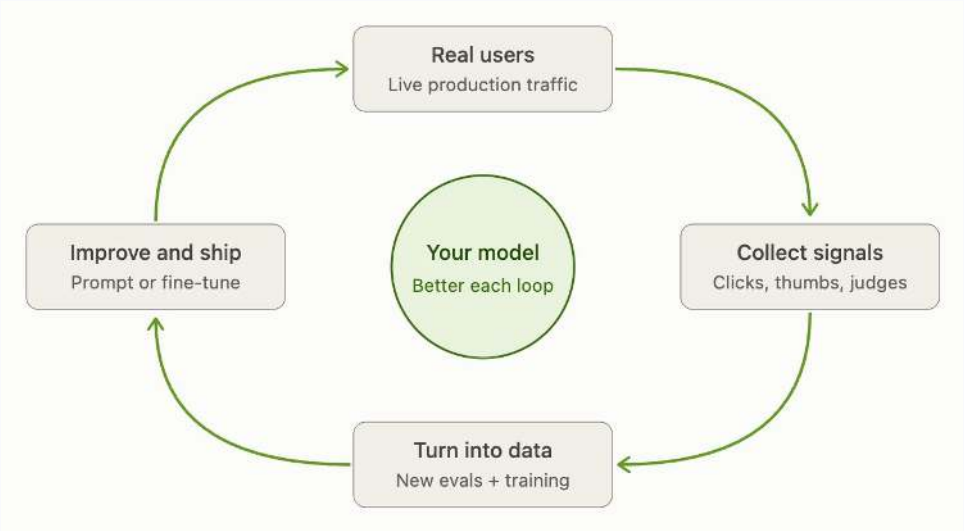
Stages of AI Product Personalization

Each step gives you a better ranking than the last, at the cost of needing more data to get there.



BUCKET 03 · AFTER LAUNCH

Continuous Feedback, with Privacy & Trust

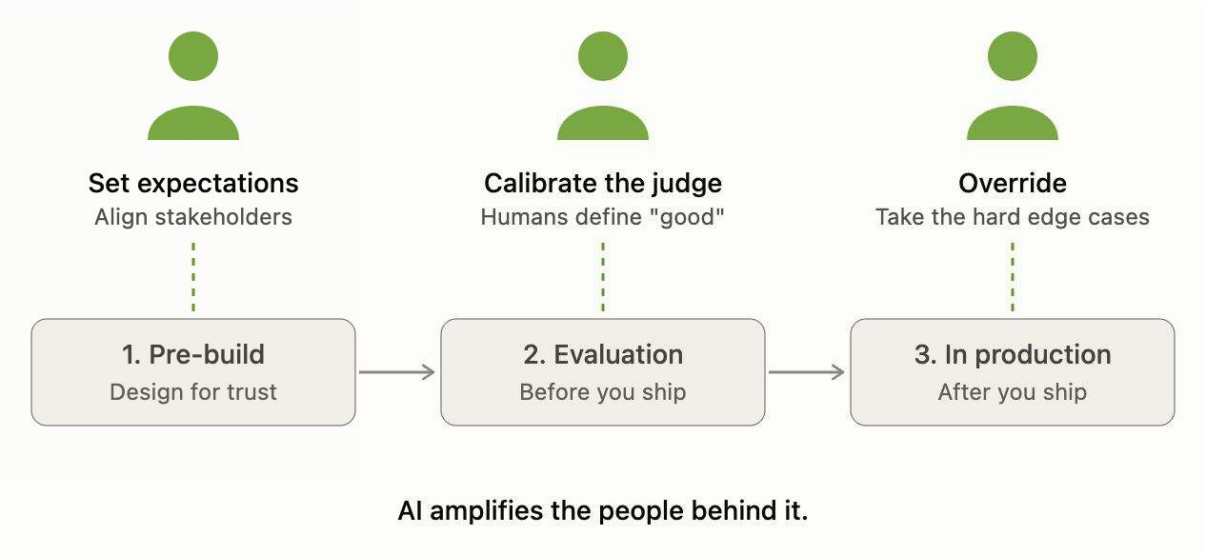


PART 3

The role of builders

PART 3 · THE ROLE OF BUILDERS

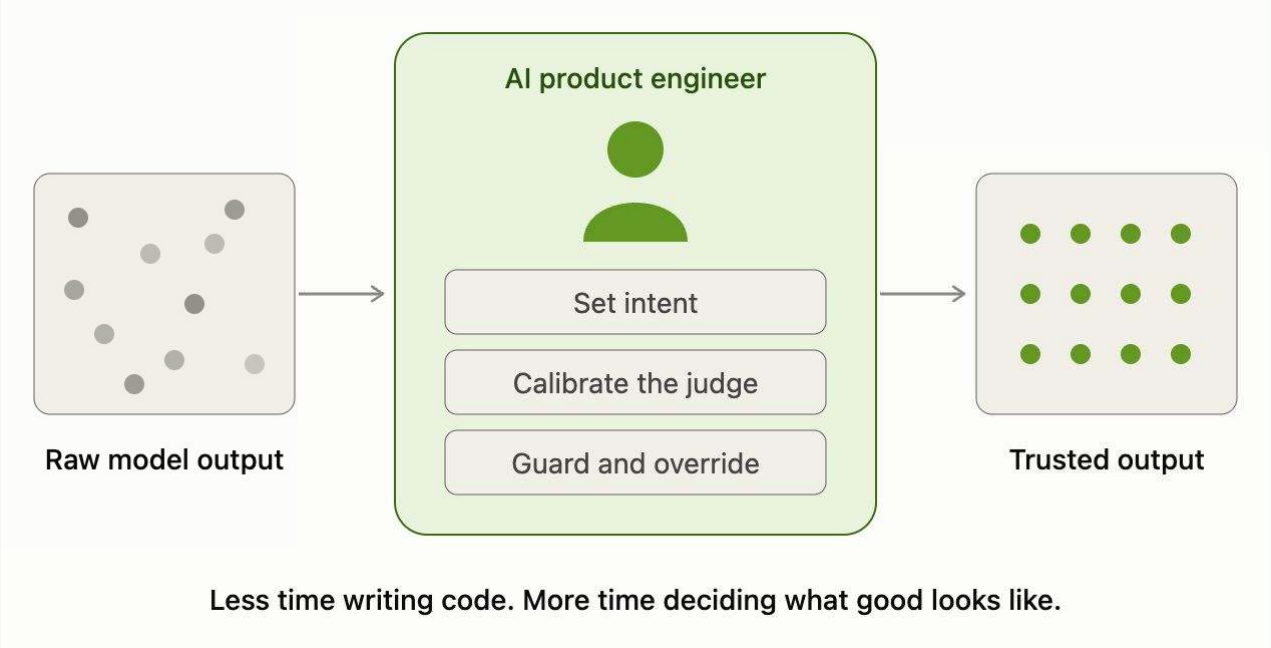
Human in the Loop



PARADIGM SHIFT

The AI Product Engineer

Shifting from implementation to intent.

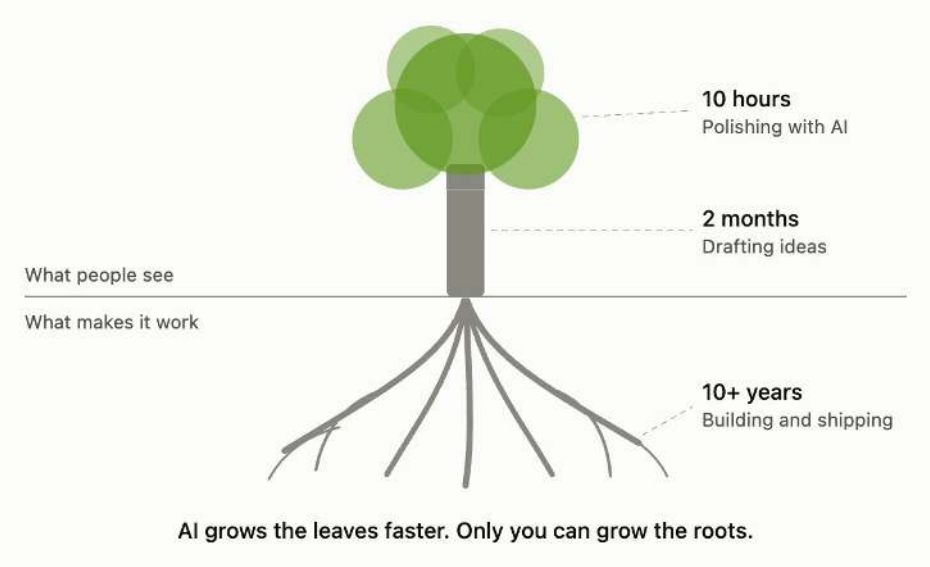


Less time writing code. More time deciding what good looks like.

WHY THE BUILDER IS VITAL

A Builder's Judgement Is Irreplaceable

AI automates execution, but hard-earned experience remains the ultimate moat.



AFROTECH Conference

SUMMARY & KEY TAKEAWAYS

Conclusion

Why AI Products Fail

AI products fail frequently, mostly due to unclear product experience, high cost, lack of trust, and lack of robust AI evaluation systems.

What We Need to Do

Essential actions required across pre-build, building, and post-deployment stages to build, create, and scale high-value AI products.

The Builder's Role

The vital role of the builder in driving judgement, oversight, and execution across all of these phases.

AFROTECH™ Conference

Thank You & Questions?



SPEAKER

Olukunle Owolabi, PhD

AI Engineer at Meta
Researcher · Builder

LET'S CONNECT

LinkedIn
Olukunle Owolabi, PhD

Website
olukunleowolabi.com

THANK YOU ✦ QUESTIONS? ✦ SCAN THE QR TO CONNECT ✦ THANK