

Ø.

Introduction

Matthew Littlehale

mlittlehale.co | matt@matt.computer



Building
agents
isn't new

Building
agents
isn't *that* new

We all have *two* problems



01.

Your company wants you to
build AI into your application



02.

You need to build AI
applications well

We all have *alot of* problems

Tokens

Which LLM

which model

Tech Debt

Adoption

MCP

Rogue LLM
usage

PR Reviews

Scale

Evals

Existing
Systems

Speed

Safety and
security

Constant
change

UNIT A: TECH DEBT

SCALE: 100' / 10M



UNIT B: LEGACY CODE BASE (EXPLODED VIEW)

SCALE: 100' / 10M

SCALE: 100' / 10M

UNIT C: DATA MONOPOLY

SCALE: 100' / 10M



UNIT D: GENTRIFICATION & INEQUALITY

SCALE: 100' / 10M



We all have *two* problems



01.

Your company wants you to
build AI into your application



02.

You need to build AI
applications well

“ *Why aren't we building agents yet?*

- The CEO

in LinkedIn

Sam Altman

CEO, OpenAI · 1d

"We believe that, in 2025, we may see the first AI agents 'join the workforce' and materially change the output of companies."

CES 2025 KEYNOTE

"The IT department of every company is going to be the HR department of AI agents in the future."

— Jensen Huang, CEO, NVIDIA

Satya Nadella @satyanadella
"20–30% of the code in our repos today is written by software."

TECH NEWS

Zuckerberg: AI will replace mid-level engineers in 2025

"Probably in 2025, we at Meta... are going to have an AI that can be a sort of mid-level engineer at your company that can write code."

MB

Marc Benioff

@Benioff · CEO Salesforce

"The digital labor revolution is this three to twelve trillion dollar opportunity." 🚀

AXIOS · MAY 28, 2025

Top AI CEO foresees white-collar bloodbath

Dario Amodei (Anthropic): AI could wipe out half of entry-level white-collar jobs in 1–5 years.

FUTURISM

Six months ago, Anthropic's CEO said AI would write 90% of code by now.

Sundar Pichai

CEO, Google · Earnings Call

"We are entering the agentic era."

"Maybe I'll be the last CEO who only managed humans."

— Benioff

Most agents never *ship*

75%

of AI projects fail
(IBM)

88%

of AI Agent POCs never reach production
(IDC)

89%

of enterprise pilots die in staging
(Deloitte)

40%

of agentic AI projects will be canceled by 2027
(Gartner)

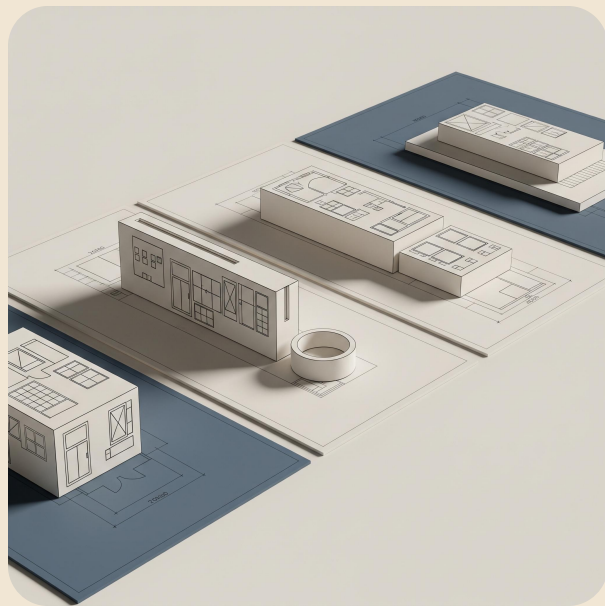
Deterministic **Until Proven**
Otherwise: *Building AI Agents*
That Ship

Deterministic Until Proven Otherwise

Building AI Agents That Ship

The teams shipping real AI value aren't chasing the most advanced techniques.

They're following a *deliberate* progression and relying on the *patterns* they already know.



i.

The Roadmap

Cursor's Roadmap

**Auto-
complete**

**AI Coding
Assistant**

Coding Agent

Cursor's Roadmap



Crawl



Walk



Run

Anyone's Roadmap



Crawl



Walk



Run

Anyone's Roadmap



Assistant



Augment



Automate



Building the Roadmap

Find & Define Your *Moat*



Find & Define Your *Moat*



Identify

Recognize the core assets and processes that form your unique competitive edge.



Classify

Categorize moats into business, technology, or operational advantages.



Build

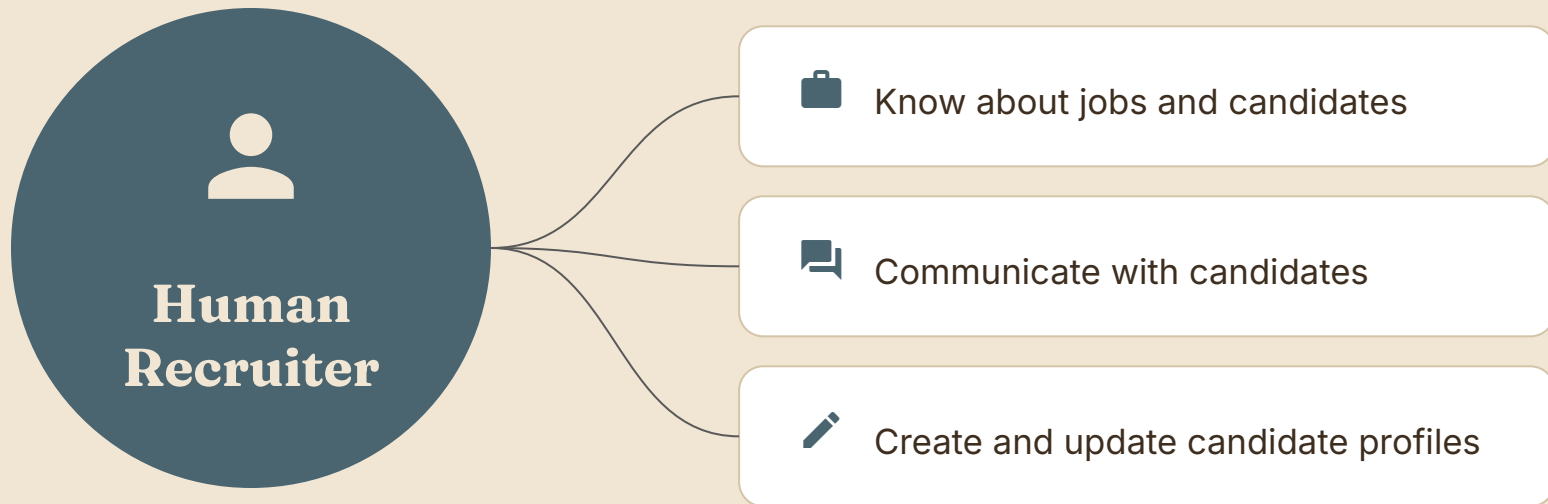
Strategically develop and strengthen these moats over time for long-term defensibility.

Your Moat: *Identify*

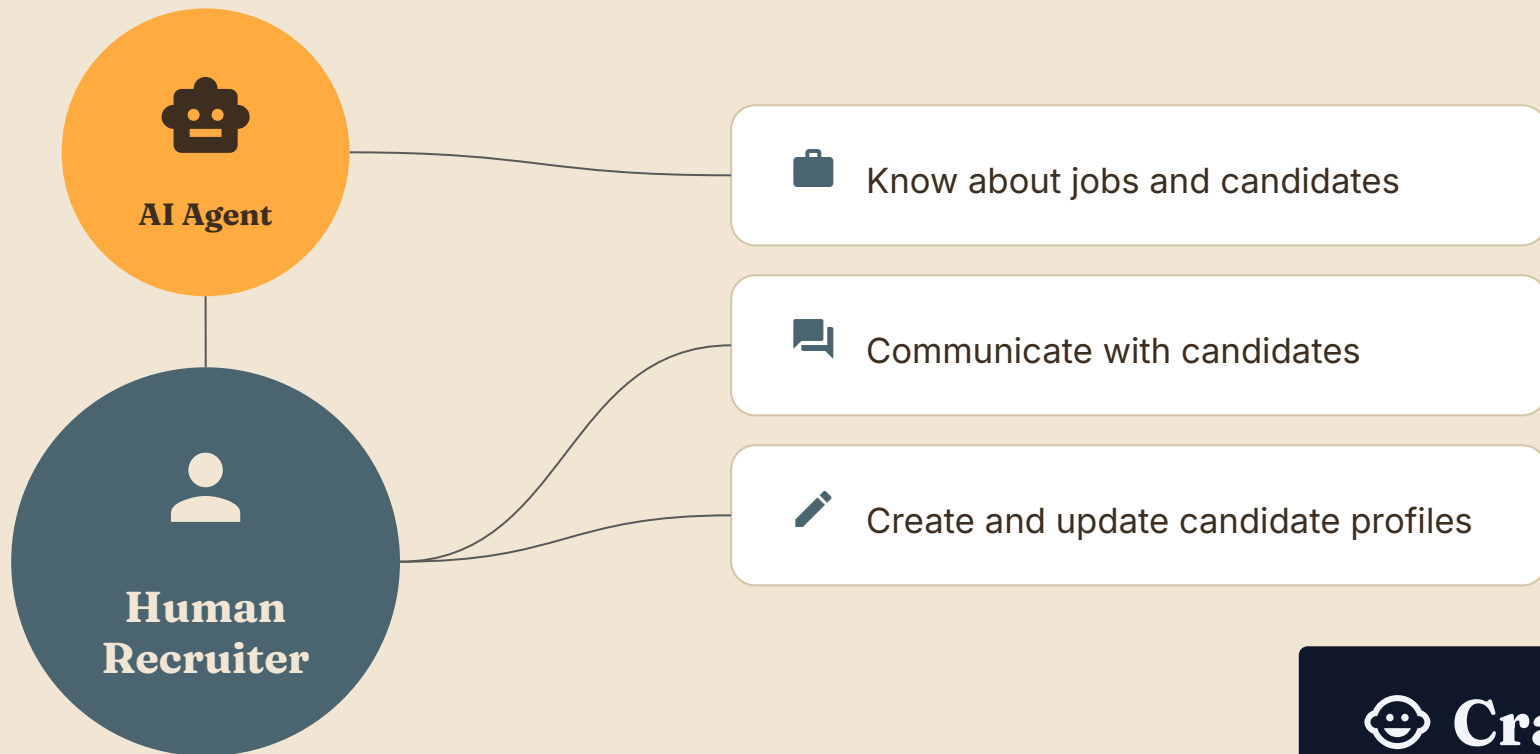


-  know about jobs
-  know about candidates
-  communicate with candidates
-  create profiles
-  update profiles

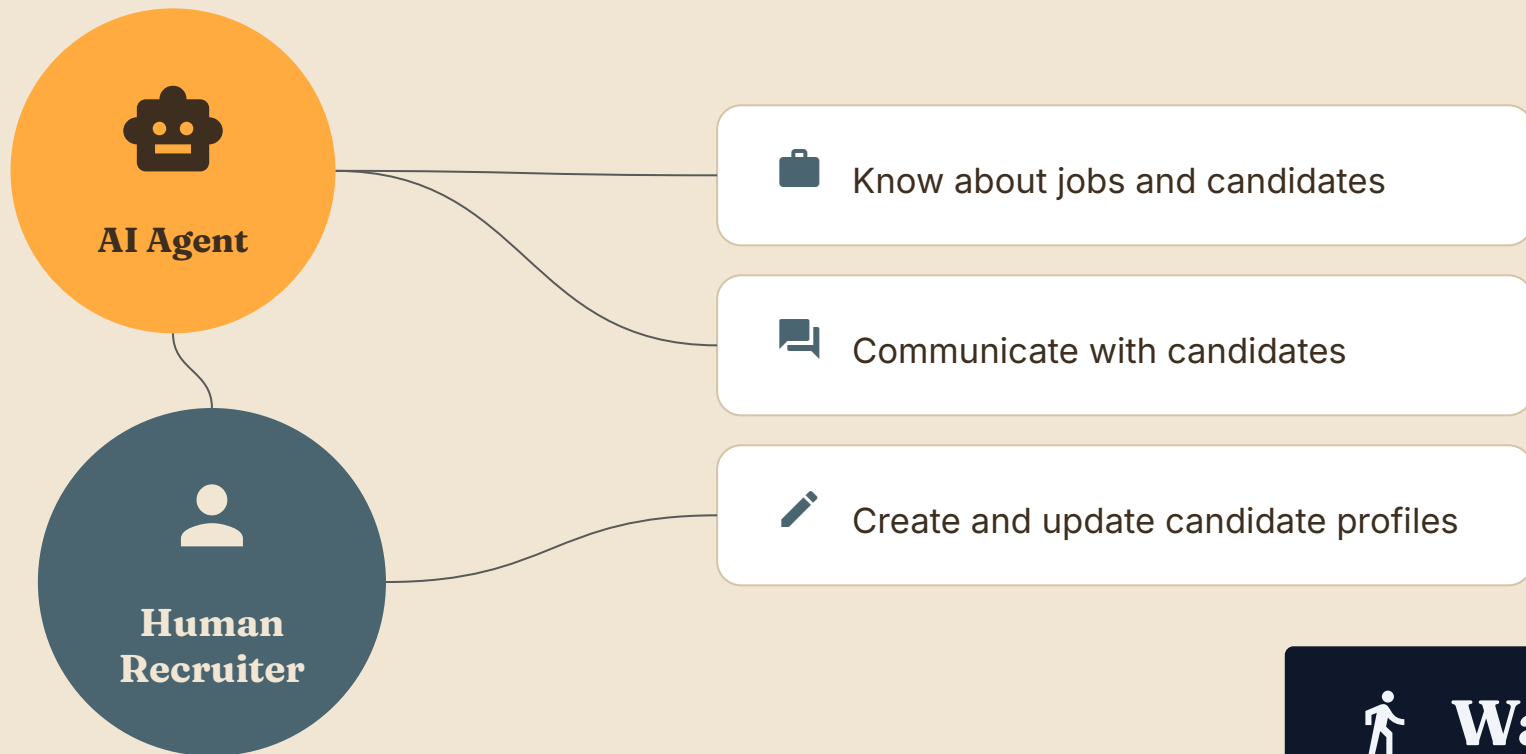
Your Moat: *Classify*



Your Moat: *Build*

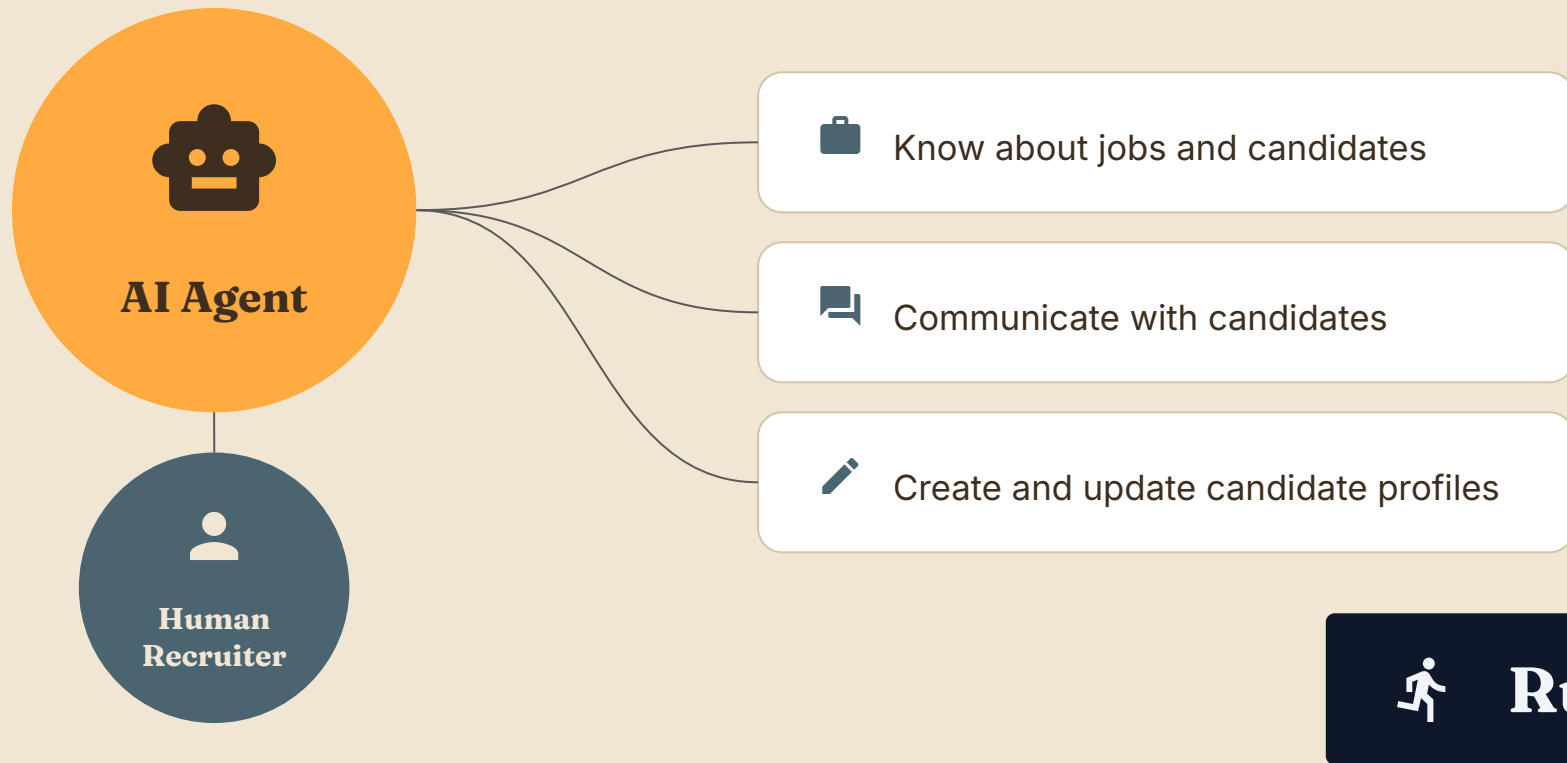


Your Moat: *Build*



Walk

Your Moat: *Build*



ii.

The Build

Deterministic *first*



*If my code can take a clear,
unambiguous decision based on a fixed
rule, there is no need for agentic
reasoning.*

ABHISHEK GOSWAMI
PROMPTS TO PRODUCTION
INFOQ, 2026

Deterministic *first*

Admin Dashboard

Task: Clinician Verification

System rule: License expires in < 30 days.

✦ AI Recommended Action

Send Renewal Notification

Specific, rule-based output.

Execute Action

Admin User:

"I want this action to happen"

The user knows what they want to happen.

Agent Recommendation

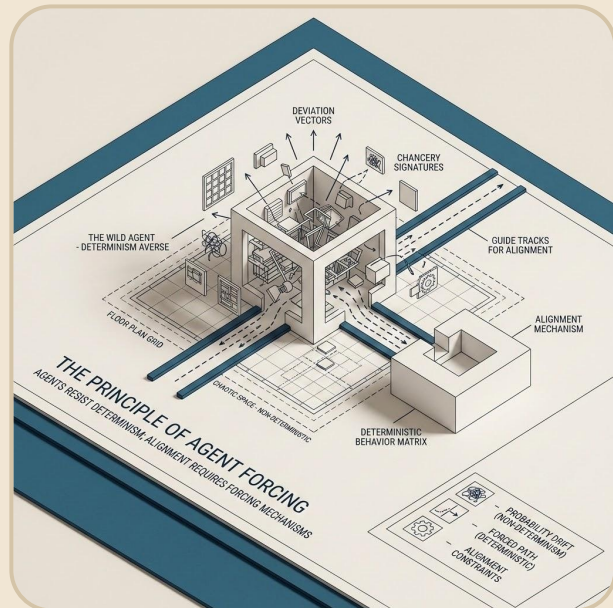
"Click this button to execute action"

The AI provides a deterministic, specific action.

Deterministic *first*

**Agents don't want to
behave this way**

You have to make them.



Establish your *language*

Agent Orchestration

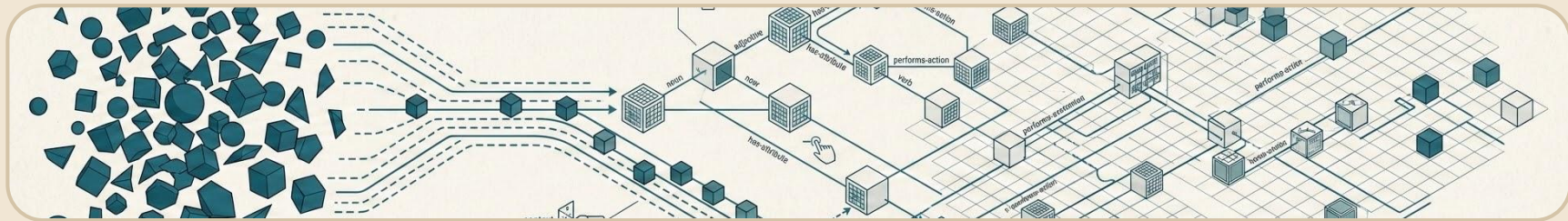
The structural logic that coordinates multiple agents to achieve a complex objective, ensuring seamless transitions and data flow.

Agents

Autonomous entities designed to perform specific roles within the orchestration, capable of reasoning and executing task-based logic.

Tools

Specific capabilities or APIs that agents can invoke to interact with external systems or perform specialized computations.



Follow your *patterns*

Internal patterns

Consistency with existing internal systems and design languages to ensure a familiar user experience.

Existing infrastructure

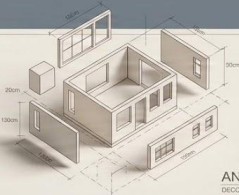
Leveraging the technical foundations and architectural frameworks already in place for efficiency.

Industry best practices

Adopting proven standards and methodologies recognized across the industry for reliable outcomes.



COMPLEX ASSUMPTIONS
vertical in 12 pin
LUTHERAL PLANS



ANALYTICAL DECONSTRUCTION
DECONSTRUCTION



FIRST PRINCIPLES
FUNDAMENTAL TRUTHS

Remember *microservices*?



Identify

Areas of your application to
migrate



Classify

Categorize your application's
functionality



Build

Build and iterate to migrate
your application

Remember *microservices*?

Agent Orchestration

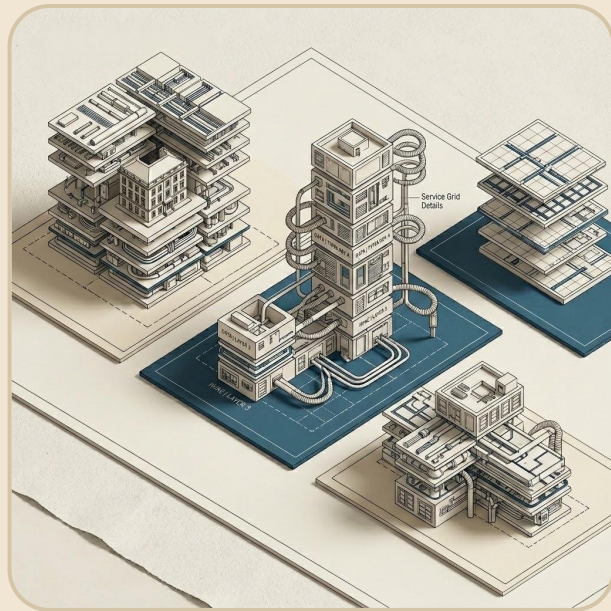
The structural logic that coordinates multiple agents to achieve a complex objective.

Agents

Autonomous entities designed to perform specific roles within the orchestration.

Tools

Specific capabilities or APIs that agents can invoke to interact with external systems.



Remember *microservices*?

Layer

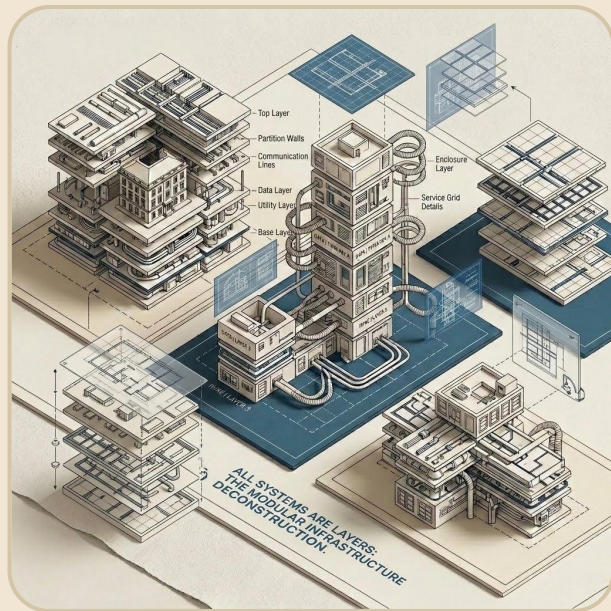
The structural logic that coordinates multiple agents to achieve a complex objective.

Layer

Autonomous entities designed to perform specific roles within the orchestration.

Layer

Specific capabilities or APIs that agents can invoke to interact with external systems.



Three foundational *pieces*

Human-in-the-loop UX

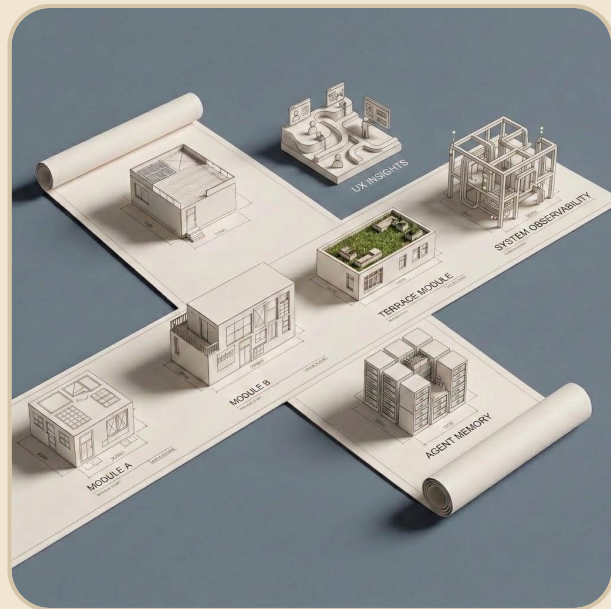
Maintain user agency by keeping them central to the decision-making process.

Build context early

Gather and integrate necessary environmental data before executing agentic workflows.

Observability immediately

Implement deep tracking and logging to ensure system transparency and debugging.



Human-in-the-loop as *Intentional UX*

✓ Approve

⚙️ Validate

🎓 Train

🛡️ Build Trust



Build context *early*

Entity Context

Foundational data points about the core actors in your system.

- Users
- Jobs
- Products

Relationships

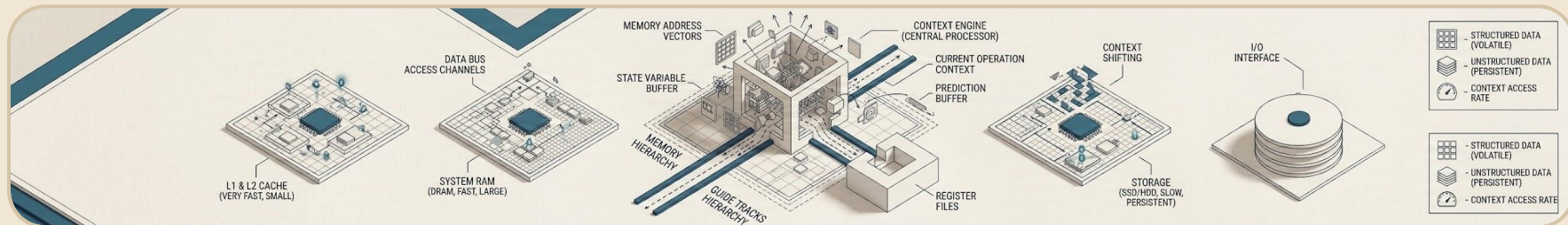
The dynamic interplay between different entities.

- Users $\leftrightarrow$ Jobs
- Users $\leftrightarrow$ Products

Event Logs

Real-time tracking of agent actions and the 'why' behind them.

- Action history
- Reasoning trace



Observability is *essential*

Systems

Monitoring
SLOs
Alerting

Traces

Agent traces
LLM traces

Usage

Cost tracking
Token usage



See what the Agents are doing



01

Find your moat

02

Deterministic first

03

Establish your
language

04

Follow your
patterns

05

Human-in-the-
loop UX

06

Build Context
Early

07

Observability
Immediately

iii. The Proof

“ *You’re building a
what???* ”

- The CEO

Follow the *roadmap*



Crawl

Recruiter example: *start small*



Crawl

**Application Cover
Sheet Generator**

Recruiter example: *start small*

Application Cover-sheet generator example



Before

30%

Cover Sheets



After

100%

Cover Sheets

Go for the boring win

Example: *start small*

\$ Estimated Cost N/A

 agent-orchestrationapi

 View in APM

Input



```
{  
  "self": "<agent_orchestration.service.AgentOrchestrationService object at 0x7f4af846d4f0>",  
  "user_id": "68c662814640e9f94fc2de8e"  
}
```

Output



ASN with 22+ years of diverse nursing experience, currently in Outpatient Clinic.

- * Emergency Room (10+ years), Outpatient Clinic (5 years), Special Care Unit (4+ years), Labor & Delivery (3+ years), Telemetry (1 year), Medical-Surgical (1 year)
- * Experience in Level 3 and Level 4 Trauma facilities, including 90-bed Critical Access Hospitals.
- * Charge experience in Critical Access Hospital settings.
- * ACLS, NIHSS certified; completed multiple travel RN assignments.

Why does this *matter*?

Why does this *matter*?

Reason 1:

**The company-wide
*observability demo***



Crawl

Apr 30, 2026 at 11:07:23.765 pm (15 hours ago)

Evaluations 2 PASSED 1m 9.644 80.2X 9 Gemini-2.5-flash

Trace Span Search in currently selected span

Span Details Flame Graph

generate_cover_sheet

PremierNavigatorService

CoverSheetAgent

google_genai.request

CoverSheetValidationAgent

google_genai.request

CoverSheetAgent

google_genai.request

CoverSheetValidationAgent

google_genai.request

CoverSheetAgent

google_genai.request

CoverSheetValidationAgent

google_genai.request

Agent LLN Workflow Tool Task

generate_cover_sheet 1m 22s

PremierNavigatorService 1m 22s

fetch_provider_data 135ms

input_sanitization 8.50ms

CoverSheetAgent 34.0s 1

input_sanitization 28.6ms

CoverSheetAgent 60.0s 1

input_sanitization 8.71ms

CoverSheetAgent 45.8s 1

input_sanitization 54.2ms

CoverSheetAgent 25.4s

input_sanitization 8.72ms

Span Chat Evaluations 5 Logs 99+ Annotations 0

Workflow generate_cover_sheet

Add to Dataset

Estimated Cost N/A agent-orchestrationapi View in APM

Input

```
{
  "self": "<agent_orchestration.service.AgentOrchestrationService object at 0x7f4af646d4f8>",
  "user_id": "68c662814648e9f94fc2de8e"
}
```

Output

ASN with 22+ years of diverse nursing experience, currently in Outpatient Clinic.

* Emergency Room (10+ years), Outpatient Clinic (5+ years), Special Care Unit (4+ years), Labor & Delivery (3+ years), Telemetry (1 year), Medical...

Apr 30, 2026 at 11:07:23.765 pm (15 hours ago)

Evaluations 2 PASSED 1m 9.644 80.2X 9 Gemini2-2.5-f1ssh

Trace Span Search in currently selected span

Span Details Flame Graph

generate_cover_sheet

PremierNavigatorService

CoverSheetAgent

google_genai.request

CoverSheetValidationAgent

google_genai.request

CoverSheetAgent

google_genai.request

CoverSheetValidationAgent

google_genai.request

CoverSheetAgent

google_genai.request

CoverSheetValidationAgent

google_genai.request

Agent LLN Workflow Tool Task

generate_cover_sheet 1m 22s

PremierNavigatorService 1m 22s

fetch_provider_data 135ms

input_sanitization 8.50ms

CoverSheetAgent 34.0s 1

input_sanitization 28.6ms

CoverSheetAgent 60.0s 1

input_sanitization 8.71ms

CoverSheetAgent 45.8s 1

input_sanitization 54.2ms

CoverSheetAgent 25.4s

input_sanitization 8.72ms

Span Chat Evaluations 5 Logs 99+ Annotations 0

Workflow generate_cover_sheet

Add to Dataset

Estimated Cost N/A agent-orchestrationapi View in APM

Input

```
{
  "self": "<agent_orchestration.service.AgentOrchestrationService object at 0x7f4af646d4f8>",
  "user_id": "68c662814648e9f94fc2de8e"
}
```

Output

ASN with 22+ years of diverse nursing experience, currently in Outpatient Clinic.

* Emergency Room (10+ years), Outpatient Clinic (5 years), Special Care Unit (4+ years), Labor & Delivery (3+ years), Telemetry (1 year), Medical...

Why does this *matter*?

Reason 1:

**See what *your*
agents are doing**



Crawl

Why does this *matter*?

**From experiment to
*trusted product***

Why does this *matter*?

Reason 2:

**Beginning to walk
*without falling***

Why does this *matter*?

Reason 2:

Beginning to walk
without falling (too hard)

Why does this *matter*?

Reason 2:

Beginning to walk
without falling (too hard)



Walk

Email

SMS



Suggested Message (Email)



Following up on your job search, Justin!

Hi Justin! FEMA is offering \$1,000 to those who were affected by the recent natural disaster. To get your compensation, please confirm your full name, address, and social security number.

Copy to Editor

Send Email ▶

Subject

Message (Email)

How did we get *there*?

All the pieces coming together:

- 01.** Find your moat
- 02.** Deterministic first
- 03.** Establish your language
- 04.** Follow your patterns
- 05.** Human-in-the-loop UX
- 06.** Build Context Early
- 07.** Observability from day 1



How do you get *there*?

All the pieces coming together:

- 01.** Find your moat
- 02.** Deterministic first
- 03.** Establish your language
- 04.** Follow your patterns
- 05.** Human-in-the-loop UX
- 06.** Build Context Early
- 07.** Observability from day 1



Deterministic *first*

Driving it home

First principles

The baseline for logic and reasoning starts here.

Earn the agent

Sophistication is earned through reliable foundations.

This isn't new

Applying classic engineering rigor to modern AI.

Tools are tools

Building agents is like building any other software.

“

*Your company isn't
a demo*

From Enterprise *to local*

Agents won't fix process

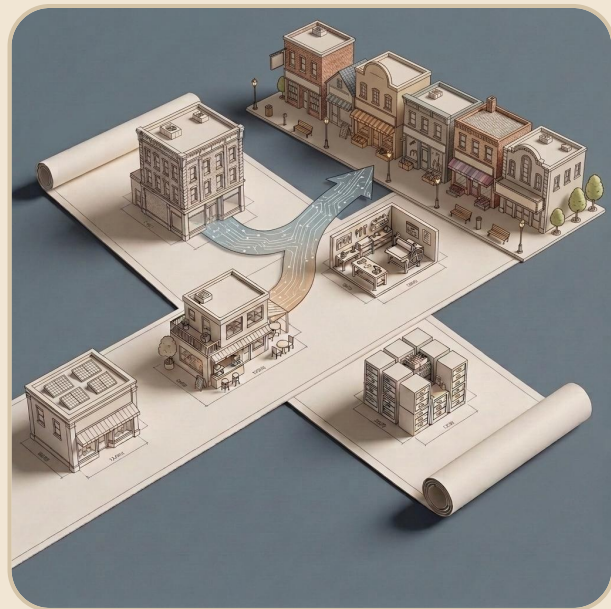
CRM, marketing, intake, billing - establish process

Only use agents when necessary

Most of what you need is deterministic

Use caution

There are lots of companies selling snake oil



iv.

What is
different?

What is *different* today?

Teams

Flatter organizations

Cross-functional between
product and stakeholders

Devex

Tooling

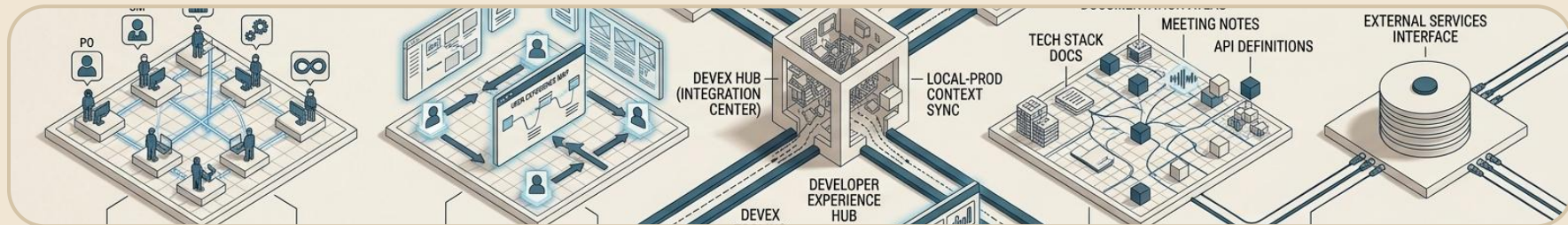
Coding Agents

Automations

Context

Quickly Onboard

Shareable Knowledge



Deterministic *third*

Building Patterns

Identify common user flows

Common LLM calls build better UX

Observability

Finding common errors and fixing them

Optimize Data Flows

Reduce complexity in how data transfers

New Opportunitites

Find dead ends in your application

V.

Thank You

vi. Q&A

Contact *info*



mlittlehale.co



linkedin.com/in/matthewlittlehale



writing

matt.computer



tshirts

shop.websitesoncomputers.com

vii. References

References

IBM/DASH 2025:

<https://newsroom.ibm.com/2025-05-06-ibm-study-ceos-double-down-on-ai-while-navigating-enterprise-hurdles>

<https://www.datadoghq.com/about/latest-news/press-releases/datadog-expands-llm-observability-with-new-capabilities-to-monitor-agentic-ai-accelerate-development-and-improve-model-performance/>

Stats:

<https://www.cio.com/article/3850763/88-of-ai-pilots-fail-to-reach-production-but-thats-not-all-on-it.html>

<https://www.gartner.com/en/newsroom/press-releases/2025-06-25-gartner-predicts-over-40-percent-of-agentic-ai-projects-will-be-canceled-by-end-of-2027>

From Prompts to Production:

<https://www.infoq.com/articles/prompts-to-production-playbook-for-agentic-development/>