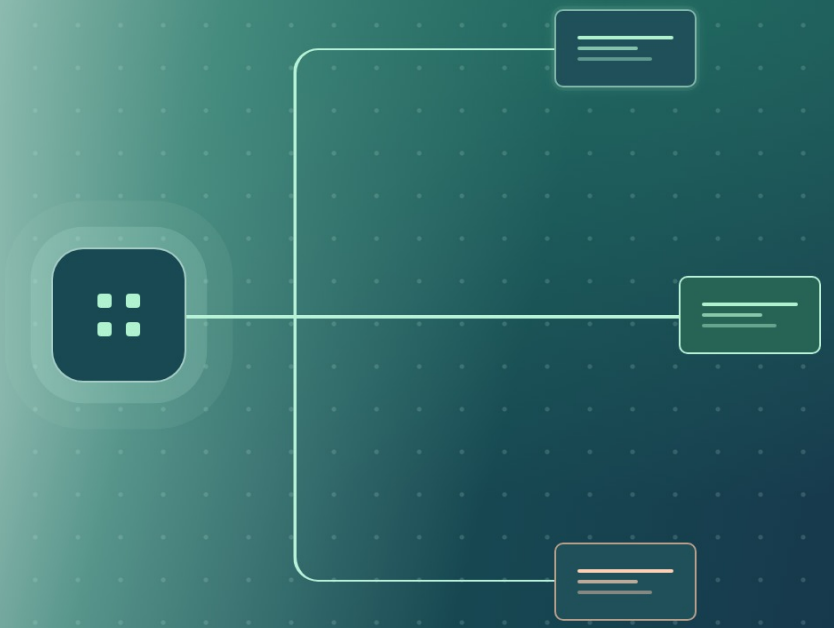
Beyond One-Off Agents:

Building Reusable Enterprise AI Capabilities



Manohar Boorlu

Associate Director – AI Forward Deployed Engineering Lead

Novartis | Ex-Merck

Views expressed are my own and do not necessarily reflect those of Novartis or Merck.

Four years of building blocks

From late 2022 to today, each step made another part of an AI application reusable.

Late 2022

ChatGPT

Chat makes prompting familiar.

REUSABLE

The prompt

2023

Chat model APIs

ChatGPT-class models reach any app.

Function calling

Apps ask the model which tool to use.

REUSABLE

The model call

2024

MCP

One protocol for tools and context.

Structured outputs

Responses that match a schema.

REUSABLE

The connection

2025

Coding agents

Agents that build and change software.

Agent Skills

Packaged know-how for repeated work.

REUSABLE

The know-how

2026

Managed harnesses

The agent loop as a service.

Agent registries

Find and govern tools, skills and agents.

REUSABLE

The execution loop

Next

Still open

Shared context, evaluation and ownership across teams.

OUR WORK

The ecosystem

Selected product and protocol milestones

The trap of one-off agents

Coding assistants make the prototype fast. Every deployment still needs an operating foundation.

FIRST PROTOTYPE · ONE TEAM

Fast, and it works

- **Problem and sample inputs**
  From the business expert
- **Coding assistant**
  Builds the first version
- **Working prototype**
  Something people can try
- **Expert feedback**
  While change is still cheap

ITERATE

Ready for a demo. Not yet ready for production.

Three teams later

AT SCALE

Every team rebuilds the same foundation

	TEAM A	TEAM B	TEAM C
Domain logic	Payer rules	Brief prompts	Forecast logic
Approved sources	Payer policies	Interviews	Sales data
Data connectors	↻	↻	↻
Identity and permissions	↻	↻	↻
Guardrails and tests	↻	↻	↻
Telemetry	↻	↻	↻

Different by design

12 rebuilds
4 components × 3 teams

THE LIMIT

The prototype is the fast part. The operating foundation is the part worth sharing.

Forward deployed engineers find what repeats

Embedded in each unit, we observe the work, agree acceptance and build with the team. What repeats across units becomes shared.

Market access

- Payer medical policies and formulary documents — Document extraction
- Payer advisory board recordings — Audio-to-text
- Formulary coverage data — Data agent
- Payer landscape and policy changes — Deep research agent

Market research

- Syndicated reports and past study decks — Document extraction
- Physician and patient interviews — Audio-to-text
- Patient interview transcripts — PHI/PII detection, Adverse event detection
- Survey results for stakeholders — Visualization agent

Strategy planning

- Competitor publications and analyst reports — Document extraction
- Market and competitor landscape — Deep research agent
- Sales, share and forecast data — Data agent
- Performance views for leadership — Visualization agent

Field intelligence

- Free-text call notes from the field — PHI/PII detection, Adverse event detection
- Field rep voice notes and debriefs — Audio-to-text
- CRM call activity — Data agent
- Field insight dashboards — Visualization agent

WHAT REPEATS

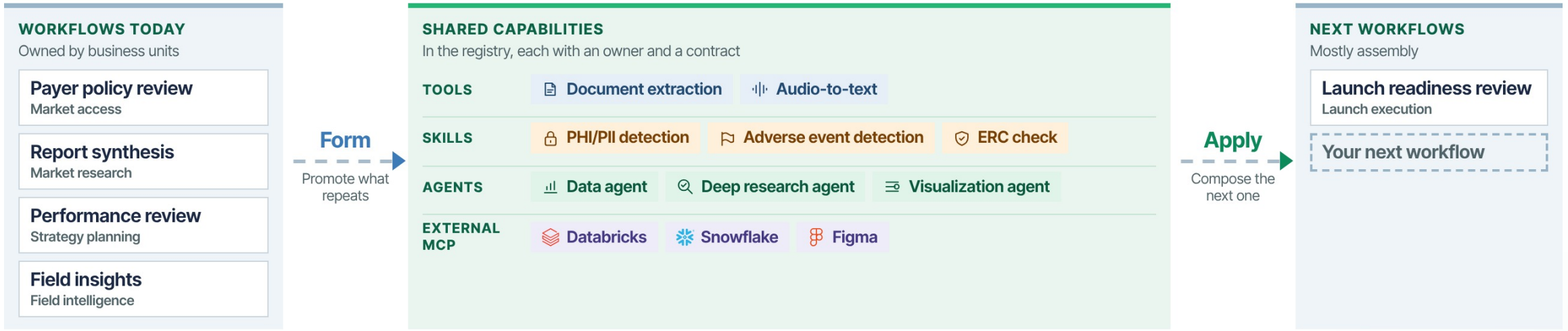
- Document extraction
- Data agent
- Visualization agent
- Audio-to-text
- Deep research agent
- PHI/PII detection
- Adverse event detection

Tool · Skill · Agent

Examples of recurring work in each unit

How capabilities form and get applied

What repeats becomes a shared capability. The next workflow composes what already exists.



Read left to right. Today's workflows feed the shared catalog. The catalog feeds the next workflows.

PLATFORM SERVICES: Identity | Policy | Gateways | Registry | Runtime | Memory | Observability | Evaluations

BUILD ON

MANAGED

- **AWS AgentCore** — Primary foundation in our work
- **Microsoft Foundry** — Foundry Agent Service
- **Google Cloud** — Gemini Enterprise Agent Platform

OPEN SOURCE

- LangGraph
- CrewAI
- Langfuse

Workflows are illustrative. Products are examples, not an endorsement.

What makes a component reusable

Four kinds of capability, and six things a second team needs before it can depend on one.

Tool
A callable operation behind a stable interface

Document extraction · Audio-to-text

Skill
Instructions and resources an agent loads when needed

PHI/PII detection · Adverse event detection · ERC check

Agent
A model with its own tools for a bounded job

Data agent · Deep research agent · Visualization agent

External MCP
Another system, reached through one protocol

Databricks · Snowflake · Figma

document-extraction — Tool · MCP — PUBLISHED

Contract	In: PDF, Word or scanned document. Out: text, tables and source locations.	✓
Partial results	Flags missing pages, tables or footnotes instead of failing silently.	✓
Owner and version	A named team, versioned releases, change notes for consumers.	✓
Registry entry	Discoverable by builders and coding assistants.	✓
Gateway route	Identity and policy on every call.	✓
Evaluation set	Representative documents, rerun on every change.	✓

LIFECYCLE: Propose → Build → Evaluate → Register → Monitor → Version

Example registry entry, illustrative

Context decides how useful the agent is

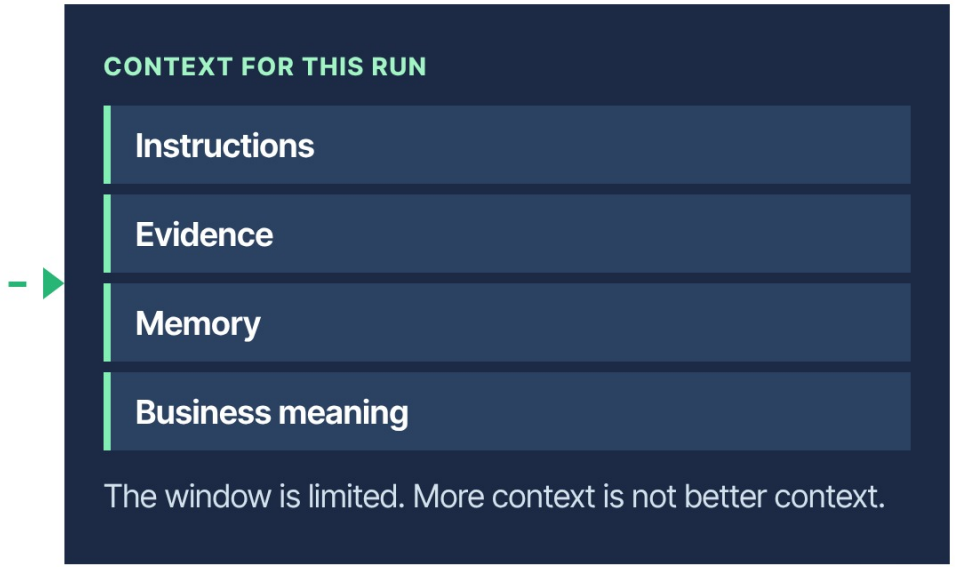
Context engineering is choosing what the model sees on each run. A semantic layer gives that context business meaning.

Instructions and skills
How to do the task

Tool results
Evidence, with source locations

Memory
Scoped to a user, a task and a time

Semantic layer
Business terms mapped to governed data



Select
Give each step only what it needs.

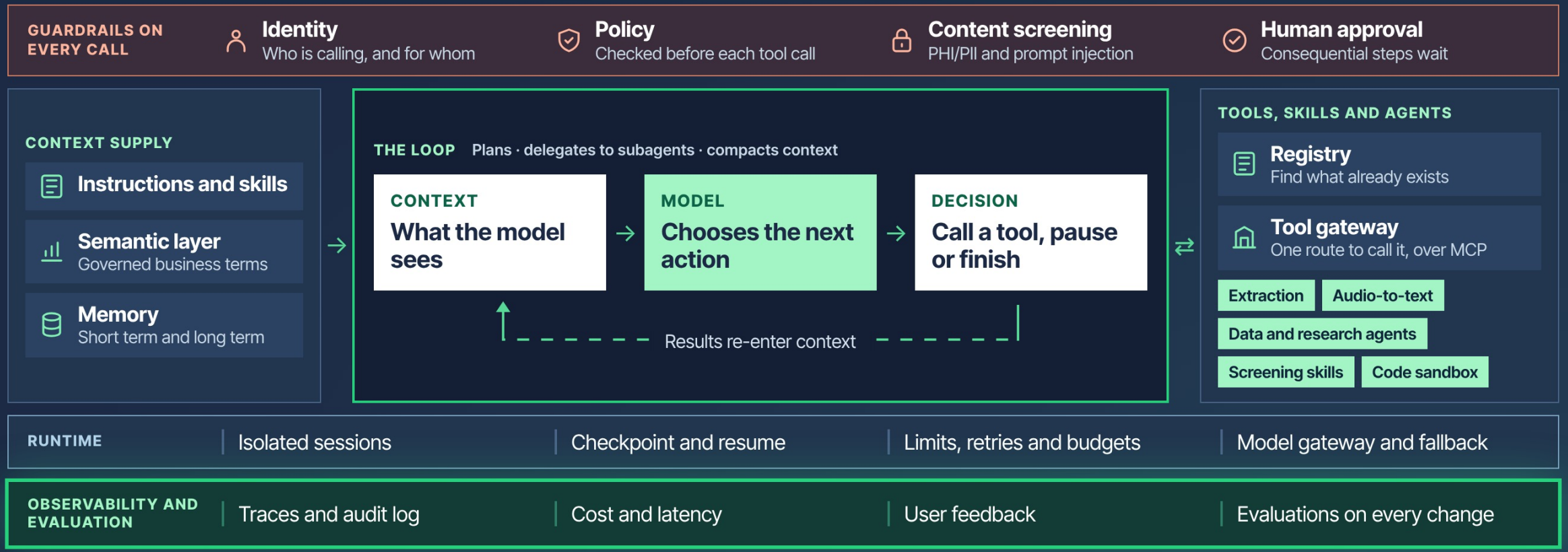
Ground
Keep source and permissions attached.

Define
One meaning per business term, such as “market share”.

High-level view. Structured context reaches agents through external MCP tools such as Databricks and Snowflake.

Inside the agent harness

The loop is small. What surrounds it is what makes an agent enterprise grade.



Roles are vendor neutral. Managed options are in Appendix E.

Governance is designed in, not bolted on

Every question governance asks has an answer in the platform, and every answer leaves evidence.

1 Before it runs
Design and approval

Is this use case approved?
Registry entry with an owner, a purpose and a risk tier

Inventory of every agent, tool and skill

Who can use what?
Identity on every call, permissions checked at the source

Who called what, on whose behalf

2 While it runs
Every call, every output

Is personal data protected?
PHI/PII skill, plus policy at the gateway

Screening result on every run

Does it meet ERC and pharma rules?
ERC skill checks outputs. Possible adverse events go to safety, promotional use to MLR.

ERC flags, safety referrals, review records

Who decides?
Approval checkpoints written into the workflow

Reviewer decisions, with the evidence shown

3 After it runs
Audit and monitoring

Can we show what happened?
Traces and an audit log for every run

The full history of each run

Does it still work after a change?
Evaluations and regression checks on every change

Evaluation history for each version

Illustrative mapping; controls scale with risk tier. ERC: ethics, risk and compliance. MLR: medical, legal and regulatory review.

What to standardize and customize

One test settles most cases: does more than one workflow need it?

YES Standardize across teams		NO Customize by workflow
Tool contracts, registry entries and error conventions	INTEGRATION	Domain APIs, allowed sources and partial-result handling
Identity, policy enforcement and the audit trail	SECURITY	Who may use which source, and where a person approves
Trace format, evaluation runner and release checks	QUALITY	Representative inputs, acceptance rules and prompts
Source locations in the extractor's contract	EXAMPLE	**When a missing coverage footnote holds a finding**

Decision rule Share the behavior several consumers need. Keep interpretation with the workflow that owns the decision.

The full ownership matrix is in Appendix D

COMMERCIAL WORKFLOW

Field insights, built from shared parts

The decision: in lagging territories, should the field team focus on access barriers or HCP awareness?

STEP 1
Frame
Pick the territories, the time window and the two hypotheses.

Scope and approved sources

STEP 2
Gather
Structured: sales, share, CRM activity.
Unstructured: payer policies, debrief recordings.

Document extraction
Audio-to-text
Data agent

STEP 3
Screen
Check free-text call notes for personal data and possible adverse events.

PHI/PII detection
Adverse event detection

STEP 4
Draft
Join both signals into a cited summary with charts, and flag gaps.

Visualization agent
Evaluations
Footnote missing: hold

STEP 5
Decide
An ERC check, then the analyst checks fit and the field lead chooses.

ERC check
Human approval

Shared capability
Owned by this workflow

Test this failure The extractor returns a regional payer policy but drops the footnote that limits coverage to one patient subgroup. The call succeeds. The workflow should hold the access finding and show the reviewer the gap.

Illustrative workflow using capabilities our team has built

IN PRODUCTION

Field insights in production

One question, end to end, across people, the app, the agent platform, context, capabilities and data.

1 Sign in and ask — 2 Start a session — 3 Call capabilities — 4 Screen — 5 Decide — 6 Observe and improve

From question to decision. Each click follows one request through the system below.

PEOPLE	**Insights analyst** Asks, reviews	**Field lead** Decides	**Safety team** Receives referrals		
APPLICATION	**Field insights app** Web or Teams, single sign-on	**Review screen** Sources, gaps, approve			
AGENT PLATFORM AWS AgentCore	**Identity** On behalf of the user	**Runtime session** Field-insights agent on the harness	**Memory** Stores persona and session memory	**Gateway and policy** Every tool call checked	
CONTEXT Loaded per session	**KPI definitions** Uptake, share, reach, via the semantic layer	**Territory details** Accounts, coverage status, targets	**Persona memory** Analyst's role, territories, past questions	**Instructions and skills** How to build a field insights summary	
CAPABILITIES AND DATA	**Document extraction**	**Audio-to-text**	**Data agent**	**Databricks, Snowflake** Sales and CRM, via MCP	**Screening skills** PHI/PII · adverse events · ERC
OBSERVE AND IMPROVE	**Traces** Every step, with cost	**Evaluations** Rerun on every change	**Reviewer feedback** Feeds the next version		

Illustrative flow on AWS AgentCore services. Other managed platforms offer equivalents.

BACK TO THE OPENING STORY

What happened when the second team showed up

A second workflow reused our document extractor. Here is what it inherited, and what it still owned.

WORKFLOW 1
Needs extraction
We build the document extractor

→

PUBLISHED ONCE
An MCP tool
Registered in the AWS Agent Registry

→

BUILD TIME
Workflow 2 finds it
A builder searches the registry

→

RUN TIME
Workflow 2 calls it
Through the gateway, on every run

→

NEXT
More workflows
Same extractor, no new build

INHERITED

What the second team didn't have to build

- A working extraction capability
- A place to find it: the registry
- One route to call it: the gateway
- No second extractor to maintain

STILL OWNED

What only the second team could decide

- Where the call fits in its process
- Which sources it may use
- How to handle partial results
- What passes review

HOW WE WOULD MEASURE IT

1. **Time to acceptance**
2. **Valid results and rework**
3. **Repeat use per cycle**

Observed reuse of one capability. The measures are proposed, not yet reported.

KEY TAKEAWAYS

Be ready when the second team shows up.

1

Find what repeats

Embed with the business. A second consumer is the signal to build a capability.

NEXT PILOT **Name the second consumer**

2

Build it once, properly

Contract, owner, registry, gateway and evaluations, on a harness with governance designed in.

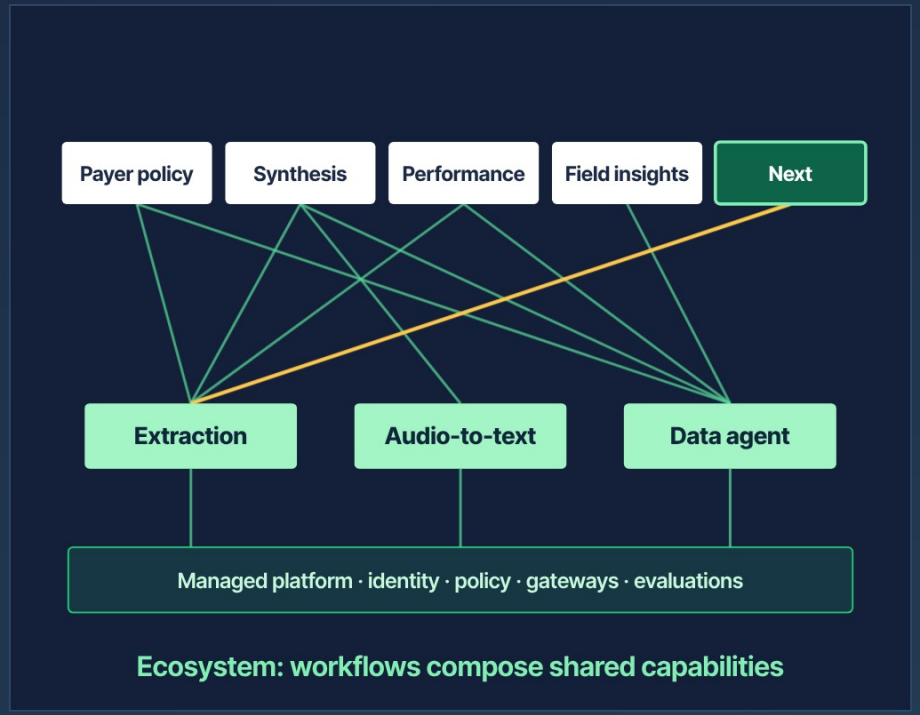
NEXT PILOT **Write the shared contract**

3

Keep the decision with the workflow

Share the operation. Each workflow owns its sources, approvals and failure tests.

NEXT PILOT **Test one failure that should stop it**



Ecosystem: workflows compose shared capabilities

> The first workflow shows that an agent can work. **The second shows whether you built a capability.**

Thank you.

The loop is paused for human input.

Your questions?

Manohar Boorlu
Associate Director – AI Forward Deployed Engineering Lead
Novartis | Ex-Merck

