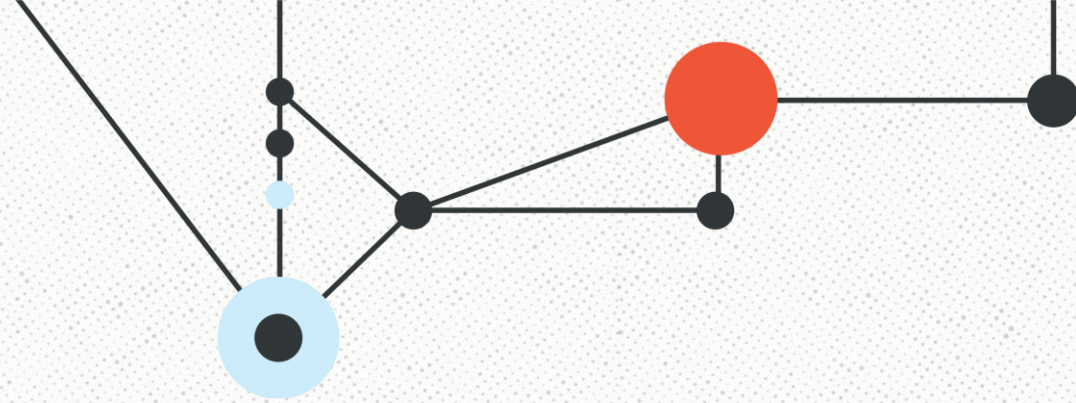
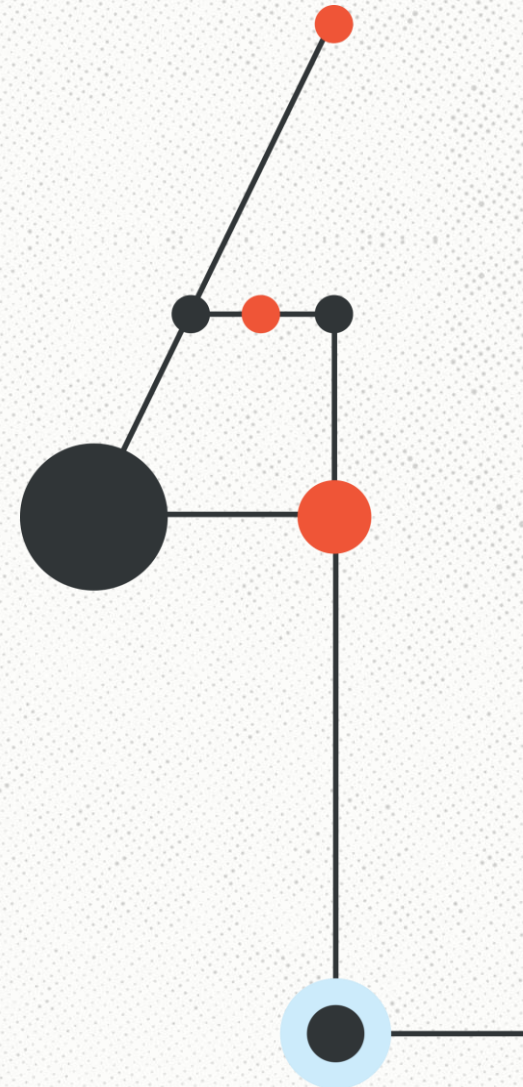


A decorative network diagram in the top left corner consisting of black lines connecting several nodes. One node is a large light blue circle with a black center, and another is a large red circle.

Dark Testing Factory

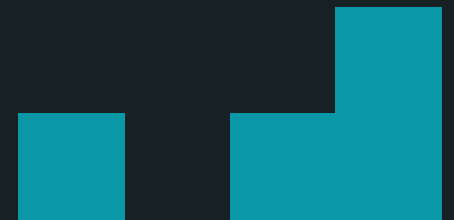
**A self-operating quality system,
led by humans, to ship value fast**

Ingo Philipp & Horea Soanca, UiPath



Turning testing from something you run to something that runs

Always-on testing, strategically led by humans





UiPath Test Cloud

The home of the dark testing factory

UiPath Test Cloud

OPERATIONS



Dark Factory

The operating model for autonomous testing

LAYERS

Governance — control, oversight, feedback · **Execution** — cross-app operations · **Knowledge** — data, history

TRAITS

Human-above-the-loop · Always-on operations · Risk-driven autonomy · Self-improving system · Fixed-cost scaling

LOOP

Detect change ▶ Assess risk ▶ Generate tests ▶ Prepare context ▶ Execute tests ▶ Analyze results ▶ Learn & adapt

CAPABILITIES



Agentic AI

The agentic capabilities powering the factory

PRODUCTS

Autopilot · Delegate · Coding Agents · Agent Builder · Maestro · Test Manager · Studio · Insights · Data Fabric

PILLARS

Leverage out-of-the-box AI agents · Build custom AI agents · Orchestrate AI agents, automations, and humans

FEATURES

Generate tests · Generate automation · Generate test data · Self-healing automation · Autonomous execution

FOUNDATION



Quality Core

The core system for self-operating quality

ACTIVITIES

Functional testing · Performance testing · Regression testing · Progression testing · Exploratory testing

AUTOMATION

Low-code · Pro-code | Web · API · Mobile · SAP · Citrix · Mainframe · Oracle · Epic · Desktop | Windows · MacOS

FEATURES

CI/CD, ALM, Playwright Integration · Testing of AI · Test Data Mgmt · Reporting & Analytics · Parallel Execution

PLATFORM



Governance

Trust Layer · Compliance · Security



Observability

Traceability · Audit · Metrics · Logs



Models

BYOM · Model Hub · Routing



Integration

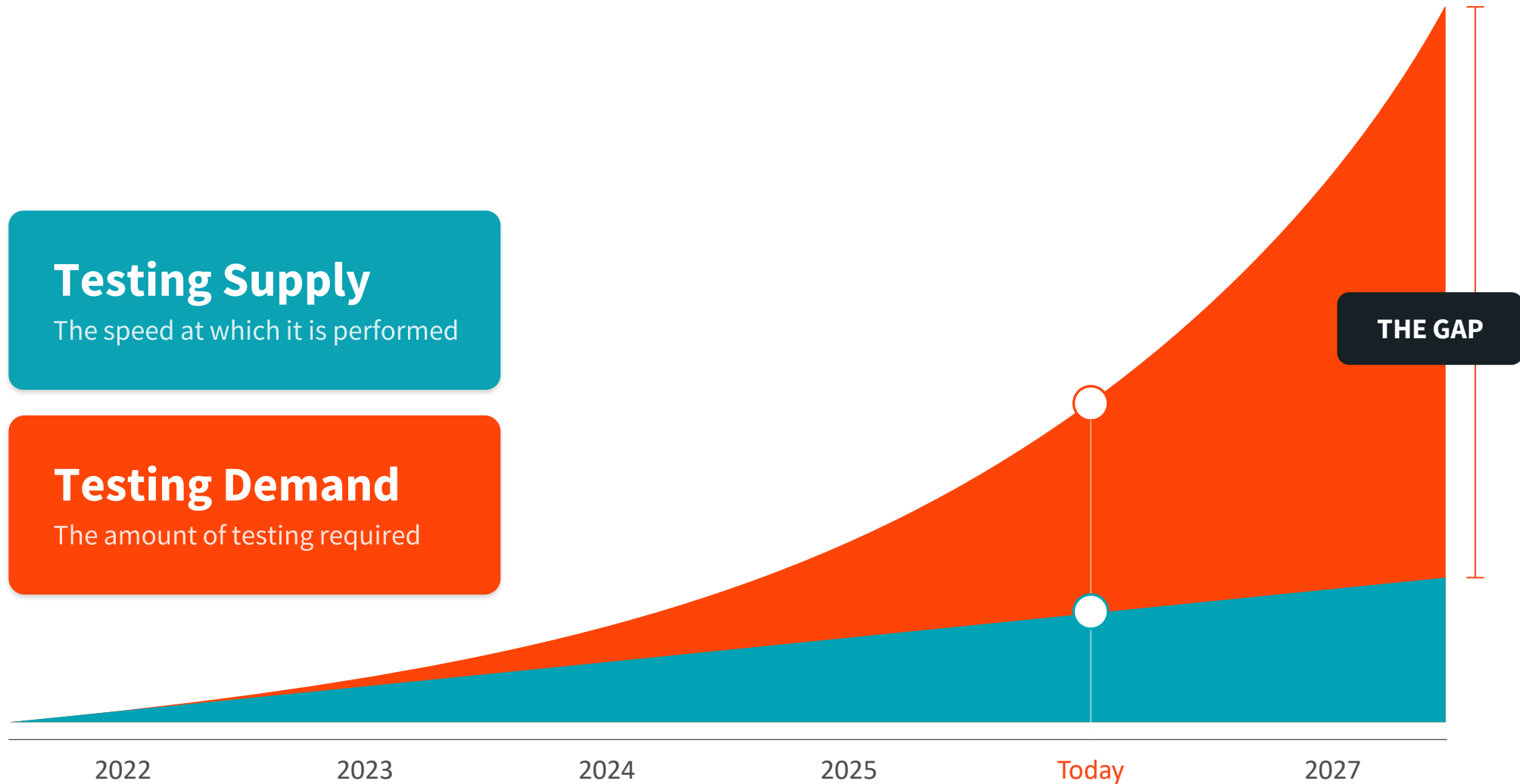
Connectors · SDKs · A2A · MCP · CLI

The time needed for
testing is **always bigger**
than the time you get

Software testing never ends, it just stops



The pressure gap, we feel every day.



More code, more change, more risk.

1

AI enables developers to create code faster → **more code**

2

Devs revise AI-generated code frequently → **more change**

3

More code + more change = **more risk to be tested**

4

QA can't keep up with **increased volume** and **velocity**

We create risk faster than we can test.

60% of AI-generated code accepted into codebases

Task completion +34%; epics per developer +66%

Quality is slipping → # of bugs per developer +54%

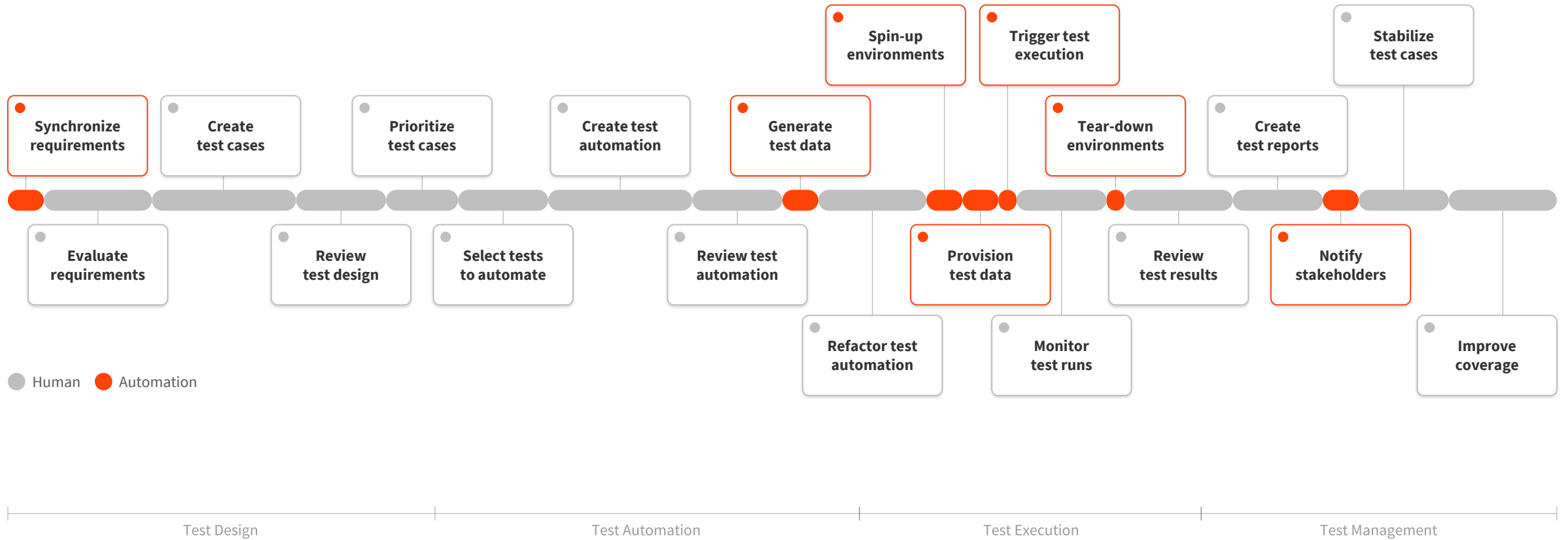
Reliability worsening: incidents-to-PR ratio tripled

Oversight weakening: +31% PRs merge w/o review

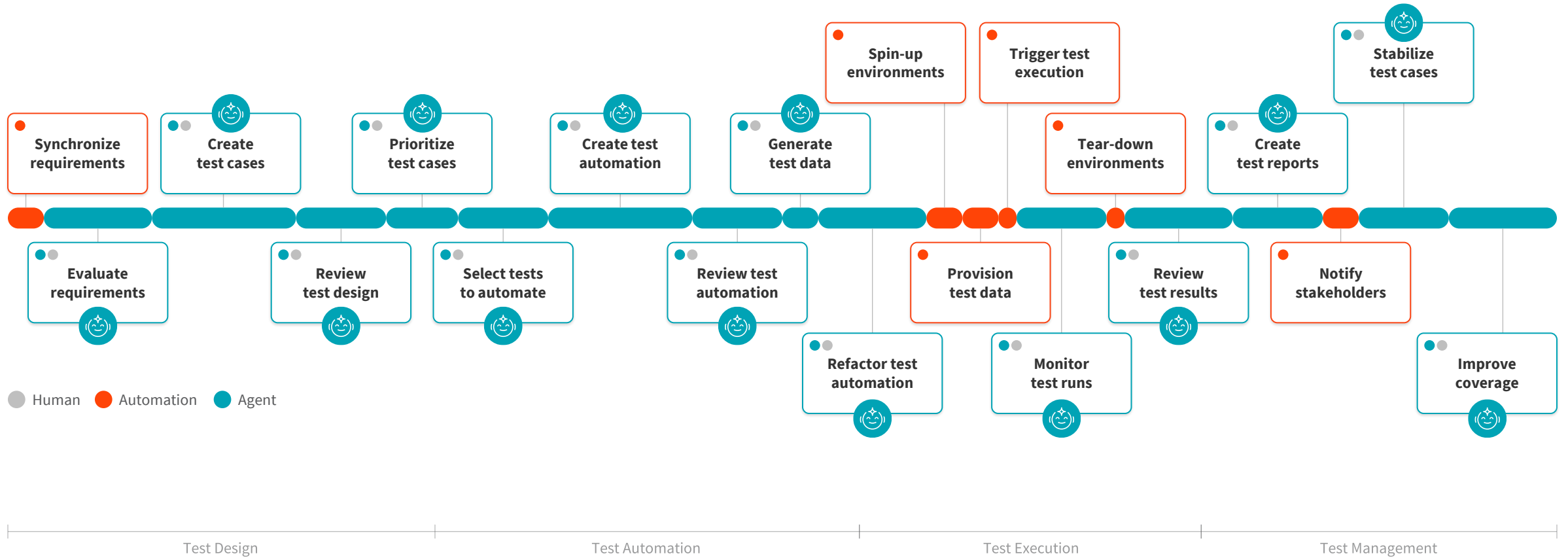
Report reflects analysis as of March 2026



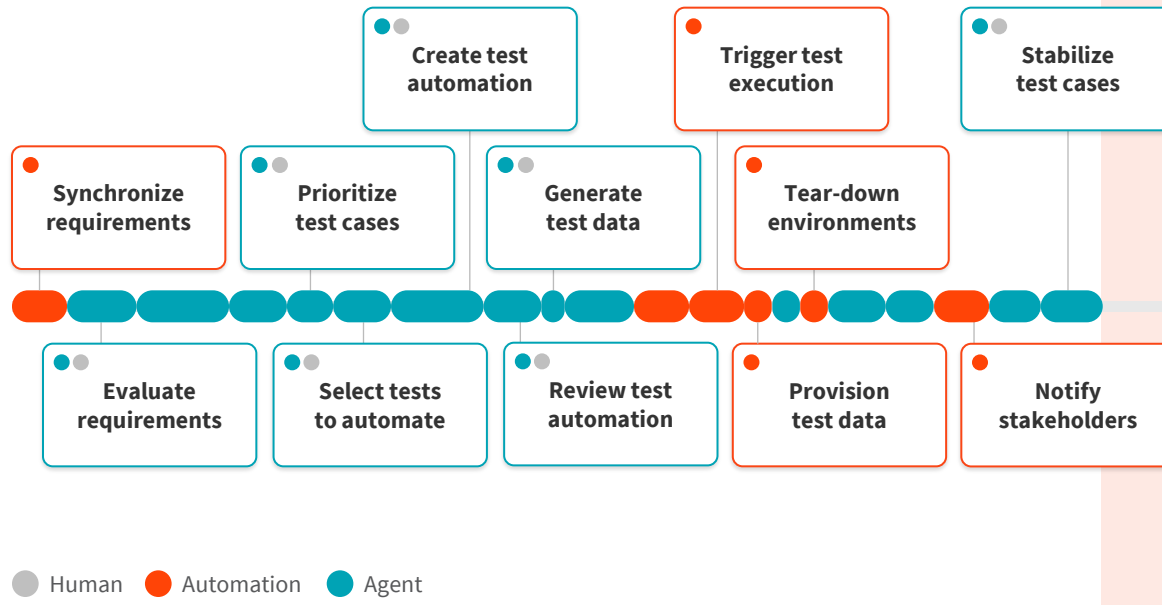
The manual grind, that slows everything down.



Agents step in, and the bottleneck breaks.



Faster cycles, fewer bottlenecks.



Test Design Test Automation Test Execution Test Management

6x

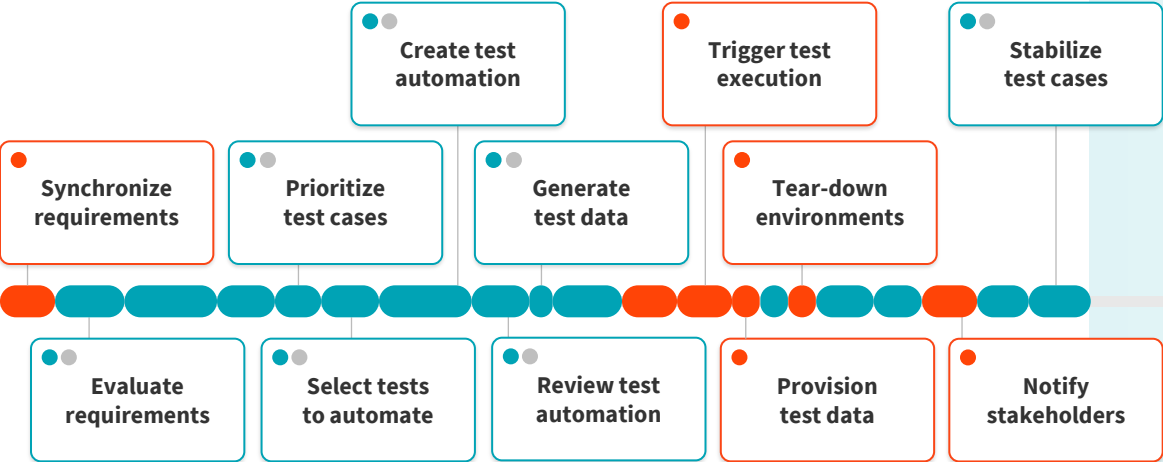
Faster Releases

Ship in days, not weeks.



→ This is only the beginning.

6x faster, 4 million reasons why.



● Human ● Automation ● Agent

Test Design Test Automation Test Execution Test Management

\$4M

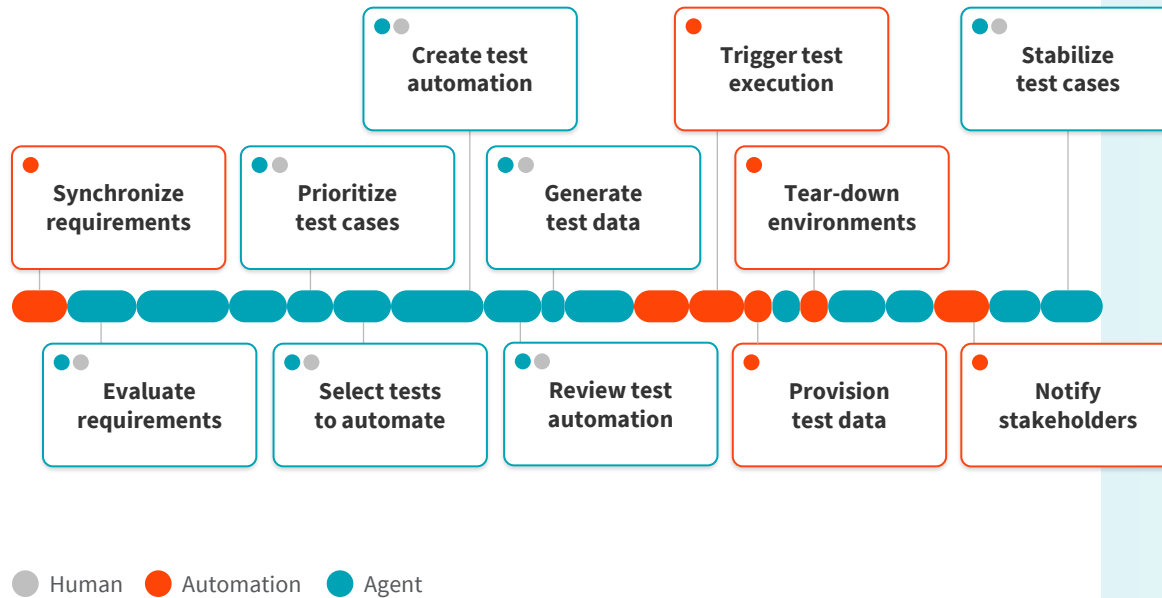
Annual Savings

Real savings, real fast.



→ This is only the beginning.

Three moves to close the gap.



HOW TEST CLOUD COMPRESSES THE CYCLE



1 Leverage Agents

Use a broad collection of out-of-the-box and customizable AI agents across the lifecycle

UiPath Autopilot™

2 Build Agents

Build your own low-code or coded AI agents tailored specifically to your unique needs

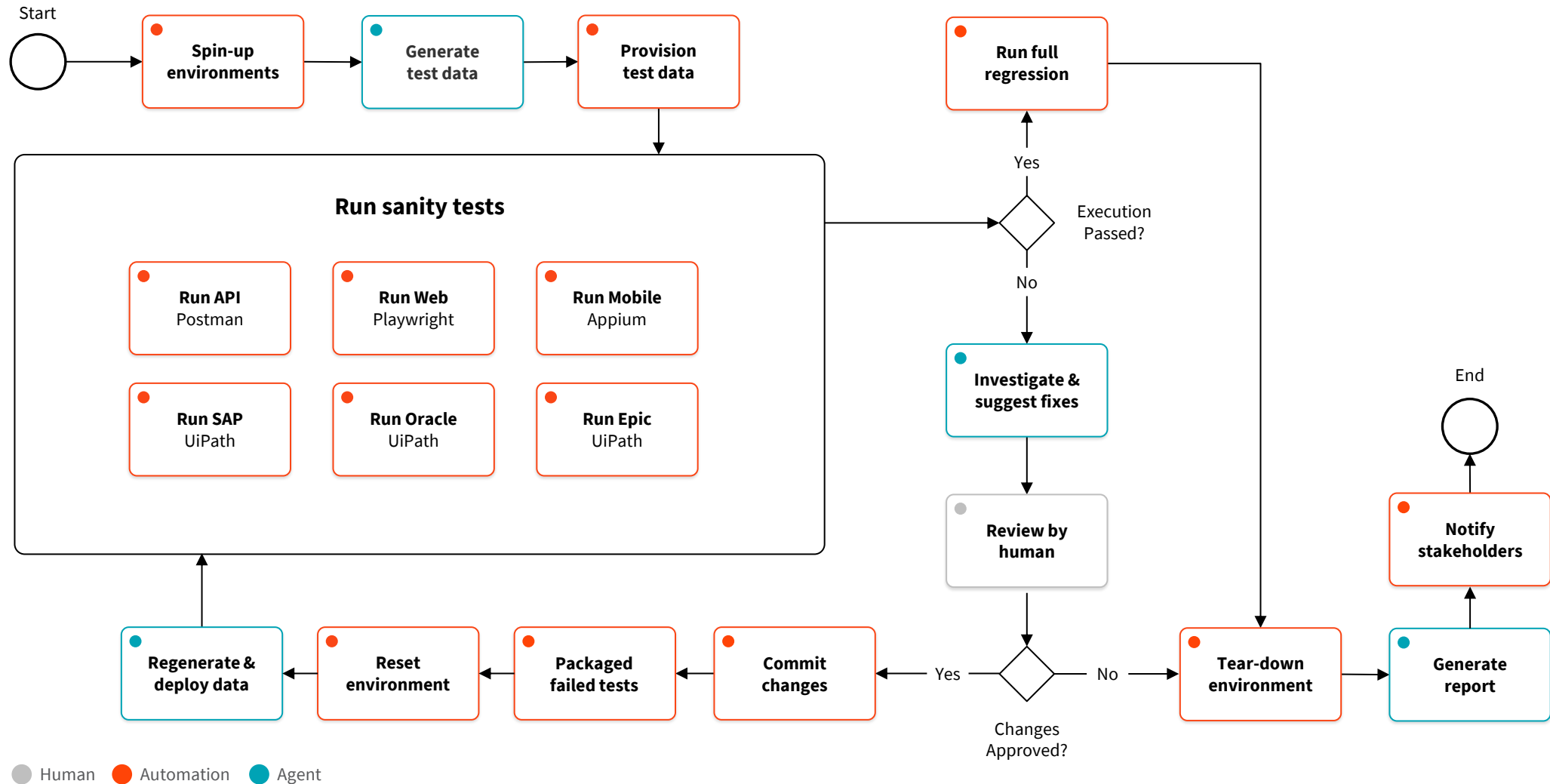
UiPath Agent Builder

3 Orchestrate Agents

Connect agents, automations, and humans to orchestrate agentic testing workflows e2e

UiPath Maestro

Autonomy, securely scaled today.

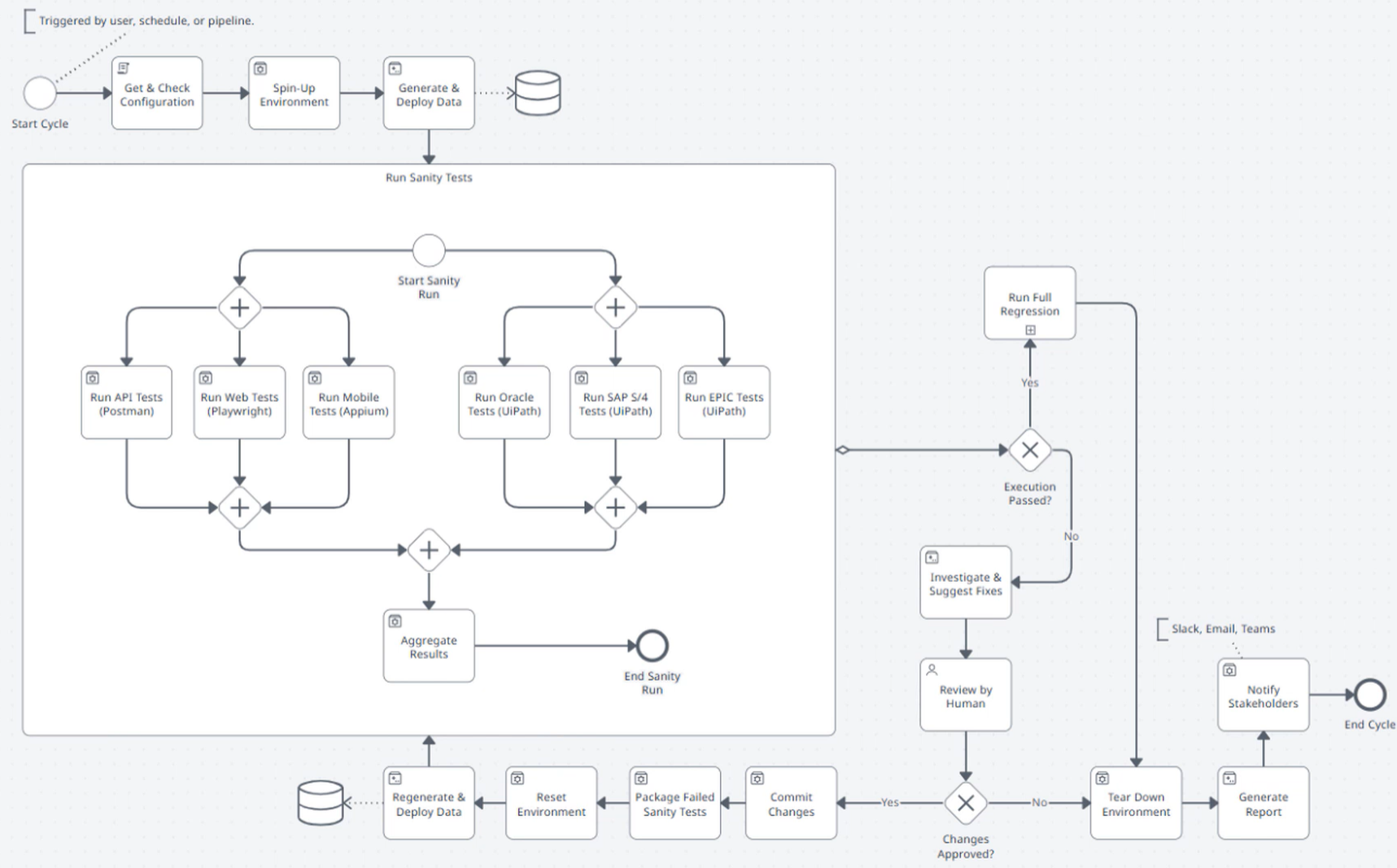


Stop

Continue

Next step

Executio... .bpmn



The lights are **dimming**:
humans are stepping back,
agents are stepping up

As autonomy rises, testing gets darker



Autonomy moves from feature to factory.

AGENTIC TESTING STARTS HERE ↓

1

Manual Testing

You perform every step by hand. No scripts, no machine, you are the test runner, observer, and judge. Every ounce of agency is human.

2

Automated Testing

The classic "if this, then that" of automation. You script every step, the machine follows, blindly. The agency sits entirely with you.

3

AI-Augmented Testing

You prompt AI, it guides you across all phases. You steer the course, decide what happens next but the primary agency remains with you.

4

Autonomous Testing

You state intent for tightly supervised, multi-step agentic flows. Agents take full agency but only inside that "narrow" lane.

5

Dark Testing Factory

Autonomous testing becomes an end-to-end operating system that turns testing from something you run into something that runs.

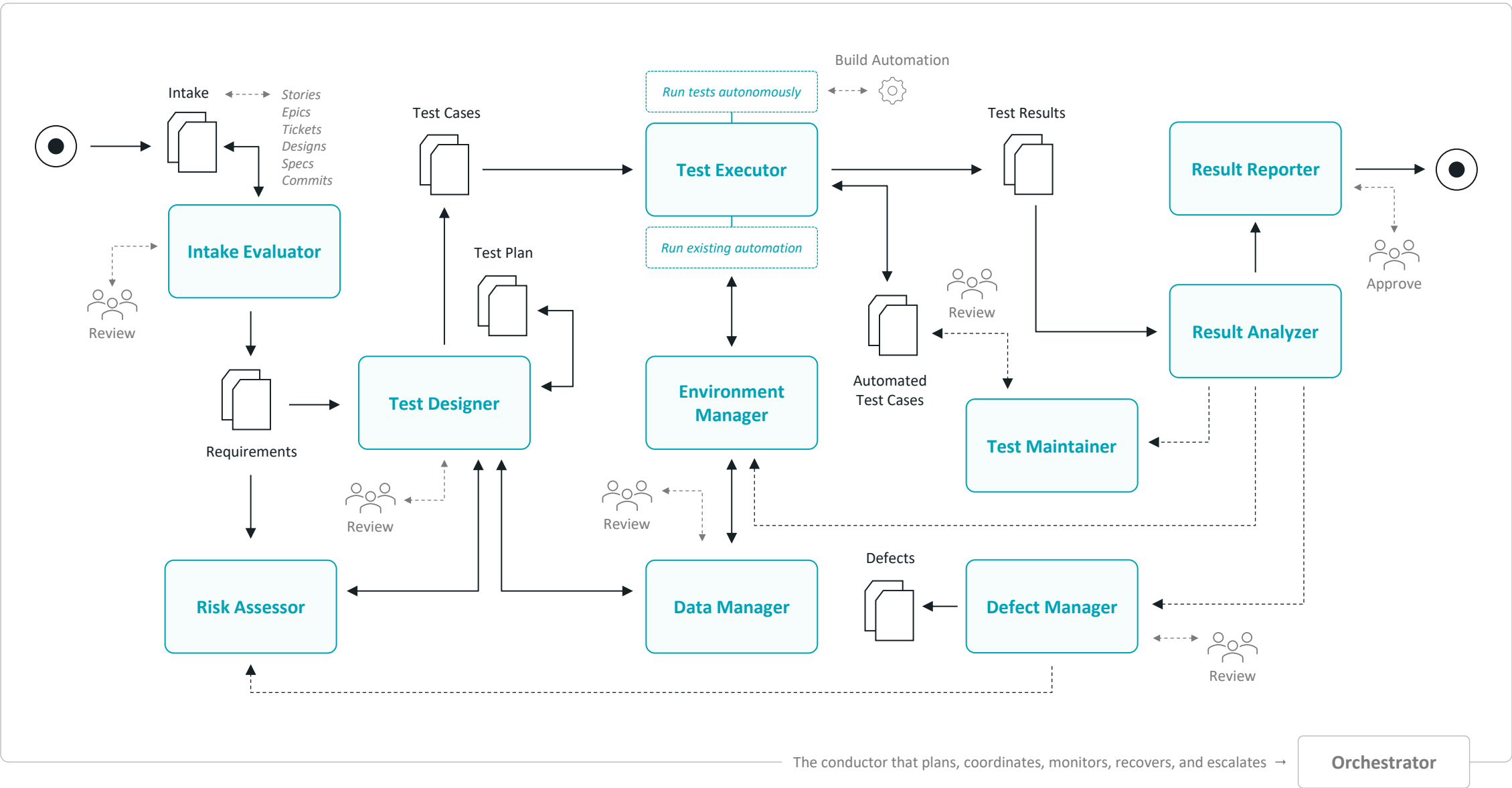
Today's reality, tomorrow's factory.



AGENTIC TESTING STARTS HERE ↓

		Human Role	Reality Check	
1	Manual Testing	Doer	Large legacy baseline well behind the current state of practice	25%
2	Automated Testing	Scripter	Where most companies sit today; broad adoption, uneven coverage	60%
3	AI-Augmented Testing	Director	The active shift, assistants now default to design, automate, maintain tests	12%
4	Autonomous Testing	Delegator	Early rollouts, tightly supervised multi-step agentic flows, not e2e yet	3%
5	Dark Testing Factory	Leader	Aspirational, only a handful of frontier teams chasing full-lifecycle autonomy	<1%

The fabric of the factory, a blueprint.



Inside the risk assessor, a fleet of agents.

Risk Assessor



Risk Extractor — What could go wrong because of this change?

Surfaces credible risks from change signals and context, deliberately blind to existing tests



Risk Consolidator — How does this fit the existing risk model?

Keeps the risk model coherent, current, and traceable (new, changed, split, merged, retired)



Risk Prioritizer — Which risks matter most, and why?

Ranks risks against the quality owner's tolerance, reasoned priority bands, no false precision



Risk Coverage Assessor — How well do current tests examine each risk?

Classifies coverage as strong, partial, missing, or uncertain, and flags redundant tests

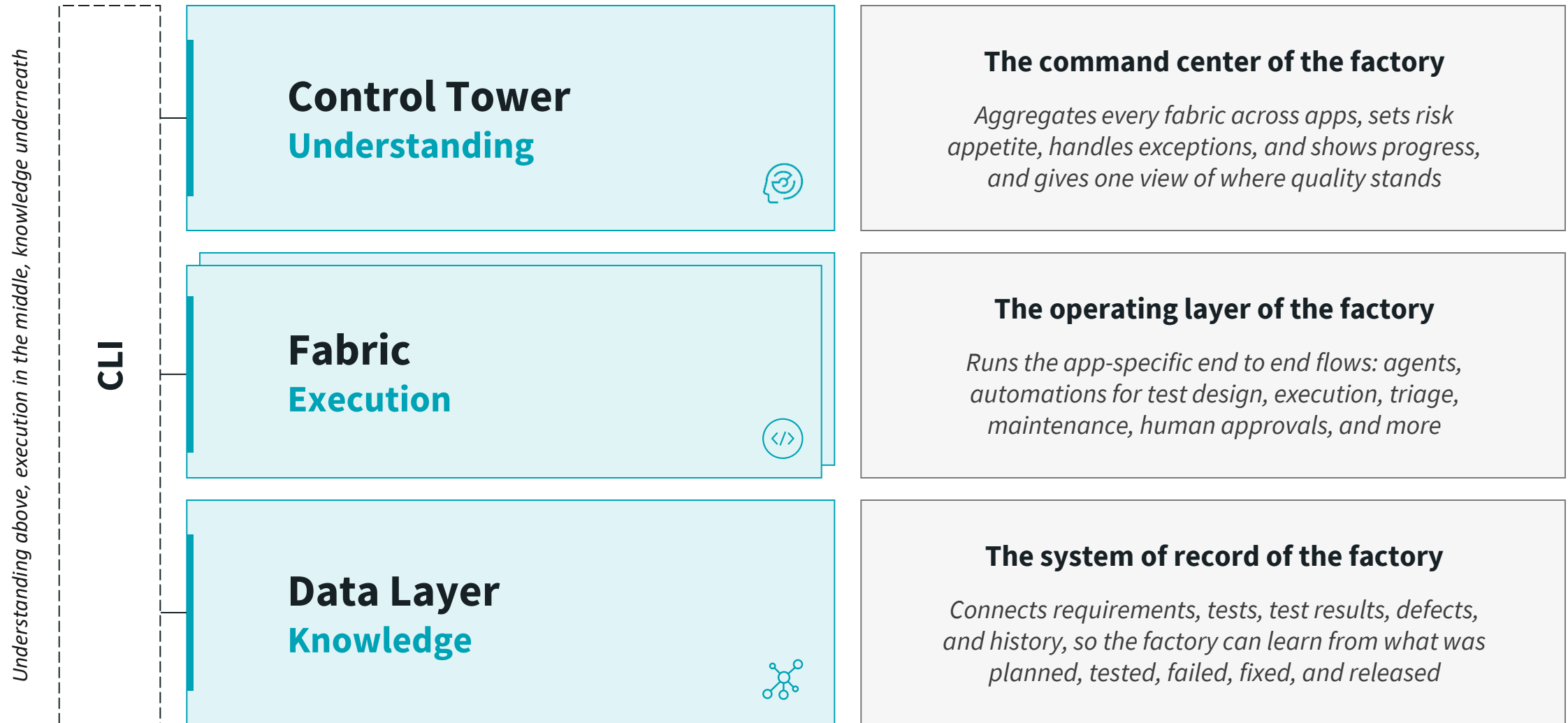


Risk Model Challenger — What might the other agents have missed?

Forms its own view first, then objects, explains, and escalates, a four-eyes principle

The orchestrator coordinates this loop, bounds the reviews, and escalates what can't be resolved.

The high-level architecture of the factory.



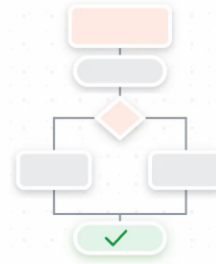
TEST FACTORY

Autonomous testing, **running**
while you ship

Connect a factory to your project. It validates every deployment automatically — scope analysis, execution, triage, quality verdict.

+ Create factory

See how it works



Factories

View all 6 →

Error

Awaiting

Payments Gateway — v3.0

Payments QA

No runs yet

5 exceptions

Running

Awaiting

SAP S/4HANA

SAP CoE

9d ago

Budget

Running

Awaiting

Identity & Access Platform

Platform Engineering

No runs yet

2 exceptions

Running

Awaiting

Billing API — v1.8

Finance QA

No runs yet

1 exception

Capped

Projects

Search projects

Owner

View all 15 →

+ New project

OMG

Order Management

Order lifecycle validation from creation to fulfilment

Updated 448d ago

S4

SAP S/4HANA

Autonomous quality assurance for SAP S/4HANA SD, MM, and FI

Updated 481d ago

ATS

Automated Tests

End-to-end automation suite for core platform flows

Updated 681d ago

P2P

Procure to Pay

Procurement and payment workflow regression coverage

Updated 751d ago

CAD

Compliance & Audit

Regulatory control validation and audit-trail integrity checks

Updated 309d ago

NTF

Notification Service

Email, SMS, and push notification delivery and template tests

Updated 323d ago

ANL

Analytics Platform

Event ingestion, aggregation, and dashboard render correctness

Updated 338d ago

LOG

Logistics & Shipping

Carrier API integration and shipment-tracking regression

Updated 356d ago

Six traits that make a dark factory.



Authorable by coding agents and accelerated by templated agents to start fast

Humans above the loop

Humans set policy, resolve ambiguity, supply missing knowledge, intervene on exceptions, but don't push buttons

Always-on, in-flow ops

The factory runs continuously, and is embedded deeply into the broader flow of software delivery

Closed, non-linear system

It's not a pipeline, but a living loop; a failed test may trigger test repair, data reset, defect filing, human escalation

Risk-driven scheduling

The factory decides what to test next based on change impact, production telemetry, business risk, and policy

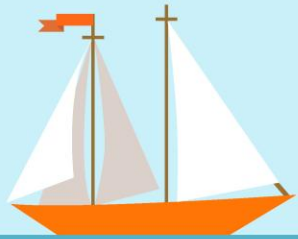
Self-improving system

Human decisions become reusable, knowledge; the factory improves through every human intervention

Fixed-cost scaling

Repeatable effort becomes reusable knowledge, so deeper coverage does not require proportional headcount

The dark factory turns software testing from labor that repeats into software that scales



TERRITORY OF HUMANS

Leadership

Define objectives
Prioritize efforts
Decide trade-offs
Own outcomes

People lead, the
factory scales.

The dark factory absorbs the
time-heavy tactical testing work so
humans can lead strategically

Agents use automations for fast, reliable, low-cost execution

Maintain tests
Execute tests
Design tests
Discover tests

Execution

TERRITORY OF AGENTS

From tactical executors to strategic leaders.



What quality leadership means.

Defining quality goals

Setting risk profiles

Handling exceptions

Designing guardrails

Supplying knowledge

Clarifying ambiguities

Deciding trade-offs

Prioritizing efforts

Owning outcomes

Being accountable

Why the dark factory enables it.

1

Leadership can't be delegated to agentic AI

2

It requires drive, tacit judgment, accountability

3

The factory frees testers from tactical execution

4

Their focus shifts to leadership responsibilities

5

This enables them to shape quality strategically

Testers become quality leaders, leading autonomous agents & shaping quality outcomes instead of being buried in tactical execution

The darker the factory, the more humans matter.



What AI won't pursue

AI can execute intent, not originate it

Human drive

The will to start, find out,
and improve when "good
enough" is not enough

What is still not good enough?

What AI can't see

AI can't act on what it can't access

Tacit judgment

The context behind context:
hidden rules, customer realities,
& risks only experience reveals

What context is missing?

What AI can't own

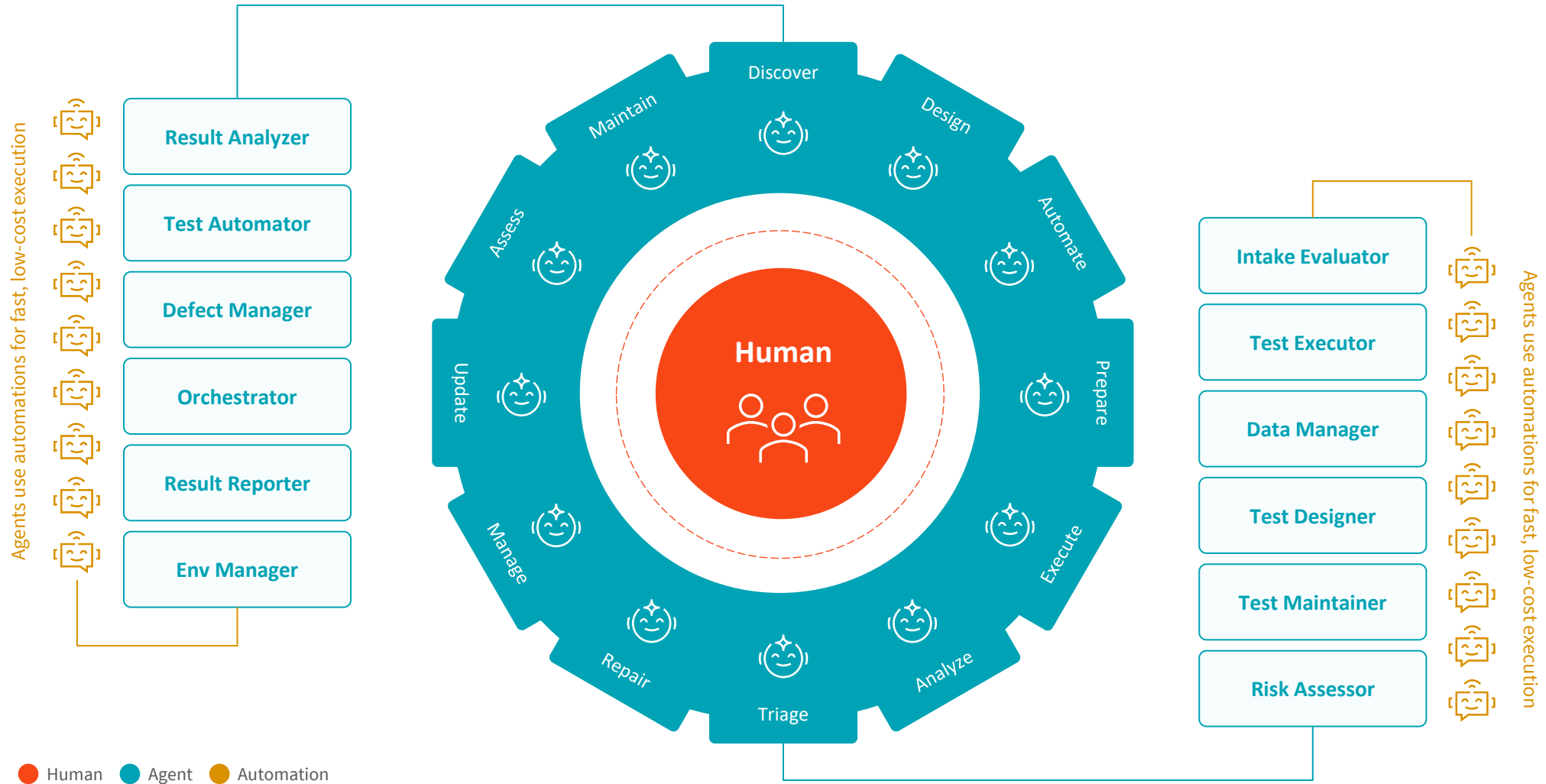
AI cannot be accountable for us

Accountability

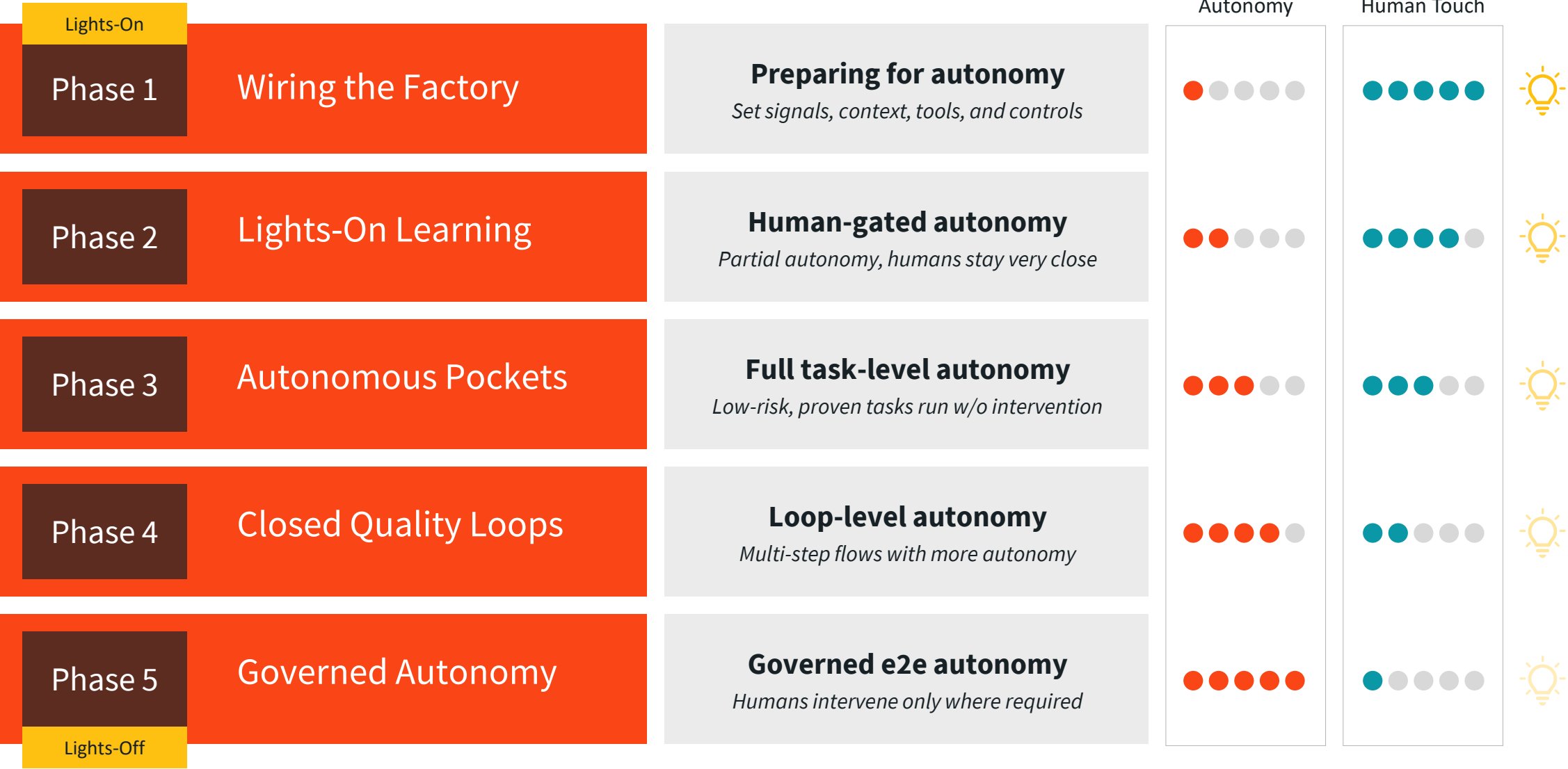
The responsibility to decide
which risks the business can
accept & which it must reject

Who is answerable?

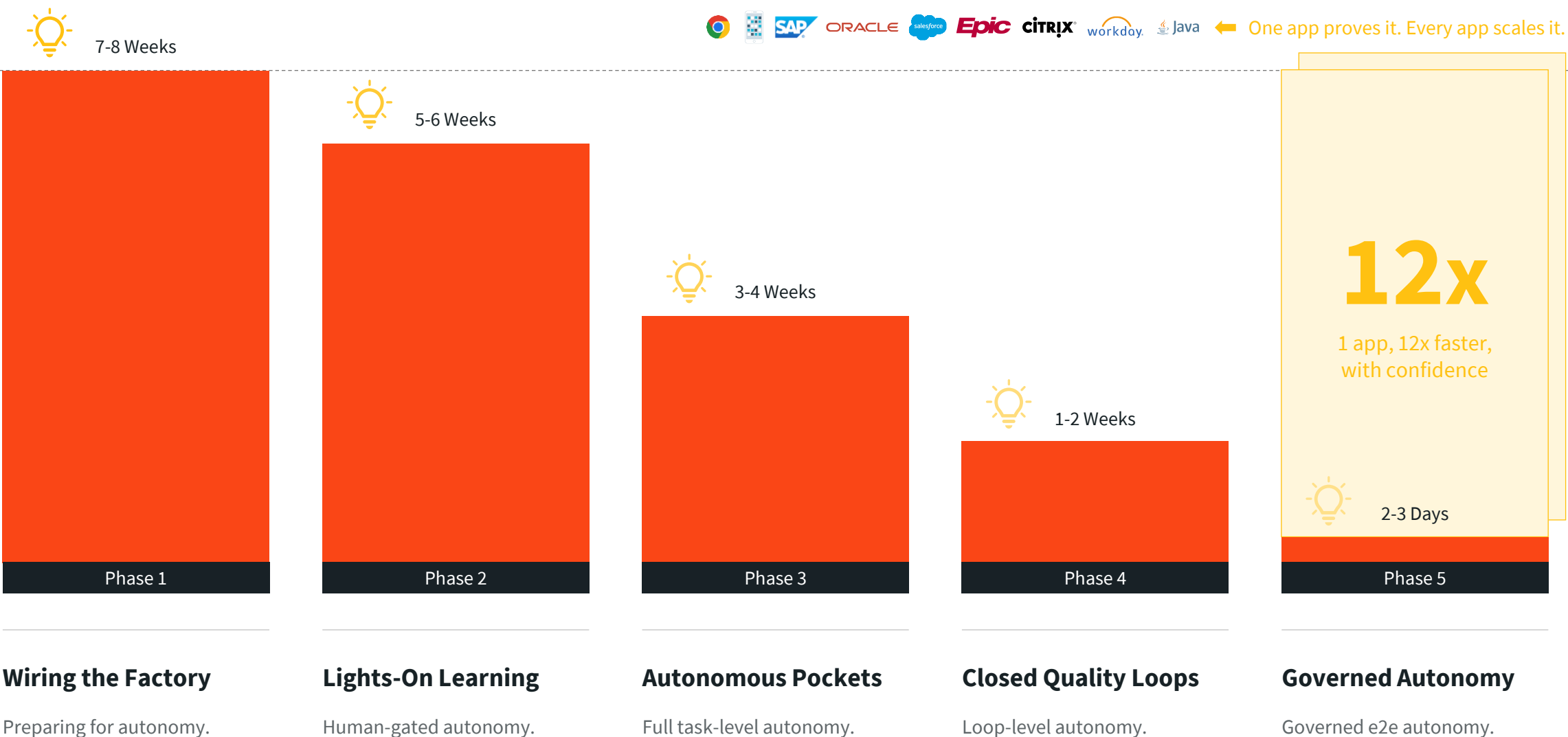
You lead the fleet, the agents do the work.



It's not built in one step, you earn it loop by loop.



Impact scales factory by factory.



The factory is the means. Your confidence is the point.



Giving you confidence



You can answer "Are we ready to ship?" with evidence, not instinct, so every release moves fast with trust

Giving you leverage



Agents handle repeatable work, so you can focus on judgment, risk, ambiguity, and quality leadership

Giving you momentum



Turn tests, defects, decisions into reusable knowledge, so each release makes the next one smarter & faster

Giving you control



Ground agents in your biz context, guide them with policies, guardrails, and judgment where it matters most

Giving you impact



Move from being perceived as the bottleneck to someone helping your business ship faster w/o flying blind

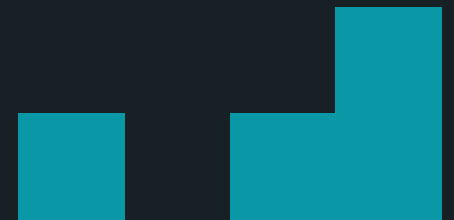
Giving you scale



Expand coverage and speed without expanding manual effort at the same rate to keep pace with rapid change

Turning testing from something you run to something that runs

Scale testing, not the team. Scale what your team can achieve.





UiPath Test Cloud

The home of the dark testing factory