

Bounded Autonomy

Agentic work for development, DevOps, infrastructure and operations

Giving agents freedom to act without abandoning organisational control.

METHODOLOGY · CONTROLS · ORGANISATIONAL LEARNING

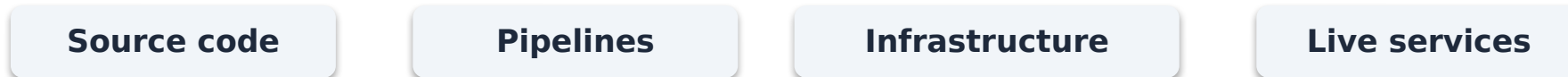


From AI assistance to delegated action

Agents do not only produce answers. They can change the systems for which the organisation remains accountable.

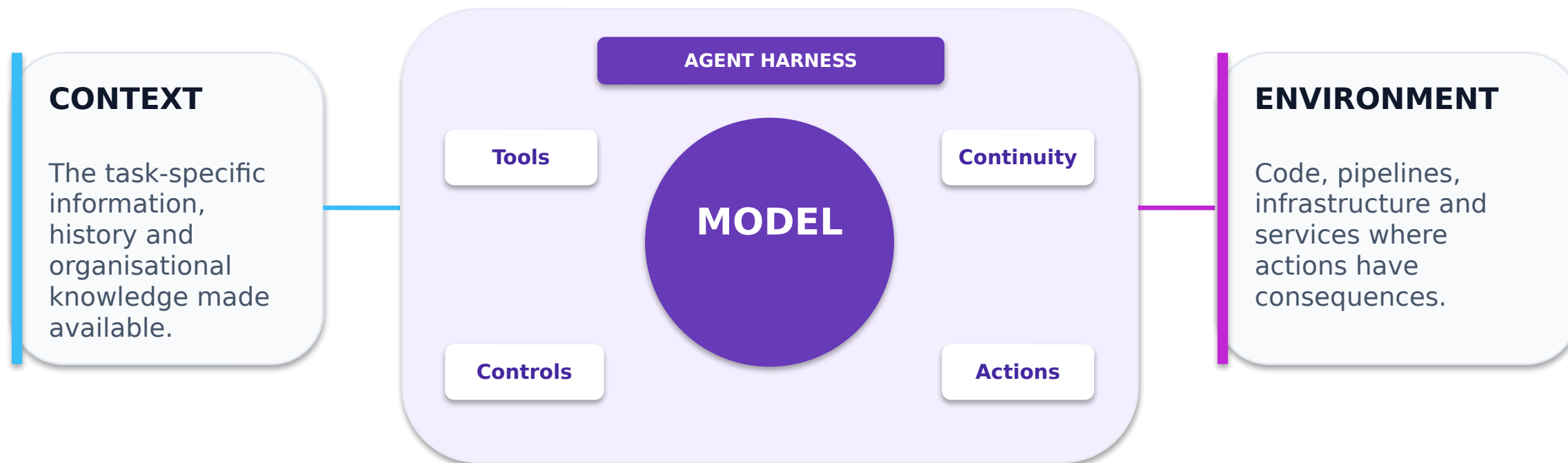


WHERE ACTION BECOMES REAL



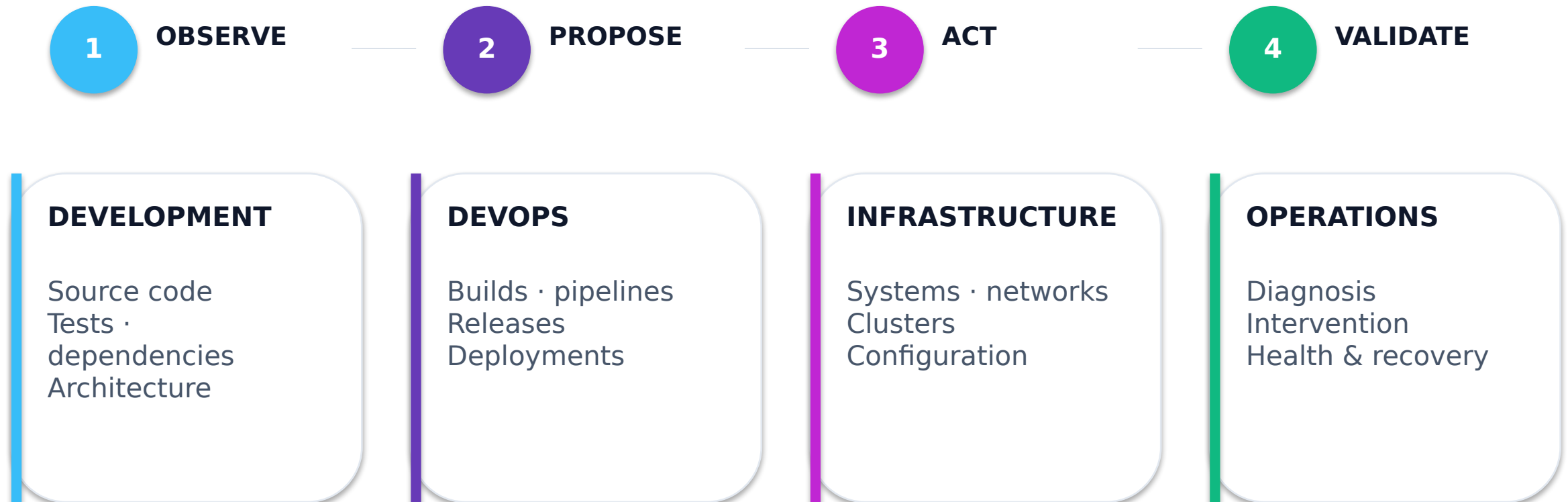
The central question is not whether an agent can act — but what authority it should receive, under which controls, and for how long.

The model is only one component



Models may change. Organisational knowledge, working methods and controls should remain durable.

One agentic cycle, different environments

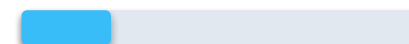


The cycle remains familiar. The authority envelope changes with the environment and its consequences.

Challenge: where should human control be placed?

Human-led assistance

The person remains in the loop continuously.



AUTONOMY

LOW AUTONOMY

Action-by-action approval

Visible control — with growing approval fatigue.



AUTONOMY

FRAGMENTED CONTROL

Request to completion

High speed — with a broad and long-lived mandate.



AUTONOMY

LONG HORIZON

Bounded autonomy

Human decisions between meaningful increments.



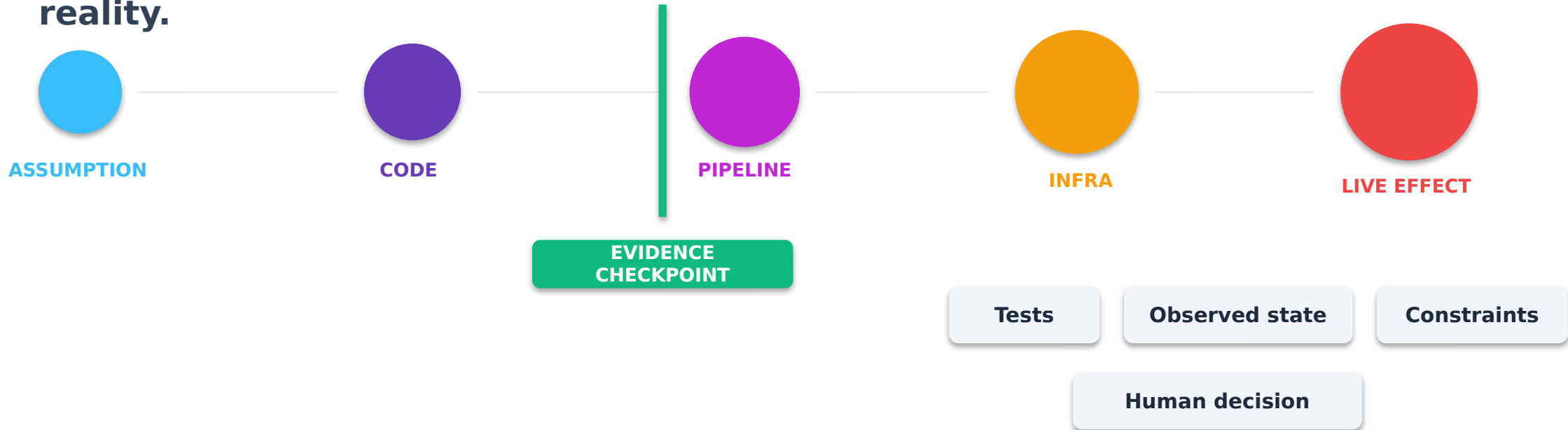
AUTONOMY

CONTROL AT DECISIONS

The objective is not to remove human control — it is to move it to the moments where judgment matters.

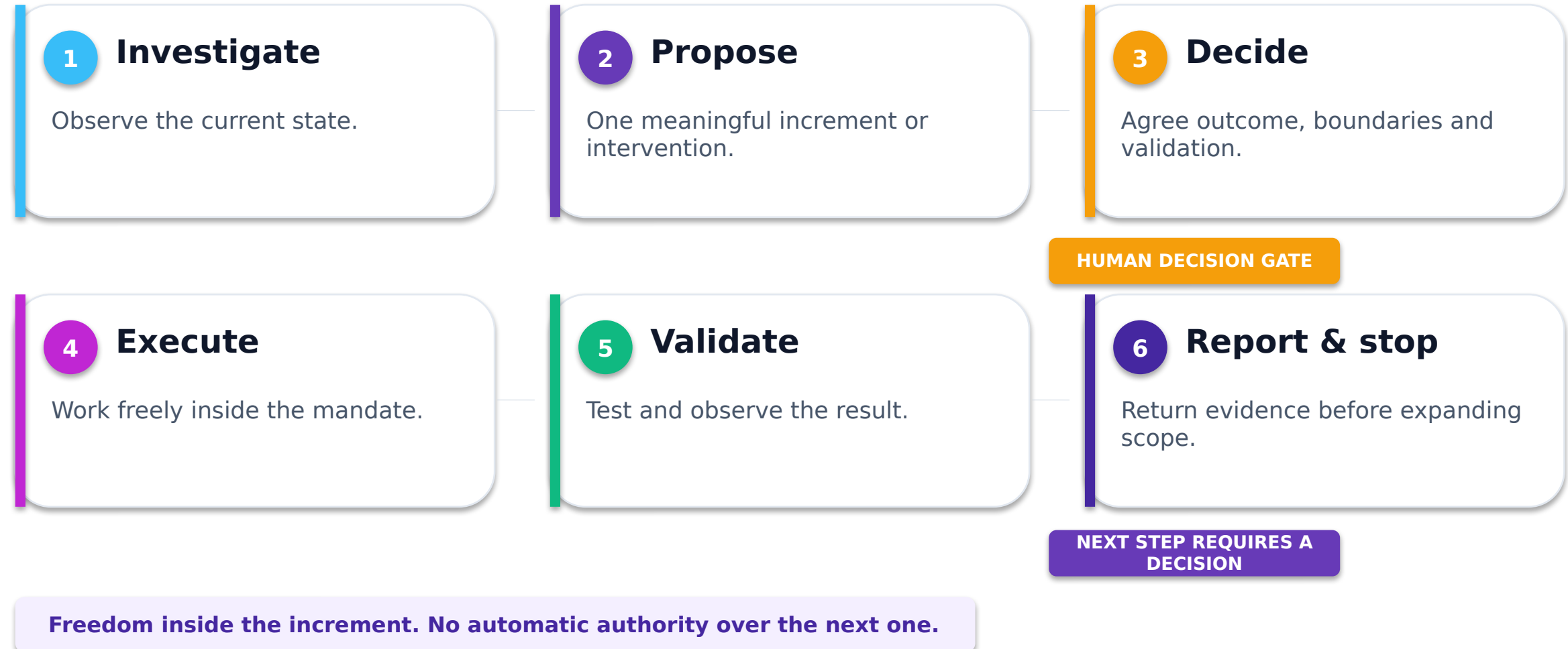
Challenge: how far should we let AI predict?

An early assumption can travel through several layers before anyone reconnects the work with reality.



A convincing result is not the same as a verified result.

Response: bounded incremental autonomy



Challenge: people must learn a new way of working

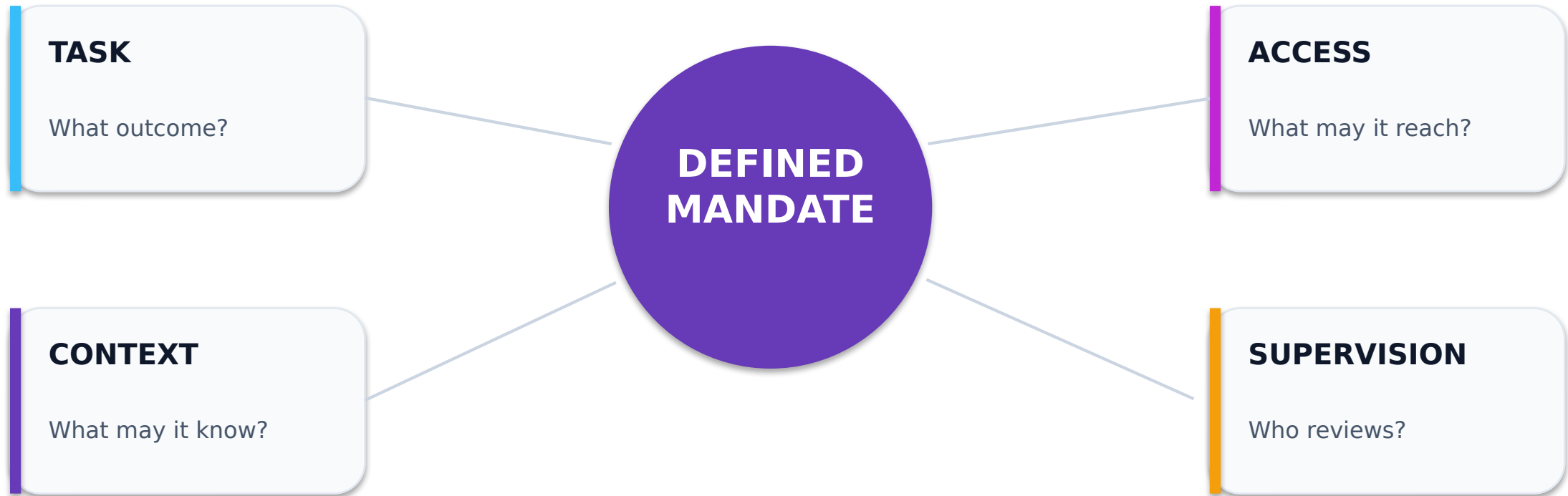
AI access alone does not create organisational capability.

Teams learn to delegate, constrain, review and challenge agents through bounded real work.



Each increment is a unit of work, authority, risk — and organisational learning.

Response: treat agents as delegated organisational actors



The organisation delegates authority. The organisation remains accountable.

Response: combine several layers of control

DIRECTION

Objectives · procedures · runbooks · organisational knowledge

SUPERVISION

Human checkpoints · independent review

BOUNDARIES

Sandboxing · controlled access · restricted communication

EVIDENCE & RESILIENCE

Tests · health checks · records · rollback · recovery

Important boundaries must remain effective even when the model misunderstands its instructions.

Pi-Run as a practical exploration

Profiles & skills

Purpose-specific roles, instructions and organisational knowledge.

Bounded workflow

Investigate, agree one increment, execute, validate and stop.

Guardian

Independent review without constant human permission prompts.

Sandbox & sidecar

Filesystem and network boundaries enforced outside the model.

Call a Friend

Controlled consultation with another model and limited context.

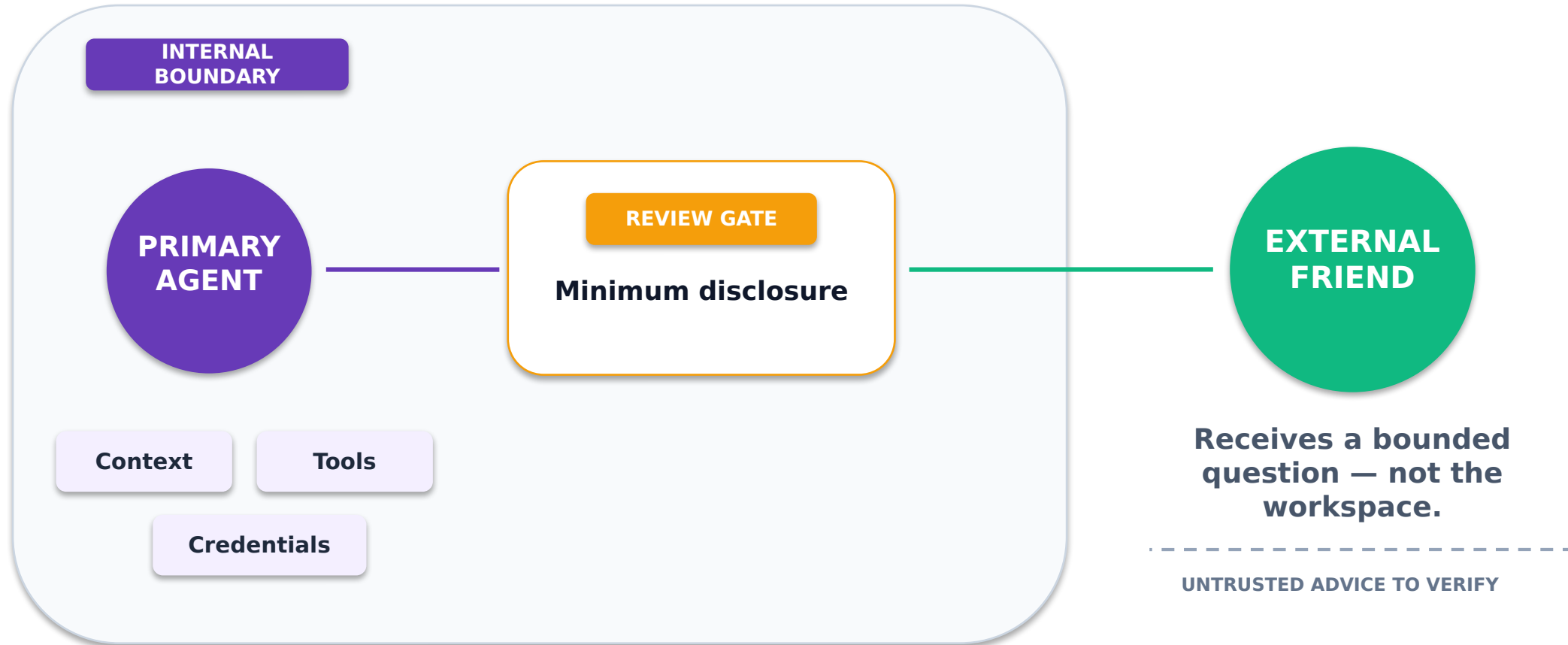
Sessions & recovery

Continuity, evidence, Git history and encrypted snapshots.

A case study in useful autonomy — not a claim that one tool solves every risk.

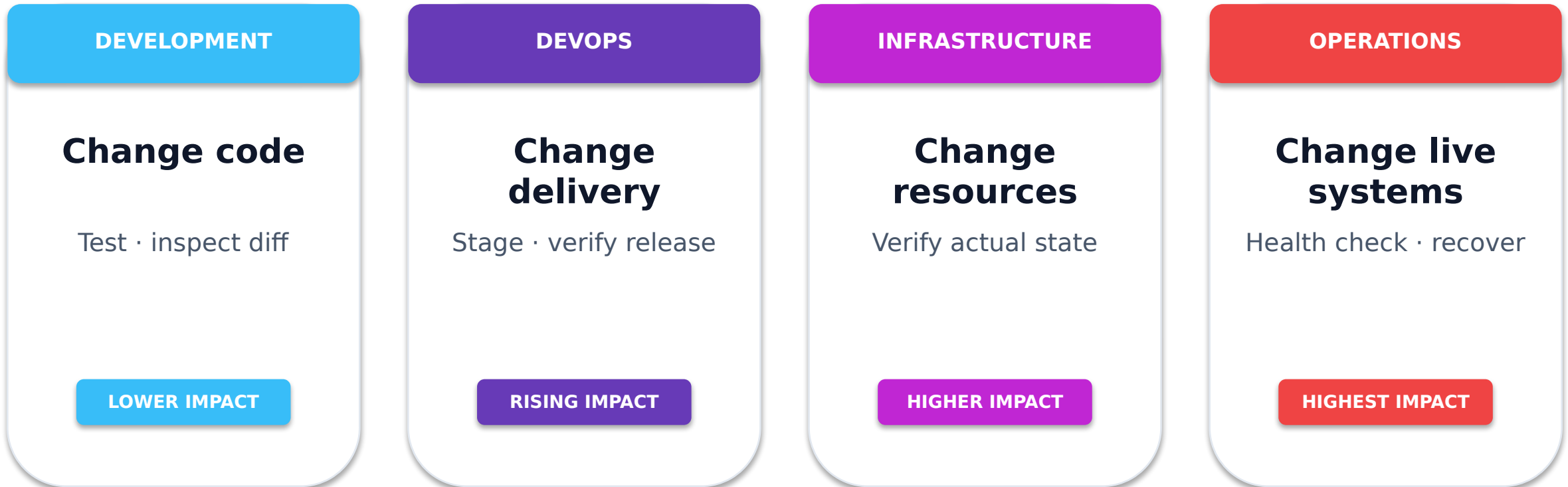
github.com/aprilsoftware/pi-run

Controlled use of external models



Share the minimum context required for the specific purpose — and keep authority internal.

Adapt the method to the consequences



Greater blast radius → smaller steps, stronger boundaries and clearer recovery.

Benefits and trade-offs

BENEFITS

- Useful autonomy with clearer accountability
- Less approval fatigue and shorter feedback loops
- One methodology across development and operations
- Progressive adoption aligned with organisational risk

TRADE-OFFS

- Disciplined task definition and validation are required
- Human checkpoints can become a bottleneck
- Controls may reduce maximum execution speed
- Risk is reduced and managed — never eliminated

The method is designed for organisations that value speed, responsibility and continuity together.

Further reading

Six articles exploring the methodology, technology and organisational controls behind bounded autonomy.

- 1** [Beyond Permission Prompts: Building a Workstation for AI Agents](#)
PI-RUN · BOUNDED WORKSTATIONS [READ ↗](#)
- 2** [AI Coding in YOLO Mode](#)
DEVELOPMENT METHODOLOGY [READ ↗](#)
- 3** [How Far Should We Let AI Predict?](#)
PREDICTION HORIZON · EVIDENCE [READ ↗](#)
- 4** [Beyond YOLO Mode: Building Boundaries for Enterprise AI Agents](#)
ORGANISATIONAL CONTROLS [READ ↗](#)
- 5** [Sovereign AI: The Model Can Change. Your Knowledge Shouldn't Have To.](#)
KNOWLEDGE · MODELS · SOVEREIGNTY [READ ↗](#)
- 6** [Inside an AI Agent Harness](#)
MODELS · CONTEXT · TOOLS · AGENCY [READ ↗](#)

**PROGRESSIVE, GOVERNABLE
AUTONOMY**

Freedom inside the task. Constraints around the task.

DECISIONS Human attention on
consequential choices

BOUNDARIES Controls independent
of model behaviour

EVIDENCE Reality checks across
code and live systems

LEARNING Trust expanded
through experience

**Neither constant permission prompts nor an unlimited
mandate.**