

How Agents Make a Country Grow

The binding constraint on growth was always that capable human attention is scarce and expensive. **AI agents make competence cheap to manufacture** — and turn 24 stalled growth levers into buildable programmes.

+34%

Productivity gain for the *least-experienced* workers given an AI agent — capability, diffused to where it was scarcest.

Source: NBER (Brynjólfsson, Li & Raymond)

ABOUT THIS REPORT

Capability was the bottleneck. Agents break it.

For two centuries, "development" meant slowly accumulating human capability and hoping institutions held long enough to use it. Agents invert the sequence: capability becomes abundant first, and judgment, legitimacy and discipline become the scarce, decisive inputs.

This is the companion to ENSI's *What Actually Makes a Country Grow*, which showed growth comes from a buildable **state-capacity + industrial-learning stack with a hard discipline test** — not the policy menu. This report re-asks every lever in that stack through one question: **what could a fleet of disciplined AI agents do here that scarce humans never could?**

It is grounded in two libraries: **142 documents** on how 24 countries grew, and **53 reports** on how AI makes states and economies more effective (OECD, IMF, World Bank, WEF, NIST, the EU AI Act, Stanford HAI, NBER, ILO, UNESCO, WHO, and national AI strategies).

How to read this report

Depth 1 — 90 seconds. Cover, this page, the executive summary, the Agentic Dashboard, and the Stack diagram. You own the thesis and the headline numbers.

Depth 2 — 15 minutes. Add the reframe, the five layers, the 24-at-a-glance matrix, and the build roadmap.

Depth 3 — the full read. The 24 agentic opportunity areas, the readiness landscape, the risks, and the evidence base.

THE TWO THREADS IN EVERY AREA

THE DISCIPLINE TEST

Every agent fleet is tied to an objective external metric it cannot fake — the agentic descendant of the export test that separated Korea from Brazil.

THE GUARDRAILS

Human-in-the-loop on consequential decisions, audit logs, evaluation and provenance — without them the programme is politically dead.

142

country-growth documents, across 24 countries

53

agentic-government reports, across 10 themes

24 · 5 · 7

opportunity areas · layers · lenses per area

CONTENTS

What's inside

A · THE 90-SECOND READ

01	Cover	1
02	About / how to read	2
03	Contents	3
04	Executive summary	4
05	The Agentic Dashboard	6
06	The Agentic Growth Stack — the thesis in one diagram	8
07	Twelve findings	9

B · THE FRAMEWORK & THE LANDSCAPE

08	The reframe: capability was the bottleneck	10
09	The five layers explained	12
10	The 24 opportunity areas at a glance	13
11	The agent archetypes & the seven-lens method	14
12	The build sequence	15

C · THE 24 AGENTIC OPPORTUNITY AREAS

13	Layer I — Agentic State Capacity (areas 1–4)	16
14	Layer II — Agentic Administration (5–10)	21
15	Layer III — Transformation Machinery (11–16)	28
16	Layer IV — Quality & Resilience (17–21)	35
17	Layer V — Enablers & Guardrails (22–24)	41

D · EVIDENCE, ROADMAP & RISKS

18	The agent archetypes catalog	45
19	The discipline & guardrails model	46
20	Country readiness landscape	47
21	The build roadmap (recommendations)	48
22	Risks & failure modes	50

E · BACK MATTER

23	Methodology	51
24	The evidence base: 195-source library	52
25	Back cover	53

EXECUTIVE SUMMARY · 1 OF 2

The price of state capacity is collapsing

Growth is built, not bought — and the thing that builds it is a capable state that can pick productive bets, finance them, and discipline them against a test it cannot fake. The reason most states under-deliver is not the wrong policy. It is that competent execution has always been scarce and expensive, because it came bundled inside trained, honest, attentive human beings.

That is the constraint agents break. Not chatbots or dashboards — **agents**: autonomous digital workers that take a goal, plan, call tools, act on live systems, and can be held to a measured result. An agent is the first technology that manufactures competent execution at near-zero marginal cost and near-infinite parallelism. A country that could never hire ten thousand world-class analysts, auditors, caseworkers and tutors can field ten million agentic ones.

This makes the agentic moment a **growth** story, not merely an efficiency one. Every lever in the stack was throttled by the human-attention constraint, so relaxing it lifts every lever at once — the pilot-agency that can finally model every bet, the tax authority that can check every return, the SME that can reach frontier capability, the clinic that can see every patient.

~40%

Of global jobs are exposed to AI — ~60% in advanced economies (IMF)

+34%

Productivity gain for novice workers given an AI assistant (NBER)

~41%

Of public-sector tasks could be supported by generative AI (Alan Turing Institute)

The catch is the catch that always applied. Capability without discipline is not growth; it is rent-seeking at scale, executed faster. An agent fleet pointed at the wrong target, or shielded from a hard test, will scale mistakes with terrifying confidence. So discipline and guardrails matter *more* in the agentic era, not less.

THE ONE-LINE VERSION

Growth was never throttled by ideas or capital. It was throttled by the supply of capable hours. Agents end that scarcity — so the game becomes pointing the new abundance at the right targets, and holding it to account.

EXECUTIVE SUMMARY · 2 OF 2

This report enumerates the **24 areas** where disciplined agents pay off most against a real growth lever, grouped into five layers that mirror the Growth Stack. Read in order, they are a build sequence: stand up the substrate and guardrails, then the engine of state capacity, then administration, the transformation machinery, and the human systems.

WHAT TO KNOW IN FIVE POINTS

- 1 **Capability is becoming abundant.** Agents manufacture competent execution cheaply — the input that constrained every growth lever.
- 2 **So the scarce inputs flip.** Judgment, legitimacy and discipline — not compute — now decide who wins.
- 3 **It's a growth story.** Relaxing the human-attention constraint lifts the whole stack at once, from the pilot-agency to the clinic.
- 4 **Discipline matters more, not less.** Tie every agent fleet to an objective external test, and keep humans on consequential decisions.
- 5 **Order is the strategy.** Substrate & guardrails → engine → administration → machinery → human systems. Skipping a layer is the recurring failure.



The agentic era does not reward the country with the most models. It rewards the one that best decides what to point its agents at — and best holds them to account.

— ENSI, *HOW AGENTS MAKE A COUNTRY GROW*

abundant

capability — the old bottleneck on every growth lever

scarce

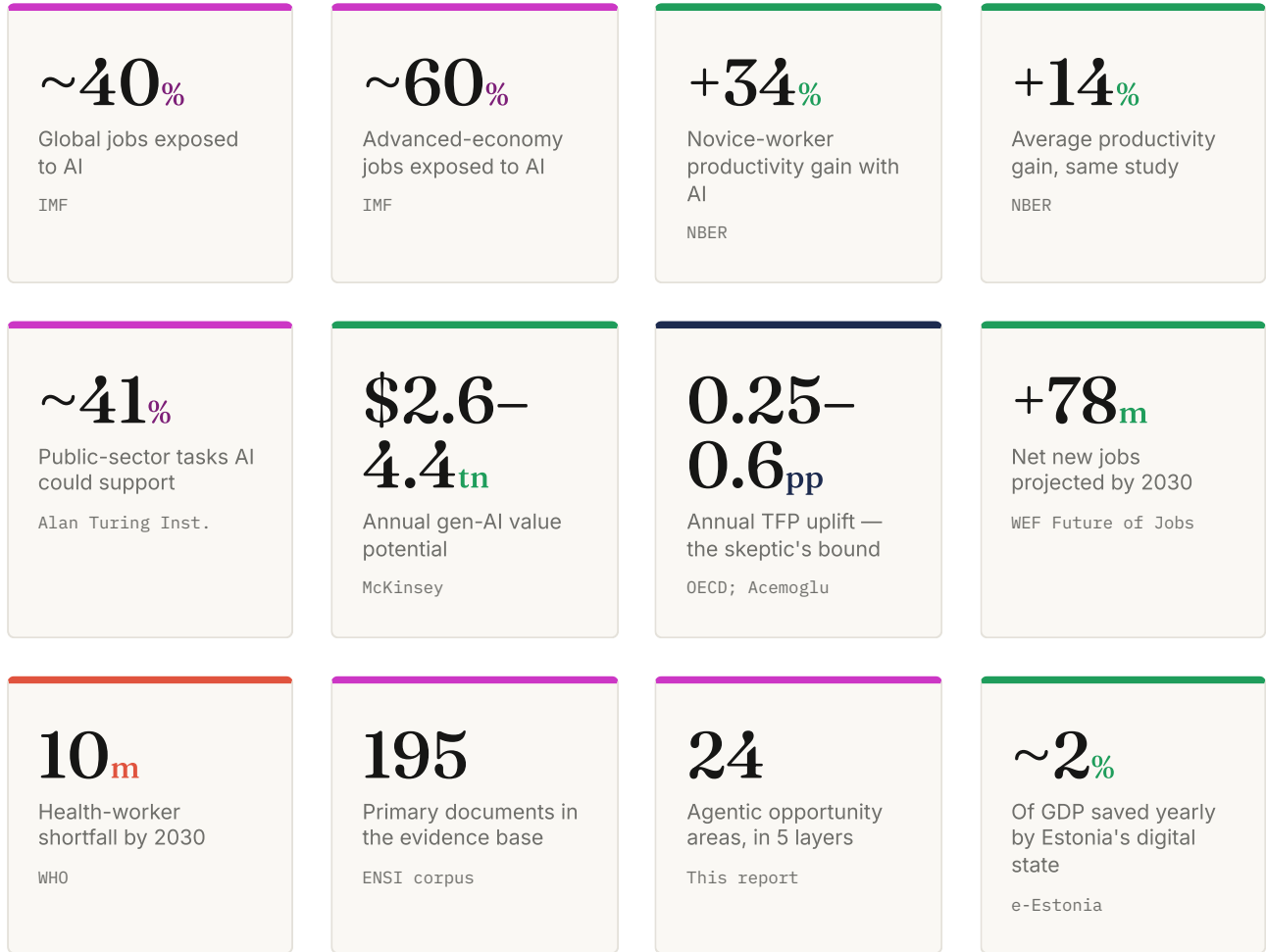
judgment, legitimacy, discipline — the new constraints

buildable

all of it is a choice, not a fate

THE AGENTIC DASHBOARD · 1 OF 2

The thesis, in numbers



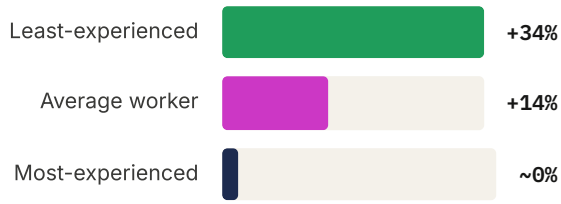
WHAT THESE NUMBERS SAY

Read together: AI's reach is economy-wide (40–60%), its gains are largest where skill was scarcest (+34%), the state is the single biggest opportunity (~41%), and the evidence is deep (195 sources). The charts overleaf show why — and the \$2.6–4.4tn annual prize depends on real adoption, not announcements.

THE AGENTIC DASHBOARD · 2 OF 2

FIG. 1

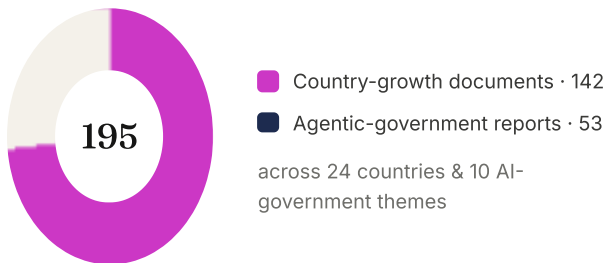
Agents diffuse capability to where it was scarcest



Takeaway: the gain is largest for novices and near-zero for experts — agents are a diffusion machine, not just a speed-up. **Source:** NBER (Brynjolfsson, Li & Raymond), customer-support field study.

FIG. 2

The evidence base



195 primary sources. See the source library, §24.

FIG. 5

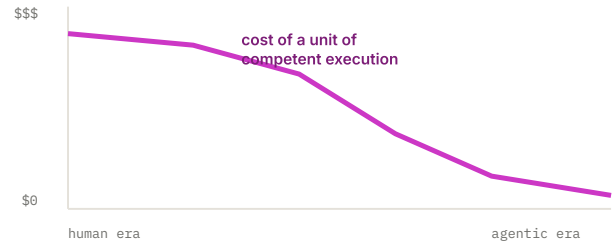
Exposure to AI is highest inside the state



Share of work AI could support, by domain. Source: Alan Turing Institute; IMF.

FIG. 3

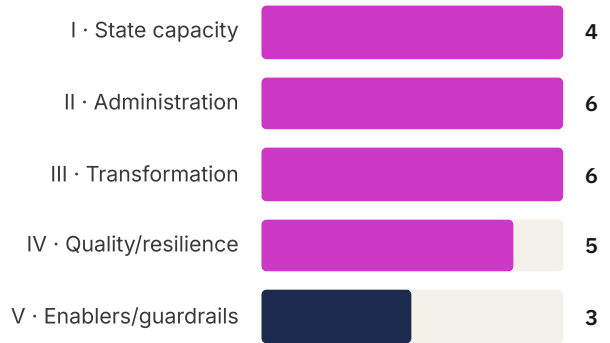
The collapsing price of competent execution



Takeaway: as agents arrive, the cost of competent execution — the input behind every growth lever — falls toward zero. **Illustrative,** after NBER/OECD productivity evidence.

FIG. 4

Where the 24 areas sit



24 areas across 5 layers — the body of this report (§13–17).

READER'S NOTE

Capability is becoming abundant (Figs. 1 & 3), the evidence is broad (Fig. 2), and the opportunity is concentrated in the state (Figs. 4 & 5). The rest of this report is how to capture it — without losing the discipline test.

THE CATCH

Value lands only where adoption is real, not announced. Most public-sector AI pilots never reach production (OECD).

THE THESIS IN ONE DIAGRAM

The Agentic Growth Stack

The 24 areas sort into five layers. Build bottom-up: the substrate and guardrails make agents possible and legitimate; the engine decides what to point them at; the upper layers are where the growth shows up.

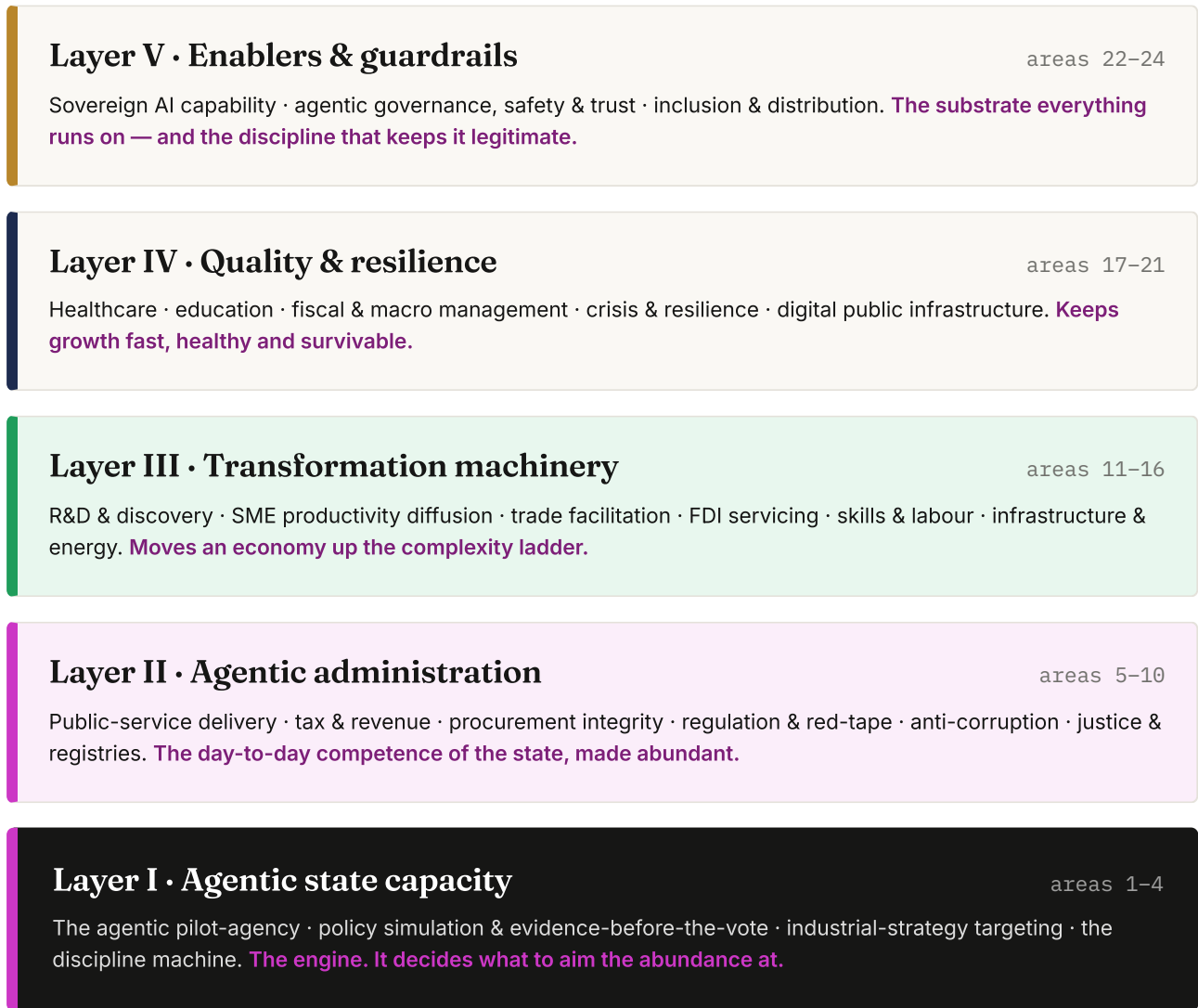


Fig. 5 · The Agentic Growth Stack. Build order runs bottom-to-top; the 24 areas in §13–17 follow this sequence.

<p>Build ↑</p> <p>Each layer is the precondition for the one above — read bottom-to-top</p>	<p>Layer I</p> <p>The engine: decides what to point the agentic abundance at</p>	<p>Layer V</p> <p>Substrate & guardrails — laid first, or nothing above is legitimate</p>
--	---	--

HOW TO USE THIS DIAGRAM

The 24 areas in §13–17 are dissected in this exact order. Where your country already has a layer working, move up; where it doesn't, build down to it first. The most common failure — in the human era and the agentic one — is skipping a layer.

TWELVE FINDINGS

What the evidence shows

- 1 Capability — not capital — was the binding constraint on growth; agents make it abundant. ENSI
- 2 The productivity gain is largest for the **least-skilled** (+34%) and near-zero for experts — a diffusion engine. NBER
- 3 ~**40%** of global jobs (60% in advanced economies) are exposed to AI — the reach is economy-wide. IMF
- 4 ~**41%** of public-sector tasks could be agent-supported — the state is the biggest single opportunity. Turing
- 5 Gen-AI's annual value potential is **\$2.6–4.4tn** — but only where adoption is real, not announced. McKinsey
- 6 Skeptics still model real gains (**0.25–0.6pp** TFP/yr) — the floor is positive, the ceiling large. OECD
- 7 Agents without a discipline test scale mistakes faster — discipline matters **more**, not less. ENSI
- 8 Digital public infrastructure is the precondition: Estonia's saves ~**2% of GDP** a year. e-Estonia
- 9 The labour transition is net-positive but disruptive: **+78m** net jobs by 2030 with mass reskilling. WEF
- 10 Human systems are starved of capacity — e.g. a **10m** health-worker shortfall — exactly where agents help most. WHO
- 11 National AI strategies are converging on compute, data and talent as sovereign assets. OECD
- 12 Order is the strategy: substrate & guardrails → engine → administration → machinery → human systems. ENSI

40% / 60%

Global / advanced-economy jobs exposed to AI (IMF)

+34%

Productivity gain for the least-skilled worker (NBER)

195

Primary documents behind these findings

THE THROUGH-LINE

Every finding points the same way: capability is becoming abundant and cheap, so the scarce, decisive inputs flip to judgment, legitimacy and discipline. The 24 areas that follow are where that shift turns into growth — and the guardrails that keep it from turning into rent-seeking at scale.

THE REFRAME · 1 OF 2

Capability was the bottleneck all along

The development debate spent fifty years arguing about *which policies* cause growth. It mostly missed that the same policies succeed or fail depending on whether a state has the capacity to execute them — and that capacity was always rationed by the supply of capable humans.

THE CONVENTIONAL VIEW	THE MECHANISM THE EVIDENCE SHOWS
"Get the policies right and growth follows."	Identical policies succeed in capable states and fail in weak ones. Execution capacity is the hidden variable.
"AI is an efficiency tool — do the same, cheaper."	Agents change <i>what is possible</i> : every lever throttled by scarce attention is released at once. It is a capacity story, not a cost story.
"Build human capital slowly, then grow."	Agents invert the sequence — capability becomes abundant first; judgment and discipline become scarce.
"More AI = more growth."	Capability without a discipline test is rent-seeking at scale. The test, not the tool, decides the outcome.
"Frontier models are the prize."	The prize is what you point agents at and how you hold them to account — a governance choice, not a purchase.



Identical reforms built Korea and wrecked their imitators. The difference was never the reform. It was the apparatus that could execute and discipline it.

— ENSI SYNTHESIS, AFTER WADE, GOVERNING THE MARKET

WHAT THE OLD FRAME COST

Fifty years of identical reforms exported to states that could not execute them — privatisation without capacity became capture; liberalisation without a test became rent-seeking.

WHAT THE NEW FRAME ENABLES

If capacity is the variable and agents supply it, the question changes from "which reform" to "what do we point the abundance at, and how do we hold it to account."

same policy

succeeds in capable states, fails in weak ones

the variable

was always execution capacity, not the menu

now buildable

agents manufacture the capacity that was scarce

THE REFRAME · 2 OF 2

Why this is a growth story, not an IT project

Treat agents as a procurement line item and you get a faster call-centre. Treat them as a way to **manufacture state capacity** and you get the thing that actually drives growth: a state that can pick bets, finance them, and cut the losers — at a scale and tempo no human bureaucracy could reach.

Era-2 thinking (a tool)

- ▶ A chatbot bolted onto a portal

- ▶ One model, many demos, little adoption

- ▶ Pilots that never reach production

- ▶ Success measured in activity, not outcomes

- ▶ Capability rented; dependence grows

Era-3 thinking (a workforce)

- ▶ Fleets of agents doing end-to-end work

- ▶ Every fleet tied to an external test it can't fake

- ▶ Human-in-the-loop on consequential calls

- ▶ Success measured in growth outcomes

- ▶ Sovereign-enough capability; control retained

10m → ∞

From thousands of scarce officials to unbounded agentic capacity

24/7

Always-on, every language, near-zero marginal cost

1 test

Each fleet disciplined by one objective metric it cannot fake

THE BOTTOM LINE

The countries that win the next era are not the ones with the most agents. They are the ones that best decide what to aim them at — and best hold them to account.

tool →

workforce

From AI-as-feature to agents-as-workers

activity →

outcome

Measure growth delivered, not pilots launched

rented →

sovereign

Retain control of the capability you depend on

WHY IT MATTERS HERE

The country that treats agents as procurement gets a faster call-centre. The country that treats them as manufactured state capacity gets the engine of growth. Every one of the 24 areas in this report is written for the second country.

THE FIVE LAYERS EXPLAINED

How the stack is built

Each layer is the precondition for the one above. The order below is the recommended build sequence.

LAYER	NAME	AREAS	WHAT IT DOES & WHY IT COMES WHEN IT DOES
V	Enablers & guardrails	22-24	Sovereign-enough compute/data/talent, the governance layer, and inclusion. Built (at least in part) <i>first</i> — without the substrate and the guardrails, every other agent is impossible or illegitimate.
I	Agentic state capacity	1-4	The engine: the pilot-agency, policy simulation, industrial targeting, and the discipline machine. Decides <i>what to aim the abundance at</i> .
II	Agentic administration	5-10	The day-to-day competence of the state — services, tax, procurement, regulation, integrity, justice — made abundant and clean.
III	Transformation machinery	11-16	The levers that move an economy up the complexity ladder: R&D, SME diffusion, trade, FDI, skills, infrastructure.
IV	Quality & resilience	17-21	The human systems and shock-absorbers — health, education, macro management, crisis response, and the DPI rails agents run on.

THE BUILD PRINCIPLE

Stand up Layer V's substrate and guardrails, then the Layer I engine that points the agents at the right targets — only then scale II, III and IV. Skipping a layer is the recurring cause of failure, in the human era and the agentic one alike.

5

Layers, built bottom-up

24

Opportunity areas

7

Analytical lenses per area

THE SEQUENCING LAW

Layer V is drawn last but built first: without the rails (DPI, area 21), the substrate (sovereign AI, 22) and the guardrails (governance, 23), every agent above is impossible or illegitimate. Only then build the Layer I engine that decides what to aim the abundance at — then scale upward.

0-18 mo

Lay rails & guardrails (Layer V)

0-18 mo

Stand up the engine (Layer I)

12-36 mo

Scale II-IV, where growth shows up

THE 24 OPPORTUNITY AREAS AT A GLANCE

The whole report in one table

#	AREA	LAYER	THE AGENTIC MOVE	HEADLINE SIGNAL
1	The Agentic Pilot-Agency	I	Analyst/planning fleet = an EDB-grade strategy unit	41% public tasks AI-supportable (Turing)
2	Policy Simulation & Evidence	I	War-game & cost a policy before the vote	Forecast-vs-outcome tracking
3	Industrial-Strategy Targeting	I	Map complexity, design the productive bets	Korea/Taiwan, done by search
4	The Discipline Machine	I	Instrument every subsidy vs an external test	Publish-before-you-cut
5	Public-Service Delivery	II	End-to-end agentic casework, 24/7	\$90bn legacy/75% (Ash Center)
6	Tax & Revenue Administration	II	Risk-based audit, fraud, pre-filled service	55-88% tax-admin AI use (IMF)
7	Procurement & Spend Integrity	II	Design tenders, detect collusion, audit delivery	Attacks the largest budget leak
8	Regulation & Red-Tape Removal	II	Draft/test rules, auto-issue permits	Months → minutes
9	Anti-Corruption & Integrity	II	Cross-check transactions, declarations, conflicts	\$80bn illicit flows (World Bank)
10	Justice, Courts & Registries	II	Triage cases, draft rulings, run registries	30k claims in 2 weeks (World Bank)
11	R&D & Scientific Discovery	III	Synthesis, hypotheses, lab automation	AlphaDev/GNoME (Stanford HAI)
12	SME Productivity Diffusion	III	Frontier capability inside every small firm	+34% for laggards (NBER)
13	Export & Trade Facilitation	III	Customs, trade-finance, market intelligence	Extends the export test downward
14	FDI Attraction & Servicing	III	Always-on IDA/CINDE + supplier matching	Ireland/Costa Rica, scaled
15	Skills & Labour-Market Engine	III	AI tutors + real-time skills-to-jobs matching	+78m net jobs by 2030 (WEF)
16	Infrastructure & Energy Ops	III	Optimise grids, ports, logistics, water	More output per unit of capital
17	Healthcare System	IV	Triage, diagnosis, admin, surveillance	10m worker shortfall (WHO)
18	Education System	IV	AI tutors + teacher copilots to mastery	Bloom's 2-sigma, cheaply
19	Fiscal, Debt & Macro	IV	Monitor revenue, debt, the real exchange rate	Catch overvaluation early
20	Crisis, Risk & Resilience	IV	Early warning, scenarios, coordinated response	Lag is the real damage
21	Digital Public Infrastructure	IV	ID, payments, data-exchange: the rails	~2% GDP/yr (e-Estonia)
22	Sovereign AI Capability	V	National compute, data, talent, models	UK/Singapore/SDAIA strategies
23	Governance, Safety & Trust	V	HITL gates, audit, eval, provenance	NIST RMF · EU AI Act
24	Inclusion & Distribution	V	Reskilling, rural/informal reach, language	Keeps growth survivable (ILO)

HOW TO READ THIS TABLE

Each area is dissected in full in §13-17 across seven lenses. "Layer" gives the build order (I first); the "headline signal" is the single most telling number or proof from the evidence base. Ordered by foundational priority, not by size of prize.

THE AGENT ARCHETYPES & THE SEVEN-LENS METHOD

How to read each area

The five agent archetypes

The same handful of agent types recur across all 24 areas. Spotting them is how you see the pattern.

- ▶ **Mapper / Analyst** — scans, structures and models a domain (the pilot-agency, complexity targeting).

- ▶ **Monitor / Scorekeeper** — instruments activity against a test and flags drift (the discipline machine, integrity).

- ▶ **Caseworker / Doer** — executes end-to-end transactions (services, tax, permits, registries).

- ▶ **Advisor / Drafter** — proposes options and drafts for human sign-off (policy, rulings, regulation).

- ▶ **Guardian / Auditor** — checks, logs and explains every action (the guardrail layer).

The seven lenses

Every area is dissected the same way, so they can be compared like-for-like.

- 1 **In short** — the opportunity and the agentic reframe.
- 2 **The growth lever** — what a country needs to grow here.
- 3 **The human bottleneck** — why scarce attention limited it.
- 4 **The agentic mechanism** — which agents, doing what.
- 5 **What it makes possible** — the impact at scale.
- 6 **How to build it** — the sequence, the test, the guardrails.
- 7 **Evidence & risks** — proof and the strongest objection.

READ THE COLOUR

Magenta marks the focal point and the agentic move; **green** marks a positive outcome; **coral** marks a risk or the steelman. Every area ends on a risk on purpose.

5 archetypes

recombine into all 24 areas

7 lenses

applied identically, so areas compare like-for-like

1 spine

every fleet tied to a test + a human gate

THE PATTERN TO WATCH FOR

In every area, a Mapper or Caseworker does the volume work, a Monitor scores it against the discipline test, and a Guardian logs and explains it — with a human on the consequential decision. Spot that loop and you have seen the agentic state.

THE BUILD SEQUENCE

The order is the strategy

A country that fields agents everywhere at once, with no sequence and no test, reproduces the failures of the human era — faster. Build in this order.

1 · Lay the substrate & guardrails (areas 21, 22, 23)

Digital public infrastructure, sovereign-enough capability, and the governance layer. Without these, every other agent is impossible or illegitimate.



2 · Build the engine (areas 1-4)

The pilot-agency and the discipline machine — the capacity to decide what to point the abundance at, and to cut what fails the test.



3 · Scale administration (areas 5-10)

Make the day-to-day state competent and clean — fast wins that fund and legitimise the rest.



4 · Run the transformation machinery & human systems (areas 11-20)

R&D, SME diffusion, trade, FDI, skills, infrastructure, health, education, macro and resilience — where the growth shows up.

THE FAILURE MODE

Every stalled programme skips a layer: agents scaled before the guardrails (illegitimate), or before the engine (pointed at the wrong target), or before the DPI rails (nothing to act through).

Step 1

Rails & guardrails — the foundation
(areas 21-23)

Step 2

The engine — what to aim the agents at
(areas 1-4)

Steps 3-4

Administration, then machinery & human
systems (5-20)

LAYER I · AREAS 1-4

Agentic State Capacity

The engine of agentic growth: the capacity to pick productive bets, finance them, and discipline them against a test the state cannot fake — built from fleets of agents rather than a few thousand scarce officials.

1

2

3

4

01 The Agentic Pilot-Agency

Every country that escaped poverty did it with an elite strategy unit — Singapore's EDB, Korea's EPB, Ireland's IDA. Almost no government can build one, because it needs the country's scarcest fifty people. The agentic reframe: a pilot-agency is not a payroll but a *function* — scanning, option-drafting, trade-off modelling, cross-ministry coordination — and that function can now be manufactured as a standing fleet of agents.

<p>41% of public-sector working time current AI could support Alan Turing Institute</p>	<p>~0marginal cost of the next analysis once the fleet runs ENSI</p>	<p>£10sbn/yr est. savings for a UK-sized state from agent throughput Tony Blair Institute</p>
--	---	--

THE GROWTH LEVER & THE BOTTLENECK

State capacity ranks first in the stack because nothing else is buildable without it (ENSI, Priority 1). The lever is **throughput**: the rate at which the system can analyze, decide, and execute without the decision being captured or reversed. The bottleneck is brutal — capable human attention is scarce and expensive. A poor state cannot hire a Singapore-grade bureaucracy; a rich state cannot scale its best fifty analysts across every decision. You can buy a port; you cannot buy fifty incorruptible polymaths on a government salary. The constraint was never the ideas — it was the number of competent hours (ENSI, *Czech State: A Systematic Analysis of Malfunction*).

WHAT IT MAKES POSSIBLE

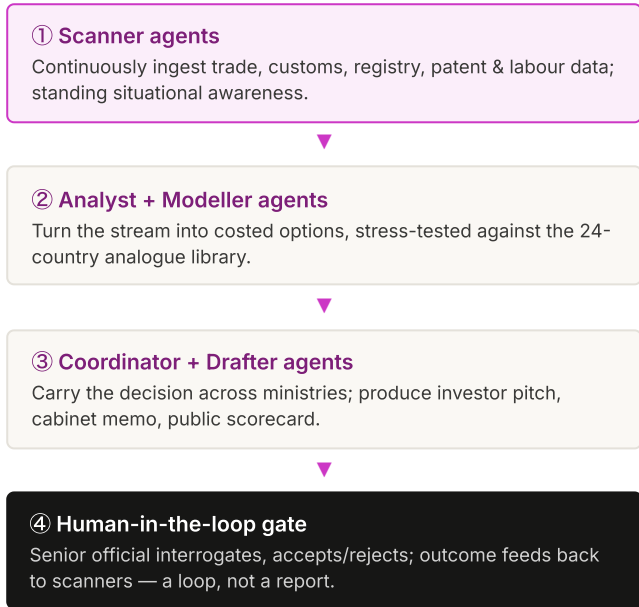
<p>50→0 scarce analysts no longer the binding input</p>	<p>24/7 continuous review of the complexity space</p>
--	--

The pilot-agency stops being a luxury good and becomes infrastructure any government can stand up — scarce humans elevated from doing the analysis to *judging* it (Tony Blair Institute).

HOW TO BUILD IT

- Secure the principal, narrow the mandate.** A head of government who protects autonomy and a concrete, measurable mission — not another dashboard (ENSI, Priority 1).
- Lay the rails.** Agents need DPI — consented data-exchange, machine-readable registries (World Bank, *GovTech Maturity Index*; UCL IIPP).
- Discipline test.** Every bet carries a falsifiable forecast and the export test — earn foreign sales in a window or lose support (World Bank/Stiglitz, *East Asian Miracle*).
- Guardrail.** HITL on every consequential call; immutable audit log of recommendation, evidence & approver (NIST, *AI RMF*; EU AI Act).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- Korea** — the EPB super-ministry ran the heavy-and-chemical push above the line ministries (ADB, *Chaebol and Industrial Policy*).
- Ireland** — the IDA identified, courted and landed specific multinationals over decades (OECD, *Economic Survey of Ireland 2025*).
- Estonia** — a small, capable administration on the right digital rails outperformed far larger bureaucracies (OECD, *Economic Survey 2024*).

THE RISK (STEELMAN)

The bottleneck was never analysis — it was political will; a flawless fleet for a government that won't cut its losers changes nothing. Agents do not supply will, but a published, falsifiable scorecard makes protecting a zombie firm a documented choice rather than a quiet drift. Pointed at a captured state, the same fleet industrialises the capture — the discipline test and the named accountable human at every gate are what stop that.

02 Agentic Policy Simulation & Evidence-Before-the-Vote

Almost every consequential policy is made blind: a subsidy, tariff or tax ships with no costed, falsifiable forecast against which the outcome can later be checked. Policy is made the way medicine was before trials — on authority and anecdote. The agentic reframe turns policymaking from opinion into experiment: a fleet models a policy *before* the vote, returns a costed forecast with a confidence range, then tracks the outcome against it and publishes the gap.

<p>3→300</p> <p>policies a state can model per year once the fleet runs</p> <p>ENSI</p>	<p>41%</p> <p>of public-sector time AI can support — incl. analysis & drafting</p> <p>Alan Turing Institute</p>	<p>2nd</p> <p>rank of the discipline mechanism in the growth stack</p> <p>ENSI · Priority 2</p>
--	--	--

THE GROWTH LEVER & THE BOTTLENECK

The discipline mechanism ranks second because it converts capacity from a danger into an asset (ENSI, Priority 2). The East Asian states did not pick winners well — they had a mechanism to **kill losers fast**, the export test, an external referee that cannot be lobbied (World Bank/Stiglitz). The value was never in selection; it was in the *exit*. The bottleneck: rigorous ex-ante analysis is staggeringly expensive in the scarcest input — capable analytic hours. No state can model more than a handful of flagship decisions a year, so the thousand others ship on whoever argued most persuasively, gathering only the evidence that agrees (ENSI, *Evidence over Anecdote*).

WHAT IT MAKES POSSIBLE

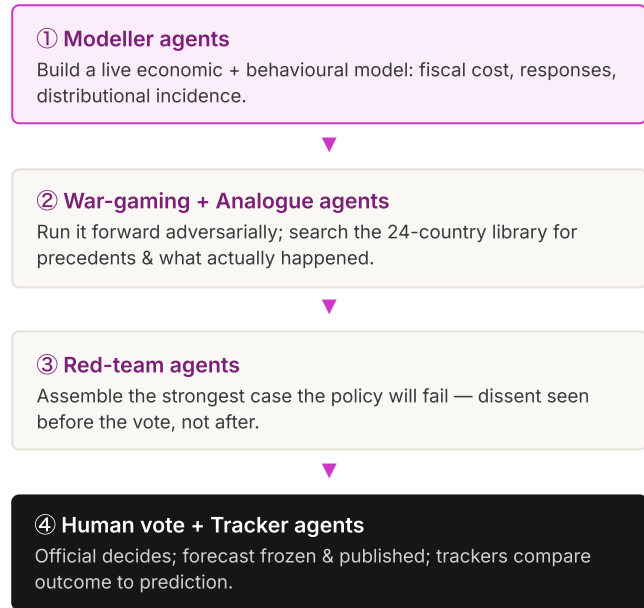
<p>frozen</p> <p>ex-ante forecast on every significant decision</p>	<p>falling</p> <p>mean forecast error as the fleet calibrates</p>
--	--

The state begins to *learn*: every forecast frozen, every outcome tracked, so it accumulates an evidence base instead of a pile of war stories (ENSI, *Evidence over Anecdote*).

HOW TO BUILD IT

- 1 **Start where forecasts are expected.** Begin with budget & fiscal decisions, then widen the frozen-forecast rule to regulation and subsidies.
- 2 **Build the model layer.** Data rails + DPI plus the *National Policy Twin*, a live queryable model of the policy estate (Tony Blair Institute; World Bank, *GovTech Maturity Index*).
- 3 **Discipline test.** The frozen ex-ante forecast itself: miss it by a defined margin and a mandatory review defaults toward exit — the generalised export test (ENSI, Priority 2).
- 4 **Guardrail.** Agents inform, an accountable human votes; immutable audit log of every model, assumption & analogue (WEF; NIST, *AI RMF*; EU AI Act).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Korea vs Brazil** — the export test separated Korea's directed credit from Brazil's (World Bank/Stiglitz, *East Asian Miracle*).
- ▶ **OECD governments** — already running AI-based simulation, scenario analysis & decision support across functions (OECD, *Governing with AI*).
- ▶ **TBI / WEF** — a worked National Policy Twin and a near-term readiness framework for agentic systems in government (Tony Blair Institute; WEF).

THE RISK (STEELMAN)

Garbage-in, confident-out: a forecast wraps contestable behavioural assumptions in a veneer of objectivity that switches off scrutiny (NIST, automation bias). And the moment a forecast becomes a target it gets gamed — officials lobby for soft predictions. The defence is the mandated red team and analogue agents kept structurally separate from the proposing ministry, plus the published assumption list — though the independence is a design problem, not a solved one.

03 Agentic Industrial-Strategy Targeting

Choosing what a country should try to make next — reading the complexity space, finding the adjacent possible, designing the bet — was the work of a handful of irreplaceable planners. Underneath, it is a search problem over a measurable space of capabilities and products, and search is exactly what a standing fleet of agents is for.

<p>24countries</p> <p>Historical record each bet is war-gamed against ENSI Growth Stack</p>	<p>monthlyrefresh</p> <p>Capability inventory update vs. once-a-decade plans World Bank, AI in the Public Sector</p>	<p>2axes</p> <p>Reachability x value, ranked per candidate rung Hausmann-Rodrik complexity</p>
--	---	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is Priority 4 of the Growth Stack — pick tradables and climb the complexity ladder — fused with Priority 13, structural transformation. The climb's hidden killer is *distance*: ambitious leaps that ignore the capability gap fail; disciplined short jumps compound. Targeting that judgment well demands holding three fast-moving things in one head — current capabilities, the global product space, and the adjacency between them — and redoing it every quarter. That attention is the scarcest, least-exportable ingredient in development; most states cannot buy a Singapore-grade planning unit, so targeting collapses into drift on static comparative advantage or capture by the loudest incumbents.

WHAT IT MAKES POSSIBLE

any

state can stand up targeting, in every sector at once

0

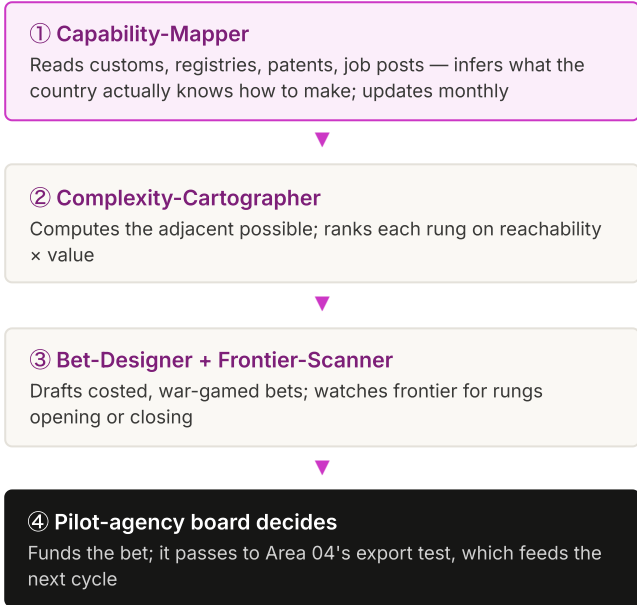
cronies hidden in a published, sourced complexity ranking

Capability distance becomes a measured quantity rather than a hunch — the difference between a disciplined short jump that compounds and an ambitious leap that fails.

HOW TO BUILD IT

- 1 **Data substrate first.** Consented, machine-readable customs/export records, firm and supplier registries, patent and standards filings; where DPI is thin, building it is the first deliverable.
- 2 **Map before you bet.** Run the Mapper and Cartographer for a full cycle and ship a ranked, sourced adjacent-possible map — and nothing else — before any money moves.
- 3 **Discipline test.** Does each targeted bet, within its window, earn hard currency from foreign buyers who owe the country nothing? Pass → support continues; fail → the fleet climbs back down and proposes another rung.
- 4 **Guardrail.** Human-in-the-loop at the bet, not the analysis; full audit trail and a *published* ranking on every recommendation to keep cronies out.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Korea** — channeled subsidized credit, then forced a climb from textiles to ships to semiconductors within a generation via the Economic Planning Board (ADB; PIIE).
- ▶ **China** — *Made in China 2025*, the most explicit published map of the climb, from low-value assembly into robotics, aerospace and advanced semiconductors (Gov. of China; MERICS).
- ▶ **Taiwan / Ireland / Vietnam** — SME-led ascent into semiconductor indispensability; FDI used as a capability-transfer machine (NDC; OECD; ERIA).

THE RISK (STEELMAN)

The future is genuinely uncertain, so a fleet will produce confident, well-sourced bets that are still wrong — only faster and at scale. Answer: the machine is not a winner-picker but a hypothesis-generator wired to Area 04's export test; without that leash it just industrializes rent-distribution, which is why the two areas are one system.

04 The Agentic Discipline Machine

Every fast grower intervened heavily; what set the winners apart was tying support to an objective external test the state could not fake — usually exports. Discipline, not direction, is the scarce ingredient, and continuous, unlobbyable, fatigue-proof measurement is exactly what an agent fleet is for.

<p>100%</p> <p>Target share of state support carrying an attached falsifiable test</p> <p>ENSI Growth Stack, Priority 2</p>	<p>Ozombies</p> <p>Supported, non-exporting, never-exited firms the machine drives toward zero</p> <p>ENSI Growth Stack</p>	<p>customs-grade</p> <p>Recipient-proof data only — no self-reports counted</p> <p>World Bank; OECD</p>
--	--	--

THE GROWTH LEVER & THE BOTTLENECK

The lever is Priority 2 of the Growth Stack: subsidize nothing without a hard external performance test. Support is conditional on selling in world markets within a defined window — foreign buyers cannot be lobbied, bribed, or charmed. The rule is easy to write and almost impossible for humans to enforce: measurement degrades into self-report theater; the verdict to cut a firm is a confrontation with a constituency that employs voters and funds campaigns; and the referee's goalposts get lobbied soft. The binding constraint is the throughput of honest, unconflicted monitoring and the political shielding of the verdict — a capacity problem, not a character problem.

WHAT IT MAKES POSSIBLE

exit

the one thing every failed industrial policy could not do

full

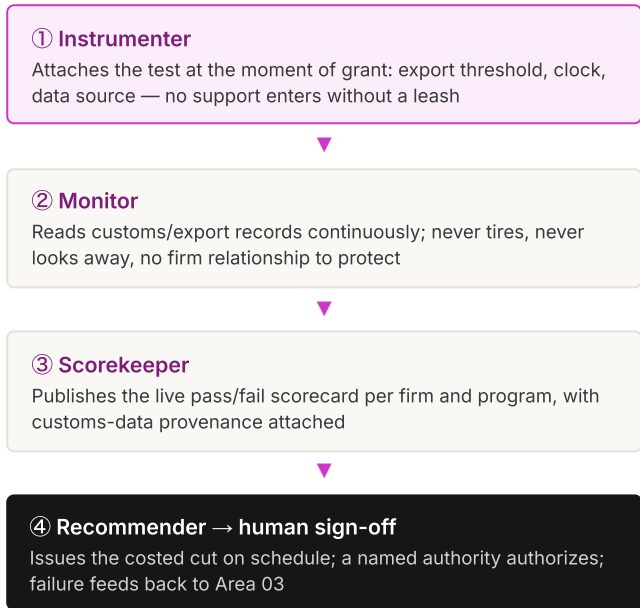
coverage — every subsidy, SEZ tenant, state loan, at once

A live, customs-data-backed scorecard makes sparing a failing crony a public act against visible evidence — honesty by publication.

HOW TO BUILD IT

- 1 **Data substrate first.** Machine-readable, consented customs and export records, the registry of every support instrument, the firm/supplier graph — recipient-proof, or there is no discipline machine.
- 2 **Instrument, then publish early.** Stand up the Instrumenter so no new support enters untested, back-fill the existing portfolio, and publish the scorecard from the start — the public verdict is the mechanism, not a flourish.
- 3 **Discipline test.** Does each supported firm, within its window, earn hard currency from foreign buyers who owe the country nothing? Pass → support continues; fail → it ends, automatically and visibly.
- 4 **Guardrail.** Human-in-the-loop on the cut with an appeal path; audit trail on every verdict; protect the referee — every change to a threshold logged and published; falsifiable conditions only.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Korea** — withdrew subsidized credit from chaebol that missed export benchmarks, forcing a climb from textiles to ships to semiconductors in a generation (ADB; PIIE).
- ▶ **China / Singapore** — SEZ firms competed for global supply-chain orders with officials promoted on export performance; incentives routed to export-purposed MNCs (PIIE; IMF/Ghesquiere).
- ▶ **Botswana** — excellent fiscal discipline over diamond rents, but no manufacturing export contest, so it struggled to diversify — absence proves the rule (World Bank; NRG).

THE RISK (STEELMAN)

The machine never failed for lack of measurement — it failed for lack of political will, and an agent recommending a powerful firm's cut will simply be overruled or quietly defunded. Answer: it cannot manufacture will, but it moves the verdict from private discretion to a public, sourced record rendered *before* any official can bury it, making not-cutting expensive, visible, and attributable.

LAYER II · AREAS 5-10

Agentic Administration

The day-to-day competence of the state — services, tax, procurement, regulation, integrity, justice — made abundant, fast, and clean. The fast wins that fund and legitimise everything above.

5

6

7

8

9

10

05 Agentic Public-Service Delivery

Every government runs on one scarce input — capable human attention — and competent casework does not scale. The agentic move is to manufacture the caseworker rather than hire one: an autonomous digital worker that reads the request, retrieves the facts, applies the rule, decides or drafts, and explains itself, in any language, at near-zero marginal cost.

<p>27%</p> <p>US citizens satisfied with government digital services Harvard Ash Center</p>	<p>92%</p> <p>say better digital service would improve their view of government Harvard Ash Center</p>	<p>90\$bn/yr</p> <p>US federal tech spend — ¾ just keeping legacy alive Harvard Ash Center</p>
--	---	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is state effectiveness, and through it, legitimacy: a state that can reliably register a business, pay a benefit or issue a permit earns the consent that lets reform survive an electoral cycle. The bottleneck is structural — casework has four steps (intake, retrieval, reasoning, resolution) that each consume scarce attention, capacity is bound to labour, and a poor state cannot recruit a Singapore-grade bureaucracy. With one-third of the US SSA's ~22,000 staff due to retire within five years as claims rose (Harvard Ash Center), the scissors of rising need and falling capacity are universal.

WHAT IT MAKES POSSIBLE

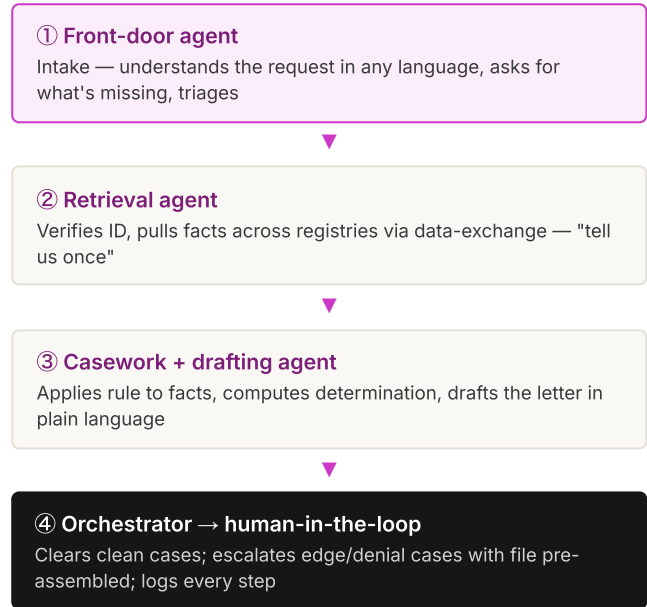
<p>24/7</p> <p>always-on service without proportional headcount</p>	<p>~0</p> <p>marginal cost per extra language or channel</p>
--	---

Capacity expands without expanding labour (TBI); personalisation and prevention replace standardised, reactive provision; inclusion of minority-language and rural citizens becomes a side effect, not a special programme (World Bank).

HOW TO BUILD IT

- 1 **Lay the rails first.** Digital ID, a data-exchange layer, and readable/writable systems of record — without DPI an agent has nothing to act through (World Bank).
- 2 **Pick & instrument one service.** A single bounded, high-volume, rule-governed benefit; measure the baseline end-to-end before automating, then widen autonomy case-type by case-type.
- 3 **Discipline test.** Resolution time, accuracy vs. human adjudication, appeal-and-overturn rate, cost per case and satisfaction — published against the baseline; an agent that misses is switched off, not defended.
- 4 **Guardrail.** Human-in-the-loop on denials and high-stakes/low-explainability cases (IMF); immutable audit log of inputs, rule, reasoning, actor; consent- and purpose-limited data exchange.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **US cities & Japan** — five core capabilities (answering, document-handling, routing, translation, drafting) already operating in citizen services (Harvard Ash Center, 2017).
- ▶ **World Bank survey** — governments using AI across citizen engagement, fraud/anti-corruption, business-process automation and service delivery — the exact agentic stack.
- ▶ **Leading digital states** — UN e-government & World Bank GovTech work show the rails (digital ID, data exchange, systems of record) already built, making the agentic layer buildable.

THE RISK (STEELMAN)

A wrongly denied benefit is not a bad film recommendation — at population scale a biased model produces wrong denials systematically and invisibly, and most World Bank client countries are "not ready," lacking the data quality and frameworks the model assumes. The answer is the design itself: rails and data quality first, human-in-the-loop at high-stakes decisions, audit + appeal mandatory — the same discipline that separated Korea from Brazil, riding on top of the cost collapse.

06 Agentic Tax & Revenue Administration

A state is what it can collect — not the tax rate but the tax machine. Broad, fair collection requires competent attention applied to millions, so most states collect from the few they can see and let the base leak. Agents — autonomous workers that select audits, detect fraud in real time, nudge compliance and serve taxpayers — make that attention cheap to manufacture.

<p>55%</p> <p>surveyed tax administrations already self-report some AI use IMF, 2024</p>	<p>88%</p> <p>of EU member states' tax administrations use AI (24 of 27) Univ. of Antwerp</p>	<p>18/priority</p> <p>fiscal capacity in the Growth-Stack — funds everything downstream ENSI Growth-Stack</p>
---	--	--

THE GROWTH LEVER & THE BOTTLENECK

The lever is fiscal capacity and the accountability loop it disciplines: "Coercion collects pennies; legitimacy collects the GDP." A broad, even-handed base delivers predictable finance, cheaper borrowing and consent for voluntary compliance. The bottleneck is a vast sorting problem on scarce attention — an administration can audit only a tiny fraction of returns, so the deterrent rests on selecting the right fraction, while fraud moves at transaction speed and the informal base is simply invisible. "A taxpayer the state cannot identify is a taxpayer the state cannot tax" — the narrow-base trap is the shadow of scarce attention.

WHAT IT MAKES POSSIBLE

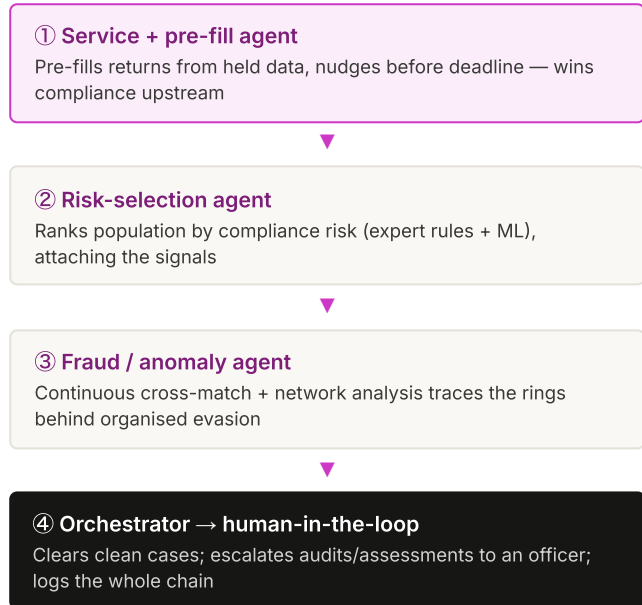
<p>broad</p> <p>base becomes collectable — "broad base, moderate rate" made executable</p>	<p>real-time</p> <p>carousel & refund fraud caught in the stream, not months later</p>
---	---

Compliance by design (OECD 3.0 "making tax just happen"); enforcement that is sharper and fairer — fewer pointless audits of the honest, and the powerful scrutinised exactly as everyone else.

HOW TO BUILD IT

- 1 **Build the spine.** Universal taxpayer identity on the national digital ID + data-exchange rails for third-party data — without legibility there is nothing to reason over (World Bank).
- 2 **Win the upstream, then promote.** Deploy service + pre-fill first (cheapest revenue is the correct-by-default return), run risk-selection in shadow mode, promote to live only where it beats baseline; add real-time fraud last.
- 3 **Discipline test.** Audit hit-rate and revenue-per-audit vs. human selection, plus false-positive rate on the compliant; an agent that misses on effectiveness *or* fairness is switched off (IMF: "Evaluate Use Cases for Performance and Intent").
- 4 **Guardrail.** Human-in-the-loop on every punitive determination (IMF; legally required); published AI-use inventory + disclosure + audit log; bias-testing on selection models trained on historical audits; consent- and purpose-limited data exchange.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Estonia** — taxation built on national digital ID and the X-Road backbone makes filing near-automatic and evasion structurally hard; among the lowest compliance costs in the OECD.
- ▶ **India** — taxation built atop the Aadhaar identity and digital-payments stack, formalising millions of new taxpayers and widening the taxable base.
- ▶ **Rwanda** — rebuilt a revenue authority from near-zero and sharply raised its domestic-revenue share, cutting aid dependence (Priority 18).

THE RISK (STEELMAN)

Tax is where an opaque, biased, fast machine does the most damage — a selection model trained on historical audits encodes old bias and aims the state's coercive power at the politically weak, and the IMF rates several high-value enforcement cases only *medium* on explainability. The answer is the build's own order: legibility and the upstream service win before enforcement, shadow-mode and a published fairness test, human-in-the-loop and the audit log tightest where the machine is least transparent. Worth building only if it collects the GDP the way legitimacy does —

07 Agentic Procurement & Spend Integrity

Procurement is the single largest controllable leak in most state budgets — lost not to dramatic theft but to overpriced contracts, rigged tenders and phantom deliveries. The agentic move: a standing fleet that designs tenders against the market, screens every bid for collusion, and audits whether what was paid for was delivered.

<p>500 firms</p> <p>owned by the civil servants supervising their contracts (Brazil) World Bank</p>	<p>37%</p> <p>of integrity bodies rank evidence-gathering & document review the top GenAI use OECD 2024</p>	<p>198 economies</p> <p>tracked on 48 indicators by the GovTech Maturity Index World Bank</p>
--	--	--

THE GROWTH LEVER & THE BOTTLENECK

Clean procurement maps to two Growth-Stack priorities: **fiscal efficiency** (18) — every unit overpaid for a road is a road un-built — and **clean governance** (11), where rigged contracts sever the link between merit and reward a market needs. The constraint is throughput, not character: a state already has rules against overpaying; it lacks the bandwidth to check, contract by contract, whether they were followed. Cross-checking a tender means reconciling bidders against ownership registries, prior awards, price benchmarks and conflict declarations — expensive, repetitive work done by sampling, late, or never. Armenia gave its anti-corruption commission an asset-declaration register but its *limited resources* made *meaningful checks* impossible (OECD). The data was filed; nobody could read it.

WHAT IT MAKES POSSIBLE

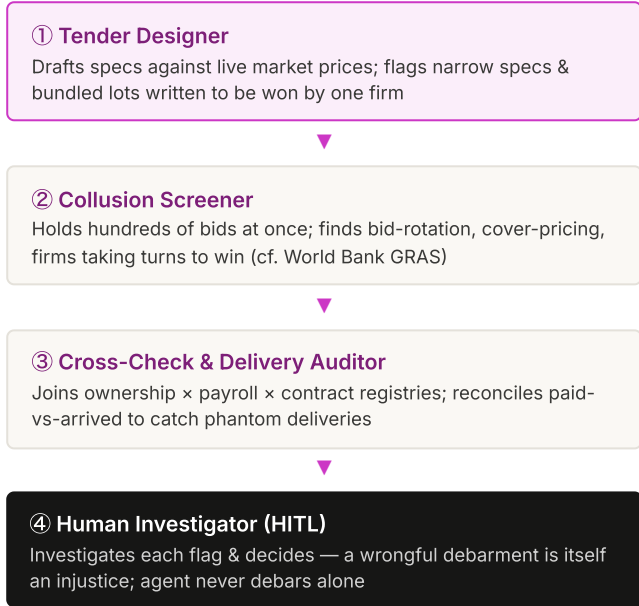
<p>100%</p> <p>of tenders screened, not a tip-off sample</p>	<p>live</p> <p>detection at transaction, not the post-mortem</p>
---	---

Population-scale screening replaces sampling; the buyer gains total recall of every price paid and every firm caught; designed-out discretion becomes the cheap default rather than the costly aspiration.

HOW TO BUILD IT

- 1 Fix the data before the model.** Structure and link supplier-ownership, payroll, awards, benchmarks and contracts first — Armenia's declarations weren't even machine-readable; an agent on garbage produces confident garbage (OECD).
- 2 Start where the join is unambiguous.** Begin with conflict-of-interest detection (the Brazil pattern); crisp wins buy cover for fuzzier collusion work, then move screening upstream from awards to bids to tender design.
- 3 Discipline test.** A published scorecard: tenders screened, value of confirmed collusion/conflicts, price anomalies caught, delivery shortfalls recovered, and unit prices now-vs-before. No catch in money = theatre.
- 4 Guardrail.** Human-in-the-loop on every consequence; tamper-resistant audit logs & provenance on every flag; transparency of method with a privacy boundary; independence from the awarding authority.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **World Bank** — the Governance Risk Assessment System (GRAS): *advanced data analytics for detecting fraud, corruption, and collusion in public expenditures* (2023).
- ▶ **Brazil** — on a World Bank-funded project, AI detected 500 firms owned by the civil servants supervising their own contracts — a join of three registries.
- ▶ **Spain** — the General Comptroller (IGAE) uses supervised ML, trained on proven cases, to detect fraud in public grants (OECD 2024).

THE RISK (STEELMAN)

Unguarded anti-fraud has a body count: the Netherlands' *Toeslagenaflaare* and Australia's *Robodebt* (~470,000 wrong debt notices, ~€775m demanded) wrongly accused people at scale; a captured insider can poison the model to lock out rivals. The answer is not non-deployment but the guardrails — HITL adjudication, provenance, independence, and the published scorecard that turns "no one did" into "no one could"

08 Agentic Regulation & Red-Tape Removal

Regulation is the operating system of an economy — and most states run the slow, corrupt version by default, where every approval queue becomes a toll booth. The agentic move: a fleet that drafts and stress-tests rules before they ship, auto-issues permits against the published rules, and supervises markets continuously — turning months into minutes.

<p>80%</p> <p>of benefit enquiries resolved without a civil servant by Norway's NAV agent <i>Frida</i> OECD</p>	<p>70%</p> <p>of surveyed countries have used AI to enhance internal government operations OECD</p>	<p>£10sbn/yr</p> <p>potential savings for a UK-sized state from governing with AI Tony Blair Inst.</p>
--	--	---

THE GROWTH LEVER & THE BOTTLENECK

Red tape maps to two priorities: **property rights & ease of doing business** (10) — an unprovable title is dead capital, an un-registered idea is no firm — and **competition** (23), because *capture operates through rules, not just market power*, and red tape is the incumbent's moat. The constraint is throughput, not the rulebook's length. Attention is scarce in both directions: writing rules well means tracing every interaction with the existing thicket; running them well means processing a flood of permits no office can sustain. The cost is regressive — a large firm amortises a compliance department; the entrant faces the same fixed wall with no volume, so it stays home. Friction is *a direct tax on production* (OECD).

WHAT IT MAKES POSSIBLE

afternoon

idea → registered company, not a month

0

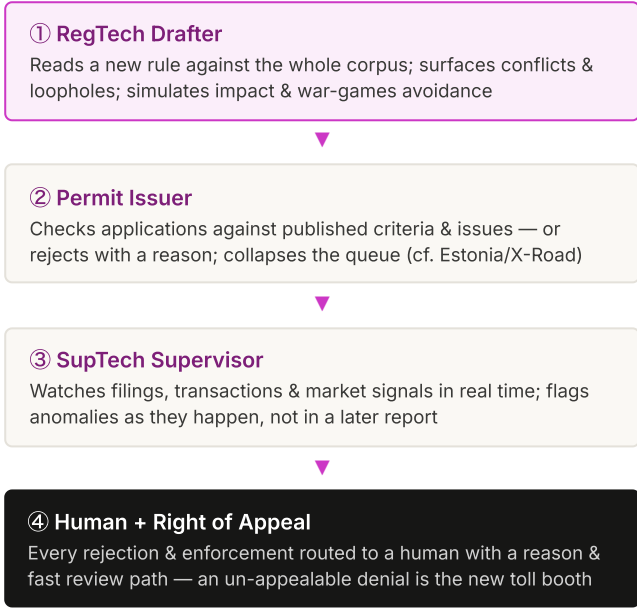
discretion left to sell for a faster "yes"

Time-to-firm collapses; de-moating becomes tractable when an agent can read all the rules at once; supervision shifts from periodic autopsy to live monitor; rules ship with side-effects modelled, not discovered in the field.

HOW TO BUILD IT

- 1 **Fix data & digitise rules first.** An agent cannot issue against rules locked as prose in a drawer; build the digital rights stack — registries, identity, structured rules — the Estonia/X-Road precondition, before any ML (OECD).
- 2 **Start with permit automation.** Auto-issue the most rule-bound, highest-volume registrations (company registration, routine licences) where criteria are objective, the win is visible, and a wrong issuance is low-stakes and reversible.
- 3 **Discipline test.** Publish median time-to-permit and time-to-firm before/after, share of approvals issued without human touch, rules rationalised away, and supervision flags that became enforced actions.
- 4 **Guardrail.** HITL on every rejection & enforcement; a right to explanation and appeal; transparency of method; audit logs & provenance; an independent supervisor the regulated industry cannot quietly retrain.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Estonia** — digitised the entire land/business/contract rights stack into a tamper-evident e-state via X-Road; *a leader in e-government* (OECD, Economic Surveys 2024).
- ▶ **Rwanda** — 3rd-best country for business in Africa and among the world's most-improved (World Bank Doing Business); targets the top 10 by 2035 via the Irembo platform.
- ▶ **Singapore** — ranked #1 for ease of doing business (IMF), uses regulatory sandboxes to *support innovation and risk-taking* — the template for testing agentic regulation safely.

THE RISK (STEELMAN)

Automated regulation run unguarded builds a faster, more opaque toll booth: *Robodebt* issued ~470,000 wrong debt notices (~€775m demanded) with no real appeal. The same agent that can de-moat a thicket can be tuned to protect incumbents and bury entrants in flags. The answer is the guardrails — a human and an appeal behind every denial, transparency, independence, and the scorecard that turns "automated" into "faster, fairer, and provably so."

09 Agentic Anti-Corruption & Integrity

Corruption is not first a morality problem but a *predictability* problem — a variable tax that severs merit from reward. A standing fleet of always-on agents cross-checks every declaration, tender and ownership flow against the public record, making graft visible and expensive the moment it happens, not years later in a report no one reads.

<p>500firms owned by the civil servants supervising their own contracts — surfaced by one AI join World Bank</p>	<p>\$80bn/yr illicit financial flows out of Africa alone, reconciled across terabytes UNECA / World Bank</p>	<p>59bodies integrity actors surveyed; 88% enforcement/audit/prosecution OECD 2024</p>
---	---	---

THE GROWTH LEVER & THE BOTTLENECK

Clean governance sits in the same load-bearing tier as property rights: control corruption and a country becomes faster, cheaper to operate in, and more meritocratic (Singapore — IMF/Ghesquiere; Botswana — Acemoglu, Johnson & Robinson). The report's prescription for priority 11 is to *attack discretion first* — every point where an official says yes-or-no for a price is a corruption site. But anti-corruption is monitoring work, and monitoring is exactly what scarce human attention cannot do at scale. Armenia's Corruption Prevention Commission collected every asset declaration and then could not read them: "given the large number of officials... and the CPC's limited resources, it was difficult to perform any meaningful checks" (OECD 2024).

WHAT IT MAKES POSSIBLE

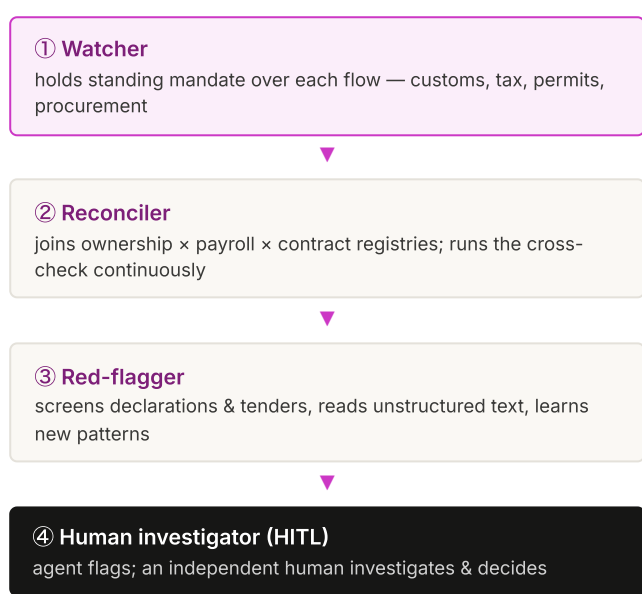
<p>100% declarations screened, not a sampled few</p>	<p>real-time detection at the transaction, not post-mortem audit</p>
---	---

Agents change the *price of vigilance*: a poor state cannot recruit a Singapore-grade integrity corps, but it can rent or run a fleet that does the cross-checking that corps would have done.

HOW TO BUILD IT

- 1 Fix the data first.** Structure and link registries (ownership, payroll, contracts, declarations, land) before pointing agents at them — Armenia's decisive early lesson; declarations were not even machine-readable (OECD 2024).
- 2 Start where the join is crisp, then learn.** Begin with Brazil-type conflict-of-interest detection; deploy rule-based screening; layer pattern-learning once trusted.
- 3 Discipline test.** Publish the catch — rate of declarations screened, value of flagged-and-confirmed conflicts, procurement price anomalies caught. A fleet that cannot show its catch is theatre.
- 4 Guardrail.** HITL on every consequence, tamper-resistant audit logs & provenance, transparency about method (not the algorithm), independence from the executive (NIST AI RMF; OECD; EU AI Act).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Armenia** — CPC built a linked data platform, deployed automated red-flag screening of asset declarations, now piloting pattern-learning AI (OECD 2024, Box 2.5).
- ▶ **Brazil** — World Bank-funded AI detected 500 firms owned by the civil servants supervising their contracts (World Bank).
- ▶ **Spain** — tax authority uses machine learning to assess public-grant fraud risk (OECD, *Countering Public Grant Fraud in Spain*).

THE RISK (STEELMAN)

NSW ICAC warns the same tools cut both ways: an actor could poison training data or manipulate models to turn the watchdog into the predator, and a false positive against an honest official is itself an injustice (OECD 2024). The answer is not abstention but the guardrails — HITL adjudication, tamper-proof provenance, transparency about method, independence — since the status quo, where the state cannot read the declarations it already collects, simply keeps graft cheap.

10 Agentic Justice, Courts & Registries

Property rights and contract enforcement that investors *actually trust* are not the plumbing of growth but its foundation. The binding constraint is throughput: courts clog and registries fall years behind because legal attention is scarce. A standing fleet of agents triages dockets, drafts rulings for human sign-off, and keeps registries current — turning justice that arrives too late into justice worth having.

<p>30,000 claims</p> <p>UK pension backlog cleared by RPA in two weeks World Bank</p>	<p>2 weeks</p> <p>vs. an estimate of thousands of hours over many months of manual work World Bank</p>	<p>2m pages</p> <p>UK government web pages classified & streamlined for citizen-centric delivery World Bank</p>
--	---	--

THE GROWTH LEVER & THE BOTTLENECK

Secure, trusted rights are the precondition that makes every other priority *bankable*: without them FDI stays footloose, firms stay small and informal, and credit stays scarce because no one will lend against collateral a court won't protect (Botswana — Acemoglu, Johnson & Robinson; Mauritius — Subramanian, UNU-WIDER). The report's priority-10 build order reads as an agentic stack: registries first, then enforcement speed, all digitised so records cannot quietly disappear. The constraint is human throughput — you cannot mint experienced judges or ask clerks to read faster, so dockets clog and a title you cannot prove to a stranger is *dead capital*.

WHAT IT MAKES POSSIBLE

minutes

company registration & tax filing
(Estonia proof-of-concept)

months

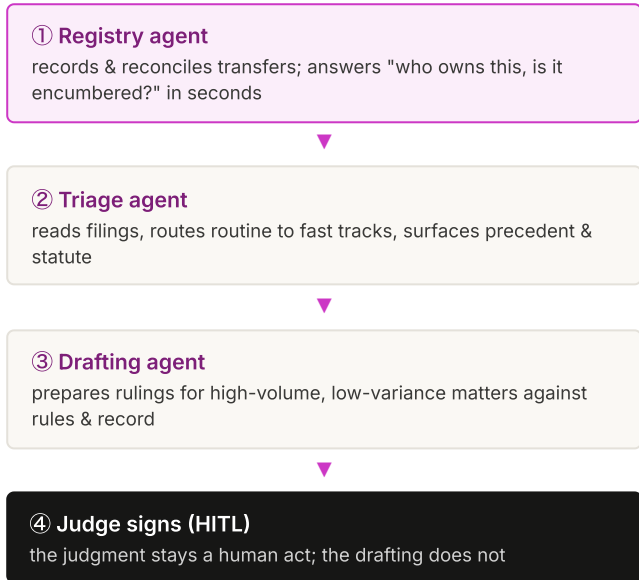
not years — median dispute resolution, the report's target

A current, queryable registry turns dead capital live so banks lend against it; collapsing the docket re-prices every long-horizon bet in the economy.

HOW TO BUILD IT

- 1 **Registries before rulings.** Digitise & structure land, company and collateral registries, then run agents to keep them current — the RPA-grade work with the cleanest payoff and lowest sensitivity.
- 2 **Triage before drafting; drafting narrow, supervised, last.** Multiply judicial throughput before any agent touches the substance; expand the envelope only as the review record proves reliability.
- 3 **Discipline test.** Publish median time-to-judgment, registry currency & search latency, the share of agent-drafted rulings accepted unedited vs. overturned, and the reversal rate of triaged cases.
- 4 **Guardrail.** Judgment stays human; provenance & explainability on every output; bias auditing; tamper-resistant logs; a right to human review (EU AI Act classes justice AI high-risk; NIST AI RMF; OECD).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **United Kingdom** — RPA cleared 30,000 pension claims in two weeks vs. months of manual work (World Bank) — the docket/registry backlog twin.
- ▶ **France** — LLaMandement summarises legislative text, demonstrating the supervised legal-drafting layer (OECD 2024).
- ▶ **Estonia / Rwanda** — minutes-long company & tax portals; aggressive property-registration & contract-enforcement reforms (growth report; World Bank, *Rwanda CEM*).

THE RISK (STEELMAN)

Justice is where automation's failure modes are least tolerable: a hallucinated precedent or a biased training record corrupts the rule of law with the false authority of a machine, and a litigant who senses a machine-made decision may not accept it (World Bank: the legal environment for AI "does not yet exist" in most countries). The answer is the boundary, not abstention — human judgment at the point of decision, narrow supervised drafting, provenance and bias auditing — since the status quo is a backlogged judiciary where rights sit unprovable as dead capital.

LAYER III · AREAS 11-16

Transformation Machinery

The levers that move an economy up the complexity ladder: research, the diffusion of capability to small firms, trade, investment, skills and the physical capital that carries it all.

11

12

13

14

15

16

11 Agentic R&D & Scientific Discovery

A country that only assembles other people's designs captures the thin assembly margin; a country that *invents* captures the durable technology rents. The national innovation system is the engine that makes the difference — and its binding constraint was never the research budget but the scarce senior attention that runs the discovery cycle serially. Agentic R&D turns that cycle into a parallel, around-the-clock process.

<p>14/32 Innovation system — Growth-Stack priority rank ENSI 32 Priorities</p>	<p>2023 Year AI moved from periphery to engine of science (AlphaDev, GNoME) Stanford HAI</p>	<p>0.25–0.6pp/yr Skeptics' aggregate TFP ceiling — why discipline is load-bearing OECD / NBER</p>
---	---	--

THE GROWTH LEVER & THE BOTTLENECK

For a country that has already built state capacity and an industrial-learning engine, the innovation system (priority 14) converts temporary catch-up into permanent frontier membership. Aghion, Jones & Jones (NBER WP 23928) show ideas are the binding input to growth and are getting *harder* to produce; the scarce input to ideas is a handful of senior minds who read the literature, frame hypotheses, design experiments and read results — serially, in working hours. Automating idea-production relaxes the one constraint growth theory says binds.

WHAT IT MAKES POSSIBLE

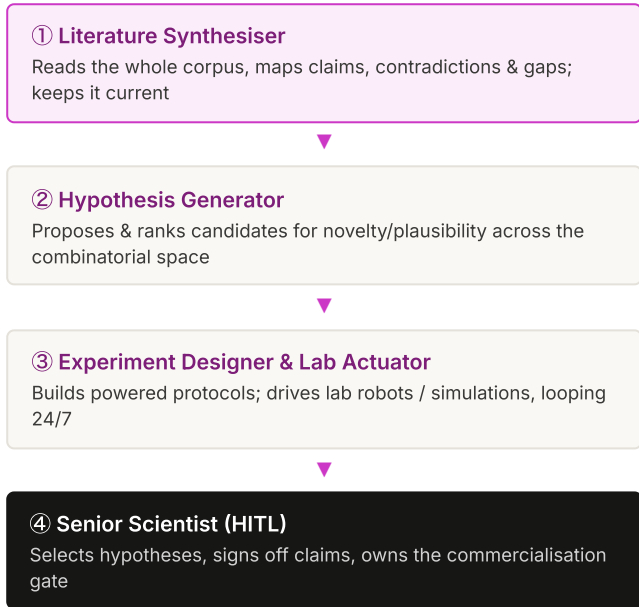
- | | |
|--|---|
| <p>1
Run an innovation system <i>before</i> a country is rich enough to staff one — an agentic ITRI</p> | <p>2–3
National missions become affordable to coordinate & run</p> |
|--|---|

The bridging capability — the connective tissue between lab and firm that the report names the true constraint — is partly manufactured, raising the *rate* of indigenous innovation that beats the middle-income trap.

HOW TO BUILD IT

- 1 Don't build it early.** A layer-three move; before state capacity (P1) & the learning engine (P6) it is an automated paper mill.
- 2 Bridge before brilliance.** Stand up the frontier-scanning, product-translating fleet first; concentrate it on 2–3 national missions, then close the loop to the lab and diffuse to the supplier base.
- 3 Discipline test.** Not "did we publish?" but *did the discovery get commercialised by a domestic firm into something sold on a world market?* Score in products shipped & export revenue; kill missions that fail.
- 4 Guardrail.** Human judgment at the gate (models "cannot reliably deal with facts" — HAI); full provenance/reproducibility audit trail; sovereign control of the discovery estate.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Stanford HAI** — *AI Index 2024* records AlphaDev finding faster sorting algorithms & GNoME accelerating materials discovery; names 2023 the year AI became an engine of science.
- ▶ **Taiwan / Korea** — ITRI seeded a semiconductor industry across thousands of SMEs; *Innovative Korea* (World Bank) shows a bridging institution, not a budget, converts knowledge into rents.
- ▶ **Israel** — Office of the Chief Scientist & the Yozma programme built one of earth's densest concentrations of high-wage technical employment by design, not wealth.

THE RISK (STEELMAN)

Acemoglu's *Simple Macroeconomics of AI* (NBER WP 32487) puts the decade TFP gain near half a point — discovery may compress at the bench while the aggregate barely moves, because the constraint shifts to the *un*-automated steps of commercialisation. The answer is the discipline test: automate the lab *and* drag the discovery across the valley to a domestic firm that the world market tests.

12 Agentic SME Productivity & Capability Diffusion

The biggest aggregate-productivity prize in a developing economy is not the frontier firm — it is the long tail of small firms operating far below it. The industrial-learning engine beats the middle-income trap by diffusing capability from the few firms that have it to the many that do not. Agentic SME productivity puts a frontier-grade consultant, marketing department, financial analyst and process engineer *inside every small firm at once*.

<p>6/32</p> <p>Industrial-learning engine — ranks above innovation (14)</p> <p>ENSI 32 Priorities</p>	<p>+34%</p> <p>Productivity gain for novice / low-skill workers (vs ~0 for experts)</p> <p>NBER WP 31161</p>	<p>5,179</p> <p>Live workers in the first at-scale generative-AI workplace study</p> <p>Brynjolfsson et al.</p>
--	---	--

THE GROWTH LEVER & THE BOTTLENECK

Aggregate productivity is a weighted average dragged down by an enormous unproductive tail; the bottom half of firms getting 30% better moves the national average enormously. The OECD's *Impact of AI on Productivity, Distribution and Growth (2024)* makes closing the frontier-laggard gap the pivotal variable. Capability has stayed trapped because its carrier is a competent person — and there is one McKinsey consultant per several-million small firms; none will visit a workshop in a secondary city. Diffusion has always been rate-limited by scarce human carriers.

WHAT IT MAKES POSSIBLE

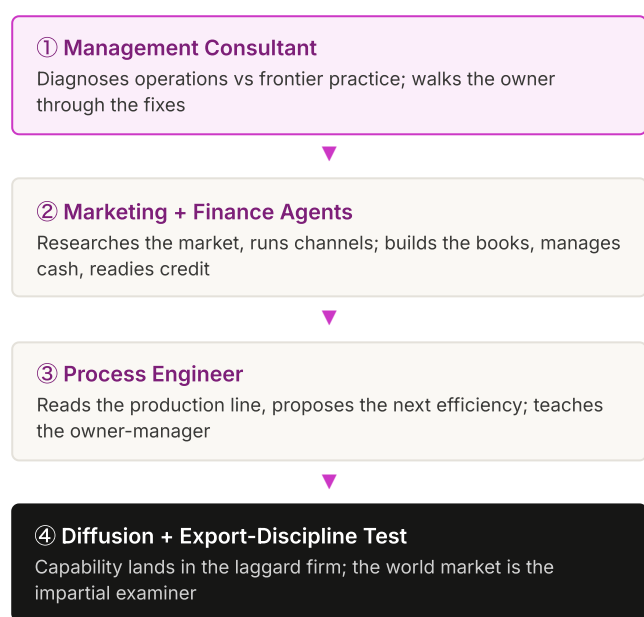
<p>+14%</p> <p>Average productivity lift; gains <i>persisted</i> through outages — real learning</p>	<p>base</p> <p>Supplier networks upgradeable at population scale; informal sector reachable</p>
---	--

Because gains concentrate on the least-skilled, diffusion is distributionally progressive *by its own mechanism* — it raises the floor, keeping the politics of growth survivable.

HOW TO BUILD IT

- 1 **Fix the foundation.** A P6 move — capability poured into firms under an overvalued currency (P5) or no contract enforcement (P10) is wasted.
- 2 **Point at the laggards.** Left to the market, the frontier adopts first and the gap *widens* (OECD's bad outcome); route capability to the bottom, anchored to existing supplier-development & standards bodies; build for learning, not lock-in.
- 3 **Discipline test.** Does the firm raise measured productivity and win sales on a competitive (ideally export) market within a defined window? If not, the support stops — the agents that diffuse also instrument the test.
- 4 **Guardrail.** Firm owns its data (no foreign-platform dependency); high-stakes outputs (credit, filings, structural eng.) route through human sign-off; resource rural/informal/female-led firms by design (ILO).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **NBER WP 31161** — *Generative AI at Work*: +14% avg, +34% novice across 5,179 live workers; mechanism is disseminating top performers' tacit knowledge; gains persisted after the tool was removed.
- ▶ **Taiwan / Costa Rica** — semiconductor capability diffused across thousands of exporting SMEs via ITRI; a dense supplier base built around medical-device anchors, not a single champion.
- ▶ **OECD (2024)** — frames closing the frontier-laggard firm gap as the variable that decides whether AI grows an economy broadly or merely concentrates it.

THE RISK (STEELMAN)

OECD's *Miracle or Myth?* and Acemoglu (NBER WP 32487) warn firm-level gains may not aggregate — and worse, capability diffusion *without* the test postpones the exit of firms that should fail, locking resources in low-productivity uses. Pointed nowhere and gated by nothing, agentic capability is a subsidy; pointed at the laggards and gated by the export test, it is the cheapest route to the largest productivity prize a country has.

13 Agentic Export & Trade Facilitation

Exporting is buried under specialist administrative work — HS classification, customs, trade finance, foreign certification, buyer discovery — that small firms cannot afford to staff. Agents do that work end-to-end at near-zero marginal cost, extending the export discipline test, the most powerful disciplining force in development, to the whole economy rather than a handful of champions.

<p>5 expert tasks</p> <p>Separate specialist jobs per first-time export shipment ENSI corpus</p>	<p>Largest gains</p> <p>Productivity uplift accrues to least-experienced workers NBER</p>	<p>1996 → now</p> <p>Customs admins already deploy AI for fraud & clearance IMF / OECD</p>
---	--	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is strategic trade openness on the country's own terms (priority 12) and the export discipline mechanism beneath it (priority 2): world demand is the impartial examiner a protected home market never is — Korea tied subsidized credit to export quotas firms could not fake or lobby away (World Bank/Stiglitz, *East Asian Miracle*). The bottleneck is bandwidth, not character: a small firm faces the same fixed wall of specialist transaction work as a large one, with no volume to amortize it — so it never reaches the referee at all.

WHAT IT MAKES POSSIBLE

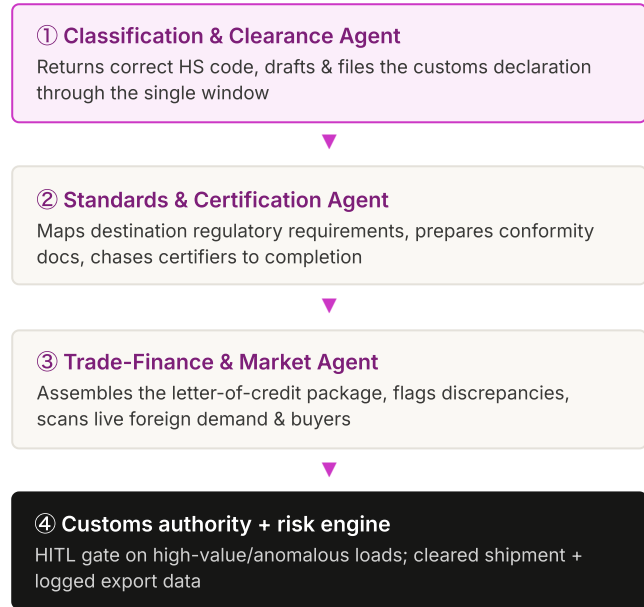
<p>+base</p> <p>Small, young, rural firms pulled into tradables</p>	<p>universal</p> <p>Discipline test applied to thousands of firms at once</p>
--	--

Cheap standards-compliance also makes the *climb* up the value chain cheap, and every cleared shipment yields the unfakeable customs data that instruments every subsidy.

HOW TO BUILD IT

- 1 Lay the rails first.** A national single window, digital customs interface, machine-readable tariff & standards schedules — agents need something to act through.
- 2 Start with classification & clearance, then finance and standards, then market intelligence last.** Highest-volume, most measurable step first; demand-scanning once narrow agents earn trust.
- 3 Discipline test.** Does it raise non-commodity exports, the number of exporting firms, and the small-firm share among them — customs-verified, on a clock, cut if not.
- 4 Guardrail.** HITL gates for high-value/anomalous consignments; full audit logs; agent outputs are recommendations subject to the customs risk engine, never unchecked authority.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Mauritius** — used export-processing zones to enter the global textiles chain on its own terms, then diversified into tourism, finance and ICT (Harvard/Frankel; ODI; UNU-WIDER).
- ▶ **Vietnam** — sequenced integration into global value chains, entering at assembly and climbing, anchored by Samsung (Govt. of Vietnam SEDS 2021–2030; CEPPI; ERIA).
- ▶ **Korea** — reciprocal credit contingent on export targets measured in customs data firms could not manipulate (World Bank/Stiglitz; ADB).

THE RISK (STEELMAN)

Trade barriers are often political, not informational: cheap paperwork does nothing about a tariff wall, a non-tariff barrier, or a customs officer whose income *is* the friction — and cheap exporting could lock firms at the lowest rung. The answer: bolt the programme to the discipline machine and the climb-the-chain intent, so it teaches with a test attached rather than running as a frictionless chute.

14 Agentic FDI Attraction & Investor Servicing

Great investment agencies — Ireland's IDA, Costa Rica's CINDE, Singapore's EDB — buy a country a shortcut through the slow accumulation of know-how, but they run on sustained expert attention few states can staff. An always-on agentic IDA manufactures that attention: targeting investors continuously, running the one-stop pipeline end-to-end, and matching local suppliers to force the spillovers that turn a foreign factory from a tenant into a teacher.

<p>Anchor</p> <p>One Intel plant (1996–97) catalysed Costa Rica's high-tech base World Bank</p>	<p>3,000 engineers</p> <p>Capability leaked per anchor decade = an economy, not just payroll ENSI corpus</p>	<p>200 suppliers</p> <p>Local firms to qualify to global standard for real spillover ENSI corpus</p>
--	---	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is FDI with spillover conditions — import capability, not just capital (priority 9): "the jobs are not the asset; the capability the jobs leave behind is the asset." The bottleneck is that the institution which buys the shortcut is built almost entirely from the scarce resource the country lacks — analysts who target, deal-makers who court for years, case officers running the one-stop pipeline, and the most-neglected supplier-development teams who qualify local firms. That attention-hungry work is exactly what gets cut, and exactly what agents are best at.

WHAT IT MAKES POSSIBLE

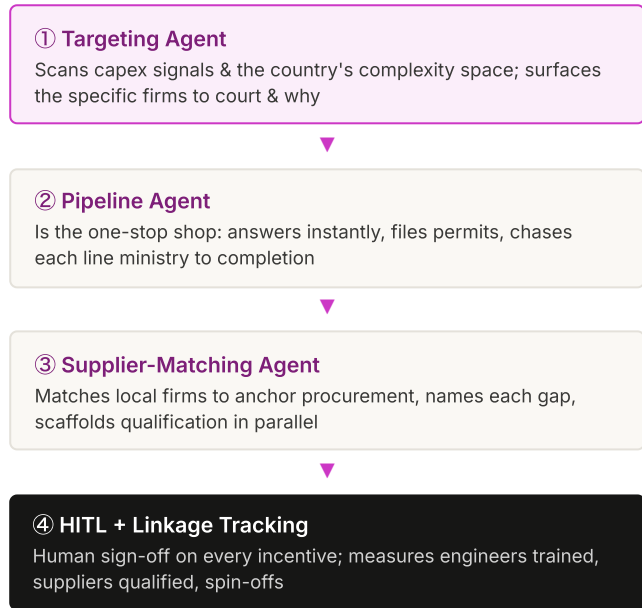
<p>overnight</p> <p>EDB-grade strategy unit for any country</p>	<p>at scale</p> <p>Supplier development run across hundreds of firms</p>
--	---

The agency shifts from chasing capital on the lowest tax rate to engineering measured capability — and the anchor strategy compounds without a staffing ceiling.

HOW TO BUILD IT

- 1 **Secure the principal, narrow the mandate.** A political patron protects autonomy; target two or three concrete activities, not "promote investment" — agents inherit the focus.
- 2 **Pipeline agent first, then targeting, then supplier matching before the anchor lands.** The spillover window opens the day the investor arrives, so qualify local firms in advance, not years late.
- 3 **Discipline test.** Judge on spillover, not splash: local-supplier content as a share of anchor procurement, locals promoted into technical roles, exports, spin-offs — on a clock, redesigned if it delivers payroll without capability.
- 4 **Guardrail.** HITL sign-off on every incentive & material concession; full audit logs; transparency on what each investor got and committed to; agents recommend, never unilaterally grant.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Costa Rica / CINDE** — landed Intel as an anchor in 1996–97; one plant catalysed a high-tech & medical-devices cluster (World Bank, *Impact of Intel*; Harvard CID WP 58).
- ▶ **Ireland / IDA** — targeted, courted and embedded Intel, pharma and tech; the cluster's Irish managers seeded a domestic ecosystem (Enterprise Ireland; OECD 2025).
- ▶ **Singapore / EDB & Vietnam** — EDB sustained selective capability-import for decades (Govt. Singapore CFE; ADBI); Samsung pulled suppliers into Vietnam's export base (ERIA).

THE RISK (STEELMAN)

The binding constraint is rarely processing capacity but fundamentals — power, ports, trusted courts; an always-on IDA marketing a country with unreliable electricity just loses deals faster when due diligence finds the gaps. And machine-scale supplier matching could degrade into box-ticking. The answer: the discipline test must measure outcomes — real exports from qualified suppliers, real locals in technical roles, real spin-offs — not the count of suppliers stamped

15 Agentic Skills & Labour-Market Engine

The workforce is the input every other growth lever consumes — yet schools teach the average student to a credential, guessing years ahead what the economy will need. The agentic reframe: AI tutors that teach each learner to demonstrated mastery, plus agents that read live employer demand, route learners into the exact gap-closing training, and broker the match — in every language, at near-zero marginal cost.

<p>39%</p> <p>of workers' skill sets transformed or outdated by 2030 WEF</p>	<p>2σ</p> <p>one-to-one tutoring lift over a classroom (Bloom) Bloom / ed. research</p>	<p>40%</p> <p>of global employment exposed to AI (60% in advanced economies) IMF</p>
---	--	---

THE GROWTH LEVER & THE BOTTLENECK

Human capital aligned to production (priority 7) — not credentials, but demonstrated capability that maps onto what firms actually hire for; the German–Swiss dual model is the alignment benchmark (OECD). The bottleneck is the price of one-to-one human attention: tutoring every learner to mastery and reading every employer's demand and brokering every match is caseworker work, and skilled human hours are scarce. So countries ration capable attention — and the rationing shows up as credential inflation beside unfilled vacancies.

WHAT IT MAKES POSSIBLE

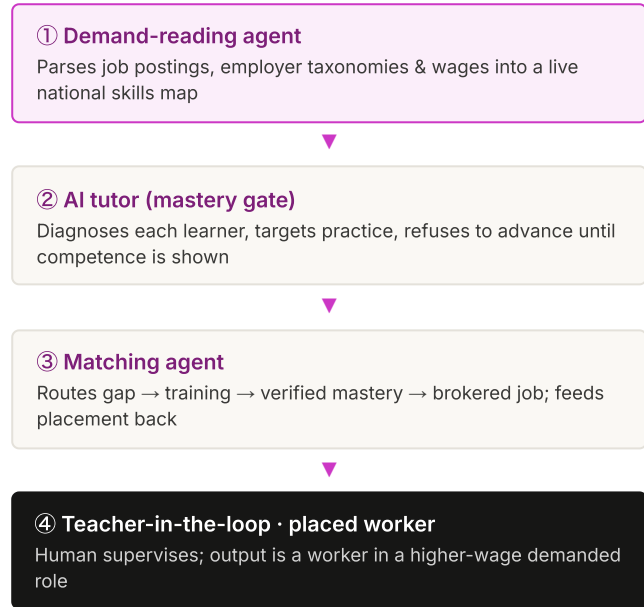
<p>63%</p> <p>of employers name skill gaps the top barrier to transformation (WEF)</p>	<p>11/100</p> <p>workers who need training are unlikely to get it (WEF)</p>
---	--

Mastery at population scale, a labour market with a live nervous system, and the dual system's alignment without Germany's century-old institutions — reaching the informal, rural, low-literacy cohort voice-first at the cost of a phone.

HOW TO BUILD IT

- 1 **Instrument demand first.** Build the live skills-demand map before tutors, so they teach toward signal, not a guess.
- 2 **Tutors to mastery, not seat-time.** Advancement gate is demonstrated competence, validated against the demand map (UNESCO: validate before deploy).
- 3 **Discipline test.** Score on placement & wage outcomes the system cannot fabricate — never enrolments or certificates issued; cut modules that do not move placements.
- 4 **Guardrail.** Human teacher in the loop; audit the demand map for bias; no opaque automated denial of opportunity; verify mastery against on-the-job performance.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Singapore** — SkillsFuture: national employer-linked, lifelong skills funding; the engine is this with tutoring + matching run continuously at the margin for free.
- ▶ **South Korea** — scaled engineering & technical education in step with its move into electronics and heavy industry; agents do that demand-tracking live, not in five-year plans.
- ▶ **ILO** — augmentation dominates: only 24% of clerical tasks highly exposed; in low-income countries 10.4% of jobs face augmentation vs 0.4% automation — a training opportunity.

THE RISK (STEELMAN)

The "GenAI divide" (UNESCO): tools reach the connected and literate first, and a gameable mastery test just issues a faster diploma that still means nothing. The answer is build voice-first/offline-tolerant, validate mastery externally against on-the-job performance, and audit the demand map — outcomes, not certificates.

16 Agentic Infrastructure & Energy Operations

Every other lever runs on two physical systems — the infrastructure that moves goods, power and data, and the energy that drives it. The binding constraint is rarely how much capital a country owns but how well it is operated: identical cranes turn ships in 18 hours or three days; identical wires lose 3% or 20%. The agentic reframe: instead of pouring concrete, put a fleet of agents inside existing assets to forecast, schedule, balance and pre-empt failure in real time.

<p>8th</p> <p>rank of hard infrastructure in the growth stack (energy 26th)</p> <p>Growth corpus</p>	<p>~1/3</p> <p>of treated water lost to leaks across typical pipe networks</p> <p>World Bank</p>	<p>decay</p> <p>neglected maintenance, not under-building, is the dominant failure</p> <p>Growth corpus</p>
---	--	--

THE GROWTH LEVER & THE BOTTLENECK

Hard infrastructure plus cheap, reliable energy (priorities 8 & 26) — but the corpus's own diagnosis is operation, not capacity: "an unmaintained asset is a stranded cost," and "a competitive currency and a slow port cancel out." The bottleneck is expert operational attention — continuous, real-time, quantitative control of coupled systems, done today by a few human controllers. There are never enough dispatchers, predictive-maintenance engineers or network planners; the rationing shows up as low throughput, high losses, and factories self-providing diesel backup.

WHAT IT MAKES POSSIBLE

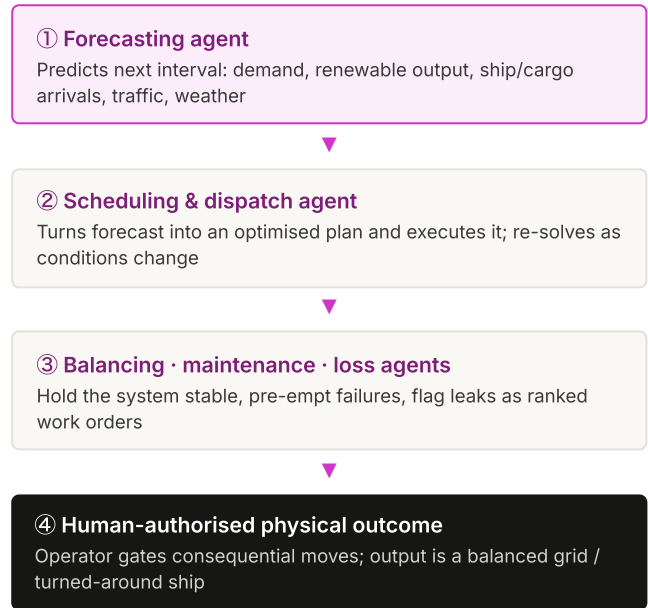


Reliability that ends the self-provision tax, recovered water/power/fuel, and a grid that can absorb intermittent renewables while keeping firmness.

HOW TO BUILD IT

- 1 **Instrument & observe first.** Build sensing/data, run agents in shadow mode; start with the grid, the most binding constraint.
- 2 **Automate the reversible, recommend the consequential.** Execute low-stakes moves; keep a human for load-shedding or taking a substation offline.
- 3 **Discipline test.** Tie every deployment to unfakeable physical output vs a pre-registered baseline — throughput, uptime, losses, landed cost; decommission what doesn't move the box or keep the lamp on.
- 4 **Guardrail.** HITL for safety-critical control; full audit logs & explainability; graceful degradation to manual fallback; secure the control stack as critical national infrastructure.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Singapore** — a leading container port with no domestic energy, competing on grid reliability and a transparent power market: reliability substitutes for endowment (IMF).
- ▶ **China** — "massive ahead-of-demand investment in generation and transmission" underwrote three decades of expansion (MGI); the return now depends on operating it better, not building twice.
- ▶ **Vietnam · Costa Rica** — power + port expansion sequenced behind FDI factories (ADB); near-total renewable electricity via state utility ICE — keeping a high-renewable grid firm is the agent's home turf.

THE RISK (STEELMAN)

Critical infrastructure is the worst place to let autonomous software act — a control error costs lives and can cascade, and adds a cyber-attack surface; some constraints are genuinely about building, not dispatch, and macro TFP gains may be modest (OECD, *Miracle or Myth?*; Acemoglu, NBER WP 32487). Answer: shadow mode, hard external bounds, human authorisation, audit logs, manual fallback, and diagnose whether operation or capacity binds before deploying.

LAYER IV · AREAS 17-21

Quality & Resilience

The human systems and shock-absorbers that keep growth fast, healthy and survivable — health, education, macro management, crisis response, and the digital rails the agents run on.

17

18

19

20

21

17 Agentic Healthcare System

A country does not grow on sick workers — health is productive infrastructure, not a cost. AI agents do not replace the clinician; they manufacture the scarce capability *around* the doctor — always-on triage, first-pass diagnosis where no specialist exists, keyboard liberation, and population surveillance no ministry can staff.

<h2>10M</h2> <p>Global health-worker shortfall by 2030 WHO</p>	<h2>72%</h2> <p>Medical assns: AI benefits outweigh risks OECD</p>	<h2>63%</h2> <p>Clinician emotional exhaustion (was 36%) OECD</p>
--	--	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is **health as productive infrastructure (priority 22)**: a population's health is an input to production exactly as a road or power grid is. The bottleneck is the doctor's *attention* — the most expensive, least scalable input in the system. It is throttled by absolute scarcity (WHO's 10M gap, OECD's 3M deficit), by misallocation (clinicians buried in documentation and coding), and by burnout that corrodes the attention already paid for: depersonalisation rose from 25% to 44% during COVID (OECD, 2024).

WHAT IT MAKES POSSIBLE

<h2>3M</h2> <p>OECD physician deficit the model relieves without a decade of training</p>	<h2>94%</h2> <p>assns worried on ethics — why guardrails are decisive (OECD)</p>
---	--

A competent first contact in every village without a clinic in every village — more healthy working years per citizen, bought cheaply (OECD, 2024).

HOW TO BUILD IT

- 1 **Back-office first.** Note-drafting, record completion, translation, scheduling — lowest risk, immediate workforce relief, builds trust (WHO lists these as most mature).
- 2 **Then judgment, last.** Triage with a clear human route; clinician decision support delivered *to* a human who signs off; ministry-run population surveillance.
- 3 **Discipline test.** Every agent measured against confirmed clinical outcomes — diagnostic accuracy, under-triage rate, error vs human baseline; publish the scorecard, cut what fails.
- 4 **Guardrail.** Human clinician is the accountable decider, always; informed consent, audit logs, built-in refusal/escalation, and a named liability owner (WHO six principles).

THE AGENTIC MECHANISM

① Triage Agent
Always-on, local-language symptom intake; routes reassurance / nurse / hospital-now

▼

② Diagnostic Co-pilot
Reads image + lab + history; first-pass opinion where no specialist is present

▼

③ Documentation + Surveillance Agent
Drafts notes (keyboard liberation); watches population for outbreak & AMR signals

▼

④ Accountable Clinician
Signs off every recommendation; named liability owner with consent & audit log

EVIDENCE & EXAMPLES

- ▶ **OECD survey (18 assns)** — 70% agree the physician's role stays central; *no* respondent believed AI would replace doctors.
- ▶ **WHO LMM guidance** — catalogues virtual health assistants, complex-case & routine-diagnosis support, language translation as near-term uses.
- ▶ **OECD** — AI already streamlining screening & diagnosis, advancing antimicrobial-resistance work and personalised monitoring.

THE RISK (STEELMAN)

Medicine is the worst place for probabilistic systems: automation bias, skill degradation, and models that "encode bias towards high-income countries" (WHO) could automate misdiagnosis at population scale — 74% of assns name weak data governance as the binding barrier (OECD). The answer is the discipline: back-office-first, outcome-tested, clinician-signed, fenced by the WHO principles.

18 Agentic Education System

Youth converts into growth only when turned into human capital — and that conversion has one ancient bottleneck: a good teacher's attention does not scale. AI agents are the first technology that dissolves it, delivering the one-to-one mastery teaching Bloom proved works (1984) at a cost that collapses toward zero, while teacher-copilots hand back the hours grading and admin consume.

<h2>2σ</h2> <p>Bloom: tutored beat 98% of classroom peers World Bank</p>	<h2>2\times</h2> <p>AI-tutored learned more, in less time (Harvard) World Bank</p>	<h2>31M+</h2> <p>LAC tertiary enrolment 2023 (7.4M in 1990) World Bank</p>
--	---	--

THE GROWTH LEVER & THE BOTTLENECK

This serves **human capital aligned to production (priority 7)** and the **demographic dividend (priority 19)** — the same lever from two angles. The dividend is conditional, and the condition is learning. The bottleneck is teaching attention, worse than the doctor shortage because demand is universal and simultaneous. One-to-one is the only proven intervention and the one a teacher-short, fast-massifying system cannot afford — the bottom income quintile is still just 24% of enrolment (World Bank, 2025).

WHAT IT MAKES POSSIBLE

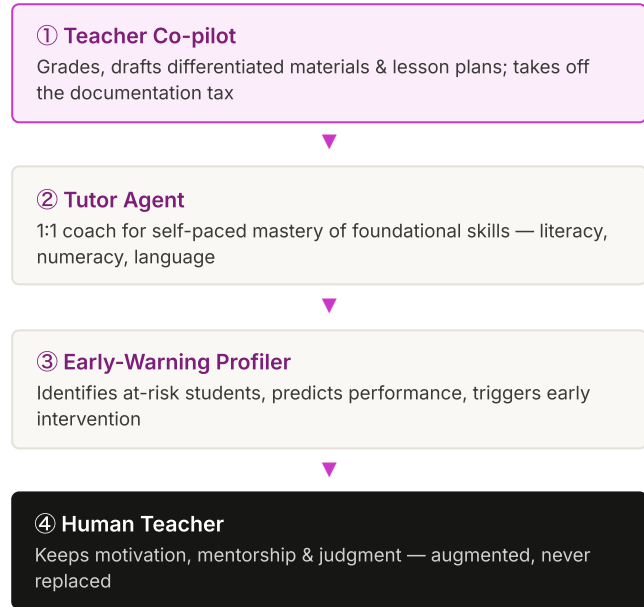
<h2>\$20</h2> <p>per tutor / yr to scale expert teaching — Stanford (World Bank)</p>	<h2>+20%</h2> <p>placement efficiency from AI assignment platforms (+38% for under-assigned)</p>
--	--

Bloom's 2 Sigma finally delivered to a whole population — reaching the rural, non-native and special-needs learners the system was failing.

HOW TO BUILD IT

- 1 **Copilot first.** Grading, materials, lesson drafting, at-risk flagging — lowest risk, immediate relief, and it puts teachers (not vendors) in control of the tech.
- 2 **Then tutors, foundational skills first.** Literacy, numeracy, language where mastery is measurable; higher-order & project-based support last, where the agent risks doing the thinking *for* the student.
- 3 **Discipline test.** Does it raise *measured learning* — not engagement or screen-time — without degrading the capability it builds? Controlled comparison vs the classroom baseline; publish; cut what flatlines (UNESCO).
- 4 **Guardrail.** Data-privacy and an age limit; protect human agency; inclusion, equity, linguistic diversity; teacher in the loop — UNESCO's human-centred floor.

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Harvard (Kestin et al., 2024)** — AI-tutored students learned more than twice as much in less time than active-learning peers.
- ▶ **Stanford (Wang et al., 2025)** — AI-enhanced tutoring scaled expert teaching practices at ~\$20 per tutor per year.
- ▶ **UNESCO GenAI guidance** — GenAI as a 1:1 coach for foundational skills, inquiry & project-based learning, and support for special needs.

THE RISK (STEELMAN)

An education agent can raise scores while destroying independent thinking: 83% of faculty worry students can't critically evaluate AI outputs (World Bank), and badly built tools "may inadvertently amplify existing educational inequities" via the digital divide. The answer: copilot-first, validated against measured learning, fenced by UNESCO's guardrails — and aimed deliberately at the bottom quintile and the rural learner.

19 Agentic Fiscal, Debt & Macro Management

Macro stability is the part of growth that is invisible until it fails. Agents are the first technology that can watch a dozen interacting prices — the exchange rate, debt path, inflation, reserves — continuously and without fatigue, flagging the slow drift that kills tradables while it is still cheap to fix.

<p>1/yr</p> <p>Cadence of typical debt-sustainability analysis (vs. nightly with agents) IMF Article IV</p>	<p>8economies</p> <p>East Asian high-performers whose defining trait was a persistently competitive rate World Bank</p>	<p>Cognitive tasks</p> <p>Where generative AI delivers its largest productivity gains IMF</p>
--	--	--

THE GROWTH LEVER & THE BOTTLENECK

A country grows on a competitive real exchange rate and the fiscal capacity that funds a capable state (Growth-Stack priorities 5 and 18). An overvalued currency is a tax on every export and a subsidy on every import — the most common, most invisible way promising economies kill their tradable sector before it learns. The constraint is not data but *attention*: scarce, senior judgment applied continuously across a dozen interacting variables. So debt analyses run yearly, exchange-rate reviews are episodic, and the slow drift lives precisely in the gap human vigilance cannot cover.

WHAT IT MAKES POSSIBLE

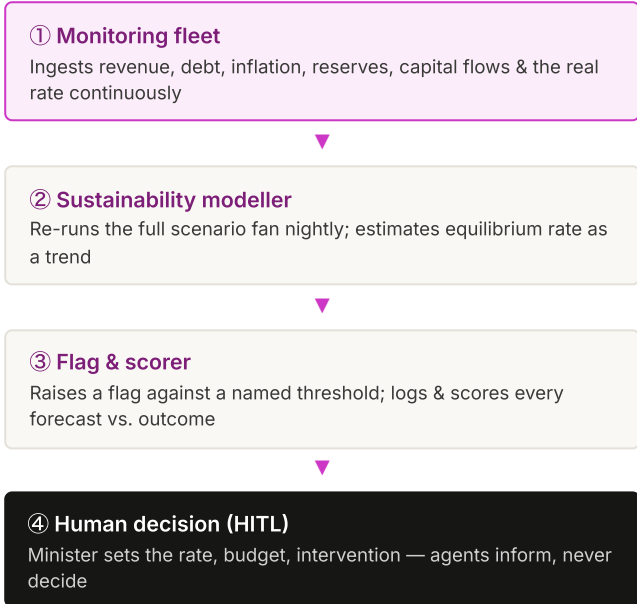
Q2	200bp
overvaluation flagged in quarter two, not year five	nightly rate-shock test on the debt path

A poor treasury gets the continuous structural-balance estimation a rich one buys with hundreds of economists — separating windfall from wallet in real time, and manufacturing credible commitment lenders can price.

HOW TO BUILD IT

- 1 **Instrument & observe.** Connect agents to existing feeds and have them reproduce official analyses — a mismatch is a bug, not a result.
- 2 **Add cadence & breadth.** Move validated runs from annual to nightly across the full interacting board, with named thresholds; then publish to earn credibility.
- 3 **Discipline test.** Not accuracy but whether the state *acts on the flag* — every flag, response and outcome logged, mirroring the corpus "did you export?" external test (priority 2).
- 4 **Guardrail.** HITL on every decision; full auditability of model version & inputs; contestable, traceable flags (EU AI Act; NIST AI RMF).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **Taiwan** — held money cheap and the rate competitive while thousands of small manufacturers built export capacity (World Bank).
- ▶ **Botswana** — managed diamond rents so the windfall never appreciated the pula into a tradables-killer (UNRISD).
- ▶ **IMF deployments** — AI in tax & customs already doing compliance nudging, audit-target selection and real-time fraud detection (IMF; OECD *Tax Admin 3.0*).

THE RISK (STEELMAN)

Agentic surveillance is only as good as its feeds; automating a thin, lagged statistical pipeline industrializes its errors, and nightly model runs can manufacture a precision equilibrium-rate economics does not support. The deepest failure is political — an agent cannot supply the will to tighten before an election. The guardrail: validate before you trust, score every forecast, keep humans deciding — the tool removes the excuse that no one saw it coming, not the choice itself.

20 Agentic Crisis, Risk & Resilience

Growth is a compounding process, and compounding is destroyed by interruption. Agents are the first technology that can hold the early-warning watch continuously, war-game the shock before it lands, and coordinate the response faster than a crisis cell can convene.

<p>4%</p> <p>An uninterrupted path beats a 6% path that resets — the magic is unbroken multiplication</p> <p>ENSI priority 28</p>	<p>50 cases</p> <p>Outbreak caught early vs. 5,000 — detection shifted left in time</p> <p>ENSI</p>	<p>Across agencies</p> <p>Core failure: fragmented info & slow inter-agency coordination</p> <p>kGovernment</p>
--	--	--

THE GROWTH LEVER & THE BOTTLENECK

Resilience and policy continuity (priority 28) is the meta-discipline that lets every other priority survive long enough to work — the largest source of wasted development is good policy abandoned before it matured. The damage is rarely the shock itself; it is the lag between arrival and a coherent answer. The binding constraint is capable attention, demanded continuously in calm and all-at-once in catastrophe — the two things human teams are structurally worst at: monotonous vigilance, and synthesis under a firehose.

WHAT IT MAKES POSSIBLE

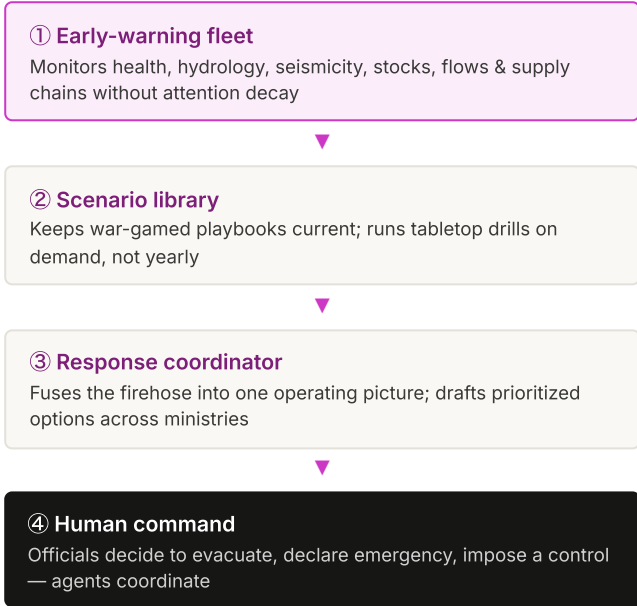
<p>Hours</p> <p>decisions a bureaucracy makes in months</p>	<p>1 qtr</p> <p>shock that costs a quarter, not a decade</p>
--	---

Lead time turns catastrophe into incident; a poor state gets the war-gamed readiness a rich one staffs with a permanent planning unit; and the strategy survives the shock instead of resetting to zero.

HOW TO BUILD IT

- 1 **Instrument the watch.** Connect early-warning agents to existing signal streams with named thresholds and a clear escalation path.
- 2 **Build & exercise.** Generate war-gamed scenarios, validate against historical events and the 24 comparators, drill relentlessly; wire coordination last, routed through human command.
- 3 **Discipline test.** The exercise, before the crisis — every flag scored against what followed, every drill an after-action record; an un-drilled library fails automatically (corpus priority 2).
- 4 **Guardrail.** Human command on consequential action; auditability through the fog; graceful degradation that never shows a confident picture on stale data (NIST AI RMF; arXiv oversight structures).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **World Bank** — survey of AI in the public sector locates clearest wins in continuous monitoring of large, messy flows (*AI in the Public Sector*).
- ▶ **Alan Turing Institute** — mapping of generative AI to government work finds the same early-warning sweet spot (*Mapping the Potential*).
- ▶ **WEF** — readiness framework for agents that plan, decide and act across organizational boundaries under oversight (*Making Agentic AI Work for Government*).

THE RISK (STEELMAN)

Garbage in, catastrophe out: a detector that cries wolf trains the state to ignore it; one that misses the real signal manufactures false security. And crises are exactly when power, connectivity and data feeds fail — a state that lets human capability atrophy because "the agents handle it" builds a brittle single point of failure. The answer: validated scored thresholds with human triage, graceful degradation, and relentless human-led drills — agents give lead time, the playbook and a coherent picture, not the judgment or courage a crisis demands.

21 Agentic Digital Public Infrastructure — the rails agents run on

Digital ID, instant payments and consented data-exchange are no longer the thing citizens click — they are the thing agents *call*. They are the substrate that decides whether every other agentic area is deployable at all: the road the whole fleet drives on.

<p>64 countries have DPI-like digital ID systems UCL IIPP 2025</p>	<p>103 countries have DPI-like data-exchange UCL IIPP 2025</p>	<p>1.73bn still on wrong side of digital divide UN DESA 2024</p>
---	---	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is **the digital state**: the capacity to act on citizens and firms accurately, instantly and at scale. DPI collapses the three transaction costs that tax every state–economy interaction — knowing who you deal with (identity), moving money (payment), finding a fact the state already holds (consented data-exchange, under the once-only principle). The binding constraint was never the decision; it was the clerical labour wrapped around it. Identity was re-verified at every counter, payment batch-processed over days, facts siloed across ministries — forcing the state to ration transactions and stay small relative to its economy (UCL IIPP 2025).

WHAT IT MAKES POSSIBLE

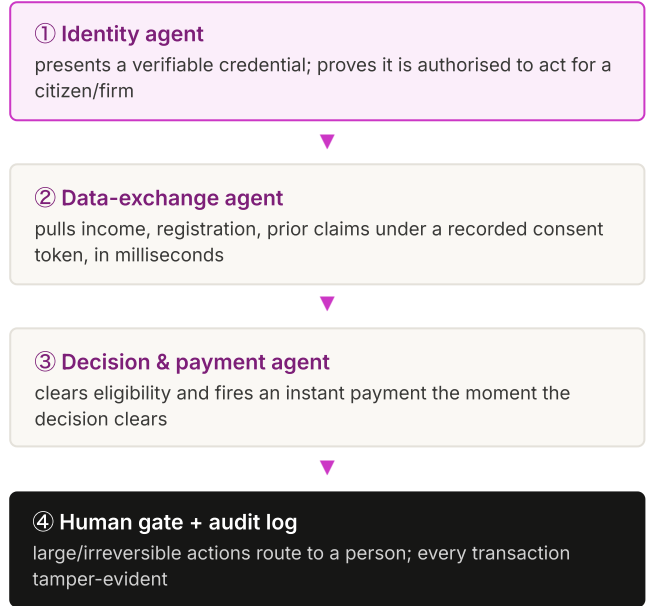
~0	secs
marginal cost of the millionth agentic transaction	authenticate, decide and pay end-to-end

Each remaining area becomes one agent composing the same three primitives in a different order — benefit delivery (Area 5), tax (Area 6), procurement (Area 7), FDI & trade (Areas 13–14) — so the report becomes a deployment schedule, not a wish-list (UCL IIPP 2025).

HOW TO BUILD IT

- 1 **Identity first.** Machine-verifiable through an open interface, not a plastic card — nothing else authenticates without it.
- 2 **Then payments, then data-exchange.** 97 countries have payment rails (Pix, UPI, CoDi); add consented exchange (X-Road, Trembita) under the once-only principle (UCL IIPP 2025).
- 3 **Discipline test.** Can a third-party agent authenticate, retrieve a verified fact and move value through documented open APIs, no human, in <10s — with that rate published monthly? Tie it to the World Bank GovTech Maturity Index (World Bank).
- 4 **Guardrail.** Recorded consent token + tamper-evident log per transaction; purpose limitation; revocable, privacy-preserving identity (80% of ID systems meet privacy variables today; target 100%) (UCL IIPP 2025).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **India** — India Stack (Aadhaar + UPI + consented data) turned account-opening, subsidy disbursement and merchant payment from multi-day paper processes into instant API calls (UCL IIPP 2025).
- ▶ **Estonia / Ukraine** — X-Road and Trembita wired registries into one secure data-exchange backbone, making the once-only principle the state's default behaviour (UCL IIPP 2025).
- ▶ **Brazil / Mexico** — Pix and CoDi put bank-grade instant settlement into the informal economy at near-zero fee (UCL IIPP 2025).

THE RISK (STEELMAN)

DPI is *necessary, not sufficient* — the IIPP's own Finding #6 is that it is unclear DPI correlates with performance, and one identity layer is a single point of failure and surveillance. The answer is the discipline test plus non-negotiable guardrails: instrument everything, publish the scorecard, keep identity revocable and route high-impact actions through a human gate (UCL IIPP 2025).

LAYER V · AREAS 22-24

Enablers & Guardrails

The substrate everything runs on — sovereign-enough capability — and the discipline that keeps it legitimate: governance, safety, trust, and an inclusion that keeps growth politically survivable.

22

23

24

22 Sovereign AI Capability — the substrate for every other area

National compute, governed data estates, talent and — only where it matters — sovereign models. The point is not a flag on the leaderboard; it is to be an "AI maker, not an AI taker" by controlling the substrate that carries the state's critical load, rather than renting capability and dependence (UK Govt).

<p>20x</p> <p>AI Research Resource compute expansion by 2030 UK Govt</p>	<p>15,000 people</p> <p>AI practitioner pool target Singapore Govt</p>	<p>2030 goal</p> <p>become a leading exporter of Data & AI SDAIA (Saudi)</p>
---	---	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is unusual: it is **the substrate for every other lever** — compute to run the fleet, data it learns from and acts on, talent to build and *supervise* it. The bottleneck is one level up: not scarce attention to do the work, but scarce capacity to *control the workers*. In the pre-sovereign world a country could only rent — model, compute and data sent abroad. That is fine for a chatbot and catastrophic for a state: rented capability can be repriced, rate-limited or cut off; data abroad is sovereignty surrendered; and a borrowed, opaque model cannot be audited, red-teamed or governed (OECD).

WHAT IT MAKES POSSIBLE

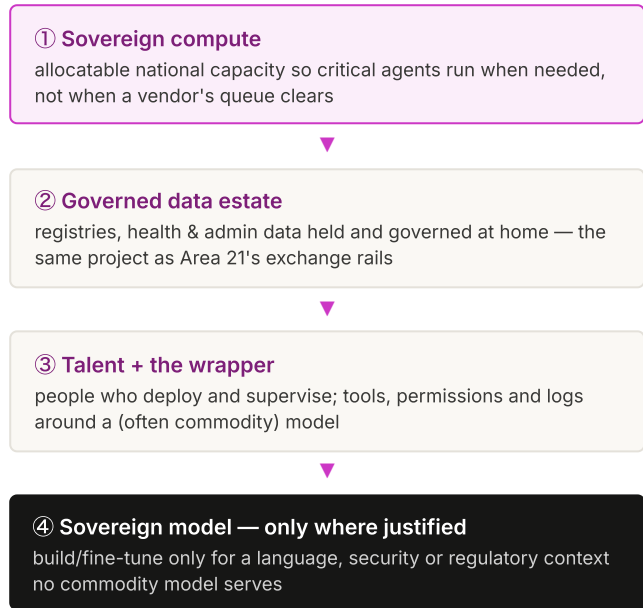
4	opt
diverse states converging on the same 3 pillars	optionality — the power to say no to any vendor

Every other area (tax, courts, health) runs on allocatable national compute over the state's own data, supervised by its own talent — auditable, dependable, and able to keep running even if a foreign supplier is cut off (UK Govt; Singapore Govt).

HOW TO BUILD IT

- 1 Compute first, allocatable.** Secure guaranteed-allocation compute — the UK's 20x AIRR and AI Growth Zones fast-tracking land and power are the template (UK Govt).
- 2 Govern data, then build talent.** Treat the public data estate as a strategic asset; train and anchor the people who supervise the fleet (Singapore Govt; SDAIA). Sovereign models last, only where the market fails.
- 3 Discipline test.** If our largest foreign supplier cut us off tomorrow, could critical agentic services keep running on infrastructure and data we control, staffed by people we employ? If no, it is rented, not sovereign — publish the dependency map and continuity plan (UK Govt).
- 4 Guardrail.** State models audited and red-teamed before deployment; compute and data carry access controls and logs; all of it operates inside Area 23's HITL governance so "the state controls the AI" never becomes "the AI runs the state" (OECD).

THE AGENTIC MECHANISM



EVIDENCE & EXAMPLES

- ▶ **United Kingdom** — AI Opportunities Action Plan: 20x sovereign compute, AI Growth Zones, data unlocking and a talent pipeline, framed as "AI maker, not AI taker" (UK Govt).
- ▶ **Singapore** — NAIS 2.0 organises the strategy around Compute, Data and Talent in a "Trusted Environment," with a 15,000-practitioner target and selective peaks of excellence, not autarky (Singapore Govt).
- ▶ **Saudi Arabia / India** — SDAIA puts data first and aims to export Data & AI by 2030; India's NITI Aayog strategy and the OECD overview show the same compute–data–talent shortages recurring (SDAIA; OECD).

THE RISK (STEELMAN)

"Sovereign" misread as "build a national frontier model" burns billions on a model obsolete on arrival and still dependent on foreign chips; advanced silicon is chokepointed by export controls, so full autarky is impossible. The defence is graduated sovereignty — control the substrate (compute, data, talent, supervision), diversify suppliers, build models only where the commodity market fails — paired with Area 23 governance, since sovereign capability without sovereign governance is more dangerous than dependence (OECD).

23 Agentic Governance, Safety & Public Trust

Every other area asks what a disciplined agent fleet can build; this one decides whether any of it is allowed to stand. The agentic move: manufacture oversight the way the rest of the report manufactures competence — watcher agents, consequence-sized human gates, tamper-evident logs and continuous evaluation against a published test — pointed at the state's own agents. Governance is not the brake on agentic growth; it is the licence to operate.

<p>193states adopted UNESCO's AI Ethics Recommendation — first global normative instrument UNESCO 2021</p>	<p>59rules AI regulations from US federal agencies in 2024, up from 25 in 2023 Stanford HAI 2025</p>	<p>1,889mentions of AI in legislative proceedings across 75 countries — >9x since 2016 Stanford HAI 2025</p>
---	---	--

THE GROWTH LEVER & THE BOTTLENECK

The lever is **legitimacy** — the precondition that lets every hard reform survive contact with politics. The OECD frames its AI standard around exactly this: to foster innovation *and trust* by promoting responsible stewardship of trustworthy AI (OECD/LEGAL/0449). The bottleneck is scarce, expensive human attention at its worst: oversight is *pure* attention, producing nothing but judgement on whether something else was done right. A poor state cannot staff a competent front line, let alone a second line of auditors to watch it — so review collapses into rubber-stamping. Worse, a tired human stapled to a fast agent's output succumbs to **automation bias**, which the EU AI Act (Reg. (EU) 2024/1689, Art. 14) names explicitly. Naive "human in the loop" can launder machine decisions, not check them.

WHAT IT MAKES POSSIBLE

<p>100% of the consequential fraction reviewed, not a 0.5% sample a year later</p>	<p>live discipline — contemporaneous, not retrospective</p>
--	---

Oversight finally keeps pace with action; every consequential decision is logged with its inputs, model version and approver — legible government, and reform that survives its own scandals because it can show what went wrong, who owned it, and what changed.

HOW TO BUILD IT

- 1 **Govern first, then classify.** Stand up the NIST RMF Govern function as the spine — who is accountable, what risks are tolerable, what the published test is — then map every deployment to an EU-AI-Act risk tier.
- 2 **Instrument before launch; gate, then loosen.** Logging, provenance and continuous evaluation are launch blockers, not roadmap items. Start consequential agents in human-approval mode and widen autonomy only as the measured record earns it.
- 3 **Discipline test.** Is it instrumented against an objective external standard, is the standard published, and are failures acted on — including shutting the agent down? When the scorecard goes red, an automatic consequence, not a meeting.
- 4 **Guardrail.** No silent autonomy on rights; independence of the watchers from the unit they police; provenance mandatory, and tamper-evident, an act with no reconstructable record should not have happened.

THE AGENTIC MECHANISM

① **Continuous Evaluator**
 Runs NIST AI RMF (Govern-Map-Measure-Manage); scores accuracy, error, disparate impact & drift against a published test, not at launch only

▼

② **Watcher Agent**
 Inspects the other agents; cross-checks outputs vs rules, samples by *risk* not at random, flags drift from the evaluated baseline (Area 9 logic, turned inward)

▼

③ **Provenance & Audit Log**
 Tamper-evident record of every consequential act — inputs, model & policy version, approver (AI Act Art. 12) — so conduct is reconstructable

▼

④ **Consequence-Sized Gate (HITL)**
 Rights-affecting acts halt for a named accountable officer to approve, disregard or stop (Art. 14); the regress ends in a person who can be fired

EVIDENCE & EXAMPLES

- ▶ **NIST** — the AI Risk Management Framework (AI RMF 1.0) supplies the operational loop, Govern-Map-Measure-Manage, and seven trustworthiness characteristics used as a scorecard.
- ▶ **EU** — the AI Act (Reg. 2024/1689) makes human oversight (Art. 14), event logging (Art. 12) and transparency binding obligations for high-risk systems.
- ▶ **UNESCO / OECD** — 193 states adopted the Ethics Recommendation (life-cycle accountability); 40+ governments signed the OECD principles. Four independent bodies converged on the same controls.

24 Agentic Inclusion & Broad-Based Distribution

Every other area makes competence cheap to manufacture; this one decides who gets it. Left to its gradient, an agentic economy concentrates — disrupting the rural, informal, non-dominant-language speaker and the woman before it reaches them. The agentic move: deliberately commission a fraction of the fleet as an *inclusion engine* — reskilling, rural reach, benefits take-up, language access — so AI raises the floor rather than tilting the table. The final area, because it keeps the other twenty-three politically survivable.

<p>40% of global employment exposed to AI (~60% in advanced economies) <small>IMF SDN/2024/001</small></p>	<p>4.7% vs 2.4% female vs male employment in the highest GenAI-exposure category <small>ILO WP 140, 2025</small></p>	<p>~4m women pulled into formal wage work by Bangladesh's garment sector <small>World Bank</small></p>
---	---	---

THE GROWTH LEVER & THE BOTTLENECK

The lever is **social cohesion** (priority 27) — reforms that hurt before they help only get the time they need if people trust the help will reach them too. Inclusion is the *precondition* of sustained growth, not its byproduct: a strategy that visibly enriches a narrow group is abandoned before it compounds. Its sharpest face is **female participation** (29) — "not a redistribution of the pie; an enlargement of the pie itself." The bottleneck is scarce attention at its most unevenly spread: reaching the people the formal economy already serves is cheap; the expensive last mile — a caseworker in dialect, an extension officer two valleys out — is what strapped states cut first. And the technology's natural gradient runs *against* inclusion: the IMF warns inequality rises where AI complements high-income workers and capital; UNESCO names a widening "GenAI divide."

WHAT IT MAKES POSSIBLE

<p>2x female participation "roughly doubles the effective dividend"</p>	<p>~0 marginal cost to serve the regions — enclave pattern broken</p>
--	--

The same women most exposed to clerical disruption become the prime beneficiaries of augmentation — threat and remedy land on one population; only sequence decides which arrives first. Regional balance shifts from costly redistribution to a near-free property of deployment, and reform survives its own disruption.

HOW TO BUILD IT

- 1 **Commission inclusion as an owned objective with a budget line.** Following the female-participation playbook — "publish a number with a deadline" — name take-up, regional and participation targets and ring-fence the fleet fraction.
- 2 **Point reskilling at the most-exposed first; lead with language & the informal sector.** Use the ILO exposure index as a targeting map; deploy cheapest, highest-reach speech-first agents first to unlock the largest excluded population.
- 3 **Discipline test.** The report's test, turned on distribution: is the engine instrumented against a published standard and are failures acted on? The standard is the *gap*, not the average — a rising headline with a widening gap has failed.
- 4 **Guardrail: Bias audit as a launch blocker** (else "the inclusion case inverts into an exclusion machine"); measure access at the hardest-to-serve margin; no quiet redirection of the ring-fenced inclusion budget.

THE AGENTIC MECHANISM

① Reskilling Tutor
Area 15 engine, re-pointed at most-exposed first; GenAI's largest gains fall on the least-skilled — routes the highest-return training to clerks, informal traders, non-graduates

▼

② Rural & Informal Reach Agent
Casework, advisory & extension at the same marginal cost in a secondary town as the centre — collapses the last-mile economics (Area 5, pointed outward)

▼

③ Take-Up & Language Agent
Proactively finds eligible people, explains entitlements in plain, any-language, speech-first form — turns passive opt-in benefits into active outreach

▼

④ Gap Instrumentation (HITL owner)
Adoption & outcomes broken down by gender, region, income & language; a named owner acts when the published gap-metric goes red

EVIDENCE & EXAMPLES

- ▶ **Bangladesh** — RMG pulled ~4m women into formal wage work; the World Bank credits it with the historic poverty decline — the multiplier on the demographic dividend, not charity.
- ▶ **Saudi Arabia** — Vision 2030's legal-first sequence (2018–21 reforms preceded the target) plus a published "22%→30%" female-participation number the state blew past.
- ▶ **Korea / Taiwan** — land reform *before* industrial takeoff built the cohesion to impose painful adjustment and stay the course for decades; Mauritius diversified on broad, rule-based institutions.

The five agent archetypes

Across all twenty-four areas the same five workers recur. Naming them turns the report from a list of opportunities into a build manual: a government does not procure twenty-four bespoke systems, it stands up five reusable agent *types*, each with a fixed job and a fixed human gate. Every area is some weighting of these five.

<p>5</p> <p>archetypes that compose all 24 areas</p> <p>ENSI synthesis</p>	<p>100%</p> <p>of archetypes carry a named human-in-the-loop gate</p> <p>ENSI / NIST AI RMF</p>	<p>24</p> <p>areas mapped to the same reusable fleet</p> <p>ENSI synthesis</p>
---	--	---

ARCHETYPE	WHAT IT DOES	WHERE IT APPEARS (AREAS)	HUMAN-IN-THE-LOOP GATE
① Analyst / Mapper read-only	Ingests data, builds the picture, costs options, finds the analogue. Turns raw streams into decision-ready briefs.	Pilot-agency scanning (01), bottleneck diagnosis (02), industrial mapping (11), market scans (13), evidence reviews (17).	Official interrogates the brief before it informs a decision; agent never decides.
② Monitor / Scorekeeper continuous	Watches a metric against an external test 24/7; flags drift, fires the cut/scale signal, keeps the public scorecard.	Discipline test (03), spend tracking (06), KPI dashboards (07), SLA monitoring (09), resilience signals (20).	Human owns the cut/scale decision; agent only surfaces the threshold breach.
③ Caseworker / Doer acts on records	Executes the transaction end-to-end — eligibility, processing, routing, fulfilment — at marginal cost ≈ zero.	Benefits & permits (05), service delivery (08), procurement ops (10), tax administration (06), registries (14).	Sampled + exception review; high-stakes or contested cases auto-escalate to a person.
④ Advisor / Drafter generates	Produces the artefact — memo, regulation, pitch, curriculum, legal text — to a reviewable first draft in minutes.	Policy drafting (12), investment promotion (13), skills & curricula (15), R&D translation (16), comms (08).	Named author signs; draft is a starting point, never published unreviewed.
⑤ Guardian / Auditor checks the others	Watches the watchers: provenance, bias, hallucination, capture; writes the immutable log; runs red-team evals.	Anti-corruption (04), audit & assurance (18), safety & QA (19), legitimacy (21), the guardrail layer (22–24).	Independent oversight body reviews findings; auditor cannot be silenced by the audited.

WHY FIVE AND NOT TWENTY-FOUR

Reusability is the economic argument. A bespoke system per area means twenty-four procurements, twenty-four failure modes, twenty-four oversight regimes. Five typed agents — each with one job, one data scope and one gate — collapse the surface area the state has to govern. The hard part was never building the agent; it was building the **discipline that makes its output trustworthy**, and discipline scales only if the patterns repeat.

THE ESCALATION LADDER

- ▶ **Read → write → decide** — autonomy rises left to right; gates tighten in lockstep.
- ▶ **Analyst & Monitor** never act on the world; safe to run broadly.
- ▶ **Caseworker & Advisor** touch citizens; sampled review + escalation are mandatory.
- ▶ **Guardian** is the precondition for trusting the other four at scale.

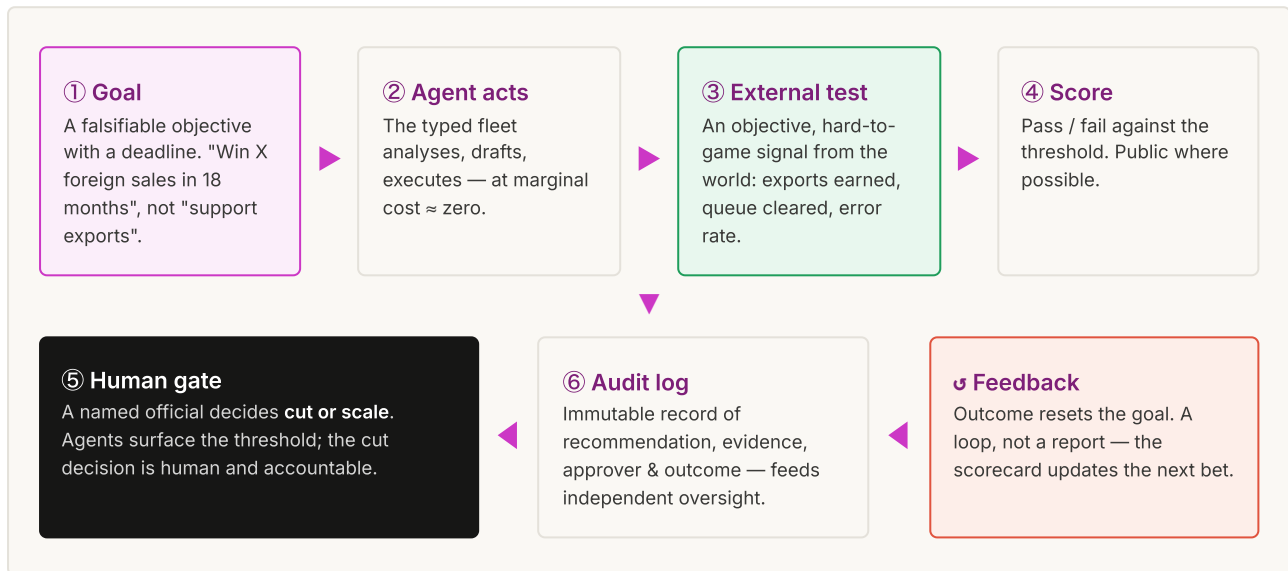
THE COMPOSITION RULE

Most areas chain the archetypes into a loop: **Analyst** frames → **Advisor** drafts → human decides → **Caseworker** executes → **Monitor** scores against the external test → **Guardian** logs it. The loop, not any single agent, is the unit of value.

The discipline & guardrails model

This is the spine of the whole report. An agent that makes competence cheap is dangerous unless every output is tied to an objective external test and a named human gate. The control loop below is identical in every area — only the goal, the test and the metric change.

FIG. - THE AGENTIC CONTROL LOOP



The discipline test (3④5) is the East-Asian export discipline made continuous and machine-monitored; the guardrail spine (6) is the NIST / EU AI Act layer that makes it auditable.

GUARDRAIL	WHAT IT REQUIRES	WHY IT MATTERS	STANDARD ANCHOR
Human-in-the-loop	A named person owns every consequential decision; agents recommend, humans decide.	Prevents automation bias and confident-wrongness at scale; keeps accountability locatable.	EU AI Act Art. 14; OECD AI Principle 1.3
Audit & provenance	Immutable, tamper-evident log of inputs, recommendation, evidence and approver.	Makes capture a documented choice; enables independent oversight and redress.	NIST AI RMF (GOVERN/MAP)
Evaluation	Pre-deployment + continuous evals against held-out tests; red-teaming for high-risk use.	Catches drift and silent failure before it reaches citizens.	NIST AI RMF (MEASURE)
Data quality	Consented, representative, documented data; known gaps surfaced, not hidden.	Garbage-in is the dominant real-world failure mode; bias enters here.	OECD AI Principle 1.4; EU AI Act Art. 10
Provenance / sourcing	Every claim traceable to a source; outputs labelled illustrative vs evidenced.	Stops hallucinated fact entering the record of state.	NIST AI RMF (MANAGE)

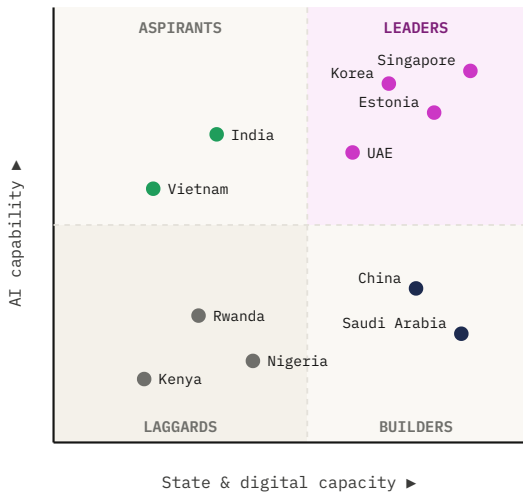
THE NON-NEGOTIABLE

An agent without an external test is just a faster way to be confidently wrong; an external test without a human gate is automation of the cut decision — and that is where legitimacy dies. The model only holds when both are present on every loop.

Country readiness landscape

Agentic government compounds two stocks a country already has or lacks: *state & digital capacity* (the rails) and *AI capability* (the engine). Plotting both reveals who can deploy this stack now, who must build rails first, and who risks importing dependency without sovereignty. Positions below are **illustrative**, triangulated from public indices.

FIG. - READINESS 2x2 (ILLUSTRATIVE)



Illustrative positioning triangulated from World Bank GovTech, UN E-Government, OECD and Stanford HAI indices — directional, not a ranking.

THE READ

Leaders (top-right) have both rails and engine — Singapore (NAIS 2.0), Estonia (X-Road), Korea, UAE — and can deploy the loop today. **Builders** have AI ambition but must deepen institutional rails. **Aspirants** like India hold strong digital public infrastructure (Aadhaar, UPI) and can leapfrog if AI capability follows. **Laggards** face the hardest path — but also the largest upside, since the binding constraint (scarce competent attention) bites hardest where it is scarcest.

THE LEAPFROG THESIS

Rails beat engine. A country with consented data-exchange and machine-readable registries can rent frontier AI capability; a country with frontier models and no rails has nowhere to plug them in. The policy implication: **build the rails first** (Areas 22–24), then layer the fleet.

~198
 economies scored on the World Bank GovTech Maturity Index
 World Bank, GTMI 2022

193
 member states ranked in the UN E-Government Survey
 UN DESA, EGDI 2024

2
 stocks that gate readiness — rails & engine; rails come first
 ENSI synthesis

WHAT TO DO AT EACH READINESS LEVEL

POSITION	FIRST MOVE	THE WATCH-OUT
Leaders	Deploy the full control loop now; scale the engine (areas 1–4) and administration (5–10).	Complacency — keep the discipline test live as throughput rises.
Builders	Deepen institutional rails (DPI, area 21) while standing up the guardrails (23).	Strong AI ambition outrunning weak rails → illegitimate pilots.
Aspirants	Leverage existing DPI (Aadhaar/UPI-class) to leapfrog; add AI capability (area 22) on top.	Rails without the engine — capability bought but not aimed.
Laggards	Build the rails first (areas 21–23); start with one high-volume administration win.	Largest upside but slowest path — sequence over breadth.

THE RULE
 Readiness is not a ranking to envy; it is a diagnosis of which layer to build next. The binding constraint bites hardest where capability is scarcest — so the largest upside is at the bottom of the table, not the top.

The build roadmap — Layers I–III

The actionable payoff. One move per area, with its owner, horizon, precondition, the discipline metric that proves it worked, and the guardrail that keeps it safe. Read top to bottom: nothing in a later layer is buildable without the layer above it.

AREA	MOVE	OWNER	HORIZON	PRECONDITION	DISCIPLINE METRIC	GUARDRAIL
LAYER I · AGENTIC STATE CAPACITY						
01 Pilot-agency	Stand up a standing agent fleet for scanning & option-drafting.	Head of govt	0–12 mo	Protected mandate; data rails.	Bets carry falsifiable forecasts; hit-rate tracked.	HITL + audit log on every call.
02 Bottleneck diagnosis	Continuous machine map of where the system actually jams.	Centre of govt	0–12 mo	Cross-ministry data access.	Diagnosed bottleneck cleared within window.	Provenance on every flagged cause.
03 Discipline test	Tie every state bet to an external pass/fail signal.	Finance min.	6–18 mo	Pilot-agency live (01).	Share of programmes with a live external test.	Cut decision is human, logged.
04 Anti-corruption	Guardian agents on procurement & spend anomalies.	Audit office	6–18 mo	Open registries; independence.	Recovered funds; flagged-case conviction rate.	Auditor independent of audited.
LAYER II · AGENTIC ADMINISTRATION						
05 Benefits & permits	Caseworker agents clear eligibility & processing end-to-end.	Service min.	12–24 mo	Digital identity; registries.	Cycle time & backlog cleared; appeal rate.	Sampled review; contested cases escalate.
06 Tax & spend	Monitor agents track collection, fraud & expenditure live.	Revenue min.	12–24 mo	Transaction data access.	Tax gap closed; false-positive rate.	HITL on enforcement; redress route.
07 KPI dashboards	Scorekeeper agents keep one live cross-govt scorecard.	Centre of govt	6–18 mo	Shared data definitions.	Decisions traceable to a tracked metric.	Public where possible; provenance.
08 Citizen service	Advisor + Caseworker agents on front-line channels.	Service min.	12–24 mo	Knowledge base; identity.	Resolution rate; satisfaction; error rate.	Human handover on demand.
09 SLA & ops	Monitor agents enforce service-level promises in real time.	Line ministries	12–24 mo	Instrumented services.	SLA breach rate; time-to-resolve.	Exceptions auto-escalate.
10 Procurement ops	Caseworker agents run sourcing; Guardian watches it.	Procurement off.	12–24 mo	Open contracting data.	Cost vs benchmark; competition rate.	Two-agent + human sign-off.
LAYER III · AGENTIC TRANSFORMATION MACHINERY						
11 Industrial mapping	Analyst agents map capabilities, gaps & adjacencies.	Industry min.	12–24 mo	Trade & firm data.	Identified bets that earn foreign sales.	Illustrative vs evidenced labelled.
12 Policy drafting	Drafter agents produce reviewable regulation drafts.	Line ministries	6–18 mo	Legal corpus access.	Time-to-draft; post-enactment effect.	Named human author signs.
13 Investment promotion	Advisor agents court & land specific investors (IDA model).	Investment agcy	12–36 mo	Pilot-agency live (01).	Foreign sales / FDI landed in window.	Export test; HITL on incentives.
14 Registries & rails	Caseworker agents maintain machine-readable registries.	Digital agcy	0–18 mo	Data-exchange standard.	Registry coverage & freshness.	Consent & access controls audited.
15 Skills & curricula	Drafter agents build & refresh demand-linked curricula.	Education min.	18–36 mo	Labour-market signals.	Placement rate of trained cohorts.	Educator review; bias checks.
16 R&D translation	Analyst agents move research toward commercial use.	Science min.	18–36 mo	Research & patent data.	Spin-outs / licences that reach market.	Provenance; IP guardrails.

HOW TO READ THE HORIZON

Horizons are **illustrative sequencing**, not promises: 0–12 mo = foundations, 12–24 mo = administration at scale,

The build roadmap — Layers IV–V

The second half completes the stack: quality & resilience make the output trustworthy, and the enablers & guardrails make the whole thing legitimate. These layers are not optional polish — they are what separates agentic government from automated capture.

AREA	MOVE	OWNER	HORIZON	PRECONDITION	DISCIPLINE METRIC	GUARDRAIL
LAYER IV · AGENTIC QUALITY & RESILIENCE						
17 Evidence review	Analyst agents synthesise evidence for every major decision.	Centre of govt	6–18 mo	Document corpus access.	Decisions citing a synthesised base.	Source traceability mandatory.
18 Audit & assurance	Guardian agents run continuous, not annual, assurance.	Audit office	12–24 mo	Immutable logs (22).	Findings closed; recurrence rate.	Independent oversight reviews.
19 Safety & QA	Guardian agents red-team services before & during use.	Digital agcy	12–24 mo	Eval harness.	Pre-launch eval pass rate; incidents.	NIST AI RMF MEASURE loop.
20 Resilience signals	Monitor agents watch for shocks & cascading failure.	Centre of govt	12–36 mo	Cross-sector data feeds.	Lead-time before a shock manifests.	HITL on response activation.
21 Legitimacy	Publish what agents do; contestable decisions & redress.	Head of govt	ongoing	Audit logs (22).	Trust & appeal-upheld rates.	Right to human review.
LAYER V · ENABLERS & GUARDRAILS (THE FOUNDATION)						
22 Data & audit rails	Lay consented data-exchange + immutable audit infrastructure.	Digital agcy	0–18 mo	Political backing.	Coverage of consented, logged exchange.	Privacy-by-design; access audited.
23 Talent & institutions	Build the human layer that judges, not does, the work.	Centre of govt	0–36 mo	Mandate & budget.	Officials trained to interrogate outputs.	Skills gate before autonomy rises.
24 Sovereignty	Govern dependency: portability, exit, domestic capability.	Head of govt	ongoing	Procurement leverage.	Vendor-switching cost; data portability.	No single-vendor lock on core rails.

THE SEQUENCING LAW

Layer V is drawn last but built **first**. Rails (22), people who can judge (23) and a dependency strategy (24) are preconditions for everything above. The most expensive mistake in the corpus is deploying caseworker agents on citizens before the audit rails and the human-review right exist — it converts an efficiency gain into a legitimacy crisis.

OWNER & BUDGET REALITY

- ▶ **Head of govt** owns the four moves that need protected autonomy (01, 21, 24).
- ▶ **Digital agency** carries the rails (14, 19, 22) — the spine everyone else plugs into.
- ▶ **Audit office** owns the Guardian moves (04, 18) and must stay independent.
- ▶ **Line ministries** deliver the citizen-facing wins (05, 08, 09) once rails exist.

THE PAYOFF LINE

Twenty-four moves, five reusable agents, one loop. A government that builds the rails, keeps the human gate and publishes the scorecard turns scarce competent attention from the binding constraint into an abundant one — which is the whole growth thesis, made operational.

Risks & failure modes

Making competence cheap is not free. The same fleet that manufactures capability can industrialise the worst tendencies of the state — error, capture and surveillance — at the same marginal cost \approx zero. Every risk below has a mitigation already in the model; the danger is deploying the agent and skipping the guardrail. The mitigations are not aspirational add-ons — they are the conditions under which the thesis holds at all.

RISK / FAILURE MODE	MITIGATION (ALREADY IN THE MODEL)
Automation bias — humans rubber-stamp agent output because it is fast and fluent.	Train officials to interrogate, not approve ; surface confidence + dissenting evidence; make the human gate a real decision with a logged rationale, not a click.
Confident wrongness at scale — a plausible error replicated across millions of cases before anyone notices.	External test on every loop; continuous evals (NIST MEASURE); canary + sampled review; kill-switch and rollback on threshold breach.
Capture — the fleet is pointed at protecting incumbents, not at the public interest.	Independent Guardian agents the audited cannot silence; immutable audit log; published scorecard turns protecting a zombie firm into a documented choice.
Surveillance & privacy — continuous monitoring slides into a panopticon.	Purpose limitation; consented, minimised data; privacy-by-design rails (22); independent data-protection oversight; transparency on what is monitored.
Dependency & sovereignty — core state functions locked to a single foreign vendor or model.	Portability and exit terms; domestic capability investment; no single-vendor lock on core rails (24); open standards for data-exchange.
Job displacement — administrative roles automated faster than people can transition.	Reframe humans from <i>doing</i> to <i>judging</i> ; fund reskilling (23); phased rollout; the scarce-attention thesis means redeployment, not pure replacement.
Data quality — biased, stale or unrepresentative data poisons every downstream decision.	Documented provenance; representativeness checks; known gaps surfaced not hidden; data-quality is a precondition gate, not a post-hoc fix.
Legitimacy — citizens reject decisions they cannot understand, contest or appeal.	Right to human review (21); explainable, contestable decisions; public reporting; redress routes designed in, not bolted on.

8

distinct failure modes, each with a model-native mitigation

0

that lack a guardrail — the danger is skipping it, not absence

\approx 0

marginal cost — of the harm as much as the good

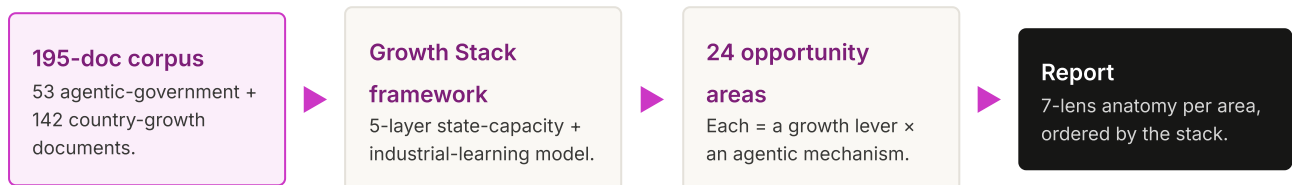
THE HONEST STEELMAN

The strongest objection is that guardrails are easy to write and easy to quietly drop under political pressure — and a captured state will drop them first. There is no purely technical fix for that. The defence is structural: **independent oversight**, an **immutable log** that makes the dropping visible, and a **named human** accountable at every gate. Agents do not supply political will. They do make its absence impossible to hide.

Methodology

How this report was built: a defined corpus, a fixed framework, a repeatable per-area anatomy, and an explicit ordering rule — with limitations stated plainly. The aim was a build manual, not a survey; every area had to survive the same test before it earned a page.

FIG. — FROM CORPUS TO REPORT



THE CORPUS

195 documents in two streams. The **agentic stream** (53) covers AI-in-government practice, standards and national strategies; the **growth stream** (142) covers what actually made countries rich — industrial policy, state capacity, development economics. The areas live at the intersection: a growth lever that an agent stream shows is now manufacturable.

DERIVING THE 24 AREAS

Each area had to clear three filters: it is grounded in a **real growth lever** from the growth stream; it has a **concrete agentic mechanism** from the agentic stream; and it survives a **discipline test** — an external pass/fail signal exists. Candidates without all three were cut. The survivors mapped cleanly onto the five-layer stack.

THE 7-LENS ANATOMY

- 1 **The lever & the bottleneck** — why it matters, what jams.
- 2 **What it makes possible** — the unlock once attention is cheap.
- 3 **The agentic mechanism** — which archetypes, in what loop.
- 4 **How to build it** — concrete first moves.
- 5 **The discipline test** — the external pass/fail signal.
- 6 **Evidence & examples** — real country precedents.
- 7 **The risk (steelman)** — the strongest objection, answered.

SCORING & ORDERING

Areas are ordered by the **stack's dependency logic**, not by impact score: State Capacity (I) first because nothing else is buildable without it, then Administration (II), Transformation (III), Quality & Resilience (IV), and Enablers & Guardrails (V) drawn last but built first. Within layers, areas run from foundational to derived. This ordering is the report's central argument made structural.

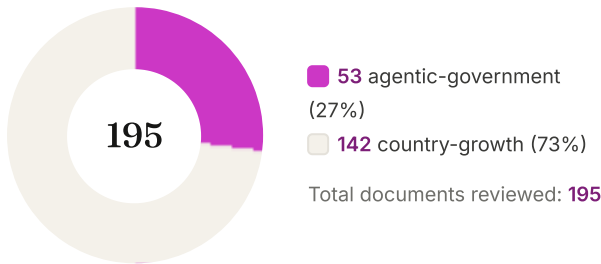
LIMITATIONS

This is a strategic synthesis, not an empirical study. Country positions and horizons are **illustrative** and directional. The corpus is curated, not exhaustive, and skews toward English-language and OECD sources. Agentic mechanisms describe what is technically feasible and disciplined — not a guarantee any government will keep the guardrails. Numbers are real and sourced where given; any figure without a hard citation is labelled illustrative.

The evidence base

The source library behind the report: 195 documents across two streams, ten agentic themes and twenty-four growth countries. Representative providers below; the report cites real, defensible sources only and never invents precise citations.

FIG. - CORPUS BY STREAM



Two streams. The agentic stream supplies the mechanism; the growth stream supplies the lever and the discipline.

10 AGENTIC THEMES

- State capacity
- Service delivery
- Public-sector AI
- Standards & risk
- Digital ID & DPI
- Anti-corruption
- National AI strategy
- Procurement
- Audit & assurance
- Skills & labour

AGENTIC STREAM – REPRESENTATIVE PROVIDERS

- ▶ **OECD** — AI Principles; government & public-sector AI surveys.
- ▶ **NIST** — AI Risk Management Framework (GOVERN/MAP/MEASURE/MANAGE).
- ▶ **EU** — AI Act (Arts. 10, 14; high-risk regime).
- ▶ **UK** — AI Opportunities Action Plan; Alan Turing Institute; Tony Blair Institute.
- ▶ **Singapore** — National AI Strategy (NAIS 2.0).
- ▶ **Saudi Arabia** — SDAIA national AI programme.
- ▶ **Estonia** — e-Estonia / X-Road data-exchange.
- ▶ **Stanford HAI** — AI Index; WEF; ILO; UNESCO; WHO.

GROWTH STREAM – REPRESENTATIVE PROVIDERS

- ▶ **World Bank** — East Asian Miracle; GovTech Maturity Index.
- ▶ **IMF** — fiscal & structural country surveys.
- ▶ **OECD** — Economic Surveys (Ireland, Estonia, Korea...).
- ▶ **NBER** — industrial policy & development economics.
- ▶ **ADB** — chaebol & East-Asian industrial policy.
- ▶ **UCL IIPP** — mission-oriented state capacity.
- ▶ **UN DESA** — E-Government Survey (EGDI).
- ▶ **ENSI** — Czech State malfunction analysis; the Growth Stack.

24 GROWTH COUNTRIES REFERENCED

- Singapore
- Korea
- Estonia
- Ireland
- UAE
- Saudi Arabia
- China
- Japan
- Taiwan
- India
- Vietnam
- Rwanda
- Kenya
- Nigeria
- Israel
- Finland
- Denmark
- Netherlands
- Poland
- Czechia
- Chile
- Costa Rica
- Botswana
- Mauritius

SOURCING DISCIPLINE

Where a hard citation exists it is named; where a figure is directional it is labelled **illustrative**. Precise quotes and page numbers are never fabricated — a claim either has a real source or is flagged as the authors' synthesis.

Agents make capability cheap; judgment, legitimacy and discipline make it growth.

Twenty-four agentic opportunity areas, five reusable agent archetypes, one disciplined control loop — the binding constraint on state capacity was never ideas, it was the number of competent hours. That constraint is now negotiable. Whether it becomes growth depends on the guardrails a country chooses to keep.

HOW AGENTS MAKE A COUNTRY GROW — 24 AGENTIC OPPORTUNITY AREAS. SYNTHESIS OF 195 DOCUMENTS (53 AGENTIC-GOVERNMENT + 142 COUNTRY-GROWTH). SOURCES: OECD · IMF · WORLD BANK · WEF · NIST AI RMF · EU AI ACT · STANFORD HAI · NBER · ILO · UNESCO · WHO · UK AI OPPORTUNITIES ACTION PLAN · SINGAPORE NAIS 2.0 · SDAIA · E-ESTONIA / X-ROAD.