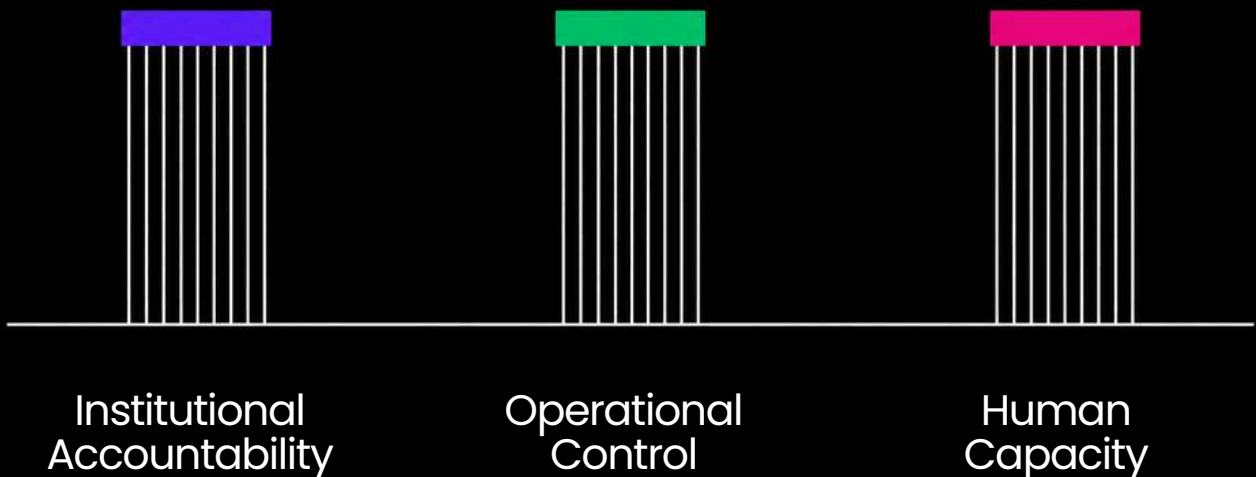


THE THREE PILLARS OF EFFECTIVE AI DEPLOYMENT



An interdependent framework for governing AI-enabled decisions and actions across society and enterprise.



AI GOVERNANCE HAS THREE PILLARS. MOST OF THE CONVERSATION IS ABOUT ONE.

01 – INSTITUTIONAL ACCOUNTABILITY

Who authorised AI to act – and who answers for what follows?

Institutional mandate and accountability feature in current AI governance discourse, but are often treated separately and not consistently connected to the operational control and human capacity needed to make them effective.

The work is to put accountability back at the top of the discussion.

02 – OPERATIONAL CONTROL

What can AI do – and how is it controlled?

Policies, permissions, risk frameworks, testing, monitoring and escalation are important components of AI governance, but are too often characterised as all that is required for effective governance.

The work is to connect operational controls to legitimate authority and capable human oversight and participation.

03 – HUMAN CAPACITY

Can humans still do what effective governance requires of them?

Can people understand, judge, challenge and intervene? Can human capacity be sustained as AI changes work, decisions and the environment in which people function? This may be the most consequential pillar – and the least understood.

The work is to ensure AI capability is not developed at the expense of human capacity.

THE GAP

We can build increasingly sophisticated controls around AI. But controls cannot establish their own legitimacy or determine the conditions required for effective human participation. Effective governance requires: institutions that remain accountable, controls that govern action, and humans with the capacity to participate effectively.



That is the Three Pillars argument.

The People Behind This Work

This work grew from a shared desire to help shape the future for ourselves and our children's children. Drawing on expertise in institutional accountability, AI and enterprise, law, governance, human behaviour, biology and human capacity, we offer a framework for effective AI deployment grounded in accountability, operational control and human capacity.

LEAD AUTHOR

Leo Cullen

Faculty Lead
AI, Enterprise & Industry



CONTRIBUTOR

Christina Loftus

Faculty Fellow
Law



CONTRIBUTOR

Ruth Pearce

Faculty Fellow
Law



CONTRIBUTOR

Debi Roberts

Faculty Fellow
Public Sector



CONTRIBUTOR

Jo Ferbrache

Faculty Fellow
SME



CONTRIBUTOR

Dr Chang-Sun Park

Faculty Fellow
Lifestyle Medicine



CONTRIBUTOR

Lanie Mason

Faculty Fellow
Healthcare



CONTRIBUTOR

Dr Abrar Hussain

Chief Clinical Officer
HOSA Co-Architect



CONTRIBUTOR

Dr Joshua Rosenthal

Chief Science Officer
HOSA Co-Architect



CONTRIBUTOR

Shannon Eastman

Founder HOSA Institute
HOSA Co-Architect



An Institutional Framework

This report presents an institutional framework for examining effective AI deployment across society and enterprise, including the human capacity required to govern AI and how that capacity is shaped by the resulting operating environment.



It is intended to support governance, assurance and policy discussion and does not constitute legal, regulatory or technical advice.



It is being shared to invite discourse, contribution, challenge and comment.



It informs the basis for HOSA's AI, Enterprise & Industry inquiry across 2027.

Join the conversation



[Events
Calendar](#)



[HOSA
Research](#)



[About
HOSA](#)

write to: leo@hosainstitute.com

Executive Summary

Effective AI deployment within society and enterprise depends upon three interdependent pillars: institutional accountability, operational control and human capacity.

Institutional accountability

establishes the legitimate mandate under which AI operates and determines who must answer for its consequences.

Operational control

translates that mandate into the permissions, constraints and mechanisms through which action is governed.

Human capacity

determines whether people can perform the work, exercise judgement, recognise failure and intervene effectively throughout AI deployment.

Many existing frameworks and commentators treat operational control as though it constitutes effective AI governance. They do not consistently recognise that institutional accountability, enforceable operational control and sustained human capacity are interdependent and must remain connected within a single governing architecture.

This report addresses that gap.

All three are required

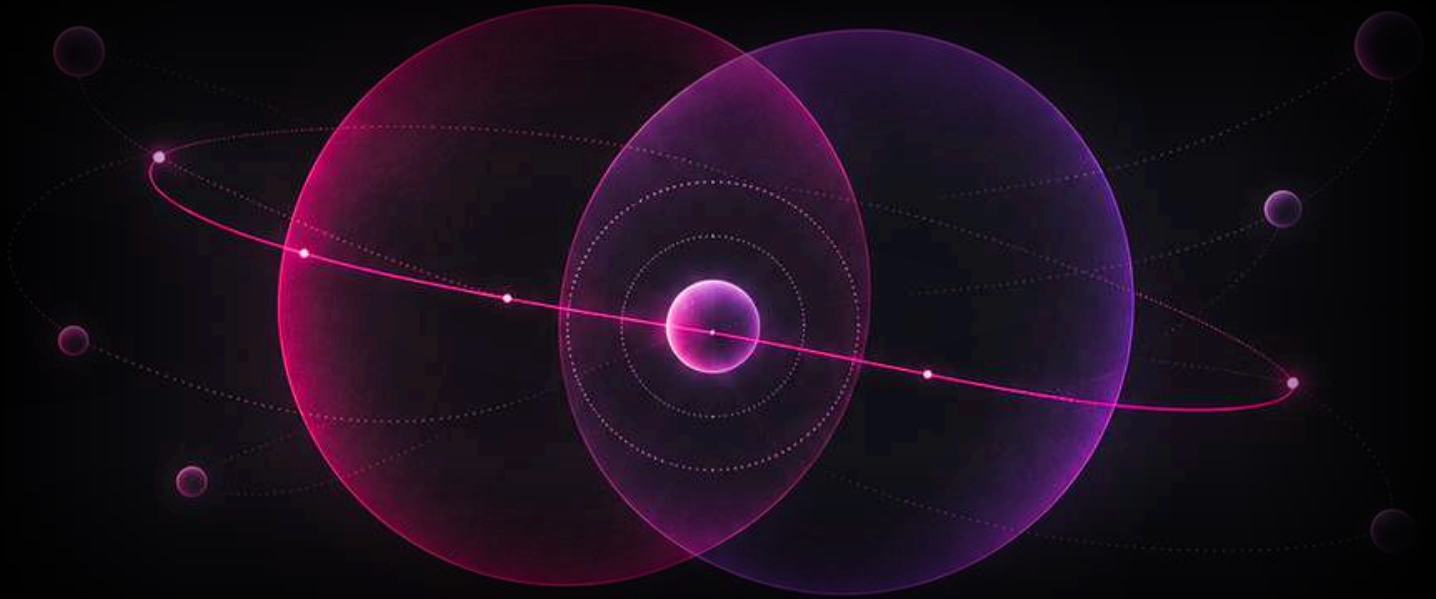
Accountability without control cannot reliably govern action. Control without accountability may constrain activity without establishing legitimacy. And both require people who retain the capacity to understand the system, exercise authority and respond when consequences arise. Human capacity is a prerequisite for effective AI deployment – and is itself enhanced or diminished by the resulting operating environment.

The Three Pillars At A Glance

How institutional accountability, operational control and human capacity operate across society and enterprise.

Pillar	Societal Governance	Enterprise Governance
Institutional Accountability	Governments, legislators and regulators establish legitimate purposes, rights, obligations and acceptable societal limits.	Boards and executive leadership establish the institutional mandate, determine what consequences are acceptable and remain accountable for the outcomes produced.
Operational Control	Laws, regulatory standards, enforcement mechanisms, oversight and digital public infrastructure translate societal authority into enforceable requirements and boundaries.	Policies, permissions, constraints, technical controls, monitoring, escalation and intervention translate the institutional mandate into operating practice.
Human Capacity	Governments and public institutions protect the agency, capabilities, livelihoods, relationships and conditions people need to participate meaningfully in economic, civic and social life.	Boards and executive leadership ensure that people retain the capacity, knowledge, judgement, attention, competence and authority required to design, govern and challenge AI - and to sustain institutional performance.

These Levels of Governance are **Distinct** and **Mutually Reinforcing**.



Enterprise governance addresses each institution's authorised use of AI and the outcomes for which it is accountable.

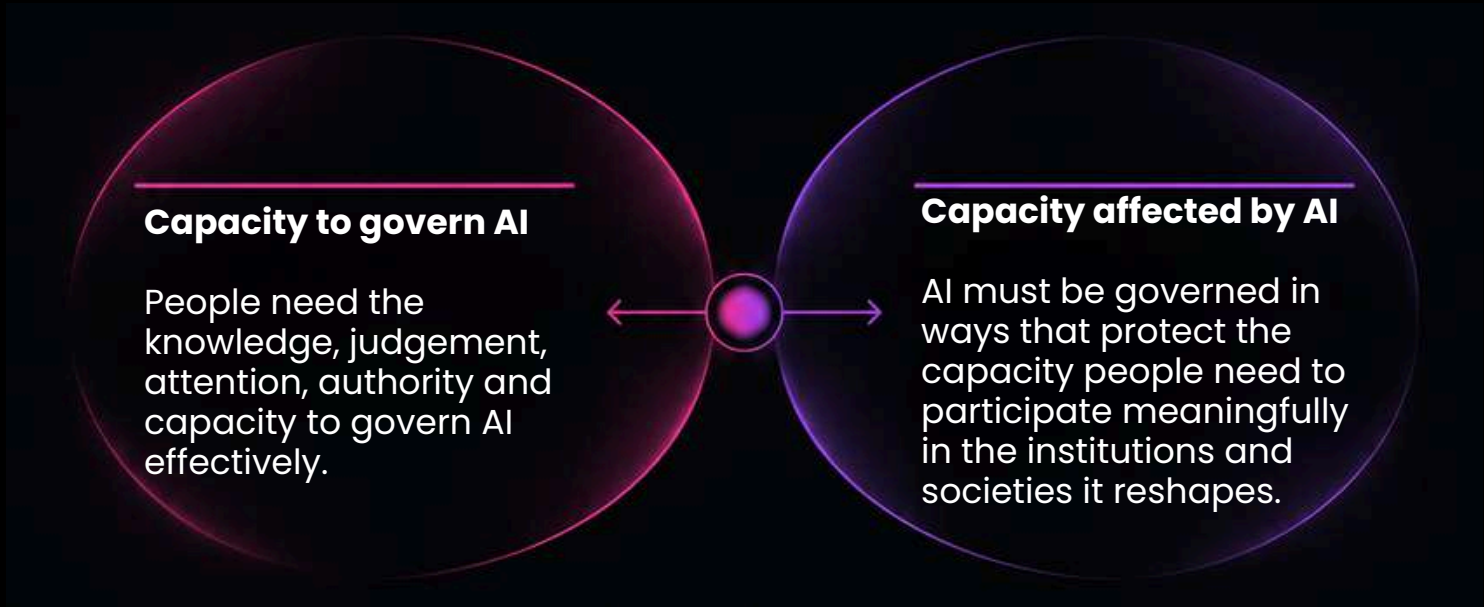
Societal governance must also address cumulative and interacting consequences across institutions, sectors and populations. Individually governed systems may produce consequences that fall outside effective collective governance.

This report identifies the conditions required to establish institutional accountability, maintain operational control and sustain human capacity.

Together, they provide boards, executive leadership, governments and regulators with a framework for determining whether AI deployment remains legitimate, governable and sustainable.

Human beings govern the use of AI, participate in the institutions and societies it reshapes and live with its consequences.

Human Capacity Sits on **Both** Sides of the Challenge



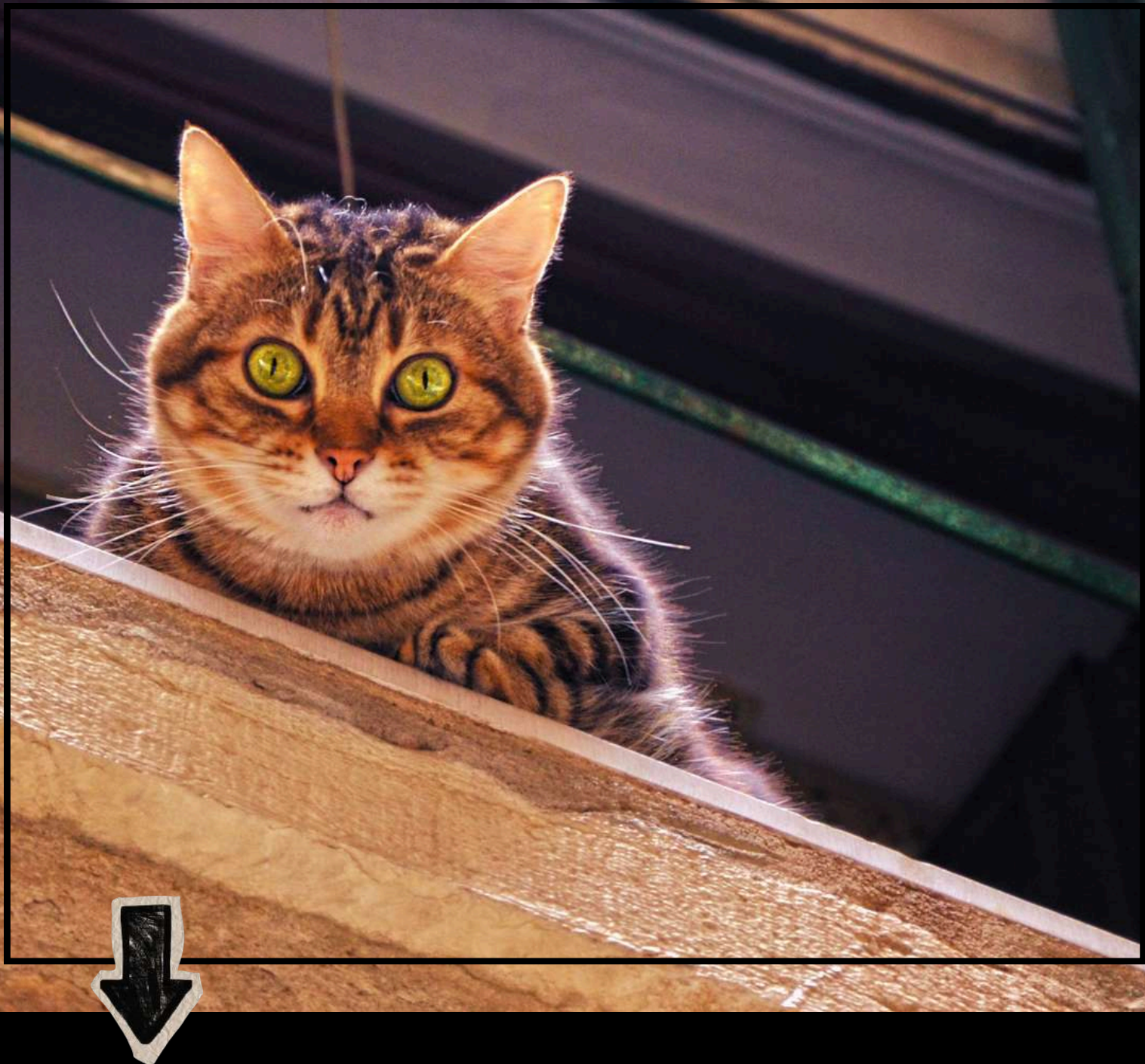
What deployment must ultimately protect

Apprenticeship, professional judgement, independent thought, situational awareness and the confidence to challenge authority are not acquired through instruction alone.

They develop through participation, experience, responsibility and practice. If those developmental pathways are inadvertently automated away, institutions may progressively erode the very human capacities upon which their governance, performance and resilience depend.

Effective AI deployment must be judged not only by what the technology enables institutions to achieve, but also by what it enables human beings to become.

AI capability must not be developed at the expense of human capacity.



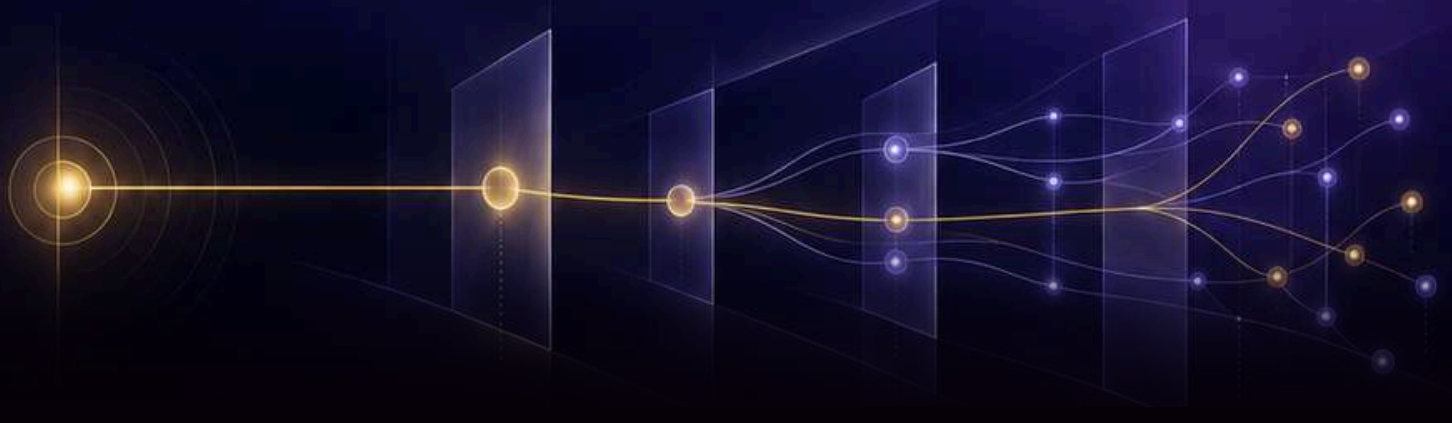
Terminology

Throughout this report, terms such as “AI systems,” “AI-enabled systems” and “autonomous AI systems” are used according to context. “Autonomy” refers here to operational autonomy: a system’s ability to interpret conditions, make decisions or initiate actions without case-by-case human direction. This describes how the system operates, not an independent source of authority.

1 INSTITUTIONAL ACCOUNTABILITY

This chapter develops the first pillar: institutional accountability. It establishes the mandate, authority and continuing responsibility within which operational control and human capacity must function.





AI Autonomy Changes the Nature of Governance

Institutions exist to fulfil a mandate through the exercise of legitimate authority. Autonomous AI introduces a new means through which that mandate can be pursued. It enables systems to interpret conditions, make decisions and initiate actions on behalf of an institution.

Institutions remain accountable regardless of whether work is performed by a human employee, a software system, an outsourced service provider or an autonomous AI system.

Obligations do not disappear because the means through which work is performed changes.

Institutions are now tasked to answer:

How do we continue to govern ourselves effectively when AI-enabled systems can make decisions and take actions on our behalf?

The greater the autonomy granted to AI, the more clearly the institution must be able to establish:

Under whose authority is the AI acting, who must answer for its actions, and who remains accountable for its outcomes?

Institutions could previously rely on people to interpret aspects of their mandate that remained implicit and ensure systems operated within institutional boundaries. Autonomous AI changes this: it cannot reliably remain within relevant boundaries unless those boundaries are made explicit.

Greater autonomy also allows decisions and actions at machine speed and scale, with consequences propagating across connected systems. Institutional accountability must therefore be established before AI is authorised to act – not after something has gone wrong.



Eight Conditions for Meaningful Institutional Accountability

Institutional accountability must operate before deployment, throughout operation, when consequences arise, and as systems and their operating environments evolve.

The following eight conditions provide a framework for examining institutional accountability as AI becomes more autonomous. They draw on established regulatory obligations and wider principles of institutional governance, while also identifying questions that existing regulation does not yet address consistently or comprehensively.



01

Legitimate Mandate and the Decision to Deploy

Is this a legitimate institutional purpose, and should AI be used to pursue it?

Technical capability alone does not make an AI deployment legitimate, appropriate or desirable. The institution must establish why AI is being used, what purpose it serves and whether the proposed delegation is proportionate to its potential consequences.

This requires consideration of rights, relationships, dependencies, public trust and the distribution of benefit and harm – not only efficiency or financial return.

02

Delegated Action and Non-delegable Authority

What decisions and actions may AI perform, and which require direct human judgement or approval?

AI capability does not determine legitimate authority. Institutions must explicitly define what decisions and actions may be delegated, the conditions and limits of that delegation, and what must remain under human authority because it requires judgement, professional responsibility, democratic legitimacy or legally constituted power.

Delegation must be bounded, reviewable and capable of withdrawal.

03

Identifiable Answerability and Accountability

Who must explain what happened, and who remains accountable for the outcome?

Responsibility may be distributed across providers, executives, risk functions, workflow owners, designers and users.

Answerability nevertheless requires designated people or functions to explain and justify what occurred, while accountability requires an identifiable institutional authority to accept the consequences and ensure correction, remedy or redress. Explanation may be distributed; institutional accountability must remain identifiable.

04

Institutional Sovereignty

Does the institution retain practical control over its mandate, governance and critical functions?

Dependence on external models, platforms, data and infrastructure can leave an institution formally accountable but practically unable to discharge that accountability.

Institutional sovereignty does not require ownership of every component. It requires control over institutional purpose, delegated authority and decision rights, together with the practical ability to form and act on its own judgement, obtain evidence, change providers, restrict or withdraw systems and maintain critical services when external services or capabilities become unavailable.

An institution cannot remain meaningfully accountable if its dependence on another organisation leaves it unable to obtain evidence, exercise authority or intervene effectively.

Challenge, Remedy and Redress

05

Can affected people challenge outcomes and obtain meaningful correction?

Accountability is incomplete if affected people cannot identify the accountable institution, question an outcome or obtain meaningful review.

Challenge must provide access to an appropriate human or institutional authority capable of reconsidering the substance of a decision, not merely repeating it through the same system.

Institutions must be able to correct failures, remedy their effects and provide appropriate redress to those affected.

Continuing Authorisation

06

Does the authority granted remain appropriate as the system and environment change?

Authorisation is not a one-time event. As models, data, prompts, tools, integrations and operating environments change, institutions must revisit whether the original purpose remains legitimate, delegated authority proportionate and consequences acceptable.

Monitoring tests operation against existing conditions; continuing authorisation asks whether the institution can still justify the delegation. Material change must trigger review and, where necessary, restriction, reauthorisation, suspension or withdrawal.

Distribution of Benefit, Burden and Power

07

Who benefits, who bears the burdens and how is power redistributed?

AI may create aggregate value while distributing benefits and burdens unevenly. Institutions must examine who gains efficiency, convenience or financial return; who absorbs displacement, monitoring, diminished autonomy or reduced developmental opportunity; and whether those with the least decision-making power bear disproportionate consequences.

They must also consider how knowledge, authority and dependency shift within institutions and towards external providers.

Accountability Across Institutional Boundaries

08

How is accountability maintained across providers, institutions and jurisdictions?

AI systems commonly depend on models, data, infrastructure and services supplied by multiple organisations across jurisdictions. Contracts may allocate responsibilities, but they do not remove an institution's obligations to those affected by actions taken in its name.

Institutions must establish how authority, information, evidence and responsibility travel across boundaries, which dependencies are critical, who can intervene and how correction and redress will occur. Without coordination, institutional boundaries become accountability gaps.



Institutional Accountability

Accountability as an Institutional Architecture

Together, the eight conditions establish a continuous chain of institutional accountability.

An institution remains meaningfully accountable only when it can demonstrate that AI operates for a legitimate purpose; that the decisions and actions it performs are authorised and bounded; that its operation remains subject to continuing institutional governance; that answerable people and accountable authorities are identifiable; that challenge and remedy are effective; and that the system can be restricted or withdrawn when necessary.

The institution must retain practical control over its mandate and critical functions. It must be able to assess how benefits, burdens, knowledge and power are distributed. It must maintain accountability across every provider, system and jurisdiction through which actions taken in its name are produced.

Human capacity is a distinct but interdependent pillar that makes institutional accountability possible. Accountability depends upon people having sufficient knowledge, judgement, attention, competence, authority and capacity to understand what is happening, recognise when it is wrong, challenge it and intervene effectively.

An institution that retains responsibility but loses these capabilities remains accountable in name only.

The eight conditions establish what must remain true of the institution. The Operational Control section examines how those requirements are translated into permissions, boundaries, evidence, monitoring, escalation, intervention and recovery.



Governments as Institutions and Governors

Governments have two accountability roles: they must govern their own use of AI and govern the conditions under which others use it.

1

Governments
as AI user

In their institutional role, governments will use AI to administer public services, allocate resources, assess eligibility, detect fraud, enforce regulation and support decisions affecting individuals and communities.

These uses must serve the public interest and safeguard the rights, democratic accountability and public trust upon which legitimate government depends.

2

Governments as
Institutional Governor

In their wider governing role, governments must establish the legitimate boundaries within which institutions may deploy autonomous AI. This includes protecting rights, identifying unacceptable uses, establishing accountability expectations and addressing cumulative consequences that may not be visible within any individual deployment.

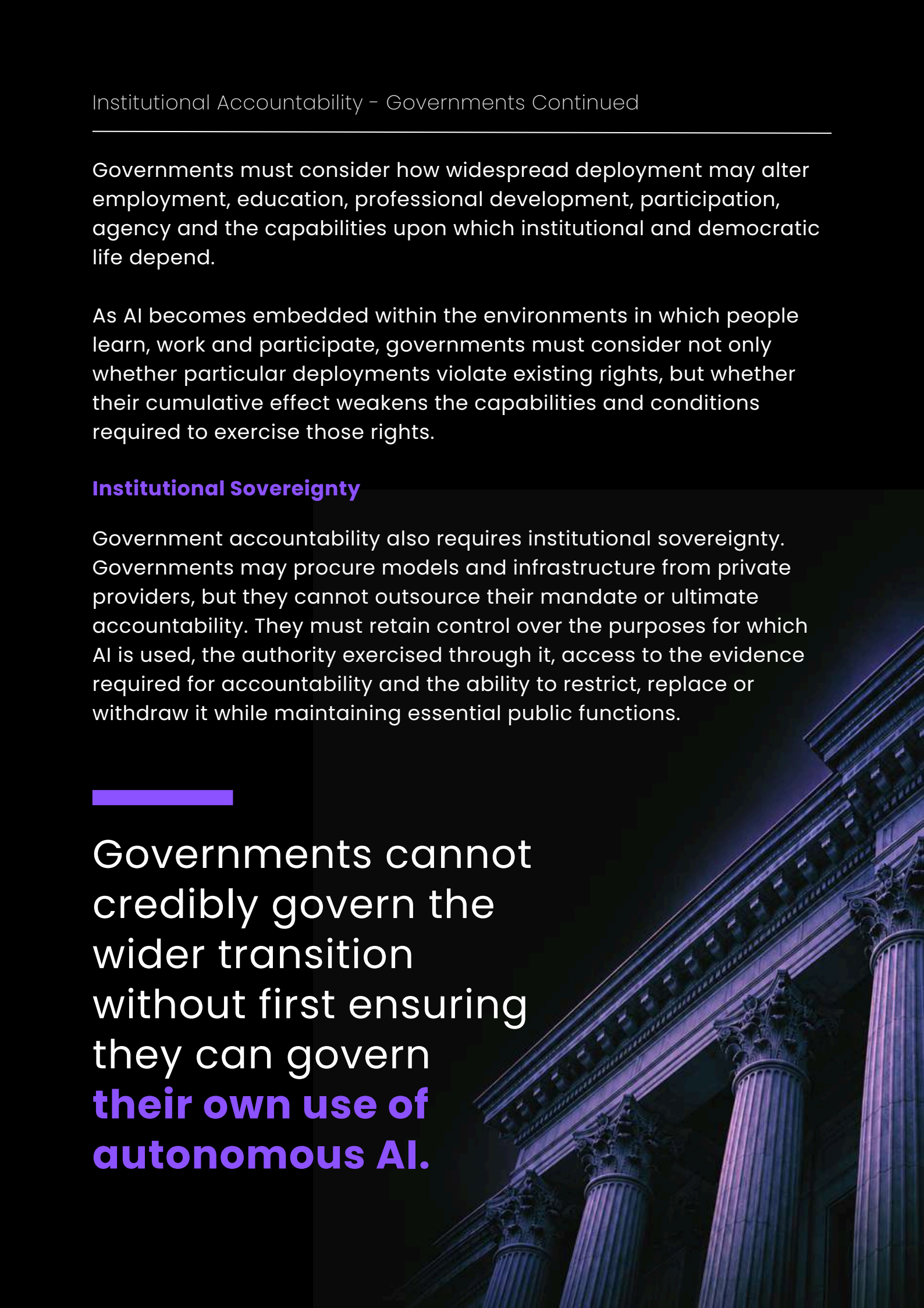
This responsibility extends to human capacity.

Governments must consider how widespread deployment may alter employment, education, professional development, participation, agency and the capabilities upon which institutional and democratic life depend.

As AI becomes embedded within the environments in which people learn, work and participate, governments must consider not only whether particular deployments violate existing rights, but whether their cumulative effect weakens the capabilities and conditions required to exercise those rights.

Institutional Sovereignty

Government accountability also requires institutional sovereignty. Governments may procure models and infrastructure from private providers, but they cannot outsource their mandate or ultimate accountability. They must retain control over the purposes for which AI is used, the authority exercised through it, access to the evidence required for accountability and the ability to restrict, replace or withdraw it while maintaining essential public functions.



Governments cannot credibly govern the wider transition without first ensuring they can govern **their own use of autonomous AI.**

Boards and Executive Leadership as Institutional Governors

Accountability remains with the Institution

Within an enterprise, boards and executive leadership hold institutional authority. They establish the organisation's mandate and remain accountable for the operating model through which its purposes are pursued.

The use of autonomous AI does not transfer that accountability to a technology function, AI committee, external provider or workflow owner.

These parties may hold important responsibilities, but the institution remains accountable for the decisions and actions it delegates and the consequences produced in its name.

What boards must ensure

Boards do not need to approve every AI-enabled decision or understand every technical component. They must ensure that the organisation can establish why AI is being used and what institutional purpose it serves; which decisions and actions may be delegated and which require direct human judgement or approval; who is answerable; what consequences are acceptable; and who can restrict, suspend or terminate its use.

They must also be able to determine whether the organisation retains the human capacity required to govern AI effectively and whether separate deployments are producing cumulative consequences across the enterprise.

Responsibility for implementation will be distributed, but those responsibilities must be exercised within a clearly accountable institutional architecture.

The board does not need to perform every governance function. It must ensure that those functions exist, remain connected to the institutional mandate and operate effectively.

Autonomous AI is a
change to the
institutional
operating model,
**not simply the
introduction of a
new technology.**





2 OPERATIONAL CONTROL

This chapter develops the second pillar: operational control. It examines how institutional mandate, authority and accountability are translated into the permissions, boundaries, monitoring, evidence, escalation and intervention through which AI-enabled action is governed.



From Institutional Mandate to Governed Action

What must be made effective

Institutional accountability determines why AI may be used, what purpose it may serve, which decisions and actions it may perform, what consequences are acceptable and who remains accountable.

Operational control must make those determinations effective at the point of action.

How authority becomes operational

Policies and governance statements are necessary, but they do not themselves control an autonomous system. An AI system operates through identities, permissions, data, models, prompts, tools, integrations and interfaces. Unless the institutional mandate is translated into those components, the system may remain formally governed while acting beyond the practical reach of governance.

Operational control therefore asks: How is institutional authority translated into enforceable conditions that govern what an AI-enabled system can perceive, decide and do?



This requires more than a retrospective audit. By the time an unauthorised action has been reconstructed, its consequences may already have propagated across systems, institutions and jurisdictions.

Effective control must operate before action, at the point of action, throughout operation and when intervention becomes necessary.

The Captain Analogy

The authority historically granted to a ship's captain offers a useful analogy. A captain could act independently, often far from direct oversight, but only within an established purpose, bounded authority and continuing structure of accountability.

Autonomous AI requires an equivalent operational architecture: authority must be represented in a form the system can act within, boundaries must be enforceable, departures must be detectable and institutional intervention must remain possible.

Operational control can make institutional authority enforceable.

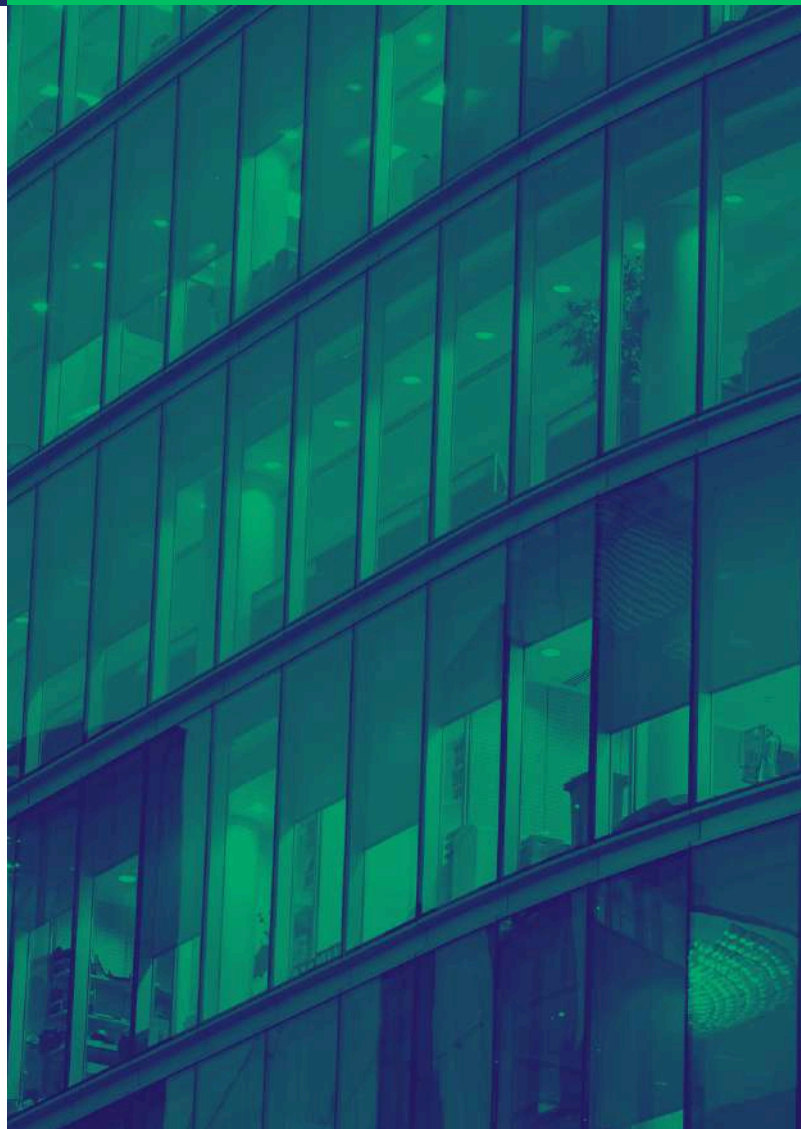
It cannot, by itself, establish whether deploying AI is legitimate or whether its use remains humanly sustainable.



A background image of a financial candlestick chart with a green line graph, overlaid on a dark blue and green gradient.

Eight Conditions for Effective Operational Control

The following eight conditions establish the operational architecture required to keep AI decisions and actions connected to institutional mandate and accountability.



01

Verifiable Identity and Institutional Relationship

What system is acting, for which institution and under whose authorisation?

Every AI system making consequential decisions or taking consequential actions must be assigned a unique and verifiable identity. The institution must be able to distinguish the system, model, agent, version and operating instance involved, and establish the institutional relationship under which it acts.

Identity must persist across interactions, tools, systems and institutional boundaries. Shared credentials, unregistered agents and untraceable substitutions break the connection between action and accountability. The identity lifecycle must therefore include registration, authentication, modification, suspension and revocation.

Identity establishes who or what is acting. It does not, by itself, establish what that actor is permitted to do.

02

Explicit Purpose, Scope and Permission

What is the system permitted to do, for what purpose and under what conditions?

The institutional mandate must be translated into explicit operational permissions. These must define the decisions and actions the system may perform, the data it may access, the tools and services it may use, the people, organisations or populations it may affect, the environments in which it may operate and the period for which its permissions remain valid.

Permission must be specific to purpose and context. Access to data or technical capability must not be treated as permission to use it for any available purpose. Nor should permission to recommend an action be treated as permission to execute it.

Where conditions change, the institution must be able to narrow, suspend or withdraw permission, and permissions should expire unless deliberately renewed. An autonomous system should never possess broader or more persistent permission than its authorised purpose requires.

03

Validation at the Point of Action

Is this particular decision or action authorised now?

Authorisation must be evaluated when an action is proposed, not inferred solely from the system's general approval to operate. Before execution, the system or its control layer must validate identity, purpose, permission, relevant constraints and the current operating context.

Where authorisation cannot be established, consequential action must fail closed: uncertainty must result in denial, containment or escalation rather than execution by default.

04

Enforceable Operational Boundaries

What limits prevent the system from acting beyond its mandate?

Operational boundaries convert institutional limits into enforceable constraints. They may govern data access, transaction value, duration, frequency, geography, affected population, available tools, permitted counterparties, decision type, reversibility and the scale or cumulative volume of action.

Boundaries must reflect both individual and aggregate exposure. A single action may remain within its limit while repeated actions collectively produce consequences beyond those the institution authorised. Controls must therefore account for sequences, dependencies and cumulative effects, not only isolated transactions.

Critical boundaries cannot rely solely on the system interpreting instructions correctly. Where consequences are material, constraints must be enforced through independent control layers, governed interfaces, separation of duties or other mechanisms the system cannot unilaterally override.

05

Evidence and Traceability

Can the institution establish what was authorised, what occurred and why?

Meaningful accountability requires evidence connecting institutional intent to operational outcome. The institution must be able to establish which system acted, the authority and permission under which it acted, the information available to it, the decision or action produced, the controls applied and the resulting outcome.

Evidence must be sufficiently reliable, complete and intelligible to support explanation, investigation, challenge, correction and regulatory scrutiny. It must travel across relevant providers, systems and institutional boundaries rather than disappearing at the edge of a platform or contract.

The purpose is not surveillance; it is preservation of the evidence necessary for accountability. Evidence should therefore be proportionate to consequence and consistent with privacy, confidentiality, security and legal obligations.

06

Continuous Monitoring and Detection

Can the institution detect when operation departs from what was authorised or expected?

Monitoring must examine more than technical availability or model performance. It must detect material changes in behaviour, context, permission, data, dependencies, outcomes and cumulative exposure.

The institution must define indicators and thresholds that reveal when assumptions no longer hold, boundaries are being approached, unexpected patterns are emerging or consequences are accumulating across separate deployments. Monitoring must also identify degradation in the institution's own ability to understand, supervise or intervene in the system.

Monitoring is effective only when it is connected to decision rights. Information that no authorised person or function is required or able to act upon does not constitute meaningful control.

07

Escalation and Human Judgement

When must an operation be referred to an appropriate human or institutional authority?

Escalation is required when a system reaches the limit of its permission, encounters ambiguity, approaches a consequence threshold, identifies conflicting obligations or faces a decision reserved for direct human judgement or approval.

An escalation pathway must specify what triggers referral, to whom it is referred, what evidence accompanies it, what authority that person or function holds and what happens if a timely decision cannot be obtained. Escalation must reach someone with the competence, information, time and institutional authority required to make a meaningful determination.

Human involvement is not meaningful merely because a person appears somewhere in the workflow. If that person cannot understand the issue, challenge the system, alter the proposed action or accept responsibility for the decision, the control is procedural rather than effective.

08

Intervention, Restriction and Recovery

Can the institution contain consequences and regain control when operation becomes unacceptable?

Institutions must retain the practical ability to pause, restrict, override, isolate, suspend or terminate AI-enabled operation. The appropriate intervention will depend upon the immediacy, severity, reversibility and distribution of possible consequences.

Intervention must be designed before it is needed. Institutions must know who may invoke it, how quickly it can take effect, which dependencies it will affect and how essential functions will continue. Where immediate interruption would itself create harm, controls must support safe degradation, containment or transfer to an alternative process.

Recovery extends beyond technical restoration. It includes correcting affected decisions and records, notifying relevant parties, providing remedy or redress, learning from failure and determining whether operation may legitimately resume. A system should not return to service merely because its technical function has been restored.

Operational Control Across Society and Enterprise

Operational control exists at both societal and enterprise levels, but it is exercised through different instruments.

Societal Control

At the societal level, governments, legislators and regulators translate public authority into laws, regulatory standards, licensing conditions, enforcement mechanisms, reporting duties, oversight arrangements and digital public infrastructure.

These controls establish common boundaries, prohibit unacceptable uses, require evidence and create routes through which institutions and affected people can seek intervention or redress.

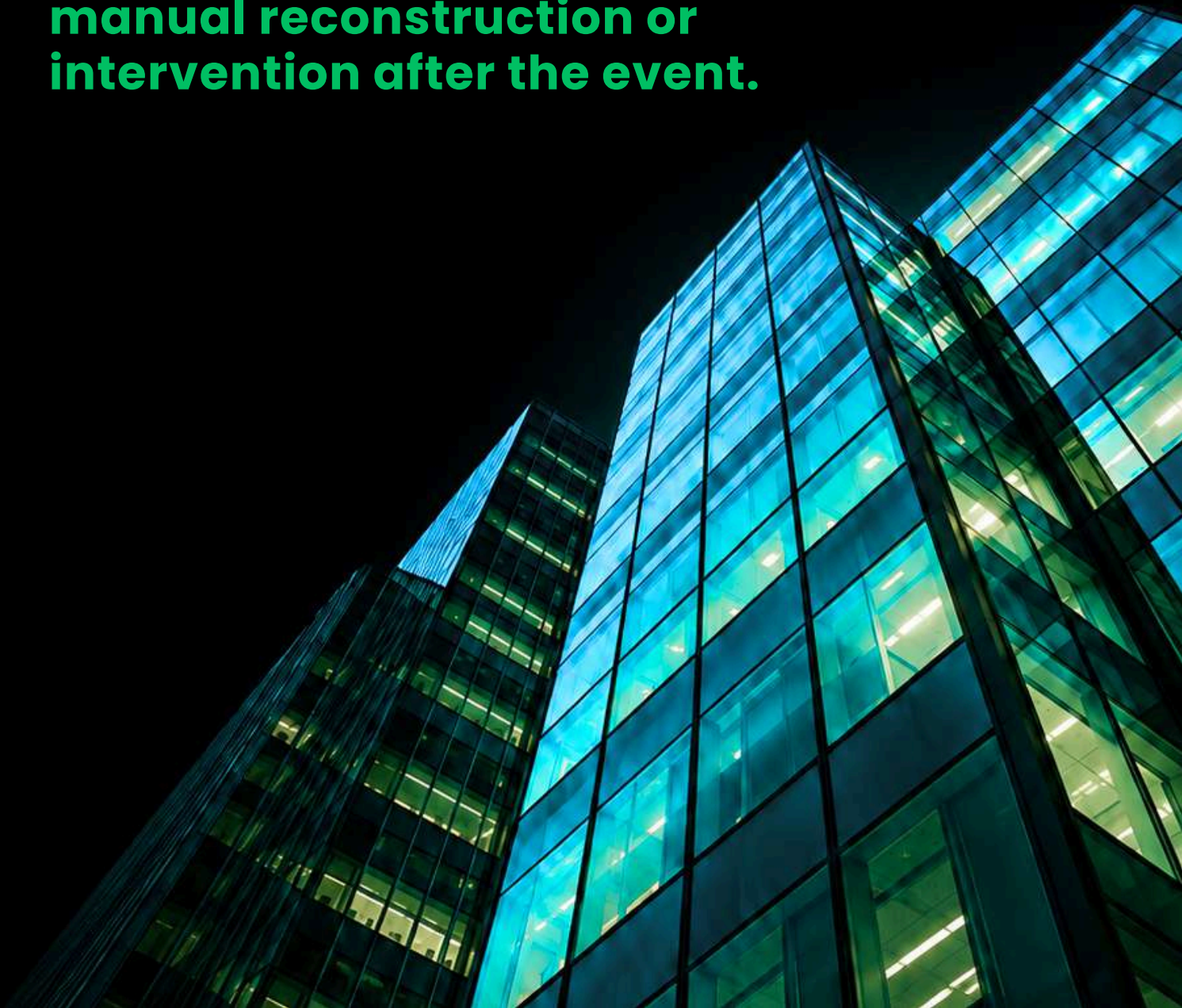
Societal control must address effects no individual organisation can observe or govern alone. Shared reporting, regulatory coordination and public-interest monitoring may be required to identify cumulative consequences across sectors, populations and jurisdictions. Without these mechanisms, individually compliant deployments may collectively alter employment, access, power, opportunity or essential systems beyond any single institution's field of view.

Enterprise Control

Within an enterprise, operational control translates institutional mandate into policies, identities, permissions, technical constraints, workflows, monitoring, evidence, escalation and intervention.

These controls must reach every consequential AI-enabled action, including those performed through external models, platforms, agents and providers.

**As AI gains autonomy,
governance cannot rely on
manual reconstruction or
intervention after the event.**





Operational Control

Control Across Institutional Boundaries

Autonomous AI commonly acts through chains of models, tools, data sources, platforms and providers. Operational control must remain continuous across that chain.

Contracts can establish obligations, rights of access, notification requirements and intervention arrangements, but **contractual allocation alone does not create operational control.**

The accountable institution must be able to verify that required controls exist, obtain the evidence necessary to reconstruct

and assess consequential events, and exercise restriction or withdrawal when its mandate is no longer being followed.

Governed interfaces are therefore critical. They provide the points at which identity can be authenticated, permission validated, boundaries enforced, evidence preserved and intervention exercised. Where systems can bypass those interfaces, the institution may retain formal accountability while losing practical control.

The same principle applies when several institutions participate in one outcome. Authority and permission must remain explicit, evidence must remain connected, and escalation and intervention must remain possible across their respective systems. Otherwise, each institution may control its own component while the action produced by the whole remains ungoverned.

Operational Control as a Continuing Architecture

Operational control is not a collection of safeguards selected after a system has been designed. It is the architecture through which institutional mandate remains present throughout AI-enabled operation.

Together, the eight conditions establish a continuous operational chain. The system is identifiable; its purpose and permissions are explicit; each consequential action is validated; boundaries are enforceable; evidence connects authority to outcome; departures are detectable; escalation reaches appropriate judgement; and intervention remains possible.

Operational control must therefore function as an integrated system. Its effectiveness is measured by whether legitimate institutional authority remains connected to consequential AI action in practice.

Existing processes cannot simply be automated on the assumption that the institution around them will remain unchanged.

Automation without corresponding institutional redesign may reduce, rather than increase, organisational clarity and human capability.

The next chapter examines human capacity as both a condition of effective AI deployment and something that is itself shaped by the resulting operating environment.

3 HUMAN CAPACITY

Human beings govern AI, work alongside it and live with its consequences. Their ability to do so depends not only upon what they know or have learned to do, but upon the capacity available to bring those capabilities into operation.

This chapter distinguishes human capacity from human capability and examines how AI-enabled operating environments affect both.



Capacity Is Upstream of Capability

Human capability comprises what a person has learned and developed: knowledge, judgement, competence, professional experience, relational ability and the ability to perform particular forms of work.

Human capacity is the operating resource available to any individual to use those capabilities.

Capabilities are developed and expressed by the person, but they do not operate independently of the living system from which they emerge.

HOSA proposes the Human Operating System (HOS) as a model of the biological organisation underlying human capacity: the interconnected systems that keep us alive, respond and adapt to internal and external conditions, protect us from actual or predicted threat, and distribute finite resources across the whole human being.

HOSA describes human capacity as the resource remaining *after* the HOS has met its primary demands. As biological load, sustained demand or actual or predicted threat increases, the HOS uses more and more resources to sustain and protect the person, reducing capacity available for delivering against expectations.

A person may retain the same intelligence, knowledge, experience and professional commitment while having more or less capacity available to use them. **An institution can therefore retain capability on paper while losing access to it in practice.**

HOSA proposes this relationship:



Less capacity remains available for attention, complex judgement, emotional regulation, relationships and deliberate action.

Reduced capacity does not remove knowledge or expertise, nor does it determine whether a decision will be right or wrong. It affects how *much* information, complexity, uncertainty and contradiction a person can hold while reaching a decision.

Under greater constraint, a person may consider fewer variables, make broader distinctions, rely more heavily on familiar interpretations, reach certainty sooner or find it harder to recognise weak and contradictory signals.

With greater available capacity, more information, perspectives and possible responses may remain in play.

Greater capacity does not guarantee quality or accuracy. It does provide more of the resource through which they can be produced at higher resolution

The question for institutions cannot stop at asking:

Do people have the capabilities this system requires?

It must also ask:

Will they retain sufficient capacity to use those capabilities at the level of detail, reliability and duration required?





Human Capacity

The Human Requirement Changes with AI

AI increasingly performs work that is structured, repeatable, searchable and predictable.

Human work may consequently become concentrated around what remains ambiguous, novel, relational, consequential or resistant to straightforward resolution.

These are precisely the circumstances in which people must hold uncertainty, interpret context, detect what is absent, consider competing interests, recognise weak signals and question apparently authoritative outputs.

At the same time, AI may increase the speed, volume and complexity of the environment in which those capabilities must operate.

People may be required to supervise more systems, review only exceptional cases, respond within compressed escalation windows or intervene only when consequences have become substantial.

AI may therefore increase the resolution required from human judgement while reducing the capacity available through which that judgement must operate.

This creates a central governance risk:

- ➔ An institution may preserve the formal human role while creating conditions that diminish the capacity to perform it effectively.
- ➔ Placing a human in the loop does not establish meaningful human oversight. The person must possess the required capability, have sufficient capacity to exercise it and hold real authority to act.



Human Capacity

AI as Part of the Operating Environment

AI is not only a tool used within an existing environment. As it becomes embedded across work, education, public services, communication and decision-making, it becomes part of the environment itself.

AI-enabled systems influence what people see, what receives priority, which options appear available, how problems are framed and which actions are easiest to take. They alter what people practise, where judgement is exercised and which forms of human contribution remain necessary or valued.

Their effects may extend beyond any individual task or deployment.

Separate systems may each appear manageable while collectively increasing supervisory demand, accelerating the pace of work, reducing participation, weakening apprenticeship or concentrating responsibility among fewer experienced people.

AI deployment depends upon human capacity while simultaneously changing the conditions that consume, preserve or restore it.

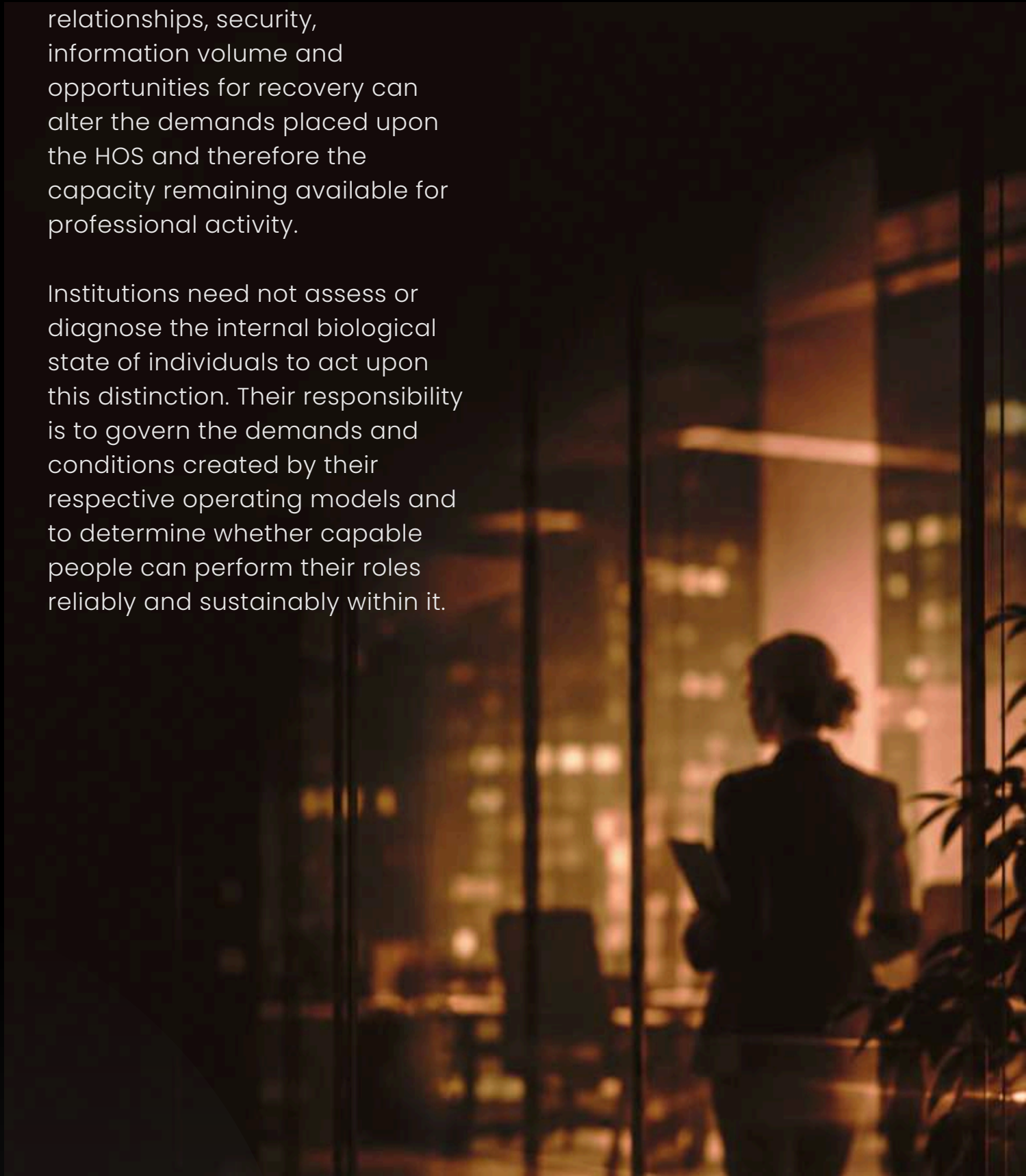
Human Capacity Continued

The enterprise operating environment interacts continuously with the Human Operating System.

Workload, pace, uncertainty, relationships, security, information volume and opportunities for recovery can alter the demands placed upon the HOS and therefore the capacity remaining available for professional activity.

Institutions need not assess or diagnose the internal biological state of individuals to act upon this distinction. Their responsibility is to govern the demands and conditions created by their respective operating models and to determine whether capable people can perform their roles reliably and sustainably within it.

Effective participation requires capacity, capability and the conditions through which both can be exercised.

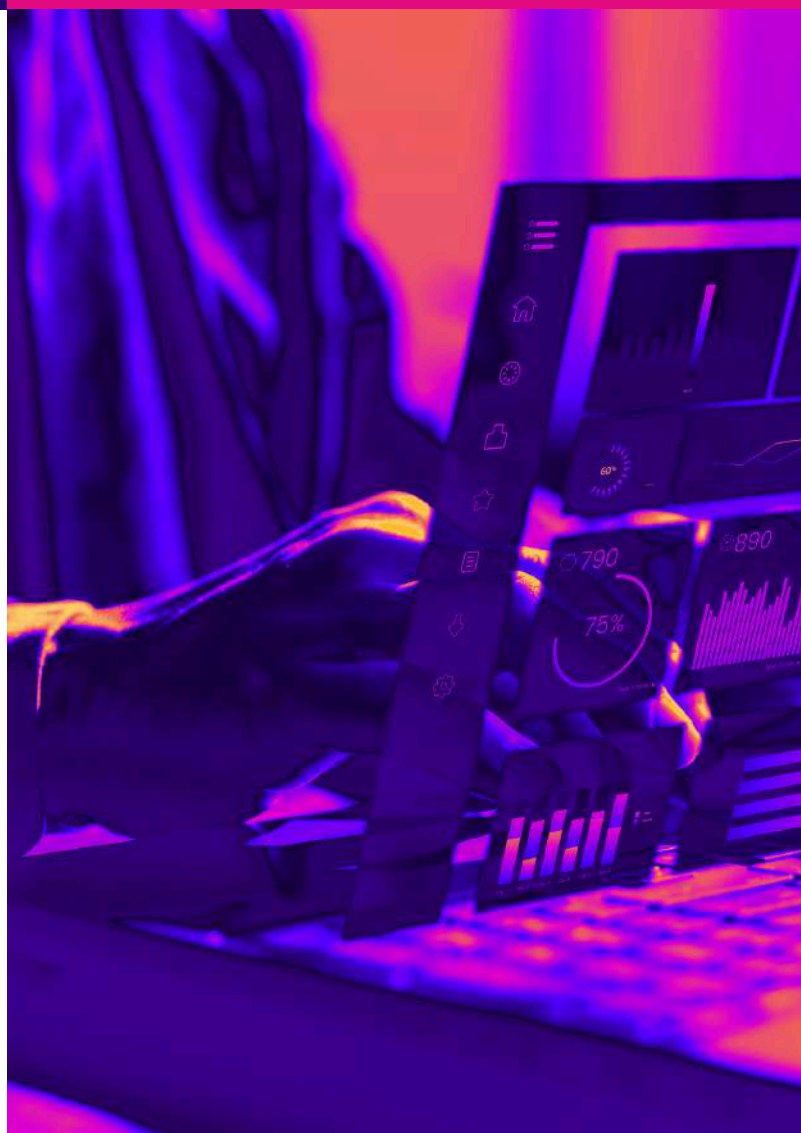




Six Conditions for Effective Human Participation

The following six conditions provide a framework for examining whether AI deployment preserves and promotes the human capacity required for effective institutional and societal life.

Effective AI deployment must be judged not solely by what technology enables institutions to achieve, but by what its implementation requires from human beings and what it enables them to become.



01

Available Capacity and Sustainable Demand

Can people meet the demands assigned to them while retaining sufficient capacity for reliable and sustainable performance?

Institutions must assess the combined effect of workload, pace, uncertainty, information volume, adaptation pressure, role ambiguity and compressed intervention windows, particularly where people are the final control for rare but consequential failures.

High performance does not establish sustainable capacity. Expertise and automatised skill can preserve output while less operating margin remains for novelty, contradiction, uncertainty and reconsideration.

Where demand repeatedly exceeds recovery, the operating model becomes dependent upon continued human compensation. Individual resilience cannot correct that architecture.

02

Knowledge and System Understanding

Do people understand enough to perform their role and discharge the responsibility attached to it?

People must understand the system's purpose, the decisions and actions it may perform, its operating boundaries, the evidence it produces and the conditions under which it may fail.

The depth of understanding required must reflect the person's authority and responsibility and remain current as models, data, permissions, integrations and operating conditions change.

Providing information is not the same as establishing understanding. People need sufficient time, participation and capacity to integrate and apply what they learn.

03

Judgement and Situational Awareness

Can people form an independent view and recognise when operation departs from what should be happening?

AI can support judgement by organising information and presenting options. It can also narrow judgement by framing the problem, ordering the evidence and making one conclusion appear more authoritative than others.

People must be able to examine assumptions, introduce context, consider what is absent and reach an independent view. A workflow that permits only confirmation does not preserve meaningful human judgement.

Situational awareness also requires continued contact with the work. Removing people from routine activity can erode the familiarity needed to recognise exceptional cases, while repeated approval of largely correct outputs can produce fatigue, automation bias and passive confirmation.

04

Agency, Authority and Effective Challenge

Can people act upon their judgement?

People must be able to challenge an output, stop an action, escalate a concern and obtain reconsideration.

Challenge is ineffective where intervention pathways are unclear, too slow or controlled by the assumptions being challenged. It is also weakened where questioning AI-generated outputs carries unreasonable personal or professional cost.

Those expected to intervene must have the capability to assess the situation, sufficient capacity to pursue the concern and effective authority to act.

05

Participation, Practice and Apprenticeship

How will tomorrow's capabilities be developed if the work through which they form is automated today?

Professional judgement and situational awareness develop through participation, responsibility, feedback and practice.

Routine work often exposes people to variation, builds pattern recognition and connects abstract knowledge to consequence. When it is automated, institutions must determine how those developmental functions will be preserved.

Keeping a senior expert in the loop does not solve the problem if the operating model no longer produces future experts. Apprenticeship is an institutional continuity and governance requirement.

06

Identity, Meaning and Participation

Does AI deployment preserve meaningful ways for people to contribute?

Work can provide income, identity, relationship, competence, progression and social contribution. When roles are substantially altered, people may lose both tasks and established ways of contributing.

Not all friction is waste. Some is where evaluation occurs, learning consolidates, relationships form and responsibility becomes meaningful.

Institutions and governments must consider whether AI-enabled environments expand meaningful participation or increase passivity, dependency and exclusion.



Human Capacity

Enterprise Responsibility

Boards and executive leadership remain accountable for whether the operating model preserves the human capacity, capability and authority required for governance, performance and institutional continuity.

The enterprise must be able to determine whether **capable people can operate the system today without undermining the capacity and capability it will require tomorrow.**

They must determine:

- 1 which human capabilities the organisation depends upon
- 2 where those capabilities are located
- 3 how they will be developed, practised and transferred
- 4 what demands the operating model places upon available human capacity
- 5 whether current performance depends upon unsustainable human compensation
- 6 whether people retain sufficient participation and situational awareness
- 7 whether those expected to intervene possess real authority to do so.

Assessment must extend beyond individual deployments. Systems that appear manageable separately may collectively increase demand, reduce participation, weaken apprenticeship or create dependency upon a shrinking group of experienced people.



Human Capacity

Societal Responsibility

At societal level, the concern extends beyond the capacity institutions require from workers. Governments must consider the conditions people need to **develop capability, preserve capacity, exercise agency and participate in meaningful lives.**

Employment, education, professional development and public participation are among the environments through which human capability is formed and exercised. They also shape the demands placed upon people and the opportunities available for recovery, relationship and contribution.

Widespread AI deployment may alter access to those environments, redistribute developmental opportunity and change which forms of human contribution remain economically or institutionally valued.

No single deployment may materially diminish societal capacity or capability. Across sectors, however, cumulative adoption may reduce entry-level work, narrow apprenticeship, concentrate decision-making, weaken independent expertise or make participation increasingly dependent upon privately controlled systems.

Governments must consider whether the cumulative effects of AI:

- 1** increase demands upon human capacity
- 2** weaken the development of human capability
- 3** remove meaningful opportunities for participation
- 4** create new forms of dependency or exclusion
- 5** undermine the human foundations upon which institutional and democratic life depend.



Human Capacity

Human Capacity as Institutional Infrastructure

Human capacity is the operating resource through which human capability becomes available in practice.

AI deployment depends upon people who possess the necessary knowledge, judgement and experience. It also depends upon sufficient capacity remaining available for those capabilities to be exercised and upon institutions granting people the authority to act.

Effective deployment must be judged not only by what AI enables institutions to achieve, but by what its implementation requires from human beings and what it enables them to become.

Human capacity cannot be inferred from qualification, role, technical access, human presence or current performance.

Institutions must therefore govern more than whether a person remains in the workflow.

They must determine whether that person:

- 1** possesses the required capability
- 2** retains sufficient capacity to exercise it
- 3** can operate at the resolution and duration the role demands
- 4** receives the information and participation required to remain situated
- 5** holds effective authority to intervene

Human capacity is both a prerequisite for effective AI deployment and something that the deployment itself can preserve or diminish.

THE TEST OF EFFECTIVE AI DEPLOYMENT



The Test of Effective AI Deployment

The three pillars of effective AI deployment are interdependent.

Institutional accountability establishes whether AI operates for a legitimate purpose, under identifiable authority and accountability.

Operational control translates that mandate into enforceable permissions, boundaries, monitoring, evidence, escalation and intervention.

Human capacity enables people to design, understand, govern and challenge AI, and to intervene effectively when circumstances require it.

Weakness in any pillar undermines the others:

Failure	Consequence
Accountability without operational control	Legitimate intention cannot reliably govern action.
Operational control without institutional accountability	Activity may be constrained without necessarily being legitimate or justified.
Accountability and control without human capacity	Governance becomes procedural rather than effective.
Human capacity without adequate institutional architecture	Capable people must continually compensate for systemic deficiencies.
Deployment-by-deployment governance alone	Individually governed systems may produce collectively ungoverned consequences.

The 3 questions that institutions must answer together:

Institutional accountability

Is the deployment operating for a legitimate institutional purpose, under identifiable authority and accountability, with consequences the institution has determined it is prepared to accept?

Operational control

Is that mandate translated into enforceable permissions, boundaries, monitoring, evidence, escalation and intervention at the point of action?

Human capacity

Do people retain the capacity required to design, understand, govern, challenge and intervene in the system – and does the resulting operating environment preserve and promote that capacity over time?

The Test of Effective AI Deployment Continued

The ultimate test is whether the institutional architecture as a whole continues to produce legitimate, governable and humanly sustainable action.

Human beings govern this transition, participate in it and live with its consequences. Effective AI deployment must therefore be judged not only by what the technology enables institutions to achieve, but by what its implementation enables human beings to become.

**AI capability must not be developed
at the expense of human capacity.**



EXECUTIVE BRIEFINGS

Bring us the decision you are facing.

You have read the framework. The next question is what it reveals about an AI system your institution is using or preparing to use.

HOSA works with leadership teams to examine consequential AI deployments across institutional accountability, operational control and human capacity. We begin with the decision in front of you: who has authority, what can be controlled, and whether the people relied upon to govern the system can genuinely do so.

Start a conversation with HOSA →



hosainstitute.com/hosa-on-ai

SELECTED FOUNDATIONS & FURTHER READING

research /rɪˈsɜːʃ/ (a)
■ n. [mass noun]
systematic investigation
of materials and
establish facts
conclusions:
research in ge

establi
conclusio
arch

Selected Foundations & Further Reading

The following publications provide context for the questions examined in this report. They are included to support further inquiry and do not imply endorsement of every position they contain.

Human Dignity and the Common Good

Pope Leo XIV. Magnifica Humanitas: On Safeguarding the Human Person in the Time of Artificial Intelligence. Vatican, 15 May 2026.

A substantial examination of artificial intelligence, human dignity, responsibility, work, freedom, education and the common good. Particularly relevant to institutional accountability and human capacity.

[Read the encyclical →](#)

AI Systems, Autonomy and Operational Control

Shavit, Yonadav, et al. Practices for Governing Agentic AI Systems. OpenAI, 14 December 2023.

An early examination of the governance challenges created by systems operating with limited direct supervision. It addresses suitability, constrained action, approvals, legibility, monitoring, attribution and interruptibility. Particularly relevant to operational control and continued institutional control.

[Read the paper →](#)

Santoni de Sio, Filippo, and Jeroen van den Hoven. Meaningful Human Control over Autonomous Systems: A Philosophical Account. Frontiers in Robotics and AI, 28 February 2018.

A philosophical and practical account of the conditions required for meaningful human control over autonomous systems. It argues that systems must respond to relevant human reasons and that their operation must remain traceable to people who understand their capabilities and recognise their responsibility for their use. Particularly relevant to legitimate authority, governed action, traceability and the human conditions required for continuing control.

[Read the paper →](#)

Kahl, Peter. The Answerability Fuse: A Statutory Trigger for Reconstructable Decisions. Lex et Ratio Ltd, 2026.

Proposes a statutory escalation trigger when a lawful AI deployment leaves consequential decisions unable to be adequately reconstructed and challenged. Particularly relevant to delegated authority, institutional answerability, evidence and intervention.

[Read the paper →](#)

Laux, Johann. Institutionalised Distrust and Human Oversight of Artificial Intelligence: Towards a Democratic Design of AI Governance under the European Union AI Act. *AI & Society*, Volume 39, 2024.

An examination of human oversight as an institutional arrangement rather than an individual act of vigilance. It argues that effective oversight requires structured scrutiny, contestability, distributed authority and the practical ability to intervene. Particularly relevant to institutional accountability, meaningful oversight and the conditions required for human authority to remain effective in practice.

[Read the article →](#)

Elish, Madeleine Clare. Moral Crumple Zones: Cautionary Tales in Human-Robot Interaction. *Engaging Science, Technology, and Society*, Volume 5, 2019.

An influential examination of how responsibility may be assigned to the human closest to an automated system even when that person possesses little meaningful authority or control. Particularly relevant to answerability, the distribution of responsibility and the distinction between formal human oversight and the practical capacity to understand, challenge and intervene.

[Read the paper →](#)

AI, Work and Human Capacity

Bainbridge, Lisanne. Ironies of Automation. *Automatica*, Volume 19, Issue 6, November 1983.

A foundational examination of how automation can make the remaining human role more difficult and consequential while reducing the opportunities through which the necessary skills are practised and maintained. Particularly relevant to skill atrophy, monitoring, intervention, apprenticeship and the preservation of human capacity within increasingly automated operating environments.

[Read the paper →](#)

Anthropic. The Anthropic Economic Index. Launched February 2025.

An ongoing research programme examining how AI is being used across occupations and economies, including the balance between augmentation and automation. Particularly relevant to work redesign, human participation and the development or displacement of human capability.

[Explore the Economic Index →](#)

Handa, Kunal, et al. *Which Economic Tasks Are Performed with AI? Evidence from Millions of Claude Conversations.* arXiv, 11 February 2025.

An empirical study of how generative AI is used across occupational tasks, distinguishing between augmentation and automation. Particularly relevant to changing roles, delegation, human capability and the limits of current evidence concerning AI's effects on work.

[Read the research paper →](#)

Frontier Capability and Deployment Risk

OpenAI. Preparedness Framework, Version 2. 15 April 2025.

A framework for identifying, evaluating and responding to advanced AI capabilities that could create severe harm. Particularly relevant to capability thresholds, deployment decisions, safeguards and the relationship between increasing capability and required control.

[Read the framework →](#)

Anthropic. Responsible Scaling Policy. First published 19 September 2023.

A framework connecting increasingly capable AI systems with progressively stronger technical and organisational safeguards. Particularly relevant to risk thresholds, deployment conditions, organisational responsibility and autonomous action contrary to institutional intent.

[Read the policy →](#)

Dragan, Anca, Helen King and Allan Dafoe. Frontier Safety Framework. Google DeepMind, 17 May 2024.

A framework for identifying advanced AI capabilities that may create severe harm and establishing corresponding safety and deployment measures. Particularly relevant to capability assessment, risk thresholds, institutional preparedness and continued control.

[Read the framework →](#)

Public Frameworks, Principles and Standards

National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1, United States Department of Commerce, 26 January 2023.

A voluntary framework for managing risks to individuals, organisations and society across the AI lifecycle. Its Govern, Map, Measure and Manage functions are particularly relevant to institutional accountability, operational control, monitoring and the continuing evaluation of consequences.

[Read the framework →](#)

Organisation for Economic Co-operation and Development. Recommendation of the Council on Artificial Intelligence: OECD AI Principles. Adopted 22 May 2019; updated May 2024.

The first intergovernmental standard on AI, establishing principles relating to human rights, democratic values, human agency, transparency, robustness, safety and accountability. Particularly relevant to societal mandate, government responsibility, institutional accountability and human capacity.

[Read the principles →](#)

International Organization for Standardization and International Electrotechnical Commission. ISO/IEC 42001:2023—Information Technology—Artificial Intelligence—Management System. First edition, December 2023.

The first international management-system standard for organisations developing, providing or using AI. It specifies requirements for establishing, implementing, maintaining and continually improving an AI management system. Particularly relevant to organisational governance, risk management, traceability, defined responsibilities and continuous operational control.

[View the standard →](#)

Infocomm Media Development Authority. Model AI Governance Framework for Agentic AI, Version 1.5. Government of Singapore, first published 22 January 2026; updated May 2026.

A voluntary framework for the responsible organisational deployment of AI systems capable of planning and acting with greater autonomy. It addresses risk assessment, limits on authority and access, meaningful human accountability, approval checkpoints, technical controls, third-party agents, continuous monitoring, automation bias, transparency and training. Particularly relevant to institutional accountability, operational boundaries, human oversight and continued institutional control.

[Read about the framework →](#)

The 2026 Singapore Consensus on Global AI Safety Research Priorities

Developed through the International Scientific Exchange on AI Safety, the 2026 Singapore Consensus sets out priority research areas for building trustworthy, reliable and secure AI systems. Its companion report identifies ten principles for managing agentic risk, including least privilege, traceable identity, auditability, validated deployment, runtime assurance, interruptibility, legibility and human oversight. It also recognises approval fatigue, cognitive overload and automation bias as threats to meaningful human control.

[Read the report →](#)

European Union and United States Approaches

The European Union and the United States currently represent materially different approaches to AI governance. The European Union has established a binding, cross-sector legal framework through the EU AI Act. The United States has no direct federal equivalent and instead relies upon national policy, existing law, agency enforcement, sector-specific regulation, voluntary standards and state legislation.

European Parliament and Council of the European Union. Regulation (EU) 2024/1689 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). Official Journal of the European Union, 12 July 2024.

The European Union's binding, risk-based legal framework for AI. It establishes obligations for providers, deployers and other actors, including requirements relating to risk management, data governance, technical documentation, record-keeping, transparency, human oversight, accuracy, cybersecurity and post-market monitoring. Particularly relevant to institutional accountability, operational control, fundamental rights and the legal limits of AI deployment.

[Read the regulation →](#)

Executive Office of the President of the United States. Winning the Race: America's AI Action Plan. The White House, 23 July 2025.

The United States' national AI policy strategy, organised around accelerating innovation, building AI infrastructure and strengthening international leadership and security. It promotes adoption, infrastructure expansion, workforce development, exports and the removal of regulatory barriers. Particularly relevant to government mandate, national capability, economic competitiveness and the contrasting institutional priorities shaping AI deployment.

[Read the action plan →](#)

HOSA Foundational Research

Shannon E. Eastman, MSc; Dr Abrar Hussain; and Dr Joshua Rosenthal.

Human Operating System Architecture Theory: A Biological Architecture of Readiness for Change That Holds; Recovery Dynamics as the Basis for Inferring Durable Change; and Biological Substrates of Change That Holds: Mechanistic Hypotheses. HOSA foundational theory series, Zenodo, 2026.

A three-paper theoretical series examining the biological conditions under which human change, adaptation and performance can stabilise over time. The series distinguishes observable capability and performance from the underlying capacity available to sustain them, and develops recovery dynamics as a basis for evaluating biological readiness and the durability of change. Originally published under Nervous System Intelligence Theory, the work is now situated within the broader Human Operating System Architecture. Particularly relevant to available capacity, sustainable demand, recovery and the conditions required for reliable human participation.

[Read Paper I: A Biological Architecture of Readiness for Change That Holds](#)

[Read Paper II: Recovery Dynamics as the Basis for Inferring Durable Change](#)

[Read Paper III: Biological Substrates of Change That Holds: Mechanistic Hypotheses](#)

Publication Note

As this report was being finalised, Anthropic published a new measurement framework for understanding the pace of AI development within frontier laboratories.

Anthropic. Measurements for Understanding the Pace of AI Development Inside Frontier Labs. Anthropic Institute, 17 September 2026.

The publication introduces three proposed measurements: the extent of AI-led AI research and development, the oversight of AI agents, and the allocation of computing resources. It provides a practical example of operational monitoring, evidence and traceability, escalation and intervention, while also illustrating the need to connect these controls to legitimate institutional authority and capable human participation. Particularly relevant to Operational Control and the cross-pillar requirements of evidence, adaptation and collective consequences.

[Read the publication →](#)

Publication & Attribution

The Three Pillars of Effective AI Deployment
Institutional Accountability × Operational Control × Human Capacity

Published by the HOSA Institute
First published September 2026
Version 1.0
© 2026 HOSA Institute. All rights reserved.

Attribution

The Three Pillars of Effective AI Deployment, including the selection, organisation and expression of its framework, conditions and visual materials, is original work developed and published by the HOSA Institute.

The official report may be shared in full and in its original, unmodified form.
When referring to the framework, please use:
The Three Pillars of Effective AI Deployment – HOSA Institute, 2026.

Recommended Citation

HOSA Institute. (2026). The Three Pillars of Effective AI Deployment: Institutional Accountability × Operational Control × Human Capacity.

Permissions

Extracts may be quoted or reproduced for non-commercial purposes with appropriate attribution and a link to the original report.

The report and its visual materials may not be modified, republished as another work or presented as the work of another person or organisation.

Requests concerning publication, collaboration, professional application or licensing should be directed to the HOSA Institute. leo@hosainstitute.com

Disclaimer

This report presents an institutional framework intended to support governance, assurance, policy and public discussion. It does not constitute legal, regulatory, technical or other professional advice.

The report reflects the HOSA Institute's analysis at the time of publication. Institutions remain responsible for obtaining appropriate professional advice and complying with all laws, regulations and standards applicable to their activities.

HOSA Institute
www.hosainstitute.com
#HOSAonAI

HOSA on AI

• **LIVE**

BIG QUESTIONS EXAMINING THE FUTURE WE ARE BUILDING



Subscribe for
event notifications



leo@hosainstitute.com
to join as a panelist or
guest speaker.



hosainstitute.com/**hosa-on-ai**



HUMAN OPERATING SYSTEM ARCHITECTURE

