Working Paper
Series B – Cognitive Architectures
WP-B-L02-001

Institutional Cognitive Architecture Model

IDHUS Institute

This publication forms part of the constitutional editorial architecture of the IDHUS Institute and belongs to the Institutional Cognition Knowledge Series.

ISBN:

Made and printed in Spain

Index

Abstract

Institutions increasingly operate in environments characterised by distributed expertise, expanding informational complexity, technological mediation and problems whose cognitive demands exceed the capacity of individual decision-makers. The theory of Institutional Cognition addresses this condition by examining how processes such as perception, memory, knowledge integration, reasoning, deliberation, decision-making, coordination, feedback and learning can become distributed across institutional systems. Yet identifying these cognitive functions does not by itself explain how they become organised into coherent forms of collective cognition. Institutions may possess substantial knowledge, analytical expertise and technological capability while remaining cognitively fragmented when the relationships connecting those capacities are weak, delayed or structurally misaligned.

This Working Paper develops the **Institutional Cognitive Architecture Model (ICAM)** as a middle-range analytical framework for investigating how cognitive functions become distributed and interconnected within institutions. Building upon the Institutional Cognition Conceptual Architecture (ICCA), ICAM shifts the object of analysis from the conceptual organisation of the scientific field towards the architecture of institutional cognition itself. The model represents institutional cognition through configurations of cognitive functions, human, organisational and computational nodes, cognitive interfaces, information and knowledge flows, coordination structures and feedback mechanisms. Particular attention is given to the relational properties through which distributed cognitive capacities become collectively consequential, including cognitive bottlenecks, redundancy, diversity, temporal synchronisation, architectural fit and the distinction between persistent structural architecture and the activated cognitive configurations mobilised around particular problems.

The paper further develops the concepts of **architectural failure** and **architectural diagnosis** to explain circumstances in which institutions possess relevant cognitive capacities but remain unable to integrate them effectively into collective judgement, legitimate action or institutional learning. It also examines the historical formation and path dependence of cognitive architectures, their capacity for reconfiguration under crisis conditions, the emergence of metacognitive mechanisms through which institutions may examine their own cognitive processes, and the increasingly important integration of human, organisational and computational nodes within shared institutional systems. Throughout the model, cognitive capability remains analytically distinct from legitimate authority, preserving the normative and constitutional dimensions through which institutional decisions acquire collective validity.

ICAM establishes institutional cognitive architecture as a researchable object positioned between individual cognition and system-level governance outcomes. Rather than proposing a universal design for intelligent institutions, it provides a framework through which different cognitive architectures may progressively be reconstructed, compared, measured, simulated and experimentally investigated. In doing so, the model creates an architectural bridge between the conceptual foundations of Institutional Cognition and the subsequent operational models of the research programme, supporting a cumulative investigation into the conditions under which **distributed cognitive capacity can become coherent, adaptive and legitimate collective intelligence**.

1. Introduction — From Conceptual Architecture to Institutional Architecture

The development of the Institutional Cognition Conceptual Architecture established an integrated framework through which the expanding theoretical domain of Institutional Cognition could be organised across foundational, functional, architectural, operational and evolutionary levels. That architecture addressed a problem internal to the scientific programme itself: as the field expanded towards computational governance, human–AI collaboration, measurement, simulation, comparative research and institutional experimentation, its concepts required a sufficiently coherent structure to remain mutually intelligible. ICCA therefore provided an architecture of the science, clarifying how different theoretical propositions and research domains could remain connected while preserving the possibility of revision as operational evidence accumulated. The next stage requires a deliberate change in the object of analysis. Rather than asking how the scientific field should organise its concepts, the present Working Paper asks how cognition itself becomes organised within institutions.

This transition is consequential because the proposition that institutions possess cognitive capacities remains incomplete unless the mechanisms through which those capacities become collectively realised can be described with greater precision. Institutions may perceive changing environments, preserve knowledge, interpret evidence, reason about alternatives, coordinate specialised actors, make collective decisions and learn from consequences, yet none of these functions exists independently of the organisational structures through which it is performed. Information is collected by particular actors and systems; knowledge is preserved within particular repositories and communities; analytical judgements travel through particular channels; decisions emerge through particular configurations of authority and deliberation; and feedback reaches some parts of the organisation more readily than others. Institutional cognition therefore possesses an architecture even when that architecture has never been designed explicitly as a cognitive system.

The **Institutional Cognitive Architecture Model (ICAM)** is proposed as a framework for investigating this architecture. It begins from the Functional and Architectural distinctions established within ICCA but changes their analytical orientation. Cognitive functions such as perception, memory, reasoning, decision-making and learning are no longer considered principally as conceptual categories within the theory; they become processes whose institutional realisation must be located, connected and examined. The central question consequently shifts from what institutional cognition consists of towards how its components become organised into systems capable of collective cognitive activity.

This shift also requires moving beyond an assumption that remains implicit in many descriptions of organisational capability: that institutional intelligence can be inferred from the quality or quantity of its individual components. An institution may employ highly capable experts, maintain sophisticated information infrastructures and possess extensive analytical resources while nevertheless producing weak collective judgement if knowledge cannot move effectively between those components. Conversely, comparatively modest individual capabilities may contribute to robust institutional cognition when specialised knowledge is connected through effective interfaces, coordination mechanisms and feedback structures. The architecture through which capacities become related may therefore matter as much as the capacities themselves.

ICAM develops this proposition by treating institutional cognition as an organised configuration of cognitive functions distributed across human, organisational and computational nodes and connected through interfaces, information flows, coordination structures and feedback mechanisms. These elements do not form a simple processing chain in which information enters an institution, is transformed into a decision and subsequently produces an output. They constitute a distributed and recursive system in which memory influences perception, reasoning may generate demands for additional information, deliberation can alter the interpretation of evidence, implementation produces new knowledge and feedback can modify the structures through which future cognition occurs.

The objective of the present Working Paper is therefore not to produce a complete formal model of institutional behaviour, nor to suggest that the cognitive architecture of complex institutions can be reduced to a single universal configuration. Its purpose is to establish an intermediate analytical model through which the internal organisation of institutional cognition can become sufficiently explicit to support later measurement, simulation, comparison and experimentation. **ICAM represents the movement from identifying institutional cognitive functions towards investigating the architectures through which those functions become collectively consequential.**

2. The Research Problem — Where Does Institutional Cognition Occur?

The claim that an institution perceives, remembers, reasons or learns contains a conceptual ambiguity that becomes increasingly important as Institutional Cognition moves towards operational research. Institutions are collective entities composed of people, organisational units, formal procedures, information systems, technological infrastructures and networks of relationships, none of which individually constitutes the institution as a whole. When an institution is described as possessing knowledge or making a decision, the language therefore compresses a distributed process into an apparently unified actor. Such compression is often useful for ordinary institutional analysis, but it becomes insufficient when the objective is to understand the mechanisms through which collective cognition actually occurs.

Institutional perception, for example, may be distributed across statistical agencies, regulatory inspectors, public consultations, scientific advisory systems, operational personnel, external researchers and automated monitoring infrastructures. Institutional memory may reside partly in archives and databases, partly in procedures and legal precedents, and partly in the accumulated expertise of personnel whose knowledge has never been formally documented. Reasoning may occur within analytical units, advisory committees, ministerial offices or informal professional networks, while the authority to convert those judgements into decisions may reside elsewhere. Learning may emerge through formal evaluation processes, but it may also occur through gradual changes in professional practice or organisational expectations that are difficult to locate within a single administrative structure.

The question of where institutional cognition occurs therefore cannot be answered simply by identifying the institution's formal organisational units. Cognitive processes frequently cross organisational boundaries, and the same organisational component may perform several cognitive functions simultaneously. A scientific advisory body may contribute to perception by identifying emerging risks, to memory by preserving specialised knowledge and to reasoning by evaluating policy alternatives. Conversely, a single cognitive function may be distributed across many organisational components whose contributions must somehow be integrated before they become institutionally meaningful.

This distributed character suggests that the appropriate unit of analysis is neither the individual actor nor the institution understood as an indivisible whole, but the **architecture connecting heterogeneous cognitive contributions**. ICAM therefore introduces the concept of the cognitive node to identify components capable of contributing to one or more institutional cognitive functions. A node may correspond to an individual official, a team, department, committee, expert network, database, analytical platform or artificial intelligence system, depending upon the scale and purpose of the analysis. The concept is deliberately functional rather than organisational: what makes something a cognitive node is not its formal position within an administrative hierarchy but its contribution to processes through which information, knowledge or judgement becomes institutionally organised.

Identifying nodes alone, however, does not resolve the research problem. A collection of cognitively capable components does not necessarily constitute a cognitively integrated institution. Their contributions must become connected through mechanisms that determine what information can travel, in which direction, with what delay, under what conditions and with what transformation along the way. A highly capable analytical node whose findings cannot reach relevant decision processes may contribute little to collective cognition, while an extensive institutional memory that cannot be retrieved when needed may exist administratively without functioning cognitively. The location of cognitive capacity must therefore be analysed together with the relationships through which that capacity becomes available to the wider institutional system.

The question *where does institutional cognition occur?* consequently leads towards a more relational answer than the question initially appears to invite. Cognition occurs within particular nodes, but **institutional cognition emerges through the architecture connecting those nodes into patterns of collective perception, memory, reasoning, decision, action and learning**. Understanding that architecture constitutes the central research problem of ICAM.

3. Institutional Cognitive Architecture as an Analytical Object

The concept of architecture provides a way of moving between two analytical extremes that would otherwise limit the study of Institutional Cognition. At one extreme lies methodological individualism, in which institutional cognition would ultimately be reduced to the mental processes and decisions of individual human actors. At the other lies an excessively holistic representation of the institution as if it possessed a unified cognitive centre capable of perceiving, reasoning and deciding in a manner analogous to an individual mind. Neither position adequately represents the distributed organisation through which contemporary institutions process knowledge.

Institutional cognitive architecture can instead be understood as **the organised configuration of cognitive functions, cognitive nodes, interfaces, information and knowledge flows, coordination mechanisms and feedback structures through which cognition becomes collectively realised within an institution or governance system**. The concept does not require institutions to resemble biological brains, nor does it imply that collective cognition reproduces individual cognition at a larger scale. Architecture refers here to organisation: the patterned relationships through which specialised cognitive contributions become connected sufficiently to generate system-level capacities.

This definition introduces an important distinction between cognitive capacity and cognitive architecture. Capacity describes what a component or system may be capable of performing, whereas architecture describes how those capabilities are organised in relation to one another. Two institutions may possess comparable analytical resources while displaying very different cognitive architectures. One may concentrate analysis within a central expert unit, while another distributes it across specialised departments; one may preserve institutional memory through formal digital repositories, while another depends heavily upon professional continuity and tacit knowledge; one may integrate evidence through hierarchical reporting, while another relies upon networked deliberation. Similar functional capacities can therefore be realised through substantially different architectures.

The reverse is also possible. Similar organisational structures may produce different cognitive consequences because the relationships operating within them differ. Two ministries may possess comparable departments, committees and reporting procedures on paper, yet one may exhibit rapid and reliable knowledge transfer while the other suffers from informational fragmentation. Formal architecture alone cannot therefore determine cognitive architecture. What matters is how organisational structures actually participate in the production, transmission, transformation and use of knowledge.

ICAM should consequently be understood as a **meso-level model** positioned between individual cognition and system-level institutional behaviour. It investigates how local cognitive contributions become organised into collective processes without assuming that those processes can be explained entirely by either individual psychology or macro-level institutional outcomes. This intermediate level is especially important for operational research because many institutional failures appear precisely within the relationships between components: information reaches the wrong actors, expertise becomes isolated, feedback arrives too late, organisational memory cannot be retrieved or decision structures become overloaded by analytical complexity.

The architectural perspective also allows institutional cognition to be compared without assuming that one organisational form represents a universal optimum. Hierarchical, networked, decentralised and hybrid configurations may each support different cognitive functions under different conditions. Centralisation may accelerate coordination while reducing informational diversity; distributed architectures may expand sensing and specialised knowledge while increasing the complexity of integration. ICAM therefore seeks to provide a language through which such configurations can be investigated rather than prescribing a predetermined model of the intelligent institution.

This distinction between architecture and capacity will become increasingly important as artificial intelligence enters institutional environments. Adding a computational system capable of performing sophisticated analysis increases a particular form of cognitive capacity, but its effect upon institutional cognition depends on where the system is positioned, which information it can access, which actors can interrogate its outputs, how uncertainty is communicated and whether its contributions can enter legitimate processes of judgement and decision. **Cognitive capability becomes institutionally meaningful through architecture.**

4. Beyond the Organisational Chart — Mapping the Cognitive Institution

The formal organisation of an institution provides an indispensable description of authority, responsibility and administrative structure, but it does not necessarily reveal how cognition is organised. Organisational charts identify departments, reporting relationships and formal positions; procedural rules identify responsibilities and decision rights; legal frameworks establish competencies and constraints. These structures shape institutional cognition profoundly, yet the pathways through which knowledge actually moves may extend across, around or occasionally outside them.

An institutional cognitive map would therefore ask different questions from an organisational chart. Rather than beginning primarily with who reports to whom, it would examine where relevant signals enter the institution, where different forms of knowledge are preserved, which actors interpret them, how competing interpretations are reconciled, where uncertainty becomes visible, how expertise reaches authorised decision-makers and through which mechanisms the consequences of decisions return to those capable of learning from them. The resulting map may overlap substantially with formal organisational structure, but it may also reveal cognitive relationships that administrative representations leave largely invisible.

Informal professional networks provide one example. Experienced officials may know whom to contact when unusual problems arise even when no formal procedure identifies that relationship. Such networks can operate as highly effective cognitive interfaces by connecting specialised knowledge across departmental boundaries, yet their dependence upon particular individuals may also create vulnerability when personnel change. Similarly, scientific advisory bodies may occupy peripheral positions within formal authority structures while remaining cognitively central because they provide access to knowledge unavailable elsewhere in the institution. Conversely, senior decision-making bodies may occupy the apex of the organisational hierarchy while depending heavily upon cognitive processes occurring throughout the wider system.

Digital infrastructures complicate the relationship further. Databases, collaborative platforms, automated monitoring systems and artificial intelligence can become important cognitive components without occupying conventional positions within organisational hierarchies. A data platform may connect previously separated departments; an automated monitoring system may become a major perceptual node; an AI-assisted retrieval system may transform institutional memory by changing which historical knowledge becomes practically accessible. The cognitive architecture of contemporary institutions therefore increasingly contains relationships that cannot be represented adequately through conventional administrative diagrams alone.

Mapping these relationships does not diminish the importance of formal authority. On the contrary, distinguishing cognitive architecture from organisational hierarchy allows their interaction to be examined more precisely. Knowledge may be distributed widely while authority remains concentrated; expertise may be decentralised while decision procedures remain hierarchical; computational systems may perform substantial analytical work while possessing no legitimate decision authority. The institutional question is therefore not simply where cognition occurs, but how cognitive contributions reach the structures through which legitimate collective action becomes possible.

This distinction also creates the possibility of identifying architectural misalignments. Formal authority may be located far from relevant expertise; information may travel through reporting chains that introduce distortion or delay; decision-makers may receive highly compressed representations of complex evidence; feedback may remain within evaluation units without reaching those responsible for policy redesign. Such problems cannot always be resolved by reorganising administrative units because the underlying difficulty may lie within the cognitive relationships connecting them.

ICAM therefore proposes that institutions can be represented simultaneously through at least two overlapping maps: an organisational map describing formal structures of responsibility and authority, and a cognitive map describing the distribution and integration of knowledge-related functions. **The difference between these maps may itself become an important object of institutional research**, revealing where formal organisation supports cognition, where informal structures compensate for architectural weaknesses and where cognitive capacity remains disconnected from legitimate decision processes.

5. Cognitive Functions and Their Architectural Realisation

ICCA established cognitive functions as one of the principal analytical dimensions of Institutional Cognition, identifying processes such as perception, memory, knowledge integration, reasoning, deliberation, decision-making, coordination, feedback, learning and adaptation. ICAM begins from these distinctions but asks a different question: through what institutional arrangements do these functions become realised, and how does their distribution influence the cognitive behaviour of the wider system?

No cognitive function should be assumed to correspond automatically to a single organisational component. Institutional perception may emerge through the combined activity of statistical systems, regulatory inspections, public engagement, scientific monitoring and frontline personnel. Memory may be distributed across formal archives, digital databases, procedural routines, legal precedent and human expertise. Reasoning may involve specialised analytical teams, political advisers, external experts and computational systems whose outputs enter deliberative processes through different channels. Decision-making may be formally concentrated while the cognitive work supporting those decisions remains highly distributed.

This distinction between function and organisational location is fundamental because it allows ICAM to investigate different architectural realisations of the same cognitive requirement. Two institutions may both possess mechanisms of environmental sensing, yet one may rely primarily upon centralised statistical monitoring while another integrates distributed observations from regional offices and external networks. Both perform perception, but their architectures may differ in sensitivity, diversity, latency and vulnerability to correlated error. Similar variation may occur across every cognitive function.

Functions may also depend upon one another in ways that cannot be understood through their individual capacities alone. Perception without accessible memory may produce repeated rediscovery rather than cumulative knowledge; memory without retrieval interfaces may preserve information without making it cognitively available; analytical reasoning without deliberative integration may remain disconnected from collective judgement; feedback without organisational mechanisms for revision may generate evaluation without learning. The architectural problem is therefore not merely whether a function exists, but whether it can participate effectively within a wider configuration of cognition.

This observation provides a more precise meaning to the proposition introduced through ICCA that institutional intelligence is relational before it is additive. Increasing the capacity of an isolated function may produce diminishing or even counterproductive returns when neighbouring parts of the architecture cannot absorb its outputs. Expanding environmental monitoring can overwhelm analytical capacity; increasing analytical production can saturate decision processes; accelerating decision-making can reduce opportunities for knowledge integration; expanding feedback can generate organisational noise when institutions lack mechanisms for distinguishing meaningful signals from routine variation. Cognitive capacity therefore depends partly upon the compatibility and connectivity of functions across the architecture.

ICAM consequently treats cognitive functions not as independent modules but as **architecturally situated processes whose institutional significance emerges through their relationships**. This perspective prepares the transition towards one of the central components of the model: the cognitive interfaces through which specialised functions become connected and through which distributed cognition acquires the possibility of becoming collective intelligence.

6. Cognitive Nodes — Locating Distributed Capacity

If cognitive functions describe what occurs within institutional cognition, cognitive nodes provide an initial means of describing where particular contributions are generated, preserved or transformed. The concept is intentionally broader than the conventional notion of organisational unit because cognition within institutions is increasingly distributed across heterogeneous components that differ substantially in their organisational status, material form and relationship to authority.

Human actors represent the most familiar cognitive nodes. Officials, analysts, scientists, advisers, managers and political leaders possess specialised knowledge and perform forms of interpretation or judgement that contribute to institutional cognition. Yet individuals frequently participate within collective nodes whose capacities cannot be understood simply as the sum of their members. Committees, analytical teams, expert panels and departments can develop shared procedures, accumulated knowledge and specialised forms of reasoning that persist despite changes in individual membership. Organisational components may therefore become cognitively significant entities in their own right when the level of analysis concerns how knowledge is distributed across the institution.

Technological systems introduce another category. Databases can function as components of institutional memory; monitoring infrastructures can contribute to perception; computational models can support reasoning; artificial intelligence systems can classify information, retrieve knowledge, generate scenarios or synthesise evidence. Describing such systems as cognitive nodes does not imply equivalence between human and artificial cognition, nor does it confer agency or authority upon technologies. The concept identifies only their functional position within a wider cognitive architecture.

Hybrid nodes are likely to become increasingly important because many institutional cognitive processes already depend upon combinations of human judgement and computational capability. An analytical team using machine-learning systems, a policy unit working with simulation tools or an official interacting with an AI-supported knowledge system may perform cognitive functions that cannot be attributed meaningfully to either the human or technological component in isolation. The relevant analytical object becomes the configuration through which their capabilities interact.

Nodes also vary in their degree of permanence. Some are embedded within stable organisational structures, while others emerge temporarily around crises, projects or particular policy problems. External expertise may enter institutional cognition through advisory arrangements without becoming permanently incorporated into the organisation. Networks connecting government agencies, universities, private organisations and civil society may function as distributed cognitive structures whose boundaries extend beyond the institution conventionally understood.

The identification of cognitive nodes therefore represents only the beginning of architectural analysis. Mapping where capacities reside does not explain whether those capacities contribute effectively to collective cognition. A node may possess exceptional expertise while remaining informationally isolated; another may occupy a central network position despite possessing relatively limited analytical capability. **The cognitive significance of a node depends not only upon what it can do, but upon how it is connected to the wider institutional architecture.** It is this relational problem that leads directly towards the study of cognitive interfaces.

7. Cognitive Interfaces — Connecting Distributed Institutional Cognition

The identification of cognitive nodes makes visible the distributed character of institutional cognition, but distribution alone cannot explain how a collection of specialised capacities becomes capable of functioning as a collective system. Institutions routinely contain substantial reservoirs of expertise, information and analytical capability that remain only partially connected to one another, and the existence of those resources does not guarantee that they will become available when and where they are needed. The architectural problem therefore shifts from the presence of cognitive capacity towards the mechanisms through which different capacities become mutually accessible. ICAM describes these mechanisms as **cognitive interfaces**, understood as the organisational, informational, procedural, human or technological arrangements through which cognitive nodes and functions exchange information, knowledge, interpretations and feedback.

The concept of the cognitive interface is deliberately broader than communication. Communication refers principally to the transmission of information between actors or systems, whereas an interface determines the conditions under which that transmission becomes cognitively consequential. A report may be delivered without being incorporated into reasoning; a database may be technically accessible while remaining practically disconnected from the workflows within which decisions are made; an advisory body may communicate sophisticated analysis whose format, timing or institutional position prevents it from influencing deliberation. The existence of communication therefore does not demonstrate the existence of effective cognitive integration. What matters is whether the architecture allows a contribution generated within one part of the institution to become usable within another.

Cognitive interfaces can take highly diverse forms. Formal reporting procedures connect operational units with managerial structures; committees bring specialised perspectives into shared deliberative environments; knowledge-management systems connect institutional memory with contemporary decision processes; digital platforms allow geographically distributed actors to work from common information; professional communities transmit tacit knowledge across organisational boundaries; and human–AI interaction mechanisms increasingly mediate access to analytical and computational capacities. Some interfaces are deliberately designed, while others emerge informally through repeated institutional practice. Their cognitive importance does not depend primarily upon their formal status but upon the role they perform in connecting otherwise separated processes.

The quality of an interface may influence institutional cognition along several dimensions simultaneously. Information can arrive too slowly to remain relevant, become excessively compressed as it travels through hierarchical structures, lose contextual meaning when translated between professional communities or acquire an authority that exceeds the reliability of the underlying evidence. Interfaces therefore do not merely transmit cognition; they transform it. Every translation from technical analysis into executive briefing, from local observation into national statistics, or from model output into policy recommendation involves decisions about what information is retained, what is simplified and how uncertainty is represented. The interface becomes part of the cognitive process because the representation received by one node is not necessarily equivalent to the representation produced by another.

This observation helps explain why institutions possessing individually capable components can nevertheless exhibit weak collective cognition. The failure may reside neither within perception nor within reasoning, neither within memory nor within decision-making, but within the interfaces through which these functions are expected to interact. An institution may detect an emerging risk accurately and possess sufficient analytical expertise to understand it, yet fail to respond because the information cannot cross the organisational boundary separating monitoring from authorised decision-making. Another may conduct rigorous evaluations whose findings remain cognitively peripheral because they are published after the institutional cycle in which policy revision occurs. These are architectural failures of connection rather than simple deficits of capability.

Cognitive interfaces consequently provide one of the principal mechanisms through which the proposition that institutional intelligence is relational before it is additive becomes operationally meaningful. The capacity of an institutional system depends partly upon the quality of the relationships connecting its specialised components, and improvements to those relationships may sometimes produce greater cognitive gains than increasing the capabilities of the components themselves. **ICAM therefore treats interfaces not as secondary channels between cognitive nodes, but as constitutive elements of institutional cognitive architecture.**

8. Cognitive Flows — What Moves through the Institution?

Once cognitive interfaces are recognised as architectural components, a further question becomes unavoidable: what exactly moves through them? Institutional cognition is frequently described in terms of information flows, yet information is not a homogeneous substance that travels unchanged across an organisation. What begins as observation may become data, what is represented as data may acquire interpretation, interpreted evidence may become institutional knowledge, knowledge may contribute to judgement, and judgement may eventually become recommendation, decision or action. The cognitive significance of what moves through an institution therefore changes as it encounters different functions and nodes.

ICAM uses the concept of **cognitive flow** to capture this broader process. A cognitive flow refers not simply to the transmission of data but to the movement and transformation of epistemically relevant content across institutional architecture. This distinction matters because institutions can possess highly efficient information infrastructures while remaining cognitively fragmented. Data may circulate extensively without becoming integrated into knowledge, knowledge may exist without influencing deliberation, and recommendations may reach decision-makers without preserving the reasoning or uncertainty from which they were produced.

Environmental sensing provides an illustrative starting point. A regulatory inspector, statistical system, scientific monitoring network or automated sensor may detect a change in the external environment, but that signal becomes institutionally meaningful only through processes of selection and representation. Some observations are recorded while others are ignored; categories determine how phenomena are classified; thresholds determine which deviations become noteworthy; and aggregation procedures influence which local variations remain visible at higher organisational levels. Institutional perception therefore begins transforming reality before information reaches formal processes of reasoning.

As cognitive flows move through the architecture, they may undergo repeated changes in resolution and meaning. Detailed technical evidence may be summarised for senior officials, competing analyses may be synthesised into a common briefing, probabilistic forecasts may be translated into qualitative risk categories and complex uncertainties may eventually appear within decision processes as apparently discrete alternatives. These transformations are often necessary because no institutional actor can process the totality of available information, yet each transformation introduces the possibility of information loss, distortion or unintended amplification.

The architecture of cognitive flows therefore influences what an institution can know about itself and its environment. Hierarchical systems may concentrate information efficiently while progressively compressing contextual detail; networked systems may preserve informational diversity while generating difficulties of synthesis; highly automated systems may process enormous quantities of data while introducing forms of opacity concerning how representations were generated. No single flow architecture is universally optimal because different governance problems impose different requirements concerning speed, diversity, traceability, contextual richness and coordination.

The analysis of cognitive flows also reveals why institutional knowledge cannot be identified simply with the amount of information available. **Knowledge becomes institutionally consequential when information has been interpreted, situated and connected sufficiently to participate in reasoning and action.** The transition between information and knowledge is therefore not merely semantic; it is architectural, because it depends upon the nodes and interfaces through which interpretation occurs.

9. From Information to Knowledge and Institutional Judgement

The transformation of information into institutional judgement represents one of the most complex processes within cognitive architecture because it involves several qualitatively different forms of cognitive activity. Institutions rarely make consequential decisions directly from raw observations. Signals must be interpreted, evidence compared, historical experience retrieved, uncertainty assessed, alternative explanations considered and possible courses of action evaluated. The resulting judgement is therefore neither a simple aggregation of information nor the product of a single cognitive node, but the outcome of an architecture through which different forms of knowledge become progressively integrated.

The distinction between information and knowledge is particularly important in this context. Information can be understood as represented observations or claims that become available to institutional processes, whereas knowledge involves information that has acquired interpretative context and can contribute meaningfully to understanding. Institutional memory plays an important role in this transformation because contemporary observations often acquire significance only when compared with previous events, established models or accumulated professional experience. A statistical deviation may become meaningful because an institution remembers similar patterns from an earlier crisis; a new technological development may become relevant because existing expertise allows its implications to be interpreted.

Yet knowledge does not automatically produce judgement. Institutions frequently confront situations in which different bodies of knowledge point towards competing conclusions, evidence remains incomplete and objectives themselves are contested. Institutional judgement therefore requires the integration of epistemic and practical considerations under conditions of uncertainty. Scientific evidence may indicate probable consequences, legal analysis may identify constraints, operational expertise may reveal implementation difficulties and political deliberation may determine which trade-offs can legitimately be accepted. The cognitive architecture must somehow bring these contributions into relation without assuming that they can be reduced to a single informational metric.

This process makes deliberation particularly significant. Deliberation is not merely the exchange of opinions but an architectural mechanism through which different representations of a problem can encounter one another. When functioning effectively, it can expose hidden assumptions, identify conflicts between forms of expertise and allow uncertainty to remain visible during the formation of judgement. When functioning poorly, it can amplify hierarchy, suppress dissenting information or transform complex evidence into premature consensus. The design of deliberative interfaces therefore influences the epistemic quality of institutional judgement.

Artificial intelligence introduces additional complexity because computational systems can increasingly participate in several stages of this transformation. They may classify observations, retrieve historical knowledge, identify patterns, synthesise evidence or generate possible scenarios, thereby becoming significant contributors to institutional reasoning. Yet their outputs remain situated within wider architectures of interpretation and authority. A computational prediction becomes institutionally meaningful only when actors understand sufficiently what it represents, how uncertainty should be interpreted and what role the output is permitted to play within legitimate decision processes.

Institutional judgement consequently emerges from a sequence of transformations that cannot be attributed exclusively to information, expertise or authority. **It is an architectural achievement through which heterogeneous cognitive contributions become sufficiently integrated to support collective choice without eliminating the uncertainty, disagreement and normative constraints inherent in governance.** Understanding how that integration occurs is therefore central to any model of intelligent institutions.

10. Coordination Architectures and the Integration of Distributed Cognition

The distributed nature of institutional cognition creates a coordination problem that is inseparable from the problem of intelligence itself. Specialisation allows institutions to develop forms of expertise that would be impossible for any single actor or organisational unit to possess, yet increasing specialisation also fragments knowledge across professional, administrative and technological boundaries. The cognitive advantage produced by distribution can therefore be realised only if institutions possess mechanisms capable of coordinating contributions without destroying the diversity from which those contributions derive.

Hierarchy represents one historically important solution. By establishing formal reporting relationships and decision authority, hierarchical structures can aggregate information and coordinate action across large organisations. They reduce ambiguity concerning responsibility and can facilitate rapid mobilisation when tasks are sufficiently well defined. Yet hierarchy also influences cognition by determining which information moves upward, how it is compressed and whose interpretations acquire organisational visibility. The same structure that enables coordination may therefore produce cognitive bottlenecks when large volumes of heterogeneous information must pass through a limited number of decision points.

Networked forms of coordination offer different possibilities. Horizontal relationships can connect specialised actors directly, allowing knowledge to move without traversing every level of formal hierarchy. Professional networks, interdepartmental working groups and cross-institutional collaborations may preserve greater informational diversity and facilitate rapid knowledge exchange across organisational boundaries. Their limitations arise elsewhere: distributed networks may struggle to establish common representations, resolve disagreement or convert shared understanding into coordinated action when authority remains fragmented.

Committees and deliberative bodies occupy an intermediate position because they create institutional spaces in which distributed knowledge can be integrated without necessarily eliminating formal hierarchy. Their cognitive value depends less upon their existence than upon their composition, information access, deliberative procedures and relationship with decision authority. A committee containing diverse expertise may expand institutional reasoning, but excessive size, unclear responsibility or poorly structured deliberation can reduce its capacity to transform diversity into coherent judgement.

Computational coordination introduces another emerging possibility. Shared data infrastructures, collaborative platforms, automated workflows and AI-supported synthesis can reduce some of the transaction costs associated with distributed cognition, making information accessible across organisational boundaries that previously constrained integration. Yet computational coordination can also standardise representations in ways that suppress contextual differences, and systems designed to optimise informational efficiency may inadvertently privilege what can be formalised over forms of tacit or qualitative knowledge that remain institutionally important.

ICAM therefore does not treat centralisation or decentralisation as inherently superior cognitive designs. Their value depends upon the nature of the environment, the distribution of expertise, the temporal demands of decision-making and the consequences of error. Stable and repetitive problems may benefit from architectures different from those required for uncertain, rapidly changing or highly complex environments. The relevant question is not whether cognition should be centralised or distributed, but **which forms of distribution and coordination allow relevant knowledge to remain diverse while becoming sufficiently integrated for collective action**.

This balance between differentiation and integration is one of the central architectural problems of Institutional Cognition. Too little differentiation can produce epistemic homogeneity and blind spots; too little integration can produce fragmentation and incoherent action. Institutional intelligence therefore depends partly upon maintaining a dynamic relationship between specialised cognition and collective coordination.

11. Cognitive Bottlenecks and Architectural Constraints

The relational perspective developed through ICAM suggests that the cognitive capacity of an institution may be constrained disproportionately by particular components or connections within its architecture. Increasing the capability of one part of the system does not necessarily increase the capability of the whole when another part remains unable to process, transmit or integrate the additional contribution. This condition can be described as a **cognitive bottleneck**, referring to an architectural constraint that limits the effective use of cognitive capacity elsewhere within the institution.

Bottlenecks can arise at many points. An institution may possess extensive sensing capacity but insufficient analytical resources to interpret incoming information. It may possess strong analytical capacity but limited deliberative bandwidth, causing sophisticated evidence to accumulate faster than decision processes can evaluate it. Institutional memory may preserve enormous quantities of historical information while weak retrieval mechanisms prevent relevant knowledge from becoming available when needed. Feedback systems may generate detailed evaluations whose findings reach policy processes only after decisions have already become institutionally entrenched.

These examples reveal why the relationship between resources and intelligence is not necessarily linear. More information can increase cognitive burden; more expertise can intensify coordination problems; more analytical outputs can overwhelm decision-makers; and faster communication can amplify irrelevant signals alongside meaningful ones. Institutional cognitive performance therefore depends partly upon the balance between capacities across interconnected functions.

Bottlenecks may also be produced by authority structures. A small number of senior actors may become required to authorise large numbers of decisions, creating cognitive congestion even when information reaches them efficiently. Conversely, excessive decentralisation may distribute authority so widely that coordination becomes the limiting factor. In both cases, the constraint emerges from the relationship between cognitive architecture and institutional authority rather than from a simple shortage of knowledge.

The concept also has important implications for technological augmentation. Introducing artificial intelligence into an institution may dramatically increase analytical or generative capacity while leaving downstream cognitive interfaces unchanged. If decision-makers previously received ten complex analyses and can now receive hundreds, the institution has increased computational capacity without necessarily increasing its capacity for judgement. The resulting architecture may become more cognitively imbalanced rather than more intelligent.

Cognitive bottleneck analysis therefore changes the way institutional capability might eventually be assessed. Rather than asking only how much capacity exists within each function, researchers may need to investigate **where the architecture limits the conversion of local capacity into system-level cognition**. This will become particularly relevant for later measurement and simulation research, where bottlenecks may provide observable and modelled explanations for differences between nominal institutional capability and realised cognitive performance.

12. Cognitive Redundancy, Diversity and Institutional Resilience

If bottlenecks reveal how insufficient capacity or connectivity can constrain institutional cognition, redundancy reveals a different architectural problem. Organisational efficiency is often associated with eliminating duplication, yet cognitive systems operating under uncertainty may benefit from maintaining partially overlapping capacities. Multiple sources of information, independent analytical perspectives and distributed repositories of knowledge can create forms of resilience that would disappear if every cognitive function were concentrated within a single optimised component.

Redundancy becomes cognitively valuable when different nodes can compensate for one another or detect errors that would otherwise propagate through the system. Independent monitoring systems may identify signals missed by a centralised sensor; parallel analytical teams may expose assumptions embedded within one another's models; overlapping institutional memories may preserve knowledge when a database fails or experienced personnel leave. The apparent duplication of capacity can therefore reduce vulnerability to single points of cognitive failure.

Yet redundancy is not automatically beneficial. Multiple units collecting identical information without coordination may generate administrative burden without increasing perceptual diversity. Parallel analytical processes may reproduce the same assumptions because they depend upon common data or methodological traditions. Several information repositories may create confusion when their contents diverge or become difficult to reconcile. The architectural value of redundancy therefore depends upon whether overlapping capacities contribute genuine diversity or merely replicate the same cognitive structure.

This distinction leads towards the concept of **cognitive diversity**, which concerns variation in the perspectives, knowledge bases, methodologies and representations through which institutional problems are interpreted. Diversity can increase the probability that weak signals, alternative explanations or unintended consequences become visible, particularly under conditions of uncertainty. At the same time, greater cognitive diversity increases the demands placed upon interfaces and coordination structures because heterogeneous representations must somehow become mutually intelligible before collective action can occur.

Institutional resilience may consequently depend upon maintaining a productive relationship between redundancy, diversity and integration. Highly centralised architectures can sometimes achieve efficiency by reducing duplication, yet they may become vulnerable when their dominant representations fail. Highly distributed architectures may preserve alternative knowledge but struggle to coordinate responses under time pressure. Resilience emerges not from maximising either property but from configuring sufficient diversity and redundancy while preserving the interfaces required for collective cognition.

This perspective also suggests that some apparent organisational inefficiencies may perform latent cognitive functions that become visible only during disruption. Informal networks, overlapping expertise and multiple channels of information can appear unnecessary during routine administration but become critical when established pathways fail. **The architecture of an intelligent institution must therefore be understood not only in relation to normal cognitive performance, but also in relation to its capacity to preserve cognition when parts of the system become unreliable.**

13. Feedback Structures and the Architecture of Institutional Learning

The capacity to act collectively does not in itself make an institution adaptive. An institutional architecture may perceive changing conditions, integrate knowledge and coordinate decisions with considerable sophistication while remaining unable to modify its behaviour when the consequences of those decisions differ from what was expected. The transition from decision-making towards learning therefore depends upon a further architectural requirement: consequences must become observable, their significance must be interpreted and the resulting knowledge must be capable of returning to those parts of the institution whose future cognition or behaviour can be altered.

Feedback should consequently be understood as more than the availability of information concerning previous outcomes. Institutions generate extensive quantities of post-decision information through implementation reports, performance indicators, audits, evaluations, public responses, judicial review, scientific monitoring and the everyday experience of operational personnel. Yet these observations become cognitively significant only when they enter structures capable of connecting consequences with previous assumptions, decisions and expectations. A system may collect evidence about its own performance continuously while learning very little if that evidence remains disconnected from policy design, organisational memory or future deliberation.

The architecture of feedback therefore determines which consequences become visible and where they become visible. Some effects appear rapidly and can be associated relatively easily with particular interventions, while others emerge gradually across long causal chains in which attribution becomes uncertain. Institutions may be highly sensitive to outcomes represented through established indicators while remaining comparatively insensitive to consequences that fall outside existing measurement frameworks. Feedback architecture thus influences not only whether institutions learn, but what they are capable of learning about.

Temporal alignment becomes particularly important. Evidence may arrive after the administrative or political window in which revision remains feasible, creating a condition in which knowledge exists but cannot influence the relevant decision cycle. Conversely, demands for rapid evaluation may encourage institutions to privilege immediate effects while overlooking slower structural consequences. Effective feedback architecture must therefore connect different temporal horizons, allowing short-term operational signals and longer-term evaluative knowledge to contribute appropriately to institutional learning.

Learning itself can occur at several depths. Some feedback leads to adjustments within existing procedures without challenging the assumptions underlying them. Other evidence may require modification of decision rules, organisational structures or the representations through which problems are understood. More consequential forms of learning may eventually alter the cognitive architecture itself by creating new sensing mechanisms, redistributing expertise, changing interfaces or establishing different forms of coordination. The relationship between feedback and learning therefore connects ICAM with the Evolutionary Layer of ICCA: repeated cognitive activity can gradually modify the architecture through which future cognition occurs.

An institution capable of systematic learning must consequently do more than preserve the memory of previous outcomes. It must possess **feedback structures that connect experience with mechanisms of interpretation and revision**. Where those connections are weak, institutional experience can accumulate without producing adaptation; where they are strong, action becomes part of a recursive cognitive process through which institutions progressively modify the conditions of their own future judgement.

14. Structural Architecture and Activated Cognitive Configurations

An institutional cognitive architecture should not be imagined as a fixed network in which every node and interface participates continuously with the same intensity. Institutions contain relatively persistent structures of authority, expertise, memory and communication, but the cognitive pathways mobilised at any particular moment depend upon the problem being addressed. Routine administration, strategic planning, scientific uncertainty, economic disruption and emergency response can activate substantially different configurations within the same institutional system.

ICAM therefore distinguishes between **structural cognitive architecture** and **activated cognitive configuration**. Structural cognitive architecture refers to the relatively persistent distribution of cognitive capacities, nodes, interfaces and relationships available within an institution. Activated cognitive configuration refers to the subset of those components and relationships that becomes mobilised around a particular problem, decision or period of institutional activity. The distinction allows the architecture to remain analytically stable enough to be mapped while recognising that actual cognition is dynamic and context-dependent.

This difference becomes visible whenever an institution encounters an unfamiliar problem. Expertise that remains peripheral during routine operations may become cognitively central when circumstances change. Scientific advisers may acquire greater influence during a public-health emergency; cybersecurity teams may move from specialised technical roles into strategic decision processes during a major digital incident; external experts may be incorporated temporarily when existing institutional knowledge proves insufficient. The underlying organisation may remain formally unchanged while the cognitive centre of gravity shifts substantially.

Activated configurations can also create temporary interfaces that do not exist during ordinary administration. Crisis committees, interdepartmental task forces, emergency data-sharing arrangements and accelerated advisory mechanisms may connect nodes that normally interact only indirectly. Such configurations can increase responsiveness by reducing informational distance and concentrating relevant expertise, but they can also bypass structures designed to preserve deliberation, accountability or informational diversity. The cognitive advantages of rapid reconfiguration must therefore be examined alongside its institutional consequences.

The distinction also reveals why formal institutional capacity may differ from realised capacity. An institution can possess extensive expertise that remains dormant because no mechanism activates it when relevant. Conversely, a small group of actors may become disproportionately influential because existing procedures repeatedly place them at the centre of activated configurations. Mapping only the structural architecture would therefore provide an incomplete picture of how cognition actually occurs.

This dynamic perspective opens an important pathway towards later simulation research. If structural architecture represents the space of available cognitive relationships, activated configurations represent the patterns through which that architecture becomes operational under particular conditions. **Institutional cognition can therefore be understood not only through the architecture an institution possesses, but through the configurations it is capable of activating when different environments demand different forms of collective intelligence.**

15. Temporal Dynamics and the Problem of Cognitive Synchronisation

Distributed cognition is organised not only across institutional space but across time. Different cognitive functions operate at different speeds, and the capacity of an institution to integrate them depends partly upon whether those temporal rhythms remain sufficiently aligned. Environmental sensing may produce information continuously, scientific analysis may require weeks or months, formal deliberation may follow predetermined procedural cycles, implementation may unfold over years and some consequences may become visible only after the institutional actors responsible for the original decision have changed.

This temporal heterogeneity creates a problem of **cognitive synchronisation**. Information that is accurate but arrives too late may become practically irrelevant; rapid analysis based upon incomplete evidence may influence decisions before more reliable knowledge becomes available; feedback may enter institutional memory without reaching the next generation of decision-makers; and political or administrative cycles may impose deadlines that differ substantially from the temporal requirements of scientific investigation. The quality of institutional cognition therefore depends partly upon the architecture through which processes operating at different speeds become coordinated.

The problem is particularly visible in environments characterised by accelerating change. Institutions designed around relatively stable informational cycles may struggle when technological, economic or geopolitical conditions evolve faster than established procedures can interpret them. Attempts to solve this problem exclusively through acceleration, however, may introduce different forms of cognitive weakness. Shortening deliberation can reduce opportunities for challenge and reflection; automated analysis can increase speed while introducing uncertainty concerning interpretation; continuous monitoring can overwhelm institutions with signals whose significance remains unclear.

Intelligent temporal architecture therefore requires more than making every process faster. Some functions benefit from rapidity, while others require time precisely because their cognitive contribution depends upon comparison, deliberation or verification. Institutional memory introduces still longer horizons by preserving knowledge across cycles of leadership and organisational change. The challenge is to create interfaces capable of connecting these different temporal scales without forcing them into artificial uniformity.

Anticipatory cognition further complicates the temporal structure because institutions must frequently reason about events that have not yet occurred. Forecasting, scenario development, strategic foresight and simulation allow possible futures to influence present decisions, effectively extending institutional cognition beyond retrospective learning. Such capacities become increasingly important when the consequences of delayed action are substantial or irreversible, but they also introduce uncertainty that must be represented carefully within decision processes.

ICAM therefore treats time as an architectural property rather than merely a contextual variable. **The effectiveness of institutional cognition depends not only upon whether relevant functions are connected, but upon whether they become connected at moments when their contributions can still influence collective judgement and action.** Cognitive synchronisation may consequently prove as important as cognitive connectivity when institutional architectures are later measured or simulated.

16. Crisis Cognition and Architectural Reconfiguration

Crises provide unusually revealing conditions for the study of institutional cognitive architecture because they place existing arrangements under pressures for which they may not have been designed. Information becomes incomplete or rapidly changing, established expertise may prove insufficient, decision horizons contract and the consequences of delay can increase dramatically. Under these conditions, institutions frequently reconfigure their cognitive architecture, altering which nodes become central, which interfaces remain active and how authority relates to expertise.

Some of these changes occur through formal emergency arrangements. Specialised crisis units may be activated, reporting procedures shortened and information-sharing mechanisms expanded. Other changes emerge informally as actors seek knowledge outside established channels or create temporary networks capable of responding more rapidly than conventional structures. The resulting cognitive configuration may differ substantially from routine governance even when the formal institution remains recognisably the same.

Crisis reconfiguration can reveal both hidden capacities and latent vulnerabilities. Informal networks that appeared peripheral may become essential sources of coordination, while procedures that function adequately under stable conditions may become severe bottlenecks when information volume increases. Highly centralised decision structures may provide coherence during initial response but become overloaded as complexity expands. Distributed systems may detect local changes rapidly while struggling to construct a common representation of the crisis. The crisis therefore functions as a form of architectural stress test in which relationships normally taken for granted become visible through their failure or adaptation.

Yet crisis cognition should not be evaluated solely by speed. Emergency conditions can create incentives to suppress uncertainty, narrow consultation and concentrate authority, producing rapid decisions at the cost of cognitive diversity or legitimacy. Temporary architectures can also become institutionally persistent after the conditions that justified them have disappeared. The study of crisis cognition must therefore distinguish between cognitive effectiveness under immediate pressure and the broader institutional consequences of architectural reconfiguration.

Another important question concerns whether institutions retain knowledge generated during crises. Temporary task forces may develop extensive expertise that disappears when they are dissolved; emergency data systems may be abandoned rather than incorporated into permanent infrastructure; successful informal relationships may never become part of organisational memory. Crisis experience does not automatically become institutional learning. The transition depends upon feedback and memory structures capable of converting temporary cognitive adaptations into durable knowledge where appropriate.

ICAM can therefore use crisis conditions as an analytical lens through which the flexibility of institutional cognitive architecture becomes visible. **An institution's cognitive capacity is partly reflected in whether it can reconfigure its architecture under changing conditions without losing the knowledge, coordination and legitimate authority required for coherent collective action.**

17. Metacognitive Architecture — Institutions Observing Their Own Cognition

The capacity to learn from the consequences of action introduces one form of institutional reflexivity, but a more developed cognitive architecture may also contain mechanisms through which the institution examines the processes by which its own knowledge and decisions are produced. This capacity can be described as **institutional metacognition**: the ability of an institutional system to generate representations of its own cognitive processes and use those representations to evaluate or modify how cognition is organised.

Metacognition differs from ordinary performance evaluation because its object is not only what the institution achieved but how the institution perceived, reasoned, coordinated and decided. A policy evaluation may conclude that an intervention failed to produce its intended outcome; a metacognitive analysis would additionally ask whether relevant information was available when the decision was made, whether uncertainty was represented adequately, whether dissenting expertise reached deliberation, whether institutional memory was consulted and whether the architecture itself contributed to the error.

Many existing institutional mechanisms already perform partial metacognitive functions even though they are rarely described in these terms. Audits examine procedural and informational processes; commissions of inquiry reconstruct decision pathways after major failures; evaluation units compare expectations with outcomes; parliamentary scrutiny can expose weaknesses in the use of evidence; judicial review may reveal deficiencies in reasoning or procedure; and internal learning exercises can identify failures of coordination. These mechanisms become metacognitively significant when their findings are capable of altering the architecture through which future cognition occurs.

The location of metacognition within the architecture is itself consequential. An institution evaluating its own decision processes may reproduce the assumptions that generated the original problem, while external scrutiny may provide independence at the cost of incomplete access to tacit organisational knowledge. Distributed metacognitive arrangements may therefore be preferable in some contexts, combining internal reflection with external challenge. The objective is not to create an omniscient supervisory node but to establish mechanisms through which cognitive processes can themselves become objects of institutional knowledge.

Artificial intelligence may eventually expand these capabilities by allowing institutions to analyse large volumes of decision records, trace information flows, identify recurring bottlenecks or compare predictions with subsequent outcomes. Such systems could support institutional self-observation at scales previously difficult to achieve, but their contribution would remain subject to the same architectural and normative constraints as other computational nodes. A system capable of detecting cognitive patterns does not determine which institutional values should guide their reform.

Metacognitive architecture therefore represents a higher-order form of institutional learning. Feedback allows institutions to learn about the consequences of their actions; metacognition allows them to learn about **the processes through which those actions became thinkable and actionable in the first place**. An institution capable of sustained metacognition may consequently modify not only policies and procedures but the architecture through which knowledge, judgement and authority become connected.

18. Cognitive Architecture and Legitimate Authority

The architectural analysis developed through ICAM makes visible a distinction that becomes increasingly important as institutional cognition grows more distributed and technologically augmented: the components that contribute most substantially to cognition need not be the components that possess legitimate authority to decide. Expertise, information-processing capacity and formal decision rights can occupy different locations within the same institution, and treating them as equivalent would obscure one of the fundamental characteristics of governance.

A scientific advisory body may possess greater knowledge of a particular problem than the political authority responsible for deciding how society should respond to it. An analytical unit may construct sophisticated assessments of policy alternatives without possessing the constitutional authority to select among them. Artificial intelligence may eventually exceed human actors in particular forms of prediction, retrieval or analytical synthesis while remaining entirely outside the legitimate structures through which collective decisions are authorised. **Cognitive contribution and legitimate authority therefore belong to different dimensions of institutional architecture.**

This distinction should not be interpreted as implying that authority can remain cognitively disconnected from expertise. Legitimate decision-making requires interfaces through which relevant knowledge can reach authorised actors in forms that preserve sufficient context, uncertainty and diversity for meaningful judgement. The architectural problem is consequently one of connection without substitution: expertise must become capable of informing authority without automatically acquiring it, while authority must remain sufficiently connected to cognitive resources to avoid becoming epistemically isolated.

Different institutional systems resolve this relationship through different configurations. Some establish formal advisory bodies whose recommendations remain explicitly non-binding; others embed technical expertise within administrative decision structures; regulatory institutions may delegate substantial discretion to specialised professionals; democratic systems combine expert knowledge with political accountability through multiple layers of representation and review. ICAM does not prescribe a universal arrangement, but it allows the cognitive consequences of these configurations to become more visible.

Artificial intelligence intensifies the importance of the distinction because computational capability can create pressures towards implicit delegation. If decision-makers routinely accept model recommendations that they cannot meaningfully interrogate, formal authority may remain human while effective cognitive influence shifts towards computational systems. Conversely, requiring humans to approve outputs mechanically does not necessarily restore meaningful authority if the surrounding architecture provides no practical capacity for independent judgement. The location of legitimate authority must therefore be examined together with the interfaces through which cognitive influence is exercised.

The relationship between cognition and authority also creates productive constraints. Democratic deliberation, legal procedure and institutional accountability may slow decision-making or require knowledge to pass through multiple interfaces, but these processes cannot be evaluated solely as cognitive inefficiencies. They embody normative commitments concerning who may decide, how reasons must be justified and how power remains contestable. ICAM must therefore avoid equating the most informationally efficient architecture with the most desirable institutional architecture.

The objective is instead to understand how **epistemic capability and legitimate authority can remain distinct while becoming sufficiently integrated for informed collective judgement**. This principle establishes an important boundary around ICAM and ensures that the study of institutional cognitive architecture remains embedded within the wider constitutional and normative foundations of Institutional Cognition.

19. Nested Cognitive Architectures and Multiple Institutional Scales

Institutions rarely operate as cognitively self-contained systems. Departments participate within organisations, organisations interact through governance networks, national institutions exchange knowledge through international structures and increasingly complex problems connect local, national and planetary forms of observation, expertise and decision-making. The cognitive architecture of any particular institution must therefore be understood partly through the wider architectures in which it is embedded.

This condition can be described through the concept of **nested cognitive architecture**. A team may function as a cognitive node within a department while simultaneously possessing its own internal cognitive architecture. A government agency may contain multiple distributed nodes while becoming a single specialised node within a broader governance network. International scientific bodies may connect national institutions into systems of distributed perception and knowledge integration whose cognitive functions cannot be located entirely within any participating organisation.

Nestedness complicates the boundaries of ICAM because the appropriate unit of analysis depends upon the research question. When investigating internal policy learning, a ministry may be treated as the principal cognitive system. When investigating national crisis response, the same ministry becomes one component within a wider architecture involving other agencies, scientific organisations and political authorities. At still larger scales, national governance systems may themselves become nodes within transnational or planetary networks addressing climate change, pandemics, technological risk or other problems whose informational and causal structures exceed territorial boundaries.

The interfaces connecting these levels become particularly significant because knowledge may change as it crosses institutional scales. Local observations can lose contextual detail when aggregated nationally; international scientific assessments may require translation into domestic policy categories; global indicators may shape local priorities even when they represent conditions imperfectly. Multi-level cognition therefore depends upon interfaces capable of preserving sufficient meaning while allowing knowledge to travel across heterogeneous institutional environments.

Nested architectures also distribute authority differently from cognition. Knowledge may be generated transnationally while legitimate decisions remain national or local; local institutions may possess detailed situational awareness while strategic coordination occurs elsewhere. The resulting architecture cannot be represented adequately through a single hierarchy because cognitive and authoritative relationships cross scales in different ways.

ICAM should therefore remain compatible with analysis at multiple levels without attempting to model planetary governance comprehensively within this Working Paper. That extension belongs principally to the later research on governance networks. What must be established here is the architectural principle: **institutional cognition is frequently nested within wider systems of cognition, and the boundaries of the institution do not necessarily coincide with the boundaries of the cognitive processes in which it participates.**

20. Towards a Dynamic Model of Institutional Cognitive Architecture

The preceding analysis allows ICAM to move beyond a static representation of institutional cognition as a collection of functions and components. Cognitive nodes possess differentiated capacities; interfaces determine how those capacities become connected; cognitive flows carry and transform information and knowledge; coordination structures integrate distributed contributions; bottlenecks constrain system-level capacity; redundancy and diversity provide resilience; feedback enables learning; activated configurations reorganise cognition around particular problems; temporal dynamics determine whether contributions arrive when they remain consequential; and metacognitive structures allow institutions to examine aspects of their own cognition.

Taken together, these relationships suggest that institutional cognitive architecture is better understood as a **dynamic configuration of interacting capacities than as a fixed organisational design**. The architecture contains relatively persistent structures, but cognition emerges through changing patterns of activation within them. Different problems mobilise different nodes, alter the centrality of interfaces and create different informational demands. Learning can modify the architecture over time, while crises may temporarily accelerate or restructure relationships that remain comparatively stable during routine governance.

This dynamic conception also helps avoid the temptation to search for a universal architecture of the intelligent institution. Institutional environments differ too substantially in scale, purpose, complexity, legitimacy requirements and exposure to uncertainty for one configuration to optimise every form of cognition. Architectures should instead be investigated in relation to the problems they must address and the trade-offs their structures create. A configuration optimised for rapid emergency response may be poorly suited to long-term deliberative governance; an architecture capable of integrating enormous informational diversity may require coordination processes inappropriate for routine administration.

ICAM therefore provides a language for representing relationships rather than a blueprint for institutional design. Its scientific value will depend upon whether those relationships can eventually support comparison, measurement and simulation without reducing the complexity of institutions to an artificially closed system. The model must remain sufficiently abstract to travel across institutional contexts while sufficiently structured to generate researchable questions concerning connectivity, bottlenecks, activation, redundancy, temporal alignment and learning.

The transition towards such operational investigation will require the architecture to become progressively more explicit. Researchers will need to determine which elements can be observed, how different architectural configurations might be compared and which properties provide meaningful indicators of institutional cognition. Later Working Papers will address those questions directly. At this stage, the principal contribution of ICAM is to establish the analytical object from which they can proceed.

Institutional cognitive architecture can therefore be understood as the evolving configuration through which distributed cognitive capacities become connected, activated and coordinated across time, enabling institutions to transform information into knowledge, knowledge into judgement, judgement into legitimate action and experience into the conditions of future cognition.

21. Representing Institutional Cognitive Architecture

The transition from conceptual description towards operational research requires some means of representing institutional cognitive architecture without mistaking the representation for the institution itself. ICAM has so far identified functions, nodes, interfaces, cognitive flows, coordination structures, feedback mechanisms and dynamic configurations as analytically relevant components, but these elements become scientifically useful only when researchers can describe their relationships with sufficient consistency to support comparison across cases. Representation therefore constitutes an intermediate methodological problem: the architecture must become visible enough to investigate while remaining open enough to preserve institutional complexity.

A first representation might resemble a network in which cognitive nodes are connected through interfaces and differentiated according to the functions they perform. Such a representation would immediately reveal features that conventional organisational charts tend to obscure, including nodes that participate in several cognitive functions, interfaces that cross formal administrative boundaries and components that occupy cognitively central positions despite possessing limited formal authority. Yet network representation alone would remain insufficient because institutional relationships differ qualitatively. A reporting channel, a deliberative forum, a shared database and a human–AI interaction mechanism may all connect nodes while performing substantially different cognitive roles.

ICAM therefore requires a multidimensional representation in which relationships can be characterised not only by their existence but by their function. An interface may transmit information, retrieve memory, integrate interpretations, coordinate action or return feedback. Some relationships may operate continuously, while others activate only under particular conditions. Directionality may also vary: sensing systems primarily transmit observations towards analytical processes, while feedback structures return information from implementation towards evaluation and learning. Other interfaces, particularly deliberative ones, may involve reciprocal exchanges in which representations are repeatedly modified.

Temporal properties add another dimension. Two nodes may be strongly connected in organisational terms while remaining cognitively misaligned because their operating cycles differ. A scientific advisory process that produces rigorous analysis after a decision window has closed may be structurally connected but temporally ineffective. Representation must therefore eventually incorporate not only connectivity but latency, activation conditions and temporal synchronisation.

The same caution applies to cognitive content. A map showing that information moves from one component to another says little about whether meaning is preserved during transmission. Representations may eventually need to distinguish between raw observations, structured information, contextualised knowledge, analytical judgement, authorised decisions and feedback, because the transformation occurring at an interface may be as significant as the connection itself. A highly connected architecture can remain cognitively weak when information is repeatedly distorted, excessively compressed or detached from its original uncertainty.

The purpose of representation within ICAM is consequently not to produce a complete digital replica of institutional cognition. It is to establish a disciplined abstraction capable of making relationships visible and generating questions that can later be examined empirically. **A useful cognitive map should reveal not only where institutional capacities are located, but how they become connected, transformed, activated and constrained within the wider architecture.**

22. Architectural Properties of Institutional Cognition

Once institutional cognitive architectures can be represented consistently, comparison becomes possible at the level of architectural properties rather than merely organisational form. Institutions that appear administratively different may share important cognitive characteristics, while organisations with similar formal structures may differ substantially in the way knowledge moves through them. ICAM therefore requires a vocabulary for describing properties that emerge from relationships across the architecture.

Connectivity represents one such property, although its meaning must be treated carefully. Greater connectivity can increase access to distributed knowledge and reduce informational isolation, but excessive connectivity can generate cognitive overload, coordination costs and diffusion of responsibility. The relevant question is therefore not how connected an institution is in the abstract, but whether cognitively relevant nodes possess appropriate pathways to the functions with which they must interact.

Integration describes a related but distinct property. An architecture may contain many connections while remaining weakly integrated if those connections do not enable heterogeneous knowledge to contribute to shared representations or coordinated judgement. Integration concerns the capacity of distributed cognitive contributions to become mutually consequential. It therefore depends upon the quality of interfaces as much as their number.

Diversity concerns the range of knowledge, perspectives and interpretative approaches available within the architecture. As previously established, diversity can reduce collective blind spots and strengthen reasoning under uncertainty, but its benefits depend upon mechanisms capable of preserving difference while enabling coordination. An architecture that eliminates diversity prematurely may become coherent but epistemically fragile, while one that preserves unlimited diversity without integration may struggle to produce collective action.

Modularity provides another potentially significant property. Cognitive functions may be organised into partially autonomous subsystems capable of performing specialised tasks without requiring continuous coordination with the entire institution. Modularity can reduce complexity and contain failures, but excessive separation may create informational silos. The cognitive value of modularity therefore depends upon the interfaces connecting specialised components to the wider architecture.

Redundancy, centrality, latency and adaptability provide further dimensions through which architectures may eventually be described. Centrality can identify nodes or interfaces whose position gives them disproportionate influence over cognitive flows; latency concerns the time required for relevant information or feedback to travel through the architecture; redundancy concerns the presence of alternative pathways or overlapping capacities; and adaptability concerns the capacity of the architecture to modify its configuration when circumstances change.

These properties should not yet be treated as formal indicators. Their operational definitions, measurement procedures and empirical validity belong to later research. Within ICAM their purpose is conceptual: they identify **architectural characteristics that may help explain why similarly resourced institutions display different cognitive behaviours**. The transition from properties to indicators will require substantial methodological work, but the properties themselves provide an important bridge between architectural theory and future measurement.

23. Cognitive Topologies and Institutional Variation

The possibility of comparing architectural properties raises a broader question concerning whether recurring forms of institutional cognitive organisation can be identified across different contexts. Institutions vary enormously in purpose, scale and constitutional structure, yet their cognitive architectures may display patterns that can be described as different **cognitive topologies**: recurrent configurations through which distributed functions and nodes become organised.

A strongly centralised topology might concentrate knowledge integration and decision authority within a limited number of nodes, with information moving predominantly through hierarchical interfaces. Such architectures can support coherent action and clear responsibility, particularly where tasks are stable and authority must be exercised rapidly, but they may become vulnerable to bottlenecks, informational compression and failures concentrated at cognitively central nodes.

Distributed topologies may instead locate sensing, expertise and reasoning across multiple semi-autonomous components connected through horizontal interfaces. These arrangements can preserve local knowledge and increase cognitive diversity, particularly when institutional environments are heterogeneous, but they place greater demands upon coordination and may struggle to produce shared representations when perspectives diverge.

Networked topologies may connect specialised nodes through overlapping relationships that do not correspond neatly to formal organisational boundaries. Knowledge can travel through multiple pathways, creating resilience and opportunities for cross-domain integration, while simultaneously increasing complexity concerning accountability and authority. Hybrid architectures may combine these forms, centralising some functions while distributing others according to the cognitive requirements of particular tasks.

These distinctions should not become rigid institutional categories. Most complex institutions contain several topologies simultaneously, and activated cognitive configurations may temporarily alter their relative importance. A decentralised administrative system may centralise crisis cognition; a hierarchical ministry may rely upon highly networked scientific expertise; an international organisation may combine distributed sensing with concentrated analytical synthesis. Cognitive topology therefore describes patterns within architecture rather than fixed organisational identities.

The concept is useful precisely because it allows comparative research to move beyond conventional classifications of institutions. Two organisations from different political systems might exhibit similar cognitive topologies in particular functional domains, while two agencies within the same administrative tradition may organise knowledge in fundamentally different ways. Future empirical research could therefore investigate whether particular topological characteristics are associated with different forms of cognitive strength or vulnerability under specific environmental conditions.

ICAM remains deliberately non-prescriptive at this stage. No topology should be assumed to represent a universally superior architecture of intelligence. **The scientific question concerns how different configurations shape cognition under different conditions, not which institutional form should be declared cognitively optimal in advance.**

24. Architectural Trade-offs and the Absence of a Universal Optimum

The analysis of cognitive topologies reinforces a principle that should remain central to ICAM: institutional cognitive architecture is characterised by trade-offs rather than by a single hierarchy of superior and inferior designs. Properties that strengthen cognition under one set of conditions may weaken it under another, and attempts to maximise individual architectural characteristics can create unintended consequences elsewhere within the system.

Centralisation illustrates this tension clearly. Concentrating knowledge integration and decision authority can reduce coordination costs and accelerate collective action, yet the same concentration may create bottlenecks, increase dependence upon a small number of nodes and suppress information that does not fit dominant representations. Decentralisation can expand local responsiveness and preserve diversity while making system-level coordination more difficult. The appropriate balance depends upon the complexity of the problem, the distribution of relevant knowledge, the consequences of delay and the institutional requirements of legitimate authority.

Speed creates a similar trade-off. Rapid information flows can improve responsiveness, particularly during crises, but acceleration can reduce opportunities for verification, deliberation and contextual interpretation. Slower processes may support deeper reasoning while becoming ineffective when environments change before decisions are reached. Institutional intelligence therefore cannot be equated with processing speed; temporal architecture must align different cognitive rhythms with the demands of the problem.

Redundancy and efficiency generate another tension. Eliminating overlapping functions can reduce cost and simplify coordination, yet it may also remove alternative sources of knowledge and create single points of failure. Increasing redundancy can strengthen resilience while generating duplication and informational complexity. Diversity similarly expands the range of possible interpretations while increasing the cognitive effort required to integrate them.

Transparency can itself involve architectural trade-offs. Extensive documentation and traceability can support accountability, memory and metacognition, yet excessive procedural requirements may increase cognitive burden and slow informal knowledge exchange. Highly flexible architectures can adapt rapidly while becoming difficult to audit; highly formalised architectures can preserve consistency while reducing the capacity for improvisation when novel problems arise.

The existence of these tensions means that cognitive architecture should be evaluated relative to environmental and institutional requirements rather than against an abstract ideal. **An intelligent architecture is not one that maximises every desirable property, but one capable of sustaining appropriate relationships among competing cognitive requirements while remaining adaptable as those requirements change.** This proposition will become particularly important when later research attempts to develop performance indicators, because measurement systems that reward individual properties without recognising architectural trade-offs could encourage precisely the forms of imbalance that ICAM seeks to make visible.

25. The Problem of Architectural Fit

If no universal cognitive architecture exists, the relationship between institutional configuration and environment becomes a central object of analysis. An architecture that performs effectively under stable conditions may struggle in environments characterised by rapid change, uncertainty or high informational diversity. Conversely, structures designed for continuous crisis response may impose unnecessary cognitive costs when problems are repetitive and predictable. ICAM describes this relationship through the provisional concept of **architectural fit**.

Architectural fit concerns the degree to which the distribution and coordination of cognitive capacities correspond to the informational, temporal and decision requirements of the environment in which an institution operates. Institutions facing geographically heterogeneous conditions may require distributed sensing and substantial local knowledge, while institutions responsible for highly interconnected systemic risks may require stronger mechanisms of central integration. Problems characterised by deep uncertainty may benefit from cognitive diversity and iterative learning, whereas stable administrative functions may be performed effectively through more standardised architectures.

Fit is not static because environments change. Technological transformation can alter the speed and volume of information available to institutions; geopolitical changes can create new forms of uncertainty; demographic and environmental developments can generate problems that cross administrative boundaries established for earlier conditions. Cognitive architectures that were historically effective may therefore become progressively misaligned without experiencing any obvious internal failure.

This possibility is particularly important because institutions often evaluate themselves through measures derived from their existing architecture. A system may continue to perform efficiently according to established indicators while becoming increasingly incapable of perceiving emerging problems that those indicators were never designed to detect. Architectural misfit can consequently remain invisible until disruption exposes the limitations of existing cognitive arrangements.

Learning and metacognition provide mechanisms through which fit can be reassessed. Institutions capable of examining not only their outcomes but the adequacy of their cognitive processes may detect when established sensing systems, knowledge structures or decision interfaces no longer correspond to environmental demands. Architectural adaptation can then involve modifying relationships rather than merely increasing resources: creating new interfaces, decentralising particular functions, strengthening integration across previously separated domains or incorporating new forms of expertise.

The concept of architectural fit therefore connects ICAM with institutional evolution without implying that adaptation always produces improvement. Institutions can respond to environmental pressure in ways that create new vulnerabilities, and short-term adjustments may undermine longer-term cognitive capacity. **Fit must be understood as a dynamic relationship between architecture and environment rather than as a permanent state that an institution can finally achieve.**

26. Path Dependence and the Historical Formation of Cognitive Architecture

Institutional cognitive architectures are rarely designed from first principles. They emerge through historical processes in which organisational structures, legal responsibilities, professional communities, technologies and accumulated responses to previous problems become layered over time. Understanding contemporary architecture therefore requires recognising that many of its characteristics reflect historical trajectories rather than deliberate optimisation.

Path dependence can influence cognition in several ways. Information systems may preserve categories developed for earlier administrative purposes; organisational boundaries may separate domains that have subsequently become deeply interconnected; professional traditions may shape which forms of evidence receive credibility; and decision procedures created under previous technological conditions may continue to determine how contemporary knowledge reaches authority. Institutional memory itself can preserve not only useful knowledge but inherited assumptions about what information matters and how problems should be represented.

Historical layering can produce architectures containing components from different periods whose relationships were never designed as a coherent cognitive system. A contemporary digital monitoring platform may feed information into administrative procedures developed decades earlier; artificial intelligence may be introduced into organisations whose decision structures remain designed around scarcity of analysis rather than abundance; networked scientific expertise may coexist with highly hierarchical reporting arrangements. The resulting architecture can function through combinations of old and new mechanisms whose cognitive consequences are difficult to infer from formal design alone.

Path dependence does not imply immobility. Institutions continuously adapt, but adaptation usually begins from existing structures and therefore tends to modify rather than replace them. New interfaces are added to established networks, new technologies become integrated into inherited workflows and new forms of expertise must find positions within existing authority structures. Architectural evolution is consequently constrained by what already exists.

This historical perspective provides an important safeguard against treating cognitive architecture as an engineering problem that can simply be redesigned around abstract principles of efficiency. Institutions embody legal settlements, political compromises, professional identities and accumulated forms of legitimacy whose significance extends beyond cognitive performance. Changes intended to improve information flow or coordination may therefore encounter constraints that are not reducible to technical design problems.

ICAM should consequently understand architectural analysis as both structural and historical. **The cognitive architecture observed at any particular moment represents a temporary configuration produced through previous rounds of institutional development**, and its strengths and weaknesses can often be understood only by reconstructing how particular relationships came to exist. This proposition prepares an important connection with later comparative and civilisational research, where the historical evolution of governance cognition becomes a central object of investigation.

27. Human, Organisational and Computational Nodes within a Shared Architecture

The increasing incorporation of artificial intelligence into institutional environments makes it necessary to clarify how heterogeneous cognitive nodes can be represented within a common architectural model without implying that they possess equivalent forms of cognition, agency or authority. ICAM addresses this problem functionally: nodes are included because they contribute to particular cognitive processes, while their ontological, normative and institutional characteristics remain distinguishable.

Human nodes contribute forms of cognition shaped by experience, tacit knowledge, interpretation, social understanding and normative judgement. Their capacities are limited by attention, memory and processing constraints, yet they remain capable of contextual reasoning and forms of responsibility that cannot simply be inferred from computational performance. Organisational nodes, by contrast, emerge through coordinated groups whose procedures, shared knowledge and institutional continuity allow cognitive functions to persist beyond individual participants.

Computational nodes contribute different capabilities. Databases preserve and retrieve information at scales inaccessible to unaided human memory; statistical systems identify patterns across large datasets; simulations explore possible system behaviour; and artificial intelligence can increasingly classify, synthesise, predict and generate representations relevant to institutional reasoning. These capacities can substantially alter the architecture by changing the speed, scale and distribution of cognitive activity.

The importance of computational nodes nevertheless depends upon their interfaces with the wider institution. A highly capable model whose outputs cannot be interpreted, challenged or integrated into legitimate deliberation may increase nominal analytical capacity while contributing little to collective intelligence. Conversely, comparatively narrow computational systems can become cognitively significant when embedded within workflows that connect their outputs effectively to human expertise and institutional memory.

Hybrid configurations therefore deserve particular attention. Many future cognitive functions are unlikely to be performed exclusively by humans or artificial systems; they will emerge through interaction between them. Analysts may use AI to explore evidence while retaining responsibility for interpretation, decision-makers may interrogate computational scenarios during deliberation and knowledge systems may combine automated retrieval with expert contextualisation. The cognitive node in such cases may itself become a socio-technical configuration whose capabilities arise from interaction.

ICAM establishes only the architectural foundations for this analysis. The detailed study of human–AI cognitive distribution belongs to the dedicated L02 research on human–AI collaboration. What matters here is the principle that **institutional cognitive architecture can incorporate heterogeneous nodes without collapsing the distinctions between capability, agency, responsibility and legitimate authority**.

28. Architectural Failure as a Distinct Form of Institutional Failure

The framework developed through ICAM makes it possible to identify a category of institutional weakness that may remain obscured when analysis focuses primarily upon resources, individual competence or formal organisational design. **Architectural failure** occurs when the relationships through which cognitive capacities should become integrated prevent the institution from using capabilities that are otherwise present within the system.

Such failures can take many forms. Relevant information may be collected but fail to reach analytical processes; expertise may exist but remain isolated from deliberation; institutional memory may preserve previous lessons without making them retrievable at the point of decision; feedback may identify recurrent problems without reaching mechanisms capable of revising policy. In each case, the institution possesses at least part of the cognitive capacity required to respond effectively, but its architecture prevents that capacity from becoming collectively consequential.

Interface failure constitutes one important form of architectural failure, but the category is broader. Bottlenecks can concentrate excessive cognitive demand within particular nodes; poorly aligned temporal cycles can cause accurate knowledge to arrive too late; excessive centralisation can suppress distributed perception; excessive fragmentation can prevent specialised knowledge from becoming integrated; and insufficient redundancy can allow the failure of a single component to disrupt entire cognitive pathways.

Architectural failure can also emerge through interactions among individually reasonable arrangements. A reporting procedure designed to prevent information overload may combine with hierarchical filtering and limited deliberative time to remove precisely the uncertainty that senior decision-makers need to understand. A highly specialised division of expertise may improve local analytical quality while making cross-domain problems increasingly difficult to interpret. Institutional failure may therefore emerge not because any component is obviously defective but because their relationships generate system-level consequences that were not anticipated.

This perspective has significant implications for institutional reform. Responses to failure often concentrate upon replacing personnel, increasing resources, creating additional units or introducing new technologies. Such interventions may produce limited benefits when the underlying problem concerns architecture. Adding more data to a system constrained by analytical bottlenecks, for example, may intensify rather than resolve the failure. Effective reform may instead require changing interfaces, redistributing cognitive functions or modifying the pathways through which knowledge reaches authority.

Architectural analysis therefore complements rather than replaces conventional explanations of institutional failure. Institutions can fail because of inadequate resources, conflicting interests, political incentives, corruption, legal constraints or many other factors that cannot be reduced to cognition. ICAM identifies an additional explanatory possibility: **institutions may fail cognitively not because they lack relevant capacities, but because their architecture prevents those capacities from functioning as a coherent collective system.**

29. From Architectural Failure to Architectural Diagnosis

Recognising architectural failure creates the possibility of a more systematic form of institutional diagnosis. Rather than beginning with the assumption that poor outcomes reveal insufficient capability, researchers could examine how cognitive functions were distributed and connected during the processes that produced those outcomes. Diagnosis would therefore reconstruct the architecture through which information became knowledge, knowledge became judgement and judgement became action.

Such reconstruction would need to distinguish between failures of nodes and failures of relationships. Did relevant information exist but remain inaccessible? Was expertise available but excluded from deliberation? Did decision-makers receive evidence without sufficient representation of uncertainty? Did feedback reveal the consequences of previous actions but fail to alter subsequent policy? Each question identifies a different architectural mechanism and therefore suggests different explanations.

Diagnosis would also need to consider activated configurations rather than relying exclusively upon formal structures. The architecture that actually produced a decision may differ substantially from the institution's nominal design. Informal advisers may have become central, formal committees may have played little substantive role and emergency procedures may have bypassed normal knowledge pathways. Reconstructing cognition therefore requires attention to what the institution actually activated around a problem.

Temporal analysis is equally important. A decision may appear poorly informed when judged retrospectively even though relevant knowledge became available only later. Conversely, information may have existed early enough to influence action but remained trapped within slow institutional pathways. Diagnostic analysis must therefore reconstruct not only who knew what, but **when particular forms of knowledge became available and when they became institutionally consequential**.

The objective of architectural diagnosis should not be to produce an illusion that complex institutional outcomes can be attributed to a single cognitive mechanism. Governance decisions emerge through interactions among knowledge, interests, authority, law and political judgement. Architectural diagnosis contributes by making one dimension of those interactions more explicit and by identifying where cognitive processes may have constrained the range or quality of choices available.

This diagnostic potential provides one of the strongest bridges between ICAM and applied institutional research. If cognitive architecture can be reconstructed consistently, later research may compare successful and unsuccessful decision processes, identify recurring patterns of bottleneck or interface failure and eventually test whether targeted architectural interventions alter institutional performance. **ICAM thus begins to transform Institutional Cognition from a language for describing collective intelligence into a framework capable of investigating why that intelligence sometimes fails to emerge.**

30. The Boundaries of ICAM

The increasing explanatory reach of architectural analysis makes it particularly important to define what ICAM does not claim to explain. Institutions are cognitive systems only in a partial analytical sense: they are simultaneously political, legal, economic, social and historical formations shaped by interests, conflict, authority and cultural meaning. A model of institutional cognitive architecture can illuminate how knowledge becomes organised within these systems, but it cannot substitute for theories addressing the other dimensions through which governance operates.

Architectural explanations should therefore not be used automatically when institutional behaviour differs from what available evidence might appear to support. Decision-makers may possess accurate information and understand its implications while legitimately choosing among competing values in ways that cannot be determined by evidence alone. Political actors may reject expert recommendations because they represent different priorities rather than because the cognitive architecture failed. Institutions may also maintain procedures that appear cognitively inefficient because those procedures protect rights, distribute authority or preserve accountability.

Power creates a particularly important boundary. Cognitive architecture can reveal which actors possess access to information and which nodes occupy influential positions within knowledge flows, but the reasons those arrangements exist may require political and historical explanation. Exclusion from a cognitive interface may reflect architectural oversight, but it may also represent deliberate control over whose knowledge becomes institutionally recognised. ICAM can make such structures visible without claiming that their causes are purely cognitive.

The model must also avoid excessive functionalism. The fact that an institutional component can be described as performing a cognitive function does not imply that it was designed for that purpose or that its contribution benefits the institution as a whole. Informal networks may facilitate knowledge exchange while undermining formal accountability; redundancy may increase resilience while preserving inefficient organisational arrangements; metacognitive mechanisms may identify errors while being used selectively according to political incentives. Cognitive function and institutional desirability therefore remain distinct.

Another boundary concerns representational completeness. Any cognitive map necessarily simplifies relationships that are fluid, contested and partially invisible. Tacit knowledge, informal influence and rapidly changing configurations may resist systematic representation. The apparent precision of a model should therefore never be mistaken for exhaustive knowledge of the institution it describes.

ICAM is consequently best understood as a **middle-range analytical framework**: sufficiently structured to generate researchable propositions, but deliberately limited in its claim to explain institutional reality as a whole. Its contribution lies in making the architecture of collective cognition visible while preserving the conceptual boundaries established within ICCA between intelligence, governance and legitimate authority.

31. ICAM as a Research Framework

The development of ICAM transforms a set of theoretical propositions into an increasingly structured research agenda. If institutional cognition depends partly upon the configuration of functions, nodes, interfaces, flows, coordination structures and feedback mechanisms, then institutional comparison can begin to examine how those configurations differ and whether recurring architectural characteristics correspond to different patterns of cognitive behaviour.

The first methodological task will be descriptive reconstruction. Researchers must develop procedures for identifying cognitive functions within real institutions, locating the nodes through which they are performed and tracing the interfaces connecting them. Documentary analysis, organisational records, interviews, process tracing, digital interaction data and direct observation may each reveal different aspects of architecture. No single method is likely to capture the whole system because formal and informal cognition leave different empirical traces.

A second task concerns comparative analysis. Once architectures can be reconstructed with sufficient consistency, researchers can examine variation across institutions, policy domains and historical periods. Comparative work may reveal whether certain bottlenecks recur, whether particular forms of redundancy strengthen resilience or whether different topologies become advantageous under particular environmental conditions. Such findings could progressively refine the conceptual distinctions introduced within ICAM.

Measurement represents a later methodological stage. Architectural properties such as connectivity, latency, centrality, diversity, redundancy and adaptability may eventually support indicators, but their operationalisation must remain theoretically disciplined. A quantity that can be measured easily is not necessarily a valid representation of the cognitive property it is intended to capture. The dedicated performance-indicator research within L02 will therefore inherit from ICAM a set of candidate properties rather than a completed measurement system.

Simulation offers another pathway. Representing nodes, interfaces, flows and activation conditions computationally may allow researchers to explore how alternative configurations behave under different informational environments or disruptions. Simulation could help identify bottlenecks, examine the consequences of centralisation or test the resilience of alternative pathways before institutional interventions are attempted. Yet simulated behaviour will remain dependent upon assumptions that require empirical grounding.

Applied experimentation provides the most direct route towards testing architectural interventions. Institutions might modify particular interfaces, introduce new feedback mechanisms or redesign knowledge flows and then observe whether cognitive outcomes change. Such experimentation must proceed cautiously because institutional architectures are embedded within legitimate structures of authority and real governance processes. Nevertheless, carefully designed pilots could eventually provide evidence concerning whether specific architectural changes produce the effects predicted by the theory.

ICAM therefore functions neither as a finished operational model nor as a purely conceptual description. **It establishes a research framework through which institutional cognitive architecture can progressively become observable, comparable, measurable, simulatable and experimentally investigable while remaining connected to the theoretical foundations from which the model emerged.**

32. ICAM within the Operational Research Programme

The Institutional Cognitive Architecture Model occupies a strategically important position within the wider Operational Research Programme because it translates the conceptual architecture established by ICCA into a model centred upon the institution as an organised cognitive system. Whereas ICCA provided the theoretical coordinates through which the field of Institutional Cognition could remain conceptually integrated, ICAM begins to specify the structures through which cognition becomes distributed, connected and dynamically realised within institutional environments. The transition is therefore not simply from one model to another, but from the organisation of theoretical knowledge towards the investigation of institutional mechanism.

This position gives ICAM a particularly strong relationship with the Working Papers that follow. Computational policy design requires an explicit understanding of where policy-relevant cognition occurs, how evidence becomes integrated and how decision processes connect to feedback. Human–AI collaboration requires a model capable of identifying the cognitive functions and interfaces through which artificial systems may participate without conflating contribution with authority. Measurement requires candidate properties of cognitive architecture that can eventually become observable, while simulation requires sufficiently explicit representations of nodes, flows, activation conditions and bottlenecks to explore alternative configurations. Applied pilots, in turn, require architectural hypotheses precise enough that interventions can target particular mechanisms rather than merely introduce new technologies or organisational reforms without a clear account of what is expected to change.

ICAM therefore acts as an architectural bridge across several Research Series without becoming a substitute for them. Its purpose is not to anticipate the detailed work of computational modelling, indicator construction or experimental design, but to provide a sufficiently coherent representation of the institutional system from which those specialised approaches can proceed. The value of this role will become increasingly visible as the L02 generation develops, because each subsequent model will be able to inherit a common vocabulary concerning cognitive functions, nodes, interfaces, flows, feedback, activation and architectural fit while refining those concepts according to its own methodological requirements.

The relationship is equally important in the opposite direction. Later Working Papers will almost certainly expose weaknesses or ambiguities within ICAM. Measurement may reveal that some architectural properties cannot be observed reliably; simulation may show that apparently distinct mechanisms interact in unexpected ways; comparative research may demonstrate that certain configurations behave differently across institutional contexts; and pilot programmes may reveal that interventions affecting one interface produce consequences elsewhere in the architecture. ICAM should therefore remain open to revision as these findings accumulate, preserving the same versioned logic established through ICCA while operating at a more specific analytical level.

Within the Operational Research Programme, ICAM thus represents **the first sustained attempt to convert the idea of distributed institutional cognition into an**

explicit architectural research object, creating a shared foundation from which the programme can move progressively towards comparison, measurement, simulation and experimentation.

33. Progressive Validation of Institutional Cognitive Architecture

The scientific value of ICAM ultimately depends upon whether the architectural distinctions proposed in this Working Paper help explain observable differences in how institutions process knowledge, coordinate judgement and adapt over time. Because ICAM is a middle-range framework rather than a single empirical hypothesis, validation will necessarily occur progressively across multiple methods rather than through one decisive test. The relevant question is not whether every element of the model can be confirmed simultaneously, but whether the concepts and relationships identified here generate patterns that remain meaningful when applied to real institutional systems.

The first stage of validation is descriptive. Researchers should be able to reconstruct cognitive architecture with sufficient consistency that different analysts examining the same institution identify broadly comparable functions, nodes and interfaces. If the model depends entirely upon subjective interpretation, its comparative value will remain limited. Descriptive validation therefore requires clearer methodological protocols concerning how architectural components are identified, how informal relationships are represented and how activated configurations are distinguished from persistent structures.

A second stage concerns comparative validity. If ICAM captures meaningful properties of institutional cognition, then differences in architecture should correspond to differences in the way institutions acquire, integrate and use knowledge under comparable conditions. Institutions with highly fragmented feedback structures, for example, might exhibit recurring difficulties in translating evaluation into adaptation, while architectures containing severe decision bottlenecks might display characteristic patterns of delay or overload. These expectations remain hypotheses rather than conclusions, but comparative research can begin determining whether the proposed relationships possess explanatory value beyond individual cases.

Temporal analysis provides another important validation pathway. Institutional architectures are not static, and some of the strongest evidence for their significance may appear when configurations change. Organisational reforms, crises, technological adoption or personnel transitions can alter interfaces and cognitive flows, creating opportunities to observe whether corresponding changes occur in collective cognition. Longitudinal research may therefore help distinguish enduring architectural effects from correlations produced by particular environments.

Simulation can complement empirical comparison by examining the internal logic of proposed mechanisms. A model of information flow may explore how increasing latency affects decision quality, while another may test how redundancy changes resilience under disruption. Such simulations cannot validate ICAM independently because their results remain conditional upon assumptions embedded in the model, yet they can expose contradictions, identify sensitive relationships and generate hypotheses that empirical research can subsequently investigate.

Applied institutional experiments may provide the strongest evidence concerning particular architectural mechanisms. If an institution modifies a feedback interface, redesigns an analytical workflow or introduces a new mechanism for integrating distributed expertise, researchers can examine whether the resulting changes correspond to the predictions of ICAM. Such experiments will necessarily remain context-dependent, but replication across different institutions could eventually help distinguish mechanisms with broader explanatory potential from those tied to specific organisational environments.

Progressive validation should therefore be understood as an accumulation of converging and diverging evidence across methods. **ICAM will become scientifically stronger not when every institutional case fits the model perfectly, but when the model makes sufficiently precise distinctions that researchers can identify where, why and under what conditions those distinctions succeed or fail.**

34. Methodological Cautions in Architectural Research

The effort to reconstruct institutional cognitive architecture introduces methodological difficulties that should remain visible throughout future research. One of the most significant arises from the partial invisibility of cognition itself. Formal procedures leave documentary traces, but many of the relationships through which knowledge becomes institutionally consequential are informal, tacit or activated only under particular conditions. Interviews may reveal some of these pathways, yet retrospective accounts are affected by selective memory, organisational narratives and the tendency to reconstruct decisions more coherently than they occurred.

The distinction between formal and realised architecture is therefore especially important. Official documentation may describe reporting procedures that function very differently in practice, while informal networks can become cognitively central without appearing in organisational records. Architectural research will likely require methodological triangulation, combining documents, interviews, observational evidence, process tracing and, where available, digital interaction data to reconstruct the pathways through which cognition actually moved.

Another difficulty concerns causal attribution. Poor institutional outcomes may coincide with identifiable bottlenecks or interface failures without those architectural properties necessarily being the primary cause. Political conflict, resource constraints, external shocks or legal limitations may dominate the outcome while cognitive architecture plays only a secondary role. Researchers must therefore avoid interpreting every failure through the architecture simply because ICAM makes architectural mechanisms visible.

Measurement itself can create additional problems. Once properties such as centrality, latency or connectivity become quantifiable, there may be a temptation to treat them as direct indicators of cognitive quality. Yet their meaning depends upon context. High centrality may represent effective integration in one institution and dangerous concentration in another; low latency may indicate responsiveness or insufficient deliberation; high connectivity may expand knowledge exchange while generating overload. Quantification can therefore increase precision without eliminating the need for theoretical interpretation.

The analysis of AI-enabled architecture introduces further caution. Computational systems may leave detailed logs that make some cognitive flows easier to observe than human or informal processes, creating an asymmetry in available evidence. Researchers could consequently overestimate the importance of computational nodes simply because their activity is more measurable. Architectural analysis must therefore preserve attention to forms of tacit expertise, professional judgement and informal coordination that remain difficult to capture digitally.

Finally, the act of mapping architecture may itself affect institutional behaviour. Actors who become aware that their information flows or decision roles are being analysed may alter those relationships, while institutions may resist representations that expose informal dependencies or concentrations of influence. Architectural research therefore contains both methodological and political sensitivities that must be handled carefully.

These cautions do not diminish the value of ICAM. They clarify the conditions under which architectural analysis can become scientifically credible. **The model is most useful when it increases visibility without creating the illusion that institutional cognition can ever be mapped completely or interpreted independently of the wider contexts through which governance operates.**

35. ICAM and the Emerging Operational Model Suite

The development of ICAM also clarifies the possibility, introduced in the previous Working Paper, that the L02 generation may gradually produce a family of interoperable models rather than a collection of isolated frameworks. ICCA provides the conceptual architecture of the field, while ICAM represents the first major translation of that architecture into a model centred upon the institutional system itself. Subsequent Working Papers can then select particular dimensions of ICAM and develop them further according to their own research objects.

Computational policy design will inherit the distinction between cognitive functions, interfaces and feedback structures while applying them to policy processes. Human–AI collaboration will examine how human and computational nodes can be configured within the same architecture, particularly where cognitive contribution and legitimate authority become distributed differently. Governance performance indicators will require ICAM's candidate architectural properties as a basis for identifying what might eventually be observed and compared, while simulation platforms will rely upon representations of nodes, flows, activation and bottlenecks to construct dynamic computational environments.

The relationship with historical and comparative research is equally important. Cognitive topologies, architectural fit and path dependence provide concepts through which institutional architectures can be examined across time and across different governance traditions. Applied pilot programmes can then use architectural diagnosis to identify specific mechanisms suitable for intervention, while future institutional design can explore alternative configurations without assuming that new technological capability automatically produces better architecture.

The emerging operational model suite should therefore be understood as a system of conceptual inheritance and methodological specialisation. Each model will simplify some dimensions while elaborating others, and their interoperability will depend upon preserving clear relationships with the concepts from which those models derive. The objective is not to force every Research Series into a common formal representation, but to ensure that findings developed through different methods can continue to speak to one another.

ICAM contributes to this possibility by establishing a shared architectural language for institutional cognition. Whether that language remains sufficiently stable to support a more formal model suite will become clearer only after the remaining L02 Working Papers have been developed. For the moment, **the importance of ICAM lies less in formal consolidation than in creating the first common representation of the institution as a distributed cognitive architecture capable of being examined from multiple methodological perspectives.**

36. Scientific Contribution of ICAM

The central contribution of ICAM is to move the study of Institutional Cognition from the identification of cognitive functions towards the analysis of the architectures through which those functions become collectively realised. The model argues that institutional intelligence cannot be inferred solely from the amount of knowledge, expertise or technology available within an organisation, because the effectiveness of those resources depends upon how they are distributed, connected and activated. Cognition becomes institutional when specialised capacities can influence one another through structures that preserve enough information, context and authority to support coherent judgement and action.

This architectural perspective introduces several concepts that can support subsequent research. Cognitive nodes provide a functional way of locating distributed capacities without assuming that cognition maps directly onto organisational charts.

Cognitive interfaces identify the relationships through which those capacities become mutually accessible. Cognitive flows describe the transformation of information as it moves through the institution, while coordination structures explain how specialised contributions become integrated into collective action. Bottlenecks, redundancy, diversity and architectural fit then provide ways of describing properties that may shape system-level cognition even when individual components appear capable in isolation.

The distinction between structural architecture and activated configuration adds a dynamic dimension, recognising that the cognitive organisation of an institution changes according to the problem being addressed. Temporal synchronisation makes visible the importance of when knowledge becomes available, while crisis cognition reveals the capacity of architectures to reconfigure under pressure. Metacognitive architecture extends the model further by identifying mechanisms through which institutions can examine not only outcomes but the processes through which their own knowledge and decisions are produced.

Perhaps most importantly, ICAM makes possible a distinct category of explanation: architectural failure. Institutions may possess relevant information, expertise and authority while failing cognitively because those capacities remain poorly connected. This perspective does not replace conventional explanations of institutional weakness, but it provides an additional mechanism through which repeated patterns of fragmentation, delay or non-learning can be investigated systematically.

The broader scientific contribution is therefore methodological as much as conceptual. **ICAM establishes institutional cognitive architecture as a researchable object positioned between individual cognition and system-level governance outcomes**, creating a level of analysis at which distributed knowledge, organisational structure and collective intelligence can become connected without reducing institutions either to individual actors or to abstract unified entities.

37. Future Research Directions

The development of ICAM opens a substantial research agenda because nearly every concept introduced in this Working Paper remains provisional until it has been examined across real institutional environments. One of the most immediate tasks will be to develop methodological protocols for cognitive mapping, allowing researchers to reconstruct functions, nodes and interfaces consistently enough that different institutions can be compared without erasing contextual differences. Such protocols will need to combine formal organisational evidence with methods capable of capturing informal and activated configurations.

A second direction concerns the empirical investigation of architectural properties. Connectivity, integration, centrality, latency, redundancy, diversity and adaptability may eventually become measurable dimensions, but their relationships with institutional cognition remain open questions. Research should therefore examine when these properties improve collective performance, when they create trade-offs and how their effects depend upon environmental complexity or institutional purpose.

Architectural fit represents another major area for investigation. Comparative studies could examine whether institutions facing similar environmental demands converge towards similar cognitive configurations or whether multiple architectures can perform effectively under comparable conditions. Longitudinal research could investigate how misfit emerges as environments change and how institutions recognise or fail to recognise the need for architectural adaptation.

Human–AI integration will provide an increasingly important extension. ICAM has established the possibility of computational and hybrid nodes within institutional architecture, but the distribution of tasks, authority, verification and oversight requires a dedicated model. Future research should therefore investigate not merely whether AI systems improve particular tasks, but how their incorporation alters interfaces, cognitive flows and dependencies across the institution.

Simulation offers another pathway, particularly for exploring mechanisms that are difficult to isolate empirically. Models of bottlenecks, redundancy, activation or temporal synchronisation may allow researchers to examine how alternative configurations behave under disruption or changing information loads. These simulations can then help identify hypotheses suitable for real-world testing.

Applied research should eventually examine whether targeted architectural interventions produce meaningful changes in institutional cognition. New feedback mechanisms, redesigned knowledge interfaces, alternative deliberative structures or modified human–AI configurations could each provide experimental opportunities. Such research will require careful safeguards because changing cognitive architecture can influence authority, responsibility and institutional legitimacy as well as analytical performance.

Finally, historical and comparative research will be essential for testing the portability of ICAM. The model has been developed within a contemporary context in which digital systems and formal organisations play substantial cognitive roles, yet institutions have organised collective knowledge through many different architectures across history. Examining these variations may reveal which concepts possess broader explanatory value and which reflect assumptions specific to contemporary governance.

The future research agenda is therefore inherently plural. **ICAM provides a common architectural object around which qualitative, comparative, computational, historical and experimental methods can begin to converge, while leaving open the possibility that the model itself will require substantial revision as those methods generate new evidence.**

38. Closing Perspective — From Distributed Capacity to Collective Intelligence

The central problem addressed by this Working Paper arises from a tension that lies at the heart of complex governance. Institutions depend upon cognitive specialisation because the problems they confront exceed the knowledge and attention of any individual actor, yet the very distribution that makes specialised cognition possible creates the risk of fragmentation. Expertise can become isolated, information can lose meaning as it moves through organisational hierarchies, memory can remain inaccessible, analytical systems can overwhelm decision processes and feedback can accumulate without producing learning. The capacity of institutions to govern intelligently therefore depends not only upon the resources they possess but upon the architectures through which those resources become connected.

ICAM has sought to make this architecture visible by identifying cognitive functions, nodes, interfaces, flows, coordination structures and feedback mechanisms as interconnected components of institutional cognition. The model emphasises that these relationships are dynamic rather than static, that different configurations become activated under different conditions and that institutions must coordinate cognition across multiple temporal scales. It also recognises that architecture exists within broader structures of authority and legitimacy, ensuring that the study of cognitive capability does not become detached from the constitutional conditions through which collective decisions acquire institutional meaning.

The model further suggests that some of the most consequential institutional weaknesses may be relational. An organisation can possess excellent expertise and sophisticated technology while remaining cognitively fragmented because those capacities cannot influence one another effectively. Conversely, relatively modest resources may support robust collective intelligence when interfaces preserve knowledge, feedback reaches learning processes and distributed expertise can become integrated without eliminating diversity. Architectural analysis therefore shifts attention from the possession of capability towards the conditions under which capability becomes systemically consequential.

This perspective does not imply that institutional intelligence can be engineered through architecture alone. Institutions remain historical and political systems whose cognitive arrangements emerge through legal structures, organisational traditions, power relationships and accumulated adaptation. Any attempt to redesign cognition must therefore remain attentive to path dependence and legitimate authority. The contribution of ICAM is more limited but also more precise: it provides a framework through which the architecture of collective cognition can become an explicit object of scientific investigation.

The transition established by this paper is consequently from **distributed cognitive capacity towards the architecture of collective intelligence**. ICAM does not claim to identify a universal form of the intelligent institution, nor does it assume that one topology, level of centralisation or technological configuration can optimise governance across every context. It instead provides a language for examining how different architectures organise the relationships through which information becomes knowledge, knowledge becomes judgement, judgement becomes legitimate action and experience becomes the basis of future learning.

In this sense, the Institutional Cognitive Architecture Model represents a necessary step in the operational development of Institutional Cognition. **Once the architecture of institutional cognition becomes sufficiently explicit to be reconstructed, compared and challenged, the programme can begin to investigate not only whether institutions are capable of collective intelligence, but why some configurations allow that intelligence to emerge more reliably than others.**

About the IDHUS Working Papers

This publication forms part of the constitutional publication ecosystem of the IDHUS Institute. It should not be interpreted as an isolated academic paper but as one component within a progressively expanding scientific architecture dedicated to the development of the emerging discipline of **Institutional Cognition**.

Within the organisational architecture of the IDHUS Institute, the Working Paper occupies a distinctive position. It comprises the elementary cognitive module through which operational research is conducted. At the structural level, however, the Working Paper also functions as the smallest autonomous organisational unit capable of generating identifiable scientific contributions within the broader research ecosystem. For the way we conduct our research, it represents the point at which constitutional concepts, Research Series, Research Lines, methodologies, computational developments and empirical investigations converge into a coherent research product capable of contributing directly to the cumulative evolution of institutional knowledge.

Series A — Foundations of Institutional Cognition

Series A preserves the constitutional foundations of the entire research ecosystem. Rather than constructing new constitutional theories, it continuously refines, clarifies and strengthens the conceptual architecture established throughout *The IDHUS Method*, the *Research Validation Papers* and the *Research Atlas*. It provides the common scientific language that allows every other Research Series to remain conceptually coherent while supporting the long-term evolution of Institutional Cognition.

Research Lines

A1 — Constitutional Evolution of Institutional Cognition
A2 — Conceptual Architecture and Theory Integration
A3 — Epistemology of Institutional Intelligence
A4 — Philosophy of Adaptive Governance
A5 — Scientific Foundations of Planetary Cognitive Civilisation

Series B — Cognitive Architectures

Series B investigates how cognition emerges and operates across multiple organisational scales, from individuals and institutions to artificial intelligence systems and distributed governance environments. Research within this Series explores organisational learning, collective intelligence, institutional memory, cognitive networks and adaptive decision-making, providing many of the theoretical and computational models that support the broader research programme.

Research Lines

B1 — Individual and Collective Cognition
B2 — Institutional Cognitive Architectures
B3 — Distributed Intelligence Systems

B4 — Human-AI Cognitive Collaboration
B5 — Organisational Learning Architectures

Series C — Computational Governance

Series C examines how artificial intelligence, computational modelling and digital infrastructures are transforming governance and public administration. Its objective is to understand how computational systems become active components of institutional cognition through intelligent decision-support, AI-native public administration, digital governance architectures and adaptive regulatory systems capable of responding to increasing societal complexity.

Research Lines

C1 — AI-Augmented Public Administration
C2 — Computational Policy Design
C3 — Intelligent Decision-Support Systems
C4 — Algorithmic Governance Architectures
C5 — AI-Native Institutional Ecosystems

Series D — Planetary Systems

Series D expands the scope of research from individual institutions towards planetary-scale systems of governance and coordination. Drawing upon systems thinking, complexity science and computational simulation, it investigates global governance networks, civilisational dynamics, resilience, distributed coordination and the cognitive infrastructures required to understand increasingly interconnected planetary systems.

Research Lines

D1 — Global Cognitive Systems
D2 — Planetary Governance Networks
D3 — Civilisational Systems Dynamics
D4 — Global Coordination Mechanisms
D5 — Planetary Resilience Architectures

Series E — AI and Institutional Intelligence

Series E explores the relationship between human intelligence, institutional cognition and artificial intelligence. Rather than studying AI as an isolated technology, it investigates how hybrid cognitive systems emerge through the interaction of people, organisations and intelligent computational agents. This Series provides the principal research environment for understanding the future evolution of AI-enabled institutions and distributed governance.

Research Lines

E1 — Institutional Intelligence
E2 — Human-AI Collaboration
E3 — AI Governance Systems
E4 — Autonomous Institutional Agents
E5 — Constitutional AI for Public Governance

Series F — Cognitive Metrics and Measurement

Series F develops the methodological foundations required to observe and evaluate Institutional Cognition. It designs metrics, indicators, assessment frameworks and measurement methodologies capable of transforming constitutional concepts into operational research variables. These tools support empirical research, comparative analysis, validation programmes and evidence-based governance across the entire research ecosystem.

Research Lines

F1 — Institutional Intelligence Metrics
F2 — Governance Performance Indicators
F3 — Cognitive Capacity Assessment
F4 — Adaptive Systems Measurement
F5 — Planetary Governance Metrics

Series G — Simulation and Digital Twins

Series G establishes the computational laboratories of the Institute. Through simulation environments, digital twins, systems dynamics and agent-based modelling, it creates experimental platforms for exploring institutional behaviour under controlled conditions. These environments allow researchers to investigate alternative governance architectures, analyse complex adaptive systems and evaluate future scenarios before they emerge in practice.

Research Lines

G1 — Institutional Digital Twins
G2 — Governance Simulation Platforms
G3 — Multi-Agent Governance Models
G4 — Scenario Generation and Strategic Foresight
G5 — Computational Experimentation

Series H — Comparative Civilisational Studies

Series H investigates how different societies, institutions and historical periods have organised governance, knowledge and collective intelligence. By comparing diverse civilisational experiences, governance models and institutional traditions, it seeks to identify recurring cognitive principles while understanding the contextual diversity through which Institutional Cognition evolves across cultures and historical epochs.

Research Lines

H1 — Comparative Institutional Cognition
H2 — Civilisational Governance Evolution
H3 — Cross-Cultural Adaptive Systems
H4 — Historical Institutional Intelligence
H5 — Comparative Planetary Governance

Series I — Applied Governance Laboratories

Series I serves as the operational bridge between scientific research and institutional practice. It develops collaborative research environments in which conceptual models, computational systems and governance methodologies can be explored through pilot projects, organisational experimentation, innovation programmes and long-term partnerships with public institutions. Practical experience generated through these laboratories continuously feeds back into the broader research programme.

Research Lines

I1 — Public Sector Innovation Laboratories
I2 — AI Governance Pilot Programmes
I3 — Institutional Transformation Frameworks
I4 — Governance Prototyping and Testing
I5 — Living Governance Laboratories

Series J — Future Research Frontiers

Series J ensures that the research ecosystem remains permanently open to future scientific developments. It explores emerging technologies, new governance paradigms, evolving AI capabilities and other frontier topics whose long-term significance may not yet be fully understood. By institutionalising organised exploration, this Series enables the IDHUS Institute to adapt continuously to future scientific opportunities while preserving the constitutional coherence of its broader research architecture.

Research Lines

J1 — Emerging AI Paradigms
J2 — Future Institutional Architectures
J3 — Post-Governance Systems
J4 — Civilisational Foresight
J5 — Unknown Research Horizons

Internal Constitutional References

The present Working Paper forms part of the constitutional and operational architecture of the IDHUS Institute. It should be read in conjunction with the following foundational publications, which collectively establish the constitutional framework within which the Working Papers Programme develops.

Constitutional Framework

- **The IDHUS Canon**
- **The IDHUS Method**

Programme I Research Validation Papers

- **RVP-001 Operational Cognitive Architecture**
- **RVP-002 Operational Consciousness**
- **RVP-003 Collective Cognition**
- **RVP-004 Institutional Cognition**
- **RVP-005 Governance Intelligence**
- **RVP-006 Adaptive Public Administration**
- **RVP-007 AI-Augmented Public Administration**
- **RVP-008 Adaptive Public Policy**
- **RVP-009 Cognitive Public Services**
- **RVP-010 Self-Evolving Governmental Systems**

Research Atlas

- **Volume 0 Research Ecosystem Architecture**
- **Volume 1 Research Architecture**
- **Volume 2 Planetary Research Ecosystems**
- **Volume 3 Planetary Cognitive Systems**
- **Volume 4 Planetary Governance Architecture**
- **Volume 5 Civilisational Evolution**
- **Volume 6 Planetary Cognitive Civilisation**

Operational Research Programme

- **Operational Research Programme- Working Papers Framework**

These publications together establish the constitutional foundations upon which the operational research developed throughout the IDHUS Working Papers Programme is constructed.

Glossary

Activated Cognitive Configuration

The temporary configuration of cognitive nodes, interfaces, functions and flows mobilised around a particular problem, decision or institutional condition. Activated cognitive configurations represent the dynamic use of an institution's broader structural cognitive architecture and may change substantially during crises, strategic transitions or unfamiliar governance situations.

Architectural Diagnosis

The systematic reconstruction and analysis of institutional cognitive architecture in order to identify where cognitive capacities become constrained, disconnected, delayed or otherwise unable to contribute effectively to collective cognition. Architectural diagnosis examines relationships among functions, nodes, interfaces, flows and authority rather than assuming that poor outcomes necessarily result from insufficient individual capability.

Architectural Failure

A form of institutional cognitive failure arising when the organisation of relationships among available cognitive capacities prevents those capacities from functioning effectively as a collective system. Architectural failure may emerge through interface failures, bottlenecks, fragmentation, temporal misalignment, inadequate feedback, excessive concentration or other relational weaknesses.

Architectural Fit

The dynamic relationship between an institutional cognitive architecture and the informational, temporal, environmental and decision requirements confronting the institution. Architectural fit does not imply the existence of a universally optimal configuration; it describes the degree to which a particular architecture remains appropriate to the conditions under which it must operate.

Cognitive Architecture

The organised configuration through which cognitive functions and processes become distributed and interconnected within a system. Within ICAM, cognitive architecture refers specifically to the relationships among functions, nodes, interfaces, flows, coordination structures and feedback mechanisms through which institutional cognition becomes collectively realised.

Cognitive Bottleneck

An architectural constraint through which the effective cognitive capacity of an institutional system becomes disproportionately limited by a particular node, interface, process or relationship. Bottlenecks may occur even when substantial cognitive capacity exists elsewhere within the institution.

Cognitive Capability

The capacity of a human, organisational or computational component to perform or contribute to a cognitive function such as perception, memory, interpretation, reasoning, prediction or learning. Cognitive capability remains analytically distinct from legitimate authority and does not by itself imply the right to determine collective decisions.

Cognitive Diversity

Variation in the knowledge, perspectives, methodologies, representations and interpretative approaches available within an institutional cognitive architecture. Cognitive diversity can reduce collective blind spots and strengthen reasoning under uncertainty, although its contribution depends upon interfaces capable of integrating heterogeneous perspectives without eliminating meaningful differences prematurely.

Cognitive Flow The movement and transformation of epistemically relevant content through institutional cognitive architecture. Cognitive flows may involve observations, data, information, contextualised knowledge, analytical interpretations, recommendations, decisions or feedback, and their cognitive significance may change as they pass through different nodes and interfaces.

Cognitive Function

A distinguishable process contributing to institutional cognition, including perception, memory, knowledge integration, reasoning, deliberation, decision-making, coordination, feedback, learning and adaptation. ICAM examines how these functions become architecturally realised rather than assuming that each corresponds to a single organisational component.

Cognitive Interface

An organisational, informational, procedural, human or technological mechanism through which cognitive nodes or functions become connected and capable of influencing one another. Cognitive interfaces determine not only whether information can move across an institution, but also how it is translated, contextualised, compressed or transformed during transmission.

Cognitive Map

An analytical representation of the distribution and relationships of cognitive functions, nodes, interfaces and flows within an institution or governance system. Unlike a conventional organisational chart, a cognitive map seeks to represent how knowledge is generated, transformed, integrated and used rather than only where formal administrative authority is located.

Cognitive Node

A human actor, collective organisational component, technological system or hybrid configuration capable of performing or contributing to one or more cognitive functions within an institutional architecture. The cognitive significance of a node depends both upon its local capabilities and upon its relationships with the wider institutional system.

Cognitive Redundancy

The presence of partially overlapping cognitive capacities, information sources, knowledge repositories or pathways within an institutional architecture. Redundancy may strengthen resilience and error detection when it provides meaningful alternative capacities, but may also generate unnecessary duplication when overlapping components do not contribute cognitive diversity or functional resilience.

Cognitive Synchronisation

The temporal coordination of cognitive processes operating at different speeds or institutional cycles. Cognitive synchronisation concerns whether information, analysis, deliberation, decision-making and feedback become connected at moments when their contributions can still influence collective judgement and action.

Cognitive Topology

A recurring pattern through which cognitive nodes, interfaces and functions become organised within an institutional architecture. Centralised, distributed, networked and hybrid topologies provide analytical descriptions of architectural patterns rather than fixed categories of institutions, and multiple topologies may coexist or become differently activated within the same institutional system.

Collective Intelligence

A system-level cognitive capacity emerging when distributed contributions become sufficiently connected and integrated to support forms of perception, reasoning, coordination, learning or problem-solving that cannot be attributed to isolated components alone. Within ICAM, collective intelligence is understood as an architectural achievement rather than the simple accumulation of individual capabilities.

Coordination Architecture

The configuration of mechanisms through which distributed cognitive contributions become integrated sufficiently to support collective judgement and action. Coordination may occur through hierarchy, networks, deliberative bodies, protocols, professional communities, computational systems or hybrid combinations of these mechanisms.

Distributed Cognition

The organisation of cognitive processes across multiple human, organisational and technological components rather than within a single decision-maker or cognitive centre. Distributed cognition provides the foundational condition from which ICAM investigates the architectural emergence of institutional cognition.

Feedback Architecture

The configuration of mechanisms through which information concerning the consequences of institutional actions returns to processes capable of evaluation, learning or revision. Feedback architecture determines whether experience becomes cognitively connected to future perception, reasoning, decision-making and adaptation.

Human–AI Cognitive Configuration

An arrangement in which human and artificial cognitive capabilities participate within shared institutional processes. ICAM recognises such configurations as part of contemporary cognitive architecture while leaving the detailed distribution of human and artificial functions, oversight and authority to dedicated research on human–AI collaboration.

Hybrid Cognitive Node

A cognitive node whose functional capability emerges through sustained interaction between human actors and technological or artificial systems. Hybrid nodes are analytically useful when the relevant cognitive contribution cannot be attributed meaningfully to either the human or computational component in isolation.

ICAM — Institutional Cognitive Architecture Model

The middle-range analytical model developed in this Working Paper for representing how cognitive functions become distributed, connected, activated and coordinated within institutions. ICAM translates selected relationships established within ICCA into an explicit framework centred upon institutional cognitive architecture and provides a foundation for subsequent comparison, measurement, simulation and experimentation.

ICCA — Institutional Cognition Conceptual Architecture

The conceptual architecture of the scientific field of Institutional Cognition. ICCA organises the theoretical relationships among foundational, functional, architectural, operational and evolutionary dimensions of the field, while ICAM shifts the unit of analysis towards the architecture through which cognition occurs within institutions themselves.

Institutional Cognition

The distributed processes through which institutions and governance systems perceive relevant environments, preserve and integrate knowledge, reason about complex conditions, deliberate, coordinate specialised actors, make collective decisions, learn from consequences and adapt over time.

Institutional Cognitive Architecture

The organised configuration of cognitive functions, human, organisational and computational nodes, cognitive interfaces, cognitive flows, coordination mechanisms and feedback structures through which institutional cognition becomes collectively realised.

Institutional Learning

The process through which institutions modify knowledge, interpretations, procedures, decision rules or behaviour in response to experience and feedback. Within ICAM, learning depends upon architectural pathways capable of connecting consequences with mechanisms of interpretation and revision.

Institutional Memory

The distributed capacity through which institutional knowledge, experience, decisions, procedures and historical understanding are preserved and remain potentially retrievable across time. Institutional memory may reside in human expertise, organisational routines, archives, databases, legal precedent and technological knowledge systems.

Institutional Metacognition

The capacity of an institution or governance system to generate representations of its own cognitive processes and use them to evaluate or modify how perception, knowledge integration, reasoning, coordination, decision-making or learning are organised.

Legitimate Authority

The institutionally recognised and normatively grounded capacity to make or authorise collective decisions. ICAM distinguishes legitimate authority from cognitive capability so that expertise, analytical capacity or computational performance are not treated as automatically conferring decision rights.

Meso-Level Analysis

An analytical level positioned between individual cognitive processes and system-level institutional outcomes. ICAM operates primarily at this level by examining the structures and relationships through which local cognitive contributions become organised into collective institutional processes.

Nested Cognitive Architecture

The condition in which cognitive architectures exist within and participate in wider cognitive systems across multiple scales. A team may possess its own internal architecture while functioning as a node within an institution, and an institution may itself become a node within a broader governance network.

Path Dependence

The process through which existing institutional structures, technologies, professional traditions, legal arrangements and previous adaptations constrain or shape subsequent architectural development. Within ICAM, cognitive architectures are understood as historically layered configurations rather than systems designed entirely from first principles.

Structural Cognitive Architecture

The relatively persistent distribution of cognitive functions, nodes, interfaces, flows and relationships available within an institution. Structural cognitive architecture provides the broader configuration from which different activated cognitive configurations may emerge in response to particular problems or conditions.

Temporal Architecture

The organisation of cognitive processes across different institutional timescales, including sensing, analysis, deliberation, decision-making, implementation, feedback and long-term learning. Temporal architecture influences whether cognitively relevant contributions become available at moments when they can remain consequential for institutional action.

IDHUS
Institute for the Development
of Human Societies